

IDENTIFYING THE BROWNIAN COVARIATION FROM THE CO-JUMPS GIVEN DISCRETE OBSERVATIONS

Cecilia Mancini¹ and Fabio Gobbi²

24 March 2010

Abstract

In this paper we consider two semimartingales driven by Wiener processes and (possibly infinite activity) jumps. Given discrete observations we separately estimate the *integrated covariation* IC from the sum of the co-jumps.

The Realized Covariation (RC) approaches the sum of IC with the co-jumps as the number of observations increases to infinity. Our threshold (or truncated) estimator \hat{IC}_n excludes from RC all the terms containing jumps in the finite activity case and the terms containing jumps over the threshold in the infinite activity case, and is consistent. To further measure the dependence between the two processes also the betas, $\beta^{(1,2)}$ and $\beta^{(2,1)}$, and the correlation coefficient $\rho^{(1,2)}$ among the Brownian semimartingale parts are consistently estimated.

In presence of only finite activity jumps: 1) we reach CLTs for \hat{IC}_n , $\hat{\beta}^{(i,j)}$ and $\hat{\rho}^{(1,2)}$; 2) combining thresholding with the observations selection proposed in [16] we reach an estimate of IC which is robust to asynchronous data.

We report the results of an illustrative application, made in a web appendix³, to two very different simulated realistic asset price models and we see that the finite sample performances of \hat{IC}_n and of the sum of the co-jumps estimator are good for values of the observation step large enough to avoid

¹Department of Mathematics for Decisions, University of Florence

²Department of Mathematical Economics, University of Bologna

³see on <http://www.dmd.unifi.it/upload/sub/persona/mancini/WebAppendix3.pdf>

the typical problems arising in presence of microstructure noises in the data. However we find that the co-jumps estimators are more sensible than \hat{IC}_n to the choice of the threshold. Finding criteria for optimal threshold selection is object of further research.

Keywords: co-jumps, integrated covariation, integrated variance, finite activity jumps, infinite activity jumps, threshold estimator.

1 Introduction

We consider two state variables evolving as follows

$$dX_t^{(q)} = a_t^{(q)} dt + \sigma_t^{(q)} dW_t^{(q)} + dJ_t^{(q)}, \quad q = 1, 2$$

for $t \in [0, T]$, $T < \infty$ fixed, where a and σ are càdlàg stochastic processes; $W^{(1)}$ is a standard Brownian motion and, for all $t \in [0, T]$, $dW_t^{(2)}$ is a combination $\rho_t dW_t^{(1)} + \sqrt{1 - \rho_t^2} dW_t^{(3)}$, where $W^{(3)}$ is an independent standard Brownian motion, and ρ a stochastic process with cadlag paths, in particular $d \langle W^{(1)}, W^{(2)} \rangle_t = \rho_t dt$; $J^{(1)}$, $J^{(2)}$ are possibly correlated pure jump processes. Given discrete observations $X_{\tau_j}^{(1)}, X_{\nu_i}^{(2)}$, with observation times spanned in $[0, T]$, we are interested in the separate identification of the *integrated covariation* $IC_T := \int_0^T \rho_t \sigma_t^{(1)} \sigma_t^{(2)} dt$, the covariation between the two Brownian semimartingale parts, and of the sum of the *co-jumps* $\Delta J_t^{(1)} \Delta J_t^{(2)}$, the simultaneous jumps of $X^{(1)}$ and $X^{(2)}$, where, for each $q = 1, 2$, $\Delta J_t^{(q)}$ denotes the size $J_t^{(q)} - J_{t-}^{(q)}$ of the jump occurred at time t .

The recent empirical interest in co-jumps in financial econometrics ([9]) is motivated by the problem of a correct assets price model selection. This has important

consequences in forecasting, in option pricing, in portfolio risk management, and even in the credit risk management, since a default of a firm is interpretable as a jump in the firm value and contemporaneous defaults give a co-jump, implying default dependence (contagion, [14]).

A commonly used approach to estimate $\int_0^T \rho_t \sigma_t^{(1)} \sigma_t^{(2)} dt$ is to take synchronous and evenly-spaced observations $X_{t_0}^{(1)}, X_{t_1}^{(1)}, \dots, X_{t_n}^{(1)}, X_{t_0}^{(2)}, X_{t_1}^{(2)}, \dots, X_{t_n}^{(2)}$, with $t_n = T, n \in \mathbb{N}$ and to consider the sum of cross products $\sum_{j=1}^n \Delta_j X^{(1)} \Delta_j X^{(2)}$, where $\Delta_j X^{(q)} := X_{t_j}^{(q)} - X_{t_{j-1}}^{(q)}$; however this estimate can be highly biased when the processes $X^{(q)}$ contain jumps; in fact, as $n \rightarrow \infty$, such a sum approaches the global quadratic covariation

$$[X^{(1)}, X^{(2)}]_T = \int_0^T \rho_t \sigma_t^{(1)} \sigma_t^{(2)} dt + \sum_{0 \leq t \leq T} \Delta J_t^{(1)} \Delta J_t^{(2)},$$

which also contains the co-jumps. To our aim it is crucial to single out the time intervals where the jumps occurred.

A jump process J is said to have *finite activity* (FA) when a.s. only a finite number of jumps can occur in each finite time interval. On the contrary J is said to have *infinite activity* (IA). In the special case where J is a Lévy process and has IA then a.s. infinitely many jumps occur in *each* finite time interval.

Our estimator of IC_T is based on a threshold criterion (as introduced by [24]) allowing to identify all the time intervals $]t_{j-1}, t_j]$ where the path of a univariate semimartingale jumped, if the jump component J has FA, and the intervals where jumps over the threshold occurred, if the discretely observed realization of J has infinite activity ([26]). Extending the application of the criterion to a bivariate framework allows to derive an asymptotically unbiased estimator of IC_T as well as of the sum of the co-jumps occurred up to time T , of the regression coefficients

betas, $\beta^{(1,2)}$ and $\beta^{(2,1)}$, and of the correlation coefficient $\rho^{(1,2)}$ between the Brownian semimartingale components of the given processes. More precisely we construct the following estimator

$$\hat{IC}_{T,n} := \sum_{j=1}^n \Delta_j X^{(1)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} \Delta_j X^{(2)} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}}, \quad h = T/n,$$

where only the variations under a given threshold function r_h are taken into account. Compared to the univariate case, the bivariate one presents at least the following two further problems. First in general the observations $X_{\tau_j}^{(1)}, X_{\nu_i}^{(2)}$ are not synchronous. Secondly, in presence of infinite activity jumps the speed of convergence of our estimator turns out to depend on the two $J^{(q)}$ jump activity indexes in a more complicated way than in the univariate case. More importantly, the dependence structure between the $J^{(q)}$ s also influences the speed. We will tackle the first point in section 3.1, while the second one is dealt with in [29].

The first main result of our paper is showing the convergence in probability of $\hat{IC}_{T,n}$ to IC_T , as the number n of observations tends to infinity. Not equally spaced but synchronous observations are allowed for such a result. The second group of results is given in presence of only FA jumps. Even when we only have non-synchronous data we reach consistency by modifying our estimator in a similar way as in [16] and [15]. When observations are evenly spaced, we prove a joint CLT delivering:

1. that $\hat{IC}_{T,n}$ is asymptotically Gaussian and converges with speed \sqrt{h} , which extends results in [4] who estimated IC_T in absence of jumps;
2. consistent and asymptotically Gaussian estimators of the β s and of $\rho^{(1,2)}$.

The threshold criterion originated in [24] to separate the continuous and the jump parts of a univariate parametric Poisson-Gaussian model. The criterion was

shown to work even in nonparametric frameworks in [25], [26] and [19].

The literature on non-parametric inference for stochastic processes driven by Wiener processes plus jumps, based on discrete observations, is mainly devoted to univariate cases. As for bivariate processes, [10] proposes to estimate IC_T by a weighted sum of the cross products $\Delta_j X^{(1)} \Delta_j X^{(2)}$ with a *soft rejection* weighting function in place of our indicators; [6] and [22] explore two different tests for the presence of co-jumps basically using cross multipower variations. A discussion on such approaches follows in section 3. We adopt the threshold method here since it has more favorable asymptotic properties (efficiency), as explained in more detail in section 3.

An outline of the paper is as follows. In section 2 we illustrate the framework; in section 3 we deal with the case where each component $J^{(q)}$ of $X^{(q)}$ has finite jump activity. We show that $\hat{IC}_{T,n}$ is asymptotically Gaussian (implying that it is also consistent). We find a joint CLT allowing to estimate the β s and $\rho^{(1,2)}$ of the Brownian semimartingale parts of our processes $X^{(q)}$, and we deal even with the non-synchronous observations case. In section 4 we deal with the case where each $J^{(q)}$ can have an IA semimartingale jump component $\tilde{J}_2^{(q)}$. We show that our estimator is still consistent. Since the given theory asserts that we can asymptotically identify the quantities of our interest, in section 5 we report the results obtained in a web appendix where we simulate realistic asset price models: we find that in fact the finite sample performances of $\hat{IC}_{T,n}$ and of the sum of the co-jumps estimator are good even for time step between the observations large enough (five minutes) to avoid considering microstructure effects on the data. Section 6 concludes and section 7 contains all the proofs and technical details.

2 Framework and notation

Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)$ where $X^{(1)} = (X_t^{(1)})_{t \in [0, T]}$ and $X^{(2)} = (X_t^{(2)})_{t \in [0, T]}$ are two real processes defined by

$$X_t^{(q)} = \int_0^t a_s^{(q)} ds + \int_0^t \sigma_s^{(q)} dW_s^{(q)} + J_t^{(q)}, \quad t \in [0, T], \quad q = 1, 2 \quad (1)$$

and

A1. $dW_t^{(2)} = \rho_t dW_t^{(1)} + \sqrt{1 - \rho_t^2} dW_t^{(3)}$, $W^{(1)}$ and $W^{(3)}$ are independent standard Brownian motions.

A2. The coefficients $\sigma^{(q)} = (\sigma_t^{(q)})_{t \in [0, T]}$, $a^{(q)} = (a_t^{(q)})_{t \in [0, T]}$, $q = 1, 2$, and $\rho = (\rho_t)_{t \in [0, T]}$ are càdlàg adapted processes.

Processes $J^{(q)}$ are pure jump components. In the next section we assume they have FA, i.e.

$$J_t^{(q)} = \sum_{k=1}^{N_t^{(q)}} \gamma_{\tau_k}^{(q)}, \quad q = 1, 2,$$

as specified with more detail below, where $N^{(q)} = (N_t^{(q)})_{t \in [0, T]}$ are any counting processes with $E[N_T^{(q)}] < \infty$. In section 4 each $J^{(q)}$ is assumed to be a pure jump Itô semimartingale with possibly IA.

To begin with we assume to have equally spaced and synchronous observations. The consistency results under not equally spaced but synchronous observations are straightforward using Lemma 2.1 and Theorem 7.1 below. Generalization to not equally spaced and not synchronous observations are dealt with in section 3. Let, for each n , $\pi_n = \{0 = t_0^{(n)} < t_1^{(n)} < \dots < t_n^{(n)} = T\}$ be a partition of $[0, T]$. For simplicity let us write t_j in place of $t_j^{(n)}$. Define $h := t_j - t_{j-1} = \frac{T}{n}$, for every $j = 1, \dots, n$

and $n = 1, 2, \dots$. Note that $h \rightarrow 0$ if and only if $n \rightarrow \infty$, so when computing the limits of our interest we indifferently indicate one of the two.

A3. We choose a deterministic function, $h \mapsto r_h$, satisfying the following properties

$$\lim_{h \rightarrow 0} r_h = 0, \quad \lim_{h \rightarrow 0} \frac{h \log \frac{1}{h}}{r_h} = 0.$$

For each $q = 1, 2$, the *Brownian semimartingale* part of $X^{(q)}$ is denoted by

$$D_t^{(q)} := \int_0^t a_s^{(q)} ds + \int_0^t \sigma_s^{(q)} dW_s^{(q)}.$$

As a consequence of the Paul Lévy result about the modulus of continuity of the Brownian motion paths, we can control how quickly the increments of each $D^{(q)}$ tend to zero. This is the key point to understand when $\Delta_j X^{(q)}$ is likely to contain some jumps. More precisely, the Paul Lévy law (see e.g. [23], p.114, Theorem 9.25) implies that

$$\text{a.s.} \quad \lim_{h \rightarrow 0} \sup_{j \in \{1, \dots, n\}} \frac{|\Delta_j W^{(p)}|}{\sqrt{2h \log \frac{1}{h}}} \leq 1, \quad p = 1, 3. \quad (2)$$

However even the modulus of continuity of the $D^{(q)}$ s paths share a similar property, as it is stated in the following lemma (the proof, given in [26] is reported in appendix).

Lemma 2.1. *Under A2 we have that, given an arbitrary (possibly random) partition $\{t_0 = 0, t_1, \dots, t_n = T\}$ of $[0, T]$, then for sufficiently small $h := \sup_{j=1..n} |t_j - t_{j-1}|$ we have a.s.*

$$\sup_{j=1..n} \frac{|\Delta_j D^{(q)}|}{\sqrt{2h \log \frac{1}{h}}} \leq K_q(\omega), \quad q = 1, 2$$

where $K_1(\omega) := \sqrt{M_1(\omega)} + 1$, $K_2(\omega) := \sqrt{2M_2(\omega)} + 1$ are finite random variables, $M_1(\omega) := \sup_{s \in [0, T]} (\sigma_s^{(1)})^2(\omega)$, $M_2(\omega) := \sup_{s \in [0, T]} ((\sigma_s^{(2)})^2 \rho_s^2 + \sup_{s \in [0, T]} ((\sigma_s^{(2)})^2 (1 - \rho_s^2)))(\omega)$.

The last result implies that a.s. if r_h is larger than $2K_q^2 h \log \frac{1}{h}$ (as it happens under **A3** for sufficiently small h) and if $(\Delta_j X^{(q)})^2 > r_h$ then we have $(\Delta_j X^{(q)})^2 > 2K_q^2 h \log \frac{1}{h}$, and it is likely that some jumps occurred within $]t_{j-1}, t_j]$ and made $|\Delta_j X^{(q)}|$ large. Application of Lemma 2.1 gives us the main tool for the construction of our estimators in the next section.

Notation.

- For any semimartingale Z , $\Delta Z_s = Z_s - Z_{s-}$ denotes the size of the jump of Z at time s , while $\Delta_j Z = Z_{t_j} - Z_{t_{j-1}}$ denotes the increment of process Z in the time interval $]t_{j-1}, t_j]$

- $IC_t = \int_0^t \rho_s \sigma_s^{(1)} \sigma_s^{(2)} ds$ denotes the integrated covariation up to time t , $\hat{IC}_{t,n} = \sum_{j=1..n: t_j \leq t} \Delta_j X^{(1)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} \Delta_j X^{(2)} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}}$, with $h = T/n$, is its threshold estimator

- $IV_t^{(q)} = \int_0^t (\sigma_s^{(q)})^2 ds$ denotes the integrated variance of process $X^{(q)}$, $q = 1, 2$, up to time t and $\hat{IV}_{t,n}^{(q)} = \sum_{j=1..n: t_j \leq t} (\Delta_j X^{(q)})^2 1_{\{(\Delta_j X^{(q)})^2 \leq r_h\}}$ is its threshold estimator

- sometimes $\Delta_j X^{(q)} 1_{\{(\Delta_j X^{(q)})^2 \leq r_h\}}$ is indicated briefly with $\Delta_j X_\star^{(q)}$

- sometimes we write Plim to indicate the limit in probability. \xrightarrow{st} indicates stable convergence in law of processes. See [21], ch. 8, sec. 5c, for the definition and properties of stable convergence in law, and [18] for statements of further useful properties. We recall that stable convergence in law implies convergence in distribution.

