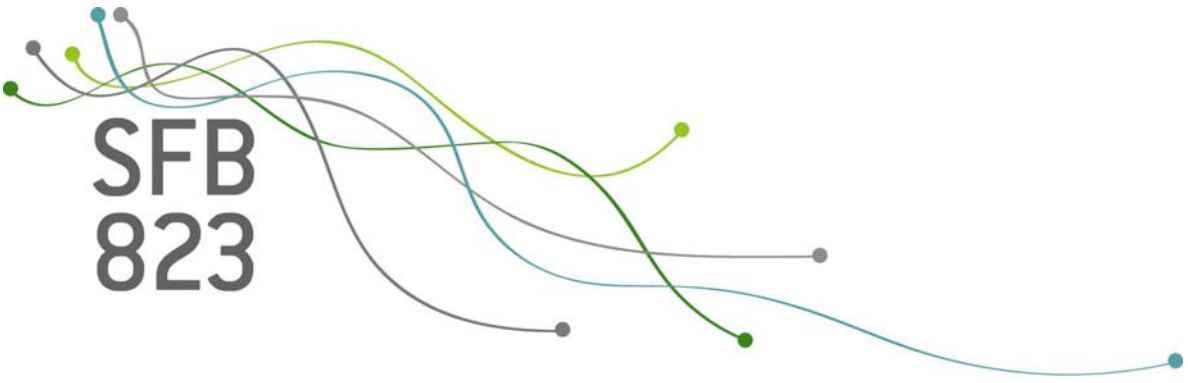
SFB 823

Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps

Markus Bibinger, Mathias Vetter

Nr. 20/2013

Discussion Paper



Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps

Markus Bibinger* and Mathias Vetter†

May 14, 2013

Abstract

We consider estimation of the quadratic (co)variation of a semimartingale from discrete observations which are irregularly spaced under high-frequency asymptotics. In the univariate setting, results from Jacod (2008) are generalized to the case of irregular observations. In the two-dimensional setup under non-synchronous observations, we derive a stable central limit theorem for the estimator by Hayashi and Yoshida (2005) in the presence of jumps. We reveal how idiosyncratic and simultaneous jumps affect the asymptotic distribution. Observation times generated by Poisson processes are explicitly discussed.

Keywords: asynchronous observations, co-jumps, statistics of semimartingales, quadratic covariation

1 Introduction

Estimating the quadratic variation of a semimartingale X probably is one of the main topics in today's high frequency statistics. Starting with the pioneering work of Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2002) on the use of realized volatility (also called realized variance) as a measure for integrated volatility over a trading day, an enormous number of articles has been dedicated to the development of estimation techniques in this area. Historically first is the extension to power variations which allows for a consistent estimation of integrated quarticity as well – a necessary task when establishing a so-called feasible central limit theorem for realized volatility that allows to construct confidence sets; see Barndorff-Nielsen and Shephard (2004).

Estimation approaches for deviations from the idealized setting of observing a continuous semimartingale at equidistant times have attracted a lot of attention since then. For models incorporating jumps, for example, integrated volatility does no longer coincide with the quadratic variation of the underlying process, as it comes from the continuous martingale part of X only. Econometricians, however, are typically interested in estimating integrated volatility due to the belief that this quantity reflects cumulative intrinsic risk of an asset whereas jumps come as external shocks. In the presence of jumps the realized volatility as a discretized quadratic variation converges in probability to the entire quadratic variation

*Humboldt-Universität Berlin, Institut für Mathematik, Unter den Linden 6, 10099 Berlin, Germany. E-Mail: bibinger@math.hu-berlin.de. Financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Ökonomisches Risiko", Humboldt-Universität zu Berlin, is gratefully acknowledged.

†Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany. E-mail: mathias.vetter@rub.de. The author is thankful for financial support through the collaborative research center "Statistik nichtlinearer dynamischer Prozesse" (SFB 823) of the Deutsche Forschungsgemeinschaft.

under high-frequency asymptotics as the maximum distance between successive observation times tends to zero. This motivated estimators which filter out jumps, like bipower variation by Barndorff-Nielsen and Shephard (2004) and truncated realized volatility by Mancini (2009). Another topic is the treatment of additional microstructure noise in the data; among various proposals see e.g. Zhang et al. (2005), Barndorff-Nielsen et al. (2008) or Podolskij and Vetter (2009). Non-regular observation times have been discussed in various situations: In the univariate context, limit theorems under irregular sampling schemes have been derived both in case of deterministic and random observations times; see e.g. Mykland and Zhang (2009), Hayashi et al. (2011) or Fukasawa and Rosenbaum (2012). In multi-dimensional settings, asynchronicity comes into play which makes the situation more complicated. Let us mention here the approach involving overlapping intervals by Hayashi and Yoshida (2005) and the concept of refresh times from Barndorff-Nielsen et al. (2011).

Where typically less focus has been laid on is statistical inference on the entire quadratic variation of X when jumps are present, though the latter is not only of some importance in economics as a measure of risk comprising jumps and volatility, but also a central quantity in stochastic analysis. Asymptotics in the case of equidistant observations of X are provided as a special case in Jacod (2008) who focuses on a number of functionals of semimartingale increments. A similar result for Lévy processes dates back to Jacod and Protter (1998). Results on an estimator for the quadratic variation when jumps and noise are present are given in Jacod et al. (2010) for their pre-averaging estimator. Apart from that, at least to the best of our knowledge, no work has dealt with central limit theorems on the entire quadratic (co)variation of X, and in particular very few is known in the framework of non-regularly spaced data.

We aim at filling this gap to a certain extent. In a first step, we generalize the asymptotic theory from Jacod (2008) on realized volatility and equidistant observations to non-equidistant (univariate) schemes. As a basis for the more involved situations we illuminate how proofs of limit theorems work for general semimartingales in the vein of Podolskij and Vetter (2010) who explained limit theorem for the continuous case. In a two-dimensional setting the quadratic covariation

$$[X^{(1)}, X^{(2)}]_t = \int_0^t \rho_s \sigma_s^{(1)} \sigma_s^{(2)} \, ds + \sum_{s \leq t} \Delta X_s^{(1)} \Delta X_s^{(2)} \tag{1.1}$$

is the sum of the integrated covolatility and the sum of products of simultaneous jumps (called co-jumps). The asymptotic theory for co-jumps entails new intriguing attributes and provides deeper insight in the multi-dimensional asymptotic properties of standard estimators.

For non-synchronous observations of continuous Itô semimartingales, the prominent estimator by Hayashi and Yoshida (2005) for integrated covolatility attains the minimum variance in the general semiparametric situation. We discuss its properties extended to the case of observing a general Itô semimartingale possibly admitting jumps. Consistency for the entire quadratic covariation is established under mild regularity assumptions. We deduce sufficient conditions on the observation times design to establish a central limit theorem. In particular, we illustrate the formal expressions for the important (and included) setup of exogenous observation times generated by homogenous Poisson processes.

The paper is organized as follows: We review the one-dimensional results by Jacod (2008) for realized volatility in Section 2. The first generalization to non-equidistant observation times is pursued in Section 3. In Section 4 we develop the asymptotic theory for the Hayashi-Yoshida estimator and non-synchronous two-dimensional observations. The case of Poisson sampling is treated as an explicit example. Section 5 demonstrates the finite sample accuracy in Monte Carlo simulations. The proofs are given in the Appendix.

2 The baseline case: univariate regular observations

Let us start with revisiting the central limit theorem for realized variance in the presence of jumps for the regular univariate setting which has been found by Jacod (2008). Suppose in the sequel that X is a one-dimensional Itô semimartingale on $(\Omega, \mathcal{F}, \mathbb{P})$ of the form

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \int_0^t \int_{\mathbb{R}} \kappa(\delta(s,z))(\mu-\nu)(ds,dz) + \int_0^t \int_{\mathbb{R}} \kappa'(\delta(s,z))\mu(ds,dz), \quad (2.1)$$

where W is a standard Brownian motion, μ is a Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}$, and the predictable compensator ν satisfies $\nu(ds,dz) = ds \otimes \lambda(dz)$ for some σ-finite measure λ on $\mathbb{R}$ endowed with the Borelian σ-algebra. κ denotes a truncation function with $\kappa(x) = x$ on a neighbourhood of zero and we set $\kappa'(x) = x - \kappa(x)$, to separate the martingale part of small jumps and the large jumps. κ is assumed to be continuous here, which helps to simplify notation and further regularity conditions, and with compact support. We impose the following fairly general structural assumptions on the characteristics of X.

Assumption 2.1. *The processes b_s, σ_s and $s \mapsto \delta(s,z)$ are continuous. Furthermore, we have $|\delta(s,z)| \leq \gamma(z)$ for some bounded positive real-valued function γ which satisfies $\int (1 \wedge \gamma^2(z))\lambda(dz) < \infty$.*

Our target of inference is the quadratic variation of the semimartingale X at time $0 < t \leq 1$ which becomes

$$[X,X]_t = \int_0^t \sigma_s^2 ds + \sum_{s \leq t} (\Delta X_s)^2,$$

the sum of the integrated variance and the sum of squared jumps, in the setting above. Here, $\Delta X_s = X_s - X_{s-}, X_{s-} = \lim_{t \to s, t<s} X_t$, denotes the possible jump at time s. In the baseline case of equidistant observations, that is we observe X at the regular times i/n, $i = 0, \ldots, \lfloor nt \rfloor$, Jacod (2008) establishes a stable central limit theorem for the natural estimator realized variance. With $\Delta_i^n X = X_{i/n} - X_{(i-1)/n}$, the latter term is defined as

$$RV_t^n = \sum_{i=1}^{\lfloor nt \rfloor} (\Delta_i^n X)^2.$$

Before we state the result, let us shortly recall the notion of stable convergence. A family of random variables (Y_n) defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to converge $\mathcal{F}$-stably in law to Y defined on an extended space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$, if

$$\mathbb{E}[g(Y_n)S] \to \widetilde{\mathbb{E}}[g(Y)S]$$

for all bounded, continuous g and all bounded $\mathcal{F}$-measurable random variables S. For background information on the notion of stable (weak) convergence we refer interested readers to Jacod and Shiryaev (2003), Jacod and Protter (1998), Jacod (1997) and Podolskij and Vetter (2010).

In our context, the limiting variable depends on auxiliary random variables. We therefore consider a second space $(\Omega', \mathcal{F}', \mathbb{P}')$ supporting a standard Brownian motion W' and a sequence $(U_p')_{p \geq 1}$ of standard normal variables, all mutually independent. The extended space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ is then given by the (orthogonal) product of the two spaces where all variables above are extended to it in the canonical way. The limiting variables in the central limit theorem for quadratic variation are then defined as follows: Let $(S_p)_{p \geq 1}$ be a sequence of stopping times exhausting the jumps of X and set

$$Z_t = 2 \sum_{p: S_p \leq t} \Delta X_{S_p} \sigma_{S_p} U_p' \quad \text{and} \quad V_t = \sqrt{2} \int_0^t \sigma_s^2 dW_s'.$$

The stable limit theorem for quadratic variation adopted from Jacod (2008) now reads as follows:

Theorem 2.2. *Suppose that X is a one-dimensional Itô semimartingale with representation (2.1) for which Assumption 2.1 is satisfied. Then for each $0 < t \leq 1$ we have the $\mathcal{F}$-stable central limit theorem*

$$n^{1/2}(RV_t^n - [X,X]_t) \xrightarrow{\mathcal{L}-(s)} V_t + Z_t. \tag{2.2}$$

Remark 2.3. Even though Z_t might depend on the particular choice of the stopping times, it is shown in Jacod (2008) that its $\mathcal{F}$-conditional law does not. By definition of stable convergence, this is all that matters. Note also that the result above only holds for a fixed $t > 0$, but not in a functional sense, except X is continuous. This is due to the fact that a large jump at time t is by definition included in $[X,X]_t$, but usually not in RV_t^n, as the latter statistic only counts increments up to time $\lfloor nt \rfloor / n$. For a fixed t, this is not relevant, as the expectation of large jumps close to time t is small, but in a process sense this issue becomes important. One can account for this fact by subtracting $[X,X]_{\lfloor nt \rfloor / n}$ in (2.2) instead, however. □

We give a proof of Theorem 2.2 in Appendix B and C, basically for two reasons: First, the analogue of Theorem 2.2 is only a special case of the much more general discussion in Jacod (2008), and we believe that it is interesting to highlight how proofs of stable central limit theorems concerned with jumps work in this special (but nevertheless important) situation. In this sense, the first part of this paper can be understood as a follow-up to Podolskij and Vetter (2010) where the focus was on explaining limit theorems for continuous semimartingales. Second, the proof serves as foundation for all other setups where we employ the results provided for the baseline case discussed in this section.

Throughout the paper, we restrict ourselves to continuous σ. This condition can be weakened in the sense that it might be some Itô semimartingale itself. We refer to Jacod (2008) to an extension of (2.2) allowing even for common jumps of σ and X, in which the limit Z_t is slightly more complicated. Since we shall focus on the effects of irregular sampling, and also on the impact of jumps on the Hayashi-Yoshida estimator in the multivariate case, which furnish several new interesting effects, we believe this slight simplification helps to keep the asymptotic results readable and clear.

3 Asymptotics for irregular sampling schemes

The situation changes when the observations do not come at regular times anymore. In general, at stage n one observes a one-dimensional process X at arbitrary times $0 = t_0^n < t_1^n < \ldots$, which may either be deterministic or random (stopping) times, and a further distinction in the random case regards independent and endogenous sampling schemes. The latter are by far the most complicated, and it is well-known that already in the continuous case central limit theorems become non-standard for observations based e.g. on hitting times of X; see Fukasawa and Rosenbaum (2012) and related papers. For this reason, we restrict ourselves in this work to either deterministic observations times or those coming from independent random variables. Even in this case, it is hard to derive asymptotics in general, and this becomes particularly virulent in the multi-dimensional framework discussed in the next section.

