Quadratic covariation estimation of an irregularly observed semimartingale with jumps and noise

Yuta Koike*

*Risk Analysis Research Center, Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan, Email: kyuta@ism.ac.jp

January 26, 2015

Abstract

This paper presents a central limit theorem for a pre-averaged version of the realized covariance estimator for the quadratic covariation of a discretely observed semimartingale with noise. The semimartingale possibly has jumps, while the observation times show irregularity, non-synchronicity, and some dependence on the observed process. It is shown that the observation times' effect on the asymptotic distribution of the estimator is only through two characteristics: the observation frequency and the covariance structure of the noise. This is completely different from the case of the realized covariance in a pure semimartingale setting.

Keywords: jumps; microstructure noise; non-synchronous observations; quadratic covariation; stable limit theorem; time endogeneity.

1 Introduction

The quadratic covariation matrix of a semimartingale is one of the fundamental quantities in statistics of semimartingales. In the context of the estimation of the diffusion coefficient of an Itô process observed discretely in a fixed interval, limit theorems associated with the discretized quadratic covariation play a key role, and such research has a long history (cf. [17, 19]). Furthermore, in recent years such an asymptotic theory has been applied to measuring the covariance structure of financial assets from high-frequency data. This was pioneered by [4, 7], and has become one of the most active areas in financial econometrics. In such a context the discretized quadratic covariation is also called the *realized covariance*.

However, raw high frequency data typically deviates from the ideal situation where we observe a continuous semimartingale at equidistant times, and this motivates statisticians to develop the theory in more complicated settings. One topic is the treatment of measurement errors in the data. For financial high-frequency data such errors originate from market microstructure noise and have attracted vast attention in the past decade; among various studies see e.g. [5, 46, 47, 52, 54]. In the univariate context, central limit theorems under irregular sampling settings have also been studied by many authors, especially assuming the independence between the observed process and the observation times; see e.g. [19, 22, 42]. In the multivariate case, the irregularity of the observation times causes the non-synchronicity which makes the analysis more complicated. The prominent works on this topic are the Fourier analysis approach of [40], the sampling design kernel method of [23] and the quasi-likelihood analysis of [44]. In addition, recently various approaches to deal with these issues simultaneously have been proposed by many authors; see e.g. [1, 6, 8, 9, 15, 51].

Another important issue is incorporating jumps into the model. In such a situation interest is often paid to estimating the integrated volatility and the integrated covariance matrix, i.e. the integrated diffusion coefficient, and there are many studies on this issue in various settings. Regarding the central limit theories, see e.g. Chapters 11 and 13 of [31] for the basic setting, [45] for the noise setting, [34] for the non-synchronous observation setting, and [13] for the noisy and non-synchronous observation setting.

In contrast, turning to the *entire* quadratic covariation estimation in the presence of jumps, there are fewer works. A central limit theorem for the realized covariance of an equidistantly observed Lévy process has been proved in Jacod and Protter [30] in the context of the analysis of the Euler scheme. This result has been extended to general Itô semimartingales in Jacod [27] as a special case of the asymptotic results on various functionals of semimartingale increments. The situation where measurement errors are present has been treated by Jacod *et al.* [29] who focus on the "pre-averaging" counterparts of the functionals discussed in [27], which were introduced in Podolskij and Vetter [46] to extend classical power variation based methods to a noisy observation setting. The theory requires a different treatment in the absence of the diffusion coefficient, and this case has been studied in Diop *et al.* [16].

When we further focus on the situation where the observation times are irregular, at least to the best of the author's knowledge, there is no comprehensive study on the central limit theory for the quadratic covariation estimation, except for the recent work of Bibinger and Vetter [12] and Bibinger and Winkelmann [13]; the former have derived central limit theorems for the realized covariance and the Hayashi-Yoshida estimator of [23] for a general Itô semimartingale observed irregularly and non-synchronously, while the latter have established a central limit theorem for an adjusted version of the spectral covariance estimator of [11] in a non-synchronous and noisy observation setup, focusing on asymptotically regular observation times in the sense that they satisfy conditions in Proposition 2.54 of [43]. The aim of this study is to develop such a theory in the situation where the observation data is contaminated by noise and the observation times are as general as possible. More precisely, we derive a central limit theorem for the pre-averaged version of the realized covariance proposed in Christensen *et al.* [14] (called the *modulated realized covariance*) under an irregular sampling setting in the presence of jumps. The main finding of this paper is that in the synchronous case the observation times' effect on the asymptotic distribution of the estimator is *only* through their conditional expected durations, provided that the limit of such quantities are well-defined. In other words, the irregularity of the observation times has *no* impact on the asymptotic distribution of the estimator because the conditional expected durations of the observation times naturally link with the magnitude of the observation frequency, and thus their effect is not due to the irregularity. This is completely different from the pure semimartingale setting of [12] where the distribution of the durations around the jump times of the semimartingale directly affects the asymptotic distribution of the realized covariance.

To deal with non-synchronous observations we rely on a data synchronization method proposed in Aït-Sahalia *et al.* [1], which also matches the proposal of Section 3.6 of [14]. The non-synchronicity naturally links with the covariance structure of the noise, hence it affects the asymptotic distribution through that relation. On the other hand, the interpolations to the synchronized sampling times do not matter asymptotically. This can be seen as a counterpart of the finding of Bibinger [8] in the continuous case.

Another issue we attempt to solve is how the dependence between the observed process and the observation times (called the *time endogeneity*) affects the asymptotic theory in our setting. This issue has recently been highlighted by several authors such as [18, 37, 38, 50] in various settings, and it is indeed known that such dependence possibly causes a non-standard limit theorem even in the continuous semimartingale setting. In this paper, this issue is partly solved in the sense that we do not rule out the dependence between the continuous component of the process and the observation times, but partly rule out the dependence between the jump component and the observation times. The result shows that the time endogeneity is also immaterial in our setting.

This paper is organized as follows. Section 2 presents the mathematical model and the construction of the estimator we are focusing on. Section 3 is devoted to the main result of this paper. Section 4 provides some illustrative examples of the observation times, while Section 5 provides a simulation study. All proofs are given in Section 6.

2 The set up

Given a stochastic basis $\mathcal{B}^{(0)} = (\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t\geq 0}, P^{(0)})$, we consider a d-dimensional semimartingale $X = (X_t)_{t\in\mathbb{R}_+}$ of the form

$$X_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + (\delta 1_{\{\|\delta\|\leq 1\}}) \star (\mu - \nu)_t + (\delta 1_{\{\|\delta\|>1\}}) \star \mu_t,$$

where W is a d'-dimensional $(\mathcal{F}_t^{(0)})$-standard Brownian motion, μ is an $(\mathcal{F}_t^{(0)})$-Poisson random measure on $\mathbb{R}_+ \times E$ with E being a Polish space, ν is the intensity measure of μ of the form $\nu(\mathrm{d}t, \mathrm{d}z) = \mathrm{d}t \otimes \lambda(\mathrm{d}z)$ with λ being a σ-finite measure on E, b is an $(\mathcal{F}_t^{(0)})$-progressively measurable $\mathbb{R}^d$-valued process, σ is an $(\mathcal{F}_t^{(0)})$-progressively measurable $\mathbb{R}^d \otimes \mathbb{R}^{d'}$-valued process, and δ is an $(\mathcal{F}_t^{(0)})$-predictable $\mathbb{R}^d$-valued function on $\Omega^{(0)} \times \mathbb{R}_+ \times E$. Also, $\star$ denotes the integral (either stochastic or ordinary) with respect to some (integer-valued) random measure. Here and below we use standard concepts and notation in stochastic calculus, which are described in detail in e.g. Chapter 2 of [31]. Our aim is to estimate the quadratic covariation matrix process $[X, X] = ([X^k, X^l])_{1\leq k,l\leq d}$ of X from noisy and discrete observation data of X.

The observed process Y is subject to additional measurement errors as follows:

$$Y_t = X_t + \epsilon_t.$$

The mathematical construction of the noise process ϵ is explained later. We observe the components of the d-dimensional process $Y = (Y^1, \dots, Y^d)$ discretely and non-synchronously. For each $k = 1, \dots, d$ the observation times for Y^k are denoted by $t_0^k, t_1^k, \dots$, i.e. the observation data $(Y^k_{t_i^k})_{i\in\mathbb{Z}_+}$ is available. We assume that $(t_i^k)_{i=0}^\infty$ is a sequence of $(\mathcal{F}_t^{(0)})$-stopping times which implicitly depend on a parameter $n \in \mathbb{N}$ representing the observation frequency and satisfy that $t_i^k \uparrow \infty$ as $i \to \infty$ and $\sup_{i\geq 0}(t_i^k \wedge t - t_{i-1}^k \wedge t) \to^p 0$ as $n \to \infty$ for any $t \in \mathbb{R}_+$, with setting $t_{-1}^k = 0$ for notational convenience (hereafter we will refer to such a sequence as a *sampling scheme* for short).

Now we introduce the precise definition of the noise process ϵ. It is basically the same as the one from Chapter 16 of [31], but we need a slight modification to ensure the (joint) measurability of the process ϵ, which is necessary for us to consider variables such as $\epsilon^k_{t_i^k}$. For any $t \in \mathbb{R}_+$ there is a transition probability $Q_t(\omega^{(0)}, \mathrm{d}u)$ from $(\Omega^{(0)}, \mathcal{F}_t^{(0)})$ into $\mathbb{R}^d$ satisfying $\int u Q_t(\omega^{(0)}, \mathrm{d}u) = 0$ (this will correspond to the conditional distribution of the noise at the time t given $\mathcal{F}_t^{(0)}$). Then, at each frequency $n \in \mathbb{N}$, the stochastic basis $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$ supporting the observed process Y is constructed in the following manner (for notational simplicity we subtract the index n from $\mathcal{B}$): We endow the space $\Omega^{(1)} = (\mathbb{R}^d)^{\mathbb{N}}$ with the product Borel σ-field $\mathcal{F}^{(1)}$ and with the probability measure $Q(\omega^{(0)}, \mathrm{d}\omega^{(1)})$ which is the product $\otimes_{i\in\mathbb{N}} Q_{\mathcal{T}_i^n(\omega^{(0)})}(\omega^{(0)}, \cdot)$. Here, $(\mathcal{T}_i^n)_{i\geq 0}$ is the increasing reordering of total observation times $\{t_i^k : k = 1, \dots, d \text{ and } i \in \mathbb{Z}_+\}$. More formally, it is defined sequentially by $\mathcal{T}_0^n = \min_{k=1,\dots,d} t_0^k$ and $\mathcal{T}_i^n = \min_{k=1,\dots,d} \min\{t_j^k : t_j^k > \mathcal{T}_{i-1}^n\}$ for $i = 1, 2 \dots$. Note that $\mathcal{T}_i^n$ is an $(\mathcal{F}_t^{(0)})$-stopping time since $\mathcal{T}_i^n = \min_{k=1,\dots,d} \inf_{j\geq 1} \left(t_i^k\right)_{\{t_j^k > \mathcal{T}_{i-1}^n\}}$, where for an $(\mathcal{F}_t^{(0)})$-stopping time τ and a set $A \in \mathcal{F}_\tau^{(0)}$, we define τ_A by $\tau_A(\omega^{(0)}) = \tau(\omega^{(0)})$ if $\omega^{(0)} \in A$; $\tau_A(\omega^{(0)}) = \infty$ otherwise (see I-1.15 of [32]). Then, we define the probability space $(\Omega, \mathcal{F}, P)$ by

$$\Omega = \Omega^{(0)} \times \Omega^{(1)}, \qquad \mathcal{F} = \mathcal{F}^{(0)} \otimes \mathcal{F}^{(1)}, \qquad P(\mathrm{d}\omega^{(0)}, \mathrm{d}\omega^{(1)}) = P^{(0)}(\mathrm{d}\omega^{(0)}) Q(\omega^{(0)}, \mathrm{d}\omega^{(1)}). \tag{2.1}$$

After that, the noise process $\epsilon = (\epsilon_t)_{t\geq 0}$ is defined on this probability space by $\epsilon_t = \epsilon^0_{\mathsf{N}_n(t)}$, where $(\epsilon_i^0)_{i\in\mathbb{N}}$ denotes the canonical process on $(\Omega^{(1)}, \mathcal{F}^{(1)})$ and $\mathsf{N}_n(t) = \sum_{i=0}^\infty 1_{\{\mathcal{T}_i^n \leq t\}}$. Finally, the filtration $(\mathcal{F}_t)_{t\geq 0}$ is defined as the one generated by $(\mathcal{F}_t^{(0)})_{t\geq 0}$ and $(\epsilon_t)_{t\geq 0}$.

Any variable or process defined on either $\Omega^{(0)}$ or $\Omega^{(1)}$ is considered in the usual way as a variable or a process on Ω. Specifically, our noisy process $Y = (Y_t)_{t\geq 0}$ is the process defined as the sum of the latent process X on $\Omega^{(0)}$ and the noise process ϵ on Ω.

Remark 2.1. To ensure that the probability measure P in (2.1) is well-defined, we further need the measurability of the map $\omega^{(0)} \mapsto Q(\omega^{(0)}, A)$ for any Borel subset A of $\mathbb{R}^d$. This is ensured by the progressive measurability of the process $(Q_t(\cdot, A))_{t\geq 0}$ which we will assume later (see assumption [A4]). This assumption also ensures that $\mathcal{B}$ is the *very good* filtered extension of $\mathcal{B}^{(0)}$, which is necessary to apply the version of Jacod's stable limit theorem described by Theorem 2.2.15 of [31].

To deal with the non-synchronicity of the observation times we rely on a data synchronization method, which is commonly used in the literature; see e.g. [1, 6, 14, 53]. Let $(T_p)_{p=0}^\infty$ and $(\tau_p^k)_{p=0}^\infty$ $(k = 1, \dots, d)$ be sampling schemes such that

$$\tau_0^k \leq T_0 \quad \text{and} \quad T_{p-1} < \tau_p^k \leq T_p \quad \text{for any } p \geq 1 \text{ and any } k = 1, \dots, d. \tag{2.2}$$

We assume that the observation data $(Y^k_{\tau_p^k})_{p\in\mathbb{Z}_+}$ is available for every $k = 1, \dots, d$, i.e. $\{\tau_p^k : p \geq 0\} \subset \{t_i^k : i \geq 0\}$, and construct statistics based on this synchronized data set $(Y^k_{\tau_p^k})_{p\in\mathbb{Z}_+}$, $k = 1, \dots, d$. In Aït-Sahalia *et al.* [1] this type of synchronization method is called the *Generalized Synchronization method* and $(T_p)_{p=0}^\infty$ is called the *Generalized Sampling Time*. One way to implement such synchronization is the so-called *refresh time sampling method* introduced by Barndorff-Nielsen *et al.* [6] to this area. Namely, we first define the *refresh times* $T_0, T_1, \dots$ of the sampling schemes $\{(t_i^k)\}_{k=1}^d$ sequentially by $T_0 = \max\{t_0^1, \dots, t_0^d\}$ and $T_p = \max_{k=1,\dots,d} \min\{t_i^k : t_i^k > T_{p-1}\}$ for $p = 1, 2, \dots$. After that, for each k, (τ_p^k) is defined by interpolating the next-ticks into (T_p) as follows:

$$\tau_0^k = t_0^k \quad \text{and} \quad \tau_p^k = \min\{t_i^k : t_i^k > T_{p-1}\}, \qquad p = 1, 2, \dots.$$

Note that τ_p^k is an $(\mathcal{F}_t^{(0)})$-stopping time due to an analogous reason to that for $\mathcal{T}_i^n$.

Now the modulated realized covariance (henceforth MRC) estimator we focus on is constructed in the following way. First, we choose a sequence k_n of positive integers and a number $\theta \in (0, \infty)$ such that $k_n = \theta\sqrt{n} + o(n^{1/4})$ as $n \to \infty$. We also choose a continuous function $g : [0, 1] \to \mathbb{R}$ which is piecewise C^1 with a piecewise Lipschitz derivative g' and satisfies $g(0) = g(1) = 0$ and $\int_0^1 g(x)^2 \mathrm{d}x > 0$. After that, for any d-dimensional stochastic process $V = (V^1, \dots, V^d)$ we define the quantity

$$\overline{V}_i^k = \sum_{p=1}^{k_n - 1} g\left(\frac{p}{k_n}\right) \left(V^k_{\tau^k_{i+p}} - V^k_{\tau^k_{i+p-1}}\right), \tag{2.3}$$

and set $\overline{V}_i = (\overline{V}_i^1, \dots, \overline{V}_i^d)^*$ (hereafter an asterisk denotes the transpose of a matrix). The MRC estimator is defined by

$$\mathrm{MRC}[Y]_t^n = \frac{1}{\psi_2 k_n} \sum_{i=0}^{N_t^n - k_n + 1} \overline{Y}_i \left(\overline{Y}_i\right)^* - \frac{\psi_1}{2\psi_2 k_n^2} [Y, Y]_t^n,$$

where $N_t^n = \max\{p : T_p \leq t\}$, $\psi_1 = \int_0^1 g'(x)^2 \mathrm{d}x$, $\psi_2 = \int_0^1 g(x)^2 \mathrm{d}x$ and

$$[Y, Y]_t^n = \sum_{p=1}^{N_t^n} \Delta_p Y \left(\Delta_p Y\right)^*, \qquad \Delta_p Y = \left(Y^1_{\tau_p^1} - Y^1_{\tau_{p-1}^1}, \dots, Y^d_{\tau_p^d} - Y^d_{\tau_{p-1}^d}\right)^*$$

for each $t \in \mathbb{R}_+$. Here, we set $\sum_{i=p}^q \equiv 0$ if $p > q$ by convention. In the synchronous and equidistant sampling setting, the asymptotic distribution of the MRC estimator has been derived in Jacod *et al.* [29] (see Theorem 4.6 of that paper, and see also Section 4 of Hautsch and Podolskij [21]). Our purpose is to develop the asymptotic distribution of the MRC estimator in the situation where the observation times are possibly irregular, non-synchronous and endogenous.

3 Main result

3.1 Notation

In this subsection some notation is introduced in order to state our main result. First we introduce notation appearing in the assumptions stated in the next subsection. We write $X^n \xrightarrow{ucp} X$ for processes X^n and X to express shortly that $\sup_{0\leq t\leq T}|X^n_t - X_t| \to^p 0$ for any $T>0$. ϖ denotes some (fixed) positive constant. We denote by $(\mathcal{G}^{(0)}_t)$ (resp. $(\mathcal{G}_t)$) the smallest filtration containing $(\mathcal{F}^{(0)}_t)$ (resp. $(\mathcal{F}_t)$) such that $\mathcal{G}^{(0)}_0$ (resp. $\mathcal{G}_0$) contains the σ-field generated by μ, i.e. the σ-field generated by all the variables $\mu(A)$, where A ranges all measurable subsets of $\mathbb{R}_+ \times E$.

Next we introduce some quantities appearing in the representation of the asymptotic variance of the estimator. We set $\Sigma_s = \sigma_s\sigma_s^*$ for each $s\in\mathbb{R}_+$, i.e. Σ denotes the diffusion coefficient matrix process. We denote by Υ_t the covariance matrix of ϵ_t, i.e. $\Upsilon_t(\cdot)=\int uu^*Q_t(\cdot,\mathrm{d}u)$ (we will assume the existence of the second moment of the noise later, so this matrix always exists). For any real-valued bounded measurable functions u,v on $[0,1]$, we define the function $\phi_{u,v}$ on $[0,1]$ by $\phi_{u,v}(y)=\int_y^1 u(x-y)v(x)\mathrm{d}x$. Then, we put

$$\Phi_{22}=\int_0^1 \phi_{g,g}(y)^2\mathrm{d}y,\qquad \Phi_{12}=\int_0^1 \phi_{g,g}(y)\phi_{g',g'}(y)\mathrm{d}y,\qquad \Phi_{11}=\int_0^1 \phi_{g',g'}(y)^2\mathrm{d}y.$$

On the other hand, for any $k,l=1,\dots,d$ we define the process $\mathfrak{J}^{kl}$ by

$$\mathfrak{J}^{kl}_s = \Delta X^k_s \Delta X^l_s \left\{\Phi_{22}\theta\left(\Sigma^{kl}_{s-}G_{s-}+\Sigma^{kl}_s G_s\right)+\frac{\Phi_{12}}{\theta}\left(\Upsilon^{kl}_{s-}\chi^{kl}_{s-}+\Upsilon^{kl}_s\chi^{kl}_s\right)\right\}.$$

Remark 3.1 (Properties of $\phi_{u,v}$)**.** We will use the following properties of $\phi_{u,v}$: First, for any real-valued bounded measurable function u on $[0,1]$, $\phi_{u,u}$ is non-negative. In fact, setting $u(x)=0$ for $x\notin[0,1]$, we have $\phi_{u,u}(y)=\int_{-\infty}^{\infty}u(-(y-x))u(x)\mathrm{d}x$ for all $y\in[0,1]$, and we can extend the domain of $\phi_{u,u}$ to the whole real line using this expression. Then, denoting by $\hat{f}$ the Fourier transform of a function f on $\mathbb{R}$, we have $\hat{\phi}_{u,u}=|\hat{u}|^2\geq 0$. Hence $\phi_{u,u}$ is a positive definite function and, in particular, $\phi_{u,u}(y)\geq 0$ for all $y\in\mathbb{R}$. Next, we can easily check that $\phi'_{g,g}=\phi_{g,g'}=-\phi_{g',g}$ and $\phi''_{g,g}=-\phi_{g',g'}$. In particular, $\Phi_{12}=\int_0^1\phi_{g',g}(y)^2\mathrm{d}y$ due to integration by parts.