3 Finite activity jumps: consistency and central limit theorem

In this section we assume that $J^{(q)}$ is any FA jump process: for each $q = 1, 2$,

$$J_t^{(q)} = \int_0^t \gamma_s^{(q)} dN_s^{(q)} = \sum_{k=1}^{N_t^{(q)}} \gamma_{\tau_k^{(q)}}^{(q)},$$

where $N^{(q)} = (N_t^{(q)})_{t \in [0, T]}$ is a counting process with $E[N_T^{(q)}] < \infty$, $\{\tau_k^{(q)}, k = 1, \dots, N_T^{(q)}\}$ denote the instants of jump of $J^{(q)}$ and $\gamma_{\tau_k^{(q)}}^{(q)}$ denote the sizes $\Delta J_t^{(q)}$ of the jumps occurred at $\tau_k^{(q)}$. Denote

$$\underline{\gamma}^{(q)} = \min_{k=1, \dots, N_T^{(q)}} |\gamma_{\tau_k^{(q)}}^{(q)}|.$$

A4. Assume $P(\gamma_{\tau_k^{(q)}}^{(q)} = 0) = 0, \quad \forall k = 1, \dots, q = 1, 2.$

Remark 3.1. Condition **A4** implies that a.s. $\underline{\gamma}^{(q)} > 0.$

Example 3.2. If $J^{(q)}$ are FA Lévy processes, then they are of compound Poisson type ([12], Proposition 3.3, section 3.2): $N^{(q)}$ are simple Poisson processes with constant intensities $\lambda^{(q)}$ and for each q the random variables $\gamma_{\tau_k^{(q)}}^{(q)}$ are i.i.d., for $k = 1, \dots$, they are independent on $N^{(q)}$ and satisfy condition **A4**.

We recall that we are considering now deterministic equally spaced observation times $t_j = jh$, for $h = T/n, j = 1.., n$. We remark that the consistency and CLT we reach in this section are valid in presence of general finite activity jump processes, in that we do not need any assumptions on the law of the jump sizes, or of the counting processes $N^{(q)}$, nor any assumption of independence.

Now we construct our threshold estimators.

Definition 3.3. We define for $r, l \in \mathbb{N}$

$$v_{r,l}^{(n)}(X^{(1)}, X^{(2)})_t = h^{1-\frac{r+l}{2}} \sum_{j:t_j \leq t} (\Delta_j X^{(1)})^r (\Delta_j X^{(2)})^l,$$

$$w^{(n)}(X^{(1)}, X^{(2)})_t = h^{-1} \sum_{j:t_{j+1} \leq t} \prod_{i=0}^1 \Delta_{j+i} X^{(1)} \prod_{i=0}^1 \Delta_{j+i} X^{(2)}.$$

and their analogous threshold versions

$$\tilde{v}_{r,l}^{(n)}(X^{(1)}, X^{(2)})_t = h^{1-\frac{r+l}{2}} \sum_{j:t_j \leq t} (\Delta_j X^{(1)})^r \mathbf{1}_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} (\Delta_j X^{(2)})^l \mathbf{1}_{\{(\Delta_j X^{(2)})^2 \leq r_h\}},$$

$$\tilde{w}^{(n)}(X^{(1)}, X^{(2)})_t = h^{-1} \sum_{j:t_{j+1} \leq t} \prod_{i=0}^1 \Delta_{j+i} X^{(1)} \mathbf{1}_{\{(\Delta_{j+i} X^{(1)})^2 \leq r_h\}} \prod_{i=0}^1 \Delta_{j+i} X^{(2)} \mathbf{1}_{\{(\Delta_{j+i} X^{(2)})^2 \leq r_h\}}.$$

$v_{r,l}^{(n)}(X^{(1)}, X^{(2)})_T$ and $w^{(n)}(X^{(1)}, X^{(2)})_T$ are used in [4] to estimate IC_T in the case where $X^{(q)}$ are Brownian semimartingales. $\tilde{v}_{r,l}^{(n)}(X^{(1)}, X^{(2)})_T$ and $\tilde{w}^{(n)}(X^{(1)}, X^{(2)})_T$ are modified versions for the case of models with jumps: by Theorem 7.1 they exclude from the sums the terms containing jumps. Note that $\tilde{v}_{1,1}^{(n)}(X^{(1)}, X^{(2)})_t = \hat{IC}_{t,n}$, for all $t \in [0, T]$.

In view of the practical applications of our estimator we are now interested in the speed of convergence of $\hat{IC}_{T,n}$. We in fact reach even more. The first main result of this section is a joint central limit theorem for the threshold estimators

$$\begin{pmatrix} \hat{IV}^{(1)} & \hat{IC} \\ \hat{IC} & \hat{IV}^{(2)} \end{pmatrix}$$

which implies that in presence of finite activity jumps $\hat{IC}_{t,n}$ converges to IC_t for all $t \leq T$, at speed \sqrt{h} , $h = T/n$, and it allows to give estimators of standard dependence measures between the Brownian semimartingale parts $D^{(q)}$ of our processes $X^{(q)}$, such as the *Brownian parts regression coefficients* up to time t

$$\beta_t^{(1,2)} := \frac{IC_t}{IV_t^{(2)}}, \quad \beta_t^{(2,1)} := \frac{IC_t}{IV_t^{(1)}}$$

and the *Brownian parts correlation*

$$\rho_t^{(1,2)} := \frac{IC_t}{\sqrt{IV_t^{(1)} IV_t^{(2)}}}.$$

Theorem 3.4 (Joint CLT, FA jumps). *Under assumptions from **A1** to **A4**, with $h = T/n$, we have, as $h \rightarrow 0$,*

$$h^{-1/2} \begin{pmatrix} \hat{IV}_n^{(1)} - IV^{(1)} & \hat{IC}_n - IC \\ \hat{IC}_n - IC & \hat{IV}_n^{(2)} - IV^{(2)} \end{pmatrix} \xrightarrow{st} \frac{1}{\sqrt{2}} \begin{pmatrix} 2Z_{11} & Z_{12} + Z_{21} \\ Z_{12} + Z_{21} & 2Z_{22} \end{pmatrix},$$

where \mathbf{Z} is the 2×2 process with components

$$\begin{aligned} Z_{11,t} &:= \int_0^t (\sigma_s^{(1)})^2 dB_{11s} \\ Z_{12,t} &:= \int_0^t \rho_s \sigma_s^{(1)} \sigma_s^{(2)} dB_{11s} + \int_0^t \sqrt{1 - \rho_s^2} \sigma_s^{(1)} \sigma_s^{(2)} dB_{12s} \\ Z_{21,t} &:= \int_0^t \rho_s \sigma_s^{(1)} \sigma_s^{(2)} dB_{11s} + \int_0^t \sqrt{1 - \rho_s^2} \sigma_s^{(1)} \sigma_s^{(2)} dB_{21s} \\ Z_{22,t} &:= \int_0^t \rho_s^2 (\sigma_s^{(2)})^2 dB_{11s} + \int_0^t \rho_s \sqrt{1 - \rho_s^2} (\sigma_s^{(2)})^2 (dB_{12s} + dB_{21s}) + \int_0^t (1 - \rho_s^2) (\sigma_s^{(2)})^2 dB_{22s} \end{aligned} \quad (3)$$

and \mathbf{B} is a 2×2 -dimensional standard Brownian motion independent on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)$ where our model is defined.

Remark. The asymptotic conditional variances of the estimation errors $h^{-1/2}(\hat{IV}_{T,n}^{(q)} - IV_T^{(q)})$ and $h^{-1/2}(\hat{IC}_{T,n} - IC_T)$ are explicitly given by

$$Var(\sqrt{2}Z_{qq,T} | \mathcal{F}) = 2 \int_0^T (\sigma_s^{(q)})^4 ds, \quad q = 1, 2,$$

which is consistent with [26], and

$$Var\left(\frac{Z_{12,T} + Z_{21,T}}{\sqrt{2}} | \mathcal{F}\right) = \int_0^T (\sigma_s^{(1)})^2 (\sigma_s^{(2)})^2 (\rho_s^2 + 1) ds.$$

Moreover the conditional asymptotic covariance between the two estimation errors

$h^{-1/2}(\hat{IV}_{T,n}^{(q)} - IV_T^{(q)})$ and $h^{-1/2}(\hat{IC}_{T,n} - IC_T)$ is given by

$$Cov\left(\sqrt{2}Z_{qq,T}, \frac{Z_{12,T} + Z_{21,T}}{\sqrt{2}} | \mathcal{F}\right) = 2 \int_0^T (\sigma_s^{(q)})^3 \sigma_s^{(3-q)} \rho_s ds, \quad q = 1, 2.$$

Corollary 3.5 (Consistency, FA jumps). *Under **A1** to **A4**, as $n \rightarrow \infty$, for all $t \in [0, T]$*

$$\hat{IC}_{t,n} \xrightarrow{P} IC_t,$$

if a.s. $IV_t^{(j)} \neq 0$ then

$$\hat{\beta}_{t,n}^{(i,j)} := \frac{\hat{IC}_{t,n}}{\hat{IV}_{t,n}^{(j)}} \xrightarrow{P} \beta_t^{(i,j)}, \quad (i,j) = (1,2), (2,1)$$

if a.s. $IV_t^{(1)}IV_t^{(2)} \neq 0$ then

$$\hat{\rho}_{t,n}^{(1,2)} := \frac{\hat{IC}_{t,n}}{\sqrt{\hat{IV}_{t,n}^{(1)}\hat{IV}_{t,n}^{(2)}}} \xrightarrow{P} \rho_t^{(1,2)}.$$

Corollary 3.6 (Speed of convergence of β s and ρ , FA jumps). *If a.s. $IV_t^{(j)} \neq 0$ for all $t \in [0, T]$ we have, for $(i,j) = (1,2)$ or $(2,1)$,*

$$h^{-1/2} \left(\hat{\beta}_n^{(i,j)} - \beta^{(i,j)} \right) \xrightarrow{st} \frac{Z_{12} + Z_{21}}{\sqrt{2}IV^{(j)}} + \sqrt{2}Z_{jj} \frac{IC}{(IV^{(j)})^2}.$$

If a.s. $IV_t^{(1)}IV_t^{(2)} \neq 0$ for all $t \in [0, T]$ we have

$$h^{-1/2} \left(\hat{\rho}_n^{(1,2)} - \frac{IC}{\sqrt{IV^{(1)}IV^{(2)}}} \right) \xrightarrow{st} \frac{Z_{12} + Z_{21}}{\sqrt{2}IV^{(1)}IV^{(2)}} - \frac{Z_{22}IC}{\sqrt{2}IV^{(1)}(IV^{(2)})^{3/2}} - \frac{Z_{11}IC}{\sqrt{2}IV^{(2)}(IV^{(1)})^{3/2}}.$$

Remark. The asymptotic conditional variances of the estimation errors $h^{-1/2}(\hat{\beta}_{T,n}^{(i,j)} - \beta_T^{(i,j)})$ and $h^{-1/2}(\hat{\rho}_{T,n}^{(1,2)} - \rho_T^{(1,2)})$ are explicitly given by

$$\begin{aligned} & \text{Var} \left(\frac{Z_{12,T} + Z_{21,T}}{\sqrt{2}IV_T^{(j)}} + \sqrt{2}Z_{jj,T} \frac{IC_T}{(IV_T^{(j)})^2} \middle| \mathcal{F} \right) = \\ & \frac{\int_0^T (\sigma_s^{(1)})^2 (\sigma_s^{(2)})^2 (\rho_s^2 + 1) ds}{\left(\int_0^T (\sigma_s^{(j)})^2 ds \right)^2} + 2 \frac{\int_0^T (\sigma_s^{(j)})^4 ds}{\left(\int_0^T (\sigma_s^{(j)})^2 ds \right)^4} \left(\int_0^T \rho_s \sigma_s^{(1)} \sigma_s^{(2)} ds \right)^2 + \\ & 4 \frac{\int_0^T \rho_s \sigma_s^{(1)} \sigma_s^{(2)} ds}{\left(\int_0^T (\sigma_s^{(j)})^2 ds \right)^3} \int_0^T \rho_s (\sigma_s^{(j)})^3 \sigma_s^{(i)} ds, \quad (i,j) = (1,2) \text{ or } (2,1), \end{aligned}$$

and

$$\begin{aligned} \text{Var} \left(\frac{Z_{12,T} + Z_{21,T}}{\sqrt{2}\sqrt{IV_T^{(1)} IV_T^{(2)}}} - \sqrt{2}Z_{22} \frac{IC_T}{2\sqrt{IV_T^{(1)}(IV_T^{(2)})^{3/2}}} - \sqrt{2}Z_{11} \frac{IC_T}{2\sqrt{IV_T^{(2)}(IV_T^{(1)})^{3/2}}} \middle| \mathcal{F} \right) = \\ \frac{\int_0^T (\sigma_s^{(1)})^2 (\sigma_s^{(2)})^2 (\rho_s^2 + 1) ds}{IV_T^{(1)} IV_T^{(2)}} + \frac{IC_T}{IV_T^{(1)} IV_T^{(2)}} \left(\frac{IC_T \int_0^T (\sigma_s^{(1)})^4 ds}{2(IV_T^{(1)})^2} + \frac{IC_T \int_0^T (\sigma_s^{(2)})^4 ds}{2(IV_T^{(2)})^2} \right. \\ \left. - \frac{2 \int_0^T (\sigma_s^{(2)})^3 \sigma_s^{(1)} \rho_s ds}{IV_T^{(2)}} - \frac{2 \int_0^T (\sigma_s^{(1)})^3 \sigma_s^{(2)} \rho_s ds}{IV_T^{(1)}} \right) + \frac{(IC_T)^2}{(IV_T^{(1)})^2 (IV_T^{(2)})^2} \int_0^T (\sigma_s^{(1)})^2 (\sigma_s^{(2)})^2 \rho_s^2 ds. \end{aligned}$$

In [33] a threshold estimator of $\hat{\beta}^{(1,2)}$ with reduced asymptotic variance is introduced at the cost of using an auxiliary asset $X^{(3)}$.

The following proposition allows us to give a CLT for the standardized version of the estimation error $\hat{IC}_{t,n} - IC_t$.

Proposition 3.7 (Estimate of the standard error for $\hat{IC}_{t,n}$, FA jumps). *Under assumptions **A1** to **A4** we have, for all $t \in [0, T]$,*

$$\tilde{v}_{2,2}^{(n)}(X^{(1)}, X^{(2)})_t - \tilde{w}^{(n)}(X^{(1)}, X^{(2)})_t \xrightarrow{P} \int_0^t (1 + \rho_s^2) (\sigma_s^{(1)})^2 (\sigma_s^{(2)})^2 ds.$$

We now are ready to present the central limit theorem for the standardized estimation error.

Corollary 3.8 (CLT for the standardized version of $\hat{IC}_{t,n} - IC_t$, FA jumps). *Under **A1** to **A4**, for any $t \in [0, T]$, if a.s. $\int_0^t (1 + \rho_s^2) (\sigma_s^{(1)})^2 (\sigma_s^{(2)})^2 ds \neq 0$ we have*

$$\frac{\hat{IC}_{t,n} - IC_t}{\sqrt{h} \sqrt{\tilde{v}_{2,2}^{(n)}(X^{(1)}, X^{(2)})_t - \tilde{w}^{(n)}(X^{(1)}, X^{(2)})_t}} \xrightarrow{st} \mathcal{N},$$

where \mathcal{N} denotes a standard Gaussian random variable.

Remark. Under the further assumption that σ in turn is an Ito semimartingale, in [18] a CLT for $\hat{IC}_{t,n}$ is also shown and leads to consistent results.

Remark 3.9 (Comparison in efficiency with other methods in the literature). In the literature of the disentangling IC from the co-jumps at our knowledge there exist only two other approaches: the *Realized BiPower Covariation* (RBPC, [6]) and the *Realized Outlyingness Weighted Quadratic Covariation* (ROWQC, [10]). In both papers a CLT is shown to hold only in absence of jumps. A comparison in the efficiency of RBPC and ROWQC is made in [10] for the univariate case under a Brownian semimartingale model. We add here further information involving also the threshold estimator (that we indicate by RTC, Realized Threshold Covariation) to argue why the latter is preferable.

