# Microstructure noise in the continuous case: The pre-averaging approach[☆]

Jean Jacod[a,*], Yingying Li[b], Per A. Mykland[b], Mark Podolskij[c], Mathias Vetter[d]

[a] *Institut de Mathématiques de Jussieu, CNRS UMR 7586 and Université Pierre et Marie Curie, 175 rue du Chevaleret, 75013 Paris, France*
[b] *Department of Statistics, The University of Chicago, Chicago, IL 60637, USA*
[c] *CREATES, School of Economics and Management, University of Aarhus, Building 1322, DK-8000 Aarhus C, Danmark, Germany*
[d] *Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany*

## Abstract

This paper presents a generalized pre-averaging approach for estimating the integrated volatility, in the presence of noise. This approach also provides consistent estimators of other powers of volatility — in particular, it gives feasible ways to consistently estimate the asymptotic variance of the estimator of the integrated volatility. We show that our approach, which possesses an intuitive transparency, can generate rate optimal estimators (with convergence rate $n^{-1/4}$).

    * Corresponding author.
    *E-mail addresses:* jean.jacod@upmc.fr (J. Jacod), yyli@uchicago.edu (Y. Li), mykland@galton.uchicago.edu (P.A. Mykland), mark.podolskij@math.ethz.ch (M. Podolskij), mathias.vetter@rub.de (M. Vetter).

## 1. Introduction

The recent years have seen a revolution in the statistics of high frequency data. On the one hand, such data is increasingly available and needs to be analyzed. This is particularly the case for market prices of stocks, currencies, and other financial instruments. On the other hand, the technology for the analysis of such data has grown rapidly. The emblematic problem is the question of how to estimate daily volatility for financial prices (in stochastic process terms, the quadratic variation of log prices).

The early theory was developed in the context of stochastic calculus, before the financial application was apparent. The sum of squared increments of the process was shown to be consistent for the quadratic variation in [1]. A limit theory was then developed in [2,3], and later in [4].

Meanwhile, these concepts were introduced to econometrics in [5–7]. A limit theory was developed in [8,9]. Further early econometric literature includes, in particular, [10–17]. The setting of confidence intervals using bootstrapping has been considered by Goncalves and Meddahi [18] and Kalnina and Linton [19].

The direct application to data of results from stochastic calculus have, however, run into the problem of microstructure. No-arbitrage based characterizations of securities prices (as in [20]) suggest that these must normally be semimartingales. Econometric evidence, however, suggests that there is additional noise in the prices. This goes back to Roll [21] and Hasbrouck [22]. In the nonparametric setting, the deviation from semimartigales is most clearly seen through the signature plots of Andersen et al. [10], see also the discussion in [23].

Statistical and econometric research has for this reason gravitated towards the concept that the price (and log price) semimartingale is latent rather than observed. Research goes back to the work on rounding by Jacod [24] and Delattre and Jacod [25]. Additive noise is studied in [26], and a consistent estimator in the nonparametric setting is found in [27]. Issues of bias-variance tradeoff are discussed in [28]. In the nonparametric case, rate optimal estimators are found in [29–31]. A development for low frequency data is given in [32].

There are currently three main approaches to estimation in the nonparametric case: linear combination of realized volatilities obtained by *subsampling* [27,29], and linear combination of *autocovariances* [31]. The purpose of this paper is to give more insight to the third approach of *pre-averaging*, which was introduced in Podolskij and Vetter [30] for i.i.d. noise and for non overlapping intervals. The idea is as follows. We suppose that the (say) log securities price $X_t$ is a continuous semimartingale (of the form (2.1)). The observations are recorded prices at transaction times $t_i = i \Delta_n$, and what is observed is not $X_{t_i}$, but rather $Z_{t_i}$, given by

$$Z_{t_i} = X_{t_i} + \epsilon_{t_i}. \tag{1.1}$$

The noise $\epsilon_{t_i}$ can be independent of the $X$ process, or have a more complex structure, involving for example some rounding. The idea is now that if one averages $K$ of these $Z_{t_i}$'s, one is closer to the latent process. Define $\check{Z}_{t_i}$ as the average of $Z_{t_{i+j}}$, $j = 0, \ldots, K - 1$. The variance of the noise in $\check{Z}_i$ is now reduced by a factor of about $1/K$. If one calculates the realized volatility on the basis of $\check{Z}_0, \check{Z}_{t_1}, \check{Z}_{t_2}, \ldots$, the estimate is therefore closer to being based on the true underlying semimartingale. The scheme is particularly appealing since it is obviously robust to a wide variety of structures of the noise $\epsilon$.

The paper provides a way of implementing this idea. There are several issues that have to be tackled in the process. First of all, the pre-averaging brings in a particular dependence structure

which necessitate an adjustive constant in front of the sum of squared increments of the averages. Second, while the local averaging does reduce the impact of the noise $\epsilon$, it does not completely eliminate the bias. The pre-averaged realized volatility therefore has to be adjusted by an additive term to eliminate the remaining error. Third, one would not wish to only average over differences from non-overlapping intervals, but rather use a moving window. Fourth, the estimator can be generalized by the use of a general weight function. Our final estimator is thus on the form (3.6), where we note that the special case of simple averaging is given in the example following Theorem 3.1. Note that in the notation of that example, $k_n = 2K$. Finally, the method used here is amenable to much more general noise models than the most usual i.i.d. noise, independent of the process $X$: see Section 2.

Like the subsampling and the autocovariance methods, the pre-averaging approach, when well implemented, gives rise to rate optimal estimators (the convergence rate being $O_p(n^{-1/4})$). This result, along with a central limit theorem for the estimator, is given as our main result Theorem 3.1.

What is the use of a third approach to the estimation problem, when there already are two that provide good convergence? There are at least three advantages of the pre-averaging procedure:

(i) Transparency. It is natural to think of the latent process $X_t$ as the average of observations in a small interval. Without this assumption, identifiability problems may arise, as documented in [33]. Our procedure implements estimation directly based on this assumption. Also, as noted after the definition (3.6), the entire randomness in the estimator is, to first order, concentrated in a single sum of squares.

(ii) Estimation of other powers of volatility. The pre-averaging approach also provides straightforward consistent estimators of quarticity, thereby moving all the existing estimators closer to the feasible setting of confidence intervals. See [30] for results in the case of independent noise.

(iii) Edge effects. The three classes of estimators are similar also in that they are based on a weight or kernel function. To some approximation, one can rewrite all subsampling estimators as autocovariance estimators, and *vice versa*. The estimators in this paper can be rewritten, again to first order, as a class of subsampling or autocovariance estimators, cf. Remark 1. The difference between the three classes of estimators (and what is concealed by the term "to first order") lies in the treatment of edge effects. The potential impact of such effects can be considerable, cf. Bandi and Russell [34]. In some cases, the edge effects can even affect asymptotic properties. Because of the intuitive nature of our estimator, edge effects are less likely to be a problem, and they certainly do not interfere with the asymptotic results.

The plan of the paper is as follows. The mathematical model is defined in Section 2, and results are stated in Section 3. Section 4 provides a simulation study. The proofs are in Section 5.

## 2. The setting

We have a 1-dimensional underlying continuous process $X = (X_t)_{t \geq 0}$, and observation times $i \Delta_n$ for all $i = 0, 1, \ldots, k, \ldots$. We are in the context of high frequency data, that is we are interested in the situation where the time lag $\Delta_n$ is "small", meaning that we look at asymptotic properties as $\Delta_n \to 0$. The process $X$ is observed with an error: that is, at stage $n$ and instead of the values $X_i^n = X_{i\Delta_n}$ for $i \geq 0$, we observe real variables $Z_i^n$, which are somehow related to the $X_i^n$, in a way which is explained below.

Our aim is to estimate the integrated volatility of the process $X$, over a fixed time interval $[0, t]$, on the basis of the observations $Z_i^n$ for $i = 0, 1, \ldots, [t/\Delta_n]$. For this, we

need some assumptions on $X$ and on the "noise", and to begin with we need $X$ to be a continuous Itô semimartingale, so that the volatility is well defined. Being a continuous Itô semimartingale means that the process $X$ is defined on some filtered probability space $(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t \geq 0}, \mathbb{P}^{(0)})$ and takes the form

$$X_t = X_0 + \int_0^t b_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s, \tag{2.1}$$

where $W = (W_t)$ is a standard Wiener process and $b = (b_t)$ and $\sigma = (\sigma_t)$ are adapted processes, such that the above integrals make sense. In fact, we will need some, relatively weak, assumptions on these processes, which are gathered in the following assumption:

**Assumption (H).** We have (2.1) with two process $b$ and $\sigma$ which are adapted and càdlàg (= "right-continuous with left limits" in time).  $\square$

In this paper, we are interested in the estimation of the integrated volatility, that is the process

$$C_t = \int_0^t \sigma_s^2 \, \mathrm{d}s. \tag{2.2}$$

Next we turn to the description of the "noise". Loosely speaking, we assume that, conditionally on the whole process $X$, and for any given $n$, the observed values $Z_i^n$ are independent, each one having a (conditional) law which possibly depends on the time and on the outcome $\omega$, in an "adapted" way, and with conditional expectations $X_i^n$.

Mathematically speaking, this can be realized as follows: for any $t \geq 0$ we have a transition probability $Q_t(\omega^{(0)}, \mathrm{d}z)$ from $(\Omega^{(0)}, \mathcal{F}_t^{(0)})$ into $\mathbb{R}$, which satisfies

$$\int z \, Q_t(\omega^{(0)}, \mathrm{d}z) = X_t(\omega^{(0)}). \tag{2.3}$$

We endow the space $\Omega^{(1)} = \mathbb{R}^{[0,\infty)}$ with the product Borel $\sigma$-field $\mathcal{F}^{(1)}$ and with the probability $\mathbb{Q}(\omega^{(0)}, \mathrm{d}\omega^{(1)})$ which is the product $\otimes_{t \geq 0} \, Q_t(\omega^{(0)}, .)$. We also call $(Z_t)_{t \geq 0}$ the "canonical process" on $(\Omega^{(1)}, \mathcal{F}^{(1)})$ and the filtration $\mathcal{F}_t^{(1)} = \sigma(Z_s : s \leq t)$. Then we consider the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ defined as follows:

$$\left. \begin{array}{l} \Omega = \Omega^{(0)} \times \Omega^{(1)}, \qquad \mathcal{F} = \mathcal{F}^{(0)} \times \mathcal{F}^{(1)}, \qquad \mathcal{F}_t = \cap_{s>t} \mathcal{F}_s^{(0)} \times \mathcal{F}_s^{(1)}, \\ \mathbb{P}(\mathrm{d}\omega^{(0)}, \mathrm{d}\omega^{(1)}) = \mathbb{P}^{(0)}(\mathrm{d}\omega^{(0)}) \, \mathbb{Q}(\omega^{(0)}, \mathrm{d}\omega^{(1)}). \end{array} \right\} \tag{2.4}$$

Any variable or process which is defined on either $\Omega^{(0)}$ or $\Omega^{(1)}$ can be considered in the usual way as a variable or a process on $\Omega$. By standard properties of extensions of spaces, $W$ is a Wiener process on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, and Eq. (2.1) holds on this extended space as well.

In fact, here again we need a little bit more than what precedes:

**Assumption (K).** We have (2.3) and further the process

$$\alpha_t(\omega^{(0)}) = \int z^2 \, Q_t(\omega^{(0)}, \mathrm{d}z) - X_t(\omega^{(0)})^2 = \mathbb{E}((Z_t)^2 \mid \mathcal{F}^{(0)})(\omega^{(0)}) - X_t(\omega^{(0)})^2 \tag{2.5}$$

is càdlàg (necessarily $(\mathcal{F}_t^{(0)})$-adapted), and

$$t \mapsto \int z^8 \, Q_t(\omega^{(0)}, \mathrm{d}z) \quad \text{is a locally bounded process.} \tag{2.6}$$

Taking the 8th moment in (2.6) is certainly not optimal, but this condition is in fact quite mild (we need in any case the second moment to be locally bounded). The really strong requirement above is the condition (2.3) on the noise. This assumption, together with the product structure of the measure $\mathbb{Q}$, summarize three properties at once:

(1) If $Y_t = \int z\, Q_t(dz)$ denotes the conditional expectation of the observed value $Z_t$ at time $t$, conditionally on $\mathcal{F}^{(0)}$ (the $\sigma$-field generated by the latent process and all possible covariates), then *Y is a semimartingale*.
(2) Moreover, we have $Y_t = X_t$ (unbiasedness).
(3) Finally, knowing $\mathcal{F}^{(0)}$, the errors are independent at different times.

If (1) holds and (2) fails, our procedure provides an estimator for the integrated volatility of $Y_t$ instead of $X_t$ and, as explained in [33], it is probably impossible to make inference on the process $X$ itself, unless, of course, further specification of the model is assumed. If (1) fails, our procedure just breaks down, and we doubt that there might exist in general a method for retrieving the volatility of $X$ (see Example 3). As for (3), without any doubt it could be substantially weakened, to allow for some (weak enough) dependency (cf. the discussion in [35]), but the independence assumption simplifies things quite a lot !