3.2 Assumptions

We impose the following condition on the sampling schemes $(T_p)_{p\geq 0}$ and $(\tau^k_p)_{p\geq 0}$ $(k=1,\dots,d)$:

[A1] $(T_p)_{p\geq 0}$ and $(\tau^k_p)_{p\geq 0}$ $(k=1,\dots,d)$ are sequences of $(\mathcal{F}^{(0)}_t)$-stopping times and satisfy (2.2). It also holds that

$$r_n(t) := \sup_{p\geq 0}(T_p\wedge t - T_{p-1}\wedge t) = o_p(n^{-\xi}) \tag{3.1}$$

as $n\to\infty$ (note that $T_{-1}=0$ by convention) for every $t>0$ and every $\xi\in(0,1)$. Moreover, for each n we have a $(\mathcal{G}^{(0)}_t)$-progressively measurable positive-valued process G^n_t, a $(\mathcal{G}^{(0)}_t)$-progressively measurable $[0,1]^d\otimes[0,1]^d$-valued process $\chi^n_t=(\chi^{n,kl}_t)_{1\leq k,l\leq d}$ and a random subset $\mathcal{N}^n$ of $\mathbb{Z}_+$ satisfying the following conditions:

- (i) $\{(\omega,p)\in\Omega\times\mathbb{Z}_+ : p\in\mathcal{N}^n(\omega)\}$ is a measurable set of $\Omega\times\mathbb{Z}_+$. Moreover, there is a constant $\kappa\in(0,\frac{1}{2})$ such that $\#(\mathcal{N}^n\cap\{p:T_p\leq t\})=O_p(n^\kappa)$ as $n\to\infty$ for every $t>0$.
- (ii) $E[n(T_{p+1}-T_p)|\mathcal{G}^{(0)}_{T_p}]=G^n_{T_p}$ and $E[1_{\{\tau^k_{p+1}=\tau^l_{p+1}\}}|\mathcal{G}^{(0)}_{T_p}]=\chi^{n,kl}_{T_p}$ for every n, every $p\in\mathbb{Z}_+-\mathcal{N}^n$ and any $k,l=1,\dots,d$.
- (iii) There is a càdlàg $(\mathcal{F}^{(0)}_t)$-adapted positive valued process G such that
  - (iii-a) $n^\varpi(G^n-G)\xrightarrow{ucp}0$,
  - (iii-b) $G_{t-}>0$ for every $t>0$,

(iii-c) G is an Itô semimartingale of the form

$$G_t = G_0 + \int_0^t \widehat{b}_s \mathrm{d}s + \int_0^t \widehat{\sigma}_s \mathrm{d}W_s + (\widehat{\delta} 1_{\{|\widehat{\delta}|\leq 1\}}) \star (\mu - \nu)_t + (\widehat{\delta} 1_{\{|\widehat{\delta}|>1\}}) \star \mu_t,$$

where $\widehat{b}_s$ is a locally bounded and $(\mathcal{F}_t^{(0)})$-progressively measurable real-valued process, $\widehat{\sigma}_s$ is a càdlàg $(\mathcal{F}_t^{(0)})$-adapted $\mathbb{R}\otimes\mathbb{R}^{d'}$-valued process, and $\widehat{\delta}$ is an $(\mathcal{F}_t^{(0)})$-predictable real-valued function on $\Omega^{(0)} \times \mathbb{R}_+ \times E$ such that there is a sequence $(\widehat{\rho}_j)$ of $(\mathcal{F}_t^{(0)})$-stopping times increasing to infinity and, for each j, a deterministic non-negative function $\widehat{\gamma}_j$ on E satisfying $\int \widehat{\gamma}_j(z)^2 \wedge 1 \lambda(\mathrm{d}z) < \infty$ and $|\widehat{\delta}(\omega^{(0)}, t, z)| \leq \widehat{\gamma}_j(z)$ for all $(\omega^{(0)}, t, z)$ with $t \leq \widehat{\rho}_j(\omega^{(0)})$.

(iv) There is a càdlàg $(\mathcal{F}_t^{(0)})$-adapted $[0,1]^d \otimes [0,1]^d$-valued process χ such that $n^{\varpi}(\chi^n - \chi) \xrightarrow{ucp} 0$ as $n \to \infty$. Furthermore, for each $j \in \mathbb{N}$ we have a càdlàg $(\mathcal{F}_t^{(0)})$-adapted $[0,1]^d \otimes [0,1]^d$-valued process $\chi(j)$, an $(\mathcal{F}_t^{(0)})$-stopping time $\check{\rho}_j$, and a constant $\check{\Lambda}_j$ such that $\check{\rho}_j \uparrow \infty$ as $j \to \infty$ and $\chi(\omega^{(0)})_t = \chi(j)(\omega^{(0)})_t$ if $t < \check{\rho}_j(\omega^{(0)})$ and

$$E\left[\|\chi(j)_{t_1} - \chi(j)_{t_2}\|^2 | \mathcal{F}_{t_1 \wedge t_2}\right] \leq \check{\Lambda}_j E\left[|t_1 - t_2|^{\varpi} | \mathcal{F}_{t_1 \wedge t_2}\right]$$

for every j and any $(\mathcal{F}_t^{(0)})$-stopping times t_1 and t_2 bounded by j.

Remark 3.2. (i) The assumptions on (T_p) are motivated by the concept of the *restricted discretization scheme* discussed in detail in Chapter 14 of [31]. In fact, suppose that T_p's are of the form

$$T_p = T_{p-1} + \theta^n_{T_{p-1}} \varepsilon(n,p), \qquad p = 1, 2, \dots,$$

where θ^n is a càdlàg $(\mathcal{F}_t^{(0)})$-adapted process, $(\varepsilon(n,p))_{p\geq 1}$ is a sequence of i.i.d. positive variables independent of b, σ, δ, W, μ, and such that $E[\varepsilon(n,p)] = 1$ and $E[\varepsilon(n,p)^r] < \infty$ for every $r > 0$, and $T_0 = 0$. By constructing the filtration $(\mathcal{F}_t^{(0)})$ suitably, we may assume that $\epsilon(n,p)$ is independent of $\mathcal{F}^{(0)}_{T_{p-1}}$ for all n, p. Then we have [A1](i)–(ii) regarding G^n while we set $\mathcal{N}^n = \emptyset$ and $G^n = n\theta^n$. In this case [A1](iii) corresponds to (a weaker version of) Assumption (E) from [31], and (3.1) follows from Lemma 14.1.5 of [31]. Unlike their setting, however, our assumption does not rule out the dependence between $\varepsilon(n,p)$'s and X (see e.g. Example 4.1 in the next section). The importance of such dependence has recently been emphasized in econometric literature; see e.g. Renault and Werker [49].
(ii) The assumptions on the quantities $1_{\{\tau_p^k = \tau_p^l\}}$ are necessary for the treatment of the ($\mathcal{F}^{(0)}$-conditional) covariance between $\epsilon^k_{\tau_p^k}$ and $\epsilon^l_{\tau_p^l}$, which is given by $\Upsilon^{kl}_{\tau_p^k} 1_{\{\tau_p^k = \tau_p^l\}}$ (a similar kind of assumption also appears in Bibinger and Mykland [10] due to the same reason as ours). Therefore, those assumptions can be dropped when $\Upsilon^{kl} \equiv 0$ if $k \neq l$; this is often assumed in the literature on the covariance estimation of non-synchronously observed semimartingales with noise. The quantity χ^n measures the degree of the non-synchronicity, and χ^n_s is a matrix all of whose components are equal to 1 in the synchronous case while it is an identity matrix in the completely non-synchronous case. Hence [A1](iv) is satisfied in these two extreme cases.
(iii) The possibility of the set $\mathcal{N}^n$ being non-empty excludes the following trivial exception of [A1] with $\mathcal{N}^n$ being empty: if $T_0 = \log n/n$ and $T_p = T_{p-1} + 1/n$ for $p \geq 1$, [A1] with $\mathcal{N}^n = \emptyset$ is not satisfied because $G^n_{T_0} \to \infty$ as $n \to \infty$. This assumption is also useful to ensure the stability under the localization used in the proof; see Lemma 6.1.
(iv) The fact that we consider the conditional expected durations given $\mathcal{G}^{(0)}_{T_p}$'s instead of $\mathcal{F}^{(0)}_{T_p}$ rules out some dependence between the sampling schemes and the jumps of the observed process. For example, if μ is a jump measure of a one-dimensional Lévy process (i.e. $E = \mathbb{R}$) and T_p's are of the form $T_p = \inf\{t > T_{p-1} : |\int_{T_{p-1}}^t \int_{|z|\leq 1} z(\mu - \nu)(\mathrm{d}s, \mathrm{d}z)| > \eta_n\}$ for $p = 1, 2, \dots$ and for some appropriate sequence $(\eta_n)_{n\geq 1}$ of positive numbers, then [A1] obviously fails because T_p's are $\mathcal{G}_0^{(0)}$-measurable (this type of sampling scheme is well studied in Rosenbaum and Tankov [50]). On the other hand, it still allows the presence of the instantaneous causality between the sampling schemes and the jumps: see Example 4.2.

(v) Under [A1] it holds that

$$\frac{1}{n}N_t^n \to^p \int_0^t \frac{1}{G_s}\mathrm{d}s \tag{3.2}$$

as $n \to \infty$ for every $t \in \mathbb{R}_+$ (see Section 6.1 of [36] for the proof). In particular, [A1] ensures that the parameter n controls the magnitude of the number of observations.

We impose the following structural assumption on the latent process X:

[A2] The volatility process σ is an Itô semimartingale of the form

$$\sigma_t = \sigma_0 + \int_0^t \widetilde{b}_s\mathrm{d}s + \int_0^t \widetilde{\sigma}_s\mathrm{d}W_s + (\widetilde{\delta}1_{\{|\widetilde{\delta}|\leq 1\}}) \star (\mu - \nu)_t + (\widetilde{\delta}1_{\{|\widetilde{\delta}|>1\}}) \star \mu_t,$$

where $\widetilde{b}_s$ is a locally bounded and $(\mathcal{F}_t^{(0)})$-progressively measurable $\mathbb{R}^d \otimes \mathbb{R}^{d'}$-valued process, $\widetilde{\sigma}_s$ is a càdlàg $(\mathcal{F}_t^{(0)})$-adapted $\mathbb{R}^d \otimes \mathbb{R}^{d'} \otimes \mathbb{R}^{d'}$-valued process, and $\widetilde{\delta}$ is an $(\mathcal{F}_t^{(0)})$-predictable $\mathbb{R}^d \otimes \mathbb{R}^{d'}$-valued function on $\Omega^{(0)} \times \mathbb{R}_+ \times E$.

Moreover, for each j there is an $(\mathcal{F}_t^{(0)})$-stopping time ρ_j, a bounded $(\mathcal{F}_t^{(0)})$-progressively measurable $\mathbb{R}^d$-valued process $b(j)_s$, a deterministic non-negative function γ_j on E, and a constant Λ_j such that $\rho_j \uparrow \infty$ as $j \to \infty$ and, for each j,

(i) $b(\omega^{(0)})_s = b(j)(\omega^{(0)})_s$ if $s < \rho_j(\omega^{(0)})$,

(ii) $E\left[\|b(j)_{t_1} - b(j)_{t_2}\|^2|\mathcal{F}_{t_1\wedge t_2}\right] \leq \Lambda_j E\left[|t_1 - t_2|^{\varpi}|\mathcal{F}_{t_1\wedge t_2}\right]$ for any $(\mathcal{F}_t^{(0)})$-stopping times t_1 and t_2 bounded by j,

(iii) $\int \left\{\gamma_j(z)^2 \wedge 1\right\} \lambda(\mathrm{d}z) < \infty$ and $\|\delta(\omega^{(0)}, t, z)\| \vee \|\widetilde{\delta}(\omega^{(0)}, t, z)\| \leq \gamma_j(z)$ for all $(\omega^{(0)}, t, z)$ with $t \leq \rho_j(\omega^{(0)})$,

(iv) $E\left[\|\delta(t_1 \wedge \rho_j, z) - \delta(t_2 \wedge \rho_j, z)\|^2|\mathcal{F}_{t_1\wedge t_2}\right] \leq \Lambda_j\gamma_j(z)^2 E\left[|t_1 - t_2|^{\varpi}|\mathcal{F}_{t_1\wedge t_2}\right]$ for any $(\mathcal{F}_t^{(0)})$-stopping times t_1 and t_2 bounded by j.

Remark 3.3. An [A2] type assumption is commonly used in the literature of power variations (see e.g. [31]), except for assumptions (ii) and (iv), i.e. continuity assumptions on the drift and the jump coefficient. Such assumptions are necessary for the treatment of the irregularity and the non-synchronicity of the observation times as in [24].

We also impose the following regularity condition on the noise process:

[A3] There is a constant $\Gamma > 4$ and a sequence $(\rho'_j)_{j\geq 1}$ of $(\mathcal{F}_t^{(0)})$-stopping times increasing to infinity such that

$$\sup_{\omega^{(0)}\in\Omega^{(0)}, t<\rho'_j(\omega^{(0)})} \int \|z\|^{\Gamma} Q_t(\omega^{(0)}, \mathrm{d}z) < \infty.$$

Moreover, for each j there is a bounded càdlàg $(\mathcal{F}_t^{(0)})$-adapted $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process $\Upsilon(j)_t$ and a constant Λ'_j such that

(i) $\Upsilon(j)(\omega^{(0)})_t = \Upsilon(\omega^{(0)})_t$ if $t < \rho'_j(\omega^{(0)})$,

(ii) $E\left[\|\Upsilon(j)_{t_1} - \Upsilon(j)_{t_2}\|^2|\mathcal{F}_{t_1\wedge t_2}\right] \leq \Lambda'_j E\left[|t_1 - t_2|^{\varpi}|\mathcal{F}_{t_1\wedge t_2}\right]$ for any $(\mathcal{F}_t^{(0)})$-stopping times t_1 and t_2 bounded by j.

Remark 3.4. The locally boundedness of the moment process of the noise is used for verifying a Lyapunov type condition for central limit theorems and proving the negligibility of the edge effect. The continuity assumption of the covariance matrix process of the noise is necessary due to the same reason as for [A2]. If the noise is assumed to be i.i.d. and independent of $\mathcal{F}^{(0)}$, [A3] simply means the Γ-th moment of the noise is finite for some $\Gamma > 4$.

Finally, we introduce the following technical condition to avoid some measure-theoretic problems:

[A4] (i) A regular conditional probability of $P^{(0)}$ given $\mathcal{H}$ exists for any sub-σ-field $\mathcal{H}$ of $\mathcal{F}^{(0)}$.

(ii) The process $(Q_t(\cdot, A))_{t\geq 0}$ is $(\mathcal{F}_t^{(0)})$-progressively measurable for any Borel set A of $\mathbb{R}^d$.

Remark 3.5. (i) [A4](i) is satisfied, for example, when $(\Omega^{(0)}, \mathcal{F}^{(0)})$ is a standard measurable space, i.e. it is Borel isomorphic to some Polish space (see e.g. Theorem I-3.1 of [26]). In fact, this assumption is not restrictive for applications.
(ii) [A4](ii) is satisfied, for example, when $Q_t \equiv Q$ for some probability measure Q on $\mathbb{R}^d$, i.e. the noise is modeled by an i.i.d. sequence. Another example is the case where $Q_t(\omega^{(0)}, \cdot)$ has a density of the form $f(\cdot, X_t(\omega^{(0)}))$, where f is a measurable function on $\mathbb{R}^d \times \mathbb{R}^d$ into $[0,1]$ such that $\int_{\mathbb{R}^d} f(x,\theta)\mathrm{d}x = 1$ for every $\theta \in \mathbb{R}^d$. Example 16.1.5 of [31] is encompassed with this type of model. Thus, this assumption also seems to be unrestrictive for applications.

3.3 Result

To state the main result, we need the notion of *stable convergence* as common in this area. For each $n \geq 1$, let X^n be a random variable which is defined on $\mathcal{B}$ and takes values in a Polish space S. The variables X^n are said to *converge stably in law* to an S-valued random variable X defined on an extension of $\mathcal{B}^{(0)}$ if $E[Uf(X^n)] \to \widetilde{E}[Uf(X)]$ for any $\mathcal{F}^{(0)}$-measurable bounded random variable U and any bounded continuous function f on S, where $\widetilde{E}$ denotes the expectation with respect to the probability measure of the extension. We then write $X^n \to^{d_s} X$. Note that we need a slightly generalized definition of stable convergence described at the end of Section 2.2.1 of [31] because $\mathcal{B}$ changes as n varies. The most important property of stable convergence is the following: if the real-valued variables V_n defined on $\mathcal{B}$ converge in probability to a variable V defined on $\mathcal{B}^{(0)}$, then $X^n \to^{d_s} X$ implies that $(X^n, V_n) \to^{d_s} (X, V)$ for the product topology on the space $S \times \mathbb{R}$.

Theorem 3.1. *Suppose that* [A1]–[A4] *are satisfied. Then*

$$n^{1/4}\left(\mathrm{MRC}[Y]^n_t - [X,X]_t\right) \to^{d_s} \mathcal{W}_t + \mathcal{Z}_t$$

as $n \to \infty$ for any $t > 0$, where $\mathcal{W}$ and $\mathcal{Z}$ are $\mathbb{R}^d \otimes \mathbb{R}^d$-valued processes defined on an extension of $\mathcal{B}^{(0)}$, which conditionally on $\mathcal{F}^{(0)}$ are mutually independent, centered Gaussian with independent increments, the first one being continuous and the second one being purely discontinuous, and with (conditional) covariances

$$\widetilde{E}\left[\mathcal{W}^{kl}_t \mathcal{W}^{k'l'}_t \middle| \mathcal{F}^{(0)}\right] = \frac{2}{\psi_2^2}\int_0^t \Bigg[\Phi_{22}\theta\left\{\Sigma^{kk'}_s\Sigma^{ll'}_s + \Sigma^{kl'}_s\Sigma^{lk'}_s\right\}G_s + \frac{\Phi_{11}}{\theta^3}\left\{\Upsilon^{kk'}_s\chi^{kk'}_s\Upsilon^{ll'}_s\chi^{ll'}_s + \Upsilon^{kl'}_s\chi^{kl'}_s\Upsilon^{lk'}_s\chi^{lk'}_s\right\}\frac{1}{G_s}$$
$$+\frac{\Phi_{12}}{\theta}\left\{\Sigma^{kk'}_s\Upsilon^{ll'}_s\chi^{ll'}_s + \Sigma^{lk'}_s\Upsilon^{kl'}_s\chi^{kl'}_s + \Sigma^{ll'}_s\Upsilon^{kk'}_s\chi^{kk'}_s + \Sigma^{kl'}_s\Upsilon^{lk'}_s\chi^{lk'}_s\right\}\Bigg]\mathrm{d}s \tag{3.3}$$

and

$$\widetilde{E}\left[\mathcal{Z}^{kl}_t \mathcal{Z}^{k'l'}_t \middle| \mathcal{F}^{(0)}\right] = \frac{1}{\psi_2^2}\sum_{s\leq t}\left(\mathfrak{J}^{kk'}_s + \mathfrak{J}^{kl'}_s + \mathfrak{J}^{lk'}_s + \mathfrak{J}^{ll'}_s\right). \tag{3.4}$$

Here, $\widetilde{E}$ denotes the expectation with respect to the probability measure of the extension.

When further X is continuous, the processes $n^{1/4}\left(\mathrm{MRC}[Y]^n - [X,X]\right)$ converge stably in law to the process $\mathcal{W}$ for the Skorokhod topology.

Remark 3.6. The above theorem shows that the observation times' effect on the asymptotic distribution of the MRC estimator is only through the asymptotic conditional expected duration process G and the limiting process χ measuring the degree of the non-synchronicity. As was indicated in Remark 3.2(ii), χ simply reflects the covariance structure of the noise process, while G naturally affects the asymptotic distribution of the estimator because it links with the (spot) sampling frequency, as seen from (3.2). Consequently, the irregularity and the endogeneity of the observation times have no impact on the asymptotic distribution of the estimator.

Remark 3.7. In the proof of the theorem, it plays a key role to replace the duration $(T_{p+1} - T_p)$ with its conditional expectation $G^n_{T_p}$. Such replacement is possible because our estimator contains a local averaging procedure (2.3). More formally, this procedure makes it possible to apply a standard martingale argument described in Lemma 2.3 of [18] to the durations. The benefits of this fact appear in the treatments of the irregularity and the endogeneity of the observation times in Lemmas 6.4 and 6.7. Also, this is why the higher (conditional) moments of the durations do not affect the asymptotic distribution of the estimator.

Remark 3.8 (Covariance structure of $\mathcal{W}_t$)**.** It is convenient to observe that the covariance structure of $\mathcal{W}_t$ is analogous to the asymptotic covariance of the realized covariance in a standard setting. For this purpose, in the following we use some concepts from matrix algebra found in e.g. Horn and Johnson [25]. For each $s \in \mathbb{R}_+$, we denote by $\widetilde{\Upsilon}_s$ the Hadamard product of Υ_s and χ_s, i.e. $\widetilde{\Upsilon}^{kl}_s = \Upsilon^{kl}_s \chi^{kl}_s$ for $k, l = 1, \dots, d$, and set $\overline{\Sigma}_s(y) = \frac{\sqrt{2}}{\psi_2}\left\{\phi_{g,g}(y)\theta^{\frac{1}{2}}\Sigma_s\sqrt{G_s} + \phi_{g',g'}(y)\theta^{-\frac{3}{2}}\widetilde{\Upsilon}_s/\sqrt{G_s}\right\}$. Since both Υ_s and χ_s is positive semi-definite, so is $\widetilde{\Upsilon}_s$ due to the Schur product theorem, i.e. Theorem 5.2.1 of [25] (note that the positive semi-definiteness of χ_s can be checked directly using the fact that $\chi^{kk}_s = 1$ and $0 \leq \chi^{kl}_s \leq 1$ for any k, l). Therefore, $\overline{\Sigma}_s(y)$ is positive semi-definite as well because both $\phi_{g,g}$ and $\phi_{g',g'}$ are non-negative (see Remark 3.1). Then the left side of (3.3) can be rewritten as

$$\int_0^t \left[\int_0^1 \left\{\overline{\Sigma}_s(y)^{kk'}\overline{\Sigma}_s(y)^{ll'} + \overline{\Sigma}_s(y)^{kl'}\overline{\Sigma}_s(y)^{lk'}\right\} \mathrm{d}y\right] \mathrm{d}s.$$

The integrand of the above expression is nothing but the $\mathcal{F}^{(0)}$-conditional covariance between the (k,l)-th and (k',l')-th entries of the variable $\overline{\Sigma}_s(y)^{1/2}\zeta$, where ζ is a d^2-dimensional standard normal variable independent of $\mathcal{F}^{(0)}$. In other words, $\operatorname{vec}(\mathcal{W}_t)$ is centered Gaussian with covariance matrix $\int_0^1 \mathfrak{S}_s \mathrm{d}s$, where $\mathfrak{S}_s = \int_0^1 (\overline{\Sigma}_s(y) \otimes \overline{\Sigma}_s(y)) \operatorname{Cov}[\operatorname{vec}(\zeta\zeta^*)]\mathrm{d}y$ and, vec and $\otimes$ denote the vec-operator and the Kronecker product of matrices, respectively (cf. Section 2.2 of [9]). In particular, the process $\mathfrak{S}_s$ is càdlàg, $(\mathcal{F}^{(0)}_t)$-adapted and takes values in the set of $d \times d$ positive semidefinite matrices, hence we can construct the process $\mathcal{W}$ stated as in the theorem by Proposition 4.1.2 of [31]. More precisely, $\mathcal{W}$ can be realized as $\operatorname{vec}(\mathcal{W}_t) = \int_0^t \mathfrak{S}^{1/2}_s \mathrm{d}W'_s$, where W' is a d^2-dimensional standard Brownian motion defined on an extension of $\mathcal{B}^{(0)}$ and independent of $\mathcal{F}^{(0)}$.

Note that the Fisher information matrix for covariance matrix estimation of a multivariate diffusion process from non-synchronous and noisy observations is *not* analogous to that for a pure diffusion setting; see Section 2.2 of [9] for details.