Firstly note that in the univariate case and in presence of jumps an estimator of $IV^{(1)}$ is efficient (i.e. with minimum possible asymptotic variance, given n discrete observations) if $AVar$, the asymptotic variance of the quotient of the estimation error by \sqrt{h} , is $2 \times IQ_T^{(1)}$, where $IQ_T^{(1)} := \int_0^T (\sigma_t^{(1)})^4 dt$ (see [1] for the case of constant $\sigma^{(1)}$). The $AVar$ of all the estimators RBPC (in absence of jumps), ROWQC (in absence of jumps), RTC, as well as of the multipower variation estimators RMPV of $IV^{(1)}$ (even in presence of jumps but with maximum power less than 1), equals a factor θ times $IQ_T^{(1)}$, so the efficiency comparison is done through the magnitude of the factors $\theta_{RBPC}, \theta_{ROWQC}, \theta_{RTC}, \theta_{RMPV}$.

In the univariate case the threshold estimator is efficient ([26]), since $\theta_{RTC} = 2$, both in absence and in presence of jumps (both with finite activity and with infinite activity but finite variation, [26], [27]). The Bipower Variation estimator (with powers 1,1) is not efficient since in absence of jumps $\theta_{RBPC} = 2.609$ ([5]) and in presence of jumps it is even higher ([35]). As for the RMPVs that estimate $IV_T^{(1)}$ with at least three powers we have CLTs with factors $\theta_{RMPV} > 2.609$ in all cases (see

[36] and the discussion in [26], section 3.3). As for ROWQC, in the univariate case it consists in a weighted sum of squared increments of $X^{(1)}$. A measure of outlyingness (w.r.t. the Brownian motion squared increments behavior) of each $(\Delta_j X^{(1)})^2$ is given. Such a measure is a $\chi^2(1)$ random variable when the increment comes from a Gaussian process, so an increment is considered outlier (i.e. containing jumps) if its measure is above a threshold k equal to the $1 - \alpha$ quantile of the $\chi^2(1)$ law. The higher is the outlyingness measure the smaller is the weight given to the increment to contribute to ROWQC, on the basis of a chosen weighting function (called *soft rejection function*, SR). The estimator efficiency in absence of jumps (table 2 in [10]) heavily depends, in a decreasing way, on the choice of α , the minimal θ factor in absence of jumps is reached for $\alpha = 0$ and is 2. Since the chosen SR weighting function is $w(z) = 1 \wedge (k/z)$, $\alpha = 0$ means that ROWQC coincides with the realized variance estimator RV. In presence of jumps ROWQC is consistent for $IV^{(1)}$, however we cannot choose $\alpha = 0$ since RV is biased, therefore we cannot have an efficient ROWQC estimator.

In a bivariate framework, the AVar (the asymptotic variance of $(\hat{IC}_{T,n} - IC_T)/\sqrt{h}$) is a quite complicated function of $\rho_t^2, \sigma_t^{(1)}, \sigma_t^{(2)}$ for RBPC in absence of jumps. For instance it equals $8.46 \times \int_0^T (1 + \rho_t^2)(\sigma_t^{(1)})^4 dt$ if $(\sigma_t^{(1)})^2 \equiv (\sigma_t^{(2)})^2$. On the contrary for the threshold estimator, even in presence of (finite variation) jumps, we have $AVar_{RTC} = \int_0^T (1 + \rho_t^2)(\sigma_t^{(1)})^2(\sigma_t^{(2)})^2 dt$ for any values of $\rho_t, \sigma_t^{(1)}, \sigma_t^{(2)}$. This is less than $AVar_{RBPC}$ at least when $(\sigma_t^{(1)})^2 \equiv (\sigma_t^{(2)})^2$, as well as in all the other situations we report in more detail in appendix (remark 7.2). In fact in [9], where the bipower covariation test of [6] has been discussed, it is shown that, when dealing with large portfolios, it is necessary to use a different global cross-variation index to get reliable

results. For ROWQC the asymptotic variance $AVar_{ROWQC}$, in absence of jumps, depends on the weighting function w and is minimal at $AVar_{RTC}$, when $w(z) = 1$ for all $z \geq 0$. Analogously as in the univariate case the latter weighting function cannot be chosen in presence of jumps.

Remark 3.10 (Estimate of the co-jumps). By Corollary 3.5, clearly we have an estimate of the sum of the co-jumps up to T simply subtracting $\hat{IC}_{T,n}$ from the quadratic covariation estimator:

$$\sum_{j=1}^n \Delta_j X^{(1)} \Delta_j X^{(2)} - \hat{IC}_{T,n} \xrightarrow{P} \sum_{0 \leq s \leq T} \Delta J_s^{(1)} \Delta J_s^{(2)},$$

as $n \rightarrow \infty$. Analogously we can obtain an estimator of the sum of the co-jumps up to each time $t \in [0, T]$.

An estimate of each $\Delta J_s^{(1)} \Delta J_s^{(2)}$, with $s \in [0, T]$, can be obtained using

$$\Delta_j X^{(1)} \Delta_j X^{(2)} - \Delta_j X^{(1)} \mathbf{1}_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} \Delta_j X^{(2)} \mathbf{1}_{\{(\Delta_j X^{(2)})^2 \leq r_h\}}, \quad (4)$$

with j such that $s \in [t_{j-1}, t_j]$. Alternatively, as we consider one single term, and not the sum of n terms, even

$$\Delta_j X^{(1)} \mathbf{1}_{\{(\Delta_j X^{(1)})^2 > r_h\}} \Delta_j X^{(2)} \mathbf{1}_{\{(\Delta_j X^{(2)})^2 > r_h\}} \quad (5)$$

or

$$\Delta_j X^{(1)} \Delta_j X^{(2)} \quad (6)$$

estimate the co-jump $\Delta J_s^{(1)} \Delta J_s^{(2)}$, with $s \in [t_{j-1}, t_j]$, since $|\Delta_j X^{(q)} \mathbf{1}_{\{(\Delta_j X^{(q)})^2 \leq r_h\}} \Delta_j X^{(\ell)}| \leq 2\sqrt{r_h} \sup_{s \in [0, T]} |X^{(\ell)}|$, $q = 1, 2$, $\ell = 3 - q$, and $|\Delta_j X^{(1)}| \mathbf{1}_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} |\Delta_j X^{(2)}| \times \mathbf{1}_{\{(\Delta_j X^{(2)})^2 \leq r_h\}} \leq r_h$ tend to zero in probability as $h \rightarrow 0$, by the pathwise boundedness of each $X^{(\ell)}$ on $[0, T]$. However in the web appendix we show that estimator

(4) has the best finite sample properties, in terms of relative estimation error, in our simulated Model 1 with finite activity jumps.

In presence of (even infinite activity) jumps in a bivariate model, [22] use $\hat{B}_n := \sum_{j=1}^n (\Delta_j X^{(1)})^2 (\Delta_j X^{(2)})^2$ to estimate $B := \sum_{s \leq T} (\Delta X_s^{(1)})^2 (\Delta X_s^{(2)})^2$: \hat{B}_n is not directly comparable (as for the efficiency) with the estimate of the sum of the co-jumps $\sum_{s \leq T} \Delta X_s^{(1)} \Delta X_s^{(2)}$ we give here, but it is an alternative, power-based, measure of the occurred co-jumps, aimed to give a test for their presence.

3.1 Microstructures

Remark 3.11. Finite sample performance and microstructure noises. Our theoretical results allow to estimate IC_T and the sum of the co-jumps asymptotically for $h \rightarrow 0$, while in practice for very small values of h financial time series are affected by microstructure noises which introduce a bias which is larger as h is smaller. In the web appendix we implement our estimators on simulations of realistic financial time series and we find that $\hat{IC}_{T,n}$ and the sum of the co-jumps estimator have good performances already with temporal mesh h corresponding to five minutes, a time lag at which the impact of the microstructure noises on quadratic variation is known to be immaterial ([7]).

Remark 3.12. Asynchronous observations. It is known that the problem of the estimation of the covariation among two assets undergoes the so called Epps effect, i.e. in the empirical applications the estimator tends to zero as the step size h tends to zero. The asynchronicity among the observations of $X^{(1)}$ and $X^{(2)}$ is considered one of the possible causes ([31]; [7], section 2.10.2). In fact some authors

have tackled the problem of reaching a consistent estimator of the covariation even when data are asynchronous and $h \rightarrow 0$, under a Brownian semimartingale model. In particular we connect our work to [15] and [16]. At the time scale of five minutes the Epps effect probably does not affect our estimate of IC_T . However even for smaller h it is possible to make our estimator correctly converge to the integrated covariation, as detailed below.

Assume we have access to the records $\{D_{\tau_0^{(n)}}, D_{\tau_1^{(n)}}, \dots, D_{\tau_{m^{(n)}}^{(n)}}\}^{(1)}$, $\{D_{\nu_0^{(n)}}, D_{\nu_1^{(n)}}, \dots, D_{\nu_{k^{(n)}}^{(n)}}\}^{(2)}$ of observations of two Brownian semimartingales $D^{(1)}$ and $D^{(2)}$, with the two stochastic partitions $0 = \tau_0^{(n)} < \tau_1^{(n)} < \dots < \tau_{m^{(n)}}^{(n)}$ and $0 = \nu_0^{(n)} < \nu_1^{(n)} < \dots < \nu_{k^{(n)}}^{(n)}$ spanned on $[0, T]$. For simplicity let us write ν_i and τ_j in place of $\nu_i^{(n)}$ and $\tau_j^{(n)}$. The idea of Hayashi and Yoshida ([16]) is to select only some of the cross variations $(D_{\tau_j}^{(1)} - D_{\tau_{j-1}}^{(1)})(D_{\nu_i}^{(2)} - D_{\nu_{i-1}}^{(2)})$, in order to estimate IC_T , and precisely the ones for which there is an intersection between the time intervals $[\tau_{j-1}, \tau_j]$ and $[\nu_{i-1}, \nu_i]$. Hayashi and Kusuoka ([15]) extended the consistency result to the case where the observation times can be stochastic. We show here that using their result ([15], Corollary 2.2) we in fact reach the same kind of consistency in the case of asynchronous observations even in presence of finite activity jumps. The idea is very simple: we first eliminate the jumps, using threshold technique, and then apply the Hayashi and Yoshida estimator to the estimated continuous components $\hat{D}^{(q)}$.

Recall that $(X_a^{(q)} - X_b^{(q)})_\star = (X_a^{(q)} - X_b^{(q)})1_{\{(X_a^{(q)} - X_b^{(q)})^2 \leq r_h\}}$, for any two time instants a and b , set $h := \sup_{j=1..m^{(n)}}(\tau_j - \tau_{j-1}) \vee \sup_{i=1..k^{(n)}}(\nu_i - \nu_{i-1})$. We in fact have the following

Theorem 3.13 (Asynchronous observations). *Let **A1** to **A4** hold, $0 = \tau_0 < \tau_1 < \dots < \tau_{m^{(n)}}$, $0 = \nu_0 < \nu_1 < \dots < \nu_{k^{(n)}}$ be two sequences of stopping times such that*

$\tau_{m^{(n)}} \uparrow T, \nu_{k^{(n)}} \uparrow T$ a.s., as $n \rightarrow \infty$ then as $h \rightarrow 0$

$$\sum_{j=1..m^{(n)}, i=1..k^{(n)}} (X_{\tau_j}^{(1)} - X_{\tau_{j-1}}^{(1)})_{\star} (X_{\nu_i}^{(2)} - X_{\nu_{i-1}}^{(2)})_{\star} 1_{\{[\tau_{j-1}, \tau_j] \cap [\nu_{i-1}, \nu_i] \neq \emptyset\}} \xrightarrow{P} IC_T.$$

4 Infinite activity jumps: consistency

In this section we allow the jump components of processes $X^{(q)}$ to have infinite activity, and we consider the case where $X^{(q)}$ are general Itô semimartingales. Any unidimensional Itô semimartingale has a representation as in (1) with each $J^{(q)}$ decomposed as

$$\begin{aligned} J^{(q)} &= J_1^{(q)} + \tilde{J}_2^{(q)}, \\ J_{1t}^{(q)}(\omega) &= \int_0^t \int_{|\gamma^{(q)}(\omega, x, s)| > 1} \gamma^{(q)}(\omega, x, s) \underline{\mu}^{(q)}(\omega, dx, ds), \\ \tilde{J}_{2t}^{(q)}(\omega) &= \int_0^t \int_{|\gamma^{(q)}(\omega, x, s)| \leq 1} \gamma^{(q)}(\omega, x, s) \tilde{\underline{\mu}}^{(q)}(\omega, dx, ds), \end{aligned} \quad (7)$$

where $\underline{\mu}^{(q)}$ is the Poisson random measure of the jumps of $J^{(q)}$, $\tilde{\underline{\mu}}^{(q)}(\omega, dx, ds) = \underline{\mu}^{(q)}(\omega, dx, ds) - dx ds$ is its compensated measure, the coefficients $a^{(q)}, \sigma^{(q)}, \gamma^{(q)}$ are predictable and $\int 1 \wedge (\gamma^{(q)})^2(\omega, t, x) dx$ is a.s. finite (see [18], pp.3,4; [19], (2.11)). Conditions **A2** and **A4'** below guarantee local boundedness properties of such coefficients.

A4'. $\int 1 \wedge (\gamma^{(q)})^2(\omega, t, x) dx$ is locally bounded.

For each $q = 1, 2$, $J_1^{(q)}$ is a finite activity jump process of type $J_{1t}^{(q)} = \sum_{k=1}^{N_t^{(q)}} \gamma_{\tau_k}^{(q)}$, as in section 3, but now the sizes $|\gamma_{\tau_k}^{(q)}|$ are all larger than 1; on the contrary $\tilde{J}_2^{(q)}$ has generally infinite activity jumps, since generally $\int_0^T \int_{|\gamma^{(q)}(\omega, s, x)| \leq 1} dx ds = +\infty$. $\tilde{J}_2^{(q)}$ is a compensated sum of jumps which are bounded in absolute value by 1. Therefore, for each $q = 1, 2$, $J_1^{(q)}$ accounts for the "large" and rare jumps of $X^{(q)}$, while $\tilde{J}_2^{(q)}$ accounts for the frequent and small jumps.

Example 4.1. If one of the two processes $J^{(q)}$ is a pure jump Lévy process, it is always possible to decompose it as in (7) with $\gamma^{(q)}(\omega, t, x) \equiv x$ but then the compensator has the form $\nu^{(q)}(dx) \times ds$, where $\nu^{(q)}$ is the Lévy measure of $J^{(q)}$ and is a deterministic σ -finite measure such that $\int_{\mathbb{R}} 1 \wedge x^2 \nu^{(q)}(dx) < \infty$ and generally such that $\int_{|x| \leq 1} \nu^{(q)}(dx) = +\infty$.

We prove that $\hat{IC}_{t,n}$ is still a consistent estimator of IC_t , for all $t \in [0, T]$, for any cadlag adapted processes $\sigma^{(q)}$ s. For ease of notation we only consider IC up to time T and evenly spaced synchronous observations (not evenly spaced but synchronous observations, with $h = \sup_j |t_j - t_{j-1}|$, and arbitrary $t \in [0, T]$ are straightforward).

As for the speed of convergence of $\hat{IC}_{t,n}$, in the presence of infinite activity jump components, things are more complicated in that such a speed is determined both by whether $\tilde{J}_2^{(1)}$ and $\tilde{J}_2^{(2)}$ are dependent or not and by the amount of jump activity of each $\tilde{J}_2^{(q)}$. In [29] we consider two Lévy infinite activity jump components $\tilde{J}_2^{(1)}$ and $\tilde{J}_2^{(2)}$ with a dependence structure described by a convex family of Lévy copulas. We find that, when $\tilde{J}_2^{(1)}$ and $\tilde{J}_2^{(2)}$ do depend, the speed is still \sqrt{h} only when the activity of jump of both one processes is moderate (Blumenthal-Gettoor index smaller than 1), otherwise the speed is less than \sqrt{h} (consistently with [27] in the univariate case).