**Example 1.** If $Z_i^n = X_i^n + \varepsilon_i^n$, where the sequence $(\varepsilon_i^n)_{i\geq 0}$ is i.i.d. centered with finite 8th moment and independent of $X$, then (K) is obviously satisfied. $\square$

**Example 2.** Let $Z_i^n = \gamma \lfloor (X_i^n + \varepsilon_i^n)/\gamma \rfloor$ for some $\gamma > 0$ and $(\varepsilon_i^n)$ as in the previous example. This amounts to having an additive i.i.d. noise and then taking the rounded-off value with lag $\gamma$, for example $\gamma = 1$ cent. Then as soon as the $\varepsilon_i^n$ are uniform over $\lfloor 0, \gamma \rfloor$, or more generally uniform over $[-2i\gamma, (2i+1)\gamma]$ for some integer $i$, (K) is satisfied. Another example of model involving rounding with (K) satisfied is given in model 2 of Section 4. If the $\varepsilon_i^n$ have a $C^2$ density, with further a finite 8th moment and a support containing an interval of length $\gamma$, then (K) is not satisfied in general but the process $Y$ introduced above is of the form $Y = f(X)$ for a $C^2$ function $f$, and so everything goes through if we replace $X$ by $Y$ below. $\square$

**Example 3.** Let $Z_i^n = \gamma \lfloor X_i^n/\gamma \rfloor$ for some $\gamma > 0$ ("pure rounding"). Then the errors $Z_i^n - X_i^n$ are independent, conditionally on $X$, but (K) is not satisfied, and the process $Y$ is *not* a semimartingale, and is not even càdlàg: so nothing of what follows applies. In fact in this case, if we observe the whole process $Z_t = \gamma \lfloor X_t/\gamma \rfloor$ over some interval $[0, T]$, we can derive the local times $L_t^x$ for $t \in [0, T]$ of the process $X$ at each level $x = i\gamma$ for $i \in \mathbb{Z}$, but nothing else, and in particular we cannot infer the values of the process $C_t$.

## 3. The results

We need first some notation. We choose a sequence $k_n$ of integers and a number $\theta \in (0, \infty)$ satisfying

$$k_n\sqrt{\Delta_n} = \theta + \text{o}(\Delta_n^{1/4}) \tag{3.1}$$

(for example $k_n = [\theta/\sqrt{\Delta_n}]$). We also choose a function $g$ on $[0, 1]$, which satisfies

$$\left.\begin{aligned} &g \text{ is continuous, piecewise } C^1 \text{ with a piecewise Lipschitz derivative } g', \\ &g(0) = g(1) = 0, \qquad \int_0^1 g(s)^2 ds > 0. \end{aligned}\right\} \tag{3.2}$$

We associate with $g$ the following numbers and functions on $\mathbb{R}_+$:

$$g_i^n = g(i/k_n), \qquad h_i^n = g_{i+1}^n - g_i^n, \tag{3.3}$$

$$
\left.
\begin{aligned}
&s \in [0, 1] \mapsto \phi_1(s) = \int_s^1 g'(u)g'(u - s)\,\mathrm{d}u, \qquad \phi_2(s) = \int_s^1 g(u)g(u - s)\,\mathrm{d}u \\
&s > 1 \mapsto \phi_1(s) = 0, \qquad \phi_2(s) = 0 \\
&i, j = 1, 2 \Rightarrow \Phi_{ij} = \int_0^1 \phi_i(s)\phi_j(s)\,\mathrm{d}s, \qquad \psi_i = \phi_i(0).
\end{aligned}
\right\} \tag{3.4}
$$

Next, with any process $V = (V_t)_{t \geq 0}$ we associate the following random variables

$$
\left.
\begin{aligned}
&V_i^n = V_{i\Delta_n}, \qquad \Delta_i^n V = V_i^n - V_{i-1}^n, \\
&\overline{V}_i^n = \sum_{j=1}^{k_n-1} g_j^n \Delta_{i+j}^n V = -\sum_{j=0}^{k_n-1} h_j^n V_{i+j}^n
\end{aligned}
\right\} \tag{3.5}
$$

(the two versions of $\overline{V}_i^n$ are identical because $g(0) = g(1) = 0$).

Recall that in our setting, we do not observe the process $X$, but the process $Z$ only, and at times $i\Delta_n$. So our estimator should be based on the values $Z_i^n$ only, and we propose to take

$$\widehat{C}_t^n = \frac{\sqrt{\Delta_n}}{\theta\psi_2} \sum_{i=0}^{[t/\Delta_n]-k_n+1} (\overline{Z}_i^n)^2 - \frac{\psi_1\Delta_n}{2\theta^2\psi_2} \sum_{i=1}^{[t/\Delta_n]} (\Delta_i^n Z)^2. \tag{3.6}$$

The last term above is here to remove the bias due to the noise, but apart from that it plays no role in the central limit theorem given below.

As we will see, these estimators are asymptotically consistent and mixed normal, and in order to use this asymptotic result we need an estimator for the asymptotic conditional variance. Among many possible choices, here is an estimator:

$$
\begin{aligned}
\Gamma_t^n = {}& \frac{4\Phi_{22}}{3\theta\psi_2^4} \sum_{i=0}^{[t/\Delta_n]-k_n+1} (\overline{Z}_i^n)^4 \\
&+ \frac{4\Delta_n}{\theta^3}\left(\frac{\Phi_{12}}{\psi_2^3} - \frac{\Phi_{22}\psi_1}{\psi_2^4}\right) \sum_{i=0}^{[t/\Delta_n]-2k_n+1} (\overline{Z}_i^n)^2 \sum_{j=i+k_n}^{i+2k_n-1} (\Delta_j^n Z)^2 \\
&+ \frac{\Delta_n}{\theta^3}\left(\frac{\Phi_{11}}{\psi_2^2} - 2\frac{\Phi_{12}\psi_1}{\psi_2^3} + \frac{\Phi_{22}\psi_1^2}{\psi_2^4}\right) \sum_{i=1}^{[t/\Delta_n]-2} (\Delta_i^n Z)^2 (\Delta_{i+2}^n Z)^2.
\end{aligned} \tag{3.7}
$$

**Theorem 3.1.** *Assume (H) and (K). For any fixed $t > 0$ the sequence $\frac{1}{\Delta_n^{1/4}}(\widehat{C}_t^n - C_t)$ converges stably in law to a limiting variable defined on an extension of the original space, and which is of the form*

$$Y_t = \int_0^t \gamma_s \,\mathrm{d}B_s, \tag{3.8}$$

*where $B$ is a standard Wiener process independent of $\mathcal{F}$ and $\gamma_t$ is the square-root of*

$$\gamma_t^2 = \frac{4}{\psi_2^2}\left(\Phi_{22}\theta\sigma_t^4 + 2\Phi_{12}\frac{\sigma_t^2\alpha_t}{\theta} + \Phi_{11}\frac{\alpha_t^2}{\theta^3}\right). \tag{3.9}$$

*Moreover*

$$\Gamma_t^n \xrightarrow{\mathbb{P}} \int_0^t \gamma_s^2 \, ds, \tag{3.10}$$

*and therefore, for any $t > 0$, the sequence $\frac{1}{\Delta_n^{1/4} \sqrt{\Gamma_t^n}} (\widehat{C}^n - C)$ converges stably in law to an $\mathcal{N}(0, 1)$ variable, independent of $\mathcal{F}$.*

**Example.** The simplest function $g$ is probably

$$g_0(x) = x \wedge (1 - x). \tag{3.11}$$

In this case we have

$$\psi_1 = 1, \qquad \psi_2 = \frac{1}{12}, \qquad \Phi_{11} = \frac{1}{6}, \qquad \Phi_{12} = \frac{1}{96}, \qquad \Phi_{22} = \frac{151}{80\,640} \tag{3.12}$$

and also, with $k_n$ even, we have

$$\overline{Z}_i^n = \frac{1}{k_n} \left( \sum_{j=k_n/2}^{k_n-1} Z_{i+j}^n - \sum_{j=0}^{k_n/2-1} Z_{i+j}^n \right). \tag{3.13}$$

**Remark 1.** Our estimators are in fact essentially the same as the kernel estimators in [31]. With our notation the "flat top" estimators of that paper are

$$\overline{K}_t^n = \sum_{i=k_n}^{[t/\Delta_n]-k_n+1} (\Delta_i^n Z)^2$$

$$+ \sum_{k_n \le i \le [t/\Delta_n]-k_n+1, \ 1 \le j \le k_n} k\left(\frac{j-1}{k_n}\right) \left(\Delta_i^n Z \, \Delta_{i+j}^n Z + \Delta_i^n Z \, \Delta_{i-j}^n Z\right),$$

where $k$ is some (smooth enough) weight function on $[0, 1]$ having $k(0) = 1$ and $k(1) = 0$, and also $k'(0) = k'(1) = 0$. Then we see that

$$\widehat{C}_t^n = \overline{K}_t^n (1 + O(\sqrt{\Delta_n})) - \frac{\psi_1 \Delta_n}{2\theta^2 \psi_2} \sum_{i=1}^{[t/\Delta_n]} (\Delta_i^n Z)^2 + \text{border terms},$$

provided we take $k(s) = \phi_2(s)/\psi_2$, so there is a one-to-one correspondence between the weight functions $g$ and $k$. The "border terms" are terms arising near 0 and $t$, because the two sums in the definition of $\overline{K}_t^n$ do not involve exactly the same increments of $Z$. These border terms turn out to be of order $\Delta^{1/4}$, the same order than $\widehat{C}_t^n - C_t$, although they are asymptotically unbiased (but usually not asymptotically mixed normal). This explains why our CLT is somehow simpler than the equivalent results in Barndorff-Nielsen et al. [31].

**Remark 2.** Suppose that $Y_t = \sigma W_t$ and that $\alpha_t = \alpha$, where $\sigma$ and $\alpha$ are positive constants, and that $t = 1$. In this case there is an efficient parametric bound for the asymptotic variance for estimating $\sigma^2$ in presence of i.i.d. noise, which is $8\sigma^3 \sqrt{\alpha}$, see e.g. Gloter and Jacod [26]. On the other hand, the concrete estimators given in [31] or [30] or [29], in the i.i.d. additive noise case again, have an asymptotic variance ranging from $8.01 \, \sigma^3 \sqrt{\alpha}$ (and even from the optimal variance $8 \sigma^3 \sqrt{\alpha}$ in [31], although in a slightly different setting) to $26 \sigma^3 \sqrt{\alpha}$, upon using an

"optimal" choice of $\theta$ in (3.1). To compare with these results, here the "optimal" asymptotic variance in the simple case (3.11), obtained for $\theta = 4.777 \sqrt{\alpha}/\sigma$, is $8.545 \, \sigma^3 \sqrt{\alpha}$, quite close to the efficient bound. Note that other than (3.11) and more appropriate choices of $g$ would give lower asymptotic variances.

**Remark 3.** In practice, going down from 8.545 to 8 as in the previous remark is rather irrelevant, in front of the fact that we do not know how to choose $\theta$ in an optimal way (this is the drawback of all previously quoted papers as well, and especially for the efficient estimator of Gloter and Jacod [26]). More: since $\sigma = \sigma_t$ and $\alpha = \alpha_t$ are usually random and time dependent, there is no "optimal" choice of the number $\theta$; one should rather take a $\theta = \theta_t$ which is time-varying and data-dependent, and based on "local" estimates of $\sigma_t$ and $\alpha_t$.

We have not pursued this matter in this paper. However, simple compromises can be reached as follows: we usually have an idea of the "average" sizes $\alpha_{\text{ave}}$ and $\sigma_{\text{ave}}$ of $\alpha_t$ and $\sigma_t$: in this case one may take $\theta$ close to $4.8 \sqrt{\alpha_{\text{ave}}}/\sigma_{\text{ave}}$. Or, better, we can estimate $\int_0^T \sigma_t^4 \mathrm{d}t$, $\int_0^T \sigma_t^2 \alpha_t \mathrm{d}t$ and $\int_0^T \alpha_t^2 \mathrm{d}t$ from the data (see the next remark for an estimator of the first of these quantities, the two others can be similarly estimated using the convergence (5.65) in the proof section. To get these estimates, one uses a preliminary value for $\theta$). Then we plug these estimators in Eq. (3.9) to get an efficient value of $\theta$.

**Remark 4.** Note that a consistent estimator of the quarticity can be constructed using a different linear combination of the three parts in (3.7). More specifically, we have

$$
\begin{aligned}
\hat{Q}_t^n = \quad & \frac{1}{3\theta^2 \psi_2^2} \sum_{i=0}^{[t/\Delta_n]-k_n+1} (\overline{Z}_i^n)^4 \\
& - \frac{\Delta_n \psi_1}{\theta^4 \psi_2^2} \sum_{i=0}^{[t/\Delta_n]-2k_n+1} (\overline{Z}_i^n)^2 \sum_{j=i+k_n}^{i+2k_n-1} (\Delta_j^n Z)^2 \\
& + \frac{\Delta_n \psi_1^2}{4\theta^4 \psi_2^2} \sum_{i=1}^{[t/\Delta_n]-2} (\Delta_i^n Z)^2 (\Delta_{i+2}^n Z)^2. \\
& \xrightarrow{\ \mathbb{P}\ } \int_0^t \sigma_t^4 \mathrm{d}t;
\end{aligned}
\tag{3.14}
$$

see Section 5 for details.

**Remark 5.** Let us emphasize once more that, even though the structure of the noise considered here is quite more general than in the existing literature, it still leaves out the case where the noise is correlated, conditionally on $X$. It is likely that in the weakly correlated case similar results (with different asymptotic variances) are true, but it would seriously complicate the analysis.

Another point is worth mentioning: in contrast with some of the existing literature on i.i.d. noise and most of the literature on rounding, we do not consider a noise level going to 0 as the frequency increases. If it were the case, the same analysis would apply, but the rate (and the choice of $k_n$ in (3.1), to begin with) would be different.