Remark 3.9 (Covariance structure of $\mathcal{Z}_t$)**.** $\mathcal{Z}_t$ apparently has an analogous covariance structure to the asymptotic covariance of the realized covariance due to jumps in the regular sampling case (cf. Eq.(5.4.4) of [31]), and it can be realized as follows. Set $A_m = \{z : \gamma(z) > 1/m\}$ for each $m \in \mathbb{N}$, and denote by $(S(m,j))_{j\geq 1}$ the successive jump times of the Poisson process $1_{A_m \setminus A_{m-1}} \star \mu$. Let $(S_r)_{r\geq 1}$ be a reordering of the double sequence $(S(m,j))$. Suppose that sequences $(\Psi_{r-})_{r\geq 1}$ and $(\Psi_{r+})_{r\geq 1}$ of i.i.d. standard d'-dimensional normal variables and sequences $(\Psi'_{r-})_{r\geq 1}$ and $(\Psi'_{r+})_{r\geq 1}$ of i.i.d. standard d-dimensional normal variables are defined on an extension of $\mathcal{B}^{(0)}$ and that all of them are mutually independent and independent of $\mathcal{F}^{(0)}$. Now, the variable $\widetilde{\Upsilon}_s$ defined in Remark 3.8 is positive semi-definite, it admits the (positive semi-definite) square root $\widetilde{\upsilon}_s := \Upsilon^{1/2}_s$. Since the process $\widetilde{\Upsilon}$ is càdlàg and $(\mathcal{F}^{(0)}_t)$-adapted, so is $\widetilde{\upsilon}_s$. Then $\mathcal{Z}$ is realized as $\mathcal{Z}_t = \sum_{r:S_r\leq t}(\mathfrak{Z}_r + \mathfrak{Z}^*_r)$, where

$$\mathfrak{Z}_r = \frac{1}{\psi_2}\Delta X_{S_r}\left\{\sqrt{\Phi_{22}\theta}\left(\sigma_{S_r-}\sqrt{G_{S_r-}}\Psi_{r-} + \sigma_{S_r}\sqrt{G_{S_r}}\Psi_{r+}\right) + \sqrt{\frac{\Phi_{12}}{\theta}}\left(\widetilde{\upsilon}_{S_r-}\Psi'_{r-} + \widetilde{\upsilon}_{S_r}\Psi'_{r+}\right)\right\}^*.$$

This is indeed the desired one; see Proposition 4.1.4 of [31].

Remark 3.10 (Comparison with a pure semimartingale setting)**.** It would be interesting to observe how our result is different from Bibinger and Vetter [12]'s one in a pure semimartingale setting. For simplicity we focus on the univariate case, i.e. we assume that $d = d' = 1$, and assume that $T_p = t^1_p$ for every p for notational simplicity. Now let us recall their result briefly. Suppose that b, σ and δ are continuous. Suppose also that the sequence (T_p) is independent of $b, \sigma, \delta, W, \mu$ and satisfies (3.1). Then, according to Theorem 3.5 of [12], for any $t > 0$ we have the following convergence:

$$\sqrt{n}([X,X]^n_t - [X,X]_t) \to^{d_s} \sqrt{2}\int_0^t \sigma^2_s\sqrt{H'(s)}\mathrm{d}W'_s + 2\sum_{r:S_r\leq t}\Delta X_{S_r}\sigma_{S_r}\sqrt{\eta(S_r)}\Psi_r, \tag{3.5}$$

where W' is a standard Brownian motion, H is a (possibly random) C^1 function such that $n\sum_{p:T_p\leq t}(T_p-T_{p-1})^2\to^p H(t)$ for every $t\in\mathbb{R}_+$ (the existence is assumed), $(S_r)_{r\geq 1}$ is a sequence of stopping times exhausting the jumps of X, $(\Psi_r)_{r\geq 1}$ is a sequence of i.i.d. standard normal variables, and $(\eta(t))_{t\in\mathbb{R}_+}$ is a family of independent random variables with uniformly bounded first moments, and such that the processes $(n(T_+(t)-T_-(t)))_{t\in\mathbb{R}_+}$ converge finite-dimensionally in law to $(\eta(t))_{t\in\mathbb{R}_+}$ (the existence is assumed, and this condition can be weakened; see Assumption 3.1 of [12] for details). Here, $T_+(t)=\min\{T_p:T_p\geq t\}$ and $T_-(t)=\max\{T_p:T_p\leq t\}$ for any $t\in\mathbb{R}_+$ and W', (Ψ_r) and $(\eta(t))$ are defined on an extension of $\mathcal{B}$ and mutually independent as well as independent of $\mathcal{F}$. On the other hand, provided that $\Upsilon\equiv 0$ (so the noise is absent), the corresponding result to our estimator can be written as follows:

$$n^{1/4}(\mathrm{MRC}[Y]^n_t-[X,X]_t)\to^{d_s}\frac{\sqrt{2\Phi_{22}\theta}}{\psi_2}\left(\sqrt{2}\int_0^t\sigma_s^2\sqrt{G_s}\mathrm{d}W'_s+2\sum_{r:S_r\leq t}\Delta X_{S_r}\sigma_{S_r}\sqrt{G_{S_r}}\Psi_r\right),\tag{3.6}$$

where we also assume that G is continuous for simplicity. Compared with the above equation with (3.5), the quantities H' and η coming from the irregularity of the observation times in the left hand of (3.5) are replaced with G in (3.6). Since the quantity H contains the information of the second moments of the durations and η contains that of all the moments of the durations around the jump times, the distributional future of the durations strongly affects the asymptotic distribution in (3.5). In contrast, the first moments of the durations only affect the asymptotic distribution in (3.6).

Remark 3.11 (Comparison with the continuous case)**.** The result of the theorem is not new if X is continuous. In fact, in the case that X is continuous, a central limit theorem for the MRC estimator can be derived with a somewhat weaker assumption on the limiting process G; see Theorem 3.1 and assumption [A4] of Koike [36] for details. In the discontinuous case, we need some regularity of the path of the left limit process G_- to verify the approximation given in Proposition 6.6, so the structural assumption [A4](iii-c) is necessary.

It is worth mentioning that the structural assumption on G is necessary to deal with the irregularity of observation times in the discontinuous case. In contrast, such a condition is only required to handle the time endgeneity in the continuous case. In fact, if the observation times have a kind of pre-determination property (the so-called *strong predictability*), convergence in probability of G^n to G for the Skorokhod topology is sufficient to derive a central limit theorem; see Koike [35] for details.

Remark 3.12 (Feasible limit theorem)**.** In order to apply Theorem 3.1 to real statistical problems such as the construction of confidence intervals, we need an estimator for the asymptotic covariance matrix given by (3.3) and (3.4). This will be achieved by combining the technique used in the non-synchronously observed diffusion setting (e.g. a kernel approach of [24] or a histogram-type method of [8]) with the one used in the jump diffusion setting (e.g. a thresholding and locally averaging method of [2]). Or we can presumably use an estimator of Aït-Sahalia and Xiu [3] for the equidistant sampling setting without modification because the distribution of the variable $n^{1/4}\overline{Y}_i$ is, roughly speaking, approximated by the d-dimensional normal variable with mean 0 and covariance matrix $\theta\psi_2\Sigma_{T_i}G_{T_i}+\frac{\psi_1}{\theta}\widetilde{\Upsilon}_{T_i}$ in the absence of jumps conditionally on $\mathcal{F}^{(0)}_{T_i}$, where $\widetilde{\Upsilon}$ is the same one as in Remark 3.8 (this is theoretically manifested by Lemma 6.7 in a sense).

4 Examples of the observation times

In this section we give some illustrative examples of the observation times that satisfy the condition [A1]. We shall start to discuss univariate examples (i.e. we assume that $d=1$), which are not encompassed with the restricted discretization schemes.

Example 4.1. As an illustrative example of endogenous observation times, we consider a simple model generated by hitting times of the underlying Brownian motion W. This type of model is commonly used in the literature; see

[18, 37, 48] among others. Here we especially focus on a simpler version of the specification from [48]. Specifically, t^1_i's are defined as follows:

$$t^1_0 = 0, \qquad t^1_{i+1} = \inf\left\{t > t^1_i : W_t - W_{t^1_i} + \sqrt{n}\mathfrak{a}_{t^1_i}(t - t^1_i) = \mathfrak{b}_{t^1_i}/\sqrt{n}\right\},$$

where $\mathfrak{a}$ and $\mathfrak{b}$ are càdlàg $(\mathcal{F}^{(0)}_t)$-adapted processes such that $\mathfrak{a}_t\mathfrak{b}_t > 0$ and $\mathfrak{a}_{t-}\mathfrak{b}_{t-} > 0$ for every t. In this case [A1] is satisfied with setting $T_p = \tau^1_p = t^1_p$ for every p, as long as $G := \mathfrak{b}/\mathfrak{a}$ satisfies [A1](iii-c). In fact, noting that, conditionally on $\mathcal{F}^{(0)}_{T_p}$, $n(T_{p+1} - T_p)$ follows the inverse Gaussian distribution with mean G_{T_p} and variance $G^2_{T_p}/\mathfrak{a}_{T_p}$, (3.1) holds true for any $t > 0$ and $\xi \in (0, 1)$. Moreover, we have $E[n(T_{p+1} - T_p)|\mathcal{G}^{(0)}_{T_p}] = G_{T_p}$ for every p because W is independent of μ. Hence [A1](i)-(iii) are satisfied with $\mathcal{N}^n = \emptyset$. Finally, [A1](iv) is automatically satisfied.

Example 4.2. We can also accommodate observation times generated by hitting times of a Brownian motion plus finitely many jumps to our situation. For example, let us consider the observation times defined as follows:

$$t^1_0 = 0, \qquad t^1_{i+1} = \inf\left\{t > t^1_i : W_t - W_{t^1_i} + \sqrt{n}\mathfrak{a}_{t^1_i}(t - t^1_i) + \delta' \star \mu_t - \delta' \star \mu_{t^1_i} = \mathfrak{b}_{t^1_i}/\sqrt{n}\right\},$$

where $\mathfrak{a}$ and $\mathfrak{b}$ are the same one as in Example 4.1 and δ' is an $(\mathcal{F}^{(0)}_t)$-optional real-valued function on $\Omega^{(0)} \times \mathbb{R}_+ \times E$ such that $1_{\{\delta' \neq 0\}} \star \mu_t < \infty$ for all t. Therefore, the process $\delta' \star \mu$ has finitely many jumps. Then, it can easily been seen that [A1] is satisfied in this case under the same situation as that of Example 4.1, except for setting $\mathcal{N}^n = \{p \in \mathbb{Z}_+ : \delta' \star \mu_{T_{p+1}} - \delta' \star \mu_{T_p} > 0\}$.

Example 4.3. Let us consider the observation times discussed in Example 3.4 of Bibinger and Vetter [12]. Namely, $t^1_i = i/n$ if i is even and $t^1_i = (i + \alpha)/n$ if i is odd, where $\alpha \in (0, 1)$ is a constant. [12] showed that this observation times produce an additional randomness in the asymptotic distribution of the realized covariance estimator even though they are deterministic. In fact, in this case the variable $\eta(t)$ in (3.5) takes the values $(1 + \alpha)$ and $(1 - \alpha)$ with probabilities $(1 + \alpha)/2$ and $(1 - \alpha)/2$, respectively. On the other hand, setting $T_p = (p + 1)/n$ and $\tau^1_p = t^1_p$, [A1] is satisfied. Hence in our case this example has the same impact as that of the regular observation times on the asymptotic distribution.

Next we turn to the multivariate and non-synchronous examples. As the data synchronization method, we focus on the refresh sampling method.

Example 4.4. We shall discuss the Poisson sampling, which is one of the most popular models in this area; see e.g. [8, 12, 23, 53]. Let (t^k_i) be a sequence of Poisson arrival times with the intensity np_k for each k and suppose that $(t^1_i), \dots, (t^d_i)$ are mutually independent and independent of X and ϵ. Then, [A1] is satisfied with

$$G_s \equiv \sum_{k=1}^d \sum_{1 \leq l_1 < \cdots < l_k \leq d} \frac{(-1)^{k-1}}{p_{l_1} + \cdots p_{l_k}}, \qquad \chi^{kl}_s \equiv \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{otherwise.} \end{cases}$$

Example 4.5. Here we give an example of observation times which are possibly endogenous and satisfy [A1] with the explicit G and χ. More precisely, we give a continuous time analog of the Lo-MacKinlay model of [39].

Let $(\tau_i)_{i=0}^\infty$ be a sampling scheme and suppose that $\sup_{i \geq 0}(\tau_i \wedge t - \tau_{i-1} \wedge t) = o_p(n^{-\xi})$ as $n \to \infty$ for any $t > 0$ and $\xi \in (0, 1)$. For each $k = 1, \dots, d$, let $(M_k(n, i))_{i=0}^\infty$ be a sequence of $\mathbb{Z}_+$-valued variables defined on $\mathcal{B}^{(0)}$ and independent of X and (τ_i) such that $\mathcal{M}_k := (M_k(n, i + 1) - M_k(n, i))_{i=0}^\infty$ is independent and geometrically distributed with the common success probability $p_k \in (0, 1)$. Moreover, suppose that, for each i, $M_k(n, i)$ is an $(\mathcal{F}^{(0)}_{\tau_j})_{j=0}^\infty$-stopping time so that $t^k_i := \tau_{M_k(n,i)}$ is an $(\mathcal{F}^{(0)}_t)$-stopping time and that $M_k(n, i + 1) - M_k(n, i)$ is independent of $\mathcal{F}^{(0)}_{t^k_i}$. Finally, assume that $\mathcal{M}_1, \dots, \mathcal{M}_d$ are mutually independent and that [A1](i)–(iv) are satisfied with replacing (T_i) by (τ_i). Then it can easily be shown that [A1] holds true with

$$G_s \equiv \sum_{k=1}^d \sum_{1 \leq l_1 < \cdots < l_k \leq d} \frac{(-1)^{k-1} G^0_s}{1 - (1 - p_{l_1}) \cdots (1 - p_{l_k})} \qquad \text{and} \qquad \chi^{kl}_s \equiv \begin{cases} 1 & \text{if } k = l, \\ p_k p_l/(p_k + p_l - p_k p_l) & \text{otherwise.} \end{cases}$$

Here, G^0 denotes the asymptotic conditional expected duration process corresponding to (τ_i). By taking an endogenous sampling scheme as the underlying sampling scheme (τ_i), we can obtain endogenous observation times.

5 Simulation study

In this section we assess the finite sample accuracy of the central limit theory developed in this paper and confirm our theoretical findings via Monte Carlo experiments.

We simulate over the unit interval $[0,1]$, and basically follow the design of Aït-Sahalia and Xiu [3]. To simulate the latent semimartingale X, the following bivariate Heston model with jumps is considered:

$$\mathrm{d}X_t^k = \sigma_{k,t}\mathrm{d}W_t^k + \mathrm{d}Z_t^k, \qquad \mathrm{d}\sigma_{k,t}^2 = \kappa_k(\bar{\sigma}_k^2 - \sigma_{k,t}^2)\mathrm{d}t + s_k\sigma_{k,t}\mathrm{d}B_t^k + \mathrm{d}J_t^k - \lambda_k^V\tau_k^V\mathrm{d}t, \qquad k=1,2.$$

Here, W^1, W^2, B^1, B^2 are correlated standard Brownian motions such that

$$\mathrm{d}[W^1,W^2]_t = \rho_B\mathrm{d}t, \qquad \mathrm{d}[W^k,B^k]_t = \rho_k\mathrm{d}t, \qquad \mathrm{d}[W^1,B^2]_t = \mathrm{d}[W^2,B^1]_t = \mathrm{d}[B^1,B^2]_t = 0.$$

J^k is a compound Poisson process with jump size uniformly distributed on $[0, 2\tau_k^V]$ and jump intensity λ_k^V. J^1 and J^2 are assumed to be mutually independent. Z^k is a pure jump Lévy process specified as follows. First, Z^2 is linearly correlated with Z^1 as $Z^2 = \rho_J Z^1 + \sqrt{1-\rho_J^2}Z^0$, where Z^0 is another Lévy process independent of Z^1. For each $m=0,1$, Z^m is a CGMY process with Lévy density given by

$$f_m(x) = c_m\frac{e^{-\gamma_{m-}|x|}}{|x|^{1+\beta_m}}1_{\{x<0\}} + c_m\frac{e^{-\gamma_{m+}x}}{x^{1+\beta_m}}1_{\{x>0\}}.$$

The parameter values of the stochastic volatility processes used in the simulation are reported in Table 1. The initial value for the volatility processes $\sigma_{k,t}^2$ is set at $\bar{\sigma}_k^2$ for each $k=1,2$, which ensures that $E[\sigma_{k,t}^2] = \bar{\sigma}_k^2$ for all $t \in [0,1]$. The specification of the parameters in the CGMY processes is as follows. We set $\gamma_{m+}=3, \gamma_{m-}=5, \beta_m=0.5$ for every $m=0,1$. c_1 is selected such that the quadratic variation contributed by jumps in X^1 amounts to 15% in expectation, i.e. $E([Z^1,Z^1]_1)/E([X^1,X^1]_1) = 0.15$. Then c_0 is selected such that $E([Z^2,Z^2]_1)/E([X^2,X^2]_1) = 0.15$. Finally, the correlation parameter ρ_J between the jump processes are set at 0.2. Note that Z^1 and Z^2 can be exactly simulated because we only consider the situation where they are of finite variation; see e.g. [33] for details.

To generate observation times, we consider Lo-MacKinlay type sampling schemes illustrated in Example 4.5. Two kinds of sequence $(\tau_i)_{i=0}^\infty$ of latent observation times are considered: One is the equidistant sampling scheme $\tau_i = i/n$ and the other is the endogenous sampling scheme defined by

$$\tau_0 = 0, \qquad \tau_{i+1} = \inf\{t > \tau_i : W_t^1 - W_{\tau_i}^1 - 2\sqrt{n}(t-\tau_i) = -2/\sqrt{n}\}, \qquad i = 0,1,\dots, \tag{5.1}$$

where we set $n = 23,400$. Note that in the latter case the sequence $(\tau_{i+1}-\tau_i)_{i=0}^\infty$ is independent and identically distributed with the inverse Gaussian distribution with mean $1/n$ and variance $4/n^2$, thus we can exactly simulate τ_i's (and construct the exactly discretized path $\{W_{\tau_i}\}$ from $\{\tau_i\}$). Furthermore, in both cases the corresponding conditional expected duration processes G^0 are identical with 1. The parameters p^1 and p^2 from Example 4.5, which denote the probabilities of observations occurring, are assumed to be identical each other and varied thorough $1/3, 1/5, 1/10$ and $1/30$.

In constructing noisy prices Y, we first generate a discretized path $X_{\tau_0}, X_{\tau_1}, \dots$ of X using a standard Euler scheme. After that, we add simulated microstructure noise $Y_{\tau_i} = X_{\tau_i} + \epsilon_{\tau_i}$ by generating centered Gaussian i.i.d. variables $\epsilon_{\tau_0}^k, \epsilon_{\tau_1}^k, \dots$ with standard deviation 0.005. ϵ^1 and ϵ^2 are assumed to be mutually independent. Simulation results are based on 10,000 Monte Carlo iterations for each scenario.

Following [14], the MRC estimator is implemented using the weight function $g(x) = x \wedge (1-x)$ and the refresh time sampling method (the finite sample corrections explained in [14] are also included). We consider the window size k_n of the form $k_n = \lceil\theta\sqrt{N_1^n}\rceil$, and θ is selected among 1/3 and 1. The former value of θ corresponds to the

Table 1: The parameters of the stochastic volatility processes

k	κ_k	s_k	$\bar{\sigma}_k$	ρ_k	λ_k^V	τ_k^V	ρ_B
1	5	0.3	0.25	-0.6	5	0.05	0.5
2	4	0.4	0.3	-0.75	10	0.01	—

one used in Jacod *et al.* [28], while the latter one does to the one used in Christensen *et al.* [14]. We assess the accuracy of the standard normal approximation of the infeasible standardized statistic

$$n^{1/4}\frac{\mathrm{MRC}[Y]_1^{n,12} - [X^1, X^2]_1}{\sqrt{\mathbf{AVAR}}}, \tag{5.2}$$

where **AVAR** is the theoretical asymptotic variance given in Theorem 3.1. Table 2 reports the sample mean and standard deviation as well as 95% and 99% coverages of (5.2). As the table reveals, the central limit theorem for (5.2) fairly works. As was expected from the theory developed in the above, we find no significant difference of the results between the exogenous and the endogenous sampling cases. At relatively low frequencies like $p^1 = p^2 = 1/10$ or $1/30$, the results for $\theta = 1$ show the better performance than those for $\theta = 1/3$. This would be because k_n is not sufficiently large in such a situation, in order to work the averaging effect of the pre-averaging procedure explained in Remark 3.7.

Table 2: Simulation results of the standardized estimates

	$\tau_i = i/n$				τ_i's are defined by (5.1)			
	Mean	SD	Coverage (95%)	Coverage (99%)	Mean	SD	Coverage (95%)	Coverage (99%)
$\theta = 1/3$								
$p^1 = p^2 = 1/3$	-0.00	1.01	0.949	0.987	-0.00	1.02	0.948	0.987
$p^1 = p^2 = 1/5$	-0.01	1.02	0.946	0.987	-0.01	1.04	0.943	0.986
$p^1 = p^2 = 1/10$	-0.01	1.05	0.939	0.983	-0.01	1.06	0.937	0.984
$p^1 = p^2 = 1/30$	-0.03	1.09	0.928	0.979	-0.03	1.10	0.929	0.980
$\theta = 1$								
$p^1 = p^2 = 1/3$	-0.01	1.01	0.948	0.987	-0.01	1.01	0.949	0.989
$p^1 = p^2 = 1/5$	-0.01	1.01	0.948	0.987	-0.01	1.01	0.951	0.987
$p^1 = p^2 = 1/10$	-0.02	1.02	0.947	0.986	-0.02	1.02	0.948	0.987
$p^1 = p^2 = 1/30$	-0.03	1.03	0.946	0.985	-0.03	1.03	0.943	0.985

Note. We report the sample mean, standard deviation (SD) as well as the 95% and 99% coverages of the standardized statistics (5.2) included in the simulation study.

6 Proof of Theorem 3.1

6.1 Preliminaries

6.1.1 Localization

Before starting the proof, we strengthen our assumptions [A1]–[A3] by localization procedures. First, a standard localization procedure, described in detail in Lemma 4.4.9 of [31], for instance, allows us to replace the conditions [A2] and [A3] by the following strengthened versions, respectively:

[SA2] We have [A2], and the processes X_t, b_t, σ_t, $\widetilde{b}_t$ and $\widetilde{\sigma}_t$ are bounded. Also, b_t is $(\mathcal{H}_t^\wedge)$-progressively measurable and σ_t is $(\mathcal{H}_t^\wedge)$-adapted. Moreover, there are a constant Λ and a non-negative bounded function γ on E such that $\int \gamma(z)^2\lambda(\mathrm{d}z) < \infty$ and $\|\delta(\omega^{(0)}, t, z)\| \vee \|\widetilde{\delta}(\omega^{(0)}, t, z)\| \leq \gamma(z)$ and

$$E\left[\|b_{t_1} - b_{t_2}\|^2 | \mathcal{F}_{t_1\wedge t_2}\right] \leq \Lambda E\left[|t_1 - t_2|^{\varpi} | \mathcal{F}_{t_1\wedge t_2}\right],$$
$$E\left[\|\delta(t_1, z) - \delta(t_2, z)\|^2 | \mathcal{F}_{t_1\wedge t_2}\right] \leq \Lambda\gamma(z)^2 E\left[|t_1 - t_2|^{\varpi} | \mathcal{F}_{t_1\wedge t_2}\right]$$

for any bounded $(\mathcal{F}_t^{(0)})$-stopping times t_1 and t_2.

