Microstructure Noise in the Continuous Case: The Pre-Averaging Approach *

Jean Jacod †, Yingying Li ‡, Per A. Mykland §, Mark Podolskij ¶, and Mathias Vetter ‖

December 12, 2007

Abstract

This paper presents a generalized pre-averaging approach for estimating the integrated volatility. 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 possess an intuitive transparency, can generate rate optimal estimators (with convergence rate $n^{-1/4}$).

Keywords: consistency, continuity, discrete observation, Itô process, leverage effect, pre-averaging, quarticity, realized volatility, stable convergence.

AMS 2000 subject classifications. Primary 60G44, 62M09, 62M10; secondary 60G42, 62G20

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 analysed. This is particularly the case for market prices of stocks, currencies, and other financial instruments. On the

*Financial support from the Stevanovich Center for Financial Mathematics at the University of Chicago, from the National Science Foundation under grants DMS 06-04758 and SES 06-31605, and from Deutsche Forschungsgemeinschaft through SFB 475, is gratefully acknowledged

†Institut de Mathématiques de Jussieu, 175 rue du Chevaleret 75 013 Paris, France (CNRS – UMR 7586, and Université Pierre et Marie Curie - P6), Email: jj@ccr.jussieu.fr

‡Department of Statistics, The University of Chicago, Chicago, IL 60637, USA, Email; yyli@galton.uchicago.edu

§Department of Statistics, The University of Chicago, Chicago, IL 60637, USA, Email; mykland@galton.uchicago.edu

¶Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany, Email: podolski@cityweb.de

‖Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany, Email: mathias.vetter@rub.de

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 returns was shown to be consistent for the quadratic variation in Meyer (1967). A limit theory was then developed in Jacod (1994) and Jacod and Protter (1998), and later in Jacod (2006).

Meanwhile, these concepts were introduced to econometrics in Foster and Nelson (1996) and Andersen and Bollerslev (1997, 1998). A limit theory was developed in Barndorff-Nielsen and Shephard (2002), and Zhang (2001). Further early econometric literature includes, in particular, Andersen et al. (2000, 2001, 2003), Barndorff-Nielsen and Shephard (2004), Chernov and Ghysels (2000), Dacorogna et al. (2001), Engle (2000), and Gallant et al. (1999). The setting of confidence intervals using bootstrapping has been considered by Goncalves and Meddahi (2005) and Kalnina and Linton (2007).

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 Delbaen and Schachermayer (1994)) suggest that these must normally be semimartingales. Econometric evidence, however, suggests that there is additional noise in the prices. This goes back to Roll (1984) and Hasbrouck (1993). In the nonparametric setting, the deviation from semimartigales is most clearly seen through the signature plots of Andersen et al. (2000), see also the discussion in Mykland and Zhang (2005).

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 (1996) and Delattre and Jacod (1997). Additive noise is studied in Gloter and Jacod (2001), and a consistent estimator in the nonparametric setting is found in Zhang et al. (2005). Issues of bias-variance tradeoff are discussed in Bandi and Russell (2006b). In the nonparametric case, rate optimal estimators are found in Zhang (2006), Podolskij and Vetter (2006) and Barndorff-Nielsen et al. (2006). A development for low frequency data is given in Aït-Sahalia et al. (2005).

There are currently three main approaches to estimation in the nonparametric case: linear combination of realised volatilities obtained by *subsampling* (Zhang et al. (2005), Zhang (2006)), and linear combination of *autocovariances* (Barndorff-Nielsen et al. (2006)). The purpose of this paper is to give more insight to the third approach of *pre-averaging*, which was introduced in Podolskij and Vetter (2006). The idea is as follows. We suppose that the (say) log securities price X_t is a continuous semimartingale (of the form (2.1) below). 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 ϵ_{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 $\breve{Z}_{t_i}$ as the average of $Z_{t_{i+j}}$, $j = 0, ..., K-1$. The variance of the noise in $\breve{Z}_i$ is now reduced by a factor of about $1/K$. If one calculates the realised volatility on the basis of $\breve{Z}_0, \breve{Z}_{t_1}, \breve{Z}_{t_2}, ...$, 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 ϵ.

The paper provides a way of implementing this idea. There are several issues that have to be tackled in the process. First of all, while the local averaging does reduce the impact of the noise ϵ, it adds noise by time averaging the latent semimartingale X. The pre-averaged realised volatility $\sum_i (\breve{Z}_{t_{(2i+1)K}} - \breve{Z}_{t_{2iK}})^2$ therefore has to be adjusted by an additive term to eliminate the resulting bias. Second, one would not wish to only average over differences from non-overlapping intervals, but rather use a moving window. Finally, the estimator can be generalised 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, $k_n = 2K$.

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 Li and Mykland (2007). 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 Podolskij and Vetter (2006) 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 (2006a). 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 $i = 0, 1, \cdots, k, \cdots$. We are in the context of high frequency data, that is we are interested in the situation where the time lag Δ_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 $i = 0, 1, \cdots, [t/\Delta_n]$. 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

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \tag{2.1}$$

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 σ 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

$$C_t = \int_0^t \sigma_s^2 ds. \tag{2.2}$$

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 ω, 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 $Q_t(\omega^{(0)}, dz)$ from $(\Omega^{(0)}, \mathcal{F}_t^{(0)})$ into $\mathbb{R}$, which satisfies

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

We endow the space $\Omega^{(1)} = \mathbb{R}^{[0,\infty)}$ with the product Borel σ-field $\mathcal{F}^{(1)}$ and with the probability $\mathbb{Q}(\omega^{(0)}, d\omega^{(1)})$ 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:

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

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 Ω. 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 Equation (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

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

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

$$t \;\mapsto\; \int z^8\, Q_t(\omega^{(0)}, dz) \qquad \text{is a locally bounded process.} \qquad (2.6)$$

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 unbiasedness condition (2.3) of the noise. If this is not satisfied, and if we denote by Y_t the process $Y_t = \int z\, Q_t(dz)$, then as explained in Li and Mykland (2007) one cannot make inference on the process X itself, but only on the process Y, which thus should be assumed to satisfy Assumption (H). This is still a strong assumption on the noise, as we see in one of the following examples. If this is the case, one could replace everywhere below the process X by the process Y: so in a sense it is natural to assume $Y = X$.