## 4. Simulation results

In this section, we examine the performance of our estimator.

Table 1
Sample size "$n$" for four selected stocks from the New York stock exchange: Ford (F), Microsoft (MSFT), Intel (INTC), and Pfizer (PFE), and one week in April 2005.

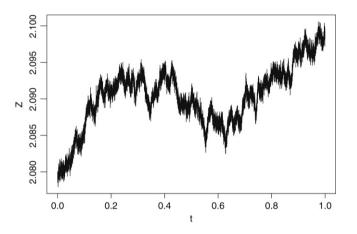|  | F | MSFT | INTC | PFE |
|---|---|---|---|---|
| 4/4/2005 | 4998 | 80 984 | 91 516 | 15 378 |
| 4/5/2005 | 4938 | 88 721 | 74 355 | 29 101 |
| 4/6/2005 | 4033 | 103 899 | 76 048 | 16 723 |
| 4/7/2005 | 5405 | 100 339 | 69 563 | 36 140 |
| 4/8/2005 | 6199 | 65 909 | 73 223 | 15 407 |



Fig. 1. One observed sample path generated by model 1.

### 4.1. Simulation design

We study the case when the weight function is taken to be $g(x) = x \wedge (1 - x)$. We simulate data for one day ($t \in [0, 1]$), and assume the data is observed once every second ($n = 23\,400$). The $X$ processes and the market microstructure noise processes are generated from the models below. 25 000 iterations were run for each model. The parameter values are chosen to be close to those used in [27].

The actual values of $n$ will vary from stock to stock and day to day. To get an idea of the range for commonly traded stock, we report some values in Table 1. Our choice to use $n = 23\,400$ is meant to represent a fairly heavily traded stock.

Model 1 — the case of constant volatility & additive noise.

$$X_t = X_0 + \sigma W_t, \qquad Z_i^n = X_i^n + \epsilon_i^n.$$

Parameters used: $\sigma = 0.2/\sqrt{252}$, $\epsilon_i^n \sim i.i.d.\ \mathcal{N}(0, 0.0005^2)$.

The observed sample path looks as in Fig. 1:

Model 2 — the case of constant volatility & rounding plus error.

$$X_t = X_0 + \sigma W_t, \qquad Z_i^n = \log\left(\gamma \left\lfloor \frac{\exp(X_i^n + \epsilon_i^n)}{\gamma} \right\rfloor\right)$$
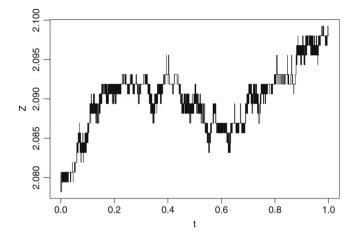
Fig. 2. One observed sample path generated by model 2.

where $\epsilon_i^n = \eta_i^n \, \log \frac{\gamma \lceil \frac{\exp(X_i^n)}{\gamma} \rceil}{\exp(X_i^n)}$ and $\eta_i^n$ is a Bernoulli variable (probabilities $p_i^n$ and $1 - p_i^n$ of taking

values 1 and 0), with $p_i^n = \log \left( \frac{\exp(X_i^n)}{\gamma \lfloor \frac{\exp(X_i^n)}{\gamma} \rfloor} \right) \Big/ \log \left( \frac{\gamma \lceil \frac{\exp(X_i^n)}{\gamma} \rceil}{\gamma \lfloor \frac{\exp(X_i^n)}{\gamma} \rfloor} \right)$.

This model is similar to the two-stage contamination model studied in Li and Mykland [33], where the observed log prices $Z_i^n$'s are the logarithm of the rounded contaminated prices. Basically this model amounts to randomly assigning one of the two nearest rounding grids of $X_t$ to its corresponding observation (with probability $p_t$ the upper grid $\log(\gamma \lceil \frac{\exp(X_t)}{\gamma} \rceil)$, $1 - p_t$ the lower grid $\log(\gamma \lfloor \frac{\exp(X_t)}{\gamma} \rfloor)$). It's easy to check that for this model, the Assumption (K) is satisfied.

Parameters used: $\sigma = 0.2/\sqrt{252}$, $X_0 = \log(8)$, $\gamma = 0.01$. Note that the initial value $X_0$ is chosen so that the averaged size of error $\alpha_{\text{ave}}$ is close to that in the simulations of model 1 and model 3.

The observed log price process looks like in Fig. 2.

Model 3 — the case of stochastic volatility & additive noise. The Heston model [36] is used to generate the stochastic volatility process.

$$dX_t = (\mu - v_t/2)dt + \sigma_t dB_t, \qquad Z_i^n = X_i^n + \epsilon_i^n$$

and

$$dv_t = \kappa(\alpha - v_t)dt + \gamma v_t^{1/2} dW_t,$$

where $v_t = \sigma_t^2$ and we assume $\text{Corr}(B, W) = \rho$.

Parameters used: $\mu = 0.05/252$, $\kappa = 5/252$, $\alpha = 0.04/252$, $\gamma = 0.05/252$, $\rho = -0.5$ and $\epsilon_i^n \sim i.i.d. \, \mathcal{N}(0, 0.0005^2)$.

## 4.2. Simulation results

Some initial simulations (not recorded here) showed that our estimator is fairly robust to the choice of $k_n$, in other words, it performs reasonably well for a large range of $k_n$. Since $\theta$ comes

Table 2
Relative bias in the estimators $\hat{C}_t^n$ and $\Gamma_t^n$, for the three models.

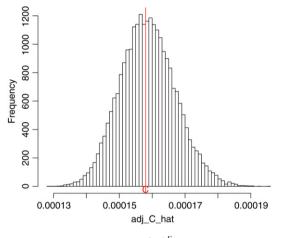|  | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Relative small-sample bias Avg$[(\hat{C}_t^n - C)/C]$ | $-0.008759$ | $-0.009784$ | $-0.008463$ |
| Relative bias in the variance estimator Avg$[(\Gamma_t^n - \int_0^t \gamma_s^2 ds)/ \int_0^t \gamma_s^2 ds]$ | $-0.014182$ | $-0.013898$ | $-0.013085$ |



Fig. 3. Histogram of $\hat{C}_t^{n,\text{adj}}$, for Model 1.

from asymptotic statistics, it doesn't give precise instruction about $k_n$ for small samples. On the other hand, when computing the true asymptotic variance $\int_0^t \gamma_s^2 ds$, the $\theta$ we should use is really $k_n \sqrt{\Delta_n}$. We decided to firstly fix $k_n$ to be close to the one suggested by the optimal $\theta$, and then re-define $\theta$ to be $k_n \sqrt{\Delta_n}$ for further computations. In all our simulations, we used $k_n = 51$, which corresponds to a $\theta \approx 1/3$.

Table 2 reports the performance of the estimator $\hat{C}_t^n$ and the variance estimator $\Gamma_t^n$.

As we will see later, the results in Model 1 (see Fig. 3), where $\hat{C}_t^n$ and $\Gamma_t^n$ are essentially normal distributed, show the importance of a correction of these estimators, when dealing with small sample sizes. We propose to replace the parameters $\psi_i$ and $\Phi_{ij}$ by their finite sample analogues, which are defined as follows:

$$\psi_1^{k_n} = k_n \sum_{j=1}^{k_n} (g_{j+1}^n - g_j^n)^2, \qquad \psi_2^{k_n} = \frac{1}{k_n} \sum_{j=1}^{k_n-1} (g_j^n)^2$$

$$\phi_1^{k_n}(j) = \sum_{j=i+1}^{k_n} (g_{i-1}^n - g_i^n)(g_{i-j-1}^n - g_{i-j}^n), \qquad \phi_2^{k_n}(j) = \sum_{j=i+1}^{k_n} g_i^n g_{i-j}^n$$

$$\Phi_{11}^{k_n} = k_n \left( \sum_{j=0}^{k_n-1} (\phi_1^{k_n}(j))^2 - \frac{1}{2}(\phi_1^{k_n}(0))^2 \right)$$

$$\Phi_{12}^{k_n} = \frac{1}{k_n} \left( \sum_{j=0}^{k_n-1} \phi_1^{k_n}(j)\phi_2^{k_n}(j) - \frac{1}{2}\phi_1^{k_n}(0)\phi_2^{k_n}(0) \right)$$

Table 3

Bias in the estimators $\hat{C}_t^{n,\text{adj}}$ and $\Gamma_t^{n,\text{adj}}$, for the three models.

|  | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Relative small-sample bias $\text{Avg}[(\hat{C}_t^{n,\text{adj}} - C)/C]$ | −0.000292 | −0.001169 | 0.000039 |
| Relative bias in the variance estimator $\text{Avg}[(\Gamma_t^{n,\text{adj}} - \int_0^t \gamma_s^2 ds)/\int_0^t \gamma_s^2 ds]$ | 0.000525 | 0.000904 | 0.001653 |

$$\Phi_{22}^{k_n} = \frac{1}{k_n^3}\left(\sum_{j=0}^{k_n-1}(\phi_2^{k_n}(j))^2 - \frac{1}{2}(\phi_2^{k_n}(0))^2\right).$$

As it can be seen in the proof, these parameters are the "correct" ones, but each of them converges at a smaller order than $n^{-\frac{1}{4}}$ and can therefore be replaced in the central limit theorem. Nevertheless, for small sizes of $k_n$ the difference between each of the parameters and its limit turns out to be substantial. A second adjustment regards the sums appearing in the estimators. The numbers of summands are implicitly assumed to be $\lfloor t/\Delta_n \rfloor$ rather than $\lfloor t/\Delta_n \rfloor - k_n + 2$, say. This doesn't matter in the limit, but it is reasonable to scale each sum by $\lfloor t/\Delta_n \rfloor$ divided by the actual number of summands to obtain better results. However, this adjustment is of minor importance. The last step is a finite sample bias correction due to the fact that $\sum_{i=1}^{\lfloor t/\Delta_n \rfloor}(\Delta_j^n X)^2$ converges to $C_t$. Therefore, the latter term in $\hat{C}_t^n$ gives a small negative bias, which we dispose of by another scaling factor. Summarized, the new statistics can be defined as follows:

$$\hat{C}_t^{n,\text{adj}} = \left(1 - \frac{\psi_1^{k_n}\Delta_n}{2\theta^2\psi_2^{k_n}}\right)^{-1}\left(\frac{\lfloor t/\Delta_n \rfloor\sqrt{\Delta_n}}{(\lfloor t/\Delta_n \rfloor - k_n + 2)\theta\psi_2^{k_n}}\sum_{i=0}^{\lfloor t/\Delta_n \rfloor - k_n + 1}(\bar{Z}_i^n)^2\right.$$
$$\left. - \frac{\psi_1^{k_n}\Delta_n}{2\theta^2\psi_2^{k_n}}\sum_{i=1}^{\lfloor t/\Delta_n \rfloor}(\Delta_j^n X)^2\right)$$

and

$$\Gamma_t^{n,\text{adj}} = \left(1 - \frac{\psi_1^{k_n}\Delta_n}{2\theta^2\psi_2^{k_n}}\right)^{-2}\left(\frac{4\Phi_{22}^{k_n}\lfloor t/\Delta_n \rfloor}{3\theta(\psi_2^{k_n})^4(\lfloor t/\Delta_n \rfloor - k_n + 2)}\sum_{i=0}^{\lfloor t/\Delta_n \rfloor - k_n + 1}(\bar{Z}_i^n)^4\right.$$
$$+ \frac{4\Delta_n\lfloor t/\Delta_n \rfloor}{\theta^3(\lfloor t/\Delta_n \rfloor - k_n + 2)}\left(\frac{\Phi_{12}^{k_n}}{(\psi_2^{k_n})^3} - \frac{\Phi_{22}^{k_n}\psi_1^{k_n}}{(\psi_2^{k_n})^4}\right)$$
$$\times \sum_{i=0}^{\lfloor t/\Delta_n \rfloor - k_n + 1}(\bar{Z}_i^n)^2\sum_{j=i+k_n}^{i+2k_n-1}(\Delta_j^n Z)^2$$
$$+ \frac{\Delta_n\lfloor t/\Delta_n \rfloor}{\theta^3(\lfloor t/\Delta_n \rfloor - 2)}\left(\frac{\Phi_{11}^{k_n}}{(\psi_2^{k_n})^2} - \frac{2\Phi_{12}^{k_n}\psi_1^{k_n}}{(\psi_2^{k_n})^3} + \frac{\Phi_{22}^{k_n}(\psi_1^{k_n})^2}{(\psi_2^{k_n})^4}\right)$$
$$\times \sum_{i=1}^{\lfloor t/\Delta_n \rfloor - 2}(\Delta_i^n Z)^2(\Delta_{i+2}^n Z)^2\right).$$

Table 3 reports the performance of the adjusted estimator $\hat{C}_t^{n,\text{adj}}$ and the variance estimator $\Gamma_t^{n,\text{adj}}$.

