Econometric Theory, **32**, 2016, 533–611.
doi:10.1017/S0266466614000954

ESTIMATION OF INTEGRATED COVARIANCES IN THE SIMULTANEOUS PRESENCE OF NONSYNCHRONICITY, MICROSTRUCTURE NOISE AND JUMPS

YUTA KOIKE
The Institute of Statistical Mathematics

We propose a new estimator for the integrated covariance of two Itô semimartingales observed at a high frequency. This new estimator, which we call the pre-averaged truncated Hayashi–Yoshida estimator, enables us to separate the sum of the co-jumps from the total quadratic covariation even in the case that the sampling schemes of two processes are nonsynchronous and the observation data are polluted by some noise. We also show the asymptotic mixed normality of this estimator under some mild conditions allowing infinite activity jump processes with finite variations, some dependency between the sampling times and the observed processes as well as a kind of endogenous observation error. We examine the finite sample performance of this estimator using a Monte Carlo study and we apply our estimators to empirical data, highlighting the importance of accounting for jumps even in an ultra-high frequency framework.

1. INTRODUCTION

In the past years there has been a considerable development in statistical inferences for the quadratic covariations of semimartingales observed at a high frequency. This was mainly motivated by financial application because price processes need to follow a semimartingale under the no-arbitrage assumption (see Delbaen and Schachermayer,1994 for instance) and technological developments made high frequency data commonly available. In general the quadratic covariation of two semimartingales consists of two sources; the continuous martingale parts and the co-jumps of the semimartingales. Recently many authors have indicated that separating these two sources benefits various areas of finance such as

I am indebted to Professor Nakahiro Yoshida for his encouragement to my research and valuable suggestions which led to considerable improvements of the paper. I am also grateful to two anonymous referees and the co-editor Eric Renault for their extremely insightful comments that significantly improved this paper. This work was partly supported by Grant-in-Aid for JSPS Fellows and the Program for Leading Graduate Schools, MEXT, Japan.
Address correspondence to Yuta Koike, Risk Analysis Research Center, The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan; e-mail: kyuta@ism.ac.jp.

volatility forecasting (Andersen, Bollerslev, and Diebold, 2007), credit risk management (Cont and Kan, 2011), the construction of a hedging portfolio (Todorov and Bollerslev, 2010), and so on. Motivated by these reasons, this paper will focus on disentangling these two components of the quadratic covariations of two semimartingales by using high-frequency observation data.

Let Z^1 and Z^2 be two Itô semimartingales and let $(S^i)_{i\in\mathbb{Z}_+}$ be a sequence of stopping times that is increasing a.s., $S^i \uparrow \infty$, and $S^0 = 0$. Then it is well known in the classic stochastic calculus that

$$\sum_{i:S^i\le t}\left(Z^1_{S^i}-Z^1_{S^{i-1}}\right)\left(Z^2_{S^i}-Z^2_{S^{i-1}}\right)\to^p\left[Z^1,Z^2\right]_t \tag{1}$$

for any $t>0$, provided $\sup_{i\in\mathbb{N}}(S^i\wedge t - S^{i-1}\wedge t)\to^p 0$. Therefore, if we observe Z^1 and Z^2 at the time S^i for every i, we can use the statistic on the left side of the equation (1) (which is called the *realized covariance*) as a consistent estimator of the quadratic covariation $[Z^1,Z^2]_t$ of Z^1 and Z^2. Given that

$$\left[Z^1,Z^2\right]_t=\left\langle Z^{1,c},Z^{2,c}\right\rangle_t+\sum_{0\le s\le t}\Delta Z^1_s\Delta Z^2_s, \tag{2}$$

our aim will be achieved by constructing an estimator for the quantity $\langle Z^{1,c},Z^{2,c}\rangle_t$ which we call the *integrated covariance* of Z^1 and Z^2. In the present situation we have the observation data $(Z^1_{S^i}+Z^2_{S^i})_{i\in\mathbb{Z}_+}$ and $(Z^1_{S^i}-Z^2_{S^i})_{i\in\mathbb{Z}_+}$, so that the problem results in the univariate case due to the polarization identity $\langle Z^{1,c},Z^{2,c}\rangle=(\langle Z^{1,c}+Z^{2,c}\rangle-\langle Z^{1,c}-Z^{2,c}\rangle)/4$. As a consequence, we can benefit from vast numbers of studies on jump detection in financial high-frequency data for a single asset. For example, Barndorff-Nielsen and Shephard (2004a) used such a method based on the *bipower technique* introduced in Barndorff-Nielsen and Shephard (2004b) and proposed an estimator called the *realized bipower covariation*. On the other hand, there are several approaches which directly treat the multivariate data; see Mancini and Gobbi (2012) and Boudt, Croux, and Laurent (2011) for examples.

In real financial markets, however, some difficulties caused by the so-called *market microstructure* confront us. In the present context there are two major topics related to them: one is the nonsynchronicity of observation times and the other is a kind of observation error called *microstructure noise*. In recent years the simultaneous treatment of these two problems, which is based on the combination of methods for dealing with each individual one, has been established by many authors in the case that jumps are absent. See Aït-Sahalia, Fan, and Xiu (2010), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011), Bibinger (2011), Christensen, Kinnebrock, and Podolskij (2010), Shephard and Xiu (2012), and Zhang (2011) for examples.

Since the main concern of the present article is estimating integrated covariances separately from jumps under the influence of the market microstructure, it is necessary to focus on the simultaneous treatments of jumps and the market microstructure problems. Regarding this issue, there are two important contributions which we will benefit from. First, in the case that observation times are nonsynchronous and observed processes are not contaminated with microstructure noise, but may contain jumps. Mancini and Gobbi (2012) combined the Hayashi–Yoshida method proposed in Hayashi and Yoshida (2005) (to deal with nonsynchonicity) with the thresholding jump-detection technique (proposed independently in Mancini, 2001 and Shimizu, 2003) to construct a consistent estimator for the integrated covariance. Second, in the univariate case Podolskij and Vetter (2009b) proposed a new method for dealing with microstructure noise called the *pre–averaging method*. Combining this method with the bipower technique, they introduced a class of consistent estimators for the integrated volatility which can withstand both jumps and microstructure noise (their method has been further investigated in Jacod, Li, Mykland, Podolskij, and Vetter, 2009 and Podolskij and Vetter, 2009 for example).

In the present article we investigate the methodology accommodated to a situation where all of the above problems are present simultaneously. That is, we consider two Itô semimartingales which are observed at stopping times in a nonsynchronous manner and contaminated by noise. Then we develop a method for estimating their integrated covariance separately from the sum of their co-jumps. For this purpose, we combine the Hayashi–Yoshida method (to deal with the nonsynchronicity of the observation times) and the preaveraging method (to remove the noise) with the thresholding technique (to separate the jumps) and consider a class of statistics called the *preaveraged truncated Hayashi–Yoshida estimator*. We prove the consistency and the asymptotic mixed normality of the preaveraged truncated Hayashi–Yoshida estimator under a very general situation allowing the presence of infinite activity jumps, some dependency between the observation times and the observed processes as well as a kind of endogenous noise.

Our new estimator is closely related to the one proposed in Wang, Liu, and Liu (2013) recently. In fact, based on the above idea, they constructed a consistent estimator for the integrated covariance in the simultaneous presence of the above three elements. The difference between their estimator and ours is that we use an additional synchronization step based on refresh times for the construction of the estimator. This type of synchronization is often used in the literature, e.g. Aït-Sahalia et al. (2010), Barndorff-Nielsen et al. (2011), Bibinger (2011), and Zhang (2011). This additional step enables us to develop fully an asymptotic theory for the estimator in a general setting, which is the main contribution of the present article. In particular, in contrast to Wang et al. (2013), we have a feasible central limit theorem for the estimator, which enables us to construct confidence intervals of the estimator.

Recent empirical studies suggest that most jumps at a relatively low frequency (e.g., 5-minute) consist of consecutive small returns if sampled at an ultra high frequency (see Christensen, Oomen, and Podolskij, 2011a for instance). However, such consecutive small returns still involve relatively large ones in some cases. We cannot capture such large returns precisely at relatively low frequencies because they are "contaminated" by many small returns. This suggests the potential usefulness of our estimator. We will give a simple empirical study to reinforce this conjecture.

This paper is organized as follows: In Section 2 we briefly review on the results about the asymptotic properties of the preaveraged Hayashi–Yoshida estimator in the continuous Itô semimartingale setting. In Section 3 we present the construction of our estimator and the main results in this paper. We discuss some topics for the statistical application to finance of our estimator in Section 4, while Section 5 provides some numerical experiments to illustrate the finite sample properties of our estimator. Section 6 presents an empirical application to high frequency stock returns data. The Appendix contains the remaining proofs of the results as well as the list of the assumptions introduced in this paper.

2. A BRIEF REVIEW OF THE CONTINUOUS CASE

We start by introducing an appropriate stochastic basis on which our observation data are defined. Let $\mathcal{B}^{(0)} = \left(\mathbf{\Omega}^{(0)}, \mathcal{F}^{(0)}, \mathbf{F}^{(0)} = (\mathcal{F}_t^{(0)})_{t\in\mathbb{R}_+}, P^{(0)}\right)$ be a stochastic basis. Namely, $(\mathbf{\Omega}, \mathcal{F}, P)$ denotes a probability space and $\mathbf{F}$ denotes a filtration of $\mathcal{F}$. For any $t \in \mathbb{R}_+$ we have a transition probability $Q_t\left(\omega^{(0)}, \mathrm{d}z\right)$ from $\left(\mathbf{\Omega}^{(0)}, \mathcal{F}_t^{(0)}\right)$ into $\mathbb{R}^2$, which satisfies $\int z Q_t(\omega^{(0)}, \mathrm{d}z) = 0$. We endow the space $\mathbf{\Omega}^{(1)} = (\mathbb{R}^2)^{[0,\infty)}$ with the product Borel σ-field $\mathcal{F}^{(1)}$ and with the probability $Q\left(\omega^{(0)}, \mathrm{d}\omega^{(1)}\right)$ which is the product $\otimes_{t\in\mathbb{R}_+} Q_t(\omega^{(0)}, \cdot)$. We also call $(\zeta_t)_{t\in\mathbb{R}_+}$ the "canonical process" on $(\mathbf{\Omega}^{(1)}, \mathcal{F}^{(1)})$ and the filtaration $\mathcal{F}_t^{(1)} = \sigma(\zeta_s; s \leq t)$. Then we consider the stochastic basis $\mathcal{B} = (\mathbf{\Omega}, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$ defined as follows:

$$\mathbf{\Omega} = \mathbf{\Omega}^{(0)} \times \mathbf{\Omega}^{(1)}, \qquad \mathcal{F} = \mathcal{F}^{(0)} \otimes \mathcal{F}^{(1)}, \qquad \mathcal{F}_t = \cap_{s>t} \mathcal{F}_s^{(0)} \otimes \mathcal{F}_s^{(1)},$$
$$P\left(\mathrm{d}\omega^{(0)}, \mathrm{d}\omega^{(1)}\right) = P^{(0)}\left(\mathrm{d}\omega^{(0)}\right) Q\left(\omega^{(0)}, \mathrm{d}\omega^{(1)}\right).$$

Any variable or process which is defined on either $\mathbf{\Omega}^{(0)}$ or $\mathbf{\Omega}^{(1)}$ can be considered in the usual way as a variable or a process on $\mathbf{\Omega}$. In terms of financial applications, the space $\mathbf{\Omega}^{(0)}$ stands for latent log-price processes, while the space $\mathbf{\Omega}^{(1)}$ stands for microstructure noise. In particular, the process $(\zeta_t)_{t\in\mathbb{R}_+}$ corresponds to the microstructure noise which is centered and mutually independent conditionally on the latent processes.

Next we introduce our observation data. There are two continuous semimartingales $X^1 = \left(X_t^1\right)_{t\in\mathbb{R}_+}$ and $X^2 = \left(X_t^2\right)_{t\in\mathbb{R}_+}$ on $\mathcal{B}^{(0)}$ with canonical decompositions

$$X^l = A^l + M^l, \qquad l = 1, 2, \tag{3}$$

where A^1 and A^2 are continuous $\mathbf{F}^{(0)}$-adapted processes with locally finite variations, while M^1 and M^2 are continuous $\mathbf{F}^{(0)}$-local martingales. In financial applications X^1 and X^2 represent latent log-price processes of two assets. Furthermore, observation times are modeled as two sequences of $\mathbf{F}^{(0)}$-stopping times $(S^i)_{i\in\mathbb{Z}_+}$ and $(T^j)_{j\in\mathbb{Z}_+}$ that are increasing a.s.,

$$S^i \uparrow \infty \qquad \text{and} \qquad T^j \uparrow \infty. \tag{4}$$

As a matter of convenience we set $S^{-1} = T^{-1} = 0$. These stopping times implicitly depend on a parameter $n \in \mathbb{N}$, which represents the observation frequency. Let ξ' be a constant satisfying $0 < \xi' < 1$. In this paper, we will always assume that

$$\begin{aligned} r_n(t) :&= \sup_{i\in\mathbb{Z}_+} [(S^i \wedge t) - (S^{i-1} \wedge t)] \vee \sup_{j\in\mathbb{Z}_+} [(T^j \wedge t) - (T^{j-1} \wedge t)] \\ &= o_p\left(n^{-\xi'}\right) \end{aligned} \tag{5}$$

as $n \to \infty$ for any $t \in \mathbb{R}_+$. Note that $r_n(t)$ represents the maximum duration of the sampling times up to the time t. The condition (5) is equivalent to assuming that

$$\sup_{i\geq 0: S^i \leq t} \left(S^i - S^{i-1}\right) \vee \sup_{j\geq 0: T^j \leq t} \left(T^j - T^{j-1}\right) = o_p\left(n^{-\xi'}\right)$$

as $n \to \infty$ for any $t \in \mathbb{R}_+$.

Remark 2.1. Condition (5) implies that $r_n(t)$ tends to zero. In real applications, this corresponds to the situation that $r_n(t)$ is sufficiently small i.e., data frequently observed in the fixed interval $[0, t]$. Such situations appear in the analysis of high frequency financial data. More concretely, let us consider a data set of intraday transactions for some stocks traded on the New York Stock Exchange (NYSE). Such stocks are often traded in seconds. We model this data as discrete observations of continuous-time stochastic processes in a fixed interval, say, [0, 1]. Since the NYSE is open for 6.5 hours, in this situation 1 second corresponds to 1/23400 of a trading day. 1/23400 seems to be sufficiently small, so we can make it valid to assume that $r_n(t)$ tends to zero. In fact, it is standard in the literature to assume that $r_n(t)$ tends to zero; see Barndorff-Nielsen et al. (2011) and Christensen, Kinnebrock, and Podolskij (2010).

In order to derive central limit theorems for the estimators introduced in the following, we will impose the following additional assumption regarding the dependency between the sampling times and the observed processes, which is an analog to the condition [A2] in Hayashi and Yoshida (2011) (known as the strong predictability condition). Let ξ be a positive constant satisfying $\frac{1}{2} < \xi < 1$.

[A1] For any $n, i \in \mathbb{N}$, S^i and T^i are $\mathbf{G}^{(n)}$-stopping times, where $\mathbf{G}^{(n)} = \left(\mathcal{G}_t^{(n)}\right)_{t\in\mathbb{R}_+}$ is the filtration given by $\mathcal{G}_t^{(n)} = \mathcal{F}^{(0)}_{(t-n^{-\xi+1/2})_+}$ for $t \in \mathbb{R}_+$.

Remark 2.2. When S^i and T^i are independent of X^1, X^2, $\underline{X}^1$, and $\underline{X}^2$ we may assume S^i and T^i are $\mathcal{F}_0^{(0)}$-measurable without loss of generality. This is

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

because X^1, X^2, $\underline{X}^1$, and $\underline{X}^2$ are still continuous semimartingales with the canonical decompositions (3) and (8) with respect to the filtration generated by $\mathbf{F}^{(0)}$ and $S^0, T^0, S^1, T^1, \ldots$ (for all n). Note that the condition [A2], which is a regularity condition on the sampling times introduced below, uses the filtration $\mathbf{H}^n$ different from $\mathbf{F}^{(0)}$, so that expansions of the filtration $\mathbf{F}^{(0)}$ causes no problem on [A2]. This implies that we can always assume that sampling times independent of X^1, X^2, $\underline{X}^1$, and $\underline{X}^2$ satisfy the condition [A1].