We now state the main result in presence of infinite activity jumps.

Theorem 4.2 (Consistency in presence of IA jumps, synchronous observations).

Let $(X_t^{(1)})_{t \in [0, T]}$ and $(X_t^{(2)})_{t \in [0, T]}$ be two processes of the form (1), with $J^{(q)}$ as in (7).

Assume **A1**, **A2**, **A3** and **A4'**. Then

$$\hat{IC}_{T,n} \xrightarrow{P} IC_T,$$

as $n \rightarrow \infty$.

The reason why the threshold estimator is still able to isolate IC is that now we have

$$[X^{(1)}, X^{(2)}]_T = \int_0^T \rho_s \sigma_s^{(1)} \sigma_s^{(2)} ds + \sum_{s \leq T} \Delta J_{1s}^{(1)} \Delta J_{1s}^{(2)} + \sum_{s \leq T} \Delta \tilde{J}_{2s}^{(1)} \Delta \tilde{J}_{2s}^{(2)},$$

and, as it is shown in the appendix, for all $\delta > 0$ the threshold r_h cuts off all the jumps of the $J_1^{(q)}$ and the jumps of the $\tilde{J}_2^{(q)}$ larger, in absolute value, than $\sqrt{4r_h + \delta}$. However as $r_h \rightarrow 0$, since δ is arbitrary, then every jump of each process $\tilde{J}_2^{(q)}$ is cut off.

Remark. Under the further assumption that σ in turn is an Ito semimartingale, in [18] the consistency of $\hat{IC}_{t,n}$ in presence of IA jumps is also proved.

Remark 4.3. [Estimate of the sum of the co-jumps] Even in this framework of infinite activity jumps, as a consequence of Theorem 4.2, the sum of the co-jumps up to T is consistently estimated by

$$\sum_{j=1}^n \Delta_j X^{(1)} \Delta_j X^{(2)} - \sum_{j=1}^n \Delta_j X^{(1)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} \Delta_j X^{(2)} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}}.$$

Estimate of a *single* co-jump $\Delta J_s^{(1)} \Delta J_s^{(2)}$, with $s \in]t_{j-1}, t_j]$, is not possible anymore, since in general in each $]t_{j-1}, t_j]$ infinitely many co-jumps occur. We only can estimate the *sum* of the co-jumps within $]t_{j-1}, t_j]$, and in principle we could make it using (4), (5) or (6). Their performance on simulations of Model 2 with infinite activity jumps (in the web appendix) is good in terms of absolute estimation error but it is not good in terms of relative error, due to the fact that the jumps are very small. However we do not present the simulation results here since we think that they are not empirically relevant. In fact if we are interested in checking the existence of correlation in the jumps of two assets we need to estimate the *sum* of

the co-jumps. If on the other hand we are interested in checking the presence of contemporaneous shocks affecting two assets, then we need to estimate the *large* co-jumps, which in fact constitute a FA process ([12], page 38, after (2.47)). \square

5 Implementation

The performance of our estimator depends on the choice of n (number of observations) and of the threshold r_h we use. These choices can be assessed using simulations of realistic models, which is done in the same spirit as for example [17] to show that, for finite samples, the log version of the bipower variation has a better performance than bipower variation; or in the same spirit as [3] who use simulations to choose the threshold parameters α and ϖ as well as the parameters k, p and k_n involved in their estimators.

5.1 Choice of the threshold

In view of (3) we can expect that the optimal threshold r_h has to depend on $\sup_{s \in [0, T]} |\sigma_s^{(1)}|$, $\sup_{s \in [0, T]} |\sigma_s^{(2)}|$, which is what we are in fact estimating. Even for the multipower variations the exponents r_i have to be chosen depending on the Blumenthal-Gettoor index α of X (see e.g. [36]), which is unknown.

In a web appendix we check on simulations how sensitive the estimates of IC are to the choice of r_h in a class of powers of h , for fixed observation step h , and we select a reasonable threshold that is then used to estimate IC . This is only a limited investigation which is given as indicative, more sophisticated choices have been done in the literature in univariate cases: in [28] ($h = 1$ day) the threshold to be compared

to each $(\Delta_j X)^2$ is 9 times a GARCH forecast of the realized squared volatility process made at time t_{j-1} ; in [13] an iterative estimation of IV is implemented, where at each step the previously estimated squared volatility is plugged into the threshold, until the estimate stabilizes; in [34] for each day t_j the bipower variation estimator of IV made in the previous day is plugged into the threshold that is compared with $|\Delta_j X|$.

Each model has its own optimal threshold, as it is typical of all non-parametric estimators. Formal study of methods for optimal threshold selection in a given model is object of further research.

Based on the figures and tables of the web appendix, taking daily units of measure and five minutes observations, $r_h = 1.01h^{0.99}$ appears to be a good choice for estimating IC , in both the two very different models we implemented. In principle there are many functions r_h that satisfy conditions **A3**, however for $h \rightarrow 0$ we have that $(2h \log \frac{1}{h})/h^\beta \rightarrow 0$ for all $\beta \in]0, 1[$, so that, consistently e.g. with [3] and [2], the choice of a power ch^β of h seemed to be natural. Among the powers h^β , the closer β is to one, the closer is the speed of convergence (to zero) of h^β to the speed of $2h \log \frac{1}{h}$. However, for values of h in the range $[1/100000, 1/50]$ in fact we have $2h \log \frac{1}{h} > h^\beta$ for all values $\beta \in [0.9, 1]$. This means that, in order to estimate IC , thresholds like $h^{0.9}$ or $h^{0.99}$ or $h^{0.999}$ exclude many observations which in fact would be pertinent. If there are only rare jumps then a higher r_h is better, since it includes more squared increments due only to the Brownian semimartingale part. If, on the contrary, $J^{(q)}$ has infinite activity then the variations $\Delta_j \tilde{J}_2^{(q)}$ are usually quite small and the relative $\Delta_j X^{(q)}$ are included in \hat{IC}_n and give a spurious estimate of IC , so r_h has not to be too large.

Despite the fact that the simulated Model 1 with FA jumps in the web appendix has stochastic volatility, the chosen r_h has a constant factor not depending on the $\sigma^{(q)}$ s. A justification is that in realistic simulated models the values of $\sigma^{(1)}$ and $\sigma^{(2)}$ usually range in a day within $[0.01, 0.03]$, in daily units of time, so the constant factor $c=1.01$ is about 6 times the mean expected value of $(\sigma^{(q)})^2$ in annual units of measure. We expect that our selected threshold would not be good if the volatilities values were of a very different magnitude order.

The simulated models on which we selected the threshold and which are presented in detail in the web appendix are very different: in Model 1, proposed in [17], each $X^{(q)}$ has stochastic volatility and a FA compound Poisson jump part while in Model 2, proposed in [11], each $X^{(q)}$ has constant volatility and the jumps follow a CGMY dynamics and have IA. Our unit of measure is one day. The simulated models are realistic since for Model 1 the parameters of the univariate $X^{(q)}$ are taken from [17]. A path of each σ varies most between 0.0128 and 0.0204 in a day. For Model 2 processes $X^{(q)}$ are constructed taking the parameters of the stocks HWP, MSFT and GE in Table 2 of [11].

In Model 1 we have $\tilde{J}_2^{(q)} \equiv 0$ while in Model 2 $J_1^{(q)} \equiv 0$, $q = 1, 2$. In order to effectively introduce non zero co-jumps, in each model $J^{(q)}$ are linear combinations of independent jump processes of the same type. On the other hand between the Brownian semimartingale parts of the $X^{(q)}$ s we have a constant correlation $\rho = \text{corr}(W^{(1)}, W^{(2)})$. We have taken the five minutes synchronous returns ($T=1$, $h=1/288$, we considered a 24 hours open market) and constructed our daily threshold estimator $\hat{IC}_{T,n}$. We simulated 1500 bivariate paths.

The average values of IC_T and of the sum of the co-jumps are 1.2567×10^{-4}

and 9.6903×10^{-6} in Model 1 (with jump intensities $\lambda^{(1)} = 0.188, \lambda^{(2)} = 0.236$) and 7.485×10^{-3} and 5.8664×10^{-4} in Model 2. For computational convenience, in Model 1 we obtained $J^{(1)}$ and $J^{(2)}$ as combination of only 2 independent compound Poisson jumps. Since such a scheme gave non-realistic proportions of the realized quadratic covariation with IC_T in Model 2, in this case we combined 3 independent CGMY processes.

First of all for each model we implemented the estimator of IC_T as r_h varies in order to select the threshold to use. We plotted the mean percentage relative bias

$$100 \frac{\hat{IC}_{T,n} - IC_T}{IC_T}$$

as β varies in $[0.71, 0.99]$ and $c = 0.01, \dots, 1.97$ with step 0.04, in Model 1 with $\lambda^{(1)} = 0.118, \lambda^{(2)} = 0.236$ and in Model 2 respectively, for h fixed. Values of c above 0.61 give the a constant relative error of 1% in Model 1 for all β s while in Model 2 a high value of β and a value of c around 1 have to be chosen. We remark that in Model 1 lower values of the relative error could be reached (but for such a choices of c, β the estimates in Model 2 would be worse). In fact values of c higher than 0.65 are such that all the observations are included in the estimator. However the jumps are few and the generated estimation error of IC_T is acceptable. We also plotted the empirical densities and the relative QQ-plots of the normalized bias

$$NB := \frac{\hat{IC}_{T,n} - IC_T}{\sqrt{h} \sqrt{\tilde{v}_{2,2}^{(n)}(X^{(1)}, X^{(2)})_T - \tilde{w}^{(n)}(X^{(1)}, X^{(2)})_T}} \quad (8)$$

as β is fixed to level 0.99 while c varies as before, in Model 1 with $\lambda^{(1)} = 0.118$ and $\lambda^{(2)} = 0.236$ and in Model 2. As further checks we implemented both the NB in Model 1 with no jumps ($\lambda^{(1)} = \lambda^{(2)} = 0$) for fixed $\beta = 0.99$ and c varying as before,

and the NB for c fixed to level 1.01 as β varies in $[0.71, 0.99]$ with step 0.02 in Model 2. Model 2 is the most sensitive to the choice of the threshold level. The web appendix also reports tables with the descriptive statistics of all the error measures we considered, for both 5 minutes and 1 minute sample frequencies. We confirm the choice of $\beta = 0.99$ and $c=1.01$.

5.2 Estimates of IC and of $\sum_{0 \leq t \leq T} \Delta J_t^{(1)} \Delta J_t^{(2)}$ on simulations

To check the efficiency of the estimators of IC_T and of $\sum_{0 \leq t \leq T} \Delta J_t^{(1)} \Delta J_t^{(2)}$ we studied the percentage relative errors, where the threshold is the one previously selected, and keeping T fixed equal to one day and h equal to five minutes. In the web appendix we show the histograms of $100 \frac{(\hat{IC}_{T,n} - IC_T)}{IC_T}$ in Model 1, with $\lambda^{(1)} = 0.118, \lambda^{(2)} = 0.236$, and in Model 2. The relative summary statistics are also reported. We conclude that $\hat{IC}_{T,n}$ has good efficiency, even in Model 1 with no jumps, since the percentage relative errors are around 1-2%. We remark that \hat{IC}_n turns out to be quite robust in Model 1, since, as we noted before, the threshold was not optimal in this case.

On the other hand estimation of the co-jumps is much more sensible to the choice of the threshold. In the web appendix we plotted the following percentage relative estimation conditional bias

$$100 \frac{(\sum_{j=1}^n \Delta_j X^{(1)} \Delta_j X^{(2)} - \hat{IC}_{T,n}) - \sum_{0 \leq t \leq T} \Delta X_t^{(1)} \Delta X_t^{(2)}}{\sum_{0 \leq t \leq T} \Delta X_t^{(1)} \Delta X_t^{(2)}}$$

in Model 1, $\lambda^{(1)} = 0.118$ and $\lambda^{(2)} = 0.236$ and in Model 2. Since for Model 1 the jump frequencies give that only 180 generated paths (among 1500) do have co-jumps, we conditioned to the occurrence of at least one co-jump within the day, i.e. what is

computed is the distribution of the relative estimation bias for 1500 generated paths of $(X^{(1)}, X^{(2)})$ where some co-jumps occurred. We remark that in fact only one co-jump occurred in each case. In Model 2, on the contrary, each generated daily path of $(X^{(1)}, X^{(2)})$ contains some co-jumps. Simulation of Model 2 is based on an approximation of the CGMY model with a suitable compound Poisson process, and we exactly know the jump instants of this last process.

The performance of the sum of the co-jumps estimator is good in Model 2, while in Model 1 it is much worse, due to the non optimality of the choice of the threshold in this model, as we already commented. We note that in a significant number of days a co-jump occurred but the estimate is zero (the relative bias equals -100%), so our estimator did not recognize the co-jump, thus for that day $\sum \Delta_j X^{(1)} \Delta_j X^{(2)} = \hat{IC}_n$ which means that all the co-increments $\Delta_j X^{(1)} \Delta_j X^{(2)}$ are below the threshold. That means that the chosen r_h is too high in this model. Also in some days we have a distortion in the estimation of the size of the occurred co-jumps (the relative bias is close to -100%) since the estimator of the sum of the co-jumps is not zero but is very small compared to the true sum of the co-jumps (one jump in fact).

We then change the threshold choice in Model 1 by exploiting that \hat{IC}_n is quite robust to an even not really optimal threshold. After some trials we found that the threshold $33\hat{IC}h^{0.99}$ (corresponding to $c = 33\sigma^{(1)}\sigma^{(2)}\rho$ if $\sigma^{(a)}$ and ρ were constant, and being roughly $(4\sigma)^2$ if we further knew that our two stocks have roughly the same volatility) seems to give the best estimates. In Model 1 this choice gives already quite good results, but as announced a further study would have to be done before reliably estimate the sum of the co-jumps of empirical data.

5.3 Estimate of the single co-jumps

As we announced in remark 4.3 we report the results on estimating each single co-jump only for Model 1 (FA jumps, $\lambda^{(1)} = 0.118$, $\lambda^{(2)} = 0.236$). We used here the threshold function $33\hat{I}Ch^{0.99}$ we selected at the end of the previous subsection. Our goal was to check which is the most informative estimator among (4), (5) and (6).

We define by $\hat{J}J_t$ (joint jump) the estimate of $\Delta J_t^{(1)}\Delta J_t^{(2)}$. As above we considered 1 day time horizon and h equal to five minutes. The mean (on the 1500 simulated paths containing co-jumps) absolute error $\hat{J}J_t - \Delta J_t^{(1)}\Delta J_t^{(2)}$ is 2×10^{-6} (with SEE 4×10^{-6}) for estimator (4), 10^{-8} (with SEE 10^{-8}) for (5) and 6×10^{-7} (with SEE 2×10^{-7}) for (6). However the relative errors $\frac{\hat{J}J_t - \Delta J_t^{(1)}\Delta J_t^{(2)}}{\Delta J_t^{(1)}\Delta J_t^{(2)}}$ display a worse behavior. In a table in the web appendix we reported the mean conditional percentage relative error for estimators (4) and (5). Estimator (6) displays an even worse performance than (5) and we did not report it. If the sizes of the jumps were higher (higher σ_J) the performance of the estimators would be better, as it would be easier to recognize a co-jump. Estimator (4) is not bad and is preferable in any case. A reason for the bias in the relative error levels is the not correct detection of jump times: for $\sigma_J = 0.015$, at a frequency of observation of 5 minutes, estimator (4) finds about the 83% of the true co-jump times while for the 17% of the paths it does not find the existing co-jumps. Estimator (6) has a much higher error in detecting the realized co-jumps times. The co-jump times detection capability would improve with a higher σ_J level for both estimators.

6 Conclusions

In this paper we introduce new estimators of the continuous part IC and of the sum of the co-jumps in the quadratic covariation of two semimartingales $X^{(q)}$. To capture the separate contributions to the quadratic covariation has important applications in finance (model selection, forecasting, option pricing, risk and credit risk management).