Table 4

Comparisons of quantiles of $N_t^n$, $N_t^{n,\mathrm{adj}}$ with $\mathcal{N}(0, 1)$.

|  | Mean | Stdv. | 0.5% | 2.5% | 5% | 95% | 97.5% | 99.5% |
|---|---|---|---|---|---|---|---|---|
| Model 1 $N_t^n$ | −0.22 | 1.04 | 1.26% | 4.72% | 8.32% | 96.86% | 98.56% | 99.82% |
| Model 1 $N_t^{n,\mathrm{adj}}$ | −0.05 | 1.02 | 0.82% | 3.20% | 6.08% | 95.55% | 97.94% | 99.68% |
| Model 2 $N_t^n$ | −0.24 | 1.06 | 1.76% | 5.03% | 8.77% | 96.92% | 98.67% | 99.93% |
| Model 2 $N_t^{n,\mathrm{adj}}$ | −0.07 | 1.04 | 1.20% | 3.66% | 6.42% | 95.73% | 98.03% | 99.85% |
| Model 3 $N_t^n$ | −0.21 | 1.05 | 1.32% | 4.86% | 8.41% | 96.8% | 98.66% | 99.82% |
| Model 3 $N_t^{n,\mathrm{adj}}$ | −0.05 | 1.03 | 0.84% | 3.42% | 6.24% | 95.58% | 97.99% | 99.73% |

Table 5

Comparisons of quantiles of $N0_t^n$, $N0_t^{n,\mathrm{adj}}$ with $\mathcal{N}(0, 1)$.

|  | Mean | Stdv. | 0.5% | 2.5% | 5% | 95% | 97.5% | 99.5% |
|---|---|---|---|---|---|---|---|---|
| Model 1 $N0_t^n$ | −0.17 | 1.03 | 0.64% | 3.27% | 6.75% | 95.75% | 97.70% | 99.47% |
| Model 1 $N0_t^{n,\mathrm{adj}}$ | −0.01 | 1.04 | 0.40% | 2.21% | 4.81% | 94.26% | 96.78% | 99.16% |
| Model 2 $N0_t^n$ | −0.19 | 1.04 | 1.00% | 3.74% | 7.10% | 95.82% | 97.70% | 99.53% |
| Model 2 $N0_t^{n,\mathrm{adj}}$ | −0.02 | 1.05 | 0.63% | 2.72% | 5.07% | 94.20% | 96.74% | 99.12% |
| Model 3 $N0_t^n$ | −0.16 | 1.04 | 0.65% | 3.50% | 6.85% | 95.71% | 97.79% | 99.53% |
| Model 3 $N0_t^{n,\mathrm{adj}}$ | 0.00084 | 1.05 | 0.38% | 2.40% | 4.94% | 94.06% | 96.76% | 99.16% |

We test the normality of the statistics $N_t^n = \dfrac{\hat{C}_t^n - C}{\Delta_n^{1/4}\sqrt{\Gamma_t^n}}$ and $N_t^{n,\mathrm{adj}} = \dfrac{\hat{C}_t^{n,\mathrm{adj}} - C}{\Delta_n^{1/4}\sqrt{\Gamma_t^{n,\mathrm{adj}}}}$, whose quantiles are compared with the $\mathcal{N}(0, 1)$ quantiles (see Table 4):

One of the reasons that the above quantiles don't look good enough is that there is a (small) positive correlation between the estimator $\hat{C}_t^n$ and $\Gamma_t^n$. One can adjust this effect by using a first order Taylor expansion of $\Gamma$: expanding $\Gamma_t^n$ or $\Gamma_t^{n,\mathrm{adj}}$ around the theoretical asymptotic variance $\Gamma_0 \left( \dfrac{1}{\sqrt{\Gamma_0}} \approx \dfrac{1}{\sqrt{\Gamma_t^n}} - \dfrac{\Gamma_0 - \Gamma_t^n}{2\Gamma_0^{3/2}} \right)$:

$$N0_t^n := N_t^n - \frac{(\Gamma_0 - \Gamma_t^n)(\hat{C}_t^n - C)}{2\Delta_n^{1/4}\Gamma_0^{3/2}}$$

and

$$N0_t^{n,\mathrm{adj}} := N_t^{n,\mathrm{adj}} - \frac{(\Gamma_0 - \Gamma_t^{n,\mathrm{adj}})(\hat{C}_t^{n,\mathrm{adj}} - C)}{2\Delta_n^{1/4}\Gamma_0^{3/2}}.$$

The quantiles of $N0_t^n$ and $N0_t^{n,\mathrm{adj}}$ are compared with the $\mathcal{N}(0, 1)$ quantiles (see Table 5):

We see from the above simulation results that our estimator works quite well for these models. We note that our theoretical results do not require the often made assumption that the rounding threshold go to zero, and this is reflected in the simulation when comparing results of model 2 with those of model 1. Comparing results of model 3 with those of model 1, we see that our approach works well for the stochastic volatility model (see Figs. 4 and 5).
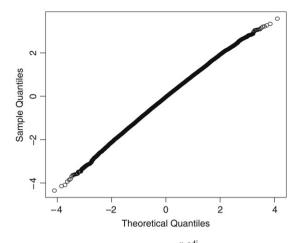
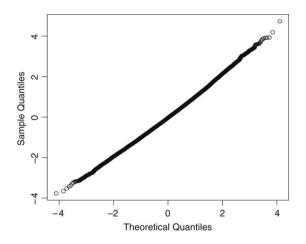Fig. 4. Normal $Q$–$Q$ plot of $N_t^{n,\text{adj}}$ for Model 1.



Fig. 5. Normal $Q$–$Q$ plot of $N0_t^{n,\text{adj}}$ for Model 1.

## 5. The proof

To begin with, we introduce a strengthened version of our assumptions (H) and (K):

**Assumption (L).** We have (H) and (K), and further the processes $b$, $\sigma$, $\int z^8 Q_t(\mathrm{d}z)$ and $X$ itself are bounded (uniformly in $(\omega, t)$) (then $\alpha$ is also bounded).    □

Then a standard localization procedure explained in details in [4] for example shows that for proving Theorem 3.1 it is no restriction to assume that (L) holds. Below, we assume these stronger assumptions without further mention.

There are two separate parts in the proof. One consists in replacing in (3.6) the observed process $Z$ by the unobserved $X$, at the cost of additional terms which involve the quadratic mean error process $\alpha$ of (2.5). The other part amounts to a central limit theorem for the sums of the variables $(\overline{X}_i^n)^2$. This is not completely standard because $(\overline{X}_i^n)^2$ and $(\overline{X}_j^n)^2$ are strongly dependent when $|i - j| < k_n$, since they involve some common variables $X_l^n$, whereas $k_n \to \infty$. So for this

we split the sum $\sum_{i=0}^{[t/\Delta_n]-k_n+1}(\overline{X}_i^n)^2$ into "big" blocks of length $pk_n$, with $p$ eventually going to $\infty$, separated by "small" blocks of length $k_n$, which are eventually negligible but ensure the conditional independence between the big blocks which we need for the central limit theorem.

Obviously, this scheme asks for somehow involved notation, which we present all together in the next subsection.

## 5.1. Some notation

First, $K$ denotes a constant which changes from line to line and may depend on the bounds of the various processes in (L), and also on $\sup_n k_n^2 \Delta_n$ (recall (3.1)), and is written $K_r$ if it depends on an additional parameter $r$. We also write $O_u(x)$ for a (possibly random) quantity smaller than $Kx$ for some constant $K$ as above.

In the following, and unless otherwise stated, $p \geq 1$ denotes an integer and $q > 0$ a real. For each $n$ we introduce the function

$$g_n(s) = \sum_{j=1}^{k_n-1} g_j^n \, 1_{((j-1)\Delta_n, j\Delta_n]}(s), \tag{5.1}$$

which vanishes for $s > (k_n - 1)\Delta_n$ and $s \leq 0$ and is bounded uniformly in $n$. We then introduce the processes

$$\left.\begin{array}{l} X(n,s)_t = \displaystyle\int_0^t b_u g_n(u-s)\mathrm{d}u + \int_0^t \sigma_u g_n(u-s)\mathrm{d}W_u \\[2mm] C(n,s)_t = \displaystyle\int_0^t \sigma_u^2 \, g_n(u-s)^2 \, \mathrm{d}u. \end{array}\right\} \tag{5.2}$$

These processes vanish for $t \leq s$, and are constant in time for $t \geq s + (k_n - 1)\Delta_n$, and

$$\overline{X}_i^n = X(n, i\Delta_n)_{(i+k_n)\Delta_n}, \qquad c_i^n := \sum_{j=1}^{k_n-1}(g_j^n)^2 \Delta_{i+j}^n C = C(n, i\Delta_n)_{(i+k_n)\Delta_n}. \tag{5.3}$$

Next, we set

$$A_{i,j}^n = \sum_{m=i\vee j}^{i\wedge j+k_n-1} h_{m-i}^n h_{m-j}^n \alpha_m^n, \qquad A_i^n = A_{i,i}^n = \sum_{m=0}^{k_n-1}(h_m^n)^2 \alpha_{i+m}^n. \tag{5.4}$$

$$\widetilde{Z}_i^{n\prime} = (\overline{Z}_i^n)^2 - A_i^n - c_i^n, \qquad \zeta(Z,p)_i^n = \sum_{j=i}^{i+pk_n-1} \widetilde{Z}_j^{n\prime}, \tag{5.5}$$

$$\zeta(X,p)_i^n = \sum_{j=i}^{i+pk_n-1}\left((\overline{X}_j^n)^2 - c_j^n\right), \qquad \zeta(W,p)_i^n = \sum_{j=i}^{i+pk_n-1}\left((\sigma_i^n \overline{W}_j^n)^2 - c_j^n\right), \tag{5.6}$$

(note the differences in the definition of $\zeta(V,p)_i^n$ when $V = Z$ or $V = X$ or $V = W$). Moreover for any process $V$ we set

$$\zeta'(V,p)_i^n = \sum_{(j,m):\, i\leq j<m\leq i+pk_n-1} \overline{V}_j^n \overline{V}_m^n \, \phi_1\left(\frac{m-j}{k_n}\right), \tag{5.7}$$

$$\zeta''(V)_i^n = (\overline{V}_i^n)^2 \sum_{j=i+k_n}^{i+2k_n-1} (\Delta_j^n V)^2. \tag{5.8}$$

Next we consider the discrete time filtrations $\mathcal{F}_j^n = \mathcal{F}_{j\Delta_n}^{(0)} \otimes \mathcal{F}_{j\Delta_n-}^{(1)}$ (that is, generated by all $\mathcal{F}_{j\Delta_n}^{(0)}$-measurable variables plus all variables $Z_s$ for $s < j\Delta_n$ and $\mathcal{F}_j'^n = \mathcal{F}^{(0)} \otimes \mathcal{F}_{j\Delta_n-}^{(1)}$ and $\mathcal{G}(p)_j^n = \mathcal{F}_{j(p+1)k_n}^n$ and $\mathcal{G}'(p)_j^n = \mathcal{F}_{j(p+1)k_n+pk_n}^n$, for $j \in \mathbb{N}$, and we introduce the variables

$$\left.\begin{array}{ll} \eta(p)_j^n = \dfrac{\sqrt{\Delta_n}}{\theta \psi_2} \zeta(Z,p)_{j(p+1)k_n}^n, & \overline{\eta}(p)_j^n = \mathbb{E}(\eta(p)_j^n \mid \mathcal{G}(p)_j^n) \\[3mm] \eta'(p)_j^n = \dfrac{\sqrt{\Delta_n}}{\theta \psi_2} \zeta(Z,1)_{j(p+1)k_n+pk_n}^n, & \overline{\eta}'(p)_j^n = \mathbb{E}(\eta'(p)_j^n \mid \mathcal{G}'(p)_j^n). \end{array}\right\} \tag{5.9}$$

Then $j_n(p,t) = \left[ \frac{t+\Delta_n}{(p+1)k_n\Delta_n} \right] - 1$ is the maximal number of pairs of "blocks" of respective sizes $pk_n$ and $k_n$ that can be accommodated without using data after time $t$, and we set

$$\left.\begin{array}{ll} F(p)_t^n = \displaystyle\sum_{j=0}^{j_n(p,t)} \overline{\eta}(p)_j^n, & M(p)_t^n = \displaystyle\sum_{j=0}^{j_n(p,t)} (\eta(p)_j^n - \overline{\eta}(p)_j^n) \\[6mm] F'(p)_t^n = \displaystyle\sum_{j=0}^{j_n(p,t)} \overline{\eta}'(p)_j^n, & M'(p)_t^n = \displaystyle\sum_{j=0}^{j_n(p,t)} (\eta'(p)_j^n - \overline{\eta}'(p)_j^n). \end{array}\right\} \tag{5.10}$$

With the notation $i_n(p,t) = (j_n(p,t)+1)(p+1)k_n$, we also have three "residual" processes:

$$\widehat{C}(p)_t^n = \frac{\sqrt{\Delta_n}}{\theta \psi_2} \sum_{i=i_n(p,t)}^{[t/\Delta_n]-k_n+1} \widetilde{Z}_i^{n\prime}, \tag{5.11}$$

$$\widehat{C}'(p)_t^n = \frac{\sqrt{\Delta_n}}{\theta \psi_2} \sum_{i=0}^{[t/\Delta_n]-k_n+1} A_i^n - \frac{\psi_1 \Delta_n}{2\theta^2 \psi_2} \sum_{i=1}^{[t/\Delta_n]} (\Delta_i^n Z)^2, \tag{5.12}$$

$$\widehat{C}_t''^n = \frac{\sqrt{\Delta_n}}{\theta \psi_2} \sum_{i=0}^{[t/\Delta_n]-k_n+1} c_i^n - C_t. \tag{5.13}$$

The key point of all this notation is the following identity, valid for all $p \geq 1$:

$$\widehat{C}_t^n - C_t = M(p)_t^n + M'(p)_t^n + F(p)_t^n + F'(p)_t^n + \widehat{C}(p)_t^n + \widehat{C}'(p)_t^n + \widehat{C}_t''^n. \tag{5.14}$$