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. $\Box$

Example 2) Let $Z_i^n = \gamma[(X_i^n + \varepsilon_i^n)/\gamma]$ for some $\gamma > 0$ and (ε_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 γ, for example $\gamma = 1$ cent. Then as soon as the ε_i^n are uniform over $[0, \gamma]$, or more generally uniform over $[-2i\gamma, (2i+1)\gamma]$ for some integer i, (K) is satisfied. If the ε_i^n have a C^2 density, with further a finite 8th moment and a support containing an interval of length γ, 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. $\Box$

Example 3) Let $Z_i^n = \gamma[X_i^n/\gamma]$ 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 an Itô semimartingale, and is not even càdlàg: so *nothing* of what follows applies. In fact in this case, if we observe the whole process $Z_t = \gamma[X_t/\gamma]$ 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

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

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

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

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

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

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

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

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

(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:

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

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*

$$Y_t \;=\; \int_0^t \gamma_s\, dB_s, \tag{3.8}$$

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

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

Moreover

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

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

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

Remark 1: Our estimators are in fact essentially the same as the kernel estimators in Barndorff-Nielsen et al. (2006). With our notation the "flat top" estimators of that paper are

$$\overline{K}_t^n = \sum_{i=k_n}^{[t/\Delta_n]-k_n+1} (\Delta_i^n Z)^2 + \sum_{k_n \le i \le [t/\Delta_n]-k_n+1,\ 1\le j\le k_n} k\Big(\frac{j-1}{k_n}\Big)\Big(\Delta_i^n Z\,\Delta_{i+j}^n Z + \Delta_i^n Z\,\Delta_{i-j}^n Z\Big),$$

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

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

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. (2006).

Remark 2: Suppose that $Y_t = \sigma W_t$ and that $\alpha_t = \alpha$, where σ and α are positive constants, and that $t = 1$. In this case there is an efficient parametric bound for the asymptotic variance for estimating σ^2, which is $8\sigma^3\sqrt{\alpha}$, see e.g. Gloter and Jacod (2001). On the other hand, the concrete estimators given in Barndorff-Nielsen et al. (2006) or Podolskij and Vetter (2006) or Zhang (2006), in the i.i.d. additive noise case, have an asymptotic variance ranging from $8.29\,\sigma^3\sqrt{\alpha}$ to $26\,\sigma^3\sqrt{\alpha}$, upon using an "optimal" choice of θ 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.

In practice we do not know how to choose θ 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 (2001)). Moreover the existence of an "optimal" choice of θ is not even very clear, since $\sigma = \sigma_t$ and $\alpha = \alpha_t$ are usually random and time dependent. Nevertheless we usually have an idea of the "average" sizes α_{ave} and σ_{ave} of α_t and σ_t: in this case one should take θ close to $4.8\ \sqrt{\alpha_{ave}}/\sigma_{ave}$.

4 Simulation results

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

4.1 Simulation Design

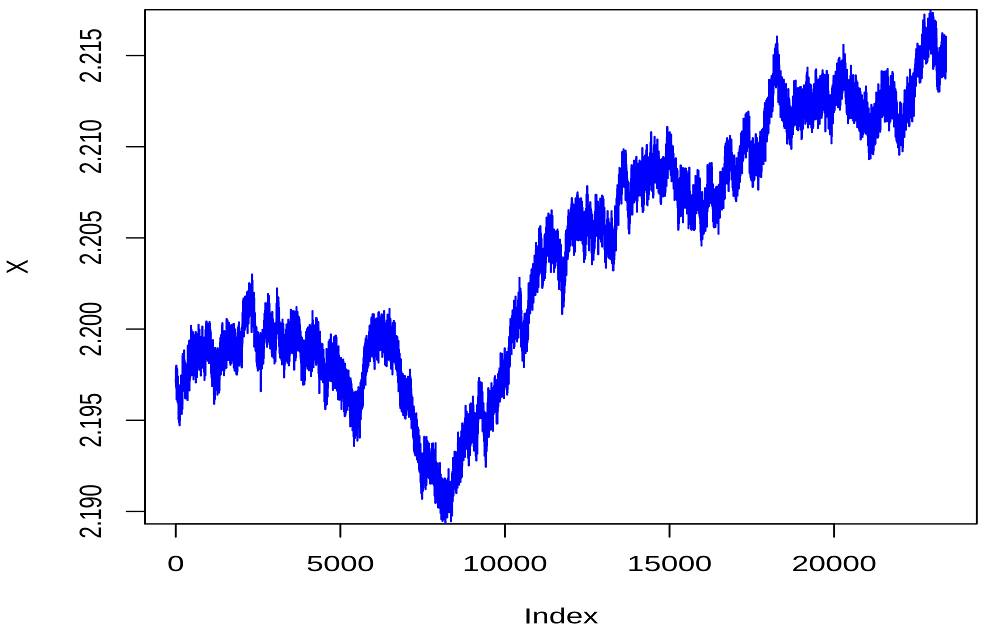
We study the case when the weight function is taken to be $g(x) = x \wedge (1-x)$. We simulate data for one day ($t \in [0,1]$), and assume the data is observed once every second (n=23400). The X processes and the market microstructure noise processes are generated from the models below. 25000 iterations were run for each model.