[SA3] There are a constant $\Gamma > 4$ and a constant Λ' such that the process $\int \|z\|^\Gamma Q_t(\mathrm{d}z)$ is bounded and

$$E\left[\|\Upsilon_{t_1} - \Upsilon_{t_2}\|^2 | \mathcal{F}_{t_1\wedge t_2}\right] \leq \Lambda' E\left[|t_1 - t_2|^{\varpi} | \mathcal{F}_{t_1\wedge t_2}\right]$$

for any bounded $(\mathcal{F}_t^{(0)})$-stopping times t_1 and t_2. Moreover, Υ_t is càdlàg and $(\mathcal{H}_t^\wedge)$-adapted.

Next we introduce a strengthened version of [A1]. In the following we fix a constant $\xi \in (0, 1)$ such that

$$\xi > \frac{7}{8} \vee \frac{1}{2}\left(\kappa + \frac{3}{2}\right) \vee (1 - \varpi), \tag{6.1}$$

and we set $\bar{r}_n = n^{-\xi}$.

[SA1] We have [A1], and for every n it holds that

$$\sup_{p\geq 0}(T_p - T_{p-1}) \leq \bar{r}_n. \tag{6.2}$$

The following lemma allows us to replace [A1] by [SA1] via another localization argument. The proof is similar to that of Lemma 6.3 from [36], so we omit it.

Lemma 6.1. *Assume* [A1]. *One can find sampling schemes* $(\widetilde{T}_p)$ *and* $(\widetilde{\tau}_p^k)$ $(k = 1, \dots, d)$ *satisfying the following conditions:*

(i) $(\widetilde{T}_p)$ *and* $(\widetilde{\tau}_p^k)$ *satisfy* [SA1] *with the same limiting processes* G *and* χ *as those of the original sampling schemes.*

(ii) For any $t > 0$ *there is a subset* $\Omega_{n,t}^{(0)}$ *of* $\Omega^{(0)}$ *such that* $\lim_n P^{(0)}(\Omega_{n,t}^{(0)}) = 1$. *Moreover, on* $\Omega_{n,t}^{(0)}$ *we have* $T_p \wedge t = \widetilde{T}_p \wedge t$ *and* $\tau_p^k \wedge t = \widetilde{\tau}_p^k \wedge t$ *for all* k, p.

6.1.2 Outline of the proof

Here we give a brief description of the scheme of the proof. First, for the proof it is convenient to realize the processes $\mathcal{W}$ and $\mathcal{Z}$ on an extension of $\mathcal{B}^{(0)}$ as in Remarks 3.8–3.9 (so we will use the notation introduced in these remarks in the following). For notational simplicity, we use the same letters P and E for the probability and the expectation with respect to this extension.

Next we introduce some notation. We denote by $\mathcal{R}_m$ the set of all indices r such that $S_r = S(m', j)$ for some $j \geq 1$ and some $m' \leq m$. Also, we set

$$\begin{cases} b(m)_t = b_t - \int_{A_m\cap\{z:|\delta(t,z)|\leq 1\}} \delta(t, z)\lambda(\mathrm{d}z), \quad B(m)_t = \int_0^t b(m)_s \mathrm{d}s, \quad M_t = \int_0^t \sigma_s \mathrm{d}W_s, \\ C(m)_t = X_0 + B(m)_t + M_t, \quad J(m)_t = \delta 1_{A_m} \star \mu_t, \quad X(m)_t = C(m)_t + J(m)_t, \\ Z(m)_t = X_t - X(m)_t = \delta 1_{A_m^c} \star (\mu - \nu)_t. \end{cases}$$

These processes are well-defined under [SA2]. Furthermore, set $I_p = [T_{p-1}, T_p)$ for every $p \in \mathbb{Z}_+$. On the other hand, for any process V and any (random) interval $I = [S, T)$, we define the random variable $V(I)$ by $V(I) = V_T - V_S$. We also set $I(t) = I \cap [0, t) = [S \wedge t, T \wedge t)$ for any $t \in \mathbb{R}_+$ and $|I| = T - S$. For any real-valued function u on $[0, 1]$, we set $u_p^n = u(p/k_n)$ for $p = 0, 1, \dots, k_n$. For any d-dimensional processes U, V, any $k, l \in \{1, \dots, d\}$ and any $u, v \in \{g, g'\}$, we define the process $\Xi_{u,v}^{(k,l)}(U, V)^n$ by

$$\Xi_{u,v}^{(k,l)}(U, V)_t^n = \frac{1}{\psi_2 k_n} \sum_{i=1}^{N_t^n - k_n + 1} \overline{U}(u)_i^k \overline{V}(v)_i^l, \qquad t \in \mathbb{R}_+,$$

where $\overline{U}(u)^k_i = \sum_{p=0}^{k_n-1} u^n_p U^k(I_{i+p})$ and $\overline{V}(v)^l_i$ is defined analogously. Moreover, we define the d-dimensional process $\mathfrak{E}$ by

$$\mathfrak{E}^k_t = -\frac{1}{k_n}\sum_{p=1}^{\infty} \epsilon^k_{\tau^k_p} 1_{\{\tau^k_p \leq t\}}, \qquad t \in \mathbb{R}_+, \quad k = 1, \dots, d.$$

It can easily been seen that $\mathfrak{E}$ is a purely discontinuous locally square-integrable martingale on $\mathcal{B}$ under [SA3]. Finally, for any d-dimensional process V we define the $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process $\mathbf{\Xi}[V]^n$ by

$$\mathbf{\Xi}[V]^{n,kl} = \Xi^{(k,l)}_{g,g}(V,V)^n + \Xi^{(k,l)}_{g,g'}(V,\mathfrak{E})^n + \Xi^{(l,k)}_{g,g'}(V,\mathfrak{E})^n + \Xi^{(k,l)}_{g',g'}(\mathfrak{E},\mathfrak{E})^n, \qquad k,l = 1,\dots,d.$$

Now we turn to the outline of the proof. In the first step we show that the errors from the interpolations to the synchronized sampling times are asymptotically negligible:

Proposition 6.1. *Assume* [SA1]–[SA3] *and* [A4](ii)*. Then* $n^{1/4}\left(\mathrm{MRC}[Y]^n - \mathbf{\Xi}[X]^n + \frac{\psi_1}{\psi_2 k_n^2}[Y,Y]^n\right) \xrightarrow{ucp} 0$.

The proof of this proposition is an easy extension of that of Proposition 6.1 from [36], so we omit it.

In the next step we decompose the quantity $\mathbf{\Xi}[X]^n$ as $\mathbf{\Xi}[X]^n = \mathbf{\Xi}[X(m)]^n_t + (\mathbf{\Xi}[X]^n - \mathbf{\Xi}[X(m)]^n)$ for each m, and show that the first term enjoys a central limit theorem for any fixed m and the second term is negligible as $m \to \infty$. More precisely, we prove the following propositions:

Proposition 6.2. *Suppose that* [SA1]–[SA3] *and* [A4] *are satisfied. Then*

$$n^{1/4}\left(\mathbf{\Xi}[X(m)]^n_t - [X(m),X(m)]_t - \frac{\psi_1}{\psi_2 k_n^2}[Y,Y]^n_t\right) \to^{d_s} \mathcal{W}_t + \mathcal{Z}(m)_t$$

as $n \to \infty$ for any $t > 0$ and any $m \geq 1$, where $\mathcal{Z}(m)_t = \sum_{r \in \mathcal{R}_m : S_r \leq t}(\mathfrak{Z}_r + \mathfrak{Z}^*_r)$.

When further X is continuous, the processes $n^{1/4}(\mathbf{\Xi}[X]^n - [X,X] - \frac{\psi_1}{\psi_2 k_n^2}[Y,Y]^n)$ converge stably in law to the process $\mathcal{W}$ for the Skorokhod topology.

Proposition 6.3. *Suppose that* [SA1]–[SA3] *and* [A4](ii) *are satisfied. Then*

$$\mathcal{Z}(m)_t \to^{d_s} \mathcal{Z}_t \tag{6.3}$$

as $m \to \infty$ and

$$\limsup_{m\to\infty}\limsup_{n\to\infty} P\left(n^{1/4}\left\|\mathbf{\Xi}[X]^n_t - \mathbf{\Xi}[X(m)]^n_t\right\| > \eta\right) = 0 \tag{6.4}$$

for any $t, \eta > 0$.

Combining Propositions 6.1–6.3 with Proposition 2.2.4 of [31], we obtain Theorem 3.1.

6.2 Proof of Proposition 6.2

Throughout the discussions, for (deterministic) sequences (x_n) and (y_n), $x_n \lesssim y_n$ means that there is a (non-random) constant $K \in [0,\infty)$ such that $x_n \leq K y_n$ for large n. We also denote by E_0 the conditional expectation given $\mathcal{F}^{(0)}$, i.e. $E_0[\cdot] = E[\cdot|\mathcal{F}^{(0)}]$.

The proof of Proposition 6.2 is divided into the following steps:

(i) Approximating the estimation error due to the diffusion part by a more tractable one,
(ii) Proving a central limit theorem for the approximation constructed in (i),
(iii) Approximating the estimation error due to the jump part by a more tractable one (Section 6.2.1),
(iv) Proving a local stable convergence result corresponding to Lemma 16.3.7 of [31] (Section 6.2.2),
(v) Proving a joint limit theorem for the pair of the above approximations and completing the proof of the proposition (Section 6.2.3).

The first two steps have already been carried out in [36]. For the later use, we summarize the result in the following. We begin by introducing some notation. For any d-dimensional processes U, V, any $k, l \in \{1, \dots, d\}$ and any real-valued functions u, v on $[0,1]$, we define the processes $\mathbb{M}^{(k,l)}_{u,v}(U,V)^n$ and $\mathbb{L}^{(k,l)}_{u,v}(U,V)^n$ by

$$\mathbb{M}^{(k,l)}_{u,v}(U,V)^n_t = \sum_{q=2}^{N^n_t+1} C^n_{u,v}(U)^k_q V^l(I_q), \qquad \mathbb{L}^{(k,l)}_{u,v}(U,V)^n_t = \mathbb{M}^{(k,l)}_{u,v}(U,V)^n_t + \mathbb{M}^{(l,k)}_{v,u}(V,U)^n_t,$$

where

$$C^n_{u,v}(U)^k_q = \sum_{p=(q-k_n)\vee 1}^{q-1} c^n_{u,v}(p,q) U^k(I_p), \qquad c^n_{u,v}(p,q) = \frac{1}{\psi_2 k_n} \sum_{i=(p\vee q-k_n+1)\vee 1}^{p\wedge q} u^n_{p-i} v^n_{q-i}.$$

Moreover, define the $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process $\mathbf{L}[M]^n$ by

$$\mathbf{L}[M]^{n,kl} = \mathbb{L}^{(k,l)}_{g,g}(M,M)^n + \mathbb{L}^{(k,l)}_{g,g'}(M,\mathfrak{E})^n + \mathbb{L}^{(l,k)}_{g,g'}(M,\mathfrak{E})^n + \mathbb{L}^{(k,l)}_{g',g'}(\mathfrak{E},\mathfrak{E})^n.$$

Then, we have the following results, which are proved as Proposition 6.2 and Eq.(6.7) from [36]:

Proposition 6.4. *Suppose that* [SA1]–[SA3] *and* [A4](ii) *are satisfied. Then*

$$n^{1/4}\left(\mathbf{\Xi}[C(m)]^n - \mathbf{L}[M]^n - [M,M] - \frac{\psi_1}{\psi_2 k_n^2}[Y,Y]^n\right) \xrightarrow{ucp} 0$$

as $n \to \infty$ for any $m \geq 1$.

Proposition 6.5. *Suppose that* [SA1]–[SA3] *and* [A4](ii) *are satisfied. Then, the processes $n^{1/4}\mathbf{L}[M]^n$ converge stably in law to $\mathcal{W}$ for the Skorokhod topology.*

From the next section we start the proofs of the remaining steps.

6.2.1 Approximation of the estimation error due to the jump part

In this subsection we fix $t > 0$ and $m \in \mathbb{N}$, and denote by $\Omega_n(t,m)$ the set on which $k_n - 1 \leq N^n_{S_r-} \leq N^n_t - k_n$ for all $r \in \mathcal{R}_m$ such that $S_r \leq t$. On this set we have

$$\begin{aligned} n^{1/4}\Xi^{(k,l)}_{g,g}(C(m),J(m))^n_t &= \frac{n^{1/4}}{\psi_2 k_n} \sum_{i=0}^{N^n_t - k_n+1} \sum_{p,q=0}^{k_n-1} g^n_p g^n_q C(m)^k(I_{i+p}) J(m)^l(I_{i+q}) \\ &= \sum_{r\in\mathcal{R}_m : S_r \leq t} \left\{\eta_+(n,r)^k + \eta_-(n,r)^k\right\} \Delta X^l_{S_r}, \end{aligned} \tag{6.5}$$

where

$$\begin{cases} \eta_+(n,r) = n^{1/4} \sum_{p=N^n_{S_r-}+1}^{N^n_{S_r-}+k_n} c^n_{g,g}(p, N^n_{S_r-}+1) C(m)(I_p), \\ \eta_-(n,r) = n^{1/4} \sum_{p=(N^n_{S_r-}-k_n+2)_+}^{N^n_{S_r-}} c^n_{g,g}(p, N^n_{S_r-}+1) C(m)(I_p). \end{cases}$$

Similarly, on $\Omega_n(t,m)$ we have

$$n^{1/4}\Xi^{(k,l)}_{g',g}(\mathfrak{E}, J(m))^n_t = \sum_{r\in\mathcal{R}_m : S_r \leq t} \left\{\eta'_+(n,r)^k + \eta'_-(n,r)^k\right\} \Delta X^l_{S_r}, \tag{6.6}$$

where

$$\begin{cases} \eta'_+(n,r)^k = -\frac{n^{1/4}}{k_n} \sum_{p=N^n_{S_r-}+1}^{N^n_{S_r-}+k_n} c^n_{g',g}(p, N^n_{S_r-}+1) \epsilon^k_{\tau^k_p}, \\ \eta'_-(n,r)^k = -\frac{n^{1/4}}{k_n} \sum_{p=(N^n_{S_r-}-k_n+2)_+}^{N^n_{S_r-}} c^n_{g',g}(p, N^n_{S_r-}+1) \epsilon^k_{\tau^k_p}. \end{cases}$$

The aim of this subsection is to approximate $\eta_\pm(n,r)$ and $\eta'_\pm(n,r)$ by more tractable quantities. Here, the major difficulty coming from the irregularity of the observation times is the fact that $N^n_{S_r-} - k_n + 1$ might not be a $(\mathcal{G}^{(0)}_{T_p})_{p=0}^\infty$-stopping time. Therefore, we first "approximate" $N^n_{S_r-} - k_n + 1$ by a $(\mathcal{G}^{(0)}_{T_p})_{p=0}^\infty$-stopping time.

More precisely, set $\underline{S}_r = (S_r - \frac{k_n}{n}\log n)_+$ and $S_r^\dagger = (S_r - \frac{k_n}{n}G^n_{\underline{S}_r} \wedge \log n)_+$. Then, $S_r^\dagger$ is a $(\mathcal{G}^{(0)}_t)$-stopping time (Lemma 6.2) and thus $N^n_{S_r^\dagger}+1$ is a $(\mathcal{G}^{(0)}_{T_p})_{p=0}^\infty$-stopping time, and this variable gives an approximation of $N^n_{S_r-}-k_n+1$ (Lemma 6.4).

Now we can define our tractable approximations of $\eta_\pm(n,r)$ and $\eta'_\pm(n,r)$ as follows. For any non-negative random variable ρ and any real-valued function ϕ on $[0,1]$, we define the d'-dimensional variable $L(\phi,\rho)_n = (L(\phi,\rho)^j_n)_{1\leq j\leq d'}$ and the d-dimensional variable $L'(\phi,\rho)_n = (L'(\phi,\rho)^k_n)_{1\leq k\leq d}$ by

$$L(\phi,\rho)^j_n = n^{1/4}\sum_{w=1}^{k_n-1}\phi^n_w W^j(I_{i(\rho)^n+w}), \qquad L'(\phi,\rho)^k_n = \frac{n^{1/4}}{k_n}\sum_{w=1}^{k_n-1}\phi^n_w \epsilon^k_{\tau^k_{i(\rho)^n+w}}$$

where we set $i(\rho)^n = N^n_\rho+1$ (recall that $\phi^n_w = \phi(w/k_n)$). We also define the function $\widetilde{\phi}$ on $[0,1]$ by $\widetilde{\phi}(x) = \phi(1-x)$. Then we set

$$\begin{cases} z^n_{r-} = \psi_2^{-1}L(\widetilde{\phi_{g,g}}, S_r^\dagger)_n, & z^n_{r+} = \psi_2^{-1}L(\phi_{g,g}, S_r)_n, \\ z'^n_{r-} = -\psi_2^{-1}L'(\widetilde{\phi_{g',g}}, S_r^\dagger)_n, & z'^n_{r+} = -\psi_2^{-1}L'(\phi_{g',g}, S_r)_n. \end{cases}$$

The aim of this subsection is to prove the following proposition:

Proposition 6.6. *Suppose that* [SA1]–[SA3] *and* [A4](ii) *are satisfied. Then*

$$\eta_-(n,r) = \sigma_{S_r^\dagger}z^n_{r-} + o_p(1), \qquad \eta_+(n,r) = \sigma_{S_r}z^n_{r+} + o_p(1), \tag{6.7}$$

$$\eta'_-(n,r) = z'^n_{r-} + o_p(1), \qquad \eta'_+(n,r) = z'^n_{r+} + o_p(1). \tag{6.8}$$

Now we start to justify that the variable $N^n_{S_r^\dagger}+1$ is an appropriate approximation of $N^n_{S_r-}-k_n+1$. In the remainder of this subsection we fix an index $r\in\mathcal{R}_m$ such that $S_r<\infty$.

Lemma 6.2. *Under* [SA1], $S_r^\dagger$ *is a* $(\mathcal{G}^{(0)}_t)$*-stopping time.*

Proof. For any $t\geq 0$, we have $\{S_r^\dagger \leq t\} = \left\{S_r \leq t + (k_n/n)G^n_{\underline{S}_r}\wedge \log n\right\}\cap\{\underline{S}_r\leq t\}$. Therefore, noting that S_r is $\mathcal{G}^{(0)}_0$-measurable, we obtain $\{S_r^\dagger\leq t\}\in\mathcal{G}^{(0)}_t$. □

Lemma 6.3. *Under* [SA1], $\sup_{0<h<h_0}|G_{S_r-} - G_{(S_r-h)_+}| = O_p(\sqrt{h_0})$ *as* $h_0\downarrow 0$.

Proof. Define the processes $G(m)$, $G'(m)$ and $G''(m)$ by $G(m)_t = \int_0^t \widehat{\sigma}_s \mathrm{d}W_s + (\widehat{\delta}1_{A_m^c\cap\{|\widetilde{\delta}|\leq 1\}})\star(\mu-\nu)_t$, $G'(m)_t = (\widehat{\delta}1_{A_m\cap\{|\widetilde{\delta}|\leq 1\}}+\widehat{\delta}1_{\{|\widetilde{\delta}|>1\}})\star\mu_t$, and $G''(m) = G - G(m) - G'(m)$. Since $G'(m)$ is piecewise constant, it is evident that $\sup_{0<h<h_0}|G'(m)_{S_r-} - G'(m)_{(S_r-h)_+}| = O_p(\sqrt{h_0})$. Moreover, since $G''(m)$ is absolutely continuous with a locally bounded derivative, it also holds that $\sup_{0<h<h_0}|G''(m)_{S_r-} - G''(m)_{(S_r-h)_+}| = O_p(\sqrt{h_0})$. On the other hand, let $(\mathcal{G}^{A_m}_t)$ be the smallest filtration containing $(\mathcal{F}^{(0)}_t)$ such that $\mathcal{G}^{A_m}_0$ contains the σ-field generated by the restriction of the measure μ to $\mathbb{R}_+\times A_m$. Then, by Proposition 2.1.10 of [31] $G(m)$ is a locally square integrable martingale with respect to $(\mathcal{G}^{A_m}_t)$ and its predictable quadratic variation is given by $\langle G(m)\rangle = \int_0^t \widehat{\sigma}_s\widehat{\sigma}_s^*\mathrm{d}s + (\widehat{\delta}^2 1_{A_m^c\cap\{|\widetilde{\delta}|\leq 1\}})\star\nu$, and $G(m)_{S_r-} = G(m)_{S_r}$. Since S_r is $\mathcal{G}^{A_m}_0$-measurable, $(S_r-h)_+$ is a $(\mathcal{G}^{A_m}_t)$-stopping time for every $h\geq 0$. Therefore, the Lenglart inequality implies that

$$P\left(\sup_{0\leq h\leq h_0}|h_0^{-1/2}\{G(m)_{S_r} - G(m)_{(S_r-h)_+}\}|^2 > K\right) \leq \frac{K'}{K} + P\left(h_0^{-1}|\langle G(m)\rangle_{S_r} - \langle G(m)\rangle_{(S_r-h_0)_+}| > K'\right)$$

for any $K,K'>0$ and thus a standard localization argument yields $\sup_{0\leq h\leq h_0}|G(m)_{S_r} - G(m)_{(S_r-h)_+}| = O_p(\sqrt{h_0})$. This completes the proof of the lemma. □

Lemma 6.4. *Under* [SA1], $N^n_{S_r-} - N^n_{S_r^\dagger} = k_n + o_p(n^{\frac{1}{2}-\alpha'})$ *for any* $\alpha'\in(0,(\xi-\kappa-\frac{1}{2})\wedge(\xi-\frac{3}{4})\wedge\varpi)$.