Another example of sampling times satisfying [A1] (and [A2]) is given as follows. Let (S^i) and (T^j) be arbitrary sequences of $\mathbf{F}^{(0)}$-stopping times satisfying [A2]. Then the sequences $\left(S^i + n^{-\xi+1/2}\right)$ and $\left(T^j + n^{-\xi+1/2}\right)$ obviously satisfy [A1] and [A2]. This example implies that [A1] intuitively means that sampling times are determined the short time $\left(= n^{-\xi+1/2}\right)$ later than changes of latent processes. In application, this type of delay will occur while a trader is asking the broker to trade in a financial market.

On the other hand, [A1] rules out the presence of the instantaneous causality between the latent process and the durations of the observation times. For example, if X^1 is a martingale, then [A1] implies that the covariance between the simultaneous latent return and duration is zero, i.e. $E\left[\left(S^i - S^{i-1}\right)\left(X^1_{S^i} - X^1_{S^{i-1}}\right)\right] = 0$. The recent work of Li, Mykland, Renault, Zhang, and Zheng (2013) pointed out that the non–zero relation of this covariance plays an important role in the volatility inference as well as provided evidence that such a relation exists in financial data. Renault and Werker (2011) also discussed the effect of the instantaneous causality between the volatility and the durations on the volatility inference. See also Fukasawa (2010), Fukasawa and Rosenbaum (2012), Li, Zhang, and Zheng (2013), and Robert and Rosenbaum (2012) for other related works. In view of these works, the condition [A1] is still restrictive in practice and needs to be weakened. This topic, however, exceeds the scope of this paper and is left to future research.

The processes X^1 and X^2 are observed at the sampling times (S^i) and (T^j) with observation errors $(U^1_{S^i})_{i\in\mathbb{Z}_+}$ and $(U^2_{T^j})_{j\in\mathbb{Z}_+}$ respectively. In this paper, we assume that the observation errors have the following representations:

$$U^1_{S^i} = \sqrt{n}\left(\underline{X}^1_{S^i} - \underline{X}^1_{S^{i-1}}\right) + \zeta^1_{S^i}, \qquad U^2_{T^j} = \sqrt{n}\left(\underline{X}^2_{T^j} - \underline{X}^2_{T^{j-1}}\right) + \zeta^2_{T^j}. \tag{6}$$

Here, ζ^k_t denotes the k-th component of the noise process ζ_t introduced in the above, while $\underline{X}^1$ and $\underline{X}^2$ are two continuous semimartingales on $\mathcal{B}^{(0)}$. After all, we have the observation data $\mathsf{X}^1 = \left(\mathsf{X}^1_{S^i}\right)_{i\in\mathbb{Z}_+}$ and $\mathsf{X}^2 = \left(\mathsf{X}^2_{T^j}\right)_{j\in\mathbb{Z}_+}$ of the form

$$\mathsf{X}^1_{S^i} = X^1_{S^i} + U^1_{S^i}, \qquad \mathsf{X}^2_{T^j} = X^2_{T^j} + U^2_{T^j}.$$

Remark 2.3. Model (6) contains the usual additive noise model of the white noise type. In fact, such a model can be realized by taking $\underline{X}^1 = \underline{X}^2 = 0$ and $Q_t = Q$ for all t, where Q is the distribution of the noise i.e, a probability distribution on $\mathbb{R}^2$ satisfying $\int z Q(\mathrm{d}z) = 0$. Unlike this type of model, our model allows the noise to depend on both time and latent processes, which is empirically important; see Hansen and Lunde (2006) and Kalnina and Linton (2008). In fact, model (6) has a structure similar to that of the noise model proposed by Kalnina and Linton (2008). However, there are some differences between them, as explained in the following.

First, Kalnina and Linton (2008) allow the noise ζ_t^l to have a nonzero mean $m^l(t)$, where m^l is a differentiable deterministic function. However, this difference is mathematically nonsense because we can always assume $m^l(t) \equiv 0$ by replacing X^l with $X^l + m^l$ as long as m is of finite variation (this is always true if m is absolutely continuous).

Second, they also allow the variance of the noise to shrink to zero as the observation frequency n increases. Although it seems not to be difficult to incorporate such a situation into our model, we omit it for simplicity. It is worth mentioning that Jacod and Protter (2012) consider such a situation for preaveraging estimators in the synchronous and equidistant sampling setting (see Chapter 16 of that book for details).

Third, they (only) allow the variance of the noise ζ_t^l to depend on the time t, so the noise ζ_t^l is a special case of the locally stationary processes of Dahlhaus (1997) in their model. In our model the distribution of ζ_t^l can depend on t, so it is not encompassed with the locally stationary processes. In fact, our model even allows the noise ζ_t^l to depend on the whole past of X up to the time t (as long as $E[\zeta_t^l|X^l] = 0$). Especially, this allows us to incorporate noise involving some rounding into the model; see Example 2 of Jacod et al. (2009) or Model 2 simulated in Section 5 for such ones.

Fourth, they only allow (possibly nonstandard) Wiener processes as the process X^l. This induces a limitation of the range of the correlation between the latent returns and the noise (it must be in the interval $[-1/\sqrt{2}, 1/\sqrt{2}]$). Our model does not induce such a restriction because we can take e.g., $\underline{X}^l = \psi X^l$ for any constant ψ. Note that the noise ζ_t^l cannot produce the correlation between the latent returns and the microstructure noise, although the noise ζ_t^l can depend on X^l; this is due to the restriction of $E[\zeta_t^l|X] = 0$. For this reason we sometimes refer to $\sqrt{n}\left(\underline{X}^1_{S^i} - \underline{X}^1_{S^{i-1}}\right)$ and $\sqrt{n}\left(\underline{X}^2_{T^j} - \underline{X}^2_{T^{j-1}}\right)$ as the endogenous noise, although ζ_t^l could be endogenous. The importance of modeling this type of endogeneity is discussed by several authors; see e.g. Hansen and Lunde (2006) and Diebold (2006).

Remark 2.4. The reader might be afraid that the variance of the noise $\sqrt{n}\left(\underline{X}^1_{S^i} - \underline{X}^1_{S^{i-1}}\right)$ would blow up because we only assume that (5) is satisfied for some $\xi' < 1$. However, this is not problematic due to the following reason. First, we need to note that $n^{-\xi'}$ is an upper bound of the maximum duration $r_n(t)$. In fact, since we will assume that the order of the number of observations is at most n (see the condition [C1] below) and there is a restriction of $\sum_{S^i \leq t}(S^i - S^{i-1}) \leq t$, there are few durations attaining the upper bound $n^{-\xi'}$ and most durations would have the same magnitude as $1/n$. See also Remark 2.6 for further discussion.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Our aim is to estimate the integrated covariance $[X^1, X^2]_t$ of X^1 and X^2 at any time $t \in \mathbb{R}_+$ from the observation data $\left(\mathsf{X}^1_{S^i}\right)_{i:S^i \le t}$ and $\left(\mathsf{X}^2_{T^j}\right)_{j:T^j \le t}$. It is necessary to deal with both observation noise and nonsynchronicity of the observation times simultaneously. As is mentioned in Section 1, we use the preaveraging technique to remove the noise, while use the Hayashi–Yoshida method to deal with the nonsynchronicity. For the preaveraging technique we introduce some notation. We choose a sequence k_n of integers and a number $\theta \in (0, \infty)$ satisfying

$$k_n = \theta\sqrt{n} + o(n^{1/4}) \tag{7}$$

(for example $k_n = \lceil \theta\sqrt{n} \rceil$). 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 $\psi_{HY} := \int_0^1 g(x)\mathrm{d}x \neq 0$ (for example $g(x) = x \wedge (1-x)$). We associate the random intervals $I^i = [S^{i-1}, S^i)$ and $J^j = [T^{j-1}, T^j)$ with the sampling scheme (S^i) and (T^j) and refer to $\mathcal{I} = (I^i)_{i\in\mathbb{N}}$ and $\mathcal{J} = (J^j)_{j\in\mathbb{N}}$ as the sampling designs for X^1 and X^2. We introduce the *preaveraged observation data* of X^1 and X^2 based on the sampling designs $\mathcal{I}$ and $\mathcal{J}$ respectively as follows:

$$\overline{\mathsf{X}}^1(\mathcal{I})^i = \sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right)\left(\mathsf{X}^1_{S^{i+p}} - \mathsf{X}^1_{S^{i+p-1}}\right),$$

$$\overline{\mathsf{X}}^2(\mathcal{J})^j = \sum_{q=1}^{k_n-1} g\left(\frac{q}{k_n}\right)\left(\mathsf{X}^2_{T^{j+q}} - \mathsf{X}^2_{T^{j+q-1}}\right), \quad i,j = 0,1,\dots.$$

The following quantity was introduced in Christensen et al. (2010) :

DEFINITION 2.1 (Preaveraged Hayashi–Yoshida estimator). *The preaveraged Hayashi–Yoshida estimator, or preaveraged HY estimator of* X^1 *and* X^2 *associated with sampling designs* $\mathcal{I}$ *and* $\mathcal{J}$ *is the process*

$$PHY(\mathsf{X}^1, \mathsf{X}^2; \mathcal{I}, \mathcal{J})^n_t = \frac{1}{(\psi_{HY} k_n)^2} \sum_{\substack{i,j=0 \\ S^{i+k_n} \vee T^{j+k_n} \le t}}^{\infty} \overline{\mathsf{X}}^1(\mathcal{I})^i \overline{\mathsf{X}}^2(\mathcal{J})^j 1_{\{[S^i, S^{i+k_n}) \cap [T^j, T^{j+k_n}) \neq \emptyset\}}, \qquad t \in \mathbb{R}_+.$$

Remark 2.5. It is worth mentioning that even in the univariate case the above estimator is different from the preaveraging version of the realized volatility proposed in Jacod et al. (2009). In particular, we need no bias-correction term which is related to noise. To see why this difference occurs, we focus on the univariate case. Therefore, the above estimator can be rewritten as

$$PHY(\mathsf{X}^1, \mathsf{X}^1; \mathcal{I}, \mathcal{I})^n_t = \frac{1}{(\psi_{HY} k_n)^2} \sum_{\substack{i,j:i,j=0 \\ |i-j|<k_n}}^{m-k_n} \overline{\mathsf{X}}^1(\mathcal{I})^i \overline{\mathsf{X}}^1(\mathcal{I})^j,$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

where $m = \max\{i : S^i \le t\}$. Since it holds that

$$\sum_{j=i-k_n+1}^{i+k_n-1} \overline{U}^1(\mathcal{I})^j = \sum_{j=i-k_n+1}^{i+k_n-1} \sum_{p=0}^{k_n-1} \left\{ g\left(\frac{p}{k_n}\right) - g\left(\frac{p+1}{k_n}\right) \right\} U^1_{S^{j+p}}$$
$$= \sum_{p=i-k_n+1}^{i+2k_n-1} \sum_{j=p-k_n+1}^{p} \left\{ g\left(\frac{p-j}{k_n}\right) - g\left(\frac{p-j+1}{k_n}\right) \right\} U^1_{S^p}$$
$$= \sum_{p=i-k_n+1}^{i+2k_n-1} \{g(0) - g(1)\} U^1_{S^p} = 0$$

if $k_n - 1 \le i \le m - 2k_n + 1$, the noise has no contribution to the estimator in asymptotics. See also Remark 3.2 of Christensen, Podolskij, and Vetter (2011b).

Remark 2.6. Regarding the issue discussed in Remark 2.4, the things become more clear when we focus on the preaveraged data. For simplicity we assume that both X^1 and $\underline{X}^1$ are standard Brownian motions. Then we have

$$E\left[\left|\overline{X}^1(\mathcal{I})^i\right|^2\right] = E\left[\left|\sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right)\left(X^1_{S^{i+p}} - X^1_{S^{i+p-1}}\right)\right|^2\right]$$
$$= E\left[\sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right)^2 \left(S^{i+p} - S^{i+p-1}\right)\right]$$

and

$$E\left[\left|\overline{U}^1(\mathcal{I})^i\right|^2\right] = E\left[\left|\sum_{p=0}^{k_n-1} \Delta(g)^n_p U^1_{S^{i+p}}\right|^2\right]$$
$$= nE\left[\sum_{p=0}^{k_n-1} \left|\Delta(g)^n_p\right|^2 \left(S^{i+p} - S^{i+p-1}\right)\right] + E\left[\sum_{p=0}^{k_n-1} \left|\Delta(g)^n_p\right|^2 \left(\zeta^1_{S^p}\right)^2\right],$$

where $\Delta(g)^n_p = g((p+1)/k_n) - g(p/k_n)$. Since $\Delta(g)^n_p = O(1/\sqrt{n})$ due to (7), $\overline{U}^1(\mathcal{I})^i$ has the same magnitude as that of $\overline{X}^1(\mathcal{I})^i$.

The main subject of the present article is to develop fully an asymptotic theory for the method in a general setting, which involves the asymptotic mixed normality of the estimator. The above estimator is, however, not proper to this purpose. In fact, the associated central limit theorem for the above estimator has been shown in Christensen et al. (2011b), but it is restricted to the case when observation times are deterministic (or random but independent of the observed processes) and some important sampling schemes in practice, like the Poisson sampling schemes, are excluded. Moreover, the asymptotic variance given in their theorem has a quite complex form which depends on the special forms of the sampling times considered in that paper, so that it seems to be impossible that one extends their result to more general sampling schemes involving the Poisson sampling schemes. For this reason, we modify the above estimator as follows. The following notion was introduced to this area in Barndorff-Nielsen et al. (2011):

DEFINITION 2.2 (Refresh time). *The first refresh time of sampling designs $\mathcal{I}$ and $\mathcal{J}$ is defined as $R^0 = S^0 \vee T^0$, and then subsequent refresh times as*

$$R^k := \min\left\{S^i | S^i > R^{k-1}\right\} \vee \min\left\{T^j | T^j > R^{k-1}\right\}, \qquad k = 1, 2, \ldots.$$

We introduce new sampling schemes by a kind of the next-tick interpolations to the refresh times. That is, we define $\widehat{S}^0 := S^0$, $\widehat{T}^0 := T^0$, and

$$\widehat{S}^k := \min\{S^i | S^i > R^{k-1}\}, \quad \widehat{T}^k := \min\{T^j | T^j > R^{k-1}\}, \qquad k = 1, 2, \ldots.$$

Then, we create new sampling designs as follows:

$$\widehat{I}^k := [\widehat{S}^{k-1}, \widehat{S}^k), \qquad \widehat{J}^k := [\widehat{T}^{k-1}, \widehat{T}^k), \qquad \widehat{\mathcal{I}} := (\widehat{I}^i)_{i \in \mathbb{N}}, \qquad \widehat{\mathcal{J}} := (\widehat{J}^j)_{j \in \mathbb{N}}.$$

For the sampling designs $\widehat{\mathcal{I}}$ and $\widehat{\mathcal{J}}$ obtained in such a manner, we will consider the preaveraged HY estimator $\widehat{PHY}\left(\mathsf{X}^1, \mathsf{X}^2\right)^n := PHY\left(\mathsf{X}^1, \mathsf{X}^2; \widehat{\mathcal{I}}, \widehat{\mathcal{J}}\right)^n$.

Now we review the results related to the consistency and the asymptotic mixed normality of the estimator $\widehat{PHY}\left(\mathsf{X}^1, \mathsf{X}^2\right)^n$. We write the canonical decompositions of $\underline{X}^1$ and $\underline{X}^2$ as follows:

$$\underline{X}^l = \underline{A}^l + \underline{M}^l, \qquad l = 1, 2. \tag{8}$$

Here, $\underline{A}^1$ and $\underline{A}^2$ are continuous $\mathbf{F}^{(0)}$-adapted processes with locally finite variations, while $\underline{M}^1$ and $\underline{M}^2$ are continuous $\mathbf{F}^{(0)}$-local martingales. Next, let $N_t^n = \sum_{k=1}^{\infty} 1_{\{R^k \leq t\}}$ for each $t \in \mathbb{R}_+$, and we introduce the following regularity conditions:

[C1] $n^{-1} N_t^n = O_p(1)$ as $n \to \infty$ for every t.

[C2] $A^1, A^2, \underline{A}^1, \underline{A}^2$, and $[V, W]$ for $V, W = X^1, X^2, \underline{X}^1, \underline{X}^2$ are absolutely continuous with locally bounded derivatives.

Furthermore, for every $r \in [2, \infty)$ we introduce the following regularity condition for the noise process:

[$\mathrm{N}_r^\flat$] $\left(\int |z|^r Q_t(\mathrm{d}z)\right)_{t \in \mathbb{R}_+}$ is a locally bounded process.