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

$$dX_t = \sigma dB_t, \;\; Z_{t_i}^n = X_{t_i}^n + \epsilon_{t_i}^n$$

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

The observed sample path looks like this:



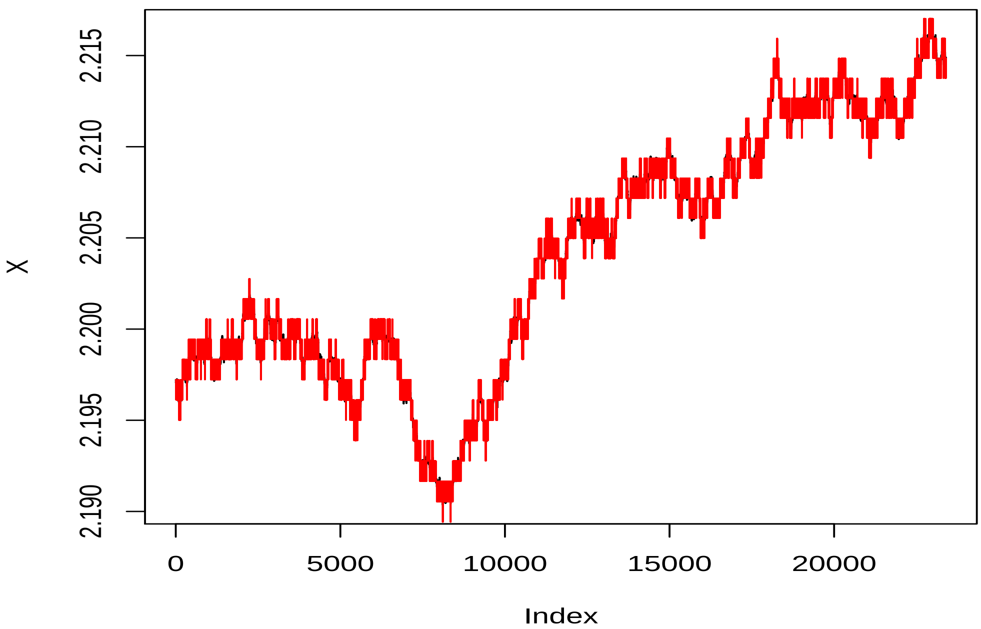
Model 2 – the case of constant volatility & rounding plus error.

$X_t = X_0 + \sigma W_t$, $U_{t_i} = \text{Uniform}(0, \log\left(\frac{\gamma\lceil\frac{exp(X_{t_i})}{\gamma}\rceil}{\gamma\lfloor\frac{exp(X_{t_i})}{\gamma}\rfloor}\right))$, $Z_{t_i} = \log(\gamma\lfloor\frac{\exp(X_{t_i}+U_{t_i})}{\gamma}\rfloor)$

This model is similar as the two-stage contamination model studied in Li and Mykland (2007), where the first stage of contamination is an additive error on the log prices, and the second stage is rounding on the prices. The observed log price Z_{t_i}'s are the logarithm of the rounded contaminated prices.

Parameters used: $\sigma = 0.2/\sqrt{252}$, $X_0 = log(9)$, $\gamma = 0.01$.

The observed log price process looks like this:



Model 3 – Model 3 – the case of stochastic volatility & additive noise. The Heston model (Heston (1993)) is used to generate the stochastic volatility process.

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

and

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

where $\nu_t = \sigma_t^2$ and we assume $Corr(B, W) = \rho$.

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

4.2 Simulation Results

Some initial simulations 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 θ comes 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 θ 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 θ, and then re-define θ 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 1 reports the performance of the estimator $\hat{C}_t^n$ and the variance estimator Γ_t^n.

	Model 1	Model 2	Model 3
Small-sample bias $Avg(\hat{C}_t^n - C)$	-1.390286e-06	-1.368032e-06	-1.329654e-06
Bias in the variance estimator $Avg(\Gamma_t^n - \int_0^t \gamma_s^2 ds)$	-1.520074e-10	-1.433976e-10	-1.385071e-10

As we will see later, the results in Model 1, where $\hat{C}_t^n$ and Γ_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 ψ_i and ϕ_{ij} by their finite sample analogues, which are defined as follows:

$$\begin{aligned}
\psi_1^{k_n} &= k_n \sum_{j=1}^{k_n} (g_{j+1}^n - g_j^n)^2, \qquad \psi_2^{k_n} = \frac{1}{k_n} \sum_{j=1}^{k_n-1} (g_j^n)^2 \\
\phi_1^{k_n}(j) &= \sum_{j=i+1}^{k_n} (g_{i-1}^n - g_i^n)(g_{i-j-1}^n - g_{i-j}^n), \qquad \phi_2^{k_n}(j) = \sum_{j=i+1}^{k_n} g_i^n g_{i-j}^n \\
\Phi_{11}^{k_n} &= k_n \Big(\sum_{j=0}^{k_n-1} (\phi_1^{k_n}(j))^2 - \frac{1}{2}(\phi_1^{k_n}(0))^2 \Big) \\
\Phi_{12}^{k_n} &= \frac{1}{k_n} \Big(\sum_{j=0}^{k_n-1} \phi_1^{k_n}(j)\phi_2^{k_n}(j) - \frac{1}{2}\phi_1^{k_n}(0)\phi_2^{k_n}(0) \Big) \\
\Phi_{22}^{k_n} &= \frac{1}{k_n^3} \Big(\sum_{j=0}^{k_n-1} (\phi_2^{k_n}(j))^2 - \frac{1}{2}(\phi_2^{k_n}(0))^2 \Big)
\end{aligned}$$