We use the notation $m_-^n(t) = \max\{i : t_i^n \leq t\}, \tau_-^n(t) = \max\{t_i^n : t_i^n \leq t\}$ and $m_+^n(t) = \min\{i : t_i^n \geq t\}, \tau_+^n(t) = \min\{t_i^n : t_i^n \geq t\}$ for an arbitrary $0 \leq t \leq 1$, referring to the number of observations around time t and to the previous and next ticks. A necessary condition in order to infer on the quadratic variation of X is to secure that the mesh of the observation times $\pi_n = \max\{|t_i^n - t_{i-1}^n| | i = 1, \ldots, m_-^n(1)\}$ tends to zero (in probability) as n increases. Standard results from stochastic analysis then ensure consistency of realized variance as an estimator for the quadratic variation, which becomes

$$RV_t^n = \sum_{i=1}^{m_-^n(t)} |\Delta_i^n X|^2 \xrightarrow{\mathbb{P}} [X,X]_t$$

in this context. Here we have set $\Delta_i^n X = X_{t_i^n} - X_{t_{i-1}^n}$ again.

In order to derive a central limit theorem for RV_t^n, we need sharper bounds on the order of π_n as well as two regularity conditions on increments of the observations times. The first assumption is concerned with the variance due to the continuous martingale part, whereas the second one is about local regularity around possible jump times. It looks rather complicated, but reflects precisely what is needed to prove stable convergence later on.

Assumption 3.1. *Suppose that the variables t_i^n are stopping times which satisfy $\mathbb{E}[\pi_n^q] = o(n^{-\alpha})$ for all $q \geq 1$ and any $0 < \alpha < q$. Furthermore, we assume*

(i) that there exists a continuously differentiable function $G : [0,1] \to [0,\infty)$, such that the convergence

$$G(t) = \lim_{n\to\infty} G_n(t) := \lim_{n\to\infty} n \sum_{i=1}^{m_-^n(t)} (t_i^n - t_{i-1}^n)^2 \tag{3.1}$$

holds pointwise (in probability)

(ii) that for any $0 < t \leq 1$ and any $k \in \mathbb{N}$ we have convergence of

$$\int_{[0,t]^k} g(x_1,\dots,x_k)\mathbb{E}\Big[\prod_{p=1}^k h_p(n(\tau_+^n(x_p) - \tau_-^n(x_p)))\Big]dx_k\dots dx_1 \tag{3.2}$$

to

$$\int_{[0,t]^k} g(x_1,\dots,x_k)\prod_{p=1}^k \int_{\mathbb{R}} h_p(y)\Phi(x_p,dy)dx_k\dots dx_1 \tag{3.3}$$

as $n \to \infty$, where the $\Phi(x,dy)$ denote a family of probability measures on $[0,\infty)$ with uniformly bounded first moment and g and h_p, $p = 1,\dots,k$, are bounded continuous functions.

Note in Assumption 3.1 (ii) that the expectation of products in (3.2) becomes a product of expectations in (3.3). This means that after standardization the lengths of the intervals around the (jump) times x_p converge to independent variables, whose distributions may in general depend on x_p. The latter property reflects for example that there might be periods in which observations come more often than in others.

Example 3.2. Suppose that the sampling scheme is deterministic with $t_i^n = f(i/n)$ for some strictly isotonic, deterministic function $f : [0,1] \to [0,1]$. If f is continuously differentiable, then Assumption 3.1 is satisfied with the deterministic limits

$$G(t) = \lim_{n\to\infty} n \sum_{i=1}^{m_-^n(t)} (t_i^n - t_{i-1}^n)^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{m_-^n(t)} (f'(\xi_i^n))^2 = \int_0^t (f'(x))^2 dx.$$

In order to prove the representation (3.3) set $\eta(x) = f'(f^{-1}(x))$. Since the design is deterministic, the expectation in (3.2) can be dropped and we obtain

$$n(\tau_+^n(x) - \tau_-^n(x)) = n\left[f\left(\left\lceil nf^{-1}(x)\right\rceil/n\right) - f\left(\left\lfloor nf^{-1}(x)\right\rfloor/n\right)\right] \to f'(f^{-1}(x)) = \eta(x).$$

Therefore (3.3) holds with the deterministic $\Phi(x,dy) = \delta_{\eta(x)}(dy)$. The bound on $\mathbb{E}[\pi_n^q]$ is trivially satisfied as well. □

Example 3.3. Alternatively, one might want to work with a random observation scheme. Classical is Poisson sampling, where $(N^n)_t$ is a Poisson process with intensity $n\lambda$ for $\lambda > 0$ fixed and each n, and the observations time t_i^n is equivalent to the time of the i-th jump of N^n. In this case, we have

$$n \sum_{i=1}^{m_-^n(t)} (t_i^n - t_{i-1}^n)^2 = n \sum_{i=1}^{\lceil n\lambda t \rceil} (t_i^n - t_{i-1}^n)^2 + o_p(1) = 2t/\lambda + o_p(1),$$

since the $t_i^n - t_{i-1}^n$ form a sequence of i.i.d. $\exp(n\lambda)$-variables. We have used both Lemma 8 in Hayashi and Yoshida (2008), which states that $\mathbb{E}[\pi_n^q] = o(n^{-\alpha})$ is satisfied, and arguments from the proof of Lemma 10, which show that $m_-^n(t)$ is close to $\lceil n\lambda t \rceil$, to obtain the first relation above. Let us now derive the limit of

$$\mathbb{E}\Big[\prod_{p=1}^{k} h_p(n(\tau_+^n(x_p) - \tau_-^n(x_p))) \Big]$$

for any fixed $x_1, \dots, x_k$, and we start with $k = 1$. First, due to memorylessness $n(\tau_+^n(x_1) - x_1) \sim \exp(\lambda)$. On the other hand, a standard result in renewal theory (see e.g. Cox (1970); page 31) gives the distribution of the backward recurrence time of the Poisson process:

$$\mathbb{P}(n(x_1 - \tau_-^n(x_1)) \le u) = \begin{cases} 1, & u = nx_1, \\ 1 - e^{-\lambda u}, & 0 < u < nx_1. \end{cases}$$

Therefore, $n(x_1 - \tau_-^n(x_1)) \xrightarrow{w} \exp(\lambda)$ as $n \to \infty$, and from the strong Markov property, which secures independence of the two summands, we have $n(\tau_+^n(x_1) - \tau_-^n(x_1)) \xrightarrow{w} \Gamma(2, \lambda)$. Similarly, for a general k, one can show that the $n(\tau_+^n(x_k) - \tau_-^n(x_k))$ are asymptotically independent, and all sequences of random variables obviously have the same limiting distribution. Condition (3.3) is therefore valid with $\Phi(x, dy)$ being the distribution of a $\Gamma(2, \lambda)$ variable for all x. □

Example 3.4. As a third example we consider a deterministic irregular scheme with a truly random limiting distribution $\Phi(x, dy)$. Consider the sequence of observation times $t_i^n = i/n, i = 2, 4, \dots$, for even numbers and $t_i^n = (i + \alpha)/n, i = 1, 3, \dots$, for odd numbers with some $0 < \alpha < 1$. For fixed x_1, interval lengths $n(\tau_+(x_1) - \tau_-(x_1))$ can alternate between $(1+\alpha)$ and $(1-\alpha)$ and do not converge. This is where a random limit comes into play: Let us again discuss $k = 1$ in detail. Setting $[0, t] = A_n \cup B_n$, where A_n denotes the subset on which $n(\tau_+(x_1) - \tau_-(x_1)) = 1 + \alpha$ and B_n the one with $n(\tau_+(x_1) - \tau_-(x_1)) = 1 - \alpha$, we obtain from continuity of g

$$\begin{aligned} \int_{[0,t]^k} g(x_1) h_1(n(\tau_+^n(x_1) - \tau_-^n(x_1)))dx_1 &= h_1(1+\alpha) \int_{A_n} g(x_1)dx_1 + h_1(1-\alpha) \int_{B_n} g(x_1)dx_1 \\ &\sim \frac{1}{t}(h_1(1+\alpha)\lambda(A_n) + h_1(1-\alpha)\lambda(B_n)) \int_0^t g(x_1)dx_1 \\ &\to (h_1(1+\alpha)(1+\alpha)/2 + h_1(1-\alpha)(1-\alpha)/2) \int_0^t g(x_1)dx_1, \end{aligned}$$

where λ denotes the Lebesgue measure. Thus, $\Phi(x, dy)$ is again independent of x and has two atoms taking the value $(1 + \alpha)$ with probability $(1 + \alpha)/2$ and the value $(1 - \alpha)$ with probability $(1 - \alpha)/2$. A generalization to arbitrary k is straightforward. Note also that the condition on π_n is satisfied by definition and that (3.1) holds with $G(t) = (1 + \alpha^2)t$. □

Let us now extend (2.2) to this framework. Thereto, denote with $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$ a probability space on which X is defined, and we assume all observation times t_i^n to live on $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$. We can now define

$(\Omega, \mathcal{F}, \mathbb{P})$ as the product space of these two, while $(\Omega', \mathcal{F}', \mathbb{P}')$ is defined similarly as before, but it is assumed to accommodate independent random variables $(\eta(x))_{0\le x\le t}$ as well, with distribution $\Phi(x, dy)$ as in (3.3). $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ finally is the orthogonal product of the latter two spaces again.

Theorem 3.5. *Suppose that X is a one-dimensional Itô semimartingale with representation (2.1) for which Assumption 2.1 is satisfied. If also Assumption 3.1 on the observation scheme holds, then for each $0 < t \le 1$ we have the $\mathcal{F}^0$-stable central limit theorem*

$$n^{1/2}(RV_t^n - [X, X]_t) \overset{\mathcal{L}-(s)}{\longrightarrow} \widetilde{V}_t + \widetilde{Z}_t, \tag{3.4}$$

where

$$\widetilde{V}_t = \sqrt{2}\int_0^t \sigma_s^2 (G'(s))^{1/2} dW'_s$$

and

$$\widetilde{Z}_t = 2\sum_{p:S_p\le t} \Delta X_{S_p} \big(\eta(S_p)\big)^{1/2} \sigma_{S_p} U'_p.$$

Here, the S_p are stopping times exhausting the jumps of X and the (U'_p) are i.i.d. standard normal on $(\Omega', \mathcal{F}', \mathbb{P}')$ as before.

This theorem has already been known in the literature, if X is a continuous process; see e.g. the survey by Mykland and Zhang (2012).

Remark 3.6. Both limiting processes look similar to the ones obtained in Theorem 2.2, apart from different standardizations due to irregular sampling. What is interesting, however, is the nature of the scaling in the part due to jumps. The schemes considered in Example 3.2 are locally regular, which leads to deterministic $\eta(S_p)$ as well. On the other hand, both the Poisson sampling and the deterministic design in Example 3.4 show local irregularities, resulting in random (but time-homogeneous) limits $\eta(S_p)$. Nevertheless, we still have regularity on a global level even for these sampling schemes, leading to a deterministic limit of $G_n(t)$ in all three cases. □

4 Asymptotics in the multivariate case

This section is devoted to non-synchronous discrete observations of a multi-dimensional Itô semimartingale with jumps. It is informative to stick to a two-dimensional setting and an underlying semimartingale of similar form as (2.1) before:

$$\begin{aligned} X_t &= \left(X^{(1)}, X^{(2)}\right)_t^\top \\ &= X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \int_0^t \int_{\mathbb{R}^2} \kappa(\delta(s,z))(\mu-\nu)(ds,dz) + \int_0^t \int_{\mathbb{R}^2} \kappa'(\delta(s,z))\mu(ds,dz), \end{aligned} \tag{4.1}$$

where $W = \left(W^{(1)}, W^{(2)}\right)^\top$ denotes a two-dimensional standard Brownian motion, and we assume without loss of generality

$$\sigma_s = \begin{pmatrix} \sigma_s^{(1)} & 0 \\ \rho_s \sigma_s^{(2)} & \sqrt{1-\rho_s^2}\sigma_s^{(2)} \end{pmatrix} \text{ such that } \sigma_s \sigma_s^\top = \begin{pmatrix} \left(\sigma_s^{(1)}\right)^2 & \rho_s \sigma_s^{(1)} \sigma_s^{(2)} \\ \rho_s \sigma_s^{(1)} \sigma_s^{(2)} & \left(\sigma_s^{(2)}\right)^2 \end{pmatrix},$$

while the other characteristics are defined analogously to Section 2 with two-dimensional jump measures. Denote with $\|\cdot\|$ the spectral norm. We develop a theory for general jump measures comprising co-jumps ($X^{(1)}$ and $X^{(2)}$ jump at the same time) and idiosyncratic jumps of the components.

We investigate the estimator by Hayashi and Yoshida (2005), called HY-estimator in the following, under the influence of jumps. The HY-estimator has been proposed and is well-studied for integrated covolatility estimation from asynchronous observations of a continuous Itô semimartingale; see Hayashi and Yoshida (2008) and Hayashi and Yoshida (2011). Our structural hypothesis for the characteristics of X reads similar as Assumption 2.1 in Section 2:

Assumption 4.1. *Assume that b_s, $\sigma_s^{(1)}$, $\sigma_s^{(2)}$, ρ_s and $s \mapsto \delta(s,z)$ are continuous and that $\|\delta(s,z)\| \leq \gamma(z)$ for a bounded positive real-valued function γ which satisfies $\int (1 \wedge \gamma^2(z))\lambda(dz) < \infty$.*

By Itô isometry, we may expect that in the presence of jumps the HY-estimator is suitable for estimating the entire quadratic covariation (1.1). Yet, there are several open questions which we address in this section and an asymptotic distribution theory of the HY-estimator with jumps is unexplored territory.