A sequence (X^n) of stochastic processes is said to converge to a process X *uniformly on compacts in probability* (abbreviated *ucp*) if, for each $t > 0$, $\sup_{0 \leq s \leq t} |X_s^n - X_s| \to^p 0$ as $n \to \infty$. We then write $X^n \xrightarrow{ucp} X$. We have the following result about the consistency of the preaveraged HY estimator (shown in Appendix A.1):

THEOREM 2.1. *Suppose (5),* [C1]-[C2] *and* $[\mathrm{N}_2^\flat]$ *are satisfied. Then*

$$\widehat{PHY}(\mathsf{X}^1,\mathsf{X}^2)^n \xrightarrow{ucp} [X^1,X^2]$$

as $n\to\infty$*, provided that* $\xi'>1/2$.

The consistency of the preaveraged HY estimator was first shown in Christensen et al. (2010) in a simpler situation.

Next we review the results related to the asymptotic mixed normality of the preaveraged HY estimator. As noted in the above, this result was first proven in Christensen et al. (2011b) for the original one from Definition 2.1 when the sampling times are deterministic transformation of equidistant ones.

Let $N_t^{n,1}=\sum_{k=1}^\infty 1_{\{\widehat{S}^k\le t\}}$ and $N_t^{n,2}=\sum_{k=1}^\infty 1_{\{\widehat{T}^k\le t\}}$ for each $t\in\mathbb{R}_+$ and

$$\Gamma^k=[R^{k-1},R^k),\qquad \check{I}^k:=[\check{S}^k,\widehat{S}^k),\qquad \check{J}^k:=[\check{T}^k,\widehat{T}^k)$$

for each $k\in\mathbb{N}$. Here, for each $t\in\mathbb{R}_+$ we write $\check{S}^k=\sup_{S^i<\widehat{S}^k}S^i$ and $\check{T}^k=\sup_{T^j<\widehat{T}^k}T^j$. Note that $\check{S}^k$ and $\check{T}^k$ may not be stopping times.

Let $\mathbf{H}^n=(\mathcal{H}_t^n)_{t\in\mathbb{R}_+}$ be a sequence of filtrations of $\mathcal{F}$ to which N^n, $N^{n,1}$, and $N^{n,2}$ are adapted, and for each n and each $\rho\ge 0$ we define the processes χ^n, $G(\rho)^n$, $F(\rho)^{n,1}$, $F(\rho)^{n,2}$, and $F(1)^{n,1*2}$ by

$$\chi_s^n=P(\widehat{S}^k=\widehat{T}^k\,|\,\mathcal{H}^n_{R^{k-1}}),\qquad G(\rho)_s^n=E\left[\left(n|\Gamma^k|\right)^\rho\,|\mathcal{H}^n_{R^{k-1}}\right],$$

$$F(\rho)_s^{n,1}=E\left[\left(n|\check{I}^k|\right)^\rho\,|\mathcal{H}^n_{\widehat{S}^{k-1}}\right],\qquad F(\rho)_s^{n,2}=E\left[\left(n|\check{J}^k|\right)^\rho\,|\mathcal{H}^n_{\widehat{T}^{k-1}}\right],$$

$$F(1)_s^{n,1*2}=nE\left[|\check{I}^k\cap\check{J}^k|+|\check{I}^{k+1}\cap\check{J}^k|+|\check{I}^k\cap\check{J}^{k+1}|\,\middle|\,\mathcal{H}^n_{R^{k-1}}\right]$$

when $s\in\Gamma^k$.

In addition to the crucial measurability condition [A1], we need a number of regularity conditions on the model, which are more or less commonly used in the literature. We detailed such conditions used in this paper in the following.

The following condition is necessary to compute the asymptotic variance of the estimation error of our estimator explicitly. In fact, this condition ensures that quantities appearing in the asymptotic variance indeed converge. For a sequence (X^n) of càdlàg processes and a càdlàg process X, we write $X^n\xrightarrow{\text{Sk.p.}}X$ if (X^n) converges to X in probability for the Skorokhod topology.

[A2] (i) For each n, we have a càdlàg $\mathbf{H}^n$-adapted process G^n and a random subset $\mathcal{N}_n^0$ of $\mathbb{N}$ such that $(\#\mathcal{N}_n^0)_{n\in\mathbb{N}}$ is tight, $G(1)^n_{R^{k-1}}=G^n_{R^{k-1}}$ for any $k\in\mathbb{N}-\mathcal{N}_n^0$, and there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process G satisfying that G and G_- do not vanish and that $G^n\xrightarrow{\text{Sk.p.}}G$ as $n\to\infty$.

(ii) There exists a constant $\rho\ge 1/\xi'$ such that $\left(\sup_{0\le s\le t}G(\rho)_s^n\right)_{n\in\mathbb{N}}$ is tight for all $t>0$.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

(iii) For each n, we have a càdlàg $\mathbf{H}^n$-adapted process χ'^n and a random subset $\mathcal{N}'_n$ of $\mathbb{N}$ such that $(\#\mathcal{N}'_n)_{n\in\mathbb{N}}$ is tight, $\chi^n_{R^{k-1}} = \chi'^n_{R^{k-1}}$ for any $k \in \mathbb{N} - \mathcal{N}'_n$, and there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process χ such that $\chi'^n \xrightarrow{\text{Sk.p.}} \chi$ as $n \to \infty$.

(iv) For each n and $l = 1, 2, 1*2$, we have a càdlàg $\mathbf{H}^n$-adapted process $F^{n,l}$ and a random subset $\mathcal{N}^l_n$ of $\mathbb{N}$ such that $(\#\mathcal{N}^l_n)_{n\in\mathbb{N}}$ is tight, $F(1)^{n,l}_{R^{k-1}} = F^{n,l}_{R^{k-1}}$ for any $k \in \mathbb{N} - \mathcal{N}^l_n$, and there exists a càdlàg $\mathbf{F}^{(0)}$-adapted processes F^l satisfying $F^{n,l} \xrightarrow{\text{Sk.p.}} F^l$ as $n \to \infty$.

(v) There exists a constant $\rho' \geq 1/\xi'$ such that $\left(\sup_{0\leq s\leq t} F(\rho')^{n,l}_s\right)_{n\in\mathbb{N}}$ is tight for all $t > 0$ and $l = 1, 2$.

The following condition is a sufficient one for the condition [A2]:

[A2$^\sharp$] (i) For every $\rho \in [0, 1/\xi']$ there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process $G(\rho)$ such that $G(\rho)^n \xrightarrow{\text{Sk.p.}} G(\rho)$ as $n \to \infty$. Furthermore, G and G_- do not vanish, where $G = G(1)$.

(ii) There exists a càdlàg $\mathbf{F}^{(0)}$-adapted process χ such that $\chi^n \xrightarrow{\text{Sk.p.}} \chi$ as $n \to \infty$.

(iii) For every $l = 1, 2$ and every $\rho' \in [0, 1/\xi']$, there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process $F(\rho)^l$ such that $F(\rho)^{n,l} \xrightarrow{\text{Sk.p.}} F(\rho)^l$ as $n \to \infty$.

(iv) There exists a càdlàg $\mathbf{F}^{(0)}$-adapted process $F(1)^{1*2}$ such that $F(1)^{n,1*2} \xrightarrow{\text{Sk.p.}} F(1)^{1*2}$ as $n \to \infty$.

Remark 2.7.

(i) Under [A2](i)–(ii) it can be shown that

$$\frac{1}{n} N^n_t \to^p \int_0^t \frac{1}{G_s} \mathrm{d}s \tag{9}$$

as $n \to \infty$ for any $t \in \mathbb{R}_+$ in a similar manner to the proof of Lemma 2.2 of Hayashi, Jacod, and Yoshida (2011). In particular, [A2] implies that the parameter n corresponds to the magnitude of the number N^n_t of the (synchronized) observation times.

(ii) An [A2$^\sharp$] type condition appears in Barndorff-Nielsen et al. (2011) and Hayashi et al. (2011), for example. The reason why we introduce a kind of exceptional sets $\mathcal{N}^l_n$ $(l = 0, 1, 2, 1*2, ')$ is that the condition [A2] without them is too local. To explain this, we focus on the univariate case. Note that in this case we have $R^k = S^k$ $(k = 0, 1, 2, \ldots)$. Let τ be a positive number and suppose that (S^i) be a sequence of Poisson arrival times whose intensity is $\underline{\lambda}$ before the time τ and $\overline{\lambda}$ after τ. Then the structure of the process $G(1)^n$ becomes very complex around the time τ (of course if

$\underline{\lambda} \neq \overline{\lambda}$), so that it will be difficult to verify the convergence $G(1)^n \xrightarrow{\text{Sk.p.}} G$ because it requires a kind of uniformity.

Since we consider sampling times which are possibly nonequidistant and nonsynchronous, it is necessary to impose a kind of continuity condition on the density processes of A^1, $[X^1]$, etc., as Hayashi and Yoshida (2011) did. Accordingly, we introduce the following conditions, which are analogs to the conditions [A3] and [A4] in Hayashi and Yoshida (2011):

[A3] For each $V, W = X^1, X^2, \underline{X}^1, \underline{X}^2$, $[V, W]$ is absolutely continuous with a càdlàg derivative, and for the density process $f = [V, W]'$ there is a sequence (σ_k) of $\mathbf{F}^{(0)}$-stopping times such that $\sigma_k \uparrow \infty$ as $k \to \infty$ and for every k and any $\lambda > 0$, we have a positive constant $C_{k,\lambda}$ satisfying

$$E\left[|f^{\sigma_k}_{\tau_1} - f^{\sigma_k}_{\tau_2}|^2 \middle| \mathcal{F}_{\tau_1 \wedge \tau_2}\right] \leq C_{k,\lambda} E\left[|\tau_1 - \tau_2|^{1-\lambda} \middle| \mathcal{F}_{\tau_1 \wedge \tau_2}\right] \tag{10}$$

for any bounded $\mathbf{F}^{(0)}$-stopping times τ_1 and τ_2, and f is adapted to $\mathbf{H}^n$.

[A4] $\xi \vee \frac{9}{10} < \xi'$ and (5) holds for every $t \in \mathbb{R}_+$.

Due to the same reason as stated in the above, conditions analogous to [A5] and [A6] in Hayashi and Yoshida (2011) are necessary to deal with the drift parts. For a (random) interval I and a time t, we write $I(t) = I \cap [0, t)$.

[A5] A^1, A^2, $\underline{A}^1$, and $\underline{A}^2$ are absolutely continuous with càdlàg derivatives, and there is a sequence (σ_k) of $\mathbf{F}^{(0)}$-stopping times such that $\sigma_k \uparrow \infty$ as $k \to \infty$ and for every k we have a positive constant $C_k 0$ and $\lambda_k \in (0, 3/4)$ satisfying

$$E\left[|f^{\sigma_k}_t - f^{\sigma_k}_\tau|^2 \middle| \mathcal{F}_{\tau \wedge t}\right] \leq C_k E\left[|t - \tau|^{1-\lambda_k} \middle| \mathcal{F}_{\tau \wedge t}\right] \tag{11}$$

for every $t > 0$ and any bounded $\mathbf{F}^{(0)}$-stopping time τ, for the density processes $f = (A^1)'$, $(A^2)'$, $(\underline{A}^1)'$, and $(\underline{A}^2)'$.

[A6] For each $t \in \mathbb{R}_+$, $nH_n(t) = O_p(1)$ as $n \to \infty$, where $H_n(t) = \sum_{k=1}^\infty |\Gamma^k(t)|^2$.

Let $r \in [2, \infty)$. The following condition is a regularity condition for the noise process:

[N_r] $\left(\int |z|^r Q_t(\mathrm{d}z)\right)_{t \in \mathbb{R}_+}$ is a locally bounded process, and the covariance matrix process

$$\Psi_t(\omega^{(0)}) = \int zz^* Q_t(\omega^{(0)}, \mathrm{d}z). \tag{12}$$

is càdlàg, quasi-left continuous and adapted to $\mathbf{H}^n$ for every n. Here, an asterisk denotes the transpose of a matrix. Furthermore, there is a sequence

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

(σ^k) of $\mathbf{F}^{(0)}$-stopping times such that $\sigma^k \uparrow \infty$ as $k \to \infty$ and for every k and any $\lambda > 0$, we have a positive constant $C_{k,\lambda}$ satisfying

$$E\left[|\Psi^{ij}_{\sigma^k \wedge t} - \Psi^{ij}_{\sigma^k \wedge (t-h)_+}|^2 \middle| \mathcal{F}_{(t-h)_+}\right] \leq C_{k,\lambda} h^{1-\lambda} \tag{13}$$

for any $i, j \in \{1, 2\}$ and any $t, h > 0$.

$[\mathrm{N}_r]$ is obviously satisfied if the noise is i.i.d. and independent of the latent processes with the moment of the r-th order i.e., $Q_t = Q$ for all t, where Q is a probability distribution on $\mathbb{R}^2$ satisfying $\int z Q(\mathrm{d}z) = 0$ and $\int |z|^r Q(\mathrm{d}z) < \infty$. We need to assume that $[\mathrm{N}_8]$ holds true for verifying a Lindeberg-type condition.

Remark 2.8. Inequalities (10), (11), and (13) are satisfied when $w(f; h, t) = O_p\left(h^{\frac{1}{2}-\lambda}\right)$ as $h \to \infty$ for every $t \in (0, \infty)$ and for every $\lambda \in (0, \infty)$, for example. Here, for a real-valued function x on $\mathbb{R}_+$, the *modulus of continuity* on $[0, T]$ is denoted by $w(x; \delta, T) = \sup\{|x(t) - x(s)|; s, t \in [0, T], |s - t| \leq \delta\}$ for $T, \delta > 0$. This is the original condition in Hayashi and Yoshida (2011). Another example where (10), (11), and (13) are satisfied is the case that there exists an $\mathbf{F}^{(0)}$-adapted process B with a locally integrable variation and a locally square-integrable martingale L such that $f = B + L$ and both of the predictable compensator of the variation process of B and the predictable quadratic variation of L are absolutely continuous with locally bounded derivatives. This type of condition is familiar in the context of the estimation of volatility-type quantities; see Hayashi et al. (2011) and Jacod, Podolskij, and Vetter (2010) for example. Furthermore, in both cases f is càdlàg and quasi-left continuous.

Remark 2.9. Conditions [A1], [A2], [A4], and [A6] are satisfied by sampling schemes that arise as realizations of two homogeneous Poisson processes which are mutually independent and independent of the processes X^1, X^2, $\underline{X}^1$ and $\underline{X}^2$. In fact, suppose that (S^i) and (T^j) are the arrival times of two mutually independent Poisson processes with intensities np^1 and np^2 $(p^1, p^2 > 0)$ which are independent of X^1, X^2, $\underline{X}^1$, and $\underline{X}^2$, respectively. Then, [A1] is satisfied due to the reason stated in Remark 2.2, while [A4] and [A6] can easily be verified by using the fact that the durations are exponentially distributed. Finally, a simple computation yields [A2] holds true with

$$G_s = \frac{1}{p^1} + \frac{1}{p^2} - \frac{1}{p^1 + p^2}, \quad \chi_s = 0, \quad F^1_s = \frac{1}{p^1}, \quad F^2_s = \frac{1}{p^2}, \quad F^{1*2}_s = \frac{2}{p^1 + p^2}.$$

This type of sampling scheme is one of the most popular model for observation times in the literature; see e.g. Bibinger (2012), Hayashi and Yoshida (2005), and Zhang (2011).

We extend the functions g and g' to the whole real line by setting $g(x) = g'(x) = 0$ for $x \notin [0, 1]$. Then we put

$$\kappa := \int_{-2}^{2} \psi_{g,g}(x)^2 \mathrm{d}x, \qquad \widetilde{\kappa} := \int_{-2}^{2} \psi_{g',g'}(x)^2 \mathrm{d}x, \qquad \overline{\kappa} := \int_{-2}^{2} \psi_{g,g'}(x)^2 \mathrm{d}x.$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Here, for each $f_1, f_2 \in \{g, g'\}$ we define the function ψ_{f_1,f_2} on $\mathbb{R}$ by $\psi_{f_1,f_2}(x) = \int_0^1 \int_{x+u-1}^{x+u+1} f_1(u) f_2(v)\mathrm{d}v\mathrm{d}u$.

We denote by $\mathbb{D}(\mathbb{R}_+)$ the space of càdlàg functions on $\mathbb{R}_+$ equipped with the Skorokhod topology. A sequence of random elements X^n defined on a probability space $(\Omega, \mathcal{F}, P)$ is said to *converge stably in law* to a random element X defined on an appropriate extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ of $(\Omega, \mathcal{F}, P)$ if $E[Yg(X^n)] \to E[Yg(X)]$ for any $\mathcal{F}$-measurable and bounded random variable Y and any bounded and continuous function g. We then write $X^n \to^{d_s} X$.