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:

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

and

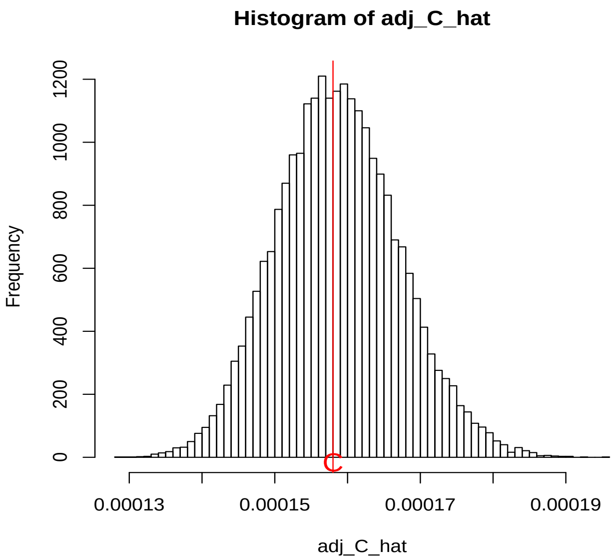
$$\Gamma_t^{n,adj} = \Big(1 - \frac{\psi_1^{k_n} \Delta_n}{2\theta^2 \psi_2^{k_n}}\Big)^{-2} \Big(\frac{4\Phi_{22}^{k_n} \lfloor t/\Delta_n \rfloor}{3\theta(\psi_2^{k_n})^4 (\lfloor t/\Delta_n \rfloor - k_n + 2)} \sum_{i=0}^{\lfloor t/\Delta_n \rfloor - k_n + 1} (\bar{Z}_i^n)^4$$

$$+\frac{4\Delta_n \lfloor t/\Delta_n \rfloor}{\theta^3\Big(\lfloor t/\Delta_n \rfloor - k_n + 2\Big)}\Big(\frac{\Phi_{12}^{k_n}}{(\psi_2^{k_n})^3} - \frac{\Phi_{22}^{k_n}\psi_1^{k_n}}{(\psi_2^{k_n})^4}\Big) \sum_{i=0}^{\lfloor t/\Delta_n \rfloor - k_n + 1} (\bar{Z}_i^n)^2 \sum_{j=i+k_n}^{i+2k_n-1} (\Delta_j^n Z)^2$$

$$+\frac{\Delta_n \lfloor t/\Delta_n \rfloor}{\theta^3(\lfloor t/\Delta_n \rfloor - 2)}\Big(\frac{\Phi_{11}^{k_n}}{(\psi_2^{k_n})^2} - \frac{2\Phi_{12}^{k_n}\psi_1^{k_n}}{(\psi_2^{k_n})^3} + \frac{\Phi_{22}^{k_n}(\psi_1^{k_n})^2}{(\psi_2^{k_n})^4}\Big) \sum_{i=1}^{\lfloor t/\Delta_n \rfloor - 2} (\Delta_i^n Z)^2(\Delta_{i+2}^n Z)^2\Big)$$

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

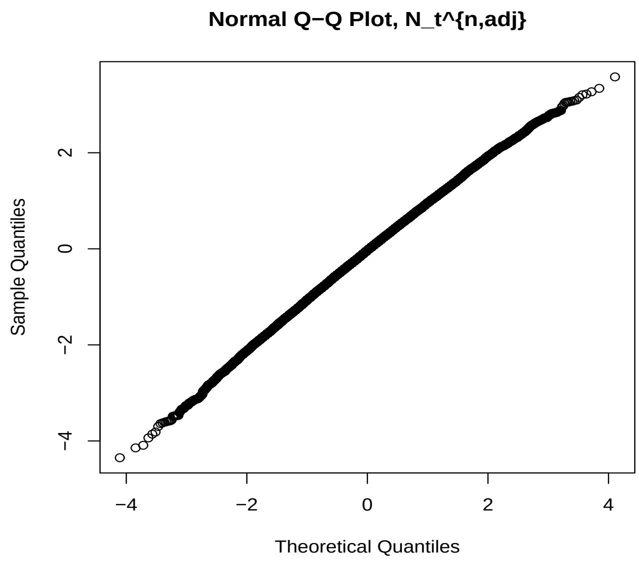
	Model 1	Model 2	Model 3
Small-sample bias $Avg(\hat{C}_t^{n,adj} - C)$	-4.641224e-08	-1.278064e-07	1.390028e-08
Bias in the variance estimator $Avg(\Gamma_t^{n,adj} - \int_0^t \gamma_s^2 ds)$	5.631088e-12	1.54554e-13	2.012485e-11



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

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

	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 $N_t^{n,adj}$	-0.05	1.02	0.82%	3.20%	6.08%	95.55%	97.94%	99.68%
Model 2 N_t^n	-0.22	1.05	1.49%	4.96%	8.28%	97.13%	98.75%	99.87%
Model 2 $N_t^{n,adj}$	-0.06	1.03	1.00%	3.65%	6.38%	95.94%	98.20%	99.80%
Model 3 N_t^n	-0.21	1.05	1.32%	4.86%	8.41%	96.8%	98.66%	99.82%
Model 3 $N_t^{n,adj}$	-0.05	1.03	0.84%	3.42%	6.24%	95.58%	97.99%	99.73%



Normal Q-Q plot of $N_t^{n,adj}$ for Model 1.