The estimator $\hat{IC}_{T,n}$ is constructed using a threshold criterion introduced in [24], and consists in summing properly selected cross products of increments of the two processes. Our estimator is consistent, as a consequence also the sum of the co-jumps occurred within $[0, T]$ is consistently estimated. In presence of only FA jumps even each single co-jump is consistently estimated, further a joint CLT for $\hat{IC}_{T,n}$ and the estimators $\hat{IV}_T^{(q)}$ of the integrated variances is proved and delivers the following important consequences:

1. estimator $\hat{IC}_{T,n}$ is also asymptotically Gaussian with speed of convergence \sqrt{h} ;
2. we construct asymptotically Gaussian estimators of the regression coefficients β s and of the correlation coefficient between the Brownian semimartingale parts of the two processes $X^{(q)}$.

We remark that a central limit theorem in presence of infinite activity jumps is studied in a further paper ([29]) where we find that the speed of convergence of $\hat{IC}_{T,n}$ is determined both by whether the two processes $J^{(q)}$ are dependent or not and by the amount of jump activity of each $J^{(q)}$.

Further we find that in presence of only FA jumps a slight modification of $\hat{IC}_{T,n}$ is consistent even when only non-synchronous observations are available.

In a web appendix we assessed the choice of the threshold and checked the performance of our estimators on two different kind of simulated models which are taken from the financial literature. Model 1 has components with stochastic volatilities and FA jumps, while Model 2 has components with constant volatilities and IA jumps. We found that with five minutes observations the estimate of IC is satisfactory in both models, using threshold $r_h = 1.01h^{0.99}$, which is "optimal" in Model 2 but not in Model 1. However \hat{IC}_n seems to be quite robust to the threshold choice in Model 1. The sum of the co-jumps estimate is on the contrary more sensible than \hat{IC}_n to the choice of the threshold, and $r_h = 1.01h^{0.99}$ gives good results in Model 2, while to reach satisfactory results for Model 1 we needed to consider $r_h = 33\hat{IC}_nh^{0.99}$. Estimation of the single co-jumps in Model 1 is better done through (4) and is not bad. Deepening the investigation on proper threshold selection is a necessary task but is beyond the scope of this paper.

Acknowledgements. The authors thank (in chronological order) Rama Cont, Roberto Renò and the anonymous referees for the important comments that helped to improve and deepen our analysis, Christian Timothy Brownlees and Peter Tankov for the help in the implementation of the CGMY model, Jean Jacod for the discussion during at the conference "Stochastic processes and financial econometrics" in Florence in 2009.

The authors benefited of financial support by the Italian Government, grant MIUR 2006 n.2006132713-004.

References

- [1] Aït-sahalia, Y. (2004) Disentangling volatility from jumps, *Journal of Financial Economics*, 74, 487-528
- [2] AÏT-SAHALIA, Y. and JACOD, J. (2007). Estimating the Degree of Activity of Jumps in High Frequency Data, forthcoming in the *Annals of Statistics*
- [3] AÏT-SAHALIA, Y. and JACOD, J. (2009). Testing for Jumps in a Discretely Observed Process, *Annals of Statistics*, 2009, 37, 184-222.
- [4] Barndorff-Nielsen, O.E. & Shephard, N. (2004) Econometric analysis of realized covariation: high frequency based covariance, regression and correlation in financial economics. *Econometrica*, (2004), 72, 885-925
- [5] Barndorff-Nielsen, O.E. & Shephard, N. (2004b) Power and bipower variation with stochastic volatility and jumps, (with discussion). *Journal of Financial Econometrics*, 2, 1-48
- [6] Barndorff-Nielsen, O.E. & Shephard, N. (2004c) Measuring the impact of jumps in multivariate price processes using bipower covariation, Working paper, Nuffield College, Oxford University
- [7] Barndorff-Nielsen, O.E. & Shephard, N. (2007) Variation, jumps and high frequency data in financial econometrics, in *Advanced in Economics and Econometrics. Theory and Applications*, Ninth World Congress eds Richard Blundell, Persson Torsten, Whitney K Newey, Econometric Society Monographs, Cambridge University Press

- [8] Barndorff-Nielsen, O.E., Graversen, S.E., Jacod, J., Podolskij, M. & Shephard, N. (2006) A central limit theorem for realised power and bipower variation of continuous semimartingales. Technical report, 2004. to appear in *From Stochastic Analysis to Mathematical Finance*, Festschrift for Albert Shiryaev, Springer.
- [9] Bollerslev, T., Law, T. H., Tauchen, G. (2008) Risk, jumps and diversification, *Journal of Econometrics*, Volume 144, Issue 1, May 2008, Pages 234-256
- [10] Boudt, K., Croux, C., Laurent, S. (2009): Outlyingness weighted quadratic covariation, working paper available on ssrn.com
- [11] Carr, P., Geman, H., Madan, D., Yor, M. (2002) The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75, 2, 305-332.
- [12] Cont, R. & Tankov, P. (2004) *Financial Modelling with jump processes*. Chapman and Hall-CRC
- [13] Corsi, F., Pirino, D., Renò, R., Threshold Bipower variation and the impact of jumps on volatility forecasting, ssrn.com
- [14] Egloff, D., Leippold, M. & Vanini, P. (2007), A simple model of credit contagion, *Journal of Banking and Finance*, 2007, (31), 2475-2492
- [15] Hayashi, T. & Kusuoka, S. (2008) Consistent estimation of covariation under nonsynchronicity, *Stat. Infer. Stoch. Process.* **11**, 93-106
- [16] Hayashi, T., Yoshida, N. (2005) On covariance estimation of non-synchronously observed diffusion processes, *Bernoulli* 11(2): 359-379

- [17] Huang, X. & Tauchen, G. (2005) The Relative Contribution of Jumps to Total Price Variance, *Journal of Financial Econometrics* 3(4):456-499
- [18] Jacod, J. (2007) *Statistics and high-frequency data*, lecture notes of SEMSTAT course in La Manga.
- [19] Jacod, J. (2008) Asymptotic properties of realized power variations and related functionals of semimartingales, *Stoch. Proc. Appl.*, 118, pp.517-559
- [20] Jacod, J. & Protter, P. (1998) Asymptotic error distributions for the Euler method for stochastic differential equations. *The Annals of Probability* **26**, 267-307
- [21] Jacod, J. & Shiryaev, A.N. (2003) *Limit theorems for stochastic processes*, 2nd edition, Springer-Verlag, Berlin
- [22] Jacod, J. & Torodov, V. (2009) Testing for common arrivals of jumps for discretely observed multidimensional processes, *Annals of Statistics*, 37, n.4, 1792-1838
- [23] Karatzas, I. & Shreve, S.E. (1999) *Brownian motion and stochastic calculus*, Springer
- [24] Mancini, C. (2001) Disentangling the jumps of the diffusion in a geometric jumping Brownian motion, *Giornale dell'Istituto Italiano degli Attuari*, Volume LXIV, Roma, 19-47
- [25] Mancini, C. (2004) Estimation of the parameters of jump of a general Poisson diffusion model. *Scandinavian Actuarial Journal*, **1**, 42-52

- [26] Mancini, C. (2009) Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps, *Scandinavian Journal of Statistics*, 36, 270-296
- [27] Mancini, C. & Cont, R. (2008) *Nonparametric tests for analyzing the fine structure of price fluctuations*, submitted paper, previous version available on ssrn.com
- [28] Mancini, C. & Renò, R. (2006) Threshold estimation of ump-diffusion models and interest rate modeling, forthcoming on *Journal of Econometrics*
- [29] Mancini, C. (2010) *Central limit theorem for integrated covariation estimate in bidimensional asset prices with stochastic volatility and infinite activity Lévy jumps*, working paper
- [30] Metivier, M. (1982) *Semimartingales: a course on stochastic processes*, De Gruyter
- [31] Renò, R. (2003) A closer look at the Epps effect, *International Journal of theoretical and applied finance* 6, 87-102
- [32] Revuz, D., Yor, M. (2001) *Continuous martingales and Brownian Motion*, Springer
- [33] Todorov, V., Bollerslev, T. (2008) Jumps and betas: a new framework for disentangling and estimating systematic risks, ssrn.com
- [34] Todorov, V., Bollerslev, T. (2009) Tails, fear and risk premia, working paper

- [35] Vetter, M. (2009): Limit theorems for bipower variation of semimartingales, working paper (available on <http://www.ruhr-uni-bochum.de/mathematik3/en/team/vetter.html>)
- [36] Woerner, J. (2006) Power and Multipower variation: inference for high frequency data, in *Stochastic Finance*, eds A.N. Shiryaev, M. do Rosário Grossinho, P. Oliveira, M. Esquivel, Springer, 343-364

7 Appendix

Proof of Lemma 2.1. A stochastic integral $\int_0^t \eta_s^{(p)} dW_s^{(p)}$, $p = 1, 3$, is a time changed Brownian motion as soon as the integrand process $\eta^{(p)}$ is cadlag with $\int_0^{+\infty} (\eta_s^{(p)})^2 ds = \infty$ ([32], Theorems 1.9 and 1.10 of ch.5), i.e. $\int_{]t_{j-1}, t_j]} \eta_s^{(p)} dW_s^{(p)} = B_{IV_{t_j}^{(p)}}^{(p)} - B_{IV_{t_{j-1}}^{(p)}}^{(p)}$, where $B^{(p)}$ are Brownian motions and $IV_t^{(p)}$ the integrated variances $\int_0^t (\eta_s^{(p)})^2 ds$ up to time t . Note that the increments of the drift part of each $X^{(q)}$ tend to zero more quickly than $\sqrt{2h \log \frac{1}{h}}$ as $h \rightarrow 0$, so for $D^{(q)}$ we can reach a result similar to (2), as soon as the boundedness of the paths of $a^{(q)}$, $\sigma^{(q)}$ and ρ is guaranteed (which is the case when they are càdlàg). In fact (as in [26])

$$\begin{aligned} \sup_{j=1..n} \frac{|\int_{t_{j-1}}^{t_j} a_s^{(1)} ds + \int_{t_{j-1}}^{t_j} \sigma_s^{(1)} dW_s^{(1)}|}{\sqrt{2h \log \frac{1}{h}}} &\leq \sup_j \frac{|\int_{t_{j-1}}^{t_j} a_s^{(1)} ds|}{\sqrt{2h \log \frac{1}{h}}} + \sup_j \frac{|\int_{t_{j-1}}^{t_j} \sigma_s^{(1)} dW_s^{(1)}|}{\sqrt{2h \log \frac{1}{h}}} \leq \\ &C_1(\omega) \sqrt{\frac{h}{\log \frac{1}{h}}} + \\ + \sup_j \frac{|B_{IV_{t_j}^{(1)}}^{(1)} - B_{IV_{t_{j-1}}^{(1)}}^{(1)}|}{\sqrt{2\Delta_j IV^{(1)} \log \frac{1}{\Delta_j IV^{(1)}}}} &\sup_j \frac{\sqrt{2\Delta_j IV^{(1)} \log \frac{1}{\Delta_j IV^{(1)}}}}{\sqrt{2M_1(\omega)h \log \frac{1}{M_1(\omega)h}}} \sup_j \frac{\sqrt{2M_1(\omega) \log \frac{1}{M_1(\omega)h}}}{\sqrt{2 \log \frac{1}{h}}}, \end{aligned} \tag{9}$$

where $C_1(\omega) := \sup_{s \in [0, T]} |a_s^{(1)}(\omega)|$, $M_1(\omega) := \sup_{s \in [0, T]} (\sigma_s^{(1)})^2(\omega)$. By [23] (Theorem 9.25) and the monotonicity of the function $x \ln \frac{1}{x}$ it follows that as $h \rightarrow 0$, the right hand side has a limsup which is bounded by $\sqrt{M_1(\omega)}$. As for the increments of $\int_0^T \sigma_t^{(2)} dW_t^{(2)} = \int_0^T \sigma_t^{(2)} \rho_t dW_t^{(1)} + \int_0^T \sigma_t^{(2)} \sqrt{1 - \rho_t^2} dW_t^{(3)}$, proceeding analogously as above for each term, we reach that

$$\limsup_h \sup_{j=1..n} \frac{|\int_{t_{j-1}}^{t_j} a_s^{(2)} ds + \int_{t_{j-1}}^{t_j} \sigma_s^{(2)} dW_s^{(2)}|}{\sqrt{2h \log \frac{1}{h}}} \leq \sqrt{2M_2'(\omega) + 2M_2''(\omega)} := \sqrt{2M_2},$$

where $M_2'(\omega) := \sup_{s \in [0, T]} ((\sigma_s^{(2)})^2 \rho_s^2)(\omega)$, $M_2''(\omega) := \sup_{s \in [0, T]} ((\sigma_s^{(2)})^2 (1 - \rho_s^2))(\omega)$.

The thesis follows. \square

The following theorem is the key result, in the finite jump activity case, validating the idea that if $(\Delta_j X^{(q)})^2$ is larger than r_h then some jumps occurred in $]t_{j-1}, t_j]$ (and vice-versa). It is based on lemma 2.1 and we state it in the general case of not equally spaced observations. This version is used for instance in the proof of Theorem 3.13.

Theorem 7.1. (*[26], FA jumps*) *Under the assumptions from **A1** to **A4**, given an arbitrary partition $\{t_0 = 0, t_1, \dots, t_n = T\}$ of $[0, T]$, then a.s. for sufficiently small, but strictly positive, $h := \sup_{j=1..n} |t_j - t_{j-1}|$ (depending on ω) we have*

$$1_{\{(\Delta_j X^{(q)})^2 \leq r_h\}} = 1_{\{\Delta_j N^{(q)} = 0\}}, \quad j = 1, 2, \dots, n, \quad q = 1, 2. \quad \square$$

Proof of Theorem 3.4 [Joint CLT, FA jumps] By Theorem 7.1 a.s. for sufficiently small h we have, for all $t \in [0, T]$,

$$h^{-1/2} \left[\sum_{t_j \leq t} \Delta_j X_\star^{(1)} \Delta_j X_\star^{(2)} - IC_t \right] = h^{-1/2} \left[\sum_{t_j \leq t} \Delta_j X^{(1)} \Delta_j X^{(2)} 1_{\{\Delta_j N^{(1)} = 0, \Delta_j N^{(2)} = 0\}} - IC_t \right]$$

$$\begin{aligned}
&= h^{-1/2} \left[\sum_{t_j \leq t} \Delta_j D^{(1)} \Delta_j D^{(2)} - \int_0^t \rho_s \sigma_s^{(1)} \sigma_s^{(2)} ds \right] - h^{-1/2} \sum_{t_j \leq t} \Delta_j D^{(1)} \Delta_j D^{(2)} 1_{\{\Delta_j N^{(2)} \neq 0\}} \\
&- h^{-1/2} \sum_{t_j \leq t} \Delta_j D^{(1)} 1_{\{\Delta_j N^{(1)} \neq 0\}} \Delta_j D^{(2)} + h^{-1/2} \sum_{t_j \leq t} \Delta_j D^{(1)} 1_{\{\Delta_j N^{(1)} \neq 0\}} \Delta_j D^{(2)} 1_{\{\Delta_j N^{(2)} \neq 0\}}.
\end{aligned}$$

Each one of last three sums tends a.s. to zero as $h \rightarrow 0$, since it contains at least one $1_{\{\Delta_j N^{(q)} \neq 0\}}$ and for any $q = 1, 2$ we have

$$\text{Plim}_{n \rightarrow \infty} \left| h^{-1/2} \sum_{t_j \leq t} \Delta_j D^{(1)} \Delta_j D^{(2)} 1_{\{\Delta_j N^{(q)} \neq 0\}} \right| \leq \text{Plim}_{n \rightarrow \infty} K_1(\omega) K_2(\omega) \sqrt{h} \log \frac{1}{h} N_T^{(q)} = 0.$$

In particular for $D^{(1)} \equiv D^{(2)}$ we have

$$\text{Plim}_{n \rightarrow \infty} h^{-1/2} \left(\sum_{t_j \leq t} (\Delta_j X^{(q)})_*^2 - IV_t^{(q)} \right) = \text{Plim}_{n \rightarrow \infty} h^{-1/2} \left(\sum_{t_j \leq t} (\Delta_j D^{(q)})^2 - IV_t^{(q)} \right)$$