4.1 Discussion of the HY-estimator and notation

The HY-estimator is the sum of products of increments with overlapping observation time instants:

$$\Big[\widehat{X^{(1)}, X^{(2)}}\Big]_t^{(HY),n} = \sum_{t_i^{(1)} \leq t} \sum_{t_j^{(2)} \leq t} \Big(X^{(1)}_{t_i^{(1)}} - X^{(1)}_{t_{i-1}^{(1)}}\Big)\Big(X^{(2)}_{t_j^{(2)}} - X^{(2)}_{t_{j-1}^{(2)}}\Big) 1_{\Big\{ \min\Big(t_i^{(1)}, t_j^{(2)}\Big) > \max\Big(t_{i-1}^{(1)}, t_{j-1}^{(2)}\Big)\Big\}}, \quad (4.2a)$$

when $X^{(l)}, l = 1,2$, is observed at times $t_i^{(l)}$. In the sequel, we introduce the notion of several interpolation functions and sequences dependent on the observation times. Let $\pi_n = \max_{i,l}\{|t_i^{(l)} - t_{i-1}^{(l)}|\}$ denote the mesh. We define

$$\tau_+^{(l)}(s) = \min_{i \in \{0,\ldots,n_l\}} \Big(t_i^{(l)} | t_i^{(l)} \geq s\Big), \quad m_+^{(l)}(s) = \min\Big(i | t_i^{(l)} \geq s\Big) \quad \text{and}$$
$$\tau_-^{(l)}(s) = \max_{i \in \{0,\ldots,n_l\}} \Big(t_i^{(l)} | t_i^{(l)} \leq s\Big), \quad m_-^{(l)}(s) = \max\Big(i | t_i^{(l)} \leq s\Big)$$

for $l = 1,2$, and $s \in [0,1]$. Let us further introduce the shortcuts

$$\tau_{++}^{(l,r)}(s) = \tau_+^{(l)}\Big(\tau_+^{(r)}(s)\Big) \quad \text{and} \quad \tau_{--}^{(l,r)}(s) = \tau_-^{(l)}\Big(\tau_-^{(r)}(s)\Big)$$

and $\tau_{++}^{(r,l)}(s), \tau_{--}^{(r,l)}(s)$, analogously. A synchronous grid serving as a reference scheme is given by the sequence of refresh times

$$T_k = \max\Big(\tau_+^{(1)}(T_{k-1}), \tau_+^{(2)}(T_{k-1})\Big), \ k = 0, \ldots, M^n(1),$$

with the convention $T_{-1} = 0$ and where we denote with $M^n(t)$ the number of refresh times smaller than or equal to $t \in [0,1]$. Each increment $T_k - T_{k-1}$ thus is the waiting time until both components of X have been observed again. The use of refresh times is adopted from Barndorff-Nielsen et al. (2011) where the same synchronous scheme is employed in a more general way. For notational convenience, indices referring to dependence on n for sampling times are often surpressed in the multi-dimensional setup.

Based on telescoping sums, the HY-estimator (4.2a) can be rewritten

$$\Big[\widehat{X^{(1)},X^{(2)}}\Big]_t^{(HY),n} = \sum_{i=1}^{m_-^{(1)}(t)-1} \Big(X^{(1)}_{t_i^{(1)}} - X^{(1)}_{t_{i-1}^{(1)}}\Big)\Big(X^{(2)}_{\tau_+^{(2)}(t_i^{(1)})} - X^{(2)}_{\tau_-^{(2)}(t_{i-1}^{(1)})}\Big) + \mathcal{O}_p(\pi_n) \tag{4.2b}$$

$$= \sum_{j=1}^{m_-^{(2)}(t)-1} \Big(X^{(2)}_{t_j^{(2)}} - X^{(2)}_{t_{j-1}^{(2)}}\Big)\Big(X^{(1)}_{\tau_+^{(1)}(t_j^{(2)})} - X^{(1)}_{\tau_-^{(1)}(t_{j-1}^{(2)})}\Big) + \mathcal{O}_p(\pi_n) \tag{4.2c}$$

$$= \sum_{k=1}^{M^n(t)-1} \Big(X^{(1)}_{\tau_+^{(1)}(T_k)} - X^{(1)}_{\tau_-^{(1)}(T_{k-1})}\Big)\Big(X^{(2)}_{\tau_+^{(2)}(T_k)} - X^{(2)}_{\tau_-^{(2)}(T_{k-1})}\Big) + \mathcal{O}_p(\pi_n). \tag{4.2d}$$

The $\mathcal{O}_p(\pi_n)$ terms in (4.2b)–(4.2d) are only due to possible end effects at time t. Apart from this, the above equalities hold exactly. Illustrations (4.2b)–(4.2d) reveal that the estimation error of the HY-estimator can be decomposed in the one of a usual synchronous-type realized covolatility and an additional error induced by non-synchronicity and interpolations. To simplify notation a bit, we write from now on

$$\Delta_i^{n_1}X^{(1)} = \Big(X^{(1)}_{t_i^{(1)}} - X^{(1)}_{t_{i-1}^{(1)}}\Big), \quad \Delta_j^{n_2}X^{(2)} = \Big(X^{(2)}_{t_j^{(2)}} - X^{(2)}_{t_{j-1}^{(2)}}\Big), \quad \Delta_k^{n}X^{(l)} = \Big(X^{(l)}_{T_k} - X^{(l)}_{T_{k-1}}\Big),\ l=1,2,$$

and for previous and next-tick interpolations with respect to the refresh time scheme

$$\Delta_k^{+,n}X^{(l)} = \Big(X^{(l)}_{\tau_+^{(l)}(T_k)} - X^{(l)}_{T_k}\Big), \quad \Delta_k^{-,n}X^{(l)} = \Big(X^{(l)}_{T_{k-1}} - X^{(l)}_{\tau_-^{(l)}(T_{k-1})}\Big),\ l=1,2.$$

Also, we denote with Δ_k^n the refresh time instants $(T_k - T_{k-1})$ and $\Delta_k^{+,n_l} = (\tau_+^{(l)}(T_k) - T_k)$ are the next- and $\Delta_k^{-,n_l} = (T_{k-1} - \tau_-^{(l)}(T_{k-1}))$ the previous-tick interpolations.

When decomposing X in different terms by the continuous part, jumps and cross terms, we can use any of the illustrations (4.2a)–(4.2d) to analyze those terms. Therefore, to gain deeper insight and to get used to the notation, let us delve into the different ways to illustrate and construct the HY-estimator:

(4.2a) This is the original idea to sum all products of increments, belonging to time intervals between adjacent observations which have a non-empty intersection.

(4.2b) Trace out all increments of $X^{(1)}$ and sum up products with the interpolated increments of $X^{(2)}$: $\Delta_i^{n_1}X^{(1)}\Big(X^{(2)}_{\tau_+^{(2)}(t_i^{(1)})} - X^{(2)}_{\tau_-^{(2)}(t_{i-1}^{(1)})}\Big),\ i=1,\dots,m_-^{(1)}(t)-1.$

(4.2c) Trace out all increments of $X^{(2)}$ and sum up products with the interpolated increments of $X^{(1)}$: $\Delta_j^{n_2}X^{(2)}\Big(X^{(1)}_{\tau_+^{(1)}(t_j^{(2)})} - X^{(1)}_{\tau_-^{(1)}(t_{j-1}^{(2)})}\Big),\ j=1,\dots,m_-^{(2)}(t)-1.$

(4.2d) Consider the refresh times grid and sum up products of interpolated increments of $X^{(1)}$ and $X^{(2)}$: $\Big(\Delta_k^{+,n}X^{(1)} + \Delta_k^{n}X^{(1)} + \Delta_k^{-,n}X^{(1)}\Big)\Big(\Delta_k^{+,n}X^{(2)} + \Delta_k^{n}X^{(2)} + \Delta_k^{-,n}X^{(2)}\Big),\ k=1,\dots,M^n(t)-1.$ At least one of the previous-tick and one of the next-tick interpolations equal zero.

Example 4.2. To illuminate the transformations between (4.2a)–(4.2d) by rearranging addends, particularly in the presence of a jump, we examine a small example displayed in Figure 1. Focusing on the increment $\Delta_4^{n_1}X^{(1)}$ that contains a jump, (4.2a) tells us that this increments is considered in the addends

$$\Delta_4^{n_1}X^{(1)}\Delta_3^{n_2}X^{(2)} + \Delta_4^{n_1}X^{(1)}\Delta_4^{n_2}X^{(2)} + \Delta_4^{n_1}X^{(1)}\Delta_5^{n_2}X^{(2)}.$$

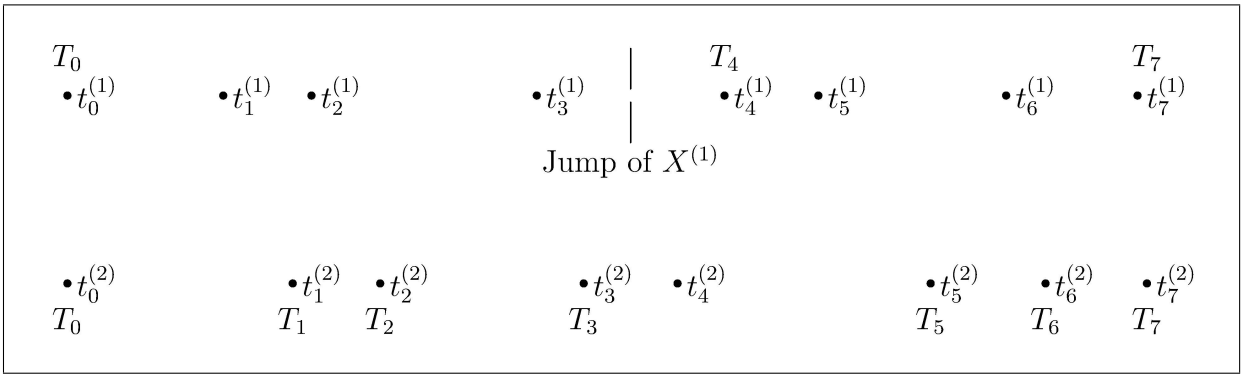


Figure 1: Example of observation times allocation.

If we start with illustration (4.2b), we directly obtain

$$\Delta_4^{n_1}X^{(1)}\Big(X^{(2)}_{\tau_+^{(2)}(t_4^{(1)})}-X^{(2)}_{\tau_-^{(2)}(t_3^{(1)})}\Big)=\Delta_4^{n_1}X^{(1)}\Big(\Delta_3^{n_2}X^{(2)}+\Delta_4^{n_2}X^{(2)}+\Delta_5^{n_2}X^{(2)}\Big)$$

as well. This is the illustration we prefer to analyze the jumps in $X^{(1)}$. Starting with the symmetric illustration (4.2c), the same terms appear, but rearranged in a different way and in several addends:

$$\begin{aligned}&\Delta_3^{n_2}X^{(2)}\Big(X^{(1)}_{\tau_+^{(1)}(t_3^{(2)})}-X^{(1)}_{\tau_-^{(1)}(t_2^{(2)})}\Big)+\Delta_4^{n_2}X^{(2)}\Big(X^{(1)}_{\tau_+^{(1)}(t_4^{(2)})}-X^{(1)}_{\tau_-^{(1)}(t_3^{(2)})}\Big)+\Delta_5^{n_2}X^{(2)}\Big(X^{(1)}_{\tau_+^{(1)}(t_5^{(2)})}-X^{(1)}_{\tau_-^{(1)}(t_4^{(2)})}\Big)\\&\quad=\Delta_3^{n_2}X^{(2)}\Big(\Delta_3^{n_1}X^{(1)}+\Delta_4^{n_1}X^{(1)}\Big)+\Delta_4^{n_2}X^{(2)}\Delta_4^{n_1}X^{(1)}+\Delta_5^{n_2}X^{(2)}\Big(\Delta_4^{n_1}X^{(1)}+\Delta_5^{n_1}X^{(1)}+\Delta_6^{n_1}X^{(1)}\Big).\end{aligned}$$

This illustration simplifies treatment of jumps in $X^{(2)}$. Finally, from the refresh time illustration (4.2d) we find the same terms in the addends

$$\begin{aligned}&\Big(\Delta_4^{-,n}X^{(1)}+\Delta_4^{n}X^{(1)}\Big)\Big(\Delta_4^{+,n}X^{(2)}+\Delta_4^{n}X^{(2)}\Big)+\Delta_3^{n}X^{(2)}\Big(\Delta_3^{+,n}X^{(1)}+\Delta_3^{n}X^{(1)}+\Delta_3^{-,n}X^{(1)}\Big)\\&\quad=\Delta_4^{n_1}X^{(1)}\Big(\Delta_4^{n_2}X^{(2)}+\Delta_5^{n_2}X^{(2)}\Big)+\Delta_3^{n_2}X^{(2)}\Big(\Delta_3^{n_1}X^{(1)}+\Delta_4^{n_1}X^{(1)}\Big).\end{aligned}$$

The effect of a jump is free from the particular illustration. It is convenient to consider the partition $[t_{i-1}^{(1)},t_i^{(1)})$, $i=1,\dots m_-^{(1)}(1)$, when we trace out jumps of $X^{(1)}$ and $[t_{j-1}^{(2)},t_j^{(2)})$, $j=1,\dots m_-^{(2)}(1)$, for jumps of $X^{(2)}$, while we use $[T_{k-1},T_k)$, $k=1,\dots,M^n(1)$, for the continuous part. The main reason for the latter is that the estimation error can be written as sum of martingale differences when using refresh times; see Bibinger (2011) for details. □

4.2 Asymptotic theory

Say a co-jump occurs at time S_p. As can be seen from representation (4.2b), the jump in $X^{(1)}$ is multiplied in the cross term with the increment of $X^{(2)}$ over the interpolated interval $[\tau_{--}^{(2,1)}(S_p),\tau_{++}^{(2,1)}(S_p)]$, and for the jump of $X^{(2)}$ symmetrically. The products are marked in Figure 2 by the arcs and dashed segments, respectively. Idiosyncratic jumps are included in the general consideration by setting the jump in one component equal to zero. Similarly to the univariate case, the part due to jumps in the limiting variable is coming from a mixture of the particular jump of one process and the continuous increment of the other. Therefore, quantities like the length of $[\tau_{--}^{(2,1)}(S_p),\tau_{++}^{(2,1)}(S_p)]$ determine the contribution of one particular jump in the asymptotic variance.