Proof. Since $\alpha' < \xi-\kappa-\frac{1}{2}$, by [SA1] we have

$$N^n_{S_r-} - N^n_{S_r^\dagger} = \sum_{p=1}^{N^n_{S_r-}+1}\frac{E\left[n|I_p||\mathcal{G}^{(0)}_{T_{p-1}}\right]}{G^n_{T_{p-1}}}1_{\{T_{p-1}>S_r^\dagger\}} + o_p(n^{1/2-\alpha'}).$$

In particular, from this expression and [SA1], we deduce $N^n_{S_r-} - N^n_{S^\dagger_r} = O_p(\sqrt{n}\log n)$. Therefore, [SA1] yields

$$\sum_{p=1}^{N^n_{S_r-}+1} E\left[\left|n^{\alpha'-\frac{1}{2}}\frac{n|I_p|}{G^n_{T_{p-1}}}1_{\{T^{p-1}>S^\dagger_r\}}\right|^2 \middle| \mathcal{G}^{(0)}_{T_{p-1}}\right] = n^{1+2\alpha'}\bar{r}_n^2 \sum_{p=1}^{N^n_{S_r-}+1}\frac{1}{G^n_{T_{p-1}}}1_{\{T^{p-1}>S^\dagger_r\}} = o_p(1),$$

hence Lemma 2.3 of [18] implies that

$$N^n_{S_r-} - N^n_{S^\dagger_r} = n\sum_{p=1}^{N^n_{S_r-}+1}\frac{|I_p|}{G^n_{T_{p-1}}}1_{\{T^{p-1}>S^\dagger_r\}} + o_p\left(n^{\frac{1}{2}-\alpha'}\right).$$

Now [SA1], Lemma 6.3 and the fact that $\alpha' < \varpi\wedge\frac{1}{4}$ yield

$$n\sum_{p=1}^{N^n_{S_r-}+1}\left(\frac{|I_p|}{G^n_{T_{p-1}}} - \frac{|I_p|}{G^n_{\underline{S}_r}}\right)1_{\{T^{p-1}>S^\dagger_r\}} = n\sum_{p=1}^{N^n_{S_r-}+1}\left(\frac{|I_p|}{G_{T_{p-1}}} - \frac{|I_p|}{G_{\underline{S}_r}}\right)1_{\{T^{p-1}>S^\dagger_r\}} + o_p\left(n^{\frac{1}{2}-\alpha'}\right) = o_p\left(n^{\frac{1}{2}-\alpha'}\right),$$

thus we have

$$N^n_{S_r-} - N^n_{S^\dagger_r} = n\sum_{p=1}^{N^n_{S_r-}+1}\frac{|I_p|}{G^n_{\underline{S}_r}}1_{\{T^{p-1}>S^\dagger_r\}} + o_p\left(n^{\frac{1}{2}-\alpha'}\right) = k_n\frac{G^n_{\underline{S}_r}\wedge\log n}{G^n_{\underline{S}_r}} + o_p\left(n^{\frac{1}{2}-\alpha'}\right).$$

Since $\lim_n P(G^n_{\underline{S}_r} > \log n) = 0$, we obtain the desired result. ∎

Now we proceed to the main body of the proof of Proposition 6.6. Denote by $\Omega_n(m)$ the set on which $|S_{r_1}-S_{r_2}| > (k_n/n)\log n$ for any $r_1, r_2 \in \mathcal{R}_m$ such that $r_1 \neq r_2$ and $S_{r_1}, S_{r_2} < \infty$. Since $S_{r_1} \neq S_{r_2}$ if $r_1 \neq r_2$ and $S_{r_1}, S_{r_2} < \infty$, we have $P(\Omega_n(m)) \to 1$ as $n \to \infty$.

Lemma 6.5. *Under* [SA2], $E[\sup_{\underline{S}_r\leq s<S_r}\|\sigma_s - \sigma_{\underline{S}_r}\|^2; \Omega_n(m)] \lesssim (k_n/n)\log n$.

Proof. Since no jump of the Poisson process $1_{A_m}\star\mu$ occurs in $[\underline{S}_r, S_r)$ on the set $\Omega_n(m)$, we have $\sigma_s = \sigma(m)_s$ for every $s \in [\underline{S}_r, S_r)$ on this set, where

$$\begin{cases} \sigma(m)_s = \sigma_0 + \int_0^s \widetilde{b}(m)_u \mathrm{d}u + \int_0^s \widetilde{\sigma}_u \mathrm{d}W_u + (\widetilde{\delta}1_{A^c_m})\star(\mu-\nu)_s, \\ \widetilde{b}(m)_u = \widetilde{b}_u - \int_{A_m\cap\{z:|\widetilde{\delta}(u,z)|\leq 1\}}\widetilde{\delta}(u,z)\lambda(\mathrm{d}z). \end{cases}$$

On the other hand, by Proposition 2.1.10 of [31] we have $E\left[\sup_{\underline{S}_r\leq s<S_r}\|\sigma(m)_s - \sigma(m)_{\underline{S}_r}\|^2\right] \lesssim \frac{k_n}{n}\log n$, which implies the desired result. ∎

Proof of Proposition 6.6. Throughout the proof we fix a constant α' such that $1-\xi < \alpha' < (\xi-\kappa-\frac{1}{2})\wedge(\xi-\frac{3}{4})\wedge\varpi$. Such an α' exists due to (6.1).

First we prove the first equation of (6.7). Set $\Omega_n = \{N^n_{S_{r-}} - k_n + 1 \geq 0\}$. By the Lipschitz continuity of g,

$$\eta_-(n,r)^k = \frac{n^{1/4}}{\psi_2}\sum_{p=N^n_{S_r-}-k_n+2}^{N^n_{S_r-}}(\phi_{g,g})^n_{N^n_{S_r-}+1-p}M^k(I_p) + o_p(1) = \frac{n^{1/4}}{\psi_2}\sum_{w=1}^{k_n-1}(\widetilde{\phi_{g,g}})^n_w M^k(I_{N^n_{S_r-}+1-k_n+w}) + o_p(1)$$

on Ω_n. On the other hand, noting that we have

$$\sup\{\|M_s - M_r\| : |s-r|\leq h, s,r\in[0,t]\} = O_p\left(\sqrt{h}|\log h|\right)$$

as $h\downarrow 0$ for any $t>0$ due to a representation of a continuous local martingale with Brownian motion and Lévy's theorem on the uniform modulus of continuity of Brownian motion, summation by parts, (6.2) and Lemma 6.4 imply that

$$n^{1/4}\sum_{w=1}^{k_n-1}(\widetilde{\phi_{g,g}})^n_w\left\{M^k(I_{N^n_{S_r-}+1-k_n+w}) - M^k(I_{i(S^\dagger_r)^n+w})\right\}$$

$$
\begin{aligned}
=&n^{1/4}\sum_{w=1}^{k_n-2}\left\{(\widetilde{\phi_{g,g}})^n_w-(\widetilde{\phi_{g,g}})^n_{w+1}\right\}\left(M^k_{T_{N^n_{S_r-}+1-k_n+w}}-M^k_{T_{i(S^\dagger_r)^n+w}}\right)\\
&+n^{1/4}(\widetilde{\phi_{g,g}})^n_{k_n-1}\left(M^k_{T_{N^n_{S_r-}}}-M^k_{T_{i(S^\dagger_r)^n+k_n-1}}\right)-n^{1/4}(\widetilde{\phi_{g,g}})^n_1\left(M^k_{T_{N^n_{S_r-}+1-k_n}}-M^k_{T_{i(S^\dagger_r)^n}}\right)\\
=&o_p\left(n^{1/4}\sqrt{n^{1/2-\alpha'-\xi}\log n}\right)=o_p\left(n^{(1-\xi-\alpha')/2}\sqrt{\log n}\right)=o_p(1)
\end{aligned}
$$

on Ω_n. Since $\lim_n P(\Omega_n)=1$, we conclude that

$$
\eta_-(n,r)^k=\frac{n^{1/4}}{\psi_2}\sum_{w=1}^{k_n-1}(\widetilde{\phi_{g,g}})^n_w M^k(I_{i(S^\dagger_r)+w})+o_p(1).
$$

Next, noting that W is a d'-dimensional $(\mathcal{G}^{(0)}_t)$-Brownian motion (recall that $(\mathcal{G}^{(0)}_t)$ is the smallest filtration containing $(\mathcal{F}^{(0)}_t)$ such that $\mathcal{G}^{(0)}_0$ contains the σ-field generated by μ), we have

$$
n^{1/4}\sum_{w=1}^{k_n-1}(\widetilde{\phi_{g,g}})^n_w M^k(I_{i(S^\dagger_r)+w})-\sum_{j=1}^{d'}\sigma^{kj}_{\underline{S}_r}z^{n,j}_{r-}=n^{1/4}\sum_{j=1}^{d'}\sum_{w=1}^{k_n-1}(\widetilde{\phi_{g,g}})^n_w\int_{T_{i(S^\dagger_r)^n+w-1}}^{T_{i(S^\dagger_r)^n+w}}\left(\sigma^{kj}_s-\sigma^{kj}_{\underline{S}_r}\right)\mathrm{d}W^j_s,
$$

hence the Lenglart inequality implies that it is enough to show that

$$
\Delta_n:=\sqrt{n}\sum_{j=1}^{d'}\sum_{w=1}^{k_n-1}\left|(\widetilde{\phi_{g,g}})^n_w\right|^2E\left[\int_{T_{i(S^\dagger_r)^n+w-1}}^{T_{i(S^\dagger_r)^n+w}}\left(\sigma^{kj}_s-\sigma^{kj}_{\underline{S}_r}\right)^2\mathrm{d}s\big|\mathcal{G}^{(0)}_{T_{i(S^\dagger_r)^n+w-1}}\right]\to^p 0.
$$

Set

$$
\Delta'_n=\sqrt{n}\sum_{j=1}^{d'}\sum_{w=1}^{k_n-1}\left|(\widetilde{\phi_{g,g}})^n_w\right|^2E\left[\int_{T_{i(S^\dagger_r)^n+w-1}}^{T_{i(S^\dagger_r)^n+w}}\left(\sigma^{kj}_s-\sigma^{kj}_{\underline{S}_r}\right)^2\mathrm{d}s\big|\mathcal{G}^{(0)}_{T_{i(S^\dagger_r)^n+w-1}}\right]1_{\left\{T_{i(S^\dagger_r)^n+w-1}\leq S_r\right\}}.
$$

Then, since $\Omega_n(m)\in\mathcal{G}^{(0)}_0$, it holds that

$$
E\left[\Delta'_n;\Omega_n(m)\right]\lesssim\sqrt{n}\sum_{j=1}^{d'}E\left[\int_{T_{i(S^\dagger_r)^n}}^{T_{i(S_r)^n}}\left(\sigma^{kj}_s-\sigma^{kj}_{S^\dagger_r}\right)^2\mathrm{d}s;\Omega_n(m)\right].
$$

Now, Lemma 6.5, the boundedness of σ and (6.2) imply that

$$
E\left[\int_{T_{i(S^\dagger_r)^n}}^{T_{i(S_r)^n}}\left(\sigma^{kj}_s-\sigma^{kj}_{S^\dagger_r}\right)^2\mathrm{d}s;\Omega_n(m)\right]\lesssim\frac{k_n}{n}(\log n)E\left[\sup_{\underline{S}_r\leq s<S_r}\left(\sigma^{kj}_s-\sigma^{kj}_{S^\dagger_r}\right)^2;\Omega_n(m)\right]+\bar{r}_n\lesssim\bar{r}_n,
$$

hence we obtain $E\left[\Delta'_n;\Omega_n(m)\right]\lesssim\sqrt{n}\bar{r}_n=o(1)$. Therefore, the equation $\lim_n P(\Omega_n(m))=1$ and the Chebyshev inequality yield $\Delta'_n=o_p(1)$. On the other hand, the boundedness of σ, (6.2) and Lemma 6.4 imply that

$$
|\Delta_n-\Delta'_n|\lesssim\sqrt{n}\bar{r}_n\sum_{w=1}^{k_n-1}1_{\left\{T_{i(S^\dagger_r)^n+w-1}>S_r\right\}}\leq\sqrt{n}\bar{r}_n\left|N^n_{S^\dagger_r}+k_n-N^n_{S_r}-1\right|=o_p(n^{1-\xi-\alpha'})=o_p(1).
$$

Consequently, we obtain $\Delta_n=o_p(1)$ and the first equation of (6.7) has been proved. On the other hand, noting that $N^n_{S_r}-N^n_{S_r-}\leq 1$ and S_r is an $(\mathcal{F}^{(0)}_t)$-stopping time, the second equation of (6.7) can be shown in a similar (and simpler) manner.

Next we prove the first equation of (6.8). By the (piecewise) Lipschitz continuity of g and g', we have on Ω_n

$$
\eta'_-(n,r)^k=-\frac{n^{1/4}}{\psi_2 k_n}\sum_{p=(N^n_{S_r-}-k_n+2)_+}^{N^n_{S_r-}}(\phi_{g',g})^n_{N^n_{S_r-}+1-p}\epsilon^k_{\tau^k_p}+o_p(1).
$$

Moreover, by Lemma 6.4, [SA3] and the Lipschitz continuity of $\phi_{g',g}$ we have

$$E_0\left[\left|\frac{n^{1/4}}{k_n}\left\{\sum_{p=(N^n_{S_r-}-k_n+2)_+}^{N^n_{S_r-}}(\phi_{g',g})^n_{N^n_{S_r-}+1-p}\epsilon^k_{\tau^k_p}-\sum_{p=i(S^\dagger_r)^n+1}^{i(S^\dagger_r)^n+k_n-1}(\phi_{g',g})^n_{i(S^\dagger_r)^n+k_n-p}\epsilon^k_{\tau^k_p}\right\}\right|^2\right]=O_p\left(n^{-\alpha'}\right)$$

on Ω_n. Since $\lim_n P(\Omega_n)=1$, we conclude that

$$\eta'_-(n,r)^k=-\frac{n^{1/4}}{\psi_2 k_n}\sum_{p=i(S^\dagger_r)^n+1}^{i(S^\dagger_r)^n+k_n-1}(\phi_{g',g})^n_{i(S^\dagger_r)^n+k_n-p}\epsilon^k_{\tau^k_p}+o_p(1)=z'^{n,k}_{r-}+o_p(1).$$

Similarly we can prove the second equation of (6.8). □

6.2.2 An auxiliary local stable convergence result

In this subsection we prove an auxiliary local stable convergence result corresponding to Lemma 16.3.7 of [31]. The proof is close to that of the aforementioned lemma, but there is a difference due to the additional randomness coming from the sampling times. Furthermore, we can also simplify some parts of the proof because it is sufficient for our purpose to prove a simpler consequence than that of the aforementioned lemma. For these reasons we give a complete proof.

The following lemma is a direct consequence of the Skorokhod representation theorem, so we omit the proof:

Lemma 6.6. *Let (f_n) be a sequence of real-valued functions on $\mathbb{R}^D$ such that there exists a constant K satisfying $|f_n(x)|\leq K$ and $|f_n(x)-f_n(y)|\leq K\|x-y\|$ for all $x,y\in\mathbb{R}^D$ and every n. If a sequence (x_n) of $\mathbb{R}^D$-valued random variables converges in law to a variable x, then $E[f_n(x_n)]-E[f_n(x)]\to 0$.*

The following lemma is the main result of this subsection. We denote by $\mathfrak{N}_D$ the D-dimensional standard normal distribution.

Lemma 6.7. *Assume that* [SA1], [SA3] *and* [A4](ii) *are satisfied. Suppose that for each n there is a $(\mathcal{G}^{(0)}_t)$-stopping time ρ_n. Suppose also that there is a finite-valued variable ρ such that $\rho_n\to\rho$ as $n\to\infty$ and one of the following two condition is satisfied:*

$$\left.\begin{array}{ll}(1) & \rho>0,\ P(T_{i(\rho_n)^n+k_n-\lfloor n^\beta\rfloor}<\rho)\to 1 \text{ as } n\to\infty \text{ for some } \beta\in(0,\xi-1/2),\\ & \text{in which case we set } G_{(\rho)}=G_{\rho-},\ \widetilde{\upsilon}_{(\rho)}=\widetilde{\upsilon}_{\rho-} \text{ and } \mathcal{G}^{(0)}_{(\rho)}=\mathcal{G}^{(0)}_{\rho-},\\ (2) & \rho_n\geq\rho \text{ for all } n, \text{ in which case we set } G_{(\rho)}=G_\rho,\ \widetilde{\upsilon}_{(\rho)}=\widetilde{\upsilon}_\rho \text{ and } \mathcal{G}^{(0)}_{(\rho)}=\mathcal{G}^{(0)}_\rho.\end{array}\right\}\tag{6.9}$$

Let ϕ_1 and ϕ_2 be continuous real-valued functions ϕ_1 and ϕ_2 on $[0,1]$. Then, for any $\mathcal{F}$-measurable bounded variable U and any bounded Lipschitz function f on $\mathbb{R}^{d'+d}$ we have

$$E[Uf(L_n,L'_n)|\mathcal{G}_{\rho_n}]\to^p E\left[U\int f\left(\|\phi_1\|\sqrt{\theta G_{(\rho)}}x,\|\phi_2\|\sqrt{\theta^{-1}}\widetilde{\upsilon}_{(\rho)}y\right)\mathfrak{N}_{d'}(\mathrm{d}x)\mathfrak{N}_d(\mathrm{d}y)|\mathcal{G}^{(0)}_{(\rho)}\right],\tag{6.10}$$

where $L_n=L(\phi_1,\rho_n)_n$, $L'_n=L'(\phi_2,\rho_n)_n$ and $\|\phi_j\|^2=\int_0^1\phi_j(x)^2\mathrm{d}x$ for $j=1,2$.

Proof. *Step 1)* For $k,l=1,\dots,d$ we set $D^{kl}_n=\frac{1}{k_n}\sum_{w=1}^{k_n-1}|(\phi_2)^n_w|^2 1_{\{\tau^k_{i(\rho_n)^n+w}=\tau^l_{i(\rho_n)^n+w}\}}$. We begin by proving $D^{kl}_n\to^p\|\phi_2\|^2\chi^{kl}_{(\rho)}$, where we set $\chi^{kl}_{(\rho)}=\chi^{kl}_{\rho-}$ in Case (1) and $\chi^{kl}_{(\rho)}=\chi^{kl}_\rho$ in Case (2). Since $i(\rho_n)$ is a $(\mathcal{G}^{(0)}_{T_p})^\infty_{p=0}$-stopping time, [SA1] and Lemma 2.3 of [18] yield $D^{kl}_n=\frac{1}{k_n}\sum_{w=1}^{k_n-1}|(\phi_2)^n_w|^2\chi^{kl}_{T_{i(\rho_n)^n+w-1}}+o_p(1)$. Since χ^{kl} is càdlàg, (6.9) implies that $D^{kl}_n=\|\phi_2\|^2\chi^{kl}_{(\rho)}+o_p(1)$.
Step 2) From Step 1, by considering an appropriate subsequence if necessary, without loss of generality we may assume that there is a subset Ω_0 of $\Omega^{(0)}$ such that $P^{(0)}(\Omega_0)=1$ and $D^{kl}_n(\omega^{(0)})\to\|\phi_2\|^2\chi^{kl}_{(\rho)}(\omega^{(0)})$ for all $\omega^{(0)}\in\Omega_0$.
Step 3) Fix $\omega^{(0)}\in\Omega_0$, and consider the probability space $(\Omega^{(1)},\mathcal{F}^{(1)},Q_0)$, where $Q_0(\cdot)=Q(\omega^{(0)},\cdot)$. Our aim in this step is to show that under Q_0

$$L'_n\to^d\|\phi_2\|\sqrt{\theta^{-1}}\widetilde{\upsilon}_{(\rho)}(\omega^{(0)})\zeta',\tag{6.11}$$

where ζ' is a standard d-dimensional normal variable independent of $\mathcal{F}$.

For each $w = 1, \dots, k_n - 1$ we define the d-dimensional variable $y^n_w = (y^{n,k}_w)_{1\leq k\leq d}$ by

$$y^{n,k}_w = \frac{n^{1/4}}{k_n}(\phi_2)^n_w \epsilon^k_{\tau^k_{i(\rho_n)^n(\omega^{(0)})+w}}.$$

Then $y^n_1, \dots, y^n_{k_n-1}$ are independent under Q_0 and we have $L'_n(\omega^{(0)}, \cdot) = \sum_{w=1}^{k_n-1} y^n_w$. Moreover, by [SA3] we have

$$E_{Q_0}(y^n_w) = 0, \qquad E_{Q_0}(\|y^n_w\|^4) \lesssim k_n^{-2}, \qquad \sum_{w=1}^{k_n-1} E_{Q_0}(\|y^n_w\|^4) \to 0,$$
$$\sum_{w=1}^{k_n-1} E_{Q_0}(y^{n,k}_w y^{n,l}_w) = \frac{n^{1/2}}{k_n^2}\sum_{w=1}^{k_n-1}|(\phi_2)^n_w|^2 \Upsilon^{kl}(\omega^{(0)})_{\tau^k_{i(\rho_n)^n(\omega^{(0)})+w}} 1_{\{\tau^k_{i(\rho_n)^n(\omega^{(0)})+w} = \tau^l_{i(\rho_n)^n(\omega^{(0)})+w}\}}.$$

Since Υ is càdlàg, (6.9) and the fact that $n^{1/2}/k_n \to \theta^{-1}$ yield $\sum_{w=1}^{k_n-1} E_{Q_0}(y^{n,k}_w y^{n,l}_w) = \theta^{-1}\Upsilon^{kl}_{(\rho)}(\omega^{(0)})D^{kl}_n(\omega^{(0)}) + o_p(1)$, where we set $\Upsilon^{kl}_{(\rho)} = \Upsilon^{kl}_{\rho-}$ in Case (1) and $\Upsilon^{kl}_{(\rho)} = \Upsilon^{kl}_\rho$ in Case (2). Since $\omega^{(0)} \in \Omega^{(0)}$, this implies that

$$\sum_{w=1}^{k_n-1} E_{Q_0}(y^{n,k}_w y^{n,l}_w) \to^p \|\phi_2\|^2\theta^{-1}\Upsilon^{kl}_{(\rho)}(\omega^{(0)})\chi^{kl}_{(\rho)}(\omega^{(0)}).$$

Now a standard central limit theorem on row-wise independent triangular arrays of infinitesimal variables (e.g. Theorem 2.2.14 of [31]) yields (6.11).

Step 4) In this step we shall show the following convergence for L_n:

$$L_n \to^d \sqrt{\Phi_{22}\theta G_{(\rho)}}\zeta, \tag{6.12}$$

where ζ is a standard d'-dimensional normal variable independent of $\mathcal{F}$. Unlike Step 3, here the limiting variable is mixed normal, so we cannot rely on the standard central limit theorem used in Step 3. Instead, we use the classic mixed normal limit theorem of Hall [20].