Now we are ready to state the result related to the asymptotic mixed normality of the preaveraged HY estimator.

THEOREM 2.2.

(a) *Suppose* [A1], [A2](i)–(iii), [A3]-[A6] *and* [N$_8$] *are satisfied. Suppose also* $\underline{X}^1 = \underline{X}^2 = 0$. *Then*

$$n^{1/4}\{\widehat{PHY}(\mathsf{X}^1,\mathsf{X}^2)^n - [X^1, X^2]\} \to^{d_s} \int_0^\cdot w_s \mathrm{d}\widetilde{W}_s \qquad \text{in } \mathbb{D}(\mathbb{R}_+) \tag{14}$$

as $n \to \infty$, *where* $\tilde{W}$ *is a one-dimensional standard Wiener process (defined on an extension of* $\mathcal{B}$*) independent of* $\mathcal{F}$ *and* w *is given by*

$$\begin{aligned} w_s^2 = \psi_{HY}^{-4}\Big[&\theta\kappa\left\{[X^1]_s'[X^2]_s' + ([X^1,X^2]_s')^2\right\}G_s \\ &+\theta^{-3}\widetilde{\kappa}\left\{\Psi_s^{11}\Psi_s^{22} + \left(\Psi_s^{12}\chi_s\right)^2\right\}G_s^{-1} \\ &+\theta^{-1}\overline{\kappa}\left\{[X^1]_s'\Psi_s^{22} + [X^2]_s'\Psi_s^{11} + 2[X^1,X^2]_s'\Psi_s^{12}\chi_s\right\}\Big]. \end{aligned} \tag{15}$$

(b) *Suppose* [A1]–[A6] *and* [N$_8$] *are satisfied. Then (14) holds true as* $n \to \infty$ *with* $\tilde{W}$ *as above and* w *is given by*

$$\begin{aligned} w_s^2 = \psi_{HY}^{-4}\Big[&\theta\kappa\left\{[X^1]_s'[X^2]_s' + ([X^1,X^2]_s')^2\right\}G_s \\ &+\theta^{-3}\widetilde{\kappa}\left\{\overline{\Psi}_s^{11}\overline{\Psi}_s^{22} + \left(\overline{\Psi}_s^{12}\right)^2\right\}G_s^{-1} \\ &+\theta^{-1}\overline{\kappa}\Big\{[X^1]_s'\overline{\Psi}_s^{22} + [X^2]_s'\overline{\Psi}_s^{11} + 2[X^1,X^2]_s'\overline{\Psi}_s^{12} \\ &-\left([\underline{X}^1,X^2]_s'F_s^1 - [X^1,\underline{X}^2]_s'F_s^2\right)^2 G_s^{-1}\Big\}\Big], \end{aligned} \tag{16}$$

where $\overline{\Psi}_s^{ll} = \Psi_s^{ll} + [\underline{X}^l]_s' F_s^l$ $(l = 1, 2)$ *and* $\overline{\Psi}_s^{12} = \Psi_s^{12}\chi_s + [\underline{X}^1, \underline{X}^2]_s' F_s^{1*2}$.

The proof is given in Appendix A.2.

Remark 2.10. Our situation contains the following specification as a special case: X^1 and X^2 are two correlated standard Wiener processes with constant correlation ρ, $\left(U_{S^i}^1\right)$, and $\left(U_{T^i}^2\right)$ are mutually independent i.i.d. centered Gaussian

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

random variables independent of X^1 and X^2, and $S^i = T^i = i/n$. In this case it is known that an upper bound of the convergence rate to estimate the parameter ρ from the observation data $\left(\mathsf{X}^1_{S^i}\right)_{i:S^i\leq 1}$ and $\left(\mathsf{X}^2_{T^j}\right)_{i:T^j\leq 1}$ is given by $n^{-1/4}$ (see Proposition 6 of Bibinger, 2011), so the estimator $\widehat{PHY}(\mathsf{X}^1,\mathsf{X}^2)^n$ achieves the optimal rate of convergence.

3. MAIN RESULTS

In this section we investigate the case that the latent processes possibly have jumps. Let $Z^1 = \left(Z^1_t\right)_{t\in\mathbb{R}_+}$ and $Z^2 = \left(Z^2_t\right)_{t\in\mathbb{R}_+}$ be two stochastic processes on $\left(\Omega^{(0)}, \mathcal{F}^{(0)}, P^{(0)}\right)$. From now on Z^1 and Z^2 will correspond the latent log-price processes, while X^1 and X^2 will correspond the continuous components of them. We have the observation data $\mathsf{Z}^1 = \left(\mathsf{Z}^1_{S^i}\right)_{i\in\mathbb{Z}_+}$ and $\mathsf{Z}^2 = \left(\mathsf{Z}^2_{T^j}\right)_{j\in\mathbb{Z}_+}$ of Z^1 and Z^2 contaminated by noise:

$$\mathsf{Z}^1_{S^i} = Z^1_{S^i} + U^1_{S^i}, \qquad \mathsf{Z}^2_{T^j} = Z^2_{T^j} + U^2_{T^j}.$$

Here, the observation noise $\left(U^1_{S^i}\right)_{i\in\mathbb{Z}_+}$ and $\left(U^2_{T^j}\right)_{j\in\mathbb{Z}_+}$ are given by (6).

The idea for the construction of our estimator is as follows. The preaveraging procedure smooths the noise and thus we can expect the preaveraged data $\bar{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i$ and $\bar{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j$ are small enough if they contain no jumps. This idea has already appeared in Aït-Sahalia, Jacod, and Li (2012) and Podolskij and Ziggel (2010) in the univariate case and Jing, Li, and Liu (2011) in the synchronous case. Following this idea, we introduce the following quantity:

DEFINITION 3.1 (Preaveraged truncated Hayashi–Yoshida estimator). *The preaveraged truncated Hayashi–Yoshida estimator, or PTHY estimator of two observation data Z^1 and Z^2 is the process*

$$\widehat{PTHY}(\mathsf{Z}^1,\mathsf{Z}^2)^n_t = \frac{1}{(\psi_{HY}k_n)^2} \times \sum_{i,j:\widehat{S}^{i+k_n}\vee\widehat{T}^{j+k_n}\leq t} \overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j\bar{K}^{ij}1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho^1_n(\widehat{S}^i),|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\leq\varrho^2_n(\widehat{T}^j)\}}, \quad \in\mathbb{R}_+,$$

where

$$\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i = \sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right)\left(\mathsf{Z}^1_{\widehat{S}^{i+p}} - \mathsf{Z}^1_{\widehat{S}^{i+p-1}}\right),$$

$$\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j = \sum_{q=1}^{k_n-1} g\left(\frac{q}{k_n}\right)\left(\mathsf{Z}^2_{\widehat{T}^{j+q}} - \mathsf{Z}^2_{\widehat{T}^{j+q-1}}\right), \quad i,j=0,1,\ldots,$$

$\bar{K}^{ij} = 1_{\{[\widehat{S}^i,\widehat{S}^{i+k_n})\cap[\widehat{T}^j,\widehat{T}^{j+k_n})\neq\emptyset\}}$ *and* $(\varrho^l_n(t))_{n\in\mathbb{N}}$, $l=1,2$, *are two sequences of positive-valued stochastic processes.*

Remark 3.1. Wang et al. (2013) proposed a truncated version of $PHY(\mathsf{X}^1, \mathsf{X}^2; \mathcal{I}, \mathcal{J})^n_t$:

$$\frac{1}{(\psi_{HY} k_n)^2} \times \sum_{i,j: S^{i+k_n} \vee T^{j+k_n} \le t} \overline{\mathsf{Z}}^1(\mathcal{I})^i \overline{\mathsf{Z}}^2(\mathcal{J})^j 1_{\left\{[S^i, S^{i+k_n}) \cap [T^j, T^{j+k_n}) \neq \emptyset, |\overline{\mathsf{Z}}^1(\mathcal{I})^i|^2 \le \varrho^1_n(S^i), |\overline{\mathsf{Z}}^2(\mathcal{J})^j|^2 \le \varrho^2_n(T^j)\right\}}. \tag{17}$$

They proved the consistency of this estimator under the conditions $[\mathrm{K}_2]$(i)–(iv) (see Section 3.2 about this condition) when $\left(U^1_{S^i}\right)$ and $\left(U^2_{T^j}\right)$ are centered i.i.d. random variables independent of Z^1 and Z^2. As noted in the previous section, in order to derive the asymptotic mixed normality in a general sampling setting, we use the additional synchronization procedure for constructing the estimator in addition to the approach of Wang et al. (2013). Moreover, there is another advantage in using the synchronization procedure for the construction: Suppose that (S^i) is sampled more frequently than (T^j), e.g. $S^i = i/n$ and $T^j = jm/n$ for some large integer m. Then, even if $[S^i, S^{i+k_n}) \cap [T^j, T^{j+k_n}) \neq \emptyset$ the nonoverlap part between $[S^i, S^{i+k_n})$ and $\cap[T^j, T^{j+k_n})$ is large. For example, in the case that $S^i = i/n$ and $T^j = jm/n$, $[S^{jm}, S^{jm+k_n}) \cap [T^j, T^{j+k_n}) \neq \emptyset$ but the nonoverlap part $[(jm+k_n)/n, (jm+k_n)/n + (m-1)k_n/n)$ is large if m is large. As a consequence, the product $\overline{\mathsf{Z}}^1(\mathcal{I})^i \overline{\mathsf{Z}}^2(\mathcal{J})^j 1_{\{[S^i, S^{i+k_n}) \cap [T^j, T^{j+k_n})\}}$ involves many cross-products of the latent returns over nonoverlap intervals such as $\left(Z^1_{S^p} - Z^1_{S^{p-1}}\right)\left(Z^2_{T^q} - Z^2_{T^{p-1}}\right) 1_{\{[S^{p-1}, S^p) \cap [T^{q-1}, T^q) \neq \emptyset\}}$. This will boost the root mean squared error of the estimator given by (17). Our synchronization procedure enables us to avoid this problem. We will check this advantage by a Monte Carlo experiment in Section 5.

We will write $\varrho^1_n[i] := \varrho^1_n(\widehat{S}^i)$ and $\varrho^2_n[j] := \varrho^2_n(\widehat{T}^j)$ for short.

3.1. Finite Activity Jump Case

First we consider the case that the observed processes have at most finite jumps. We assume the following structural assumption:

[F] For each $l = 1, 2$ we have $Z^l_t = X^l_t + \sum_{k=1}^{N^l_t} \gamma^l_k$, where X^l is a continuous semimartingale on $\mathcal{B}^{(0)}$ given by (3), N^l is a (simple) point process adopted to $\mathbf{F}^{(0)}$, and $(\gamma^l_k)_{k \in \mathbb{N}}$ is a sequence of nonzero random variables.

Moreover, we impose the following condition on the threshold processes:

[T] $\xi' > 1/2$, and for each $l = 1, 2$ we have $\varrho^l_n(t) = \alpha^l_n(t)\rho_n$, where

(i) $(\rho_n)_{n \in \mathbb{N}}$ is a sequence of (deterministic) positive numbers satisfying $\rho_n \to 0$ and

$$\frac{n^{-\xi'+1/2} \log n}{\rho_n} \to 0 \tag{18}$$

as $n \to \infty$.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

(ii) $(\alpha_n^l(t))_{n\in\mathbb{N}}$ is a sequence of (not necessarily adapted) positive-valued stochastic processes. Moreover, there exists a sequence (R_k^l) of stopping times (with respect to $\mathbf{F}$) such that $R_k^l \uparrow \infty$ and both of the sequences $\left(\sup_{0\le t<R_k^l} \alpha_n^l(t)\right)_{n\in\mathbb{N}}$ and $\left(\sup_{0\le t<R_k^l}[1/\alpha_n^l(t)]\right)_{n\in\mathbb{N}}$ are tight for all k.

Remark 3.2. Condition [T](ii) implies that the processes $\alpha_n^1(t)$ and $\alpha_n^2(t)$ have no impact on the convergence rates of the threshold estimators, so that they are controlled by the sequence ρ_n. Condition [T](i) says that ρ_n converges to zero slower than $n^{-\xi'+1/2}\log n$, which dominates the modulus of continuity of $\overline{X}^1(\widehat{\mathcal{I}})^i$ and $\overline{X}^2(\widehat{\mathcal{J}})^j$ (see Lemma A.3.2). Therefore, [T](i) can be regarded as a counterpart of Assumption 3 from Mancini and Gobbi (2012). [T](ii) allows the threshold processes to depend on both time and observation data, which is important for applications because it is natural to select the threshold processes dependently on the data; see Section 5.1.

Now we obtain the following theorem.

THEOREM 3.1. *Suppose* [F], [T], [C1]–[C2] *and* $[\mathrm{N}_r^\flat]$ *hold for some* $r \in (2,\infty)$. *Then we have*

$$\sup_{0\le s\le t} |\widehat{PTHY}(\mathsf{Z}^1,\mathsf{Z}^2)_s^n - \widehat{PHY}(\mathsf{X}^1,\mathsf{X}^2)_s^n| = o_p(n^{-\frac{1}{4}}) + O_p\left(\left(n^{-\frac{1}{2}}\rho_n^{-1}\right)^{\frac{r-2}{2}}\right)$$

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

The proof of this theorem is given in Appendix A.3. Combining this result with Theorem 2.1 or Theorem 2.2, we obtain the following results:

THEOREM 3.2 (Consistency of the PTHY estimator in finite activity case). *Suppose* [F], [T], [C1]-[C2] *and* $[\mathrm{N}_r^\flat]$ *hold for some* $r\in(2,\infty)$. *Then we have*

$$\widehat{PTHY}(\mathsf{Z}^1,\mathsf{Z}^2)^n \xrightarrow{ucp} [X^1,X^2] \tag{19}$$

as $n\to\infty$.

THEOREM 3.3 (Asymptotic mixed normality of the PTHY estimator in finite activity case).

(a) *Suppose* [A1], [A2](i)–(iii), [A3]-[A6] *and* [F] *are satisfied. Suppose also* $\underline{X}^1=\underline{X}^2=0$, $[\mathrm{N}_r]$ *holds for some* $r\in[8,\infty)$ *and* [T] *holds with* $n^{(r-3)/\{2(r-2)\}}\rho_n\to\infty$ *as* $n\to\infty$. *Then*

$$n^{1/4}\left\{\widehat{PTHY}\left(\mathsf{Z}^1,\mathsf{Z}^2\right)^n - \left[X^1,X^2\right]\right\} \to^{d_s} \int_0^\cdot w_s \mathrm{d}\widetilde{W}_s \qquad \text{in } \mathbb{D}(\mathbb{R}_+) \tag{20}$$

as $n\to\infty$, *where* $\widetilde{W}$ *is the same one as in Theorem 2.2 and* w *is given by (15).*

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

(b) *Suppose* [A1]–[A6] *and* [F] *are satisfied. Suppose also* [N_r] *holds for some* $r \in [8,\infty)$ *and* [T] *holds with* $n^{(r-3)/\{2(r-2)\}}\rho_n \to \infty$ *as* $n \to \infty$. *Then (20) holds with that* $\widetilde{W}$ *is as in the above and* w *is given by (16).*

3.2. Infinite Activity Jump Case

Next we consider the case that the observed processes are two general semimartingales contaminated by noise. We need the following structural assumption. Let $\beta \in [0,2]$.

[K_β] For each $l = 1,2$, we have

$$Z^l = X^l + \kappa\left(\delta^l\right) \star \left(\mu^l - \nu^l\right) + \kappa'\left(\delta^l\right) \star \mu^l,$$

where

(i) X^l is a continuous semimartingale given by (3).

(ii) μ^l is a Poisson random measure on $\mathbb{R}_+ \times E^l$ with intensity measure $\nu^l(\mathrm{d}t,\mathrm{d}x) = \mathrm{d}t\,F^l(\mathrm{d}x)$, where $\left(E^l,\mathcal{E}^l\right)$ is a Polish space and F^l is a σ-finite measure on $\left(E^l,\mathcal{E}^l\right)$.

(iii) $\kappa(x) = x1_{\{|x|\le 1\}}$ and $\kappa'(x) = x - \kappa(x)$ for each $x \in \mathbb{R}$.