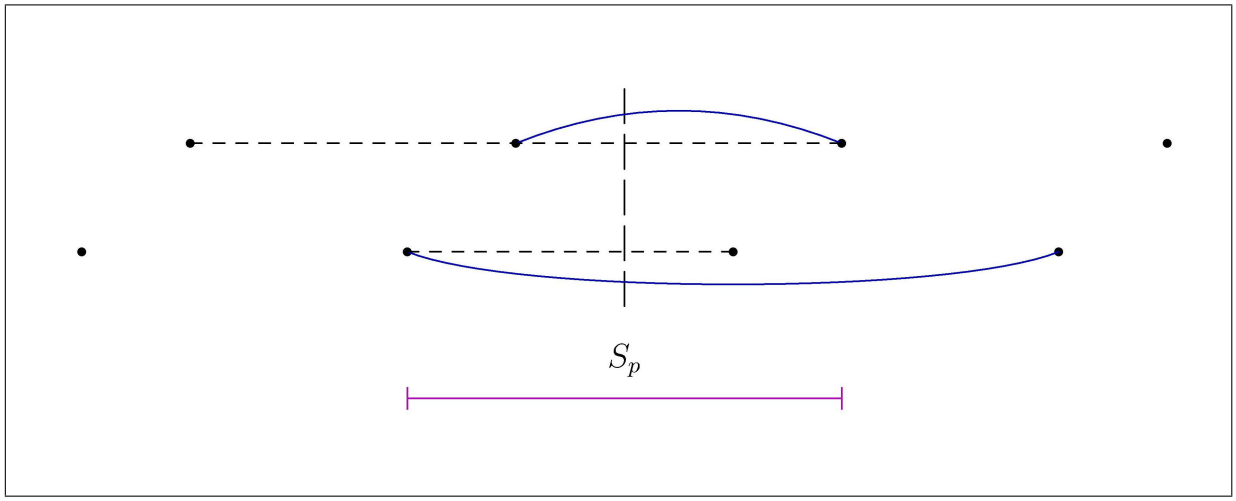


Figure 2: A co-jump and intervals that determine the covariance structure.

An intriguing effect arises by co-jumps in the multi-dimensional setting induced by the covariance of the two respective cross terms, since $d[X^{(1)}, X^{(2)}]^c_s = \sigma^{(1)}_s \sigma^{(2)}_s \rho_s \, ds$. The covariance hinges on the intersection of both interpolated intervals in the two cross terms and results in an auxiliary condition such that the variance of the HY-estimator converges. In Figure 2 this intersection is highlighted by the segment with bars. In any case, the following five intervals determine the variance of the HY-estimator by one particular co-jump at $s \in [0,1]$:

$$\left(\mathcal{R}^1_n + \mathcal{L}^1_n\right)(s) = \max\left(\tau^{(1)}_+(s), \tau^{(2)}_+(s)\right) - \min\left(\tau^{(1)}_-(s), \tau^{(2)}_-(s)\right), \tag{4.3a}$$

$$\mathcal{R}^2_n(s) = \tau^{(1,2)}_{++}(s) - \max\left(\tau^{(1)}_+(s), \tau^{(2)}_+(s)\right), \tag{4.3b}$$

$$\mathcal{R}^3_n(s) = \tau^{(2,1)}_{++}(s) - \max\left(\tau^{(1)}_+(s), \tau^{(2)}_+(s)\right), \tag{4.3c}$$

$$\mathcal{L}^2_n(s) = \min\left(\tau^{(1)}_-(s), \tau^{(2)}_-(s)\right) - \tau^{(1,2)}_{--}(s), \tag{4.3d}$$

$$\mathcal{L}^3_n(s) = \min\left(\tau^{(1)}_-(s), \tau^{(2)}_-(s)\right) - \tau^{(2,1)}_{--}(s). \tag{4.3e}$$

Either (4.3b) or (4.3c) is zero (both only in case of a synchronous observation), and the same is true for (4.3d) and (4.3e). Yet, at each jump arrival S_p, we need to distinguish if $\mathcal{R}^2_n(S_p) > 0$ or $\mathcal{R}^3_n(S_p) > 0$. The segment with bars in Figure 2 corresponds to $(\mathcal{R}^1_n + \mathcal{L}^1_n)(S_p)$.

To derive a central limit theorem for the HY-estimator, already in the purely continuous case certain regularity conditions on the sequences of observation times are required; see Hayashi and Yoshida (2011). The analogous conditions using illustration (4.2d) from Bibinger (2011) and additional conditions that ensure convergence of the variance in the presence of jumps are gathered in the next assumption:

Assumption 4.3. *Assume the $t^{(l)}_i, l = 1, 2$, are stopping times such that $\mathbb{E}[\pi^q_n] = o(n^{-\alpha})$ for all $q \geq 1$ and any $0 < \alpha < q$.*

(i) Suppose that the functional sequences

$$G_n(t) = n \sum_{T_k \leq t} (\Delta^n_k)^2, \tag{4.4a}$$

$$F_n(t) = n \sum_{T_{k+1}\le t} \Big((\Delta_k^n + \Delta_k^{-,n_2})\Delta_k^{+,n_1} + \Delta_k^{+,n_2}(\Delta_k^n + \Delta_k^{-,n_1}) + \Delta_{k+1}^n(\Delta_{k+1}^{-,n_1} + \Delta_{k+1}^{-,n_2}) \Big), \tag{4.4b}$$

$$H_n(t) = n \sum_{T_{k+1}\le t} (\Delta_k^{-,n_1}\Delta_k^{+,n_1} + \Delta_k^{-,n_2}\Delta_k^{+,n_2}), \tag{4.4c}$$

converge, i.e. satisfy $G_n(t) \to G(t)$ pointwise for some continuously differentiable limiting function G, and analogously for H_n, F_n with limits H, F.

(ii) Assume, for any $0 < t \le 1$ and any $k \in \mathbb{N}$, convergence of

$$\int_{[0,t]^k} g(x_1,\dots,x_k)\mathbb{E}\Big[\prod_{p=1}^k h_p\Big(\Big(n(\mathcal{R}_n^1 + \mathcal{L}_n^1), n\mathcal{R}_n^2, n\mathcal{L}_n^2, n\mathcal{R}_n^3, n\mathcal{L}_n^3\Big)(x_p)\Big)\Big] dx_k \dots dx_1, \tag{4.5}$$

with the expressions introduced in (4.3a)–(4.3e)*, to*

$$\int_{[0,t]^k} g(x_1,\dots,x_k) \prod_{p=1}^k \int_{\mathbb{R}} h_p(y_1,y_2,y_3,y_4,y_5)\mathbf{\Phi}(x_p, d\mathbf{y}) dx_k \dots dx_1 \tag{4.6}$$

for some family of probability measures $\mathbf{\Phi}$ on $[0,\infty)^5$ with finite first moment, holds true as $n \to \infty$, for all bounded continuous functions g and h_p, $p = 1,\dots,k$.

We now introduce limiting variables for the central limit theorem in the two-dimensional case. They rely on the functions from Assumption 4.3. Denote

$$\tilde{\mathcal{V}}_t = \int_0^t v_s \, dW_s'$$

with the variance process

$$v_s^2 = G'(s)\Big(\sigma_s^{(1)}\sigma_s^{(2)}\Big)^2(1+\rho_s^2) + F'(s)\Big(\sigma_s^{(1)}\sigma_s^{(2)}\Big)^2 + 2H'(s)\Big(\rho_s\sigma_s^{(1)}\sigma_s^{(2)}\Big)^2. \tag{4.7}$$

The limit of the cross term is

$$\begin{aligned}\tilde{\mathcal{Z}}_t = &\sum_{p:S_p\le t} \Delta X_{S_p}^{(1)}\sigma_{S_p}^{(2)}\Big(\sqrt{(\mathcal{R}^1+\mathcal{L}^1)(S_p)}U_p^{(1)} + \sqrt{\mathcal{R}^3(S_p)}U_p^{(3)} + \sqrt{\mathcal{L}^3(S_p)}Q_p^{(3)}\Big) \\ &+ \Delta X_{S_p}^{(2)}\sigma_{S_p}^{(1)}\Big(\sqrt{(\mathcal{R}^1+\mathcal{L}^1)(S_p)}\Big(\rho_{S_p}U_p^{(1)} + \sqrt{1-\rho_{S_p}^2}Q_p^{(1)}\Big) + \sqrt{\mathcal{R}^2(S_p)}U_p^{(2)} + \sqrt{\mathcal{L}^2(S_p)}Q_p^{(2)}\Big),\end{aligned} \tag{4.8}$$

where S_p denotes some enumeration of all times where at least one process jumps (so certain addends may become zero if a jump is idiosyncratic). Again, we need a second probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ on which mutually independent standard normal variables $(U_p^{(1)}, U_p^{(2)}, U_p^{(3)}, Q_p^{(1)}, Q_p^{(2)}, Q_p^{(3)})$, $p \ge 1$, and $(\mathcal{R}^1+\mathcal{L}^1, \mathcal{R}^2, \mathcal{L}^2, \mathcal{R}^3, \mathcal{L}^3)(x) \sim \mathbf{\Phi}(x, d\mathbf{y})$ for all $x \in [0,t]$ are defined.

The second and third summand of (4.7) give the limiting asymptotic variance of the error due to asynchronicity in the continuous part, i.e. the second term comes from the variance of the interpolation steps in the addends of (4.2d) and does not depend on the correlation whereas the third addend comes from the covariance between successive summands in (4.2d). We refer to Bibinger (2011) for further details and examples for the asymptotic theory concerning the continuous semimartingale part.

Theorem 4.4. *Suppose we observe a two-dimensional Itô semimartingale* (4.1) *whose characteristics fulfill the structural Assumption 4.1. If the mesh of discrete observations tends to zero in both components, for each* $0 \le t \le 1$*, the HY-estimator consistently estimates the quadratic covariation*

$$\big[\widehat{X^{(1)}, X^{(2)}}\big]_t^{(HY),n} \xrightarrow{\mathbb{P}} [X^{(1)}, X^{(2)}]_t \,. \tag{4.9}$$

On Assumption 4.3, an $\mathcal{F}^0$*-stable central limit theorem applies:*

$$\sqrt{n}\big[\widehat{X^{(1)}, X^{(2)}}\big]_t^{(HY),n} \xrightarrow{\mathcal{L}-(s)} \tilde{\mathcal{V}}_t + \tilde{\mathcal{Z}}_t \,. \tag{4.10}$$

Example 4.5. Let us discuss a particular example again. By virtue of the symmetry,

$$\mathbb{P}\Big(n\mathcal{L}_n^1(s) \le u\Big) = \begin{cases} 1, & u = ns, \\ \mathbb{P}\Big(n\mathcal{R}_n^1(s-u) \le u\Big) & 0 < u < ns, \end{cases}$$

when observation times are modeled by renewal processes (if we have time-homogeneity), which means that under high-frequency asymptotics the distributions of backward and forward waiting times are asymptotically equal. Consequently, this is valid for time-homogenous Poisson sampling. Furthermore, by the strong Markov property both the $\mathcal{R}_n$ and $\mathcal{L}_n$, and the $\mathcal{R}_n^1$ and $(\mathcal{R}_n^2, \mathcal{R}_n^3)$ variables are independent.

Precisely, suppose that $(N_l^n)_t$, $l = 1, 2$, are independent Poisson processes with intensities $n\lambda_l$ for $\lambda_l > 0$ fixed and each n, and that the observations times are equivalent to the jump times of the respective processes. We find that constantly in time

$$\mathbb{P}\Big(\mathcal{R}_n^1 \le u_1\Big) = \Big(1 - \exp(-n\lambda_1 u_1)\Big)\Big(1 - \exp(-n\lambda_2 u_1)\Big),$$

since the two Poisson processes are independent. Note further that one of the components of $(\mathcal{R}_n^2, \mathcal{R}_n^3)$ is zero, depending on which process is observed first after time $\max\big(\tau_+^{(1)}(s), \tau_+^{(2)}(s)\big) = r$. The other one follows an exponential distribution then. Precisely, we have

$$\mathbb{P}\Big(\tau_+^{(1)}(r) \ge \tau_+^{(2)}(r)\Big) = \int_0^\infty \int_0^y \lambda_1 \exp(-n\lambda_1 y)\lambda_2 \exp(-n\lambda_2 x)\, dx\, dy = \lambda_2(\lambda_1 + \lambda_2)^{-1},$$

thus

$$\mathbb{P}\Big(\mathcal{R}_n^2 \le u_2, \mathcal{R}_n^3 \le u_3\Big) = \lambda_1(\lambda_1 + \lambda_2)^{-1}\Big(1 - \exp(-n\lambda_1 u_2)\Big) + \lambda_2(\lambda_1 + \lambda_2)^{-1}\Big(1 - \exp(-n\lambda_2 u_3)\Big),$$

from which we obtain

$$\begin{aligned}\mathbb{P}\Big(\mathcal{R}_n^1 \le u_1, \mathcal{R}_n^2 \le u_2, \mathcal{R}_n^3 \le u_3\Big) &= \Big(1 - \exp(-n\lambda_1 u_1)\Big)\Big(1 - \exp(-n\lambda_2 u_1)\Big) \\ &\quad \times \Big(1 - \lambda_1(\lambda_1 + \lambda_2)^{-1}\exp(-n\lambda_1 u_2) - \lambda_2(\lambda_1 + \lambda_2)^{-1}\exp(-n\lambda_2 u_3)\Big).\end{aligned}$$

This joint law enters the random limiting variance of the HY-estimator with jumps for Poisson sampling, that is

$$\begin{aligned}\mathbb{P}\Big(\mathcal{R}^1 \le u_1, \mathcal{R}^2 \le u_2, \mathcal{R}^3 \le u_3\Big) &= \Big(1 - \exp(-\lambda_1 u_1)\Big)\Big(1 - \exp(-\lambda_2 u_1)\Big) \\ &\quad \times \Big(1 - \lambda_1(\lambda_1 + \lambda_2)^{-1}\exp(-\lambda_1 u_2) - \lambda_2(\lambda_1 + \lambda_2)^{-1}\exp(-\lambda_2 u_3)\Big)\end{aligned}$$

and the vector $(\mathcal{L}^1, \mathcal{L}^2, \mathcal{L}^3)$ follows the same distribution.