Fix $u \in \mathbb{R}^{d'}$ arbitrarily and set $y(u)^n_w = n^{1/4}(\phi_1)^n_w u^* W(I_{i(\rho_n)^n+w})$ for each $w = 1, \dots, k_n - 1$. Then $y(u)^n_w$ is $\mathcal{G}_{T_{i(\rho_n)^n+w}}$-measurable and $u^*L_n = \sum_{w=1}^{k_n-1} y(u)^n_w$. Therefore, noting that G and G_- do not vanish, it suffices to verify the following four conditions according to [20] and the Cramér-Wold method:

$$E\left[\max_{1\leq w\leq k_n-1}|y(u)^n_w|^2\right] \to 0, \tag{6.13}$$
$$\textstyle\sum_{w=1}^{k_n-1}|y(u)^n_w|^2 - \|u\|^2\|\phi_1\|^2\theta G_{\rho_n} \to^p 0, \tag{6.14}$$
$$\|u\|^2\|\phi_1\|^2\theta G_{\rho_n} \to^p \|u\|^2\|\phi_1\|^2\theta G_{(\rho)}, \tag{6.15}$$
$$\textstyle\sum_{w=1}^{k_n-1}\left|E[y(u)^n_w|\mathcal{G}_{T_{i(\rho_n)^n+w-1}}]\right| \to^p 0. \tag{6.16}$$

(6.13) follows from (6.2) and Lévy's theorem on the uniform modulus of continuity of Brownian motion. Next, [SA1] and Lemma 2.3 of [18] imply that

$$\sum_{w=1}^{k_n-1}|y(u)^n_w|^2 = \|u\|^2\sqrt{n}\sum_{w=1}^{k_n-1}(\phi_1)^n_w|I_{i(\rho_n)^n+w}| + o_p(1) = \frac{\|u\|^2}{\sqrt{n}}\sum_{w=1}^{k_n-1}(\phi_1)^n_w G_{T_{i(\rho_n)^n+w-1}} + o_p(1),$$

hence we obtain (6.14) because G is càdlàg. Finally, the fact that G is càdlàg and (6.9) yield (6.15), while we have $E[y(u)^n_w|\mathcal{G}_{T_{i(\rho_n)^n+w-1}}] = 0$ because W is a d'-dimensional $(\mathcal{F}_t)$-Brownian motion independent of $\mathcal{G}$, hence (6.16) holds true.

Step 5) We denote by $\Psi_n(U)$ and $\Psi(U)$ the left and right sides of (6.10), respectively. In this step we show that it is enough to prove

$$\Psi_n(1) \to^p \Psi(1). \tag{6.17}$$

In fact, assume this, and take an arbitrary bounded variable U. We consider the càdlàg version of the bounded martingale $U_t = E(U|\mathcal{G}^{(0)}_t)$.

First suppose that we are in Case (1). Set $\underline{k}_n = k_n - \lfloor n^\beta \rfloor$ and define the d'-dimensional variable $\underline{L}_n = (\underline{L}^j_n)_{1\leq j\leq d'}$ and the d-dimensional variable $\underline{L}'_n = (\underline{L}'_n)_{1\leq k\leq d}$ by

$$\underline{L}^j_n = n^{1/4}\sum_{w=1}^{\underline{k}_n}(\phi_1)^n_w W^j(I_{i(\rho)^n+w}), \qquad \underline{L}'^k_n = \frac{n^{1/4}}{k_n}\sum_{w=1}^{\underline{k}_n}(\phi_2)^n_w \epsilon^k_{\tau^k_{i(\rho)^n+w}}.$$

Then, since $E[\|L_n - \underline{L}_n\|^2] \lesssim \sqrt{n} n^\beta \bar{r}_n$ and $E[\|L'_n - \underline{L}'_n\|^2] \lesssim \sqrt{n} k_n^{-2} n^\beta$, by the boundedness of U and the Lipschitz continuity of f it holds that $\Psi_n(U) - \underline{\Psi}_n(U) \to^p 0$ and $\Psi_n(1) - \underline{\Psi}_n(1) \to^p 0$, where $\underline{\Psi}_n(U) = E[Uf(\underline{L}_n, \underline{L}'_n)|\mathcal{G}_{\rho_n}]$. In particular, to prove (6.17) it is enough to show that $\underline{\Psi}_n(U) \to^p \Psi(U)$. Now, since both $G_{\rho-}$ and $\widetilde{v}_{\rho-}$ are $\mathcal{G}^{(0)}_{\rho-}$-measurable, we have $\Psi(U) = U_{\rho-}\Psi(1)$ because $U_{\rho-} = E[U|\mathcal{G}^{(0)}_{\rho-}]$. Also, $f(\underline{L}_n, \underline{L}'_n)$ in restriction to the set $\Omega_n = \{\rho > T_{i(\rho_n)^n+\underline{k}_n}\}$ is $\mathcal{G}_{\rho-}$-measurable, so $\underline{\Psi}_n(U) = \underline{\Psi}_n(U_{\rho-})$ on Ω_n. We also obviously have $\underline{\Psi}_n(U_{\rho_n}) = \underline{\Psi}_n(1)U_{\rho_n} \to^p \Psi(1)U_{\rho-}$ by (6.17), $\Psi_n(1) - \underline{\Psi}_n(1) \to^p 0$ and $U_{\rho_n} \to U_{\rho-}$, while $P(\Omega_n) \to 1$ by assumption. Now, since $E[|\underline{\Psi}_n(U_{\rho_n}) - \underline{\Psi}_n(U_{\rho-})|] \leq \|f\|_\infty E[|U_{\rho_n} - U_{\rho-}|] \to 0$ by the boundedness of f and U, $U_{\rho_n} \to U_{\rho-}$ on Ω_n and the fact that $P(\Omega_n) \to 1$, we obtain the desired result.

Next suppose that we are in Case (2). Then $\Psi(U) = U_\rho\Psi(1)$ because $\Psi(1)$ is $\mathcal{G}^{(0)}_\rho$-measurable, and also $\Psi_n(U_\rho) = U_\rho\Psi_n(1)$ because $\rho_n \geq \rho$. Moreover, setting $\rho'_n = \rho_n + k_n\bar{r}_n$, $\Psi_n(1)$ is $\mathcal{G}_{\rho'_n}$-measurable due to (6.2), so $\Psi_n(U) = U_{\rho'_n}\Psi_n(1)$. Since $\rho'_n \to \rho$ and $\rho'_n > \rho$, we have $U_{\rho'_n} \to U_\rho$, and the same arguments as above shows that $\Psi_n(U_{\rho'_n}) - \Psi_n(U_\rho) \to^p 0$, thus the desired result is obtained.

Step 6) Now we finish the proof by proving the convergence (6.17). First, for each $\omega^{(0)} \in \Omega_0$ define the function $h^n_{\omega^{(0)}}$ on $\mathbb{R}^d$ by $h^n_{\omega^{(0)}}(y) = f(L_n(\omega^{(0)}), y)$. Then, noting that f is bounded and Lipschitz continuous, Lemma 6.6 and (6.11) imply that

$$\int h^n_{\omega^{(0)}}(L'_n(\omega^{(1)}))Q(\omega^{(0)}, \mathrm{d}\omega^{(1)}) - \int h^n_{\omega^{(0)}}(\|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{(\rho)}(\omega^{(0)})y)\mathfrak{N}_d(\mathrm{d}y) \to 0.$$

Since f is bounded and $P(\Omega_0) = 1$, this convergence and the bounded convergence theorem yield

$$\Psi_n(1) - E\left[\int f\left(L_n, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{(\rho)}y\right)\mathfrak{N}_d(\mathrm{d}y)|\mathcal{G}_{\rho_n}\right] \to^p 0.$$

Next, since f is Lipschitz continuous and $\widetilde{v}$ is càdlàg and bounded, by (6.9) we obtain

$$\Psi_n(1) - E\left[\int f\left(L_n, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{\rho_n}y\right)\mathfrak{N}_d(\mathrm{d}y)|\mathcal{G}_{\rho_n}\right] \to^p 0. \tag{6.18}$$

Now, noting that W is a standard d'-dimensional Brownian motion with respect to $(\mathcal{G}_t)$, by the strong Markov property of a Brownian motion $(W_{\rho_n+t} - W_{\rho_n})_{t\geq 0}$ is independent of $\mathcal{G}_{\rho_n}$, hence we have

$$E\left[\int f\left(L_n, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{\rho_n}y\right)\mathfrak{N}_d(\mathrm{d}y)|\mathcal{G}_{\rho_n}\right] = \int f\left(x, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{\rho_n}y\right)\mathbb{P}^n(\mathrm{d}x)\mathfrak{N}_d(\mathrm{d}y), \tag{6.19}$$

where $\mathbb{P}^n$ is the law of L_n under $P^{(0)}$. Then, again using the Lipschitz continuity of f and the càdlàg property of $\widetilde{v}$ as well as (6.9), we obtain

$$\int f\left(x, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{\rho_n}y\right)\mathbb{P}^n(\mathrm{d}x)\mathfrak{N}_d(\mathrm{d}y) - \int f\left(x, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{(\rho)}y\right)\mathbb{P}^n(\mathrm{d}x)\mathfrak{N}_d(\mathrm{d}y) \to^p 0,$$

hence (6.12) yields

$$\int f\left(x, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{\rho_n}y\right)\mathbb{P}^n(\mathrm{d}x)\mathfrak{N}_d(\mathrm{d}y) \to^p \int f\left(\|\phi_1\|\sqrt{\theta G_{(\rho)}}x, \|\phi_2\|\sqrt{\theta^{-1}}\widetilde{v}_{(\rho)}y\right)\mathfrak{N}_{d'}(\mathrm{d}x)\mathfrak{N}_d(\mathrm{d}y). \tag{6.20}$$

(6.18)–(6.20) imply that (6.17) holds true, and thus we complete the proof. $\square$

6.2.3 A joint convergence result and the proof of Proposition 6.2

In this subsection we prove a joint convergence result for the pair $(n^{1/4}\mathbf{L}[M]^n, (z^n_{r-}, z'^n_{r-}, z^n_{r+}, z'^n_{r+})_{r\geq 1})$ and complete the proof of Proposition 6.2.

For the proof we use some elementary results on the Skorokhod topology. For any $k \in \mathbb{N}$, denote by $\mathbb{D}^k$ (resp. $\mathbb{D}^{k\times k}$) the space of $\mathbb{R}^k$-valued (resp. $\mathbb{R}^k \otimes \mathbb{R}^k$-valued) càdlàg functions on $\mathbb{R}_+$ equipped with the Skorokhod topology. For any $x \in \mathbb{D}^k$ and any $t \in \mathbb{R}_+$, we define the function x^t by $x^t(s) = x(s \wedge t)$ for $s \in \mathbb{R}_+$. We evidently have $x^t \in \mathbb{D}^k$.

Lemma 6.8. *The map $\mathbb{R}_+ \times \mathbb{D}^k \ni (t, x) \mapsto x^t \in \mathbb{D}^k$ is continuous at every point $(t, x) \in \mathbb{R}_+ \times \mathbb{D}^k$ such that x is continuous at t, where the space $\mathbb{R}_+ \times \mathbb{D}^k$ is equipped with the product topology.*

Proof. Let $((t_j, x_j))_{j\geq 1}$ be a sequence of $\mathbb{R}_+ \times \mathbb{D}^k$ such that $(t_j, x_j) \to (t, x)$ in $\mathbb{R}_+ \times \mathbb{D}^k$ as $j \to \infty$. We need to show that $x_j^{t_j} \to x^t$ in $\mathbb{D}^k$ as $j \to \infty$. First, since $x_j \to x$ in $\mathbb{D}^k$ as $j \to \infty$, there is a sequence λ_j of strictly increasing continuous maps of $\mathbb{R}_+$ onto itself such that $\sup_{0\leq s\leq T} \|x_j(\lambda_j(s)) - x(s)\| \to 0$ and $\sup_{s\in\mathbb{R}_+} |\lambda_j(s) - s| \to 0$ as $j \to \infty$ for any $T > 0$. Then, for any $T > 0$ we have

$$
\begin{aligned}
\sup_{0\leq s\leq T} \|x_j^{t_j}(\lambda_j(s)) - x^t(s)\| \leq{}& \sup_{0\leq s\leq T} \|x_j(\lambda_j(s)) - x(s)\| + \sup_{t<s\leq\lambda_j^{-1}(t_j)\wedge T} \|x_j(\lambda_j(s)) - x(t)\| \\
&+ \sup_{\lambda_j^{-1}(t_j)<s\leq t} \|x_j(t_j) - x(s)\| + \|x_j(t_j) - x(t)\| \\
=:{}& \mathbb{I}_j + \mathbb{II}_j + \mathbb{III}_j + \mathbb{IV}_j.
\end{aligned}
$$

By the construction of λ_j we have $\mathbb{I}_j \to 0$ as $j \to \infty$. On the other hand, since x is continuous at t, $x_j(t_j) \to x(t)$ as $j \to \infty$ by Proposition VI-2.1 of [32], hence $\mathbb{IV}_j \to 0$ as $j \to \infty$. Moreover, since $\lambda_j^{-1}(t_j) \to t$ as $j \to \infty$ and x is continuous at t, we also have $\mathbb{III}_j \to 0$ as $j \to \infty$. Finally, since $\mathbb{II}_j \leq \mathbb{I}_j + \sup_{t<s\leq\lambda_j^{-1}(t_j)} \|x(s) - x(t)\|$ and x is continuous at t, we have $\mathbb{II}_j \to 0$ as $j \to \infty$. This completes the proof. $\square$

For any $S \geq 0$ we define the function π_S from $\mathbb{D}^{d\times d}$ into itself by $\pi_S(x)(t) = x(t - S)1_{\{t\geq S\}}$.

Lemma 6.9. *π_S is a continuous function of $\mathbb{D}^{d\times d}$ into itself.*

Proof. Let $(x_j)_{j\geq 1}$ be a sequence of elements of $\mathbb{D}^{d\times d}$ such that $x_j \to x$ in $\mathbb{D}^{d\times d}$ as $j \to \infty$ for some $x \in \mathbb{D}^{d\times d}$. We need to show that $\pi_S(x_j) \to \pi_S(x)$ in $\mathbb{D}^{d\times d}$ as $j \to \infty$. Since $x_j \to x$ in $\mathbb{D}^{d\times d}$, there is a sequence λ_j of strictly increasing continuous maps of $\mathbb{R}_+$ onto itself such that $\sup_{0\leq s\leq T} \|x_j(\lambda_j(s)) - x(s)\| \to 0$ and $\sup_{s\in\mathbb{R}_+} |\lambda_j(s) - s| \to 0$ as $j \to \infty$ for any $T > 0$. Then, for each j define the function $\widetilde{\lambda}_j$ on $\mathbb{R}_+$ by $\widetilde{\lambda}_j(s) = s$ if $s < S$; $\widetilde{\lambda}_j(s) = \lambda_j(s - S) + S$ otherwise. Obviously $\widetilde{\lambda}_j$ is a strictly increasing continuous map of $\mathbb{R}_+$ onto itself and satisfies $\sup_{s\geq 0} |\widetilde{\lambda}_j(s) - s| \to 0$ as $j \to \infty$. Moreover, since it holds that $\pi_S(x_j)(\widetilde{\lambda}_j(s)) = x_j(\lambda_j(s - S))1_{\{s\geq S\}}$, we also have $\sup_{0\leq s\leq T} \|\pi_S(x_j)(\widetilde{\lambda}_j(s)) - \pi_S(x)(s)\| \to 0$ as $j \to \infty$ for any $T > 0$. This implies the desired result. $\square$

Now we are ready to prove the following joint convergence result:

Proposition 6.7. *Suppose that* [SA1]–[SA3] *and* [A4] *are satisfied. Then*

$$(n^{1/4}\mathbf{L}[M]^n, (z^n_{r-}, z'^n_{r-}, z^n_{r+}, z'^n_{r+})_{r\geq 1}) \to^{d_s} (\mathcal{W}, (z_{r-}, z'_{r-}, z_{r+}, z'_{r+})_{r\geq 1})$$

as $n \to \infty$ for the product topology on the space $\mathbb{D}^{d\times d} \times (\mathbb{R}^{2(d+d')})^{\mathbb{N}}$, where

$$
\begin{cases}
z_{r-} = \psi_2^{-1}\sqrt{\Phi_{22}\theta G_{S_r-}}\Psi_{r-}, & z_{r+} = \psi_2^{-1}\sqrt{\Phi_{22}\theta G_{S_r}}\Psi_{r+}, \\
z'_{r-} = \psi_2^{-1}\sqrt{\Phi_{12}\theta^{-1}\widetilde{\upsilon}_{S_r-}}\Psi'_{r-}, & z'_{r+} = \psi_2^{-1}\sqrt{\Phi_{12}\theta^{-1}\widetilde{\upsilon}_{S_r}}\Psi'_{r+}.
\end{cases}
$$

Proof. *Step 1)* It suffices to prove

$$(n^{1/4}\mathbf{L}[M]^n, (z^n_{r-}, z'^n_{r-}, z^n_{r+}, z'^n_{r+})_{r\in\mathcal{R}}) \to^{d_s} (\mathcal{W}, (z_{r-}, z'_{r-}, z_{r+}, z'_{r+})_{r\in\mathcal{R}}) \tag{6.21}$$

in $\mathbb{D}^{d\times d}\times\mathbb{R}^{2(d'+d)\#\mathcal{R}}$ for any finite subset $\mathcal{R}$ of $\mathbb{N}$, and we prove this by induction on the number $\#\mathcal{R}$ of the elements in the set $\mathcal{R}$. First, (6.21) holds true when $\#\mathcal{R}=0$ due to Proposition 6.5. Next, let $J\in\mathbb{N}$ and assume that (6.21) holds true when $\#\mathcal{R}=J-1$. Then, we need to prove (6.21) for the case that $\#\mathcal{R}=J$. We write $\mathcal{R}=\{r_1,\dots,r_J\}$ with $S_{r_1}<\cdots<S_{r_J}$.

Step 2) Before stating the detailed proof, we briefly explain the intuition behind the proof. The basic idea is the same as in the proof of Theorem 4.3.1 from [31]. Namely, for each $\beta>0$ we set $S^{\beta-}=(S_{r_J}-\beta)_+$ and $S^{\beta+}=S^{\beta+}$, and divide $n^{1/4}\mathbf{L}[M]^n$ into the summands containing the data observed in the interval $[S^{\beta-},S^{\beta+}]$ and the remaining ones. Then we prove the negligibility of the former part (as $\beta\to 0$) and the joint limit theorem of the latter part and $(z^n_{r-},z'^n_{r-},z^n_{r+},z'^n_{r+})_{r\in\mathcal{R}}$. More formally, we set $\widehat{\mathbf{L}}(\beta)^n=n^{1/4}\left(\mathbf{L}[M]^n\right)^{S^{\beta+}}-n^{1/4}\left(\mathbf{L}[M]^n\right)^{S^{\beta-}}$ and $\widehat{\mathcal{W}}(\beta)=\mathcal{W}^{S^{\beta+}}-\mathcal{W}^{S^{\beta-}}$, and show that

$$\limsup_{\beta\to 0}\limsup_{n\to\infty}P\left(\sup_{0\leq t\leq T}\left\|\widehat{\mathbf{L}}(\beta)^n_t\right\|>\eta\right)=0,\qquad \limsup_{\beta\to 0}P\left(\sup_{0\leq t\leq T}\left\|\widehat{\mathcal{W}}(\beta)_t\right\|>\eta\right)=0 \tag{6.22}$$

for any $T,\eta>0$ and that

$$(n^{1/4}\mathbf{L}[M]^n-\widehat{\mathbf{L}}(\beta)^n,(z^n_{r-},z'^n_{r-},z^n_{r+},z'^n_{r+})_{r\in\mathcal{R}})\to^{d_s}(\mathcal{W}-\widehat{\mathcal{W}}(\beta),(z_{r-},z'_{r-},z_{r+},z'_{r+})_{r\in\mathcal{R}}) \tag{6.23}$$

in $\mathbb{D}^{d\times d}\times\mathbb{R}^{2(d'+d)J}$ as $n\to\infty$ for any fixed $\beta>0$. Then, Proposition 2.2.4 of [31] yields (6.21).

However, to prove (6.23) we need a different approach from the one of [31] because we cannot argue conditionally on the increments of W consisting of the observations in $[S^{\beta-},S^{\beta+}]$ as [31] do, which is due to the time endogeneity. For this reason we further decompose $n^{1/4}\mathbf{L}[M]^n-\widehat{\mathbf{L}}(\beta)^n$ as $n^{1/4}\mathbf{L}[M]^n-\widehat{\mathbf{L}}(\beta)^n=\check{\mathbf{L}}^n+\widetilde{\mathbf{L}}(\beta)^n$, where $\check{\mathbf{L}}^n=(n^{1/4}\mathbf{L}[M]^n)^{S^{\beta-}}$. Roughly speaking, $\check{\mathbf{L}}^n$ consists of the data observed before $S^{\beta-}$, while $\widetilde{\mathbf{L}}(\beta)^n$ consists of those observed after $S^{\beta+}$. By Proposition VI-1.23 of [32] and the continuous mapping theorem (6.23) follows once we show that

$$(\check{\mathbf{L}}^n,\widetilde{\mathbf{L}}(\beta)^n,(z^n_{r-},z'^n_{r-},z^n_{r+},z'^n_{r+})_{r\in\mathcal{R}})\to^{d_s}(\mathcal{W}^{S_{r_J}},\widetilde{\mathcal{W}}(\beta),(z_{r-},z'_{r-},z_{r+},z'_{r+})_{r\in\mathcal{R}}) \tag{6.24}$$

in $\mathbb{D}^{d\times d}\times\mathbb{D}^{d\times d}\times\mathbb{R}^{2(d'+d)J}$ as $n\to\infty$. The strategy of the proof of (6.24) is, roughly speaking, as follows. We first prove a stable limit theorem for $\widetilde{\mathbf{L}}(\beta)^n$ conditionally on $\mathcal{F}_{S^{\beta+}}$ (this will be done in Step 5; the assumption [A4](i) is necessary for this part). Then we obtain a joint stable limit theorem for $\widetilde{\mathbf{L}}(\beta)^n$ and $(z^n_{r_J-},z'^n_{r_J-},z^n_{r_J+},z'^n_{r_J+})$ conditionally on $\mathcal{G}_{S^\dagger_{r_J}}$ by virtue of Lemma 6.7 (Step 7). Finally, from the assumption of the induction we will obtain the desired result (Step 8).

Step 3) We begin with proving (6.22). The second equation immediately follows from the continuity of the process $\mathcal{W}$. On the other hand, for any $\beta>0$ we have $(S^{\beta-},S^{\beta+},n^{1/4}\mathbf{L}[M]^n)\to^{d_s}(S^{\beta-},S^{\beta+},\mathcal{W})$ as $n\to\infty$ in $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{D}^{d\times d}$ by Proposition 6.5, hence Lemma 6.8 and the continuous mapping theorem imply that $(n^{1/4}\left(\mathbf{L}[M]^n\right)^{S^{\beta+}},n^{1/4}\left(\mathbf{L}[M]^n\right)^{S^{\beta-}})\to^{d_s}(\mathcal{W}^{S^{\beta+}},\mathcal{W}^{S^{\beta-}})$ as $n\to\infty$ in $\mathbb{D}^{d\times d}\times\mathbb{D}^{d\times d}$. Therefore, Propositions VI-1.23 and VI-2.4 of [32] as well as the continuous mapping theorem yield $\sup_{0\leq t\leq T}\left\|\widehat{\mathbf{L}}(\beta)^n_t\right\|\to^{d_s}\sup_{0\leq t\leq T}\|\mathcal{W}(\beta)_t\|$ as $n\to\infty$. In particular, we have

$$\limsup_{n\to\infty}P\left(\sup_{0\leq t\leq T}\left\|\widehat{\mathbf{L}}(\beta)^n_t\right\|>\eta\right)\leq P\left(\sup_{0\leq t\leq T}\|\mathcal{W}(\beta)_t\|\geq\eta\right),$$

hence we also obtain the first equation (6.22).