(iv) δ^l is a predictable map from $\mathbf{\Omega}^{(0)} \times \mathbb{R}_+ \times E^l$ into $\mathbb{R}$. Moreover, there are a sequence $\left(R_k^l\right)$ of stopping times increasing to ∞ and a sequence $\left(\psi_k^l\right)$ of nonnegative measurable functions on E^l such that

$$\sup_{\omega^{(0)}\in\Omega^{(0)},\,t<R_k^l(\omega^{(0)})} |\delta^l(\omega,t,x)| \le \psi_k^l(x) \text{ and } \int_{E^l} 1\wedge\psi_k^l(x)^\beta F^l(\mathrm{d}x) < \infty.$$

(v) If $\beta < 1$, for the process $f_t = \int_{E^l}\kappa(\delta^l(t,x))F^l(\mathrm{d}x)$, there is a sequence (σ_k) of $\mathbf{F}^{(0)}$-stopping times such that for every k we have a positive constant C_k and $\lambda_k \in (0,3/4)$ satisfying (11) for every $t > 0$ and any bounded $\mathbf{F}^{(0)}$-stopping time τ.

Here and below $\star$ denotes the integral (either stochastic or ordinary) with respect to some (integer-valued) random measure; see Chapter II of Jacod and Shiryaev (2003) for details. Except for the condition (v), the above type of assumption appears in many articles, for example Jacod (2008). Condition [K_β](v) is necessary because sampling times are possibly nonequidistant and nonsynchronous in our situation. In fact, this condition can be seen as a jump-component counterpart of [A5] because using the notation in [K_β](v) $\int_0^t f_s\mathrm{d}s$ is the drift part of the jump-component. We also note that [K_β] implies that for each $l = 1,2$ the generalized Blumenthal–Getoor index of Z^l is less or equal than β.

THEOREM 3.4. *Suppose* [K_β] *and* [$\mathrm{N}_r^\flat$] *hold for some* $\beta \in [0,2]$ *and* $r \in (2,\infty)$. *Suppose also* [C1]-[C2], [A1], [A4], [A6], *and* [T] *are satisfied. Then we have*

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

$$\sup_{0\le s\le t} |\widehat{PTHY}(\mathsf{Z}^1,\mathsf{Z}^2)^n_s - \widehat{PHY}(\mathsf{X}^1,\mathsf{X}^2)^n_s| = o_p(n^{-\frac{1}{4}}) + O_p\left(\left(n^{-\frac{1}{2}}\rho_n^{-1}\right)^{\frac{r-2}{2}}\right) + o_p\left(\rho_n^{1-\beta/2}\right)$$

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

The proof of this theorem is given in Appendix A.4. Combining this result with Theorem 2.1 or Theorem 2.2, we obtain the following results:

THEOREM 3.5 (Consistency of the PTHY estimator in infinite activity case). *Suppose* [K$_2$] *and* [N$_r^\flat$] *hold for some $r \in (2,\infty)$. Suppose also* [C1]–[C2], [A1], [A4], [A6] *and* [T] *are satisfied. Then we have (19) as $n \to \infty$.*

THEOREM 3.6 (Asymptotic mixed normality of the PTHY estimator in infinite activity case).

(a) *Suppose* [A1], [A2](i)–(iii) *and* [A3]–[A6] *are satisfied. Suppose also* [N$_r$] *holds for some $r \in [8,\infty)$,* [K$_\beta$] *holds for some $\beta \in [0, 2-\frac{1}{2\xi'-1})$ and* [T] *holds with $n^{(r-3)/\{2(r-2)\}}\rho_n \to \infty$ and $\rho_n = O\left(n^{-1/\{2(2-\beta)\}}\right)$ as $n \to \infty$. Moreover, suppose $\underline{X}^1 = \underline{X}^2 = 0$. Then (20) holds true as $n \to \infty$ with that $\widetilde{W}$ is the same one as in Theorem 2.2 and $\mathfrak{w}$ is given by (15).*

(b) *Suppose* [A1]–[A6] *are satisfied. Suppose also* [N$_r$] *holds for some $r \in [8,\infty)$,* [K$_\beta$] *holds for some $\beta \in [0, 2-\frac{1}{2\xi'-1})$ and* [T] *holds with $n^{(r-3)/\{2(r-2)\}}\rho_n \to \infty$ and $\rho_n = O(n^{-1/\{2(2-\beta)\}})$ as $n \to \infty$. Then (20) holds true as $n \to \infty$ with that $\widetilde{W}$ is as in the above and $\mathfrak{w}$ is given by (16).*

Note that the assumptions of Theorem 3.6 require at least $\beta < 1$.

Remark 3.3 (Rate of convergence)**.** We shall briefly discuss the best rate of convergence of our estimator available from Theorem 3.6. For simplicity we assume that the noise processes have moments of all orders i.e., [N$_r^\flat$] holds true for all $r > 0$. We also assume that $r_n(t) = o_p\left(n^{-\xi'}\right)$ for any $\xi' \in (0,1)$ and any $t > 0$. In this case, under the assumptions in Theorem 3.6, the best rate of convergence for our estimator is $n^{-1/4}$ if $\beta < 1$. Since the optimal rate of convergence is given by $n^{-1/4}$ in the continuous case according to Bibinger (2011), our estimator achieves the optimal rate if $\beta < 1$. On the other hand, if $\beta \geq 1$ the theorem does not tell us what the best rate is. However, it tells us that the rate of convergence can be faster than $n^{-(1-\beta/2)\varpi}$ for any $\varpi \in (0,1)$. This result can be seen as an analog of the one given by Jacod (2008) in the absence of noise. In fact, in the absence of noise Jacod (2008) showed that the convergence rate of the truncated realized volatility can be faster than $n^{-(2-\beta)\varpi}$ for any $\varpi \in (0,1)$. In the continuous case the optimal rate changes from $n^{-1/2}$ in the absence of noise to $n^{-1/4}$ in the presence of noise, so that the above result seems to be natural.

Remark 3.4 (Positivity)**.** In the present article we focus on estimating the components of the integrated covariance matrix rather than the matrix itself. When the estimation of the matrix itself is the main interest, it is another important issue for applications whether the estimator guarantees positivity in finite samples. Indeed, as far as the author knows, no existing rate-optimal estimators guarantee positivity in the current setup without jumps, except for the realized quasi-maximum likelihood estimator proposed by Shephard and Xiu (2012) recently. Hence we shall give a brief discussion on this issue here.

Let us consider the (noncomponent-wise) estimator for the integrated covariance matrix:

$$PTHY[\mathsf{Z}]^n_t := \begin{bmatrix} PTHY(\mathsf{Z}^1,\mathsf{Z}^1;\widehat{\mathcal{I}},\widehat{\mathcal{I}})^n_t & PTHY(\mathsf{Z}^1,\mathsf{Z}^2;\widehat{\mathcal{I}},\widehat{\mathcal{J}})^n_t \\ PTHY(\mathsf{Z}^1,\mathsf{Z}^2;\widehat{\mathcal{I}},\widehat{\mathcal{J}})^n_t & PTHY(\mathsf{Z}^2,\mathsf{Z}^2;\widehat{\mathcal{J}},\widehat{\mathcal{J}})^n_t \end{bmatrix}.$$

Then, we can show that $PTHY[\mathsf{Z}]^n_t$ is not always positive-semidefinite in the following way. To simplify the problem, we focus on the synchronous case $S^i = T^i$ for every i. Note that in this case we have $\widehat{S}^i = \widehat{T}^i = R^i = S^i = T^i$. Then, in view of the fact that $\bar{K}^{ij} \neq 0$ is equivalent to $|i-j| < k_n$ (because of the synchronicity), we can rewrite our estimator as

$$PTHY[\mathsf{Z}]^n_t = \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j=0}^{N^n_t-k_n} f_0\left(\frac{i-j}{k_n}\right)\widetilde{\mathsf{Z}}_i\left(\widetilde{\mathsf{Z}}_j\right)^*.$$

Here, we define the random vector $\widetilde{\mathsf{Z}}_i = \left(\widetilde{\mathsf{Z}}^1_i,\widetilde{\mathsf{Z}}^2_i\right)^*$ by $\widetilde{\mathsf{Z}}^1_i = \overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \leq \varrho^1_n(\widehat{S}^i)\}}$ and $\widetilde{\mathsf{Z}}^2_j = \overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j 1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \leq \varrho^2_n(\widehat{T}^j)\}}$ (recall that an asterisk denotes the transpose of a matrix) and the function f_0 on $\mathbb{R}$ by $f_0(x) = 1_{(-1,1)}(x)$. Since f_0 is not a positive-definite function, $PTHY[\mathsf{Z}]^n_t$ does not guarantee positivity. Here, a function f on $\mathbb{R}$ is said to be *positive-definite* if $\sum_{i,j=1}^m a_i a_j f(x_i - x_j) \geq 0$ for any (finite) sequence $a_1,\dots,a_m,x_1,\dots,x_m$ of real numbers.

4. SOME RELATED TOPICS FOR STATISTICAL APPLICATION TO FINANCE

4.1. Estimation of the Quadratic Covariation of Jump Parts

As stated in Section 1, we are interested in the estimation of the quadratic covariation of the jump parts of two semimartingales Z^1 and Z^2. This is achieved by estimating the quadratic variation $[Z^1,Z^2]$ due to the formula (2) because we can estimate the integrated covariance $\langle Z^{1,c},Z^{2,c}\rangle_t$ by the PTHY estimator as investigated in the previous section. In the literature such estimators are usually given by consistent estimators for the integrate covariance in the absence of jumps. See Mancini and Gobbi (2012) and Podolskij and Vetter (2009a) for example. Following this approach, we consider the preaveraged HY estimator and we obtain the following result.

PROPOSITION 4.1. *Suppose* [C1]–[C2], [A1], [A4], [A6], [K_2] *and* [$\mathrm{N}^\flat_2$] *are satisfied. Then*

$$\widehat{PHY}\left(\mathsf{Z}^1,\mathsf{Z}^2\right)^n_t = \left[Z^1,Z^2\right]_t + O_p\left(n^{-\frac{1}{4}}\right)$$

as $n\to\infty$ for any $t\in\mathbb{R}_+$.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

See Appendix A.5 for a proof. Consequently, we obtain the following result on the issue of the estimation of the quadratic covariation of the jump parts:

COROLLARY 4.1. *Suppose* [C1]–[C2], [A1], [A4], [A6], [K$_2$], *and* [N$_r^\flat$] *for some* $r \in (2,\infty)$ *are satisfied. Suppose also that* [T] *holds. Then*

$$\widehat{PHY}\left(\mathsf{Z}^1,\mathsf{Z}^2\right)_t^n - \widehat{PTHY}\left(\mathsf{Z}^1,\mathsf{Z}^2\right)_t^n \to^p \sum_{0\le s\le t} \Delta Z_s^1 \Delta Z_s^2$$

as $n \to \infty$ *for any* $t \in \mathbb{R}_+$. *Furthermore, if* [K$_\beta$] *holds for some* $\beta \in \left[0, 2-\frac{1}{2\xi'-1}\right)$, *then*

$$\widehat{PHY}\left(\mathsf{Z}^1,\mathsf{Z}^2\right)_t^n - \widehat{PTHY}\left(\mathsf{Z}^1,\mathsf{Z}^2\right)_t^n = \sum_{0\le s\le t} \Delta Z_s^1 \Delta Z_s^2 + O_p(n^{-\frac{1}{4}})$$

as $n \to \infty$ *for any* $t \in \mathbb{R}_+$, *provided that* $n^{(r-3)/\{2(r-2)\}}\rho_n \to \infty$ *and* $\rho_n = O(n^{-1/\{2(2-\beta)\}})$.

4.2. Autocorrelated Noise

We have so far assumed that the observation noise is not autocorrelated asymptotically, i.e. the sample autocorrelations of the noise at nonzero lags tend to zero as $n \to \infty$. In empirical studies of financial high-frequency data, however, there is evidence that microstructure noise is autocorrelated (see Hansen and Lunde, 2006 and Ubukata and Oya, 2009 for instance). In this section we briefly discuss the case that the observation noise is (asymptotically) autocorrelated as Christensen et al. (2011b) did in the continuous case.

We focus on the synchronous case. That is, we assume that $S^i = T^i$ for all i. Note that in this case it holds that $\widehat{S}^k = \widehat{T}^k = R^k = S^k$ for all k. Let $\left(\lambda_u^l\right)_{u\in\mathbb{Z}_+}$ and $\left(\mu_u^l\right)_{u\in\mathbb{Z}_+}$ $(l = 1,2)$ be four sequences of real numbers such that

$$\sum_{u=1}^{\infty} u|\lambda_u^l| < \infty \quad \text{and} \quad \sum_{u=1}^{\infty} u|\mu_u^l| < \infty. \tag{21}$$

We assume that the observation data $\left(\mathsf{Z}_{S^i}^1\right)$ and $\left(\mathsf{Z}_{T^j}^2\right)$ are of the form

$$\left.\begin{aligned} \mathsf{Z}_{S^i}^1 &= Z_{S^i}^1 + \sum_{u=0}^{i} \lambda_u^1 \zeta_{S^{i-u}}^1 + \sqrt{n}\sum_{u=0}^{i} \mu_u^1 (\underline{X}_{S^{i-u}}^1 - \underline{X}_{S^{i-u-1}}^1), \\ \mathsf{Z}_{T^j}^2 &= Z_{T^j}^2 + \sum_{u=0}^{i} \lambda_u^2 \zeta_{T^{j-u}}^2 + \sqrt{n}\sum_{u=0}^{i} \mu_u^2 (\underline{X}_{T^{j-u}}^2 - \underline{X}_{T^{j-u-1}}^2). \end{aligned}\right\} \tag{22}$$

In other words, the observation noise follows a kind of linear process. Under such a situation the consistency of our estimators is still valid:

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

PROPOSITION 4.2. *Suppose (21) and (22) are satisfied. Suppose also* [C1]–[C2], [A1], [A4], [A6], [K$_2$], *and* [N$_2^\flat$] *are satisfied. Then* $\widehat{PHY}(\mathsf{Z}^1,\mathsf{Z}^2)^n_t \to^p [Z^1,Z^2]_t$ *as* $n\to\infty$ *for any* $t\in\mathbb{R}_+$. *Furthermore, if* [T] *and* [N$_r^\flat$] *for some* $r\in(2,\infty)$ *hold, then* $\widehat{PTHY}(\mathsf{Z}^1,\mathsf{Z}^2)^n_t \to^p [X^1,X^2]_t$ *as* $n\to\infty$ *for any* $t\in\mathbb{R}_+$.

We give a proof of Proposition 4.2 in Appendix A.6. The proof is based on a Beveridge–Nelson type decomposition for the noise.

4.3. Estimation of Asymptotic Variance

In this section we shall briefly discuss the estimation of the asymptotic variance of the PTHY estimator. This is necessary to construct feasible confidence intervals of this estimator, for example. We focus on the simple case that the endogenous terms of the microstructure noise are absent, i.e. $\underline{X}^1=\underline{X}^2=0$. In this case our aim can be achieved by a *kernel-based approach* as in Hayashi and Yoshida (2011) and Koike (2013).