For example, if $\lambda_1 = \lambda_2 = 1$, we obtain $4\Big(\sigma_{S_p}^{(1)}\Delta X_{S_p}^{(2)}\Big)^2 + 4\Big(\sigma_{S_p}^{(2)}\Delta X_{S_p}^{(1)}\Big)^2 + 6\rho_{S_p}\sigma_{S_p}^{(1)}\sigma_{S_p}^{(2)}\Delta X_{S_p}^{(1)}\Delta X_{S_p}^{(2)}$ at each jump time S_p as the conditional expectation of the induced variance of the HY-estimator. In this case the covariation of time functions (4.4a)-(4.4c) are $G(t) = (14/9)t$, $H(t) = (2/9)t$, $F(t) = (10/9)t$; for the latter compare Bibinger (2011) and Hayashi and Yoshida (2011). □

An estimator for the asymptotic variance of the HY-estimator facilitating a so-called feasible limit theorem to construct confidence intervals can be deduced similarly as in the continuous framework, pursued in Hayashi and Yoshida (2011) by a kernel-type and Bibinger (2011) by a histogram-type approach. See also Veraart (2010) for the univariate setting including jumps with regular observations.

5 Simulations

In this section we inspect the HY-estimator's finite-sample accuracy and confirm our theoretical findings in a Monte Carlo simulation study. For this purpose, we implement a two-dimensional semimartingale model with constant drift $b = (0.1, 0.1)^\top$ and $\sigma^{(1)} = \sigma^{(2)} = 1$ and a correlation parameter ρ. As a jump component we generate from a conditional Poisson process – with a fixed number of jumps at times uniformly distributed on $[0, 1]$ and jump heights 1, which allows for a simple tracking of the estimator's variance for different parameters and a comparison to the theoretical values. The non-synchronous observation times are randomly generated from two independent homogenous Poisson processes with expected time instants 1/30000 for each (such that in average $m_-^{(1)}(1) = m_-^{(2)}(1) = 30000$ observations on $[0, 1]$).
In each iteration the observation scheme is newly generated. We run 500 Monte Carlo iterations for each configuration. From the simulated continuous component we evaluate the HY-estimator without jumps and from the whole process the HY estimator with jumps. We consider three simple setups:

[Sc1] One co-jump (size one in both components) occurs.

[Sc2] One co-jump and one idiosyncratic jump in $X^{(1)}$ and one in $X^{(2)}$ (all size one) occur.

[Sc3] Only one idiosyncratic jump in $X^{(1)}$ and one in $X^{(2)}$ (all size one) occur.

The results are displayed in Table 1. First of all they confirm that the HY-estimator is eligible to consistently estimate the entire quadratic covariation in each scenario. In particular, the Monte Carlo averages show that the HY-estimator closely tracks the integrated covolatility in Scenario 3, robustly to the simulated large idiosyncratic jumps. We draw a comparison between the empirical finite-sample values and theoretical quantities obtained from the expectation of the random asymptotic variance and the convergence rate. Theoretical expected values are given in parentheses in Table 1. The variances of the continuous part and the variance due to the cross term by jumps and the continuous component are close to their theoretical counterparts. In Scenario 3 empirical variances are slightly larger than expected. The estimator's variance does not depend on the correlation here, since only idiosyncratic jumps occur. This is apparent also for the empirical figures. In Scenarios 1 and 2 one addend of the variance of the cross term increases linearly with the correlation what we witness as well for the finite-sample results.

	[Sc1]				[Sc2]		[Sc3]	
ρ	mean(HY)	mean(HY(C))	n· var(HY(J))	n· var(HY(C))	mean(HY)	n· var(HY(J))	mean(HY)	n· var(HY(J))
0	1.000	0.000	7.91 (8.00)	3.87 (4.00)	1.002	15.95 (16.00)	0.001	9.03
0.1	1.101	0.100	8.49 (8.60)	4.03 (4.03)	1.100	16.02 (16.60)	0.099	8.72
0.2	1.200	0.200	8.99 (9.20)	4.05 (4.12)	1.199	17.45 (17.20)	0.200	8.47
0.3	1.300	0.300	9.56 (9.80)	4.40 (4.27)	1.299	17.83 (17.80)	0.297	8.00
0.4	1.401	0.401	10.25 (10.40)	4.50 (4.48)	1.401	18.43 (18.40)	0.399	8.35
0.5	1.500	0.499	11.01 (11.00)	4.52 (4.75)	1.500	19.11 (19.00)	0.501	8.13
0.6	1.599	0.599	11.83 (11.60)	4.78 (5.08)	1.599	20.21 (19.60)	0.600	8.54
0.7	1.701	0.700	12.18 (12.20)	5.41 (5.47)	1.701	20.26 (20.20)	0.700	9.08
0.8	1.801	0.801	12.74 (12.80)	6.07 (5.92)	1.799	20.35 (20.80)	0.803	8.44
0.9	1.900	0.900	13.14 (13.40)	6.49 (6.43)	1.899	21.08 (21.40)	0.901	8.21
1.0	2.001	1.001	13.36 (14.00)	6.98 (7.00)	1.999	21.90 (22.00)	0.999	8.57

Table 1: Simulation results for different parameters. $\mathbb{E}[m_-^{(1)}(1)] = \mathbb{E}[m_-^{(2)}(1)] = 30000$; HY(C) refers to the estimator calculated from the simulated continuous part only; var(HY(J)) is the difference of the total Monte Carlo variance and var(HY(C)); in parentheses theoretical expectations, in [Sc3] this is constantly 8.

Appendix

A Preliminaries and some notation

Throughout the proof K and K_q denote generic constants, the latter dependent on q. On the compact time span $[0,1]$, we can reinforce the structural Assumption 2.1 replacing local boundedness by uniform boundedness which precisely means that b and σ as well as the jumps of X may be assumed to be bounded by K. Such a standard procedure is provided in Section 3.6.3. in Jacod (2012), among others.

The following notation is analogously introduced for the one-dimensional and two-dimensional setting. For any integer q we define the auxiliary (drift) process

$$b_s^q = b_s - \int (\kappa(\delta(s,z)) - \delta(s,z)1_{\{\gamma(z)\leq 1/q\}})\lambda(dz),$$

which satisfies $\|b_s^q\| \leq K_q$. We consider discretized versions $b(q,r)_s$ and $\sigma(r)_s$ of b^q and σ for some integer r. We set

$$b(q,r)_s = b^q_{(k-1)/2^r} 1_{[(k-1)/2^r, k/2^r)},\ k = 1,\dots,2^r,$$

locally constant on intervals $[(k-1)/2^r, k/2^r)$ and analogously for $\sigma(r)$. For any fixed pair (q,r), we define

$$\begin{aligned}
&B(q,r)_t = \int_0^t b(q,r)_s\,ds, \quad B'(q,r)_t = \int_0^t (b_s^q - b(q,r)_s)\,ds,\\
&C(r)_t = \int_0^t \sigma(r)_s\,dW_s, \quad C'(r)_t = \int_0^t (\sigma_s - \sigma(r)_s)\,dW_s,\\
&N(q)_t = \int_0^t\int \delta(s,z)1_{\{\gamma(z)>1/q\}}\mu(ds,dz), \quad M(q)_t = \int_0^t\int \delta(s,z)1_{\{\gamma(z)\leq 1/q\}}(\mu-\nu)(ds,dz).
\end{aligned}$$

Up to X_0, which does not matter in terms of increments anyway, X_t can be written as the sum of the six quantities above. Therefore each increment $\Delta_i^n X$ is the sum of six respective increments as well, and using the multinomial formula we see that RV_t^n (or the HY-estimator) becomes a sum of 21 addends. For the quadratic variation (for dimension one), we have

$$[X,X]_t = \int_0^t \sigma(r)_s^2\,ds + \int_0^t (\sigma_s - \sigma(r)_s)^2\,ds + 2\int_0^t \sigma(r)_s(\sigma_s - \sigma(r)_s)ds + [N(q),N(q)]_t + [M(q),M(q)]_t$$

and analogously for $[X^{(1)},X^{(2)}]$ in the two-dimensional setup. Next, we state some standard estimates for the terms from the above decomposition of X which will be used frequently in the analysis below. See e.g. Section 4.1 in Jacod (2008):

$$\forall p \geq 1, s,t \geq 0:\ \mathbb{E}\left[\|C_{s+t}(r) - C_s(r)\|^p \middle| \mathcal{F}_s\right] \leq K_p t^{p/2}, \tag{A.1a}$$

$$\begin{aligned}
\forall s,t \geq 0:\ \mathbb{E}\left[\|M(q)_{s+t} - M(q)_s\|^2 \middle| \mathcal{F}_s\right] &\leq K\,\mathbb{E}\left[\left(\int_s^{(s+t)}\int_{\{z\in\mathbb{R}^d|\gamma(z)<q^{-1}\}} \gamma^2(z)\nu(ds,dz)\right)\right]\\
&\leq Kte_q,
\end{aligned} \tag{A.1b}$$

$$\begin{aligned}
\forall s,t \in [(k-1)2^{-r}, k2^{-r}) \text{ for some } k:\ \mathbb{E}\left[\|C'_{s+t}(r) - C'_s(r)\|^2 \middle| \mathcal{F}_s\right] &\leq K\,\mathbb{E}\left[\left(\int_s^{s+t} \|\sigma_\tau - \sigma_s\|^2\,d\tau\right)\middle|\mathcal{F}_s\right]\\
&\leq Kt \sup_{\tau\in[s,t+s]} \|\sigma_\tau - \sigma_s\|^2,
\end{aligned} \tag{A.1c}$$

$$\forall s,t\geq 0:\ \mathbb{E}\left[\|N(q)_{s+t}-N(q)_s\|\Big|\mathcal{F}_s\right]\leq q\,t\int\gamma^2(z)\lambda(dz)\leq K_q\,t, \tag{A.1d}$$

where $e_q=\int_{\{z\in\mathbb{R}^d|\gamma(z)<q^{-1}\}}\gamma^2(z)\lambda(dz)$. From Assumption 2.1 we may conclude that $\int\gamma^2(z)\lambda(dz)$ is bounded, so by Lebesgue's theorem we have

$$e_q\to 0\ \text{ as }\ q\to\infty. \tag{A.1e}$$

B Proof of Theorem 2.2

We decompose the left hand side of (2.2) with the terms introduced in Appendix A. The only addends responsible for the limiting variance are

$$n^{1/2}\Big[\Big(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^nC(r))^2-\int_0^t\sigma(r)_s^2ds\Big)+2\sum_{i=1}^{\lfloor nt\rfloor}\Delta_i^nC(r)\Delta_i^nN(q)\Big]. \tag{B.1}$$

For the other terms converging to the quadratic variation the approximation errors

$$\begin{aligned}&n^{1/2}\Big[\Big(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^nC'(r))^2-\int_0^t(\sigma_s-\sigma(r)_s)^2ds\Big)+\Big(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^nN(q))^2-[N(q),N(q)]_t\Big)\\&+2\Big(\sum_{i=1}^{\lfloor nt\rfloor}\Delta_i^nC(r)\Delta_i^nC'(r)-\int_0^t\sigma(r)_s(\sigma_s-\sigma(r)_s)ds\Big)+\Big(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^nM(q))^2-[M(q),M(q)]_t\Big)\Big],\end{aligned} \tag{B.2}$$

are proved to be asymptotically negligible. The other remainder terms will be shown to be small as well. We state an overview which terms are treated jointly at this point: The pure drift parts are

$$n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^nB(q,r)+\Delta_i^nB'(q,r))^2, \tag{B.3}$$

and we also treat

$$2n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^nB(q,r)+\Delta_i^nB'(q,r))(\Delta_i^nC(r)+\Delta_i^nN(q)) \tag{B.4}$$

together. The mixed martingale part is

$$2n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\Delta_i^nM(q)(\Delta_i^nC(r)+\Delta_i^nC'(r)). \tag{B.5}$$

The remainder terms involving $C'(r)$ are now

$$2n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\Delta_i^nC'(r)(\Delta_i^nB(q,r)+\Delta_i^nB'(q,r)+\Delta_i^nN(q)), \tag{B.6}$$

and finally the remainder terms involving $M(q)$ become

$$2n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\Delta_i^nM(q)(\Delta_i^nB(q,r)+\Delta_i^nB'(q,r)+\Delta_i^nN(q)). \tag{B.7}$$

Asymptotics will always work in the sense that we let $n\to\infty$ first, then the auxiliary $r\to\infty$ and finally $q\to\infty$. The proof of Theorem 2.2 is divided in two parts which are given in the following propositions:

Proposition B.1. *On the assumptions of Theorem 2.2:*

$$(B.1) \quad \stackrel{\mathcal{L}-(s)}{\longrightarrow} \quad V_t + Z_t\,. \tag{B.8}$$

Proposition B.2. *On the assumptions of Theorem 2.2:*

$$\lim_{q\to\infty} \limsup_{r\to\infty} \limsup_{n\to\infty} \mathbb{P}(|(B.2) + (B.3) + (B.4) + (B.5) + (B.6) + (B.7)| > \epsilon) = 0$$

for any $\epsilon > 0$.