Step 4) Now we start the proof of (6.24). First, due to the property of the product topology it suffices to prove the following convergence:

$$E\left[\zeta f_1(\check{\mathbf{L}}^n)f_2(\widetilde{\mathbf{L}}(\beta)^n)\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right]\to E\left[\zeta f_1(\mathcal{W}^{S_{r_J}})f_2(\widetilde{\mathcal{W}}(\beta))\prod_{j=1}^J Y_{j-}Y_{j+}\right] \tag{6.25}$$

as $n\to\infty$, where ζ is any bounded $\mathcal{F}^{(0)}$-measurable variable, f_1 and f_2 are bounded Lipschitz functions on $\mathbb{D}^{d\times d}$, and $Y^n_{j\pm}=F_{j\pm}(z^n_{r_j\pm},z'^n_{r_j\pm})$ and $Y_{j\pm}=F_{j\pm}(z_{r_j\pm},z'_{r_j\pm})$ with F_{j-} and F_{j+} being bounded Lipschitz functions on $\mathbb{R}^{d'+d}$ for every $j=1,\dots,J$.

Step 5) We begin with proving

$$E\left[\zeta f_1(\check{\mathbf{L}}^n)f_2(\widetilde{\mathbf{L}}(\beta)^n)\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right]-E\left[\zeta f_1(\check{\mathbf{L}}^n)f_2(\widetilde{\mathcal{W}}(\beta))\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right]\to 0 \tag{6.26}$$

as $n\to\infty$. First, we introduce some notation. For any d-dimensional processes U,V, any $u,v\in\{g,g'\}$ and any $k,l=1,\dots,d$, we define the process $\widetilde{\mathbb{L}}^{(k,l)}_{u,v}(U,V)^n$ in the same way as that of $\mathbb{L}^{(k,l)}_{u,v}(U,V)^n$ with replacing $(T_p)_{p\geq 0}$ by $(\widetilde{T}_p)_{p\geq 0}:=(T_{i(S^{\beta+})^n+1+p}-S^{\beta+})_{p\geq 0}$. Also, for any process V we define the process $\mathring{V}$ by $\mathring{V}_t=V_{S^{\beta+}+t}-V_{S^{\beta+}}$, and define the $\mathbb{R}^d\otimes\mathbb{R}^d$-valued process $\widetilde{\mathbf{L}}^n$ by

$$\widetilde{\mathbf{L}}^{n,kl}=n^{1/4}\left\{\widetilde{\mathbb{L}}^{(k,l)}_{g,g}(\mathring{M},\mathring{M})^n+\widetilde{\mathbb{L}}^{(k,l)}_{g,g'}(\mathring{M},\mathring{\mathfrak{E}})^n+\widetilde{\mathbb{L}}^{(l,k)}_{g,g'}(\mathring{M},\mathring{\mathfrak{E}})^n+\widetilde{\mathbb{L}}^{(k,l)}_{g',g'}(\mathring{\mathfrak{E}},\mathring{\mathfrak{E}})^n\right\}.$$

Then, it can easily be seen that $\widetilde{\mathbf{L}}(\beta)^n-\pi_{S^{\beta+}}(\widetilde{\mathbf{L}}^n)\xrightarrow{ucp}0$ as $n\to\infty$. Therefore, by the Lipschitz continuity of f_2 as well as the boundedness of ζ, f_1 and $Y^n_{j\pm}$ we have

$$E\left[\zeta f_1(\check{\mathbf{L}}^n)f_2(\widetilde{\mathbf{L}}(\beta)^n)\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right]-E\left[\zeta f_1(\check{\mathbf{L}}^n)f_2(\pi_{S^{\beta+}}(\widetilde{\mathbf{L}}^n))\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right]\to 0 \tag{6.27}$$

as $n\to\infty$.

Now we consider a regular conditional probability $p^{(0)}(\omega^{(0)},\cdot)$ of $P^{(0)}$ given $\mathcal{F}^{(0)}_{S^{\beta+}}$. Such one exists because of the assumption [A4](i). We also consider a filtration $(\mathring{\mathcal{F}}^{(0)}_t)_{t\geq 0}$ of $\mathcal{F}^{(0)}$ defined by $\mathring{\mathcal{F}}^{(0)}_t=\mathcal{F}^{(0)}_{S^{\beta+}+t}$, and for each $\omega_0\in\Omega^{(0)}$ we introduce a stochastic basis $\mathcal{B}^{(0)}_{\omega_0}:=(\Omega^{(0)},\mathcal{F}^{(0)},(\mathring{\mathcal{F}}^{(0)}_t),p^{(0)}(\omega_0,\cdot))$. For each $t\in\mathbb{R}_+$ we also introduce a transition probability $\mathring{Q}_t(\omega^{(0)},\mathrm{d}u)$ from $(\Omega^{(0)},\mathring{\mathcal{F}}^{(0)}_t)$ into $\mathbb{R}^d$ by setting $\mathring{Q}_t(\omega^{(0)},A)=Q_{S^{\beta+}(\omega^{(0)})+t}(\omega^{(0)},A)$ for each Borel set A of $\mathbb{R}^d$. Note that the process $(\mathring{Q}_t(\cdot,A))_{t\geq 0}$ is $(\mathring{\mathcal{F}}^{(0)}_t)$-progressively measurable because of [A4](ii) and Theorem IV-57 of [41]. Now, by replacing $\mathcal{B}^{(0)}$, $Q_t(\omega^{(0)},\mathrm{d}u)$ and $(\mathcal{T}^n_i)$ with $\mathcal{B}^{(0)}_{\omega_0}$, $\mathring{Q}_t(\omega^{(0)},\mathrm{d}u)$ and the increasing reordering of $\{\widetilde{\tau}^k_p:=\tau^k_{i(S^{\beta+})^n+1+p}-S^{\beta+}:k=1,\dots,d \text{ and } p\geq 0\}$ respectively, we introduce the new stochastic basis $\mathcal{B}_{\omega_0}=(\Omega,\mathcal{F},(\mathring{\mathcal{F}}_t),P_{\omega_0})$ instead of $\mathcal{B}$.

By the strong Markov property of a Brownian motion $\mathring{W}$ is a d'-dimensional standard Brownian motion on $\mathcal{B}^{(0)}_{\omega_0}$. Moreover, defining the random measure $\mathring{\mu}$ by $\mathring{\mu}((0,t]\times A)=\mu((S^{\beta+},S^{\beta+}+t]\times A)$, the strong Markov property of a Poisson random measure implies that $\mathring{\mu}$ is a Poisson random measure on $\mathcal{B}^{(0)}_{\omega_0}$ with compensator ν. We also have

$$\begin{aligned}
\mathring{M}_t&=\int_0^t\sigma_{S^{\beta+}+s}\mathrm{d}\mathring{W}_s,\\
\sigma_{S^{\beta+}+t}&=\sigma_{S^{\beta+}}+\int_0^t\widetilde{b}_{S^{\beta+}+s}\mathrm{d}s+\int_0^t\widetilde{\sigma}_{S^{\beta+}+s}\mathrm{d}\mathring{W}_s+(\mathring{\widetilde{\delta}}1_{\{|\mathring{\widetilde{\delta}}|\leq 1\}})\star(\mathring{\mu}-\nu)_t+(\mathring{\widetilde{\delta}}1_{\{|\mathring{\widetilde{\delta}}|>1\}})\star\mathring{\mu}_t,\\
G_{S^{\beta+}+t}&=G_{S^{\beta+}}+\int_0^t\widehat{b}_{S^{\beta+}+s}\mathrm{d}s+\int_0^t\widehat{\sigma}_{S^{\beta+}+s}\mathrm{d}\mathring{W}_s+(\mathring{\widehat{\delta}}1_{\{|\mathring{\widehat{\delta}}|\leq 1\}})\star(\mathring{\mu}-\nu)_t+(\mathring{\widehat{\delta}}1_{\{|\mathring{\widehat{\delta}}|>1\}})\star\mathring{\mu}_t,
\end{aligned}$$

where for a function η on $\Omega^{(0)}\times\mathbb{R}_+\times E$ the function $\mathring{\eta}$ on $\Omega^{(0)}\times\mathbb{R}_+\times E$ is defined by $\mathring{\eta}(\omega^{(0)},t,z)=\eta(\omega^{(0)},S^{\beta+}(\omega^{(0)})+t,z)$. Therefore, noting that for any $\mathcal{F}$-measurable variable x and any sub σ-filed $\mathcal{H}$ of $\mathcal{F}$ we have $E_{P_{\omega_0}}[x|\mathcal{H}]=E[x|\mathcal{H}]$ as long as $\mathcal{F}^{(0)}_{S^{\beta+}}\subset\mathcal{H}$, it can easily been shown that the conditions [SA1]–[SA3] are satisfied with replacing $\mathcal{B}$, X, (T_p), (τ^k_p), G_t and χ_t by $\mathcal{B}_{\omega_0}$, $\mathring{M}$, $(\widetilde{T}_p)$, $(\widetilde{\tau}^k_p)$, $G_{S^{\beta+}+t}$ and $\chi_{S^{\beta+}+t}$, respectively.

Consequently, Proposition 6.5 and Lemma 6.9 as well as the continuous mapping theorem yield

$$E_{P_{\omega_0}}[\zeta f_2(\pi_{S_{r_J}(\omega_0)+\beta}(\widetilde{\mathbf{L}}^n))]\to E_{p^{(0)}(\omega_0,\cdot)}[\zeta f_2(\pi_{S_{r_J}(\omega_0)+\beta}(\mathring{\mathcal{W}}))].$$

Therefore, noting that $E_{P_{\omega_0}}[\cdot]=E[\cdot|\mathcal{F}_{S^{\beta+}}](\omega_0)$ and $E_{p^{(0)}(\omega_0,\cdot)}[\zeta f_2(\pi_{S_{r_J}(\omega_0)+\beta}(\mathring{\mathcal{W}}))]=E[\zeta f_2(\pi_{S_{r_J}(\omega_0)+\beta}(\mathring{\mathcal{W}}))|\mathcal{F}_{S^{\beta+}}](\omega_0)$ for almost all ω_0 (with respect to $P^{(0)}$) and that $f_1(\check{\mathbf{L}}^n)\prod_{j=1}^J Y^n_{j-}Y^n_{j+}$ is bounded and $\mathcal{F}_{S^{\beta+}}$-measurable (for sufficiently large n; note that $S^{\beta-}$ is an $(\mathcal{F}_{S^{\beta+}+t})$-stopping time), the bounded convergence theorem implies that

$$E\left[\zeta f_1(\check{\mathbf{L}}^n)f_2(\pi_{S^{\beta+}}(\widetilde{\mathbf{L}}^n))\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right]-E\left[\zeta f_1(\check{\mathbf{L}}^n)f_2(\pi_{S^{\beta+}}(\mathring{\mathcal{W}}))\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right]\to 0.$$

Since $\pi_{S^{\beta+}}(\mathring{\mathcal{W}})=\widetilde{\mathcal{W}}(\beta)$, by (6.27) and the above convergence we obtain (6.26).

Step 6) In this step we prove

$$E\left[\zeta' Y^n_{J-}Y^n_{J+}|\mathcal{G}_{S^\dagger_{r_J}}\right] \to^p E\left[\zeta' Y_{J-}Y_{J+}|\mathcal{G}^{(0)}_{S_{r_J-}}\right], \tag{6.28}$$

where $\zeta' = E[\zeta f_2(\widetilde{\mathcal{W}}(\beta))|\mathcal{F}^{(0)}]$. Fix a constant α' such that $1-\xi < \alpha' < (\xi-\kappa-\frac{1}{2}) \wedge (\xi - \frac{3}{4}) \wedge \varpi$. Then, set $\underline{k}_n = k_n - \lfloor n^{1/2-\alpha'} \rfloor$ and define the d'-dimensional variable $\underline{z}^n_{r_J-} = (\underline{z}^{n,j}_{r_J-})_{1\leq j\leq d'}$ and the d-dimensional variable $\underline{z}'^n_{r_J-} = (\underline{z}'^{n,k}_{r_J-})_{1\leq k\leq d}$ by

$$\underline{z}^{n,j}_{r_J-} = n^{1/4}\sum_{w=1}^{\underline{k}_n} (\widetilde{\phi_{g,g}})^n_w W^j(I_{i(S^\dagger_r)^n+w}), \qquad \underline{z}'^{n,k}_{r_J-} = -\frac{n^{1/4}}{k_n}\sum_{w=1}^{\underline{k}_n} (\widetilde{\phi_{g',g}})^n_w \epsilon^k_{\tau^k_{i(S^\dagger_r)^n+w}},$$

and put $\underline{Y}^n_{J-} = F_{J-}(\underline{z}^n_{r_J-}, \underline{z}'^n_{r_J-})$. Since $E[\|z^n_{r_J-} - \underline{z}^n_{r_J-}\|^2] \lesssim n^{1-\alpha'-\xi}$ and $E[\|z'^n_{r_J-} - \underline{z}'^n_{r_J-}\|^2] \lesssim n^{-\alpha'}$ by [SA1]–[SA3] and the optional sampling theorem, we have $Y^n_{J-} - \underline{Y}^n_{J-} \to^p 0$ due to the Lipschitz continuity of F_{J-}. Therefore, by virtue of the boundedness of ζ' and F_{J+}, for the proof of (6.28) it is enough to prove

$$E\left[\zeta' \underline{Y}^n_{J-}Y^n_{J+}|\mathcal{G}_{S^\dagger_{r_J}}\right] \to^p E\left[\zeta' Y_{J-}Y_{J+}|\mathcal{G}^{(0)}_{S_{r_J-}}\right]. \tag{6.29}$$

Now, Lemma 6.7 yields $E\left[\zeta' Y^n_{J+}|\mathcal{G}_{S_{r_J}}\right] \to^p E\left[\zeta' Y_{J+}|\mathcal{G}^{(0)}_{S_{r_J}}\right]$. Moreover, setting $\Omega_n = \{T_{i(S^\dagger_{r_J})+\bar{k}_n} < S_{r_J}\}$, we have

$$E\left[\zeta' \underline{Y}^n_{J-}Y^n_{J+}1_{\Omega_n}|\mathcal{G}_{S^\dagger_{r_J}}\right] = E\left[\underline{Y}^n_{J-}E\left[\zeta' Y^n_{J+}|\mathcal{G}_{S_{r_J}}\right]1_{\Omega_n}|\mathcal{G}_{S^\dagger_{r_J}}\right].$$

Since $\lim_n P(\Omega_n) = 1$ by Lemma 6.4, the boundedness of ζ' and $F_{J\pm}$ and the bounded convergence theorem imply that

$$E\left[\zeta' \underline{Y}^n_{J-}Y^n_{J+}|\mathcal{G}_{S^\dagger_{r_J}}\right] - E\left[\zeta'' \underline{Y}^n_{J-}|\mathcal{G}_{S^\dagger_{r_J}}\right] \to^p 0, \tag{6.30}$$

where $\zeta'' = E\left[\zeta' Y_{r_J+}|\mathcal{G}^{(0)}_{S_{r_J}}\right]$. On the other hand, Lemma 6.7 again yields $E\left[\zeta'' \underline{Y}^n_{J-}|\mathcal{G}_{S^\dagger_{r_J}}\right] \to^p E\left[\zeta'' Y_{J-}|\mathcal{G}^{(0)}_{S_{r_J-}}\right]$. Since $Y^n_{J-} - \underline{Y}^n_{J-} \to^p 0$ and ζ'' and F_{J-} are bounded, the bounded convergence theorem again implies that

$$E\left[\zeta'' \underline{Y}^n_{J-}|\mathcal{G}_{S^\dagger_{r_J}}\right] \to^p E\left[\zeta'' Y_{J-}|\mathcal{G}^{(0)}_{S_{r_J-}}\right]. \tag{6.31}$$

(6.30) and (6.31) yield (6.29).

Step 7) Set $\Omega'_n = \{S_{r_{J-1}} + k_n\bar{r}_n < S^\dagger_{r_J}\}$ if $J > 1$ and $\Omega'_n = \Omega$ otherwise. Then we have

$$E\left[\zeta f_1(\check{\mathbf{L}}^n) f_2(\widetilde{\mathcal{W}}(\beta))\textstyle\prod_{j=1}^J Y^n_{j-}Y^n_{j+};\Omega'_n\right] = E\left[f_1(\check{\mathbf{L}}^n)\textstyle\prod_{j=1}^{J-1} Y^n_{j-}Y^n_{j+} E\left[\zeta' Y^n_{J-}Y^n_{J+}|\mathcal{G}_{S^\dagger_{r_J}}\right];\Omega'_n\right].$$

Therefore, by (6.28) and the boundedness of ζ', f_1 and $F_{j\pm}$ we obtain

$$E\left[\zeta f_1(\check{\mathbf{L}}^n) f_2(\widetilde{\mathcal{W}}(\beta))\textstyle\prod_{j=1}^J Y^n_{j-}Y^n_{j+};\Omega'_n\right] - E\left[\check{\zeta} f_1(\check{\mathbf{L}}^n)\textstyle\prod_{j=1}^{J-1} Y^n_{j-}Y^n_{j+};\Omega'_n\right] \to 0,$$

where $\check{\zeta} = E\left[\zeta' Y_{J-}Y_{J+}|\mathcal{G}^{(0)}_{S_{r_J-}}\right]$. Since $\lim_n P(\Omega'_n) = 1$, we conclude that

$$E\left[\zeta f_1(\check{\mathbf{L}}^n) f_2(\widetilde{\mathcal{W}}(\beta))\textstyle\prod_{j=1}^J Y^n_{j-}Y^n_{j+}\right] - E\left[\check{\zeta} f_1(\check{\mathbf{L}}^n)\textstyle\prod_{j=1}^{J-1} Y^n_{j-}Y^n_{j+}\right] \to 0. \tag{6.32}$$

Step 8) Now we are ready to prove (6.25). From (6.26) and (6.32) it remains to prove

$$E\left[\check{\zeta} f_1(\check{\mathbf{L}}^n)\textstyle\prod_{j=1}^{J-1} Y^n_{j-}Y^n_{j+}\right] \to E\left[\zeta f_1(\mathcal{W}^{S_{r_J}}) f_2(\widetilde{\mathcal{W}}(\beta))\textstyle\prod_{j=1}^J Y_{j-}Y_{j+}\right]. \tag{6.33}$$

By the assumption of the induction we have

$$(n^{1/4}\mathbf{L}[M]^n, (z^n_{r_j-}, z'^n_{r_j-}, z^n_{r_j+}, z'^n_{r_j+})_{j=1}^{J-1}) \to^{d_s} (\mathcal{W}, (z_{r_j-}, z'_{r_j-}, z_{r_j+}, z'_{r_j+})_{j=1}^{J-1})$$

in $\mathbb{D}^{d\times d}\times\mathbb{R}^{2(d'+d)(J-1)}$ as $n\to\infty$. Therefore, by Lemma 6.8 and the continuous mapping theorem we obtain

$$(\check{\mathbf{L}}^n,(z^n_{r_j-},z'^n_{r_j-},z^n_{r_j+},z'^n_{r_j+})_{j=1}^{J-1})\to^{d_s}(\mathcal{W}^{S_{r_J}},(z_{r_j-},z'_{r_j-},z_{r_j+},z'_{r_j+})_{j=1}^{J-1})$$

in $\mathbb{D}^{d\times d}\times\mathbb{R}^{2(d'+d)(J-1)}$. This implies that $E\left[\check{\zeta}f_1(\check{\mathbf{L}}^n)\prod_{j=1}^{J-1}Y^n_{j-}Y^n_{j+}\right]\to E\left[\check{\zeta}f_1(\mathcal{W}^{S_{r_J}})\prod_{j=1}^{J-1}Y_{j-}Y_{j+}\right]$. Now, since $\zeta f_2(\widetilde{\mathcal{W}}(\beta))$ is independent of $Y_{J\pm}$ by construction, we have $\check{\zeta}=E\left[\zeta f_2(\widetilde{\mathcal{W}}(\beta))Y_{J-}Y_{J+}|\mathcal{G}^{(0)}_{S_{r_J-}}\right]$. Moreover, since $\zeta f_2(\widetilde{\mathcal{W}}(\beta))Y_{J-}Y_{J+}$ is independent of $Y_{1\pm},\dots,Y_{(J-1)\pm}$ by construction, we conclude that

$$E\left[\check{\zeta}f_1(\mathcal{W}^{S_{r_J}})\prod_{j=1}^{J-1}Y_{j-}Y_{j+}\right]=E\left[\zeta f_1(\mathcal{W}^{S_{r_J}})f_2(\widetilde{\mathcal{W}}(\beta))\prod_{j=1}^{J}Y_{j-}Y_{j+}\right],$$

hence we obtain (6.33). □

Proof of Proposition 6.2. First, by Propositions 6.4–6.7 as well as properties of stable convergence, we have

$$\left(n^{1/4}\left(\mathbf{\Xi}[C(m)]^n-[M,M]-\frac{\psi_1}{\psi_2k_n^2}[Y,Y]^n\right),(\eta_-(n,r),\eta'_-(n,r),\eta_+(n,r),\eta'_+(n,r))_{r\geq1}\right)$$
$$\to^{d_s}\left(\mathcal{W},(\sigma_{S_r-}z^n_{r-},z'^n_{r-},\sigma_{S_r}z^n_{r+},z'^n_{r+})_{r\geq1}\right).$$

The second claim immediately follows from this convergence. On the other hand, since Eqs.(6.5)–(6.6) holds on the set $\Omega_n(t,m)$ (recall that $\Omega_n(t,m)$ is defined at the beginning of Section 6.2.1), we obtain

$$n^{1/4}\left(\mathbf{\Xi}[X(m)]^n_t-\left(\Xi^{(k,l)}_{g,g}(J(m),J(m))^n_t\right)_{1\leq k,l,\leq d}-[M,M]_t-\frac{\psi_1}{\psi_2k_n^2}[Y,Y]^n_t\right)\to^{d_s}\mathcal{W}_t+\mathcal{Z}(m)_t \tag{6.34}$$

by the continuos mapping theorem and the fact that (3.2) yields $\lim_{n\to\infty}P(\Omega_n(t,m))=1$.