Therefore

$$h^{-1/2} \begin{pmatrix} \hat{IV}_n^{(1)} - IV^{(1)} & \hat{IC}_n - IC \\ \hat{IC}_n - IC & \hat{IV}_n^{(2)} - IV^{(2)} \end{pmatrix} \quad (10)$$

has the same limit in distribution as

$$h^{-1/2} \begin{pmatrix} \sum_{t_j \leq \cdot} (\Delta_j D^{(1)})^2 - IV^{(1)} & \sum_{t_j \leq \cdot} \Delta_j D^{(1)} \Delta_j D^{(2)} - IC \\ \sum_{t_j \leq \cdot} \Delta_j D^{(1)} \Delta_j D^{(2)} - IC & \sum_{t_j \leq \cdot} (\Delta_j D^{(2)})^2 - IV^{(2)} \end{pmatrix}.$$

Note that for all t

$$\sum_{t_j \leq t} \Delta_j D^{(1)} \Delta_j D^{(2)} - IC_t = \sum_{t_j \leq t} \left(\Delta_j D^{(1)} \Delta_j D^{(2)} - \Delta_j \langle D^{(1)}, D^{(2)} \rangle \right) \quad (11)$$

and, along the lines of [7] (proof of Theorem 1, sec. 3.1), using Itô formula we know that

$$d(D^{(1)} D^{(2)}) = D_-^{(1)} dD^{(2)} + D_-^{(2)} dD^{(1)} + d \langle D^{(1)}, D^{(2)} \rangle,$$

so

$$\Delta_j (D^{(1)} D^{(2)}) = \int_{t_{j-1}}^{t_j} D_{s-}^{(1)} dD_s^{(2)} + \int_{t_{j-1}}^{t_j} D_{s-}^{(2)} dD_s^{(1)} + \Delta_j \langle D^{(1)}, D^{(2)} \rangle.$$

Therefore

$$\begin{aligned} \Delta_j D^{(1)} \Delta_j D^{(2)} &= \Delta_j (D^{(1)} D^{(2)}) - D_{t_{j-1}}^{(1)} \Delta_j D^{(2)} - D_{t_{j-1}}^{(2)} \Delta_j D^{(1)} \\ &= \int_{t_{j-1}}^{t_j} D_{s-}^{(1)} dD_s^{(2)} + \int_{t_{j-1}}^{t_j} D_{s-}^{(2)} dD_s^{(1)} + \Delta_j \langle D^{(1)}, D^{(2)} \rangle - D_{t_{j-1}}^{(1)} \Delta_j D^{(2)} - D_{t_{j-1}}^{(2)} \Delta_j D^{(1)}, \end{aligned}$$

so that (11) equals

$$\begin{aligned} \int_0^t \left(D_{s-}^{(1)} - \sum_{t_j \leq t} D_{t_{j-1}}^{(1)} \mathbf{1}_{\{s \in [t_{j-1}, t_j]\}} \right) dD_s^{(2)} + \int_0^t \left(D_{s-}^{(2)} - \sum_{t_j \leq t} D_{t_{j-1}}^{(2)} \mathbf{1}_{\{s \in [t_{j-1}, t_j]\}} \right) dD_s^{(1)} \\ = A_{12,t}^{(n)} + A_{21,t}^{(n)}, \end{aligned}$$

where

$$A^{(n)} = \begin{pmatrix} \int_0^\cdot \left(D_{s-}^{(1)} - D_{\lfloor \frac{ns-1}{n} \rfloor}^{(1)} \right) dD_s^{(1)} & \int_0^\cdot \left(D_{s-}^{(1)} - D_{\lfloor \frac{ns-1}{n} \rfloor}^{(1)} \right) dD_s^{(2)} \\ \int_0^\cdot \left(D_{s-}^{(2)} - D_{\lfloor \frac{ns-1}{n} \rfloor}^{(2)} \right) dD_s^{(1)} & \int_0^\cdot \left(D_{s-}^{(2)} - D_{\lfloor \frac{ns-1}{n} \rfloor}^{(2)} \right) dD_s^{(2)} \end{pmatrix}.$$

As special cases, for each $q = 1, 2$

$$\begin{aligned} \sum_{t_j \leq t} (\Delta_j D^{(q)})^2 - IV_t^{(q)} &= \sum_{t_j \leq t} \left((\Delta_j D^{(q)})^2 - \Delta_j \langle D^{(q)}, D^{(q)} \rangle \right) = \\ &= 2 \int_0^t \left(D_{s-}^{(q)} - \sum_{j=1}^n D_{t_{j-1}}^{(q)} \mathbf{1}_{\{s \in [t_{j-1}, t_j]\}} \right) dD_s^{(q)} = 2A_{qq,t}^{(n)}. \end{aligned}$$

By Theorem 5.5 in [20] we have that

$$h^{-1/2} A^{(n)} \xrightarrow{st} \frac{\mathbf{Z}}{\sqrt{2}},$$

with \mathbf{Z} as in (3). It follows that, as $n \rightarrow \infty$, (10) converges stably in law to

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 2Z_{11} & Z_{12} + Z_{21} \\ Z_{12} + Z_{21} & 2Z_{22} \end{pmatrix}. \quad \square$$

Proof of Corollary 3.6 [Speed of convergence of β s and ρ , FA jumps]

For all $t \in [0, T]$ we have

$$h^{-1/2} \left(\hat{\beta}_{t,n}^{(i,j)} - \frac{IC_t}{IV_t^{(j)}} \right) = h^{-1/2} \frac{\hat{IC}_{t,n} - IC_t}{\hat{IV}_{t,n}^{(j)}} + h^{-1/2} IC_t \frac{IV_t^{(j)} - \hat{IV}_{t,n}^{(j)}}{\hat{IV}_{t,n}^{(j)} IV_t^{(j)}},$$

therefore

$$h^{-1/2} \left(\hat{\beta}_n^{(i,j)} - \frac{IC}{IV^{(j)}} \right) \xrightarrow{st} \frac{Z_{12} + Z_{21}}{\sqrt{2} IV^{(j)}} - IC \frac{\sqrt{2} Z_{jj}}{(IV^{(j)})^2}.$$

As for $\hat{\rho}_n^{(1,2)}$, note preliminarily that Theorem 3.4 implies that $h^{-1/2} \left(\sqrt{IV^{(j)}} - \sqrt{\hat{IV}_n^{(j)}} \right)$ converges stably, since, t by t ,

$$h^{-1/2} \left(\sqrt{IV^{(j)}} - \sqrt{\hat{IV}_n^{(j)}} \right) = \frac{h^{-1/2} \left(IV^{(j)} - \hat{IV}_n^{(j)} \right)}{\sqrt{IV^{(j)}} + \sqrt{\hat{IV}_n^{(j)}}} \xrightarrow{st} -\frac{Z_{jj}}{\sqrt{2} IV^{(j)}}.$$

As a consequence

$$\begin{aligned} & h^{-1/2} \left(\hat{\rho}_n^{(1,2)} - \frac{IC}{\sqrt{IV^{(1)} IV^{(2)}}} \right) = \\ & h^{-1/2} \frac{\hat{IC}_n - IC}{\sqrt{\hat{IV}_n^{(1)} \hat{IV}_n^{(2)}}} + h^{-1/2} IC \left(\frac{1}{\sqrt{\hat{IV}_n^{(1)} \hat{IV}_n^{(2)}}} - \frac{1}{\sqrt{IV^{(1)} IV^{(2)}}} \right). \end{aligned}$$

The first term converges stably to $\frac{Z_{12} + Z_{21}}{\sqrt{2} IV^{(1)} IV^{(2)}}$, while the second term equals

$$\begin{aligned} & \frac{h^{-1/2} IC}{\sqrt{\hat{IV}_n^{(1)}}} \left(\frac{1}{\sqrt{\hat{IV}_n^{(2)}}} - \frac{1}{\sqrt{IV^{(2)}}} \right) + \frac{h^{-1/2} IC}{\sqrt{\hat{IV}_n^{(1)} IV^{(2)}}} \left(1 - \frac{\sqrt{\hat{IV}_n^{(1)}}}{\sqrt{IV^{(1)}}} \right) = \\ & \frac{h^{-1/2} IC}{\sqrt{\hat{IV}_n^{(1)}}} \left(\frac{\sqrt{IV^{(2)}} - \sqrt{\hat{IV}_n^{(2)}}}{\sqrt{\hat{IV}_n^{(2)} IV^{(2)}}} \right) + \frac{h^{-1/2} IC}{\sqrt{\hat{IV}_n^{(1)} IV^{(2)} IV^{(1)}}} \left(\sqrt{IV^{(1)}} - \sqrt{\hat{IV}_n^{(1)}} \right) \\ & \xrightarrow{st} -\frac{Z_{22} IC}{\sqrt{2} IV^{(1)} (IV^{(2)})^{3/2}} - \frac{Z_{11} IC}{\sqrt{2} IV^{(2)} (IV^{(1)})^{3/2}}. \end{aligned}$$

□

Proof of Proposition 3.7 [Estimate of the standard error for \hat{IC}_n ,

FA jumps] For $t = T$ it is sufficient to show that as $n \rightarrow \infty$

$$\tilde{v}_{2,2}^{(n)}(X^{(1)}, X^{(2)})_T \xrightarrow{P} \int_0^T (2\rho_t^2 + 1)(\sigma_t^{(1)})^2 (\sigma_t^{(2)})^2 dt,$$

and

$$\tilde{w}^{(n)}(X^{(1)}, X^{(2)})_T \xrightarrow{P} \int_0^T \rho_t^2 (\sigma_t^{(1)})^2 (\sigma_t^{(2)})^2 dt.$$

For $t < T$ the proof is analogous with $\sum_{j=1}^n$ replaced by $\sum_{j:t_j \leq t}$. By Theorem 7.1

we can write

$$\begin{aligned}
\text{Plim}_{n \rightarrow \infty} \tilde{v}_{2,2}^{(n)}(X^{(1)}, X^{(2)})_T &= \text{Plim}_{n \rightarrow \infty} h^{-1} \sum_{j=1}^n (\Delta_j D^{(1)})^2 1_{\{\Delta_j N^{(1)}=0\}} (\Delta_j D^{(2)})^2 1_{\{\Delta_j N^{(2)}=0\}} \\
&= \text{Plim}_{n \rightarrow \infty} v_{2,2}^{(n)}(D^{(1)}, D^{(2)})_T - \text{Plim}_{n \rightarrow \infty} h^{-1} \sum_{j=1}^n (\Delta_j D^{(1)})^2 (\Delta_j D^{(2)})^2 1_{\{\Delta_j N^{(1)} \neq 0\}} \\
&\quad - \text{Plim}_{n \rightarrow \infty} h^{-1} \sum_{j=1}^n (\Delta_j D^{(1)})^2 (\Delta_j D^{(2)})^2 1_{\{\Delta_j N^{(2)} \neq 0\}} \\
&\quad + \text{Plim}_{n \rightarrow \infty} h^{-1} \sum_{j=1}^n (\Delta_j D^{(1)})^2 1_{\{\Delta_j N^{(1)} \neq 0\}} (\Delta_j D^{(2)})^2 1_{\{\Delta_j N^{(2)} \neq 0\}}.
\end{aligned}$$

By Theorem 2.1 in [8],

$$\text{Plim}_{n \rightarrow \infty} v_{2,2}^{(n)}(D^{(1)}, D^{(2)})_T = \int_0^T (2\rho_t^2 + 1)(\sigma_t^{(1)})^2 (\sigma_t^{(2)})^2 dt,$$

whereas the other terms are all zero. In fact for any $q = 1, 2$

$$\text{Plim}_{n \rightarrow \infty} h^{-1} \sum_{j=1}^n (\Delta_j D^{(1)})^2 (\Delta_j D^{(2)})^2 1_{\{\Delta_j N^{(q)} \neq 0\}} \leq \text{Plim}_{n \rightarrow \infty} K_1^2(\omega) K_2^2(\omega) h \left(\log \frac{1}{h} \right)^2 N_T^{(q)} = 0. \tag{12}$$

Now we deal with $\tilde{w}^{(n)}(X^{(1)}, X^{(2)})_T$. Analogously as before

$$\begin{aligned}
&\text{Plim}_{n \rightarrow \infty} \tilde{w}^{(n)}(X^{(1)}, X^{(2)})_T \\
&= \text{Plim}_{n \rightarrow \infty} h^{-1} \sum_{j=1}^{n-1} \left[\prod_{i=0}^1 \Delta_{j+i} D^{(1)} (1 - 1_{\{\Delta_{j+i} N^{(1)} \neq 0\}}) \prod_{i=0}^1 \Delta_{j+i} D^{(2)} (1 - 1_{\{\Delta_{j+i} N^{(2)} \neq 0\}}) \right],
\end{aligned}$$

which coincides with the sum of $\text{Plim}_{n \rightarrow \infty} w^{(n)}(D^{(1)}, D^{(2)})_T$ with a finite number of terms

which are shown to be negligible. By Theorem 2.1 in [8], $\text{Plim}_{n \rightarrow \infty} w^{(n)}(D^{(1)}, D^{(2)})_T =$

$\int_0^T \rho_t^2 (\sigma_t^{(1)})^2 (\sigma_t^{(2)})^2 dt$, while the other terms are given by the product of $\prod_{i=0}^1 \Delta_{j+i} D^{(1)}$

$\prod_{i=0}^1 \Delta_{j+i} D^{(2)}$ with at least one of the indicators $1_{\{\Delta_{j+s} N^{(q)} \neq 0\}}$, for an $s \in \{0, 1\}$.