In this section, we show the claim of the latter proposition only. We will prove an analogous claim for Theorem 3.5 later, which includes Proposition B.1 as a special case. To establish Proposition B.2 we show that it holds for each of the remainder terms (B.2) to (B.7) separately, which implies our claim. Note by Markov inequality that it is sufficient to obtain bounds for moments of the respective terms.

Proof of Proposition B.2.

- $\lim_{q\to\infty} \limsup_{r\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(|(\text{B.3})| > \epsilon\right) = 0$ for any $\epsilon > 0$:

Since we have $|\Delta_i^n B(q,r)| \leq K_q/n$ and also for $\Delta_i^n B'(q,r)$

$$n^{1/2} \sum_{i=1}^{\lfloor nt \rfloor} (\Delta_i^n B(q,r) + \Delta_i^n B'(q,r))^2 \xrightarrow{\mathbb{P}} 0 \text{ for any fixed } q \text{ and } r.$$

- $\lim_{q\to\infty} \limsup_{r\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(|(\text{B.4})| > \epsilon\right) = 0$ for any $\epsilon > 0$:

$$n^{1/2} \sum_{i=1}^{\lfloor nt \rfloor} \Delta_i^n B(q,r) \Delta_i^n C(r) = n^{-1/2} \sum_{k=1}^{\lfloor 2^r t \rfloor} b^q_{(k-1)2^{-r}} (C(r)_{k2^{-r}} - C(r)_{(k-1)2^{-r}}) + R(q,r)^n_t,$$

where the error due to increments over intervals with $(i-1)/n \leq (k-1)2^{-r} < i/n$ and boundary effects is denoted by $R(q,r)^n_t$ and satisfies $\mathbb{E}[|R(q,r)^n_t|] \leq K_r n^{-1/2}$, thus becomes small. The first term on the right hand side above is a sum of martingale differences. Therefore

$$\begin{aligned} n^{-1}\mathbb{E}\left(\sum_{k=1}^{\lfloor 2^r t \rfloor} b^q_{(k-1)2^{-r}} (C(r)_{k2^{-r}} - C(r)_{(k-1)2^{-r}}) \right)^2 &\leq K/n \sum_{k=1}^{\lfloor 2^r t \rfloor} \mathbb{E}\left(C(r)_{k2^{-r}} - C(r)_{(k-1)2^{-r}} \right)^2 \\ &\leq K/n \to 0. \end{aligned}$$

For the treatment of the analogous term involving $\Delta_i^n B'(q,r)$ let $w(f,h) = \sup\{|f(s) - f(t)|; |s-t| \leq h\}$ denote the modulus of continuity of a function f. Several applications of Cauchy-Schwarz inequality and Itô isometry give

$$n^{1/2} \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{E}|\Delta_i^n B'(q,r) \Delta_i^n C(r)| \leq n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{E}|w(b^q, 2^{-r}) \Delta_i^n C(r)| \leq K \mathbb{E}[|w(b^q, 2^{-r})|^2]^{1/2}.$$

By construction, b^q is bounded and continuous. Therefore, $w(b^q, 2^{-r})$ converges to zero for each ω as $r \to \infty$, and the entire term becomes small due to Lebesgue's theorem. The claim for (B.4) now follows using the bounds on the drift and (A.1d).

- $\lim_{q\to\infty}\limsup_{r\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(|(\text{B.5})|>\epsilon\right)=0$ for any $\epsilon>0$:

Integration by parts formula gives

$$\begin{aligned}\Delta_i^n C(r)\Delta_i^n M(q) &= \int_{(i-1)/n}^{i/n}(C(r)_{s-}-C(r)_{(i-1)/n})\,dM(q)_s\\ &\quad+\int_{(i-1)/n}^{i/n}(M(q)_s-M(q)_{(i-1)/n})\,dC(r)_s,\end{aligned}\tag{B.9}$$

with expectation zero. Applying Burkholder-Davis-Gundy inequality yields

$$\begin{aligned}&\mathbb{E}\Big(\int_{(i-1)/n}^{i/n}(C(r)_{s-}-C(r)_{(i-1)/n})\,dM(q)_s\Big)^2\le K\mathbb{E}\int_{(i-1)/n}^{i/n}(C(r)_{s-}-C(r)_{(i-1)/n})^2\,d[M(q),M(q)]_s\\ \le& K\mathbb{E}\int_{(i-1)/n}^{i/n}\int(C(r)_{s-}-C(r)_{(i-1)/n})^2\delta^2(s,z)1_{\{\gamma(z)\le 1/q\}}\lambda(dz)\,ds\\ \le& Ke_q\Big(\int_{(i-1)/n}^{i/n}\mathbb{E}(C(r)_{s-}-C(r)_{(i-1)/n})^2ds\Big)\le Ke_q/n^2\,.\end{aligned}\tag{B.10}$$

Similarly, we obtain

$$\mathbb{E}\Big(\int_{(i-1)/n}^{i/n}(M(q)_s-M(q)_{(i-1)/n})\,dC(r)_s\Big)^2\le Ke_q/n^2.$$

and altogether

$$n\mathbb{E}\Big(\sum_{i=1}^{\lfloor nt\rfloor}\Delta_i^nM(q)\Delta_i^nC(r)\Big)^2\le n\sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}\Big(\Delta_i^nM(q)\Delta_i^nC(r)\Big)^2\le Ke_q,$$

and the same bound holds for the term involving $\Delta_i^nC'(r)$. Our claim now follows by virtue of (A.1e).

- $\lim_{q\to\infty}\limsup_{r\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(|(\text{B.6})|>\epsilon\right)=0$ for any $\epsilon>0$:

$$\begin{aligned}&\Delta_i^nB(q,r)+\Delta_i^nB'(q,r)+\Delta_i^nN(q)=\int_{(i-1)/n}^{i/n}b_s^q ds+\int_{(i-1)/n}^{i/n}\int\delta(s,z)1_{\{\gamma(z)>1/q\}}\mu(ds,dz)\\ =&\underbrace{\int_{(i-1)/n}^{i/n}\Big(b_s+\int\kappa'(\delta(s,z))\lambda(dz)\Big)ds}_{=\Delta_i^n\widetilde{B}}+\underbrace{\int_{(i-1)/n}^{i/n}\int\delta(s,z)1_{\{\gamma(z)>1/q\}}(\mu-\nu)(ds,dz)}_{=\Delta_i^n\widetilde{M}(q)}.\end{aligned}$$

By two applications of Cauchy-Schwarz inequality and Itô isometry, we conclude

$$n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\Big|\Delta_i^nC'(r)\Delta_i^n\widetilde{B}\Big|\Big]\le Kn^{-1/2}\sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}[|\Delta_i^nC'(r)|^2]^{1/2}\le K\mathbb{E}[|w(\sigma^2,2^{-r})|^2]^{1/2}.$$

Convergence to zero as $r\to\infty$ can be deduced from Lebesgue's theorem again. On the other hand, we have a similar decomposition for $\Delta_i^nC'(r)\Delta_i^n\widetilde{M}(q)$ as in (B.9). The arguments from (B.10) yield

$$n\mathbb{E}\Big(\sum_{i=1}^{\lfloor nt\rfloor}\int_{(i-1)/n}^{i/n}(C'(r)_{s-}-C'(r)_{(i-1)/n})\,d\widetilde{M}(q)_s\Big)^2\le K\mathbb{E}[|w(\sigma^2,2^{-r})|^2],$$

so it remains to focus on

$$n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\int_{(i-1)/n}^{i/n}(\widetilde{M}(q)_{s-}-\widetilde{M}(q)_{(i-1)/n})(\sigma_s-\sigma(r)_s)\,dW_s.$$

We have that $\Delta_i^n\widetilde{M}(q)=\Delta_i^nN(q)-\int_{(i-1)/n}^{i/n}\int\delta(s,z)1_{\{\gamma(z)>q^{-1}\}}\lambda(dz)ds$, where the absolute value of the second addend is bounded by K_q/n. Since $\sigma_s-\sigma(r)_s\to 0$ pointwise for $r\to\infty$, all we need to discuss is

$$n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\int_{(i-1)/n}^{i/n}(N(q)_{s-}-N(q)_{(i-1)/n})(\sigma_s-\sigma(r)_s)\,dW_s.$$

Let $S_1,S_2,\dots$ be a sequence of stopping times exhausting the jumps of $N(q)$. Then

$$\begin{aligned}&n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}\Big|\int_{(i-1)/n}^{i/n}(N(q)_{s-}-N(q)_{(i-1)/n})(\sigma_s-\sigma(r)_s)\,dW_s\Big| \qquad\text{(B.11)}\\ \le\,&n^{1/2}\sum_{p\ge1}\sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}\left|1_{\{(i-1)/n<S_p\le i/n\}}|N(q)_{S_p}-N(q)_{S_p-}|\int_{S_p}^{\lceil nS_p\rceil/n}(\sigma_s-\sigma(r)_s)\,dW_s\right|.\end{aligned}$$

Using the strong Markov property of Brownian motion we obtain

$$\begin{aligned}\mathbb{E}\Big[\Big|\int_{S_p}^{\lceil nS_p\rceil/n}(\sigma_s-\sigma(r)_s)\,dW_s\Big|\Big|\mathcal{F}_{S_p}\Big]\le&\Big(\int_{S_p}^{\lceil nS_p\rceil/n}\mathbb{E}[(\sigma_s-\sigma(r)_s)^2|\mathcal{F}_{S_p}]ds\Big)^{1/2}\\ \le&Kn^{-1/2}\mathbb{E}[w(\sigma,2^{-r})^2|\mathcal{F}_{S_p}]^{1/2}.\end{aligned}$$

Therefore, using Cauchy-Schwarz inequality several times, (B.11) can be bounded by a constant times

$$\begin{aligned}&\mathbb{E}\sum_{p\ge1}\sum_{i=1}^{\lfloor nt\rfloor}\Big(1_{\{(i-1)/n<S_p\le i/n\}}1_{\{S_p\le1\}}\mathbb{E}[w(\sigma,2^{-r})^2|\mathcal{F}_{S_p}]^{1/2}\Big)\\ \le&\,\mathbb{E}\Big[\sum_{p\ge1}\Big(\sum_{i=1}^{\lfloor nt\rfloor}1_{\{(i-1)/n<S_p\le i/n\}}\Big)^2\Big]^{1/2}\mathbb{E}\Big[\sum_{p\ge1}1_{\{S_p\le1\}}\mathbb{E}[w(\sigma,2^{-r})^2|\mathcal{F}_{S_p}]\Big]^{1/2}. \qquad\text{(B.12)}\end{aligned}$$

The first factor is bounded by $\mathbb{E}[\Pi(q)]^{1/2}$, where $\Pi(q)=\int_0^1\int1_{\{\gamma(z)>1/q\}}\mu(ds,dz)$ denotes the number of large jumps on $[0,1]$. It holds that

$$\begin{aligned}\mathbb{E}[(\Pi(q))^2]\le&K\mathbb{E}\Big[\Big(\int_0^1\int1_{\{\gamma(z)>1/q\}}(\mu-\nu)(ds,dz)\Big)^2+\Big(\int_0^1\int1_{\{\gamma(z)>1/q\}}\lambda(dz)ds\Big)^2\Big]\\ \le&K\Big(\int_0^1\int1_{\{\gamma(z)>1/q\}}\lambda(dz)ds+\Big(\int_0^1\int1_{\{\gamma(z)>1/q\}}\lambda(dz)ds\Big)^2\Big)\le K_q\end{aligned}$$

by the integrability assumption on γ^2. It remains to focus on the second factor, for which

$$\begin{aligned}\sum_{p\ge1}\mathbb{E}[1_{\{S_p\le1\}}\mathbb{E}[w(\sigma,2^{-r})^2|\mathcal{F}_{S_p}]^{1/2}]\le&\sum_{p\ge1}\mathbb{E}[1_{\{S_p\le1\}}]^{3/4}\mathbb{E}(\mathbb{E}[w(\sigma,2^{-r})^2|\mathcal{F}_{S_p}]^2)^{1/4}\\ \le&\mathbb{E}[w(\sigma,2^{-r})^4]^{1/4}\sum_{p\ge1}\mathbb{P}(S_p\le1)^{3/4}.\end{aligned}$$

using Hölder and Jensen inequality, respectively. Finally,

$$\sum_{p\geq 1}\mathbb{P}(S_p\leq 1)^{3/4}=\sum_{p\geq 1}\mathbb{P}(\Pi(q)>p)^{3/4}\leq(\mathbb{E}[(\Pi(q))^2])^{3/4}\sum_{p\geq 1}p^{-3/2}\leq K_q.$$

Continuity of σ gives the claim again.

- $\lim_{q\to\infty}\limsup_{r\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(|(\text{B.7})|>\epsilon\right)=0$ for any $\epsilon>0$:

With the previous notation we have

$$n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\Big|\Delta_i^n M(q)\Delta_i^n\widetilde{B}\Big|\Big]\leq Kn^{-1/2}\sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}[(\Delta_i^n M(q))^2]^{1/2}\leq Ke_q^{1/2}\to 0$$

as $q\to\infty$. The product $M(q)\widetilde{M}(q)$ is a martingale (no common jumps) and we may proceed as in the proof for (B.5). In particular

$$\begin{aligned}&\mathbb{E}\Big(\int_{(i-1)/n}^{i/n}(M(q)_{s-}-M(q)_{(i-1)/n})\,d\widetilde{M}(q)_s\Big)^2\leq K\int_{(i-1)/n}^{i/n}\mathbb{E}(M(q)_{s-}-M(q)_{(i-1)/n})^2\Big(\int\gamma^2(z)\lambda(dz)\Big)\,ds\\ \leq&K\int_{(i-1)/n}^{i/n}\int_{(i-1)/n}^{s}\Big(\int\gamma^2(z)1_{\{\gamma(z)\leq 1/q\}}\lambda(dz)\Big)\Big(\int\gamma^2(z)\lambda(dz)\Big)duds\leq e_qK/n^2.\end{aligned}\tag{B.13}$$

Changing the roles of $M(q)$ and $\widetilde{M}(q)$ then gives the result.