Next, let $\Omega'_n(m)$ be the set on which $|S_{r_1}-S_{r_2}|>k_n\bar{r}_n$ for any $r_1,r_2\in\mathcal{R}_m$ such that $r_1\neq r_2$ and $S_{r_1},S_{r_2}<\infty$. Then we have

$$\Xi^{(k,l)}_{g,g}(J(m),J(m))^n_t=\frac{1}{\psi_2}\sum_{r\in\mathcal{R}_m:S_r\leq t}\left(\frac{1}{k_n}\sum_{i=1}^{k_n-1}g^n_ig^n_i\right)\Delta X^k_{S_r}\Delta X^l_{S_r}.$$

on the set $\Omega'_n(m)\cap\Omega_n(t,m)$. Since $\frac{1}{k_n}\sum_{i=1}^{k_n-1}g^n_ig^n_i=\psi_2+O(k_n^{-1})$ by the Lipschitz continuity of g and $\lim_nP(\Omega'_n(m)\cap\Omega_n(t,m))=1$, we obtain

$$n^{1/4}\left\{\Xi^{(k,l)}_{g,g}(J(m),J(m))^n_t-\sum_{r\in\mathcal{R}_m:S_r\leq t}\Delta X^k_{S_r}\Delta X^l_{S_r}\right\}\to^p0. \tag{6.35}$$

Finally, since

$$[X(m)^k,X(m)^l]_t=[M^k,M^l]_t+\sum_{r\in\mathcal{R}_m:S_r\leq t}\Delta X^k_{S_r}\Delta X^l_{S_r},$$

(6.34) and (6.35) imply the first claim of the proposition. □

6.3 Proof of Proposition 6.3

We decompose the target quantity as

$$\mathbf{\Xi}[X]^{n,kl}_t-\mathbf{\Xi}[X(m)]^{n,kl}_t=\left\{\Xi^{(k,l)}_{g,g}(X,X)^n-\Xi^{(k,l)}_{g,g}(X(m),X(m))^n\right\}+\Xi^{(k,l)}_{g,g'}(Z(m),\mathfrak{E})^n+\Xi^{(l,k)}_{g,g'}(Z(m),\mathfrak{E})^n. \tag{6.36}$$

We start by proving the negligibility of the second and the third terms in the right side of the above equation, which can be shown by an easy calculation:

Lemma 6.10. *Under the assumptions of Proposition 6.3, it holds that*

$$\limsup_{m\to\infty}\limsup_{n\to\infty}P\left(n^{1/4}\left|\Xi^{(k,l)}_{g,g'}(Z(m),\mathfrak{E})^n_t\right|>\eta\right)=0$$

for any $t,\eta>0$.

Proof. First, since $E_0\left[\overline{\mathfrak{E}}(g')^l_i\overline{\mathfrak{E}}(g')^l_j\right]=0$ if $|i-j|\geq k_n$ and $|E_0\left[\overline{\mathfrak{E}}(g')^l_i\overline{\mathfrak{E}}(g')^l_j\right]|\lesssim k_n^{-1}$ by [SA3] and the definition of $\mathfrak{E}$, we have

$$E_0\left[\left|n^{1/4}\Xi^{(k,l)}_{g,g'}(Z(m),\mathfrak{E})^n_t\right|^2\right]=\frac{\sqrt{n}}{\psi_2^2k_n^2}\sum_{\substack{i,j=1\\|i-j|<k_n}}^{N^n_t-k_n+1}\overline{Z(m)}(g)^k_i\overline{Z(m)}(g)^k_jE_0\left[\overline{\mathfrak{E}}(g')^l_i\overline{\mathfrak{E}}(g')^l_j\right]$$

$$\lesssim\frac{\sqrt{n}}{k_n^3}\sum_{\substack{i,j=1\\|i-j|<k_n}}^{N^n_t-k_n+1}\overline{Z(m)}(g)^k_i\overline{Z(m)}(g)^k_j\lesssim\frac{\sqrt{n}}{k_n^2}\sum_{i=1}^{N^n_t-k_n+1}\left|\overline{Z(m)}(g)^k_i\right|^2.$$

Next, the definition of $Z(m)$ and the optimal sampling theorem yield

$$E\left[\frac{\sqrt{n}}{k_n^2}\sum_{i=1}^{N^n_t-k_n+1}\left|\overline{Z(m)}(g)^k_i\right|^2\right]\leq E\left[\frac{\sqrt{n}}{k_n^2}\sum_{i=1}^{\infty}\left|\sum_{p=0}^{k_n-1}g^n_pZ(m)^k(I_{i+p}(t))\right|^2\right]$$

$$\leq\frac{\sqrt{n}}{k_n^2}\|g\|_\infty\overline{\gamma}_mE\left[\sum_{i=1}^{\infty}\sum_{p=0}^{k_n-1}|I_{i+p}(t)|\right]\leq\frac{\sqrt{n}}{k_n}t\|g\|_\infty\overline{\gamma}_m,$$

where $\overline{\gamma}_m=\int_{A^c_m}\gamma(z)^2\lambda(\mathrm{d}z)$. Since $\sqrt{n}/k_n=O(1)$ as $n\to\infty$ and $\lim_m\overline{\gamma}_m=0$ by the dominated convergence theorem, we conclude that

$$\limsup_{m\to\infty}\limsup_{n\to\infty}E\left[\frac{\sqrt{n}}{k_n^2}\sum_{i=1}^{N^n_t-k_n+1}\left|\overline{Z(m)}(g)^k_i\right|^2\right]=0.$$

Therefore, the Chebyshev inequality implies the desired result. □

Next we prove the negligibility of the term $\Xi^{(k,l)}_{g,g}(X,X)^n_t-\Xi^{(k,l)}_{g,g}(X(m),X(m))^n_t$. We further decompose it as

$$\Xi^{(k,l)}_{g,g}(X,X)^n_t-\Xi^{(k,l)}_{g,g}(X(m),X(m))^n_t=\Xi^{(k,l)}_{g,g}(Z(m),X)^n_t+\Xi^{(k,l)}_{g,g}(X,Z(m))^n_t-\Xi^{(k,l)}_{g,g}(Z(m),Z(m))^n_t.\tag{6.37}$$

Therefore, using the decomposition

$$X_t=X_0+B'_t+M_t+Z_t,\tag{6.38}$$

where $B'_t=\int_0^tb'_s\mathrm{d}s$, $b'_s=b_s+\int_{\{\|\delta(s,z)\|>1\}}\delta(s,z)\lambda(\mathrm{d}z)$ and $Z_t=\delta\star(\mu-\nu)_t$, it is enough to prove the negligibility of $\Xi^{(k,l)}_{g,g}(Z(m),V)^n_t$ for $V\in\{B',M,Z,Z(m)\}$. In the following we fix $V\in\{B',M,Z,Z(m)\}$.

Lemma 6.11. *Assume* [SA2]. *Then,*

(a) $\sup_{0\leq h\leq h_0}\|V_t-V_{(t-h)_+}\|=O_p(\sqrt{h_0})$ *as* $h_0\downarrow0$,

(b) $\sup_{0\leq h\leq h_0}\|[Z(m)^k,V^l]_t-[Z(m)^k,V^l]_{(t-h)_+}\|=O_p(h_0)$ *as* $h_0\downarrow0$ *for all* k,l.

Proof. The claim is evident if $V=B'$, so we assume that $V\neq B'$. Then, the Doob inequality and [SA2] yield $E[\sup_{0\leq h\leq h_0}\|V_t-V_{(t-h)_+}\|^2]\lesssim h_0$, which implies (a). On the other hand, the Kunita-Watanabe and Schwarz inequalities as well as [SA2] yield $E[\sup_{0\leq h\leq h_0}|[Z(m)^k,V^l]_t-[Z(m)^k,V^l]_{(t-h)_+}|]\lesssim h_0$, which implies (b). □

Lemma 6.12. $\sup_{1\leq q\leq N^n_t+1}|C^n_{g,g}(V)^k_q|=O_p(1)$ *as* $n\to\infty$ *for any* $t>0$ *and* $k=1,\dots,d$.

Proof. This can be shown in the same manner as the proof of Lemma 6.8 from [36]. □

Lemma 6.13. *Under the assumptions of Proposition 6.3, it holds that*

$$n^{1/4}\left\{\Xi^{(k,l)}_{g,g}(Z(m),V)^n_t-\mathbb{L}^{(k,l)}_{g,g}(Z(m),V)^n_t-[Z(m)^k,V^l]_t\right\}\to^p0$$

as $n\to\infty$ *for any* $t>0$.

Proof. Simple calculations and Lemma 6.11(a) yield

$$\Xi^{(k,l)}_{g,g}(Z(m),V)^n_t = \mathbb{L}^{(k,l)}_{g,g}(Z(m),V)^n_t + \sum_{p=1}^{N^n_t+1} c^n_{g,g}(p,p)Z(m)^k(I_p)_t V^l(I_p) + o_p(n^{-1/4})$$
$$=: \mathbb{L}^{(k,l)}_{g,g}(Z(m),V)^n_t + \mathbf{B} + o_p(n^{-1/4}),$$

hence it suffices to prove $\mathbf{B} = [Z(m)^k, V^l]_t + o_p(n^{-1/4})$. This can be shown analogously to the proof of Eq.(6.24) from [36], using Lemma 6.11(b). □

Lemma 6.14. *Under the assumptions of Proposition 6.3, it holds that*

$$\limsup_{m\to\infty}\limsup_{n\to\infty} P\left(n^{1/4}\left|\mathbb{L}^{(k,l)}_{g,g}(Z(m),V)^n_t\right| > \eta\right) = 0 \tag{6.39}$$

for any $t, \eta > 0$.

Proof. First, if $V = B'$, we can adopt an analogous argument to the proof of Lemma 6.10 from [36] and deduce $n^{1/4}\mathbb{L}^{(k,l)}_{g,g}(Z(m),V)^n_t \to^p 0$ as $n\to\infty$ for every m, so (6.39) holds true.

Next we suppose that $V \neq B'$. It is enough to prove

$$\limsup_{m\to\infty}\limsup_{n\to\infty} P\left(n^{1/4}\left|\mathbb{M}^{(k,l)}_{g,g}(Z(m),V)^n_t\right| > \eta'\right) = 0, \tag{6.40}$$
$$\limsup_{m\to\infty}\limsup_{n\to\infty} P\left(n^{1/4}\left|\mathbb{M}^{(k,l)}_{g,g}(V,Z(m))^n_t\right| > \eta'\right) = 0 \tag{6.41}$$

for any $\eta' > 0$. Since we can prove (6.41) in a similar manner to the proof of (6.40), we only prove (6.40).

Since V is an $(\mathcal{F}^{(0)}_t)$-martingale for any n due to [SA2], by the Lenglart inequality it suffices to show that

$$\limsup_{m\to\infty}\limsup_{n\to\infty} P\left(U(n,m)_t > \eta'\right) = 0 \tag{6.42}$$

for any $\eta' > 0$, where $U(n,m)_t = \sqrt{n}\sum_{q=2}^{N^n_t+1} E[|C^n_{g,g}(Z(m))^k_q V^l(I_q)|^2|\mathcal{F}^{(0)}_{T_{q-1}}]$. To prove (6.42), for each $j \geq 1$ we set $\Lambda(j)^n_q = \{E[n|I_q||\mathcal{F}^{(0)}_{T_{q-1}}] \leq j\}$ and decompose $U(n,m)$ as

$$U(n,m)_t = \sqrt{n}\sum_{q=2}^{N^n_t+1} E\left[\left|C^n_{g,g}(Z(m))^k_q V^l(I_q)\right|^2\middle|\mathcal{F}^{(0)}_{T_{q-1}}\right]\left(1_{\Lambda(j)^n_q} + 1_{(\Lambda(j)^n_q)^c}\right)$$
$$= U(n,m,j)_t + U'(n,m,j)_t.$$

First we prove

$$\limsup_{m\to\infty}\limsup_{n\to\infty} P\left(U(n,m,j)_t > \eta'\right) = 0 \tag{6.43}$$

for any fixed j and any $\eta' > 0$. We have

$$E\left[U(n,m,j)_t\right] = \sqrt{n}E\left[\sum_{q=2}^{N^n_t+1}\left|C^n_{g,g}(Z(m))^k_q\right|^2 E\left[\langle V^l\rangle(I_q)\middle|\mathcal{F}^{(0)}_{T_{p-1}}\right]1_{\Lambda(j)^n_q}\right]$$
$$\lesssim \frac{j}{\sqrt{n}}E\left[\sum_{q=2}^{\infty}\sum_{p=(q-k_n)\vee 1}^{q-1}\left|c^n_{u,v}(p,q)\right|^2\langle Z(m)^k\rangle(I_p(t))\right] \lesssim \overline{\gamma}_m\frac{1}{\sqrt{n}}E\left[\sum_{q=2}^{\infty}\sum_{p=(q-k_n)\vee 1}^{q-1}|I_p(t)|\right] \leq \overline{\gamma}_m k_n n^{-1/2}t,$$

hence it holds that $\limsup_m\limsup_n E\left[U(n,m,j)_t\right] = 0$. Therefore, we obtain (6.43) by the Chebyshev inequality.

Next we prove

$$\limsup_{j\to\infty}\limsup_{m\to\infty}\limsup_{n\to\infty} P\left(U'(n,m,j)_t > \eta'\right) = 0 \tag{6.44}$$

for any $\eta' > 0$. Since $E[n|I_q||\mathcal{F}^{(0)}_{T_{q-1}}] = E[E[n|I_q||\mathcal{G}^{(0)}_{T_{q-1}}]|\mathcal{F}^{(0)}_{T_{q-1}}]$, we have $\Lambda(j)^n_q \supset \{E[n|I_q||\mathcal{G}^{(0)}_{T_{q-1}}] \leq j\}$. Therefore,

$$\begin{aligned} U'(n,m,j)_t &\leq \sqrt{n} \sum_{q=2}^{N^n_t+1} E\left[\left|C^n_{g,g}(Z(m))^k_q V^l(I_q)\right|^2 \left|\mathcal{F}^{(0)}_{T_{q-1}}\right.\right] \left(1_{\{G^n_{T_{q-1}} > j\}} + 1_{\{q-1 \in \mathcal{N}^n\}}\right) \\ &=: U'(n,m,j)^{(1)}_t + U'(n,m,j)^{(2)}_t. \end{aligned}$$

Since $\{U'(n,m,j)^{(1)} > 0\} \subset \{\sup_{0 \leq s \leq t} G^n_s > j\}$, we have $\limsup_j \limsup_m \limsup_n P\left(U'(n,m,j)^{(1)}_t > 0\right) = 0$. On the other hand, [SA2] and (6.2) imply that

$$U'(n,m,j)^{(2)}_t \leq \sqrt{n}\bar{r}_n \sup_{1 \leq q \leq N^n_t+1} \left|C^n_{g,g}(Z(m))^k_q\right|^2 \#(\mathcal{N}^n \cap \{q : T_q \leq t\}),$$

hence Lemma 6.12, (6.1) and [A1](i) yield $\limsup_n P(U'(n,m,j)^{(2)}_t > \eta') = 0$. Consequently, we obtain (6.44).

From (6.43) we have $\limsup_m \limsup_n P\left(U(n,m)_t > \eta'\right) \leq \limsup_m \limsup_n P\left(U'(n,m,j)_t > \eta'\right)$ for any $j \geq 1$ and any $\eta' > 0$. Hence (6.44) yields (6.42), which completes the proof. □

Proof of Proposition 6.3. (6.4) immediately follows from Eqs.(6.38) and (6.36)–(6.37) as well as Lemmas 6.10 and 6.13–6.14. (6.3) follows from the equation $\widetilde{E}\left[\left|\mathcal{Z}(m)^{kl}_t - \mathcal{Z}^{kl}_t\right|^2 \left|\mathcal{F}^{(0)}\right.\right] = \frac{1}{\psi_2^2} \sum_{r \notin \mathcal{R}_m : S_r \leq t} \left(\mathfrak{I}^{kk}_{S_r} + 2\mathfrak{I}^{kl}_{S_r} + \mathfrak{I}^{ll}_{S_r}\right)$ and the fact that $\sum_{r \notin \mathcal{R}_m : S_r \leq t} \|\Delta X_{S_r}\|^2 \to^p 0$ as $m \to \infty$. □

Acknowledgements

I am grateful to Teppei Ogihara who pointed out a problem on the mathematical construction of the noise process in a draft of this paper. This work was supported by Grant-in-Aid for JSPS Fellows.

References

[1] Aït-Sahalia, Y., Fan, J. and Xiu, D. (2010). High-frequency covariance estimates with noisy and asynchronous financial data. *J. Amer. Statist. Assoc.* **105**, 1504–1517.

[2] Aït-Sahalia, Y., Jacod, J. and Li, J. (2012). Testing for jumps in noisy high frequency data. *J. Econometrics* **168**, 207–222.

[3] Aït-Sahalia, Y. and Xiu, D. (2014). Increased correlation among asset classes: Are volatility or jumps to blame, or both? Chicago Booth Paper 14-11, Fama-Miller Center for Research in Finance, The University of Chicago, Booth School of Business. Available at SSRN: http://ssrn.com/abstract=2425676.

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

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

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

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

[8] Bibinger, M. (2012). An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: Asymptotic distribution theory. *Stochastic Process. Appl.* **122**, 2411–2453.

[9] Bibinger, M., Hautsch, N., Malec, P. and Reiß, M. (2014). Estimating the quadratic covariation matrix from noisy observations: local method of moments and efficiency. *Ann. Statist.* **42**, 80–114.

[10] Bibinger, M. and Mykland, P. A. (2013). Inference for multi-dimensional high-frequency data: Equivalence of methods, central limit theorems, and an application to conditional independence testing. SFB 649 Discussion Paper 2013-006, Humboldt-Universität zu Berlin.

[11] Bibinger, M. and Reiß, M. (2014). Spectral estimation of covolatility from noisy observations using local weights. *Scand. J. Stat.* **41**, 23–50.

[12] Bibinger, M. and Vetter, M. (2013). Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps. SFB 649 Discussion Paper 2013-029, Humboldt-Universität zu Berlin. To appear in Ann. Inst. Statist. Math.

[13] Bibinger, M. and Winkelmann, L. (2015). Econometrics of co-jumps in high-frequency data with noise. *J. Econometrics* **184**, 361–378.

[14] Christensen, K., Kinnebrock, S. and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *J. Econometrics* **159**, 116–133.

[15] Christensen, K., Podolskij, M. and Vetter, M. (2013). On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. *J. Multivariate Anal.* **120**, 59–84.

[16] Diop, A., Jacod, J. and Todorov, V. (2013). Central Limit Theorems for approximate quadratic variations of pure jump Itô semimartingales. *Stochastic Process. Appl.* **123**, 839–886.

[17] Dohnal, G. (1987). On estimating the diffusion coefficient. *J. Appl. Probab.* **24**, 105–114.

[18] Fukasawa, M. (2010). Realized volatility with stochastic sampling. *Stochastic Process. Appl.* **120**, 829–852.

[19] Genon-Catalot, V. and Jacod, J. (1993). On the estimation of the diffusion coefficient for multidimensional diffusion. *Ann. Inst. Henri Poincaré Probab. Stat.* **29**, 119–151.

[20] Hall, P. (1977). Martingale invariance principles. *Ann. Probab.* **5**, 875–887.

[21] Hautsch, N. and Podolskij, M. (2013). Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence. *J. Bus. Econom. Statist.* **31**, 165–183.

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

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

[24] Hayashi, T. and Yoshida, N. (2011). Nonsynchronous covariation process and limit theorems. *Stochastic Process. Appl.* **121**, 2416–2454.

[25] Horn, R. A. and Johnson, C. R. (1991). *Topics in matrix analysis.* Cambridge University Press.

[26] Ikeda, N. and Watanabe, S. (1989). *Stochastic differential equations and diffusion processes.* North-Holland Publishing Company, 2nd edn.

[27] Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Process. Appl.* **118**, 517–559.

[28] Jacod, J., Li, Y., Mykland, P. A., Podolskij, M. and Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.* **119**, 2249–2276.

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

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

[31] Jacod, J. and Protter, P. (2012). *Discretization of processes.* Springer.

[32] Jacod, J. and Shiryaev, A. N. (2003). *Limit theorems for stochastic processes.* Springer, 2nd edn.

[33] Kawai, R. and Masuda, H. (2011). On simulation of tempered stable random variates. *J. Comput. Appl. Math.* **235**, 2873–2887.

[34] Koike, Y. (2014). An estimator for the cumulative co-volatility of asynchronously observed semimartingales with jumps. *Scand. J. Stat.* **41**, 460–481.

[35] Koike, Y. (2014). Limit theorems for the pre-averaged Hayashi-Yoshida estimator with random sampling. *Stochastic Process. Appl.* **124**, 2699–2753.

[36] Koike, Y. (2015). Time endogeneity and an optimal weight function in pre-averaging covariance estimation. Preprint, Available at arXiv: http://arxiv.org/abs/1403.7889v2.

[37] Li, Y., Mykland, P. A., Renault, E., Zhang, L. and Zheng, X. (2014). Realized volatility when sampling times are possibly endogenous. *Econometric Theory* **30**, 580–605.

[38] Li, Y., Zhang, Z. and Zheng, X. (2013). Volatility inference in the presence of both endogenous time and microstructure noise. *Stochastic Process. Appl.* **123**, 2696–2727.

[39] Lo, A. W. and MacKinlay, A. C. (1990). An econometric analysis of nonsynchronous trading. *J. Econometrics* **45**, 181–211.

[40] Malliavin, P. and Mancino, M. E. (2002). Fourier series method for measurement of multivariate volatilities. *Finance Stoch.* **6**, 49–61.

[41] Meyer, P. A. (1966). *Probability and potentials*. Blaisdell Publishing Company.

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

[43] Mykland, P. A. and Zhang, L. (2012). The econometrics of high-frequency data. In M. Kessler, A. Lindner and M. Sørensen, eds., *Statistical methods for stochastic differential equations*, chap. 3. CRC Press.

[44] Ogihara, T. and Yoshida, N. (2014). Quasi-likelihood analysis for nonsynchronously observed diffusion processes. *Stochastic Process. Appl.* **124**, 2954–3008.

[45] Podolskij, M. and Vetter, M. (2009). Bipower-type estimation in a noisy diffusion setting. *Stochastic Process. Appl.* **119**, 2803–2831.

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

[47] Reiß, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *Ann. Statist.* **39**, 772–802.

[48] Renault, E., van der Heijden, T. and Werker, B. J. (2014). The dynamic mixed hitting-time model for multiple transaction prices and times. *J. Econometrics* **180**, 233–250.

[49] Renault, E. and Werker, B. J. (2011). Causality effects in return volatility measures with random times. *J. Econometrics* **160**, 272–279.

[50] Rosenbaum, M. and Tankov, P. (2011). Asymptotic results for time-changed Lévy processes sampled at hitting times. *Stochastic Process. Appl.* **121**, 1607–1632.

[51] Shephard, N. and Xiu, D. (2014). Econometric analysis of multivariate realised QML: estimation of the covariation of equity prices under asynchronous trading. Chicago Booth Research Paper 12-14, The University of Chicago Booth School of Business. Available at SSRN: http://ssrn.com/abstract=2045571.

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

[53] Zhang, L. (2011). Estimating covariation: Epps effect, microstructure noise. *J. Econometrics* **160**, 33–47.

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