Therefore the limit in probability of each such term is zero as in (12). \square

Proof of Corollary 3.8 [CLT for the standardized version of $\hat{IC}_{t,n} - IC_t$, FA jumps]. By Theorem 3.4 we have

$$h^{-1/2} \left(\hat{IC}_n - IC \right) \xrightarrow{st} \frac{1}{\sqrt{2}} \left(\int_0^\cdot 2\sigma_s^{(1)}\sigma_s^{(2)}\rho_s dB_{11s} + \int_0^\cdot \sigma_s^{(1)}\sigma_s^{(2)}\sqrt{1-\rho_s^2} \left[dB_{12s} + dB_{21s} \right] \right).$$

The variance of the last term at time t is $\int_0^t (1 + \rho_s^2)(\sigma_s^{(1)})^2(\sigma_s^{(2)})^2 ds$. By Proposition 3.7 we then obtain that

$$\frac{\hat{IC}_{t,n} - IC_t}{\sqrt{h}\sqrt{\tilde{v}_{2,2}^{(n)}(X^{(1)}, X^{(2)})_t - \tilde{w}^{(n)}(X^{(1)}, X^{(2)})_t}} \xrightarrow{st} \mathcal{N},$$

where \mathcal{N} is a standard Gaussian r.v.. □

Remark 7.2. [Efficiency comparison between the Bipower and threshold covariation estimators]

Let us consider $T = 1$ (one day) and constant coefficients ρ , $\sigma^{(1)}$ and $\sigma^{(2)}$. We have $AVar_{RTC} = (1 + \rho^2)(\sigma^{(1)})^2(\sigma^{(2)})^2$, both in absence and in presence of finite variation jumps, while $AVar_{RBPC}$ is a more complicated function of $\rho^2, \sigma^{(1)}, \sigma^{(2)}$ and it has to be considered only in absence of jumps. Using that the coefficients $\sigma_1^2, \sigma_2^2, \sigma_{1,2}$ in [6] are respectively $(\sigma^{(1)})^2, (\sigma^{(2)})^2, \rho\sigma^{(1)}\sigma^{(2)}$ here, we reach

$$AVar_{RBPC}(\rho^2, \sigma^{(1)}, \sigma^{(2)}) = \mu_1^4 \left\{ 2\vartheta(1) \left[(\sigma^{(1)})^4 + (\sigma^{(2)})^4 + 2(\sigma^{(1)})^2(\sigma^{(2)})^2 + 4\rho^2(\sigma^{(1)})^2(\sigma^{(2)})^2 \right] - 2\vartheta(\rho^{+,-}) \left[(\sigma^{(1)})^4 + (\sigma^{(2)})^4 + 2(\sigma^{(1)})^2(\sigma^{(2)})^2 - 4\rho^2(\sigma^{(1)})^2(\sigma^{(2)})^2 \right] \right\},$$

where $\mu_1 = \sqrt{2/\pi}$,

$$\rho^{+,-} = \frac{\sigma^{(1)} - \sigma^{(2)}}{\sqrt{(\sigma^{(1)})^4 + (\sigma^{(2)})^4 + 2(\sigma^{(1)})^2(\sigma^{(2)})^2 - 4\rho^2(\sigma^{(1)})^2(\sigma^{(2)})^2}} \in [-1, 1],$$

and $\vartheta : [-1, 1] \rightarrow [0, 2.609]$ is an even function (specified in [6]) which is increasing in $[0, 1]$ and with $\vartheta(0) = 0, \vartheta(1) = 2.609$. We checked that

$$AVar_{RTC}(\rho^2, \sigma^{(1)}, \sigma^{(2)}) \leq AVar_{RBPC}(\rho^2, \sigma^{(1)}, \sigma^{(2)}), \quad \forall \rho \in [-1, 1] \quad (13)$$

in the following cases:

- if $(\sigma^{(1)})^2 = (\sigma^{(2)})^2$, in which case $\vartheta(\rho^{+,-}) = 0$, and

$$AVar_{RBPC}(\rho^2, \sigma^{(1)}, \sigma^{(2)}) = 8.4594(1 + \rho^2)(\sigma^{(1)})^4;$$

when $(\sigma^{(1)})^2 \neq (\sigma^{(2)})^2$ in the following cases:

- if $\vartheta(\rho^{+,-}) = 1$ or $\vartheta(\rho^{+,-}) = 2.609$;
- if $\sigma^{(1)} = 0.0981$ and $\sigma^{(2)} = 0.3815$, the constant parameters values of Model 2 in Table A: in the web appendix we plotted $AVar_{RBPC}(\rho^2, 0.0981, 0.3815)$ and $AVar_{RTC}(\rho^2, 0.0981, 0.3815)$ as functions of ρ^2 and we saw that (13) still holds for all values of $\rho^2 \in [0, 1]$.

Note that the function $AVar_{RBPC}(\rho^2, \sigma^{(1)}, \sigma^{(2)})$ keeps the same value if we exchange the roles of $\sigma^{(1)}$ and $\sigma^{(2)}$, so we verified that

- analogous plots turn out when taking all combinations of constants $(\sigma^{(1)}, \sigma^{(2)})$, with $\sigma^{(1)} > \sigma^{(2)}$ and each $\sigma^{(q)}$ ranging from 0.01 to 0.5 with step 0.01.

Proof of Theorem 3.13 [Asynchronous observations] Note that we can assume that the Brownian semimartingale parts $D^{(q)}$ of $X^{(q)}$ are bounded on $[0, T]$ ([19]), so they belong to \mathbb{L}^8 . Using Theorem 7.1 in the not evenly-spaced observations case with $h := \sup_{j=1..m^{(n)}}(\tau_j - \tau_{j-1}) \vee \sup_{i=1..k^{(n)}}(\nu_i - \nu_{i-1})$, a.s. for sufficiently small h we can write

$$\begin{aligned} & \sum_{j=1..m^{(n)}, i=1..k^{(n)}} (X_{\tau_j}^{(1)} - X_{\tau_{j-1}}^{(1)})_{\star} (X_{\nu_i}^{(2)} - X_{\nu_{i-1}}^{(2)})_{\star} 1_{\{\tau_{j-1}, \tau_j\} \cap \{\nu_{i-1}, \nu_i\} \neq \emptyset\}} = \\ & \sum_{j=1..m^{(n)}, i=1..k^{(n)}} (D_{\tau_j}^{(1)} - D_{\tau_{j-1}}^{(1)}) 1_{\{N_{\tau_j}^{(1)} - N_{\tau_{j-1}}^{(1)} = 0\}} (D_{\nu_i}^{(2)} - D_{\nu_{i-1}}^{(2)}) 1_{\{N_{\nu_i}^{(2)} - N_{\nu_{i-1}}^{(2)} = 0\}} 1_{\{\tau_{j-1}, \tau_j\} \cap \{\nu_{i-1}, \nu_i\} \neq \emptyset\}} \\ & = \sum_{j=1..m^{(n)}, i=1..k^{(n)}} (D_{\tau_j}^{(1)} - D_{\tau_{j-1}}^{(1)}) (D_{\nu_i}^{(2)} - D_{\nu_{i-1}}^{(2)}) 1_{\{\tau_{j-1}, \tau_j\} \cap \{\nu_{i-1}, \nu_i\} \neq \emptyset\}} \end{aligned}$$

$$\begin{aligned}
& - \sum_{j=1..m^{(n)}, i=1..k^{(n)}} (D_{\tau_j}^{(1)} - D_{\tau_{j-1}}^{(1)})(D_{\nu_i}^{(2)} - D_{\nu_{i-1}}^{(2)}) \left[1_{\{N_{\tau_j}^{(1)} - N_{\tau_{j-1}}^{(1)} \neq 0\}} + 1_{\{N_{\nu_i}^{(2)} - N_{\nu_{i-1}}^{(2)} \neq 0\}} \right. \\
& \quad \left. - 1_{\{N_{\tau_j}^{(1)} - N_{\tau_{j-1}}^{(1)} \neq 0, N_{\nu_i}^{(2)} - N_{\nu_{i-1}}^{(2)} \neq 0\}} \right] 1_{\{[\tau_{j-1}, \tau_j] \cap [\nu_{i-1}, \nu_i] \neq \emptyset\}},
\end{aligned}$$

The first sum of the r.h.s. tends to IC_T in probability by Corollary 2.2 in ([15]), with $f \equiv g \equiv 1$, while each sum in the second term is dominated in absolute value, for a suitable q , by

$$\sup_j |D_{\tau_j}^{(1)} - D_{\tau_{j-1}}^{(1)}| \sup_i |D_{\nu_i}^{(2)} - D_{\nu_{i-1}}^{(2)}| N_T^{(q)},$$

which tends a.s. to zero as $h \rightarrow 0$, by Lemma 2.1. \square

The following facts are used within the proof of Theorem 4.2.

Without loss of generality (as in [19], Lemma 4.6) we can assume that

A5. $\int_{x \in \mathbb{R}} 1 \wedge (\gamma^{(q)})^2(\omega, t, x) dx$ is bounded.

Lemma 7.3. *For each $q = 1, 2$ we have the following.*

1. *If processes a and σ are càdlàg then, under **A3**, a.s., for small h , $1_{\{(\Delta_j D^{(q)})^2 > r_h\}} = 0$, uniformly in j .*
2. *Under **A5** we have that, for each $j = 1, \dots, n$, $E[(\Delta_j \tilde{J}_2^{(q)})^2] \leq Kh$, for a positive constant K .*

Proof. Part 1. is a consequence of Lemma 2.1.

Part 2.

$$E[(\Delta_j \tilde{J}_2^{(q)})^2] = E\left[\int_{t_{j-1}}^{t_j} \int_{|\gamma^{(q)}| \leq 1} (\gamma^{(q)})^2 dx ds\right]:$$

since, by assumption **A4'**, $\int_{|\gamma^{(q)}| \leq 1} (\gamma^{(q)})^2 dx$ is bounded, the last term above is dominated by Kh for some positive constant K . \square

The following lemma generalizes analogous results given in [26] from the framework of Lévy jumps to the one of Itô semimartingale jumps.

Lemma 7.4. *The following facts hold.*

1. *Let us consider any sequence π_n of partitions $\{0, t_1, \dots, t_n = T\}$ of $[0, T]$, $n \in \mathbb{N}$, such that $\max_{j=1..n} |t_j - t_{j-1}| \rightarrow 0$ as $n \rightarrow \infty$. For each $q = 1, 2$, as long as $\tilde{J}_2^{(q)}$ is a semimartingale, we can find a subsequence n_k for which a.s., for any $\delta > 0$ there exists a sufficiently large k such that for all $j = 1, \dots, n_k$ on $\{(\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_{h_k}\}$ we have*

$$(\Delta_{2,s} \tilde{J}_2^{(q)})^2 \leq 4r_{h_k} + \delta, \quad \forall s \in]t_{j-1}, t_j].$$

2. *Under **A3** and **A4'**, for each $q = 1, 2$, we have $\sum_{j=1}^n P\{\Delta_j N^{(q)} \neq 0, (\Delta_j \tilde{J}_2^{(q)})^2 > 4r_h\} \rightarrow 0$ as $h \rightarrow 0$.*

Proof. *Statement 1* is a consequence of the fact that ([30], Theorem 25.1) there is a subsequence n_k such that, defined $h_k = T/n_k$, $\sum_{j=1}^{\lfloor t/h_k \rfloor} (\Delta_j \tilde{J}_2^{(q)})^2$ tends to $\sum_{s \in [0,t]} (\Delta \tilde{J}_{2,s}^{(q)})^2$ a.s. uniformly w.r.t. $t \in [0, T]$, as $k \rightarrow \infty$, where $[x]$ denotes the integer part of x . Since a.s.

$$\begin{aligned} & \sup_{j=1..n_k} |(\Delta_j \tilde{J}_2^{(q)})^2 - \sum_{s \in]t_{j-1}, t_j]} (\Delta \tilde{J}_{2,s}^{(q)})^2| \\ = & \sup_{j=1..n_k} \left\{ \left[\sum_{\ell=1}^{\lfloor t_j/h_k \rfloor} (\Delta_\ell \tilde{J}_2^{(q)})^2 - \sum_{s \in [0, t_j]} (\Delta \tilde{J}_{2,s}^{(q)})^2 \right] - \left[\sum_{\ell=1}^{\lfloor t_{j-1}/h_k \rfloor} (\Delta_\ell \tilde{J}_2^{(q)})^2 - \sum_{s \in [0, t_{j-1}]} (\Delta \tilde{J}_{2,s}^{(q)})^2 \right] \right\} \\ & \leq 2 \sup_{t \in [0, T]} \left[\sum_{j=1}^{\lfloor t/h_k \rfloor} (\Delta_j \tilde{J}_2^{(q)})^2 - \sum_{s \in [0, t]} (\Delta \tilde{J}_{2,s}^{(q)})^2 \right] \rightarrow 0, \end{aligned}$$

we in fact have that a.s. for all $j = 1, \dots, n_k$ each squared increment $(\Delta_j \tilde{J}_2^{(q)})^2$ is, uniformly on j , arbitrarily close to $\sum_{s \in]t_{j-1}, t_j]} (\Delta \tilde{J}_{2,s}^{(q)})^2$. More precisely, a.s. for all $\delta > 0$ we can find a sufficiently large k such that

$$\sup_{j=1..n_k} \left| (\Delta_j \tilde{J}_2^{(q)})^2 - \sum_{s \in]t_{j-1}, t_j]} (\Delta \tilde{J}_{2,s}^{(q)})^2 \right| < \delta,$$

so, for all j such that $(\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_h$ we have

$$\sum_{s \in]t_{j-1}, t_j]} (\Delta \tilde{J}_{2,s}^{(q)})^2 \leq \sup_{j=1..n_k} \left| (\Delta_j \tilde{J}_2^{(q)})^2 - \sum_{s \in]t_{j-1}, t_j]} (\Delta \tilde{J}_{2,s}^{(q)})^2 \right| + (\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_h + \delta.$$

In particular for any $s \in]t_{j-1}, t_j]$ with j such that $(\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_h$, each squared jump size $(\Delta \tilde{J}_{2,s}^{(q)})^2$ is bounded by $4r_h + \delta$.

Statement 2. The predictable compensator of $N_{2,t}^{(q)} := \sum_{s \leq t} 1_{|\Delta J_s^{(q)}| > 1}$ is $\int_0^t \int_{|\gamma^{(q)}(\omega, s, x)| > 1} \underline{\mu}(dx, ds)$ and the predictable quadratic variation of $\tilde{J}_2^{(q)}$ are of the form $\Lambda_t^{(q)} = \int_0^t \lambda_{2,s}^{(q)} ds$ and $\Lambda'_t{}^{(q)} = \int_0^t \lambda'_{2,s}{}^{(q)} ds$ respectively. Assumption **A4**' guarantees that both $\lambda_{2,s}^{(q)} = \int_{|\gamma^{(q)}| > 1} 1 dx$ and $\lambda'_{2,s}{}^{(q)} = \int_{|\gamma^{(q)}| \leq 1} (\gamma^{(q)})^2 dx$ are bounded, and therefore we can use exactly the same argument as in [2], eq. (62), with $\delta = 1$ and $\zeta = \sqrt{12r_h}$, and replacing $M(\delta) = \int \int_{|x| \leq 1} x \tilde{\mu}(dx, dt)$ with our $\tilde{J}_2^{(q)}$. We conclude that

$$\sum_{j=1}^n P\{\Delta_j N^{(q)} \neq 0, (\Delta_j \tilde{J}_2^{(q)})^2 > 4r_h\} = O(nh \frac{h}{r_h}). \quad \square$$

For any $\delta > 0$ denote by $Z_{h_k}^{(q), \delta}$ the following pure jump plus drift semimartingales having only jumps bounded in absolute value by $\sqrt{4r_{h_k} + \delta}$, $q = 1, 2$:

$$Z_{h_k, t}^{(q), \delta} := \int_0^t \int_{|\gamma^{(q)}| \leq \sqrt{4r_{h_k} + \delta}} \gamma^{(q)} \tilde{\underline{\mu}}^{(q)}(dx, ds) - \int_0^t \int_{\sqrt{4r_{h_k} + \delta} < |\gamma^{(q)}| \leq 1} \gamma^{(q)} dx ds, \quad t \geq 0.$$

By Lemma 7.4 we have that for any $\delta > 0$, for sufficiently large k the indices j for which $(\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_{h_k}$ are such that the increment $(\Delta_j \tilde{J}_2^{(q)})^2$ coincides with the

increment $(\Delta_j Z_{h_k}^{(q),\delta})^2$ of $Z_{h_k}^{(q),\delta}$, since $(\Delta_j \tilde{J}_2^{(q)})^2$ does not contain jumps larger than $\sqrt{4r_{h_k} + \delta}$.

Lemma 7.5. *For each $q = 1, 2$*

$$\text{Plim}_{n \rightarrow \infty} \sum_{j=1}^n (\Delta_j \tilde{J}_2^{(q)})^2 1_{\{(\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_h\}} = 0$$

Proof Consider the sequence of partitions $\pi_n = \{0, T/n, 2T/n, \dots, T\}$. Take any subsequence π_{n_ℓ} . By Lemma 7.4, point 1, a.s. there exists a sub-subsequence n_{ℓ_k} such that for any $\delta > 0$ and k sufficiently large then for all $j = 1, \dots, n_{\ell_k}$ on $(\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_{h_{\ell_k}}$ we have $(\Delta_j \tilde{J}_2^{(q)})^2 = (\Delta_j Z_{h_{\ell_k}}^{(q),\delta})^2$. Denote

$$S_n^{(q)} := \sum_{j=1}^n (\Delta_j \tilde{J}_2^{(q)})^2 1_{\{(\Delta_j \tilde{J}_2^{(q)})^2 \leq 4r_h\}}.$$

Therefore

$$\begin{aligned} 0 &\leq \text{Plim}_{k \rightarrow \infty} S_{n_{\ell_k}}^{(q)} \leq \text{Plim}_{k \rightarrow \infty} \sum_{j=1}^{n_{\ell_k}} (\Delta_j Z_{h_{\ell_k}}^{(q),\delta})^2 \\ &= \text{Plim}_{k \rightarrow \infty} \int_0^T \int_{|\gamma^{(q)}| \leq \sqrt{4r_{h_{\ell_k}} + \delta}} (\gamma^{(q)})^2 dx ds = \int_0^T \int_{|\gamma^{(q)}| \leq \sqrt{\delta}} (\gamma^{(q)})^2 dx ds. \end{aligned}$$

Since a.s. $\int_{|\gamma^{(q)}| \leq 1} (\gamma^{(q)})^2 dx < \infty$, the last term above tends a.s. to zero as $\delta \rightarrow 0$, which implies that $\text{Plim}_{k \rightarrow \infty} S_{n_{\ell_k}}^{(q)} = 0$.