- $\lim_{q\to\infty}\limsup_{r\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(|(\text{B.2})|>\epsilon\right)=0$ for any $\epsilon>0$:

We may replace $[X,X]_t$ with $[X,X]_{\lfloor nt\rfloor/n}$, since

$$n^{1/2}\mathbb{E}([X,X]_t-[X,X]_{\lfloor nt\rfloor/n})\leq Kn^{-1/2}\,.$$

Now, let $\Omega(q,r,n,t)$ denote the set on which each interval $[i/n,(i+1)/n]$ contains either one or no jump and no jump occurs on those intervals with $(i-1)/n\leq(k-1)2^{-r}<k2^{-r}\leq i/n$, $k=1,\dots,2^r$, and also not on $[\lfloor nt\rfloor/n,t]$. This is helpful, since

$$n^{1/2}\Big(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^n N(q))^2-[N(q),N(q)]_t\Big)=0$$

identically on $\Omega(q,r,n,t)$. For fixed q,r,t, we have $\Omega(q,r,n,t)\to\Omega$ as $n\to\infty$ which is why we may assume ω to live on this set. (Note that we do not need the condition involving r at this point, but the previously introduced set will be used several times later again.) Proving that

$$n^{1/2}\Big(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^n M(q))^2-[M(q),M(q)]_t\Big)=2n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\int_{(i-1)/n}^{i/n}(M(q)_{s-}-M(q)_{(i-1)/n})\,dM(q)_s$$

becomes small works similarly to (B.13). For the remaining terms in (B.2) let us exemplarily discuss

$$n^{1/2}\Big(\sum_{i=1}^{\lfloor nt\rfloor}(\Delta_i^n C'(r))^2-\int_0^t(\sigma_s-\sigma(r)_s)^2ds\Big)=2n^{1/2}\sum_{i=1}^{\lfloor nt\rfloor}\int_{(i-1)/n}^{i/n}(C'(r)_s-C'(r)_{(i-1)/n})\,dC'(r)_s$$

which is again a sum of martingale increments. Several applications of Burkholder-Davis-Gundy inequality and Cauchy-Schwarz inequality prove that the expectation of its square is bounded by $\mathbb{E}[w(\sigma,2^{-r})^4]^{1/2}$ which converges to zero as $r\to\infty$. A similar argument applies to the cross term involving $C(r)$ and $C'(r)$. This completes the proof of Proposition B.2. □

C Proof of Theorem 3.5

We have the same decomposition for $n^{1/2}(RV_t^n - [X,X]_t)$ as in Section B, where the only difference is that sums run to $m_-^n(t)$ rather than $\lfloor nt \rfloor$. Therefore, we have to prove similar claims as Proposition B.1 and Proposition B.2, and the latter works in pretty much the same way as before, up to using Assumption 3.1 in some places: To be precise, all claims follow by simply using (3.1) where necessary, apart from the first term in (B.12) which becomes the square root of

$$\mathbb{E}\Big[\sum_{p\geq 1}\Big(\sum_{i=1}^{m_-^n(t)} 1_{\{t_{i-1}^n \leq S_p \leq t_i^n\}} n^{1/2}(t_i^n - t_{i-1}^n)^{1/2}\Big)^2\Big] \leq \mathbb{E}\Big[\sum_{p=1}^{\Pi(q)} n(\tau_+^n(S_p) - \tau_-^n(S_p))\Big]$$
$$= \mathbb{E}\Big(\sum_{p=1}^{\Pi(q)} \mathbb{E}[n(\tau_+^n(S_p) - \tau_-^n(S_p))]\Big)$$

in this context, where we have used the Wald identity to obtain the latter equality. If we order the stopping times by the size of the corresponding jumps, i.e. S_1 denotes the time of the largest jump of $N(q)$, S_2 is the time of the second largest jump, and so on, then each S_p is uniformly distributed. Therefore $\mathbb{E}[n(\tau_+^n(S_p) - \tau_-^n(S_p))] \leq K$ using Assumption 3.1 (ii), and we obtain the same bound as in (B.12).

For this reason, we focus in this proof on

Proposition C.1. *On the assumptions of Theorem 3.5, we have the $\mathcal{F}^0$-stable convergence*

$$n^{1/2}\Big[\Big(\sum_{i=1}^{m_-^n(t)} (\Delta_i^n C(r))^2 - \int_0^t \sigma(r)_s^2 ds\Big) + 2\sum_{i=1}^{m_-^n(t)} \Delta_i^n C(r)\Delta_i^n N(q)\Big] \xrightarrow{\mathcal{L}-(s)} \widetilde{V}_t + \widetilde{Z}_t\,. \tag{C.1}$$

Recall that the missing proof of Proposition B.1 is included as a special case in the previous claim.

Proof of Proposition C.1. We begin with a proof of the $\mathcal{F}^0$-stable central limit theorem for realized variance in the continuous case. For fixed r, we prove

$$V(r)_t^n := n^{1/2}\Big(\sum_{i=1}^{m_-^n(t)} (\Delta_i^n C(r))^2 - \int_0^t \sigma(r)_s^2 ds\Big) \xrightarrow{\mathcal{L}-(s)} \sqrt{2}\int_0^t \sigma(r)_s^2 (G'(s))^{1/2} dW_s' =: V(r)_t \tag{C.2}$$

first. Typically, one rewrites

$$V(r)_t^n = \sum_{i=1}^{m_-^n(t)} \xi_i^n + o_p(1), \qquad \xi_i^n = n^{1/2}\Big(|\Delta_i^n C(r)|^2 - \int_{t_{i-1}^n}^{t_i^n} \sigma(r)_s^2\, ds\Big),$$

using $\mathbb{E}[\pi_n^q] = o(n^{-\alpha})$ from Assumption 3.1, and then exploits Theorem IX 7.28 in Jacod and Shiryaev (2003) for which several intermediate steps regarding the behaviour of conditional expectations of functionals of ξ_i^n with respect to $\mathcal{F}_{t_{i-1}^n}^0$ have to be shown. However, as discussed in the proof of Proposition 5.1 in Hayashi et al. (2011), one can equally well discuss conditional expectations with respect to $\mathcal{G}_{i-1}^n = \mathcal{F}_{t_{i-1}^n}^0 \vee \mathcal{F}^1$. This σ-algebra represents knowledge of the entire sampling scheme plus knowledge of

X up to the stopping time t^n_{i-1}. The proof of (C.2) then boils down to prove

$$\sum_{i=1}^{m^n_-(t)} \mathbb{E}[\xi^n_i|\mathcal{G}^n_{i-1}] \xrightarrow{\mathbb{P}} 0, \tag{C.3}$$

$$\sum_{i=1}^{m^n_-(t)} \mathbb{E}[(\xi^n_i)^2|\mathcal{G}^n_{i-1}] \xrightarrow{\mathbb{P}} 2\int_0^t \sigma(r)^4_s G'(s)ds, \tag{C.4}$$

$$\sum_{i=1}^{m^n_-(t)} \mathbb{E}[\xi^n_i \Delta^n_i M|\mathcal{G}^n_{i-1}] \xrightarrow{\mathbb{P}} 0 \text{ for all } M \in \mathcal{M}, \tag{C.5}$$

$$\sum_{i=1}^{m^n_-(t)} \mathbb{E}[(\xi^n_i)^4|\mathcal{G}^n_{i-1}] \xrightarrow{\mathbb{P}} 0, \tag{C.6}$$

where $\mathcal{M}$ is the set of all martingales that are either bounded and orthogonal to W or equal to W. Note that (C.3) is satisfied identically by construction, whereas (C.6) holds using (A.1a) and the assumption on π_n. The orthogonality condition (C.5) follows from standard arguments as e.g. laid out in Example 2 of Podolskij and Vetter (2010). What remains to show is thus (C.4). First, let i be such that $(k-1)2^{-r} < t^n_{i-1} < t^n_i < k2^{-r}$ for some k. Then

$$\mathbb{E}[(\xi^n_i)^2|\mathcal{G}^n_{i-1}] = 2n\sigma(r)^4_{(k-1)2^{-r}}(t^n_i - t^n_{i-1})^2.$$

Since all other choices of i correspond to at most 2^r summands and the length of the intervals between successive observations is bounded by π_n, we obtain

$$\sum_{i=1}^{m^n_-(t)} \mathbb{E}[(\xi^n_i)^2|\mathcal{G}^n_{i-1}] = \sum_{k=1}^{2^r} 2\sigma(r)^4_{(k-1)2^{-r}} \sum_{i=m^n_+((k-1)2^{-r})}^{m^n_-(k2^{-r})\wedge m^n_-(t)} n(t^n_i - t^n_{i-1})^2 + o_p(1),$$

where empty sums are set to be zero. Therefore, since $t^n_{m^n_-(s)} \to s$ for any s, the term becomes

$$\begin{aligned}&\sum_{k=1}^{\lfloor 2^r t\rfloor} 2\sigma(r)^4_{(k-1)2^{-r}}(G(k2^{-r}) - G((k-1)2^{-r})) + 2\sigma(r)^4_{\lfloor 2^r t\rfloor 2^{-r}}(G(t) - G(\lfloor 2^r t\rfloor 2^{-r}) + o_p(1)\\ =&2\int_0^t \sigma(r)^4_s \, dG(s) + o_p(1)\end{aligned}$$

using (3.1), which finishes the proof of (C.2) by continuous differentiability of G.

On the other hand, we need a joint result regarding $V(r)^n_t$ and the term involving the jumps. Thus, for a fixed q, let $S_1, S_2, \dots$ be a sequence of stopping times exhausting the jumps of $N(q)$. We set $\alpha(n,p) = n^{1/2}(W_{\tau^n_+(S_p)} - W_{\tau^n_-(S_p)})$, and what we discuss in a first step is the $\mathcal{F}^0$-stable convergence

$$(V(r)^n_t, \alpha(n,p)_{p\geq 1}) \xrightarrow{\mathcal{L}-(s)} (V(r)_t, (\eta(S_p)^{1/2}U'_p)_{p\geq 1}), \tag{C.7}$$

as $n \to \infty$.

The proof is close to the one of Lemma 5.8 in Jacod (2008) for most parts, which is why we do not give all details but focus on the intuition behind the steps. Let us begin with several remarks: First, we

order the stopping times by the size of the corresponding jumps again. Second, it is enough to prove the result for a fixed k, that is we show

$$(V(r)^n_t, \alpha(n,p)_{1\le p\le k}) \overset{\mathcal{L}-(s)}{\longrightarrow} (V(r)_t, (\eta(S_p)^{1/2}U'_p)_{1\le p\le k}) \tag{C.8}$$

only. Formally this means that we have to prove

$$\mathbb{E}\Big[\Psi g(V(r)^n_t)\prod_{p=1}^k h_p(\alpha(n,p))\Big] \to \widetilde{\mathbb{E}}\Big[\Psi g(V(r)_t)\prod_{p=1}^k h_p(\eta(S_p)^{1/2}U'_p)]$$

for any $\mathcal{F}^0$-measurable Ψ and for all bounded continuous functions $g, h_1, \dots, h_k$. However, it is sufficient to focus on $\mathcal{G}$-measurable variables, which is the σ-algebra generated by the measure μ and the processes b, σ, W and X, as one may otherwise replace Ψ with $\mathbb{E}[\Psi|\mathcal{G}]$ and use measurability with respect to $\mathcal{G}$ of all other variables.

The first step to obtain the previous relation is to replace $V(r)^n_t$ by some $V(r,\ell)^n_t$ which is defined over those intervals only that do not intersect with those on which the $\alpha(n,p)$ are defined. This approach secures conditional independence of the limiting Brownian motion W' and the normally distributed U'_p later. Precisely, let $S^{\ell-}_p = (S_p - 1/\ell)^+$ and $S^{\ell+}_p = S_p + 1/\ell$ denote intervals around the jump times and set $B_\ell = \cup_{p\ge1}^k [S^{\ell-}_p, S^{\ell+}_p]$. Let $\Lambda_n(\ell,t)$ denote the set of indices i such that $i \le m^n_-(t)$ and $[t^n_{i-1}, t^n_i] \cap B_\ell = \emptyset$. It is rather simple to see that both $V(r)^n_t$ and $V(r,\ell)^n_t = \sum_{i\in\Lambda_n(\ell,t)} \xi^n_i$ and $V(r)_n$ and $V(r,\ell)_n = \sqrt{2}\int_0^t 1_{B^c_\ell}(s)\sigma(r)^2_s(G'(s))^{1/2}dW'_s$ are in a suitable sense close for large ℓ, which means that it suffices to prove

$$\mathbb{E}\Big[\Psi g(V(r,\ell)^n_t)\prod_{p=1}^k h_p(\alpha(n,p))\Big] \to \widetilde{\mathbb{E}}\Big[\Psi g(V(r,\ell)_t)\prod_{p=1}^k h_p(\eta(S_p)^{1/2}U'_p)\Big]$$

for each fixed ℓ. Fix one such ℓ in the following.