We end this subsection with some miscellaneous notation:

$$\left.\begin{array}{l} \beta(p)_i^n = \displaystyle\sup_{s,t\in[i\Delta_n,(i+(p+2)k_n)\Delta_n]} (|b_s - b_t| + |\sigma_s - \sigma_t| + |\alpha_s - \alpha_t|) \\[3mm] \chi(p)_i^n = \Delta_n^{1/4} + \sqrt{\mathbb{E}((\beta(p)_i^n)^2 \mid \mathcal{F}_i^n)}. \end{array}\right\} \tag{5.15}$$

$$\Xi_{ij} = -\int_0^1 s\phi_i(s)\phi_j(s)\,\mathrm{d}s. \tag{5.16}$$

## 5.2. Estimates for the Wiener process

This subsection is devoted to proving the following result about the Wiener process:

**Lemma 5.1.** *We have*

$$\mathbb{E}((\zeta(W, p)_i^n)^2 \mid \mathcal{F}_i^n) = 4(p\,\Phi_{22} + \Xi_{22})k_n^4 \Delta_n^2 (\sigma_i^n)^4 + \mathrm{O}_u(p^2 \chi(p)_i^n), \tag{5.17}$$

$$\mathbb{E}\left(\zeta'(W, p)_i^n \mid \mathcal{F}_i^n\right) = (p\,\Phi_{12} + \Xi_{12})k_n^3 \Delta_n + \mathrm{O}_u(p\Delta_n^{-1/4}). \tag{5.18}$$

**Proof.** (1) Since $g(0) = g(1) = 0$ we have $\int_0^1 g'(s)\mathrm{d}s = 0$. We introduce the process

$$U_t = -\int_0^1 g'(s)W_{t+s}\mathrm{d}s = -\int_t^{t+1} g'(s-t)W_s\mathrm{d}s = -\int_0^1 g'(s)(W_{t+s} - W_t)\mathrm{d}s, \tag{5.19}$$

which is stationary centered Gaussian with covariance $\mathbb{E}(U_t U_{t+s}) = \phi_2(s)$, as given by (3.4). The scaling property of $W$ and (3.5) and $g(0) = g(1) = 0$ imply that

$$\left(\overline{W}_i^n\right)_{i\geq 1} \overset{\mathcal{L}}{=} \left(-\sqrt{k_n \Delta_n} \sum_{j=0}^{k_n-1} \left(g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right)\right) W_{(i+j)/k_n}\right)_{i\geq 1}.$$

Then (3.2) and the fact that $\mathbb{E}(\sup_{u\in[0,s]} |W_{t+u} - W_t|^q) \leq K_q s^{q/2}$, plus a standard approximation of an integral by Riemann sums, yield

$$\left(\overline{W}_i^n\right)_{i\geq 1} \overset{\mathcal{L}}{=} \left(\sqrt{k_n \Delta_n} U_{i/k_n} + R_i^n\right)_{j\geq 0}, \qquad \mathbb{E}(|R_i^n|^q) \leq K_q \Delta_n^{q/2} \tag{5.20}$$

where the last estimate holds for all $q > 0$. Then in view of (3.1) we get for $j \geq i$:

$$\left.\begin{aligned}
\mathbb{E}\left(\overline{W}_i^n \overline{W}_j^n \mid \mathcal{F}_i^n\right) &= k_n \Delta_n \phi_2\left(\frac{j-i}{k_n}\right) + \mathrm{O}_u(\Delta_n^{3/4}) \\
\mathbb{E}\left((\overline{W}_i^n)^4 \mid \mathcal{F}_i^n\right) &= 3k_n^2 \Delta_n^2 \psi_2^2 + \mathrm{O}_u(\Delta_n^{5/4}).
\end{aligned}\right\} \tag{5.21}$$

At this stage, (5.18) is obvious.

(2) We have

$$(\zeta(W, p)_i^n)^2 = (\sigma_i^n)^4 V_n(i, p)^2 + V_n'(i, p)^2 - 2(\sigma_i^n)^2 V_n(i, p) V_n'(i, p), \tag{5.22}$$

where

$$V_n(i, p) = \sum_{j=i}^{i+pk_n-1} (\overline{W}_j^n)^2, \qquad V_n'(i, p) = \sum_{j=i}^{i+pk_n-1} c_j^n.$$

On the one hand, we deduce from (5.3) that if $i \leq j \leq i + (p+1)k_n$,

$$c_j^n = \psi_2 k_n \Delta_n (\sigma_i^n)^2 + \mathrm{O}_u(\Delta_n + \sqrt{\Delta_n}\,\beta(p)_i^n). \tag{5.23}$$

Then obviously

$$V_n'(i, p) = \psi_2(\sigma_i^n)^2 pk_n^2 \Delta_n + \mathrm{O}_u(p\sqrt{\Delta_n} + p\,\beta(p)_i^n). \tag{5.24}$$

On the other hand, another application of (5.20) and of the approximation of an integral by Riemann sums, plus the fact that $\mathbb{E}(\sup_{u\in[0,s]} |U_{t+u} - U_t|^q) \leq K_q s^q$ (this easily follows from (5.19)), yield for any $p \geq 1$:

$$V_n(i, p) \overset{\mathcal{L}}{=} k_n^2 \Delta_n \int_0^p (U_s)^2 \mathrm{d}s + \overline{R}(p)_i^n, \qquad \mathbb{E}(|\overline{R}(p)_i^n|^q) \leq K_q p^q \Delta_n^{q/4}. \tag{5.25}$$

Since $\mathbb{E}(U_t U_{t+s}) = \phi_2(s)$, that for $p \geq 2$ the variable $\overline{U}_p = \int_0^p (U_s)^2 \mathrm{d}s$ satisfies

$$\mathbb{E}(\overline{U}_p) = p\psi_2, \qquad \mathbb{E}(\overline{U}_p^2) = p^2 \psi_2^2 + 4p\,\Phi_{22} + 4\Xi_{22}.$$

Then (5.25) yields

$$\left.\begin{aligned}
&\mathbb{E}(V_n(i, p) \mid \mathcal{F}_i^n) = p k_n^2 \Delta_n \psi_2 + \mathrm{O}_u(p\Delta_n^{1/4})\\
&\mathbb{E}(V_n(i, p)^2 \mid \mathcal{F}_i^n) = (p^2\psi_2^2 + 4p\,\Phi_{22} + 4\,\Xi_{22})k_n^4 \Delta_n^2 + \mathrm{O}_u(p^2 \Delta_n^{1/4}).
\end{aligned}\right\} \tag{5.26}$$

Combining (5.24) and (5.26) with (5.22), we immediately get (5.17). $\square$

### 5.3. Estimates for the process $X$

Here we give estimates on the process $X$. The Assumption (L) implies that for all $s, t \geq 0$ and $q > 0$,

$$\left.\begin{aligned}
&\mathbb{E}\left(\sup_{u,v\in[t,t+s]} |X_u - X_v|^q \mid \mathcal{F}_t\right) \leq K_q\, s^{q/2}\\
&|\mathbb{E}(X_{t+s} - X_t \mid \mathcal{F}_t)| \leq Ks.
\end{aligned}\right\} \tag{5.27}$$

Then, since $|h_j^n| \leq K/k_n$ and $\sum_{j=0}^{k_n-1} h_j^n = 0$ for the second inequality below, we have

$$\mathbb{E}\left(|\Delta_{i+1}^n X|^q \mid \mathcal{F}_i^n\right) \leq K_q \Delta_n^{q/2}, \qquad \mathbb{E}\left(|\overline{X}_i^n|^q \mid \mathcal{F}_i^n\right) \leq K_q \Delta_n^{q/4}. \tag{5.28}$$

An elementary consequence is the following set of inequalities (use also $|c_i^n| \leq K\sqrt{\Delta_n}$ for the first one):

$$\mathbb{E}\left((\zeta(X, p)_i^n)^4 \mid \mathcal{F}_i^n\right) \leq K_p, \qquad \mathbb{E}\left(\zeta''(X)_i^n \mid \mathcal{F}_i^n\right) \leq K\Delta_n. \tag{5.29}$$

Here and below, as mentioned before, the constant $K_p$ depends on $p$, and it typically goes to $\infty$ as $p \to \infty$ (in this particular instance, we have $K_p = Kp^4$); what is important is that it does not depend on $n$, nor on $i$.

(5.29) is not enough, and we need more precise estimates on $\zeta(X, p)_i^n$ and $\zeta'(X, p)_i^n$, given in the following two lemmas.

**Lemma 5.2.** *We have*

$$\left|\mathbb{E}(\zeta(X, p)_i^n \mid \mathcal{F}_i^n)\right| \leq Kp\Delta_n^{1/4}\chi(p)_i^n. \tag{5.30}$$

**Proof.** Observe that, similar to (5.27),

$$\mathbb{E}\left(\sup_{t\geq 0} |X(n, s)_t|^q \mid \mathcal{F}_s\right) \leq K_q \Delta_n^{q/4}, \qquad |\mathbb{E}(X(n, s)_t \mid \mathcal{F}_s)| \leq K\sqrt{\Delta_n}. \tag{5.31}$$

Let us define the processes

$$\left.\begin{aligned}
&M(n, s)_t = 2\int_0^t X(n, s)_u \sigma_u g_n(u - s)\,\mathrm{d}W_u,\\
&B(n, s)_t = 2\int_0^t X(n, s)_u b_u g_n(u - s)\mathrm{d}u.
\end{aligned}\right\}$$

Then $M(n, s)$ is a martingale, and by Itô's formula $X(n, s)^2 = B(n, s) + C(n, s) + M(n, s)$. Hence, since $\mathbb{E}(\chi(1)_j^n \mid \mathcal{F}_i^n) \le \chi(p)_i^n$ when $i \le j \le i + (p + 1)k_n$, (5.30) is implied by

$$\left| \mathbb{E}(B(n, j\Delta_n)_{(i+k_n)\Delta_n} \mid \mathcal{F}_j^n) \right| \le K \Delta_n^{3/4} \chi(1)_j^n.$$

For this we write $B(n, i\Delta_n)_{(i+k_n)\Delta_n} = U_n + V_n$, where

$$U_n = b_j^n \int_{j\Delta_n}^{(j+k_n)\Delta_n} X(n, j\Delta_n)_u \, g_n(u - j\Delta_n) \, \mathrm{d}u,$$

$$V_n = \int_{j\Delta_n}^{(j+k_n)\Delta_n} X(n, j\Delta_n)_u \, (b_u - b_j^n) \, g_n(u - i\Delta_n) \, \mathrm{d}u.$$

On the one hand, the second part of (5.31) yields that $\left| \mathbb{E}(U_n \mid \mathcal{F}_j^n) \right| \le K \Delta_n \le K \Delta_n^{3/4} \chi(p)_j^n$. On the other hand, we have $|V_n| \le K\sqrt{\Delta_n} \, \beta(1)_i^n \, \sup_{t \ge 0} |X(n, j\Delta_n)_t|$, hence the first part of (5.31) and Cauchy–Schwarz inequality yield $\mathbb{E}(|V_n| \mid \mathcal{F}_j^n) \le K \Delta_n^{3/4} \chi(1)_j^n$, and the result follows.  □

**Lemma 5.3.** *We have*

$$\left. \begin{array}{l} \left| \mathbb{E}\left( (\zeta(X, p)_i^n)^2 \mid \mathcal{F}_i^n \right) - 4(p\,\Phi_{22} + \Xi_{22}) \, k_n^4 \Delta_n^2 (\sigma_i^n)^4 \right| \le K_p \chi(p)_i^n \\[2mm] \left| \mathbb{E}\left( \zeta'(p, X)_i^n \mid \mathcal{F}_i^n \right) - (p\,\Phi_{12} + \Xi_{12}) k_n^3 \Delta_n (\sigma_i^n)^2 \right| \le K_p \Delta_n^{-1/2} \chi(p)_i^n. \end{array} \right\} \tag{5.32}$$

**Proof.** The method is rather different from the previous lemma, and based upon the property that for $i\Delta_n \le t \le s \le (i + (p + 2)k_n)\Delta_n$ we have

$$\mathbb{E}\left( \sup_{u,v \in [t,t+s]} |X_u - X_v - \sigma_t(W_u - W_v)|^q \mid \mathcal{F}_i^n \right)$$
$$\le K_{p,q} s^{q/2} \left( s^{q/2} + \mathbb{E}((\beta(p)_i^n)^q \mid \mathcal{F}_i^n) \right).$$

We deduce that for $i \le j, l \le i + (p + 2)k_n$ we have

$$\mathbb{E}\left( \left| X_j^n - X_l^n - \sigma_t(W_j^n - W_l^n) \right|^q \mid \mathcal{F}_i^n \right) \le K_{p,q} \Delta_n^{q/4} \left( \Delta_n^{q/4} + \mathbb{E}((\beta(p)_i^n)^q \mid \mathcal{F}_i^n) \right). \tag{5.33}$$

Now, $\overline{V}_j^n = \sum_{l=0}^{k_n-1} h_l^n (V_{j+l}^n - V_j^n)$ and $|h_j^n| \le K/k_n$, by using Hölder inequality and (5.28) we get for $s$ a positive integer

$$\mathbb{E}\left( \left| (\overline{X}_j^n)^s - (\sigma_i^n \overline{W}_j^n)^s \right|^q \mid \mathcal{F}_i^n \right) \le K_{p,q,s} \Delta_n^{sq/4} \left( \Delta_n^{q/4} + \mathbb{E}((\beta(p)_i^n)^q \mid \mathcal{F}_i^n) \right). \tag{5.34}$$