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 Γ_t^n. One can adjust this effect by using a first order Taylor expansion of Γ: expanding Γ_t^n or $\Gamma_t^{n,adj}$ around the theoretical asymptotic variance Γ_0 ($\frac{1}{\sqrt{\Gamma_0}} \approx \frac{1}{\sqrt{\Gamma_t^n}} - \frac{\Gamma_0 - \Gamma_t^n}{2{\Gamma_0}^{3/2}}$):

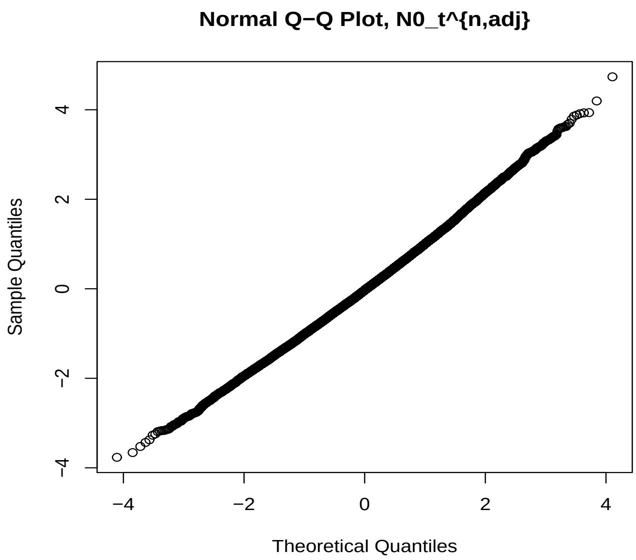
$$N0_t^n := N_t^n - \frac{(\Gamma_0 - \Gamma_t^n)(\hat{C}_t^n - C)}{2\Delta_n^{1/4}\Gamma_0^{3/2}}$$

and

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

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

	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 $N0_t^{n,adj}$	-0.01	1.04	0.40%	2.21%	4.81%	94.26%	96.78%	99.16%
Model 2 $N0_t^n$	-0.17	1.02	0.78%	3.70%	6.91%	95.90%	97.87%	99.52%
Model 2 $N0_t^{n,adj}$	-0.02	1.03	0.55%	2.64%	5.16%	94.46%	96.99%	99.24%
Model 3 $N0_t^n$	-0.16	1.04	0.65%	3.50%	6.85%	95.71%	97.79%	99.53%
Model 3 $N0_t^{n,adj}$	0.00084	1.05	0.38%	2.40%	4.94%	94.06%	96.76%	99.16%



Normal Q-Q plot of $N0_t^{n,adj}$ 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, σ,$\int z^8\, Q_t(dz)$ and X itself are bounded (uniformly in (ω, t)) (then α is also bounded). $\square$

Then a standard localization procedure explained in details in Jacod (2006) 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 α 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 ∞, 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 $\mathrm{O}_u(x)$ 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

$$\left.\begin{array}{lcl} X(n,s)_t &=& \int_0^t b_u g_n(u-s)du + \int_0^t \sigma_u g_n(u-s)dW_u \\ C(n,s)_t &=& \int_0^t \sigma_u^2 \, g_n(u-s)^2 \, du. \end{array}\right\} \tag{5.2}$$

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

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

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

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

(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

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

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

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

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

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

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

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

$$\widehat{C}(p)_t^n \;=\; \frac{\sqrt{\Delta_n}}{\theta\psi_2} \sum_{i=i_n(p,t)}^{[t/\Delta_n]-k_n+1} \widetilde{Z}_i'^n, \tag{5.11}$$

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

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

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:

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

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

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*

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

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

Proof. 1) Since $g(0) = g(1) = 0$ we have $\int_0^1 g'(s)ds = 0$. We introduce the process

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

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

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

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

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

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

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

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

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

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 $p\geq 1$:

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

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

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

Then (5.25) yields

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

Combining (5.24) and (5.26) with (5.22), we immediately get (5.17). □

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 $q>0$,

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

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

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

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

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

Here and below, as mentioned before, the constant K_p depends on p, and it typically goes to ∞ 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*

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

Proof. Observe that, similar to (5.27),

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

Let us define the processes

$$\left.\begin{array}{lcl} M(n,s)_t &=& 2\int_0^t X(n,s)_u\,\sigma_u\,g_n(u-s)\,dW_u, \\ B(n,s)_t &=& 2\int_0^t X(n,s)_u\,b_u\,g_n(u-s)du. \end{array}\right\}$$

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 $\mathbb{E}(\chi(1)_j^n \mid \mathcal{F}_i^n) \leq \chi(p)_i^n$ when $i \leq j \leq i+(p+1)k_n$, (5.30) is implied by

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

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

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

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

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

Lemma 5.3 *We have*

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

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

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

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

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

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

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

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

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

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

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

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

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

At this stage, the result readily follows from Lemma 5.1. $\Box$

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 $|h_j^n| \leq K/k_n$ and α is bounded, and by the rate of approximation of the integral of a piecewise Lipschitz function by Riemann sums, the following properties are obvious:

$$\left.\begin{array}{l} |A_{i,j}^n| \;\leq\; K\sqrt{\Delta_n} \\ |j-i| \geq k_n \quad\Rightarrow\quad A_{i,j}^n \;=\; 0 \\ i \leq j \leq m \leq i+(p+1)k_n \quad\Rightarrow \\ \qquad\qquad\qquad A_{j,m}^n \;=\; \alpha_i^n \frac{1}{k_n}\,\phi_1\Big(\frac{m-j}{k_n}\Big) + \mathrm{O}_u(p\Delta_n + \sqrt{\Delta_n}\;\beta(p)_i^n) \\ \sum_{(j,m):\; i\leq j<m\leq i+pk_n-1}(A_{j,m}^n)^2 = (\alpha_i^n)^2(p\Phi_{11}+\Xi_{11}) + \mathrm{O}_u\Big(p^3\sqrt{\Delta_n} + p\beta(p)_i^n\Big). \end{array}\right\} \tag{5.36}$$