Introduce the filtration $\mathcal{F}^{0,\ell}$ which is the smallest one containing $\mathcal{F}^0$ and is defined in such a way that $W(\ell)_t = \int_0^t 1_{B_\ell}(s)dW_s$ (and the times $S_1, \dots, S_k$) is $\mathcal{F}^{0,\ell}_0$-measurable. We will now work conditionally on $\mathcal{F}^{0,\ell}_0$, so let Q_ω denote a regular version of this conditional probability. Since $\Delta^n_i W$ is independent of $\mathcal{F}^{0,\ell}_0$ (for n large enough) and remains normally distributed for $i \in \Lambda_n(\ell,t)$, which contains all but a finite number of indices, reproducing the proof of (C.2) yields

$$\mathbb{E}_{Q_\omega}[\Psi g(V(r,\ell)^n_t)] \to \widetilde{\mathbb{E}}_{\tilde{Q}_\omega}[\Psi g(V(r,\ell)_t)].$$

Thus, using $\mathcal{F}^{0,\ell}_0$-measurablility of $\prod_{p=1}^k h_p(\alpha(n,p))$, we obtain

$$\begin{aligned}\mathbb{E}\big[\Psi g(V(r,\ell)^n_t)\prod_{p=1}^k h_p(\alpha(n,p))\big] &= \mathbb{E}\big[\mathbb{E}_{Q_\omega}[\Psi g(V(r,\ell)^n_t)]\prod_{p=1}^k h_p(\alpha(n,p))\big] \\ &\sim \mathbb{E}\big[\mathbb{E}_{\tilde{Q}_\omega}[\Psi g(V(r,\ell)_t)]\prod_{p=1}^k h_p(\alpha(n,p))\big].\end{aligned}$$

Since $\mathbb{E}_{\tilde{Q}_\omega}[\Psi g(V(r,\ell)_t)]$ is $\mathcal{F}^{0,\ell}_0$-measurable again, everything then boils down to prove

$$\mathbb{E}\big[\Gamma\prod_{p=1}^k h_p(\alpha(n,p))\big] \to \widetilde{\mathbb{E}}\big[\Gamma\prod_{p=1}^k h_p(\eta(S_p)^{1/2}U'_p)\big].$$

for any $\mathcal{F}_0^{0,\ell}$-measurable Γ. By conditioning, we may again restrict us to those variables generated by the jump times of $N(q)$ and the Brownian motion $W(\ell)$, and using Lemma 2.1 in Jacod and Protter (1998) this means to prove

$$\mathbb{E}\big[\gamma(W(\ell))\kappa(S_1,\dots,S_k)\prod_{p=1}^{k}h_p(\alpha(n,p))\big]\to\widetilde{\mathbb{E}}\big[\gamma(W(\ell))\kappa(S_1,\dots,S_k)\prod_{p=1}^{k}h_p(\eta(S_p)^{1/2}U_p')\big]$$

for any bounded continuous functions γ and κ. Since $W(\ell)$ is independent of any other quantity involved, this means finally that we have to prove

$$\mathbb{E}\Big[\kappa(S_1,\dots,S_k)\prod_{p=1}^{k}h_p(\alpha(n,p))\Big]\to\mathbb{E}\big[\kappa(S_1,\dots,S_k)\prod_{p=1}^{k}h_p(\eta(S_p)^{1/2}U_p')\big].$$

On the set $\Omega(q,r,n,t)$, introduced in the final step of the proof of Proposition B.2, the variables $\alpha(n,p)$ are defined over non-overlapping intervals which explains independence of the limiting variables. Since $\Omega(q,r,n,t)\to\Omega$ for $n\to\infty$, we may add (and subtract again) an indicator function over $\Omega(q,r,n,t)$ on the left hand side above. Then,

$$\mathbb{E}\big[\kappa(S_1,\dots,S_k)\prod_{p=1}^{k}h_p(\alpha(n,p))\big]\sim\widetilde{\mathbb{E}}\big[\kappa(S_1,\dots,S_k)\prod_{p=1}^{k}h_p((n(\tau_+^n(S_p)-\tau_-^n(S_p)))^{1/2}U_p')\big].$$

We are therefore left to prove the stable convergence of $(n(\tau_+^n(S_p)-\tau_-^n(S_p)))_{1\le p\le k}$ to $(\eta(S_p))_{1\le p\le k}$ which is exactly condition (3.3) using the joint uniform distribution of the jump times of $N(q)$.

Finally, on $\Omega(q,r,n,t)$ all of the jumps occur on intervals of piecewise constancy of $\sigma(r)$. Thus we obtain easily

$$V(r)_t^n+2n^{1/2}\sum_{i=1}^{m_-^n(t)}\Delta_i^nC\Delta_i^nN(q)\ \overset{\mathcal{L}-(s)}{\longrightarrow}\ V(r)_t+2\sum_p\Delta N(q)_{S_p}\sigma(r)_{S_p}\eta(S_p)^{1/2}U_p'.$$

The proof can be finished by first letting $r\to\infty$ and then $q\to\infty$: Since σ is continuous, we have both $\widetilde{\mathbb{E}}[|V(r)_t-V_t|^2]\to0$ and

$$\sum_p\widetilde{\mathbb{E}}[|\Delta N(q)_{S_p}||\sigma(r)_{S_p}-\sigma_{S_p}||\eta(S_p)^{1/2}||U_p'|]\le K\sum_p\mathbb{E}[1_{\{S_p\le t\}}|\sigma(r)_{S_p}-\sigma_{S_p}|]\to0$$

by successive conditioning and boundedness of the first moments of $\eta(S_p)$ and $\Pi(q)$. Finally,

$$\widetilde{\mathbb{E}}\Big[\Big|Z_t-\sum_p\Delta N(q)_{S_p}\sigma_{S_p}\eta(S_p)^{1/2}U_p'\Big|^2\Big]\le K\mathbb{E}\Big[\sum_t|\Delta X_s|^2\mathbf{1}_{|\Delta X_s|\le1/q}\Big]\to0,$$

again from Lebesgue's theorem. □

D Proof of Theorem 4.4

We decompose $X^{(1)}$ and $X^{(2)}$ in the same way as above and denote the single addends likewise. The first part of the proof establishes the limit theorem for the leading variance term:

Proposition D.1. *On the assumptions of Theorem 4.4, the term*

$$\sqrt{n}\Big(\sum_{t_i^{(1)}\le t}\sum_{t_j^{(2)}\le t} \Big(\Delta_i^{n_1}C^{(1)}(r)\Delta_j^{n_2}C^{(2)}(r) + \Delta_j^{n_2}N^{(2)}(q)\Delta_i^{n_1}C^{(1)}(r) + \Delta_i^{n_1}N^{(1)}(q)\Delta_j^{n_2}C^{(2)}(r)\Big) \\ \times 1_{\{\min(t_i^{(1)},t_j^{(2)})>\max(t_{i-1}^{(1)},t_{j-1}^{(2)})\}} - \int_0^t (\rho\sigma^{(1)}\sigma^{(2)})(r)_s\,ds\Big) \tag{D.1}$$

satisfies:

$$\text{(D.1)} \overset{\mathcal{L}-(s)}{\longrightarrow} \tilde{\mathcal{V}}_t + \tilde{\mathcal{Z}}_t\,. \tag{D.2}$$

Proof. The claim draws on the stable central limit theorem for the continuous semimartingale as one building block. For the local parametric approximation with fixed r

$$\tilde{\mathcal{V}}(r)_t^n := \sqrt{n}\Big(\sum_{t_i^{(1)}\le t}\sum_{t_j^{(2)}\le t} (\Delta_i^{n_1}C^{(1)}(r)\Delta_j^{n_2}C^{(2)}(r))1_{\{\min(t_i^{(1)},t_j^{(2)})>\max(t_{i-1}^{(1)},t_{j-1}^{(2)})\}} - \int_0^t (\rho\sigma^{(1)}\sigma^{(2)})(r)_s\,ds\Big),$$

we have that $\tilde{\mathcal{V}}(r)_t^n \overset{\mathcal{L}-(s)}{\longrightarrow} \tilde{\mathcal{V}}(r)_t$; see Hayashi and Yoshida (2011) and Bibinger (2011). Thus, the proof affiliates to the proof of Theorem 3.5 based on an extension of Lemma 5.8 in Jacod (2008). The only major difference regards the terms involving the (co-)jumps, as more quantities are of interest now. Set

$$\alpha(n,p) = n^{1/2}\Big(\Big(W^{(1)}_{\tau^{(1,2)}_{++}(S_p)} - W^{(1)}_{\max(\tau^1_+(S_p),\tau^2_+(S_p))}\Big), \Big(W^{(1)}_{\min(\tau^{(1)}_-(S_p),\tau^2_-(S_p))} - W^{(1)}_{\tau^{(1,2)}_{--}(S_p)}\Big), \\ \Big(W^{(2)}_{\tau^{(2,1)}_{++}(S_p)} - W^{(2)}_{\max(\tau^1_+(S_p),\tau^2_+(S_p))}\Big), \Big(W^{(2)}_{\min(\tau^{(1)}_-(S_p),\tau^2_-(S_p))} - W^{(2)}_{\tau^{(2,1)}_{--}(S_p)}\Big), \\ \Big(W^{(1)}_{\max(\tau^{(1)}_+(S_p),\tau^2_+(S_p))} - W^{(1)}_{\min(\tau^{(1)}_-(S_p),\tau^2_-(S_p))}\Big), \Big(W^{(2)}_{\max(\tau^{(1)}_+(S_p),\tau^2_+(S_p))} - W^{(2)}_{\min(\tau^{(1)}_-(S_p),\tau^2_-(S_p))}\Big)\Big),$$

motivated in (4.3a)–(4.3e) above. The convergence of $\alpha(n,p)$ to mixtures of independent normal limiting variables and the joint convergence follow in an analogous way as for Theorem 3.5 then, using (4.6) to establish stable convergence of the corresponding lengths of the intervals. □

Proposition D.2. *On the assumptions of Theorem 4.4:*

$$\lim_{q\to\infty}\limsup_{r\to\infty}\limsup_{n\to\infty}\sqrt{n}\left|\big[\widehat{X^{(1)},X^{(2)}}\big]_t^{(HY),n} - \text{(D.1)}\right| \xrightarrow{\mathbb{P}} 0 \tag{D.3}$$

and, moreover, the right-hand side of (D.3)$/\sqrt{n}$ *tends to zero in probability on milder assumptions as long as the mesh* $\pi_n \to 0$*.*

Proof. Based on the decomposition of X as before, the terms are treated analogously as in the proof of Theorem 2.2. Most upper bounds can be deduced along the same lines, with the exception that based on the illustration (4.2d) one has also to consider interpolated terms and dependence of adjacent addends. Yet, when denoting Δ_k^+, Δ_k^- the interpolation intervals which are non-zero, we may employ the simple estimate

$$\Delta_k^2 + \Delta_k\Delta_k^+ + \Delta_k\Delta_k^- + \Delta_k^+\Delta_k^- \le \Delta_k^2 + \Delta_{k+1}\Delta_k + \Delta_{k-1}\Delta_k + \Delta_{k-1}\Delta_{k+1}\,,$$

where $\Delta_{k+1} = T_{k+1} - T_k$ are refresh time instants as before. After an application of the Cauchy-Schwarz inequality and/or measurability arguments the addends of the remainder terms have the same structure as in the synchronous case. Hence, consistency of the HY-estimator and the CLT readily follow from the standard estimates (A.1a)–(A.1e) with the strategy of proof from the proof of Theorem 2.2. □

References

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *Int. Econ. Rev.*, 39(4):885–905.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *J. Econometrics*, 162(2):149–169.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 64(2):253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *J. Fin. Econometrics*, 2:1–48.

Bibinger, M. (2011). Asymptotics of asynchronicity. *Technical Report, Humboldt-Universität zu Berlin*, URL=`http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2011-033.pdf`.

Cox, D. (1970). *Renewal Theory*. Methuen & Co, London.

Fukasawa, M. and Rosenbaum, M. (2012). Central limit theorems for realized volatility under hitting times of an irregular grid. *Stoch. proc. appl.*, 122(12):3901–3920.

Hayashi, T., Jacod, J., and Yoshida, N. (2011). Irregular sampling and central limit theorems for power variations: The continuous case. *Ann. Inst. H. Poincaré Probab. Statist.*, 47(4):1197–1218.

Hayashi, T. and Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2):359–379.

Hayashi, T. and Yoshida, N. (2008). Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Ann. Inst. Statist. Math.*, 60(2):367–406.

Hayashi, T. and Yoshida, N. (2011). Nonsynchronous covariation process and limit theorems. *Stoch. Proc. Appl.*, 121:2416–2454.

Jacod, J. (1997). On continuous conditional gaussian martingales and stable convergence in law. *Séminaire de Probabilitiés, Strasbourg, tome*, 31:232–246.

Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Proc. Appl.*, 118(4):517–559.

Jacod, J. (2012). Statistics and high frequency data. *Proceedings of the 7th Séminaire Européen de Statistique, Cartagena, 2007: Statistical methods for stochastic differential equations, edited by M. Kessler, A. Lindner and M. Sørensen, CRC Press, Boca Raton*, pages 191–308.

Jacod, J., Podolskij, M., and Vetter, M. (2010). Limit theorems for moving averages of discretized processes plus noise. *Ann. Statist.*, 38(3):1478–1545.

Jacod, J. and Protter, P. (1998). Asymptotic error distributions for the euler method for stochastic differential equations. *Ann. Probab.*, 26:267–307.

Jacod, J. and Shiryaev, A. N. (2003). *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.*, 36(2):270–296.

Mykland, P. and Zhang, L. (2012). The econometrics of high frequency data. *Proceedings of the 7th Séminaire Européen de Statistique, Cartagena, 2007: Statistical methods for stochastic differential equations, edited by M. Kessler, A. Lindner and M. Sørensen, CRC Press, Boca Raton*, pages 109–190.

Mykland, P. A. and Zhang, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica*, 77(5):1403–1445.

Podolskij, M. and Vetter, M. (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658.

Podolskij, M. and Vetter, M. (2010). Understanding limit theorems for semimartingales: a short survey. *Stat. Neerl.*, 64(3):329–351.

Veraart, A. E. D. (2010). Inference for the jump part of quadratic variation of Itô semimartingales. *Econometric Theory*, 26(2):331–368.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.*, 100(472):1394–1411.