More precisely, let (h_n) be a sequence of positive numbers tending to 0 as $n\to\infty$. For any $s\in\mathbb{R}_+$, put

$$\widehat{[X^l,X^{l'}]'}_s = h_n^{-1}\left(\widehat{PTHY}(\mathsf{Z}^l,\mathsf{Z}^{l'})^n_s - \widehat{PTHY}(\mathsf{Z}^l,\mathsf{Z}^{l'})^n_{(s-h_n)_+}\right),$$
$$\widehat{[X^l]'}_s = \widehat{[X^l,X^l]'}_s, \qquad l,l'=1,2$$

and

$$\widehat{\Psi^{11}}_s = -\frac{1}{h_n k_n^2}\sum_{i:s-h_n<\widehat{S}^{i+1}\le s}\left(\mathsf{Z}^1_{\widehat{S}^i}-\mathsf{Z}^1_{\widehat{S}^{i-1}}\right)\left(\mathsf{Z}^1_{\widehat{S}^{i+1}}-\mathsf{Z}^1_{\widehat{S}^i}\right),$$
$$\widehat{\Psi^{22}}_s = -\frac{1}{h_n k_n^2}\sum_{j:s-h_n<\widehat{T}^{j+1}\le s}\left(\mathsf{Z}^2_{\widehat{T}^j}-\mathsf{Z}^2_{\widehat{T}^{j-1}}\right)\left(\mathsf{Z}^2_{\widehat{T}^{j+1}}-\mathsf{Z}^2_{\widehat{T}^j}\right),$$
$$\widehat{\Psi^{12}\chi}_s = -\frac{1}{2h_n k_n^2}\times\sum_{k:s-h_n<R^{k+1}\le s}\left\{(\mathsf{Z}^1_{\widehat{S}^k}-\mathsf{Z}^1_{\widehat{S}^{k-1}})(\mathsf{Z}^2_{\widehat{T}^{k+1}}-\mathsf{Z}^2_{\widehat{T}^k})+(\mathsf{Z}^1_{\widehat{S}^{k+1}}-\mathsf{Z}^1_{\widehat{S}^k})(\mathsf{Z}^2_{\widehat{T}^k}-\mathsf{Z}^2_{\widehat{T}^{k-1}})\right\}1_{\{\widehat{S}^k=\widehat{T}^k\}}.$$

Then, noting that $n^{-1}N^n_t\to^p\int_0^t 1/G_s\mathrm{d}s$ as $n\to\infty$ for any $t\in\mathbb{R}_+$ (see equation (9) in the Appendix), under the assumptions of Theorem 3.6(a) we can easily show that $\widehat{[X^l]'}_s\to^p[X^l]_{s-}$ and $\widehat{\Psi^{ll}}_s\to^p\Psi^{ll}_{s-}/\theta^2G_{s-}$ for each $l=1,2$ and that $\widehat{[X^1,X^2]'}_s\to^p[X^l,X^{l'}]_{s-}$ and $\widehat{\Psi^{12}\chi}_s\to^p\Psi^{12}_{s-}\chi_{s-}/\theta^2G_{s-}$ as $n\to\infty$ for every $s\in\mathbb{R}_+$, provided that $h_nn^{1/4}\to\infty$ (Recall that G_s and χ_s are introduced in the condition [A1$'$]). In the light of these relationships, we set

$$\widehat{w^2}_{R^k} = \sqrt{n}k_n\psi_{HY}^{-4}\left[\kappa\left\{\widehat{[X^1]'}_{R^k}\widehat{[X^2]'}_{R^k}+\left(\widehat{[X^1,X^2]'}_{R^k}\right)^2\right\}+\widetilde{\kappa}\left\{\widehat{\Psi^{11}}_{R^k}\widehat{\Psi^{22}}_{R^k}+\left(\widehat{\Psi^{12}\chi}_{R^k}\right)^2\right\}\right.$$
$$\left.+\overline{\kappa}\left\{\widehat{[X^1]'}_{R^k}\widehat{\Psi^{22}}_{R^k}+\widehat{[X^2]'}_{R^k}\widehat{\Psi^{11}}_{R^k}+2\widehat{[X^1,X^2]'}_{R^k}\widehat{\Psi^{12}\chi}_{R^k}\right\}\right]|\Gamma^{k+1}| \qquad \textbf{(23)}$$

for every $k \in \mathbb{N}$ and $\widehat{\int_0^t w_s^2 \mathrm{d}s} = \sum_{k:R^{k+1} \leq t} \widehat{w^2}_{R^k} |\Gamma^k|$ for every $t \in \mathbb{R}_+$. Then we obtain the following result:

PROPOSITION 4.3. *Under the assumptions of Theorem 3.6(a), we have*

$$\widehat{\int_0^{\cdot} w_s^2 \mathrm{d}s} \xrightarrow{ucp} \int_0^{\cdot} w_s^2 \mathrm{d}s$$

as $n \to \infty$, provided that $h_n n^{1/4} \to \infty$ and $\sup_{0 \leq t \leq T} (h_n n)^{-1} (N_t^n - N_{(t-h_n)_+}^n)$ is tight as $n \to \infty$ for any $T > 0$.

Proof. Since $\int_0^{\cdot} w_s^2 \mathrm{d}s$ is a continuous nondecreasing process, it is sufficient to prove the pointwise convergence. First, by [A1′], the Burkholder–Davis–Gundy inequality and the Lenglart inequality, we have

$$\widehat{\int_0^t w_s^2 \mathrm{d}s} - \theta \sum_{k:R^{k+1} \leq t} \widetilde{w^2}_{R^k} |\Gamma^k| \to^p 0$$

as $n \to \infty$ for every t, where $\widetilde{w^2}_{R^k}$ is defined by (23) with replacing $|\Gamma^{k+1}|$ by $G_{R^k}^n$. Then, we obtain the desired result by the assumptions and the dominated convergence theorem. ■

The above approach has the disadvantage that it depends strongly on the particular form of the asymptotic variance due to the noise. In fact, it is not proper in the case that the endogenous terms of the noise are present because we have so far known no estimator for the statistic $\int_0^t \left([\underline{X}^1, X^2]_s' F_s^1 - [X^1, \underline{X}^2]_s' F_s^2\right)^2 G_s^{-1} \mathrm{d}s$ which is the asymptotic variance due to the presence of the endogenous noise. We will also need to modify it if the noise is autocorrelated since we will need to replace the covariance matrix Ψ of the noise in the asymptotic variance with the long-run covariance matrix of the noise (see the proof of Proposition 4.2 in Appendix A.6). To avoid this problem, we might rely on the approach used in Section 4 of Christensen et al. (2011b) or the *subsampling approach* developed by Kalnina (2011) recently though it remains for further research to verify the theoretical validity of them.

5. SIMULATION STUDY

In this section, we examine the finite sample performance of our estimators by using Monte Carlo experiments.

5.1. Choice of the Threshold Processes

As is well known, the thresholding method is often sensitive to the selection of thresholds in finite samples; see Shimizu (2010) or the Web Appendix of Mancini and Gobbi (2012) for instance. Therefore, it is important to determine a reasonable rule of selecting thresholds. Here we present an easy but effective way to determine thresholds. Formal study of methods for optimal threshold selection in a given model is an important issue for the future.

We will determine the thresholds for individual processes so that we focus on the univariate case. First we compute an auxiliary estimator $\widehat{\Sigma}_t^n$ for the spot variance process $\Sigma_t = \theta\psi_2[X^1]'_t + \frac{1}{\theta}\psi_1\Psi_t^{11}$ for each sampling time t, where $\psi_1 = \int_0^1 g'(s)^2 \mathrm{d}s$ and $\psi_2 = \int_0^1 g(s)^2 \mathrm{d}s$. In this paper we will use a numerical derivative of the preaveraged bipower variation, i.e.

$$\widehat{\Sigma}_{\widehat{S}^i} = \frac{\mu_1^{-2}}{K - 2k_n + 1} \sum_{p=i-K}^{i-2k_n} |\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^p||\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^{p+k_n}|, \qquad i = K, K+1, \dots, N$$

and $\widehat{\Sigma}_{\widehat{S}^i} = \widehat{\Sigma}_{\widehat{S}^K}$ if $i < K$. Here, μ_1 is the absolute moment of the standard normal distribution, N is the number of the available preaveraged data $\left(\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i\right)$ and K is a bandwidth parameter such that $K = O(n^\alpha)$ as $n \to \infty$ for some $\alpha \in (0.5, 1)$. We will set $K = \lceil N^{3/4} \rceil$ below. Such a kind of spot variance estimator was studied in Bos, Janus, and Koopman (2012). Then we choose

$$\varrho_n^1(\widehat{S}^i) = 2\log(N)^{1+\varepsilon}\widehat{\Sigma}_{\widehat{S}^i}, \qquad i = 0, 1, \dots, N \tag{24}$$

for some $\varepsilon > 0$. We will set $\varepsilon = 0.2$ below.

The heuristic idea behind the above choice of thresholds is as follows. First we recall the following classic result:

THEOREM 5.1 (Pickands, 1967, Thm. 3.4). *Let $(X_i)_{i\in\mathbb{N}}$ be a stationary Gaussian process such that $E[X_i] = 0$, $E[X_i^2] = 1$ and $E[X_i X_{i+k}] = \gamma(k)$. If* $\lim_{k\to\infty}\gamma(k) = 0$, *then* $\max_{1\le i\le n} X_i/\sqrt{2\log n} \to 1$, *almost surely, as $n \to \infty$.*

The most important point of the above theorem is that the random variables X_i, $i = 1, 2, \dots,$ in the theorem can have a kind of dependence structure. This fact is crucial for the present situation because the preaveraged data $\left(\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i\right)$ is k_n-dependent. As a result, Theorem 5.1 has the following implication: Suppose that the observation data is given by a scaled Brownian motion with i.i.d. Gaussian noise. That is, suppose that $\mathsf{Z}_{S^i} = \sigma W_{S^i} + u_i$, $i = 0, 1, \dots,$ where $\sigma > 0$, W_t is a standard Wiener process and (u_i) is an i.i.d. random variables independent of W with $u_i \sim N(0, \omega^2)$. Suppose also that (S^i) is an equidistant sampling scheme. Then the preaveraged data $(\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i)$ is a centered stationary Gaussian process with the autocovariance function vanishing at infinity, so that Theorem 5.1 yields $\max_{0\le i\le N-1}\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i\big/\sqrt{\Sigma\cdot 2\log N} \to 1$ a.s. as $n \to \infty$, where $\Sigma = \theta\psi_2\sigma^2 + \frac{1}{\theta}\omega^2$ is the variance of $\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i$. This result suggests that we may use $\Sigma \cdot 2\log N$ as thresholds. This idea has already been introduced as the *universal threshold* by Donoho and Johnstone (1994) in the context of wavelet shrinkage. Donoho and

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Johnstone (1994) estimated the unknown parameter Σ by the square of the median absolute deviation (MAD) of $(\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i)$ divided by 0.6745, the 0.75-quantile of the standard normal distribution. In the present situation $(\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i)$ is heteroscedastic in general, hence we need to replace Σ with the spot variance process Σ_t and estimate Σ_t by $\widehat{\Sigma}_t^n$. Consequently, we have arrived at the threshold process given by (24), where we multiply the usual universal threshold by $(\log N)^\varepsilon$ to ensure the condition (18).

5.2. Simulation Design

We simulate over the interval $t \in [0,1]$. We normalize one second to be $1/23400$, so that the interval $[0,1]$ contains 6.5 hours. In generating the observation data, we discretize $[0,1]$ into a number $n = 23400$ of intervals.

In order to extract irregular, nonsynchronous observation times from n equispaced division points, we generate random observation times (S^i) and (T^j) using two independent Poisson processes with intensity n/λ_1 and n/λ_2. Here λ_l denotes the average waiting time for new data from process Z^l, so that a typical simulation will have n/λ_l observations of Z^l, $l = 1,2$. Following Barndorff-Nielsen et al. (2011), we vary $\lambda := (\lambda_1, \lambda_2)$ through the following configurations $(3,6)$, $(10,20)$, and $(30,60)$. In addition, we also consider the case that $\lambda = (3, 30)$, where (S^i) are much more frequently sampled than (T^j). This case will be useful for confirming the advantage of our estimator explained in Remark 3.1. Note that because we are simulating in discrete time, it is possible to see common points to the observation times (S^i) and (T^j).

We consider two types of bivariate Lévy processes J_t with no Brownian components to introduce jumps within the considered models. The specifics of the jump processes are as follows:

SCP1 Let L be a stratified normal inverse Gaussian compound Poisson process with a single jump per unit time (i.e., the jump time is uniformly distributed over $[0,1]$ and the jump size follows a normal inverse Gaussian distribution). The jump size is drawn from $\varepsilon\sqrt{S}$, where $\varepsilon \perp\!\!\!\perp S$, $\varepsilon \sim N(0,1)$ and $S \sim IG(c, c^2/\gamma)$, so that $\mathrm{Var}[\varepsilon\sqrt{S}] = E[S] = c$ and $\mathrm{Var}[S] = c^3/(c^2/\gamma) = c\gamma$. Then, we set $J^1 = J^2 = L$.

VG Let L^1 and L^2 be mutually independent variance Gamma processes such that $L_1^l \sim \varepsilon\sqrt{S}$, where $\varepsilon \perp\!\!\!\perp S$, $\varepsilon \sim N(0,1)$, and $S \sim \Gamma(c/\gamma, 1/\gamma)$, so that $\mathrm{Var}[\varepsilon^l\sqrt{S}] = E[S] = c$ for each $l = 1,2$ and $\mathrm{Var}[S] = (c/\gamma)/(1/\gamma)^2 = c\gamma$. Then, we set $J^1 = L^1$ and $J^2 = RL^1 + \sqrt{1-R^2}L^2$.

In the simulation we set $c = 0.1$ and $\gamma = 0.25$. The value of R is given for each model below. Note that each component of the above models coincides with the model simulated in Veraart (2010).

The observation data $\left(\mathsf{Z}^1_{S^i}\right)$ and $(\mathsf{Z}^2_{T^j})$ are generated from the models below.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Model 1 (Barndorff-Nielsen et al., 2011) — the case of stochastic volatility & noise. The following bivariate factor stochastic volatility model is used to generate the continuous semimartingales X^1 and X^2:

$$\mathrm{d}X_t^l = \mu^l \mathrm{d}t + \rho^l \sigma_t^l B_t^k + \sqrt{1-(\rho^l)^2}\sigma_t^l W_t, \qquad \sigma_t^l = \exp\left(\beta_0^l + \beta_1^l \varrho_t^l\right),$$
$$\mathrm{d}\varrho_t^l = \alpha^l \varrho_t^l \mathrm{d}t + \mathrm{d}B_t^l, \qquad l = 1, 2,$$

where (B^1, B^2, W) is a 3-dimensional standard Wiener processes. The initial values for the ϱ_t^l processes at each simulation run are drawn randomly from their stationary distribution, which is $\varrho_t^l \sim N\big(0, \big(-2\alpha^l\big)^{-1}\big)$. We carry out our numerical experiments by using the following parametrization, assumed to be identical across the two volatility factors: $(\mu^l, \beta_0^l, \beta_1^l, \alpha^l, \rho^l) = (0.03, -5/16, 1/8, -1/40, -0.3)$, so that $\beta_0^l = \big(\beta_1^l\big)^2/2\alpha^l$. This choice of parameters implies that integrated volatility has been normalized, in the sense that $E[\int_0^1 (\sigma_s^l)^2 \mathrm{d}s] = 1$. At each simulation run we add noise $\big(U_{k/n}^l\big)_{k=0}^n$ simulated as

$$U_{k/n}^l | \{\sigma, X, J\} \overset{i.i.d.}{\sim} N(0, \omega^2),$$
$$\omega^2 = \eta^2 \sqrt{\frac{1}{n}\sum_{i=1}^n \left(\sigma_{i/n}^l\right)^4}, \quad \text{and} \quad \mathrm{Corr}(U_t^1, U_s^2) = \begin{cases} R & \text{if } t = s, \\ 0 & \text{if } t \neq s, \end{cases}$$

where $R = \sqrt{1-(\rho^1)^2}\sqrt{1-(\rho^2)^2}$ and the noise-to-signal ratio, η^2 takes the value 0.001. Finally, $\mathsf{Z}_{S^i}^1$ and $\mathsf{Z}_{T^j}^2$ are given by $\mathsf{Z}_{S^i}^1 = Z_{S^i}^1 + U_{S^i}^1$ and $\mathsf{Z}_{T^j}^2 = Z_{T^j}^2 + U_{T^j}^2$, where $Z^l = X^l + J^l$, $l = 1, 2$.

Remark.

(i) The form of ω^2 is realistic in the following sense: it is known in some empirical studies that the variance of the microstructure noise is proportionate to the volatility; see Aït-Sahalia and Yu (2009) and Bandi and Russell (2006) for example. The value 0.001 given to the noise-signal ratio η^2 reflects the empirical finding of Hansen and Lunde (2006) (they found that η^2 is typically smaller than 0.001).

(ii) There are two reasons why we gave the correlation between U_t^1 and U_t^2. The first one is due to empirical evidence. Namely, some empirical studies indicate the presence of such a correlation (see e.g. Ubukata and Oya, 2009 and Voev and Lunde, 2007). The second reason is due to a mathematical motivation. If U_t^1 and U_t^2 are uncorrelated, most integrated covariance estimators are automatically unbiased because the cross product of the noise terms has zero mean. For example, the realized covariance estimator is unbiased in such a situation although it is an inconsistent estimator for the integrated covariance in the presence of noise. This fact might make the simulation results for the sample bias good because of an unexpected reason. We would like to avoid such a phenomenon.

(iii) The choice of the value R for the correlation between U_t^1 and U_t^2 came from the following consideration. The first Brownian components of X^1 and X^2 are mutually independent, so the correlation between the second ones only contributes the correlation between X^1 and X^2. Therefore, the choice means that the correlation between U_t^1 and U_t^2 is proportionate to that between X^1 and X^2. Namely, we mimic the relation between the variance of the microstructure noise and the volatility of the latent process. Unfortunately, to the author's knowledge there is no stylized fact about the relation between the correlation of microstructure noise and that of latent processes, so here we assume that a kind of similarity is present between the variance and the correlation.