By (5.7), this for $s = 1$ and $q = 2$, plus (5.28) and Cauchy–Schwarz inequality, yield

$$\mathbb{E}\left( \left| \zeta'(X, p)_i^n - (\sigma_i^n)^2 \zeta'(W, p)_i^n \right| \mid \mathcal{F}_i^n \right) \le K_p \Delta_n^{-1/2} \chi(p)_i^n.$$

In a similar way, and in view of (5.6), we apply (5.34) with $s = 2$ and $q = 2$ to get

$$\mathbb{E}\left( \left| \zeta(X, p)_i^n - \zeta(W, p)_i^n \right|^2 \mid \mathcal{F}_i^n \right) \le K_p \, (\chi(p)_i^n)^2, \tag{5.35}$$

which yields (use (5.29) and Cauchy–Schwarz inequality):

$$\mathbb{E}\left(\left|(\zeta(X,p)_i^n)^2 - (\zeta(W,p)_i^n)^2\right| \mid \mathcal{F}_i^n\right) \le K_p \, \chi(p)_i^n.$$

At this stage, the result readily follows from Lemma 5.1.　□

### 5.4. Estimates for the process Z

Now we turn to the observed process $Z$, and relate the moments of the variables $\overline{Z}_j^n$, conditional on $\mathcal{F}^{(0)}$, with the corresponding powers of $\overline{X}_j^n$. To begin with, and since $|h_j^n| \le K/k_n$ and $\alpha$ is bounded, and by the rate of approximation of the integral of a piecewise Lipschitz function by Riemann sums, the following properties are obvious:

$$\left.\begin{array}{l}
|A_{i,j}^n| \le K\sqrt{\Delta_n} \\[4pt]
|j - i| \ge k_n \quad \Rightarrow \quad A_{i,j}^n = 0 \\[4pt]
i \le j \le m \le i + (p+1)k_n \quad \Rightarrow \\[4pt]
A_{j,m}^n = \alpha_i^n \dfrac{1}{k_n} \phi_1\left(\dfrac{m-j}{k_n}\right) + O_u(p\Delta_n + \sqrt{\Delta_n}\,\beta(p)_i^n) \\[8pt]
\displaystyle\sum_{(j,m):\; i \le j < m \le i + pk_n - 1} (A_{j,m}^n)^2 = (\alpha_i^n)^2(p\,\Phi_{11} + \Xi_{11}) + O_u\left(p^3\sqrt{\Delta_n} + p\beta(p)_i^n\right).
\end{array}\right\} \tag{5.36}$$

Next, we give estimates for the $\mathcal{F}^{(0)}$-conditional expectations of various functions of $Z$. Because of the $\mathcal{F}^{(0)}$-conditional independence of the variables $Z_t - X_t$ for different values of $t$, and because of (2.3), the conditional expectation $\mathbb{E}((Z_t - X_t)(Z_s - X_s) \mid \mathcal{F}^{(0)} \otimes \mathcal{F}_{s-}^{(1)})$ vanishes if $s < t$ and equals $\alpha_t$ if $s = t$. Then, recalling (2.5) and (5.4),

$$\left.\begin{array}{l}
\mathbb{E}\left(\overline{Z}_i^n - \overline{X}_i^n \mid \mathcal{F}_i'^n\right) = 0 \\[8pt]
\mathbb{E}\left((\overline{Z}_i^n - \overline{X}_i^n)(\overline{Z}_j^n - \overline{X}_j^n) \mid \mathcal{F}_{i \wedge j}'^n\right) = A_{i,j}^n.
\end{array}\right\} \tag{5.37}$$

More generally, $\mathbb{E}\left(\Pi_{m=1}^q h_{j_m}^n (Z_{i+j_m}^n - X_{i+j_m}^n) \mid \mathcal{F}_{i+j_1}^n\right) = 0$ as soon as there is one $j_m$ which is different from all the others, and moreover $|h_j^n| \le K\sqrt{\Delta_n}$, whereas the moments (2.6) are bounded for $q \le 8$. Then if we write $(\overline{Z}_i^n - \overline{X}_i^n)^q$ as the sum of $\Pi_{m=1}^q h_{j_m}^n (Z_{i+j_m}^n - X_{i+j_m}^n)$ over all choices of integers $j_l$ between 0 and $k_n - 1$, we see that for $r, q$ integers we have

$$\begin{aligned}
&\mathbb{E}\left((\overline{Z}_i^n - \overline{X}_i^n)^q (\overline{Z}_j^n - \overline{X}_j^n)^r \mid \mathcal{F}_{i \wedge j}'^n\right) \\[4pt]
&= \begin{cases}
O_u(\Delta_n) & \text{if } q + r = 3 \\
A_i^n A_j^n + 2(A_{i,j}^n)^2 + O_u(\Delta_n^{3/2}) & \text{if } q = r = 2 \\
O_u(\Delta_n^2) & \text{if } q + r = 8.
\end{cases}
\end{aligned} \tag{5.38}$$

Now, if we expand the first members of (5.38), and in view of (5.36) and (5.5) and of $|c_i^n| \le K\sqrt{\Delta_n}$, we deduce from (5.37) and (5.38) that for $j \ge i$:

$$\left.\begin{array}{l}
\mathbb{E}\left(\widetilde{Z}_i^{n'} \mid \mathcal{F}_i'^n\right) = (\overline{X}_i^n)^2 - c_i^n, \qquad \mathbb{E}\left(|\widetilde{Z}_i^{n'}| \mid \mathcal{F}_i'^n\right) = (\overline{X}_i^n)^2 + O_u(\sqrt{\Delta_n}) \\[8pt]
\mathbb{E}\left(\widetilde{Z}_i^{n'} \widetilde{Z}_j^{n'} \mid \mathcal{F}^{(0)}\right) = ((\overline{X}_i^n)^2 - c_i^n)((\overline{X}_j^n)^2 - c_j^n) + 4\overline{X}_i^n \overline{X}_j^n A_{i,j}^n + 2(A_{i,j}^n)^2 \\[8pt]
\qquad\qquad + O_u\left(\Delta_n^{3/2} + \Delta_n|\overline{X}_i^n| + \Delta_n|\overline{X}_j^n|\right) \\[8pt]
\mathbb{E}\left((\widetilde{Z}_i^{n'})^4 \mid \mathcal{F}_i^n\right) \le K(\Delta_n^2 + |\overline{X}_i^n|^8).
\end{array}\right\} \tag{5.39}$$

Then obviously this, combined with (5.28) and (5.30), yields

$$
\left.
\begin{aligned}
&\mathbb{E}(\zeta(Z, p)_i^n \mid \mathcal{F}_i'^n) = \zeta(X, p)_i^n \\
&\mathbb{E}((\zeta(Z, p)_i^n)^4 \mid \mathcal{F}_i^n) \le K_p \\
&\left| \mathbb{E}(\zeta(Z, p)_i^n \mid \mathcal{F}_i'^n) \right| \le K_p \Delta_n^{1/4} \chi(p)_i^n
\end{aligned}
\right\}
\tag{5.40}
$$

and also, in view of (5.36),

$$
\mathbb{E}((\zeta(Z, p)_i^n)^2 \mid \mathcal{F}_i'^n) = (\zeta(X, p)_i^n)^2 + \frac{8}{k_n} \alpha_i^n \, \zeta'(X, p)_i^n + 4(\alpha_i^n)^2 (p\,\Phi_{11} + \Xi_{11})
$$
$$
+ p^3 \mathrm{O}_u \left( \left( \sqrt{\Delta_n} + \beta(p)_i^n \right) \left( 1 + \sum_{j=i}^{i+pk_n-1} |\overline{X}_j^n|^2 \right) \right).
$$

Then, using (5.28) again and (5.32) and Hölder inequality, we get

$$
\left| \mathbb{E}((\zeta(Z, p)_i^n)^2 \mid \mathcal{F}_i^n) - 4(p\,\Phi_{22} + \Xi_{22}) k_n^4 \Delta_n^2 (\sigma_i^n)^4 \right.
$$
$$
\left. - 8\alpha_i^n (\sigma_i^n)^2 (p\,\Phi_{12} + \Xi_{12}) k_n^2 \Delta_n - 4(\alpha_i^n)^2 (p\,\Phi_{11} + \Xi_{11}) \right| \le K_p \, \chi(p)_i^n.
\tag{5.41}
$$

We need some other estimates. Exactly as for (5.39) one sees that

$$
\left.
\begin{aligned}
&\mathbb{E}\left( (\overline{Z}_i^n)^4 \mid \mathcal{F}_i'^n \right) = (\overline{X}_i^n)^4 + 6(\overline{X}_i^n)^2 A_i^n + 3(A_i^n)^2 + \mathrm{O}_u \left( \Delta_n^{3/2} + \Delta_n |\overline{X}_i^n| \right) \\
&\mathbb{E}\left( (\overline{Z}_i^n)^8 \mid \mathcal{F}_i'^n \right) \le K(\Delta_n^2 + |\overline{X}_i^n|^8)
\end{aligned}
\right\}
\tag{5.42}
$$

and (using the boundedness of $X$)

$$
\left.
\begin{aligned}
&\mathbb{E}\left( \zeta''(Z)_i^n \mid \mathcal{F}_i'^n \right) = \zeta''(X)_i^n + A_i^n \sum_{j=i+k_n+1}^{i+2k_n} (\Delta_j^n X)^2 \\
&\qquad + ((\overline{X}_i^n)^2 + A_i^n) \sum_{j=i+k_n+1}^{i+2k_n} (\alpha_{j-1}^n + \alpha_j^n), \\
&\mathbb{E}((\zeta''(Z)_i^n)^2 \mid \mathcal{F}_i'^n) \le K.
\end{aligned}
\right\}
\tag{5.43}
$$

Therefore, using (5.28), (5.29), (5.36), (5.21) and (5.34) with $s = 2$, we obtain

$$
\left| \mathbb{E}((\overline{Z}_i^n)^4 \mid \mathcal{F}_i^n) - 3k_n^2 \Delta_n^2 \psi_2^2 (\sigma_i^n)^4 - 6\Delta_n (\sigma_i^n)^2 \alpha_i^n \psi_1 \psi_2 - \frac{3}{k_n^2} (\alpha_i^n)^2 \psi_1^2 \right|
$$
$$
\le K \Delta_n \chi(1)_i^n
\tag{5.44}
$$

$$
\left| \mathbb{E}(\zeta''(Z)_i^n \mid \mathcal{F}_i^n) - 2\alpha_i^n (\psi_1 \alpha_i^n + \psi_2 k_n^2 \Delta_n (\sigma_i^n)^2 \alpha_i^n) \right| \le K \chi(1)_i^n.
\tag{5.45}
$$

Finally, the following is obtained in the same way, but it is much simpler:

$$
\left.
\begin{aligned}
&\mathbb{E}((\Delta_{i+1}^n Z)^2 \mid \mathcal{F}_i'^n) = (\Delta_{i+1}^n X)^2 + \alpha_i^n + \alpha_{i+1}^n \\
&\left| \mathbb{E}((\Delta_{i+1}^n Z)^2 (\Delta_{i+3}^n Z)^2 \mid \mathcal{F}_i^n) - 4(\alpha_i^n)^2 \right| \le K \chi(1)_i^n \\
&\mathbb{E}((\Delta_{i+1}^n Z)^4 (\Delta_{i+3}^n Z)^4 \mid \mathcal{F}_i^n) \le K.
\end{aligned}
\right\}
\tag{5.46}
$$

## 5.5. Proof of the theorem

We begin the proof of Theorem 3.1 with an auxiliary technical result.