Since then from any subsequence of $S_n^{(q)}$ we can extract a sub-subsequence tending to zero in probability, we in fact have that the whole sequence $S_n^{(q)} \rightarrow 0$ in probability, as we need. \square

Proof of Theorem 4.2. We decompose $\hat{IC}_{T,n} - IC_T$ into the sum of five terms and we show that each term tends a.s. to zero, as $n \rightarrow \infty$. We need some further notation. Recall that for each $q = 1, 2$

$$D_t^{(q)} = \int_0^t a_s^{(q)} ds + \int_0^t \sigma_s^{(q)} dW_s^{(q)},$$

and denote $Y_t^{(q)} := D_t^{(q)} + J_{1t}^{(q)}$, so that we have $X_t^{(q)} = Y_t^{(q)} + \tilde{J}_{2t}^{(q)}$, $q = 1, 2$.

We now add and subtract $\sum_{j=1}^n \Delta_j Y^{(1)} 1_{\{(\Delta_j Y^{(1)})^2 \leq 9r_h\}} \Delta_j Y^{(2)} 1_{\{(\Delta_j Y^{(2)})^2 \leq 9r_h\}}$ from $\hat{I}C_{T,n} - IC_T$. The threshold $9r_h$ is needed to reach, after (15), that on $\{(\Delta_j X^{(q)})^2 \leq r_h, (\Delta_j Y^{(q)})^2 > 9r_h\}$ we have $|\Delta_j \tilde{J}_2^{(q)}| \geq 2\sqrt{r_h}$, and then that on $\{(\Delta_j X^{(q)})^2 \leq r_h, (\Delta_j \tilde{J}_2^{(q)})^2 > 4r_h\}$, through (16), we have $\Delta_j N^{(q)} \neq 0$. We reach

$$\begin{aligned}
& |\hat{I}C_{T,n} - IC_T| \\
&= \left| \sum_{j=1}^n (\Delta_j Y^{(1)} + \Delta_j \tilde{J}_2^{(1)}) 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} (\Delta_j Y^{(2)} + \Delta_j \tilde{J}_2^{(2)}) 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}} - IC_T \right| \\
&\leq \left| \sum_{j=1}^n \Delta_j Y^{(1)} 1_{\{(\Delta_j Y^{(1)})^2 \leq 9r_h\}} \Delta_j Y^{(2)} 1_{\{(\Delta_j Y^{(2)})^2 \leq 9r_h\}} - IC_T \right| \\
&+ \left| \sum_{j=1}^n \Delta_j Y^{(1)} \Delta_j Y^{(2)} \left(1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}} - 1_{\{(\Delta_j Y^{(1)})^2 \leq 9r_h\}} 1_{\{(\Delta_j Y^{(2)})^2 \leq 9r_h\}} \right) \right| \\
&\quad + \left| \sum_{j=1}^n \Delta_j Y^{(1)} \Delta_j \tilde{J}_2^{(2)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}} \right| \tag{14} \\
&\quad + \left| \sum_{j=1}^n \Delta_j \tilde{J}_2^{(1)} \Delta_j Y^{(2)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}} \right| \\
&\quad + \left| \sum_{j=1}^n \Delta_j \tilde{J}_2^{(1)} \Delta_j \tilde{J}_2^{(2)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}} \right|.
\end{aligned}$$

The first term tends to zero in probability by Corollary 3.5. The second term coincides with

$$\begin{aligned}
& \left| \sum_{j=1}^n \Delta_j Y^{(1)} \Delta_j Y^{(2)} \left[1_{\{(\Delta_j X^{(1)})^2 \leq r_h, (\Delta_j X^{(2)})^2 \leq r_h, (\Delta_j Y^{(1)})^2 > 9r_h\}} \right. \right. \\
&\quad \left. \left. + 1_{\{(\Delta_j X^{(1)})^2 \leq r_h, (\Delta_j X^{(2)})^2 \leq r_h, (\Delta_j Y^{(2)})^2 > 9r_h\}} \right. \right. \\
&\quad \left. \left. - 1_{\{(\Delta_j X^{(1)})^2 \leq r_h, (\Delta_j X^{(2)})^2 \leq r_h, (\Delta_j Y^{(1)})^2 > 9r_h, (\Delta_j Y^{(2)})^2 > 9r_h\}} \right. \right. \\
&\quad \left. \left. - 1_{\{(\Delta_j X^{(1)})^2 > r_h, (\Delta_j Y^{(1)})^2 \leq 9r_h, (\Delta_j Y^{(2)})^2 \leq 9r_h\}} - 1_{\{(\Delta_j X^{(2)})^2 > r_h, (\Delta_j Y^{(1)})^2 \leq 9r_h, (\Delta_j Y^{(2)})^2 \leq 9r_h\}} \right. \right.
\end{aligned}$$

$$+1_{\{(\Delta_j X^{(1)})^2 > r_h, (\Delta_j X^{(2)})^2 > r_h, (\Delta_j Y^{(1)})^2 \leq 9r_h, (\Delta_j Y^{(2)})^2 \leq 9r_h\}} \Big]. \quad (15)$$

All these terms tend a.s. to zero. In fact for the first three ones note that on $\{(\Delta_j X^{(a)})^2 \leq r_h, (\Delta_j Y^{(a)})^2 > 9r_h\}$ we have $\sqrt{r_h} \geq |\Delta_j X^{(a)}| \geq |\Delta_j Y^{(a)}| - |\Delta_j \tilde{J}_2^{(a)}|$ and thus $|\Delta_j \tilde{J}_2^{(a)}| \geq |\Delta_j Y^{(a)}| - \sqrt{r_h} > 3\sqrt{r_h} - \sqrt{r_h} = 2\sqrt{r_h}$. Moreover on $\{(\Delta_j X^{(a)})^2 \leq r_h, (\Delta_j \tilde{J}_2^{(a)})^2 > 4r_h\}$ we have that $\Delta_j N^{(a)} \neq 0$. Actually, since

$$2\sqrt{r_h} - |\Delta_j Y^{(a)}| < |(\Delta_j \tilde{J}_2^{(a)})^2| - |\Delta_j Y^{(a)}| \leq |\Delta_j X^{(a)}| \leq \sqrt{r_h}$$

then

$$K_q \sqrt{2h \log \frac{1}{h}} + |\Delta_j J_1^{(a)}| \geq |\Delta_j D^{(a)}| + |\Delta_j J_1^{(a)}| \geq |\Delta_j Y^{(a)}| > \sqrt{r_h}, \quad (16)$$

so

$$|\Delta_j J_1^{(a)}| > \sqrt{r_h} \left(1 - K_q \sqrt{\frac{2h \log \frac{1}{h}}{r_h}} \right) :$$

since a.s. for sufficiently small h the quantity $1 - K_q \sqrt{2h \log \frac{1}{h} / r_h}$ is positive, then in fact $|\Delta_j J_1^{(a)}| > 0$, so that $\Delta_j N^{(a)} \neq 0$.

As a consequence the probability that each one of the first three terms of (15) is non zero is dominated by $\sum_{j=1}^n P\{\Delta_j N^{(a)} \neq 0, (\Delta_j \tilde{J}_2^{(a)})^2 > 4r_h\}$ which tends to zero as $h \rightarrow 0$ by Lemma 7.4, part 2.

As for the last three terms of (15) note that on $\{(\Delta_j Y^{(a)})^2 \leq 9r_h\}$ we have a.s., for h small such that $\Delta_j N^{(a)} \in \{0, 1\}$,

$$\begin{aligned} \Delta_j N^{(a)} &\leq |\Delta_j J_1^{(a)}| = |\Delta_j Y^{(a)} - \Delta_j D^{(a)}| \leq |\Delta_j D^{(a)}| + |\Delta_j Y^{(a)}| \\ &\leq |\Delta_j D^{(a)}| + 3\sqrt{r_h} \leq K_q \sqrt{2h \log \frac{1}{h}} + 3\sqrt{r_h} \rightarrow 0, \quad q = 1, 2, \end{aligned}$$

hence, for small h on $\{(\Delta_j Y^{(a)})^2 \leq 9r_h\}$ we have $\Delta_j N^{(a)} = 0$, $j = 1, \dots, n$. Therefore $\{(\Delta_j X^{(a)})^2 > r_h, (\Delta_j Y^{(a)})^2 \leq 9r_h\} \subset \{(\Delta_j D^{(a)} + \Delta_j \tilde{J}_2^{(a)})^2 > r_h\} \subset \{(\Delta_j D^{(a)})^2 >$

$\frac{r_h}{4}\} \cup \{(\Delta_j \tilde{J}_2^{(q)})^2 > \frac{r_h}{4}\}$, $q = 1, 2$; however, by Lemma 7.3 part 1, a.s., for small h , $1_{\{(\Delta_j D^{(q)})^2 > \frac{r_h}{4}\}} = 0$, thus the last three terms of (15) are dominated by

$$\sum_{j=1}^n |\Delta_j D^{(1)} \Delta_j D^{(2)}| 1_{\{(\Delta_j \tilde{J}_2^{(q)})^2 > \frac{r_h}{4}\}}$$

for a suitable q . However this last term tends to zero in probability, since

$$\sum_{j=1}^n |\Delta_j D^{(1)} \Delta_j D^{(2)}| 1_{\{(\Delta_j \tilde{J}_2^{(q)})^2 > \frac{r_h}{4}\}} \leq K_1 K_2 2h \log \frac{1}{h} \sum_{j=1}^n 1_{\{(\Delta_j \tilde{J}_2^{(q)})^2 > \frac{r_h}{4}\}}$$

and, using lemma 7.3 part 2, $E[2h \log \frac{1}{h} \sum_{j=1}^n 1_{\{(\Delta_j \tilde{J}_2^{(q)})^2 > \frac{r_h}{4}\}}] = 2h \log \frac{1}{h} \sum_{j=1}^n P\{(\Delta_j \tilde{J}_2^{(q)})^2 > \frac{r_h}{4}\} = O(\frac{2h \log \frac{1}{h}}{r_h} nh)$.

We now show that the third and fourth terms of the right hand side of (14), which are similar, tend to zero in probability. We have

$$\begin{aligned} & \sum_{j=1}^n \Delta_j Y^{(1)} \Delta_j \tilde{J}_2^{(2)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h\}} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h\}} \\ &= \sum_{j=1}^n \Delta_j Y^{(1)} \Delta_j \tilde{J}_2^{(2)} \left[1_{\{|\Delta_j X^{(1)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(1)}| \leq 2\sqrt{r_h}\}} 1_{\{|\Delta_j X^{(2)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(2)}| \leq 2\sqrt{r_h}\}} \right. \\ & \quad + 1_{\{|\Delta_j X^{(1)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(1)}| > 2\sqrt{r_h}\}} 1_{\{|\Delta_j X^{(2)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(2)}| > 2\sqrt{r_h}\}} \\ & \quad + 1_{\{|\Delta_j X^{(1)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(1)}| > 2\sqrt{r_h}\}} 1_{\{|\Delta_j X^{(2)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(2)}| \leq 2\sqrt{r_h}\}} \\ & \quad \left. + 1_{\{|\Delta_j X^{(1)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(1)}| \leq 2\sqrt{r_h}\}} 1_{\{|\Delta_j X^{(2)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(2)}| > 2\sqrt{r_h}\}} \right]. \end{aligned} \tag{17}$$

As before on $\{(\Delta_j X^{(q)})^2 \leq r_h, (\Delta_j \tilde{J}_2^{(q)})^2 > 4r_h\}$ we have that $\Delta_j N^{(q)} \neq 0$, so, for each one of the last three terms of (17), the probability it is different from zero is dominated by

$$\sum_{j=1}^n P\{\Delta_j N^{(q)} \neq 0, (\Delta_j \tilde{J}_2^{(q)})^2 > 4r_h\} \rightarrow 0.$$

Now we show that the first term of (17) is asymptotically negligible. Note that on $\{|\Delta_j X^{(q)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(q)}| \leq 2\sqrt{r_h}\}$ a.s. for small h we have $\Delta N_j^{(q)} = 0$; in fact a.s., for small h we have $\Delta_j N^{(q)} \in \{0, 1\}$, and

$$\begin{aligned} \Delta_j N^{(q)} &\leq |\Delta_j J_1^{(q)}| = |\Delta_j X^{(q)} - \Delta_j D^{(q)} - \Delta_j \tilde{J}_2^{(q)}| \\ &\leq \sqrt{r_h} + \sup_j |\Delta_j D^{(q)}| + 2\sqrt{r_h} \rightarrow 0, \end{aligned}$$

for all $j = 1, \dots, n$, for $q = 1, 2$. So we have

$$\begin{aligned} \text{Plim}_{n \rightarrow \infty} \left| \sum_{j=1}^n \Delta_j Y^{(1)} \Delta_j \tilde{J}_2^{(2)} 1_{\{|\Delta_j X^{(1)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(1)}| \leq 2\sqrt{r_h}\}} 1_{\{|\Delta_j X^{(2)}| \leq \sqrt{r_h}, |\Delta_j \tilde{J}_2^{(2)}| \leq 2\sqrt{r_h}\}} \right| \\ \leq \text{Plim}_{n \rightarrow \infty} \sum_{j=1}^n |\Delta_j D^{(1)} \Delta_j \tilde{J}_2^{(2)}| 1_{\{|\Delta_j \tilde{J}_2^{(2)}| \leq 2\sqrt{r_h}\}}. \end{aligned}$$

By the Cauchy-Schwarz inequality, last term is dominated by

$$\begin{aligned} \text{Plim}_{n \rightarrow \infty} \sqrt{\sum_{j=1}^n (\Delta_j D^{(1)})^2} \sqrt{\sum_{j=1}^n (\Delta_j \tilde{J}_2^{(2)})^2 1_{\{|\Delta_j \tilde{J}_2^{(2)}| \leq 2\sqrt{r_h}\}}} \\ \leq \sqrt{\int_0^T (\sigma_s^{(1)})^2 ds} \text{Plim}_{n \rightarrow \infty} \sqrt{S_n^{(2)}} = 0, \end{aligned}$$

by Lemma 7.5.

It remains to consider the last term of (14), which is rewritten as in (17) with $\Delta_j \tilde{J}_2^{(1)} \Delta_j \tilde{J}_2^{(2)}$ in place of $\Delta_j Y^{(1)} \Delta_j \tilde{J}_2^{(2)}$, so that last three terms converge to zero in probability as before. As for the first term

$$\text{Plim}_{n \rightarrow \infty} \sum_{j=1}^n \Delta_j \tilde{J}_2^{(1)} \Delta_j \tilde{J}_2^{(2)} 1_{\{(\Delta_j X^{(1)})^2 \leq r_h, (\Delta_j \tilde{J}_2^{(1)})^2 \leq 4r_h\}} 1_{\{(\Delta_j X^{(2)})^2 \leq r_h, (\Delta_j \tilde{J}_2^{(1)})^2 \leq 4r_h\}}, \quad (18)$$

we remark that it is bounded in absolute value by

$$\text{Plim}_{n \rightarrow \infty} \sum_{j=1}^n |\Delta_j \tilde{J}_2^{(1)}| 1_{\{(\Delta_j \tilde{J}_2^{(1)})^2 \leq 4r_h\}} |\Delta_j \tilde{J}_2^{(2)}| 1_{\{(\Delta_j \tilde{J}_2^{(2)})^2 \leq 4r_h\}} \leq \text{Plim}_{n \rightarrow \infty} \sqrt{S_n^{(1)}} \sqrt{S_n^{(2)}} = 0,$$

by Lemma 7.5. □