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 α_t if $s = t$. Then, recalling (2.5) and (5.4),

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

More generally, $\mathbb{E}\Big(\Pi_{m=1}^q h_{j_m}^n(Z_{i+j_m}^n - X_{i+j_m}^n) \mid \mathcal{F}_{i+j_1}^n\Big) = 0$ as soon as there is one j_m which is different from all the others, and moreover $|h_j^n| \leq K\sqrt{\Delta_n}$, whereas the moments (2.6) are bounded for $q \leq 8$. 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

$$\mathbb{E}\Big((\overline{Z}_i^n - \overline{X}_i^n)^q(\overline{Z}_j^n - \overline{X}_j^n)^r \mid \mathcal{F}_{i\wedge j}'^n\Big) = \begin{cases} \mathrm{O}_u(\Delta_n) & \text{if } q+r=3 \\ A_i^n A_j^n + 2(A_{i,j}^n)^2 + \mathrm{O}_u(\Delta_n^{3/2}) & \text{if } q=r=2 \\ \mathrm{O}_u(\Delta_n^2) & \text{if } q+r=8. \end{cases} \tag{5.38}$$

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

$$\left.\begin{array}{lcl} \mathbb{E}\Big(\widetilde{Z}_i'^n \mid \mathcal{F}_i'^n\Big) &=& (\overline{X}_i^n)^2 - c_i^n, \qquad \mathbb{E}\Big(|\widetilde{Z}_i'^n| \mid \mathcal{F}_i'^n\Big) \;=\; (\overline{X}_i^n)^2 + \mathrm{O}_u(\sqrt{\Delta_n}) \\ \mathbb{E}\Big(\widetilde{Z}_i'^n \widetilde{Z}_j'^n \mid \mathcal{F}^{(0)}\Big) &=& ((\overline{X}_i^n)^2 - c_i^n)((\overline{X}_j^n)^2 - c_j^n) + 4\overline{X}_i^n \overline{X}_j^n A_{i,j}^n + 2(A_{i,j}^n)^2 \\ && +\mathrm{O}_u\Big(\Delta_n^{3/2} + \Delta_n|\overline{X}_i^n| + \Delta_n|\overline{X}_j^n|\Big) \\ \mathbb{E}\Big((\widetilde{Z}_i'^n)^4 \mid \mathcal{F}_i^n\Big) &\leq& K(\Delta_n^2 + |\overline{X}_i^n|^8), \end{array}\right\} \tag{5.39}$$

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

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

and also, in view of (5.36),

$$\begin{array}{lcl} \mathbb{E}((\zeta(Z,p)_i^n)^2 \mid \mathcal{F}_i'^n) &=& (\zeta(X,p)_i^n)^2 + \dfrac{8}{k_n}\;\alpha_i^n\,\zeta'(X,p)_i^n + 4(\alpha_i^n)^2(p\Phi_{11} + \Xi_{11}) \\ && +p^3\mathrm{O}_u\Big(\Big(\sqrt{\Delta_n} + \beta(p)_i^n\Big)\Big(1 + \displaystyle\sum_{j=i}^{i+pk_n-1} |\overline{X}_j^n|^2\Big)\Big). \end{array}$$

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

$$\begin{array}{l} \Big|\mathbb{E}((\zeta(Z,p)_i^n)^2 \mid \mathcal{F}_i^n) - 4(p\Phi_{22} + \Xi_{22})k_n^4\Delta_n^2(\sigma_i^n)^4 \\ \qquad -8\alpha_i^n(\sigma_i^n)^2(p\Phi_{12} + \Xi_{12})k_n^2\Delta_n - 4(\alpha_i^n)^2(p\Phi_{11} + \Xi_{11})\Big| \;\leq\; K_p\;\chi(p)_i^n. \end{array} \tag{5.41}$$

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

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

and (using the boundedness of X)

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

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

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

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

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

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

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*

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

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

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

Let $\varepsilon > 0$ and denote by $N(\varepsilon)_t$ the number of jumps of any of the three processes b, σ or α, with size bigger than ε, 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, σ or α of size bigger than ε. Then obviously, since all three processes b, σ, α are bounded,

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

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:*

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

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

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

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

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

$$\begin{aligned}\sum_{i=0}^{[t/\Delta_n]-k_n+1} c_i^n &= \sum_{i=0}^{[t/\Delta_n]-k_n+1}\sum_{l=i+1}^{i+k_n-1}(g_{l-i}^n)^2\Delta_l^n C\\ &= \sum_{l=1}^{[t/\Delta_n]}\Delta_l^n C\sum_{j=1\vee(l+k_n-1-[t/\Delta_n])}^{l\wedge(k_n-1)}(g_j^n)^2 \;=\; a_n\sum_{l=k_n-1}^{[t/\Delta_n]-k_n+2}\Delta_l^n C+\mathrm{O}_u(1).\end{aligned}$$

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

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

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 $1 \leq i \leq j-2$:

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

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

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

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

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

$$\begin{aligned}V_n \;=\;& \frac{\Delta_n^{1/4}}{\theta\psi_2}\left(\sum_{i=0}^{k_n-1}\alpha_i^n\ \sum_{l=0}^{i}(h_l^n)^2 + \sum_{i=i_n(p,t)}^{i_n(p,t)+k_n-2}\alpha_i^n\sum_{l=i+1-i_n(p,t)}^{k_n-1}(h_l^n)^2\right)\\ &-\frac{\psi_1\Delta_n^{3/4}}{2\theta^2\psi_2}\left(\alpha_0^n + 2\sum_{i=1}^{k_n-1}\alpha_i^n + 2\sum_{i=i_n(p,t)}^{[t/\Delta_n]-1}\alpha_i^n + \alpha_{[t/\Delta_n]}^n\right).\end{aligned}$$

On the one hand, since α_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, $\sum_{l=0}^{k_n-1}(h_l^n)^2 = \frac{\psi_1}{k_n} + \mathrm{O}(\Delta_n)$, 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 $M(p)^n_t$ and $M'(p)^n_t$ in (5.14). Those terms are (discretised) sums of martingale differences (note that $\eta(p)^n_j$ and $\eta'(p)^n_j$) are measurable with respect to $\mathcal{G}(p)^n_{j+1}$ and $\mathcal{G}'(p)^n_{j+1}$ respectively).

By Doob's inequality we have

$$\mathbb{E}\Big(\sup_{s\leq t}|M'(p)^n_s|^2\Big) \;\leq\; 4\sum_{j=0}^{j_n(p,t)}\mathbb{E}(|\eta'(p)^n_j|^2).$$

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

$$\mathbb{E}\Big(\sup_{s\leq t}|M'(p)^n_s|^2\Big) \;\leq\; \frac{Kt}{p}\,\sqrt{\Delta_n}. \tag{5.53}$$

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*

$$Y(p)_t \;=\; \int_0^t \gamma(p)_s dB_s, \tag{5.54}$$

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

$$\gamma(p)^2_t \;=\; \frac{4}{\psi_2^2}\,\Big(\Big(\frac{p}{p+1}\,\Phi_{22}+\frac{1}{p+1}\,\Psi_{22}\Big)\theta\sigma^4_t+2\Big(\frac{p}{p+1}\,\Phi_{12}+\frac{1}{p+1}\,\Psi_{12}\Big)\frac{\sigma^2_t\alpha_t}{\theta} + \Big(\frac{p}{p+1}\,\Phi_{11}+\frac{1}{p+1}\,\Psi_{11}\Big)\frac{\alpha^2_t}{\theta^3}\Big) \tag{5.55}$$

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

$$\frac{1}{\sqrt{\Delta_n}}\sum_{j=0}^{j_n(p,t)}\Big(\mathbb{E}((\eta(p)^n_j)^2\mid\mathcal{G}(p)^n_j)-(\overline{\eta}(p)^n_j)^2\Big)\;\stackrel{\mathbb{P}}{\longrightarrow}\;\int_0^t\gamma(p)^2_s\,ds, \tag{5.56}$$

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

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

where $\Delta(V,p)^n_j=V_{j(p+1)k_n\Delta_n}-V_{(j-1)(p+2)k_n\Delta_n}$ 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 Jacod and Shiryaev (2003). However, a look at the proof in Jacod and Shiryaev (2003) 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

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

We have $\zeta(W,p)^n_i=(\sigma^n_i)^2V_n(i,p)-V'_n(i,p)$ (see after (5.22)), and we set

$$\begin{aligned}\delta_j^n &= \mathbb{E}(V_n(j(p+1)k_n,p)\;\Delta(N,p)^n_j\mid\mathcal{G}(p)^n_j),\\ \delta_j'^n &= \mathbb{E}(V'_n(j(p+1)k_n,p)\;\Delta(N,p)^n_j\mid\mathcal{G}(p)^n_j).\end{aligned}$$

When $N=W$, the variable δ^n_j 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 $(\overline{W}^n_j)^2$ 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 δ^n_j 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 $\delta^n_j=0$ again. Hence in both cases we have $\delta^n_j=0$.

Since $a^n_j=(\sigma^n_{j(p+1)k_n})^2\delta^n_i-\delta'^n_i$, (5.59) will follow if we prove

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

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

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

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

$$\mathbb{E}(F_t)\;\mathbb{E}\Big(\sqrt{\Delta_n}\sum_{j=0}^{j_n(p,t)}\mathbb{E}((\beta(p)^n_{j(p+1)k_n})^2\Big),$$

and we conclude by Lemma 5.4.

4) In this step we prove that

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

Then by Cauchy-Schwarz inequality and (5.35) we see that $|a_j'^n-a^n_j|$ 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 $N_\infty = f(Z_{t_1}, \cdots, Z_{t_q})$, where f is any Borel bounded on $\mathbb{R}^q$ and $t_1 < \cdots < t_q$ and $q \geq 1$.