**Lemma 5.4.** *For any $p \geq 1$ we have*

$$\left.\begin{array}{l} \mathbb{E}\left(\sqrt{\Delta_n} \sum_{j=0}^{j_n(p,t)} \sqrt{\mathbb{E}((\beta(p)_{j(p+1)k_n}^n)^2 \mid \mathcal{F}_{j(p+1)k_n}^n)}\right) \to 0 \\[2.5em] \mathbb{E}\left(\sqrt{\Delta_n} \sum_{j=0}^{j_n(p,t)} (\beta(p)_{j(p+1)k_n}^n)^2\right) \to 0. \end{array}\right\} \tag{5.47}$$

**Proof.** We have $j_n(p,t) \leq K_p t / \sqrt{\Delta_n}$. Then the first expression in (5.47) is smaller than a constant times the square-root of the second expression, and thus for (5.47) it suffices to prove that

$$\mathbb{E}\left(\sqrt{\Delta_n} \sum_{j=0}^{j_n(p,t)} (\beta(p)_{j(p+1)k_n}^n)^2\right) \to 0. \tag{5.48}$$

Let $\varepsilon > 0$ and denote by $N(\varepsilon)_t$ the number of jumps of any of the three processes $b$, $\sigma$ or $\alpha$, with size bigger than $\varepsilon$, over the interval $[0, t]$, and set $\rho(\varepsilon, t, \eta)$ to be the supremum of $|b_s - b_r| + |\sigma_s - \sigma_r| + |\alpha_s - \alpha_r|$ over all pairs $(s, r)$ such that $s \leq r \leq s + \eta \leq t$ and such that the interval $(s, r]$ contains no jump of $b$, $\sigma$ or $\alpha$ of size bigger than $\varepsilon$. Then obviously, since all three processes $b, \sigma, \alpha$ are bounded,

$$\sqrt{\Delta_n} \sum_{j=0}^{j_n(p,t)} (\beta(p)_{j(p+1)k_n}^n)^2 \leq \left(K N(\varepsilon)_t \sqrt{\Delta_n}\right) \wedge (K_p t) + K_p\, t\, \rho(\varepsilon, t, (p+1)k_n \Delta_n)^2.$$

Moreover $\limsup_{\eta \to 0} \rho(\varepsilon, t, \eta) \leq 3\varepsilon$. Then Fatou's lemma yields that the $\limsup$ of the left side of (5.48) is smaller than $K_p t \varepsilon^2$, and the result follows.  $\square$

The proof of the first part of the theorem is based on the identity (5.14), valid for all integers $p \geq 1$. The right side of this decomposition contains two "main" terms $M(p)_t^n$ and $M'(p)_t^n$, and all others are taken care of in Lemmas 5.5 and 5.6 below:

**Lemma 5.5.** *For any fixed $p \geq 1$ we have:*

$$\Delta_n^{-1/4}\, F(p)_t^n \xrightarrow{\mathbb{P}} 0 \tag{5.49}$$

$$\Delta_n^{-1/4}\, F'(p)_t^n \xrightarrow{\mathbb{P}} 0 \tag{5.50}$$

$$\Delta_n^{-1/4}\, \widehat{C}(p)_t^n \xrightarrow{\mathbb{P}} 0 \tag{5.51}$$

$$\Delta_n^{-1/4}\, \widehat{C}_t''^n \xrightarrow{\mathbb{P}} 0. \tag{5.52}$$

**Proof.** In view of (5.9) and (5.10), the proof of both (5.49) and (5.50) is a trivial consequence of (5.40) and of Lemmas 5.2 and 5.4. Since the right side of (5.11) contains at most $K_p / \sqrt{\Delta_n}$ summands, each one having expectation less than $K \Delta_n^{1/2}$ by the last part of (5.39) and (5.28), we immediately get (5.51).

In view of (5.3), and with the notation $a_n = \sum_{j=1}^{k_n-1}(g_j^n)^2$, we see that

$$
\sum_{i=0}^{[t/\Delta_n]-k_n+1} c_i^n = \sum_{i=0}^{[t/\Delta_n]-k_n+1} \sum_{l=i+1}^{i+k_n-1} (g_{l-i}^n)^2 \Delta_l^n C
$$

$$
= \sum_{l=1}^{[t/\Delta_n]} \Delta_l^n C \sum_{j=1\vee(l+k_n-1-[t/\Delta_n])}^{l\wedge(k_n-1)} (g_j^n)^2 = a_n \sum_{l=k_n-1}^{[t/\Delta_n]-k_n+2} \Delta_l^n C + O_u(1).
$$

It follows that $\widehat{C}_t'^n = \left(\frac{\sqrt{\Delta_n}}{\theta\psi_2} a_n - 1\right) + O_u(\sqrt{\Delta_n})$. Since by Riemann approximation we have $a_n = k_n\psi_2 + O_u(1)$, we readily deduce (5.52) from (3.1).  □

**Lemma 5.6.** *For any fixed $p \geq 1$ we have $\Delta_n^{-1/4} \widehat{C}'(p)_t^n \xrightarrow{\mathbb{P}} 0$.*

**Proof.** Let $\zeta_i^n = (\Delta_i^n Z)^2 - (\alpha_{i-1}^n + \alpha_i^n)$. We get by (5.28) and (5.46), and for $1 \leq i \leq j-2$:

$$
\mathbb{E}(\zeta_i^n) = \mathbb{E}((\Delta_i^n X)^2) = O_u(\Delta_n), \qquad \mathbb{E}(\zeta_i^n \zeta_j^n) = \mathbb{E}((\Delta_i^n X)^2(\Delta_j^n X)^2) = O_u(\Delta_n^2),
$$

and also $\mathbb{E}(|\zeta_i^n|^2) \leq K$. Then obviously $\mathbb{E}\left(\left(\sum_{i=1}^{[t/\Delta_n]} \zeta_i^n\right)^2\right) \leq K/\Delta_n$, and it follows that

$$
G_n := \frac{\psi_1 \Delta_n^{3/4}}{2\theta^2\psi_2} \sum_{i=1}^{[t/\Delta_n]} \zeta_i^n \xrightarrow{\mathbb{P}} 0.
$$

It is then enough to prove that $\frac{1}{\Delta_n^{1/4}} \widehat{C}'(p)_t^n + G_n \xrightarrow{\mathbb{P}} 0$. Observe that by an elementary calculation, $\frac{1}{\Delta_n^{1/4}} \widehat{C}'(p)_t^n + G_n = U_n + V_n$, where

$$
U_n = \left(\frac{\Delta_n^{1/4}}{\theta\psi_2}\left(\sum_{l=0}^{k_n-1}(h_l^n)^2\right) - \frac{\psi_1\Delta_n^{3/4}}{\theta^2\psi_2}\right)\left(\sum_{i=k_n}^{i_n(p,t)-1} \alpha_i^n\right),
$$

$$
V_n = \frac{\Delta_n^{1/4}}{\theta\psi_2}\left(\sum_{i=0}^{k_n-1} \alpha_i^n \sum_{l=0}^{i}(h_l^n)^2 + \sum_{i=i_n(p,t)}^{i_n(p,t)+k_n-2} \alpha_i^n \sum_{l=i+1-i_n(p,t)}^{k_n-1}(h_l^n)^2\right)
$$

$$
- \frac{\psi_1\Delta_n^{3/4}}{2\theta^2\psi_2}\left(\alpha_0^n + 2\sum_{i=1}^{k_n-1}\alpha_i^n + 2\sum_{i=i_n(p,t)}^{[t/\Delta_n]-1}\alpha_i^n + \alpha_{[t/\Delta_n]}^n\right).
$$

On the one hand, since $\alpha_t$ is bounded and $|h_l^n| \leq K\sqrt{\Delta_n}$ it is obvious that $|V_n| \leq K\Delta_n^{1/4}$. On the other hand, $\sum_{l=0}^{k_n-1}(h_l^n)^2 = \frac{\psi_1}{k_n} + O(\Delta_n)$, whereas $\sum_{i=k_n}^{i_n(p,t)-1} \alpha_i^n \leq K/\Delta_n$, so by (3.1)), we see that $U_n \to 0$ pointwise. Then it finishes the proof.  □

Now we study the main terms $M(p)_t^n$ and $M'(p)_t^n$ in (5.14). Those terms are (discretised) sums of martingale differences (note that $\eta(p)_j^n$ and $\eta'(p)_j^n$ are measurable with respect to $\mathcal{G}(p)_{j+1}^n$ and $\mathcal{G}'(p)_{j+1}^n$ respectively).

By Doob's inequality we have

$$
\mathbb{E}\left(\sup_{s\leq t} |M'(p)_s^n|^2\right) \leq 4 \sum_{j=0}^{j_n(p,t)} \mathbb{E}(|\eta'(p)_j^n|^2).
$$

Now, (5.41) for $p = 1$ and the boundedness of $\chi(1)_i^n$ imply $\mathbb{E}(|\eta'(p)_j^n|^2) \leq K\Delta_n$, and thus (recall $j_n(p, t) \leq Kt/p\sqrt{\Delta_n}$):

$$\mathbb{E}\left(\sup_{s \leq t} |M'(p)_s^n|^2\right) \leq \frac{Kt}{p}\sqrt{\Delta_n}. \tag{5.53}$$

**Lemma 5.7.** *For any fixed $p \geq 2$, the sequence $\frac{1}{\Delta_n^{1/4}} M(p)^n$ of processes converges stably in law to*

$$Y(p)_t = \int_0^t \gamma(p)_s \mathrm{d}B_s, \tag{5.54}$$

*where $B$ is like in Theorem 3.1 and $\gamma(p)_t$ is the square root of*

$$\gamma(p)_t^2 = \frac{4}{\psi_2^2}\left(\left(\frac{p}{p+1}\Phi_{22} + \frac{1}{p+1}\Psi_{22}\right)\theta\sigma_t^4 + 2\left(\frac{p}{p+1}\Phi_{12} + \frac{1}{p+1}\Psi_{12}\right)\frac{\sigma_t^2\alpha_t}{\theta}\right.$$
$$\left. + \left(\frac{p}{p+1}\Phi_{11} + \frac{1}{p+1}\Psi_{11}\right)\frac{\alpha_t^2}{\theta^3}\right). \tag{5.55}$$

**Proof.** (1) In view of a standard limit theorem for triangular arrays of martingale differences, it suffices to prove the following three convergences:

$$\frac{1}{\sqrt{\Delta_n}}\sum_{j=0}^{j_n(p,t)}\left(\mathbb{E}((\eta(p)_j^n)^2 \mid \mathcal{G}(p)_j^n) - (\overline{\eta}(p)_j^n)^2\right) \xrightarrow{\mathbb{P}} \int_0^t \gamma(p)_s^2 \, \mathrm{d}s, \tag{5.56}$$

$$\frac{1}{\Delta_n}\sum_{j=0}^{j_n(p,t)}\mathbb{E}((\eta(p)_j^n)^4 \mid \mathcal{G}(p)_j^n) \xrightarrow{\mathbb{P}} 0, \tag{5.57}$$

$$\frac{1}{\Delta_n^{1/4}}\sum_{j=0}^{j_n(p,t)}\mathbb{E}(\eta(p)_j^n \, \Delta(N, p)_j^n \mid \mathcal{G}(p)_j^n) \xrightarrow{\mathbb{P}} 0, \tag{5.58}$$

where $\Delta(V, p)_j^n = V_{j(p+1)k_n\Delta_n} - V_{(j-1)(p+2)k_n\Delta_n}$ for any process $V$, and where (5.58) should hold for all bounded martingales $N$ which are orthogonal to $W$, and also for $N = W$. The last property is as stated as in [37]. However, a look at the proof in [37] shows that it is enough to have it for $N = W$, and for all $N$ in a set $\mathcal{N}$ of bounded martingales which are orthogonal to $W$ and such that the family $(N_\infty : N \in \mathcal{N})$ is total in the space $\mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. A suitable such set $\mathcal{N}$ will be described later.

(2) Since $j_n(p, t) \leq Kt/p\sqrt{\Delta_n}$, (5.57) trivially follows from (5.40), whereas (5.56) is an immediate consequence of (5.41) and of a Riemann sums argument.

(3) The proof of (5.58) is much more involved, and we begin by proving that

$$\Delta_n^{1/4}\sum_{j=0}^{j_n(p,t)} a_j^n \xrightarrow{\mathbb{P}} 0, \quad \text{where } a_j^n = \mathbb{E}(\zeta(W, p)_{j(p+1)k_n}^n \, \Delta(N, p)_j^n \mid \mathcal{G}(p)_j^n). \tag{5.59}$$

We have $\zeta(W, p)_i^n = (\sigma_i^n)^2 V_n(i, p) - V_n'(i, p)$ (see after (5.22)), and we set

$$\delta_j^n = \mathbb{E}(V_n(j(p+1)k_n, p) \, \Delta(N, p)_j^n \mid \mathcal{G}(p)_j^n),$$
$$\delta_j'^n = \mathbb{E}(V_n'(j(p+1)k_n, p) \, \Delta(N, p)_j^n \mid \mathcal{G}(p)_j^n).$$

When $N = W$, the variable $\delta_j^n$ is the $\mathcal{F}_{j(p+1)k_n\Delta_n}$-conditional expectation of an odd function of the increments of the process $W$ after time $j(p+2)k_n\Delta_n$, hence it vanishes. Suppose now that $N$ is a bounded martingale, orthogonal to $W$. By Itô's formula we see that $(\overline{W}_j^n)^2$ is the sum of a constant (depending on $n$) and of a martingale which is a stochastic integral with respect to $W$, on the interval $[j\Delta_n, (j+k_n)\Delta_n]$. Hence $\delta_j^n$ is the sum of a constant plus a martingale which is a stochastic integral with respect to $W$, on the interval $[j(p+1)k_n\Delta_n, (j+1)(p+1)k_n\Delta_n]$. Then the orthogonality of $N$ and $W$ implies $\delta_j^n = 0$ again. Hence in both cases we have $\delta_j^n = 0$.

Since $a_j^n = (\sigma_{j(p+1)k_n}^n)^2\delta_i^n - \delta_i'^n$, (5.59) will follow if we prove

$$\Delta_n^{1/4} \sum_{j=0}^{j_n(p,t)} |\delta_j'^n| \xrightarrow{\mathbb{P}} 0. \tag{5.60}$$

For this we use (5.23). Since $N$ is a martingale, we deduce (using Cauchy–Schwarz inequality) that

$$|\delta_j'^n| \le K_p \, \chi(p)_{j(p+1)k_n}^n \sqrt{\mathbb{E}(\Delta(F, p)_j^n \mid \mathcal{G}(p)_j^n)}, \tag{5.61}$$

where $F = \langle N, N \rangle$ (the predictable bracket of $N$). Then the expected value of the left side of (5.60) is smaller than the square-root of

$$\mathbb{E}(F_t) \, \mathbb{E}\sqrt{\Delta_n} \sum_{j=0}^{j_n(p,t)} \mathbb{E}\left((\beta(p)_{j(p+1)k_n}^n)^2\right),$$

and we conclude by Lemma 5.4.