Model 2 (Jacod et al., 2009) — the case of constant volatility & rounding plus error.

$$Z_t^l = X_t^l + \sigma^l J_t^l, \qquad X_t^l = X_0^l + \sigma^l W_t^l,$$

$$\mathsf{Z}_{k/n}^l = \log\left(\alpha^l \left\lfloor \frac{\exp(Z_{k/n}^l + u_{k/n}^l)}{\alpha^l} \right\rfloor\right), \qquad u_{k/n}^l = \eta_{k/n}^l \log \frac{\alpha^l \lceil \frac{\exp(Z_{k/n}^l)}{\alpha^l} \rceil}{\exp\left(Z_{k/n}^l\right)},$$

where W^1 and W^2 are correlated standard Winer processes independent of J with $\mathrm{d}\left[W^1, W^2\right]_t = R\mathrm{d}t$ and $\left(\eta_{k/n}^l\right)$ is a sequence of independent Bernoulli variables (probabilities $p_{k/n}^l$ and $1 - p_{k/n}^l$ of taking values 1 and 0), with

$$p_{k/n}^l = \log\left(\frac{\exp\left(Z_{k/n}^l\right)}{\alpha^l \lfloor \frac{\exp\left(Z_{k/n}^l\right)}{\alpha^l} \rfloor}\right) \Bigg/ \log\left(\frac{\alpha^l \lceil \frac{\exp(Z_{k/n}^l)}{\alpha^l} \rceil}{\alpha^l \lfloor \frac{\exp(Z_{k/n}^l)}{\alpha^l} \rfloor}\right).$$

We assume that the sequences $\left(\eta_{k/n}^1\right)$ and $\left(\eta_{k/n}^2\right)$ are mutually independent as well as independent of W and J. Parameters used: $\sigma^l = 0.2/\sqrt{252}$, $X_0^l = \log(8)$, $\alpha^l = 0.01$, and $R = 0.5$. Figure 1 shows typical observed paths generated by this model.

Model 3 — the case of stochastic volatility & endogenous noise . The model of the continuous semimartingales X^1 and X^2 is the same one as in Model 1, but the noise processes $U_{S^i}^1$ and $U_{T^j}^2$ are given by

$$U_{S^i}^1 = \delta^1 \sqrt{n/\lambda^1}\left(X_{S^i}^1 - X_{S^{i-1}}^1\right), \qquad U^2 T^j = \delta^2 \sqrt{n/\lambda^2}\left(X_{T^j}^2 - X_{T^{j-1}}^2\right).$$

Here we set $\delta^1 = \delta^2 = -0.01$, so that the microstructure noise is negatively correlated with the returns of the latent continuous semimartingale processes X^1 and X^2. This choice reflects the empirical findings reported in Hansen and Lunde (2006). The factors $\sqrt{n/\lambda^1}$ and $\sqrt{n/\lambda^2}$ are necessary to adjust the effect of the observation frequency: roughly speaking, the variance of $\left(X_{S^i}^1 - X_{S^{i-1}}^1\right)$ is proportionate to (λ^1/n), so it decreases as the observation frequency increases.

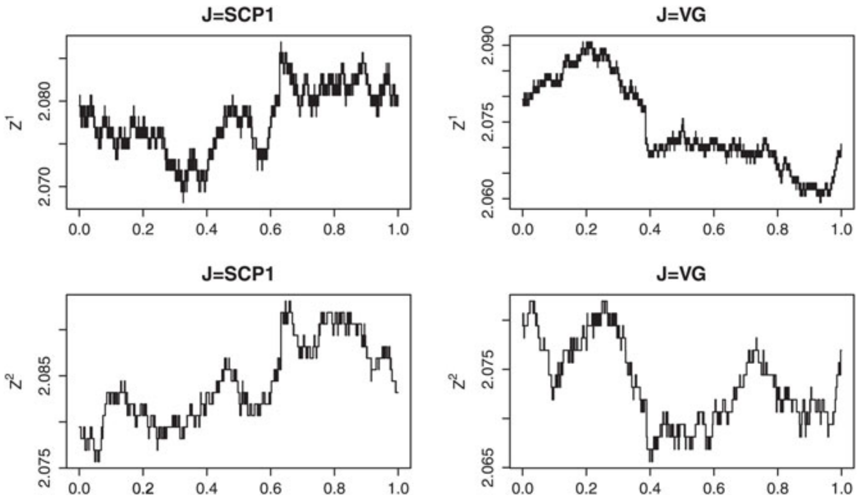

FIGURE 1. Typical observed paths generated by Model 2. The sampling frequency λ is set as $\lambda = (3, 30)$. The jumps of the paths in the left two panels are generated by the SCP1 specification, whereas those in the right two panels are generated by the VG specification. The paths in the top two panels are for Z^1, while those in the bottom two panels are for Z^2.

However, this is opposite to real data. Note that the magnitude of the noise processes in this model is smaller than the one in Model 1. Finally, as in Model 1 we set $\mathsf{Z}^1_{S^i} = Z^1_{S^i} + U^1_{S^i}$ and $\mathsf{Z}^2_{T^j} = Z^2_{T^j} + U^2_{T^j}$, where $Z^l = X^l + J^l, l = 1, 2$.

1000 iterations were run for each model. The simulation of the model paths of X^1 and X^2 has been made using the Euler–Maruyama scheme with n equi–spaced division points.

The tuning parameters for preaveraging are selected as follows. We use $\theta = 0.15$ and $g(x) = x \wedge (1 - x)$ following Christensen et al. (2011b), and set $k_n = \lceil \theta \sqrt{m} \rceil$. Here, m represents the number of refresh times minus 1 i.e., $m = N^n_1$.

5.3. Simulation Results

Table 1 presents the results of estimating the integrated covariance. We report the bias and the root mean squared error (rmse) for our PTHY estimator. As a comparison, we also computed the Wang-Liu-Liu (WLL) estimator defined by (17) and the subsampled realized bipower covariation (BPV) based on 5-minute returns. The threshold processes and the tuning parameters for the WLL estimator are the same as those for the PTHY estimator, except for setting $k_n = \lceil \theta \sqrt{n_1 + n_2} \rceil$ following Wang et al. (2013) as well as Christensen et al. (2011b). Here, n_1 and n_2 represent the numbers of the observed returns for Z^1 and Z^2, respectively. Note that the reported numbers for Model 2 are divided by $(0.2/\sqrt{252})^2$ for normalization.

It is clear from the rmse reported here that the PTHY, the WLL, and the BPV perform well from the simulation studies. Moreover, the bias is very modest when

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

TABLE 1. Simulation results of estimating the integrated covariance

	SCP1			VG		
Estimator	PTHY	WLL	BPV	PTHY	WLL	BPV
Model 1						
$\lambda = (3, 6)$	.003 (.104)	.003 (.106)	.012 (.148)	.005 (.103)	.004 (.106)	.017 (.150)
$\lambda = (10, 20)$	−.006 (.137)	−.008 (.143)	−.028 (.166)	−.003 (.137)	−.005 (.143)	−.022 (.165)
$\lambda = (30, 60)$	−.033 (.203)	−.029 (.198)	−.116 (.236)	−.030 (.200)	−.024 (.199)	−.111 (.235)
$\lambda = (3, 30)$	−.009 (.150)	−.014 (182)	−.056 (.175)	−.007 (.150)	−.012 (.181)	−.155 (.259)
Model 2						
$\lambda = (3, 6)$	.011 (.085)	.011 (.090)	.013 (.122)	.002 (.086)	.001 (.091)	.000 (.118)
$\lambda = (10, 20)$	.005 (.115)	.008 (.121)	−.011 (.124)	−.003 (.118)	−.000 (.122)	−.019 (.121)
$\lambda = (30, 60)$	−.025 (.161)	−.003 (.171)	−.076 (.145)	−.038 (.159)	−.020 (.168)	−.086 (.152)
$\lambda = (3, 30)$	−.008 (.127)	−.008 (141)	−.033 (.128)	−.010 (.130)	−.012 (.142)	−.041 (.127)
Model 3						
$\lambda = (3, 6)$	.000 (.101)	.001 (.105)	−.016 (.149)	.002 (.100)	.003 (.105)	.003 (.156)
$\lambda = (10, 20)$	−.014 (.138)	−.013 (.144)	−.037 (.139)	−.013 (.139)	−.010 (.144)	−.031 (.171)
$\lambda = (30, 60)$	−.053 (.199)	−.039 (.202)	−.141 (.253)	−.051 (.196)	−.036 (.202)	−.013 (.247)
$\lambda = (3, 30)$	−.020 (.150)	−.021 (182)	−.066 (.193)	−.018 (.148)	−.017 (.181)	−.155 (.259)

Note: We report the bias and rmse of the estimators for the integrated covariance included in the simulation study. The number reported in parenthesis is rmse. The reported numbers for Model 2 are divided by $(0.2/\sqrt{252})^2$.

the observation frequency is relatively large. When the observation frequency is small, they are downward biased due to the nonsynchronicity of the observation times and the loss of summands induced by the pre–averaging (for the PTHY and the WLL). The PTHY tends to behave better than the WLL as the observation frequency increases, but the differences seem to be insignificant, except for the case that $\lambda = (3, 30)$. In this case the rmse of the PTHY is significantly smaller than that of the WLL, which is what we would expect due to the argument from Remark 3.1. Besides, the PTHY is superior to the BPV in most of the scenarios.

Table 2 provides the results of estimating the sum of co–jumps $\sum_{0\le s\le 1} \Delta Z_s^1 \Delta Z_s^2$ (i.e., the quadratic covariation of the jump processes). We report the bias and the root mean squared error for the PTHY-based estimator developed in Section 4.1 and compare it to the WLL and BPV-based estimator. The reported numbers for Model 2 are divided by $(0.2/\sqrt{252})^2$ for normalization as shown previously. Similarly to the case that estimating the integrated covariance, the PTHY-based estimator tends to perform slightly better than the others in the higher frequencies, while the WLL-based estimator tends to behave the best in the lower frequencies.

Finally, we turn to the finite sample performance of our feasible central limit theorem. We compute the feasible standardized statistic

$$n^{\frac{1}{4}} \frac{\widehat{PTHY}(\mathsf{Z}^1, \mathsf{Z}^2)_1^n - [X^1, X^2]_1}{\sqrt{\widehat{\int_0^1 w_s^2 \mathrm{d}s}}}, \tag{25}$$

where the asymptotic variance estimator $\widehat{\int_0^1 w_s^2 \mathrm{d}s}$ is given in Section 4.3. In the current choice of g, it holds that $\kappa = 7585/1161216$, $\overline{\kappa} = 151/20160$, and $\widetilde{\kappa} = 1/24$. We use $h_n = 0.4(N_1^n)^{-0.24}$ as a bandwidth parameter to calculate $\widehat{\int_0^1 w_s^2 \mathrm{d}s}$. Table 3 reports the sample mean, standard deviation (SD), and the 95% coverage of (25). Generally speaking, the results are in line with the theory developed in the present article. In terms of the standard deviation, the central limit theorem starts to work for relatively large sample sizes, i.e. it does not work for $\lambda = (30, 60)$. We also observe that the results for Model 3 are worse than those for Model 1. It is not surprising because the estimator of the asymptotic variance used in Model 3 is wrong. In the light of the theoretical asymptotic variance given in (16), the misspecification that (23) does not contain the terms involving the endogenous error becomes important as the difference between λ^1 and λ^2 increases, which is consistent with the simulation results.

Overall, our new method promises to work well at sufficiently high frequencies and it provides a useful tool for analyzing high frequency financial data.

6. EMPIRICAL STUDY

In this section we apply our new method to a set of market data consisting of high-frequency transactions of 5 assets. The 5 assets we will focus on are American Express Company (AXP), Bank of America Corporation (BAC), JPMorgan

TABLE 2. Simulation results of estimating the sum of co-jumps

	SCP1			VG		
Estimator	PTHY	WLL	BPV	PTHY	WLL	BPV
Model 1						
$\lambda = (3, 6)$	−.003 (.037)	−.003 (.041)	−.022 (.077)	−.007 (.041)	−.008 (.044)	−.030 (.097)
$\lambda = (10, 20)$	.002 (.055)	−.000 (.056)	−.020 (.130)	−.003 (.053)	−.005 (.057)	−.028 (.100)
$\lambda = (30, 60)$	.021 (.089)	.012 (.084)	−.016 (.096)	.014 (.092)	.004 (.087)	−.026 (.121)
$\lambda = (3, 30)$	.005 (.065)	.002 (066)	−.022 (.084)	.001 (.063)	−.004 (.069)	−.031 (.109)
Model 2						
$\lambda = (3, 6)$	−.009 (.037)	−.009 (.041)	−.025 (.082)	−.009 (.034)	−.008 (.036)	−.021 (.071)
$\lambda = (10, 20)$	−.005 (.054)	.007 (.051)	−.024 (.087)	−.004 (.046)	−.007 (.049)	−.021 (.072)
$\lambda = (30, 60)$	.023 (.086)	−.001 (.070)	−.022 (.104)	.024 (.085)	.002 (.072)	−.013 (.084)
$\lambda = (3, 30)$	.001 (.057)	−.005 (.055)	−.028 (.087)	.001 (.060)	−.003 (.057)	−.020 (.076)
Model 3						
$\lambda = (3, 6)$	−.001 (.027)	−.001 (.041)	−.010 (.065)	−.004 (.041)	−.006 (.045)	−.032 (.098)
$\lambda = (10, 20)$	.009 (.059)	.004 (.059)	−.024 (.083)	.005 (.058)	−.001 (.059)	−.033 (.100)
$\lambda = (30, 60)$	.038 (.093)	.020 (.089)	−.005 (.069)	.003 (.103)	−.014 (.097)	−.033 (.118)
$\lambda = (3, 30)$	.010 (.069)	.002 (.063)	−.032 (.104)	.010 (.069)	−.001 (.069)	−.035 (.111)

Note: We report the bias and rmse of the estimators for the sum of co-jumps included in the simulation study. The number reported in parenthesis is rmse. The reported numbers for Model 2 are divided by $(0.2/\sqrt{252})^2$.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

TABLE 3. Simulation results of the standardized estimates

	SCP1			VG		
	Mean	SD	Coverage (95%)	Mean	SD	Coverage (95%)
Model 1						
$\lambda = (3, 6)$	−.063	1.009	95.2%	−.015	1.000	95.2%
$\lambda = (10, 20)$	−.174	1.009	93.5%	−.124	.996	94.6%
$\lambda = (30, 60)$	−.328	.936	94.2%	−.290	.920	95.5%
$\lambda = (3, 30)$	−.225	1.029	93.3%	−.192	1.022	94.4%
Model 2						
$\lambda = (3, 6)$	−.042	.988	96.2%	−.055	.984	95.7%
$\lambda = (10, 20)$	−.109	.988	95.6%	−.122	.983	95.6%
$\lambda = (30, 60)$	−.331	.916	94.2%	−.327	.902	94.4%
$\lambda = (3, 30)$	−.168	.977	95.3%	−.184	.998	94.3%
Model 3						
$\lambda = (3, 6)$	−.097	1.020	94.6%	−.055	1.012	94.8%
$\lambda = (10, 20)$	−.256	1.031	93.3%	−.217	1.027	93.3%
$\lambda = (30, 60)$	−.489	.979	91.7%	−.457	.968	92.9%
$\lambda = (3, 30)$	−.329	1.067	92.2%	−.294	1.067	92.6%

Note: We report the sample mean, standard deviation (SD), and the 95% coverage of the standardized estimates (25) included in the simulation study.

Chase & Co. (JPM), The Travelers Companies, Inc. (TRV), and S&P 500 Depository Receipts (SPY). The first four stocks are chosen from the financial sector of the 30 Dow Jones Industrial Average (DJIA) stocks in April, 2010. We involve the SPY because it is often used as a benchmark for empirical work in this area. The sample period runs from 1 April 2010 to 28 February 2011, delivering 230 distinct days. The data is the collection of trades recorded on the New York Stock Exchange (NYSE), taken from the Trades and Quotes (TAQ) database. Prior to the analysis, we apply the filters presented in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) to remove the obvious outliers from the data. Table 4 reports the numbers of observations per day for each asset (after the data cleaning), averaged across the 230 day period. It implies the average waiting times for new observations are between about 3-seconds and 20-seconds. We also note that all the computations below are based on log prices multiplied by a factor of 100.

The aim of this empirical study is to compare the proposed method with a traditional one based on moderate frequency data (here we consider the 5-minute frequency). In particular, we focus on how nonsynchronicity and jumps play

TABLE 4. Average over daily number of high frequency observations

AXP	BAC	JPM	TRV	SPY
2332	1078	2939	1295	6957

different roles at different frequencies and whether the difference of observation frequencies delivers different economic results. Specifically, we consider the subsampled realized bipower covariation (BPV) based on 5-minute returns as the traditional one.