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 $t_{q+1} = \infty$):

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

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

$$M(l; z_1, \cdots, z_l)_t = \mathbb{E}^{(0)}\Big(\int \prod_{r=l+1}^{q} Q_{t_r}(dz_r) f(z_1, \cdots, z_l, z_{l+1}, \cdots, z_q) \mid \mathcal{F}_t^{(0)}\Big)$$

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

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

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

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

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

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

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 $Y(p)_t \stackrel{\mathbb{P}}{\longrightarrow} Y_t$ (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 $r = 1, 2, 3$:

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

where

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

and

$$u(1)^n_i = (\overline{Z}^n_i)^4, \qquad u(2)^n_i = \Delta_n \zeta''(Z)^n_i, \qquad u(3)^n_i = \Delta_n (\Delta^n_{i+1} Z)^2 (\Delta^n_{i+3} Z)^2.$$

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

Then a simple computation shows that (3.10) is implied by

$$\Gamma(r)^n_t \;\stackrel{\mathbb{P}}{\longrightarrow}\; \Gamma(r)_t \;:=\; \int_0^t \overline{\gamma}(r)_s ds \tag{5.65}$$

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

$$\begin{aligned} \overline{\gamma}(1)_t &= 3\theta^2 \psi_2^2 \sigma_t^4 + 6\psi_1 \psi_2 \sigma_t^2 \alpha_t + \frac{3}{\theta^2}\, \psi_1^2 \alpha_t^2 \\ \overline{\gamma}(2)_t &= 2\theta^2 \psi_2 \sigma_t^2 \alpha_t + 2\psi_1 \alpha_t^2 \\ \overline{\gamma}(3)_t &= 4\alpha_t^2. \end{aligned}$$

We set $u'(r)^n_i = \mathbb{E}(u(r)^n_i \mid \mathcal{F}^n_i)$, and we denote by $\Gamma'(r)^n_t$ for $r = 1, 2, 3$ the processes defined by (5.64), with $u(r)^n_i$ substituted with $u'(r)^n_i$. Then we have $\Gamma'(r)^n_t \stackrel{\mathbb{P}}{\longrightarrow} \Gamma(r)_t$ for $r = 1, 2, 3$: this is a trivial consequence of (5.44), (5.45) and (5.46) and of an approximation of an integral by Riemann sums. Hence it remains to prove that $\Gamma(r)^n_t - \Gamma'(r)^n_t \stackrel{\mathbb{P}}{\longrightarrow} 0$, a result obviously implied by the following convergence:

$$\sum_{i,j \in I(r,n,t)} v(r,n,i,j) \to 0, \;\; \text{where } \; v(r,n,i,j) = \Big((u(r)^n_i - u'(r)^n_i)(u(r)^n_j - u'(r)^n_j) \Big). \tag{5.66}$$

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

Aït-Sahalia, Y., Mykland, P. A. and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies* **18** 351–416.

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

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

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

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

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

Bandi, F. M. and Russell, J. R. (2006a). Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. Tech. rep., University of Chicago Graduate School of Business.

Bandi, F. M. and Russell, J. R. (2006b). Separating microstructure noise from volatility. *Journal of Financial Economics* **79** 655–692.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2006). Designing realised kernels to measure ex-post variation of equity prices in the presence of noise. Tech. rep.

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

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

Chernov, M. and Ghysels, E. (2000). 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** 407–458.

Dacorogna, M. M., Gençay, R., Müller, U., Olsen, R. B. and Pictet, O. V. (2001). *An Introduction to High-Frequency Finance.* Academic Press, San Diego.

DELATTRE, S. and JACOD, J. (1997). A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors. *Bernoulli* **3** 1–28.

DELBAEN, F. and SCHACHERMAYER, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* **300** 463–520.

ENGLE, R. F. (2000). The econometrics of ultra-high frequency data. *Econometrica* **68** 1–22.

FOSTER, D. and NELSON, D. (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica* **64** 139–174.

GALLANT, A. R., HSU, C.-T. and TAUCHEN, G. T. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *The Review of Economics and Statistics* **81** 617–631.

GLOTER, A. and JACOD, J. (2001). Diffusions with measurement errors. ii - optimal estimators. *ESAIM* **5** 243–260.

GONCALVES, S. and MEDDAHI, N. (2005). Bootstrapping realized volatility. Tech. rep., Université de Montréal.

HASBROUCK, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *Review of Financial Studies* **6** 191–212.

HESTON, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bonds and Currency Options *Review of Financial Studies* **6** 327-343.

JACOD, J. (1994). Limit of random measures associated with the increments of a Brownian semimartingale. Tech. rep., Université de Paris VI.

JACOD, J. (1996). La variation quadratique du Brownien en présence d'erreurs d'arrondi. *Astérisque* **236** 155–162.

JACOD, J. (2006). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and Applications* (to appear).

JACOD, J. and PROTTER, P. (1998). Asymptotic error distributions for the euler method for stochastic differential equations. *Annals of Probability* **26** 267–307.

JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes.* 2nd ed. Springer-Verlag, New York.

KALNINA, I. and LINTON, O. (2007). Inference about realized volatility using infill subsampling. Tech. rep., London School of Economics.

LI, Y. and MYKLAND, P. A. (2007). Are volatility estimators robust with respect to modeling assumptions? *Bernoulli* **13** 601–622.

Meyer, P.A. (1967). Intégrales stochastiques - 1. In *Séminaire de probabilités de Strasbourg* **I** Lectures Notes in Math., **39** 72–94 Springer Verlag: Berlin.

Mykland, P. A. and Zhang, L. (2005). Discussion of paper “a selective overview of nonparametric metods in financial econometrics” by J. Fan. *Statistical Science* **20** 347–350.

Podolskij, M. and Vetter, M. (2006). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. Tech. rep., Ruhr-Universität Bochum.

Roll, R. (1984). A simple model of the implicit bid-ask spread in an efficient market. *Journal of Finance* **39** 1127–1139.

Zhang, L. (2001). *From Martingales to ANOVA: Implied and Realized Volatility*. Ph.D. thesis, The University of Chicago, Department of Statistics.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli* **12** 1019–1043.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* **100** 1394–1411.