(4) In this step we prove that

$$\Delta_n^{1/4} \sum_{j=0}^{j_n(p,t)} a_j'^n \xrightarrow{\mathbb{P}} 0, \quad \text{where } a_j'^n = \mathbb{E}(\zeta(X, p)_{j(p+2)k_n}^n \, \Delta(N, p)_j^n \mid \mathcal{G}(p)_j^n). \tag{5.62}$$

Then by Cauchy–Schwarz inequality and (5.35) we see that $|a_j'^n - a_j^n|$ satisfies the same estimate than $|\delta_j'^n|$ in (5.61). Hence we deduce (5.62) from (5.59) like in the previous step.

(5) It remains to deduce (5.58) from (5.62), and for this we have to specify the set $\mathcal{N}$. This set $\mathcal{N}$ is the union of $\mathcal{N}^0$ and $\mathcal{N}^1$, where $\mathcal{N}^0$ is the set of all bounded martingales on $(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)}), \mathbb{P}^{(0)})$, orthogonal to $W$, and $\mathcal{N}^1$ is the set of all martingales having $N_\infty = f(Z_{t_1}, \ldots, Z_{t_q})$, where $f$ is any Borel bounded on $\mathbb{R}^q$ and $t_1 < \cdots < t_q$ and $q \ge 1$.

When $N$ is either $W$ or is in $\mathcal{N}^0$, then by (5.40) the left sides of (5.58) and of (5.62) agree, so in this case (5.58) holds. Next, suppose that $N$ is in $\mathcal{N}^1$, associated with the integer $q$ and the function $f$ as above. In view of (2.4) it is easy to check that $N$ takes the following form (by convention $t_0 = 0$ and $t_{q+1} = \infty$):

$$t_l \le t < t_{l+1} \Rightarrow N_t = M(l; Z_{t_1}, \ldots, Z_{t_l})_t$$

for $l = 0, \ldots, q$, and where $M(l; z_1, \ldots, z_l)$ is a version of the martingale

$$M(l; z_1, \ldots, z_l)_t = \mathbb{E}^{(0)}\left(\int \prod_{r=l+1}^{q} Q_{t_r}(\mathrm{d}z_r) f(z_1, \ldots, z_l, z_{l+1}, \ldots, z_q) \mid \mathcal{F}_t^{(0)}\right)$$

(with obvious conventions when $l = 0$ and $l = q$), which is measurable in $(z_1, \ldots, z_l, \omega^{(0)})$. Then

$$\mathbb{E}((\zeta(Z, p)_{j(p+1)k_n} - \zeta(X, p)_{j(p+1)k_n}) \, \Delta(N, p)_j^n \mid \mathcal{G}(p)_j^n) = 0 \tag{5.63}$$

by (5.40) when the interval $(j(p+1)k_n \Delta_n, (j(p+1)+1)k_n \Delta_n]$ contains no point $t_l$. Furthermore, the left side of (5.63) is always smaller in absolute value than $K_p$ (use (5.29) and (5.40) and the boundedness of $N$). Since we have only $q$ intervals $(j(p+2)k_n \Delta_n, (j(p+1)+2)k_n \Delta_n]$ containing points $t_l$, at most, we deduce from this fact and from (5.63) that

$$\left| \frac{\theta \psi_2}{\Delta_n^{1/4}} \sum_{j=0}^{j_n(p,t)} \mathbb{E}(\eta(p)_j^n \, \Delta(N, p)_j^n \mid \mathcal{G}(p)_j^n) - \Delta_n^{1/4} \sum_{j=0}^{j_n(p,t)} a_j'^n \right| \le q K_p \Delta_n^{1/4},$$

and (5.58) readily follows from (5.62).  □

Now we can proceed to the proof of the first claim of Theorem 3.1. We have

$$\frac{1}{\Delta_n^{1/4}} \, (\widehat{C}_t^n - C_t) = \frac{1}{\Delta_n^{1/4}} \, M(p)_t^n + V(p)_t^n,$$

where

$$V(p)_t^n = \frac{1}{\Delta_n^{1/4}} \, \left( M'(p)_t^n + F(p)_t^n + F'(p)_t^n + \widehat{C}(p)_t^n + \widehat{C}'(p)_t^n + \widehat{C}_t''^n \right).$$

On the one hand, Lemmas 5.5 and 5.6 and (5.53) yield

$$\lim_{p \to \infty} \limsup_{n \to \infty} \mathbb{P} \left( |V(p)_t^n| > \varepsilon \right) = 0$$

for all $\varepsilon > 0$. On the other hand, we fix the Brownian motion $B$, independent of $\mathcal{F}$. Since $\gamma(p)_t(\omega)$ converges pointwise to $\gamma_t(\omega)$ and stays bounded by (5.55), it is obvious that $Y(p)_t \xrightarrow{\mathbb{P}} Y_t$ (recall (3.8) and (5.54) for $Y$ and $Y(p)$). Then the result follows from (5.54) in a standard way.

It remains to prove (3.10). We set for $r = 1, 2, 3$:

$$\Gamma(r)_t^n = \sum_{i \in I(r,n,t)} u(r)_i^n, \tag{5.64}$$

where

$$I(r, n, t) = \begin{cases} \{0, 1, \ldots, [t/\Delta_n] - k_n + 1\} & \text{if } r = 1 \\ \{0, 1, \ldots, [t/\Delta_n] - 2k_n + 1\} & \text{if } r = 2 \\ \{0, 1, \ldots, [t/\Delta_n] - 3\} & \text{if } r = 3 \end{cases}$$

and

$$u(1)_i^n = (\overline{Z}_i^n)^4, \qquad u(2)_i^n = \Delta_n \zeta''(Z)_i^n, \qquad u(3)_i^n = \Delta_n (\Delta_{i+1}^n Z)^2 (\Delta_{i+3}^n Z)^2.$$

(Note the different summations ranges $I(r, n, t)$, which ensure that we take into account all variables $\zeta(r)_i^n$ which are observable up to time $t$, and not more.)

Then a simple computation shows that (3.10) and (3.14) are implied by

$$\Gamma(r)_t^n \xrightarrow{\mathbb{P}} \Gamma(r)_t := \int_0^t \overline{\gamma}(r)_s \, \mathrm{d}s \tag{5.65}$$

for $r = 1, 2, 3$, where

$$\overline{\gamma}(1)_t = 3\theta^2 \psi_2^2 \sigma_t^4 + 6\psi_1 \psi_2 \sigma_t^2 \alpha_t + \frac{3}{\theta^2} \psi_1^2 \alpha_t^2$$

$$\overline{\gamma}(2)_t = 2\theta^2 \psi_2 \sigma_t^2 \alpha_t + 2\psi_1 \alpha_t^2$$

$$\overline{\gamma}(3)_t = 4\alpha_t^2.$$

We set $u'(r)_i^n = \mathbb{E}(u(r)_i^n \mid \mathcal{F}_i^n)$, and we denote by $\Gamma'(r)_t^n$ for $r = 1, 2, 3$ the processes defined by (5.64), with $u(r)_i^n$ substituted with $u'(r)_i^n$. Then we have $\Gamma'(r)_t^n \overset{\mathbb{P}}{\longrightarrow} \Gamma(r)_t$ for $r = 1, 2, 3$: this is a trivial consequence of (5.44)–(5.46) and of an approximation of an integral by Riemann sums. Hence it remains to prove that $\Gamma(r)_t^n - \Gamma'(r)_t^n \overset{\mathbb{P}}{\longrightarrow} 0$, a result obviously implied by the following convergence:

$$\sum_{i,j \in I(r,n,t)} v(r, n, i, j) \to 0,$$

where $v(r, n, i, j) = \left( (u(r)_i^n - u'(r)_i^n)(u(r)_j^n - u'(r)_j^n) \right).$ \hfill (5.66)

We have $|v(r, n, i, j)| \leq K \Delta_n^2$ by (5.42) for $r = 1$, by (5.43) for $r = 2$ and by (5.46) for $r = 3$. Further $v(1, n, i, j) = 0$ when $|j - i| \geq k_n$, and $v(2, n, i, j) = 0$ when $|j - i| \geq 2k_n$, and $v(3, n, i, j) = 0$ when $|j - i| \geq 5$, so (5.66) holds in all cases.

## References

[1] P.A. Meyer, Intégrales stochastiques — 1, in: Séminaire de probabilités de Strasbourg I, in: Lectures Notes in Math., vol. 39, Springer Verlag, Berlin, 1967, pp. 72–94.

[2] J. Jacod, Limit of random measures associated with the increments of a Brownian semimartingale, Tech. Rep., Université de Paris VI, 1994.

[3] J. Jacod, P. Protter, Asymptotic error distributions for the euler method for stochastic differential equations, Annals of Probability 26 (1998) 267–307.

[4] J. Jacod, Asymptotic properties of realized power variations and related functionals of semimartingales, Stochastic Processes and Applications 118 (2008) 517–559.

[5] D. Foster, D. Nelson, Continuous record asymptotics for rolling sample variance estimators, Econometrica 64 (1996) 139–174.

[6] T.G. Andersen, T. Bollerslev, Intraday periodicity and volatility persistence in financial markets, Journal of Empirical Finance 4 (1997) 115–158.

[7] T.G. Andersen, T. Bollerslev, Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, International Economic Review 39 (1998) 885–905.

[8] O.E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realised volatility and its use in estimating stochastic volatility models, Journal of the Royal Statistical Society, B 64 (2002) 253–280.

[9] L. Zhang, From martingales to ANOVA: Implied and realized volatility, Ph.D. Thesis, The University of Chicago, Department of Statistics, 2001.

[10] T.G. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, Great realizations, Risk 13 (2000) 105–108.

[11] T.G. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, The distribution of realized exchange rate volatility, Journal of the American Statistical Association 96 (2001) 42–55.

[12] T.G. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, Modeling and forecasting realized volatility, Econometrica 71 (2003) 579–625.

[13] O.E. Barndorff-Nielsen, N. Shephard, Power and bipower variation with stochastic volatility and jumps (with discussion), Journal of Financial Econometrics 2 (2004) 1–48.

[14] M. Chernov, E. Ghysels, A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation, Journal of Financial Economics 57 (2000) 407–458.

[15] M.M. Dacorogna, R. Gençay, U. Müller, R.B. Olsen, O.V. Pictet, An Introduction to High-Frequency Finance, Academic Press, San Diego, 2001.
[16] R.F. Engle, The econometrics of ultra-high frequency data, Econometrica 68 (2000) 1–22.
[17] A.R. Gallant, C.-T. Hsu, G.T. Tauchen, Using daily range data to calibrate volatility diffusions and extract the forward integrated variance, The Review of Economics and Statistics 81 (1999) 617–631.
[18] S. Goncalves, N. Meddahi, Bootstrapping realized volatility, Tech. Rep., Université de Montréal, 2005.
[19] I. Kalnina, O. Linton, Inference about realized volatility using infill subsampling, Tech. Rep., London School of Economics, 2007.
[20] F. Delbaen, W. Schachermayer, A general version of the fundamental theorem of asset pricing, Mathematische Annalen 300 (1994) 463–520.
[21] R. Roll, A simple model of the implicit bid-ask spread in an efficient market, Journal of Finance 39 (1984) 1127–1139.
[22] J. Hasbrouck, Assessing the quality of a security market: A new approach to transaction-cost measurement, Review of Financial Studies 6 (1993) 191–212.
[23] P.A. Mykland, L. Zhang, Discussion of paper "a selective overview of nonparametric metods in financial econometrics" by J. Fan, Statistical Science 20 (2005) 347–350.
[24] J. Jacod, La variation quadratique du Brownien en présence d'erreurs d'arrondi, Astérisque 236 (1996) 155–162.
[25] S. Delattre, J. Jacod, A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors, Bernoulli 3 (1997) 1–28.
[26] A. Gloter, J. Jacod, Diffusions with measurement errors. ii — Optimal estimators, ESAIM 5 (2001) 243–260.
[27] L. Zhang, P.A. Mykland, Y. Aït-Sahalia, A tale of two time scales: Determining integrated volatility with noisy high-frequency data, Journal of the American Statistical Association 100 (2005) 1394–1411.
[28] F.M. Bandi, J.R. Russell, Separating microstructure noise from volatility, Journal of Financial Economics 79 (2006b) 655–692.
[29] L. Zhang, Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach, Bernoulli 12 (2006) 1019–1043.
[30] M. Podolskij, M. Vetter, Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps, Tech. Rep., Ruhr-Universität Bochum, 2006.
[31] O.E. Barndorff-Nielsen, P.R. Hansen, A. Lunde, N. Shephard, Designing realised kernels to measure ex-post variation of equity prices in the presence of noise, Tech. Rep, 2006.
[32] Y. Aït-Sahalia, P.A. Mykland, L. Zhang, How often to sample a continuous-time process in the presence of market microstructure noise, Review of Financial Studies 18 (2005) 351–416.
[33] Y. Li, P.A. Mykland, Are volatility estimators robust with respect to modeling assumptions? Bernoulli 13 (2007) 601–622.
[34] F.M. Bandi, J.R. Russell, Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. Tech. Rep., University of Chicago Graduate School of Business, 2006.
[35] Y. Aït-Sahalia, P.A. Mykland, L. Zhang, Ultra high frequency volatility estimation with dependent microstructure noise, University of Chicago, in revision for Journal of Econometrics, 2005.
[36] S. Heston, A closed-form solution for options with stochastic volatility with applications to bonds and currency options, Review of Financial Studies 6 (1993) 327–343.
[37] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, 2nd ed., Springer-Verlag, New York, 2003.