First, we estimate the integrated covariances for every pair of the 5 assets each day. Table 5 contains the integrated covariance estimates averaged across the 230 trading days. They are computed using the PTHY estimator and the BPV estimator. The threshold processes and the tuning parameters for preaveraging are the same as in the simulation. In Table 5, the numbers above the leading diagonal are results for the PTHY estimator, the numbers below are for the BPV estimator. Asterisks indicate the BPV estimates that are outside the 95% or 99% confidence intervals for the corresponding PTHY estimates, constructed from the asymptotic variances in Section 4.3 with the bandwidth parameter h_n selected in a similar manner to the simulation. As the table reveals, the estimates of the BPV are significantly smaller than those of the PTHY, except for the AXP-SPY pair. This is consistent with the simulation study in the previous section. This result suggests that the BPV estimator suffers from the Epps effect, and it is not surprising because even in the 10 minute frequency Epps (1979) found such a phenomenon.

Next, we turn to the question of whether our new method delivers a different economic result compared with the traditional one. To accomplish this purpose we compute the average of the sums of co-jumps between the SPY and the other four stocks each day. This corresponds to estimating the sum of co-jumps between the S&P500 index and the equal-weighted portfolio of the financial sector of the DJIA stocks. Figure 2 depicts PTHY (black solid lines) and BPV (gray crosses)-based estimates for each day. We find a large variation on May 6, which is presumably due to the Flash Crash (see the left panel of Figure 3). If the Flash Crash caused this variation, the asset prices supposedly jumped to the same direction on that day, so that the sum of co-jumps would be positive. The PTHY-based estimate is consistent with this intuition, while the BPV-based one is not. Our method is based

TABLE 5. Summary statistics for integrated covariance estimates

		PTHY				
		AXP	BAC	JPM	TRV	SPY
BPV	AXP		1.61	1.50	0.79	0.96
	BAC	1.49**		1.93	0.82	1.05
	JPM	1.39**	1.74**		0.77	0.95
	TRV	0.72**	0.73**	0.70**		0.52
	SPY	0.95	0.97**	0.92*	0.49*	

Note: We report average integrated covariance estimates. The upper diagonal is based on the PTHY estimator, whereas the lower diagonal is based on the BPV estimator.
* BPV estimates are outside the 95% confidence intervals for the corresponding PTHY estimates.
** BPV estimates are outside the 99% confidence intervals for the corresponding PTHY estimates.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

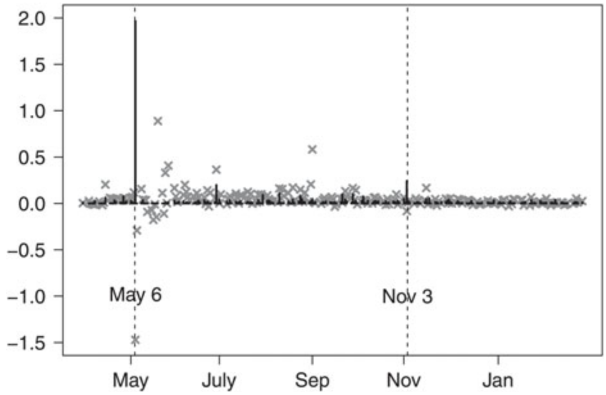

FIGURE 2. Daily estimated sum of co-jumps between the S&P500 index and the financial sector of the DJIA stocks. The black solid lines denote the PTHY-based estimates, whereas the gray crosses donate the BPV-based estimates. The two vertical dashed lines correspond to 6 May 2010 and 3 November 2010 respectively.

on the assumption that there is an underlying continuous time stochastic process driving the dynamics of the data, so the co-jumps are driven by this process and its sign does not depend on which frequency we observe the data at. To explain this finding, we zoom in on the relevant sub-period (see the right panel of Figure 3). Then, we find that the violent fluctuation observed in a relatively large scale turns out to consist of many consecutive small variations. Indeed, this fact was already emphasized in Christensen et al. (2011a). Although a few relatively large returns still seem to remain, the 5-minute frequency is apparently insufficient to capture them precisely. This might give a reasonable explanation for our finding.

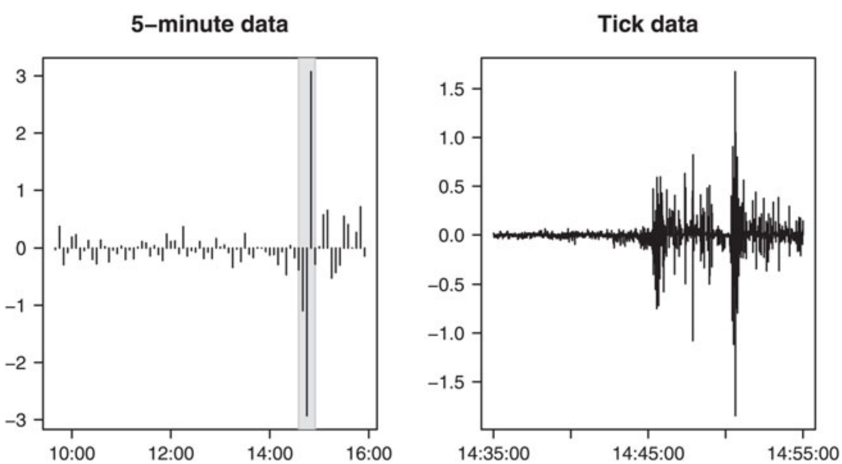

FIGURE 3. Intraday log transaction price returns of SPY on 6 May 2010 (multiplied by a factor of 100).

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

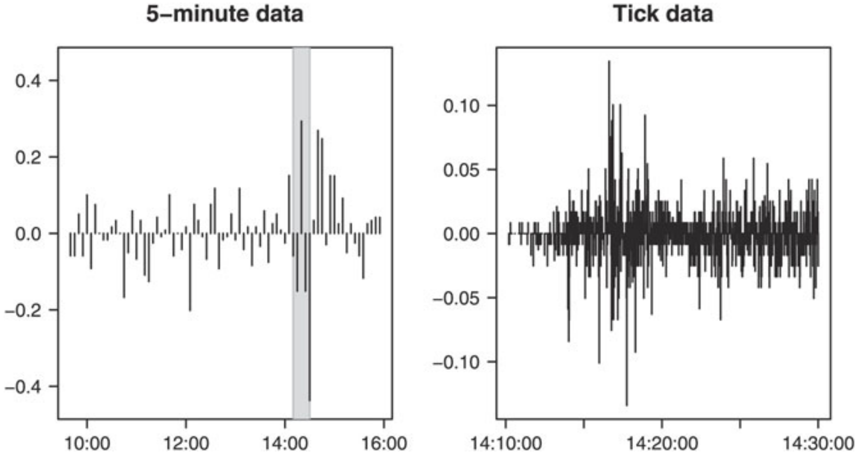

FIGURE 4. Intraday log transaction price returns of SPY on 3 November 2010 (multiplied by a factor of 100).

Namely, the BPV estimator identifies the signs of the co-jumps wrong due to its use of the low resolution data (however, note that practically one can also interpret this finding as the co-jumps tending to have opposite signs at the 5-minute frequency, especially for a market participant who trades at this frequency). A similar phenomenon is found on November 3, when the PTHY-based estimate has the second largest value among the period considered here. In fact, the market fluctuated violently on that day due to the Federal Open Market Committee (FOMC) statement (see Figure 4). Interestingly, from the right panels of Figures 3–4 the relative contribution of jumps to the total variation on November 3 looks larger than that on May 6. In fact, when we compute the relative contribution of the sum of co-jumps to the entire quadratic covariation i.e., $\sum_{0\leq s\leq 1}\Delta Z^1_s \Delta Z^2_s/[Z^1, Z^2]_1$ using our estimator, we get the values 0.13 on May 6 and 0.19 on November 3, which supports this observation. We also find some moderate negative co-jumps in the BPV-based estimates, which are contrary to the economic intuition. These findings suggest the potential usefulness of our new method for empirical studies.

REFERENCES

Aït-Sahalia, Y., J. Fan, & D. Xiu (2010) High-frequency covariance estimates with noisy and asynchronous financial data. *Journal of the American Statistical Association* 105, 1504–1517.

Aït-Sahalia, Y., J. Jacod, & J. Li (2012) Testing for jumps in noisy high frequency data. *Journal of Econometrics* 168, 207–222.

Aït-Sahalia, Y. & J. Yu (2009) High frequency market microstructure noise estimates and liquidity measures. *Annals of Applied Statistics* 3, 422–457.

Andersen, T.G., T. Bollerslev, & F.X. Diebold (2007) Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics* 89, 701–720.

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

Barndorff-Nielsen, O.E., S.E.G. Graversen, J. Jacod, M. Podolskij, & N. Shephard (2006) A central limit theorem for realised power and bipower variations of continuous semimartingales. In Y. Kabanov, R. Liptser, & J. Stoyanov (eds.), *From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift*. Springer-Verlag, pp. 33–69.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde, & N. Shephard (2009) Realized kernels in practice: Trades and quotes. *Econometrics Journal* 12, C1–C32.

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

Barndorff-Nielsen, O.E. & N. Shephard (2004a) Measuring the Impact of Jumps in Multivariate Price Processes using Bipower Covariation. Discussion paper, Nuffield College, Oxford University.

Barndorff-Nielsen, O.E. & N. Shephard (2004b) Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2, 1–37.

Beveridge, S. & C.R. Nelson (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'buisiness cycle'. *Journal of Monetary Economics* 7, 151–174.

Bibinger, M. (2011) Efficient covariance estimation for asynchronous noisy high-frequency data. *Scandinavian Journal of Statistics* 38, 23–45.

Bibinger, M. (2012) An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: Asymptotic distribution theory. *Stochastic Processes and their Applications* 122, 2411–2453.

Bos, C.S., P. Janus, & S.J. Koopman (2012) Spot variance path estimation and its application to high-frequency jump testing. *Journal of Financial Econometrics* 10, 354–389.

Boudt, K., C. Croux, & S. Laurent (2011) Outlyingness weighted covariation. *Journal of Financial Econometrics* 9, 657–684.

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

Christensen, K., R. Oomen, & M. Podolskij (2011a) Fact or friction: Jumps at ultra high frequency. CREATES Research Paper 2011–19, Aarhus University.

Christensen, K., M. Podolskij, & M. Vetter (2011b) On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. CREATES Research Paper 2011–53, Aarhus University.

Cont, R. & Y.H. Kan (2011) Dynamic hedging of portfolio credit derivatives. *SIAM Journal on Financial Mathematics* 2, 112–140.

Dahlhaus, R. (1997) Fitting time series models to nonstationary processes. *Annals of Statistics* 25, 1–37.

Delbaen, F. & W. Schachermayer (1994) A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* 300, 463–520.

Diebold, F.X. (2006) On market microstructure noise and realized volatility. *Journal of Business and Economic Statistics* 24, 181–183. Discussion of Hansen & Lunde (2006).

Donoho, D.L. & I.M. Johnstone (1994) Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81, 232–246.

Epps, T.W. (1979) Comovements in stock prices in the very short run. *Journal of the American Statistical Association* 74, 291–298.

Freedman, D.A. (1975) On tail probabilities for martingales. *Annals of Probability* 3, 100–118.

Fukasawa, M. (2010) Realized volatility with stochastic sampling. *Stochastic Processes and their Applications* 120, 829–852.

Fukasawa, M. & M. Rosenbaum (2012) Central limit theorems for realized volatility under hitting times of an irregular grid. *Stochastic Processes and their Applications* 122, 3901–3920.

Hansen, P.R. & A. Lunde (2006) Realized variance and market microstructure noise. *Journal of Business & Economic Statistics* 24, 127–161.

Hayashi, T., J. Jacod, & N. Yoshida (2011) Irregular sampling and central limit theorems for power variations: The continuous case. *Annales de l'Institut Henri Poincaré-Probabilités et Statistiques* 47, 1197–1218.

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

Hayashi, T. & N. Yoshida (2011) Nonsynchronous covariation process and limit theorems. *Stochastic Processes and their Applications* 121, 2416–2454.

Jacod, J. (1997) On continuous conditional Gaussian martingales and stable convergence in law. In Jacques Azéma, Marc Yor and Michel Emery (eds.), *Séminaire de probabilitiés xxxi*. Lecture Notes in Mathematics, vol. 1655, pp. 232–246. Springer.

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

Jacod, J., Y. Li, P.A. Mykland, M. Podolskij, & M. Vetter (2009) Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications* 119, 2249–2276.

Jacod, J., M. Podolskij, & M. Vetter (2010) Limit theorems for moving averages of discretized processes plus noise. *Annals of Statistics* 38, 1478–1545.

Jacod, J. & P. Protter (2012) *Discretization of processes*. *Stochastic Modelling and Applied Probability*, vol. 67. Springer.

Jacod, J. & A.N. Shiryaev (2003) *Limit theorems for stochastic processes*, 2nd ed. Springer.

Jing, B.-Y., C.-X. Li, & Z. Liu (2011) On Estimating the Integrated Co-Volatility using Noisy High Frequency Data with Jumps. Working paper.

Kalnina, I. (2011) Subsampling high frequency data. *Journal of Econometrics* 161, 262–283.

Kalnina, I. & O. Linton (2008) Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. *Journal of Econometrics* 147, 47–59.

Koike, Y. (2013) An estimator for the cumulative co-volatility of asynchronously observed semimartingales with jumps. *Scandinavian Journal of Statistics* 41, 460–481.

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

Li, Y., Z. Zhang, & X. Zheng (2013) Volatility inference in the presence of both endogenous time and microstructure noise. *Stochastic Processes and their Applications* 123, 2696–2727.

Mancini, C. (2001) Disentangling the jumps of the diffuson in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari* 64, 19–47.

Mancini, C. & F. Gobbi (2012) Identifying the Brownian covariation from the co-jumps given discrete observations. *Econometric Theory* 28, 249–273.

Pickands, J., III (1967) Maxima of stationary Gaussian processes. *Probability Theory and Related Fields* 7, 190–223.

Podolskij, M. & M. Vetter (2009a) Bipower-type estimation in a noisy diffusion setting. *Stochastic Processes and their Applications* 119, 2803–2831.

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

Podolskij, M. & D. Ziggel (2010) New tests for jumps in semimartingale models. *Statistical Inference for Stochastic Processes* 13, 15–41.

Renault, E. & B.J. Werker (2011) Causality effects in return volatility measures with random times. *Jornal of Econometrics* 160, 272–279.

Robert, C.Y. & M. Rosenbaum (2012) Volatility and covariation estimation when microstructure noise and trading times are endogenous. *Mathematical Finance* 22, 133–164.

Shephard, N. & D. Xiu (2012) Econometric Analysis of Multivariate Realised QML: Estimation of the Covariation of Equity Prices under Asynchronous Trading. Technical report, University of Oxford and University of Chicago.

Shimizu, Y. (2003) Estimation of diffusion processes with jumps from discrete observations. Master's thesis, University of Tokyo.

Shimizu, Y. (2010) Threshold selection in jump-discriminant filter for discretely observed jump processes. *Statistical Methods & Applications* 19, 355–378.
Todorov, V. & T. Bollerslev (2010) Jumps and betas: A new framework for disentangling and estimating systematic risks. *Journal of Econometrics* 157, 220–235.
Ubukata, M. & K. Oya (2009) Estimation and testing for dependence in market microstructure noise. *Journal of Financial Econometrics* 7, 106–151.
Veraart, A.E. (2010) Inference for the jump part of quadratic variation of Itô semimartingales. *Econometric Theory* 26, 331–368.
Voev, V. & A. Lunde (2007) Integrated covariance estimation using high-frequency data in the presence of noise. *Journal of Financial Econometrics* 5, 68–104.
Wang, K., J. Liu, & Z. Liu (2013) Disentangling the effect of jumps on systematic risk using a new estimator of integrated co-volatility. *Journal of Banking & Finance* 37, 1777–1786.
Zhang, L. (2011) Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics* 160, 33–47.

APPENDIX

A.1. Proof of Theorem 2.1

First note that for the proof we can use a localization procedure, and which allows us to systematically replace the conditions [C1], [C2], and [$\mathrm{N}_r^\flat$] by the following strengthened version:

[SC1] There is a positive constant K such that $n^{-1}N_t^n \leq K$ for all n and t.

[SC2] [C2] holds, and $(A^1)'$, $(A^2)'$, $(\underline{A}^1)'$, $(\underline{A}^2)'$ and $[V, W]'$ for each $V, W = X^1, X^2, \underline{X}^1, \underline{X}^2$ are bounded.

[$\mathrm{SN}_r^\flat$] $(\int |z|^r Q_t(\mathrm{d}z))_{t\in\mathbb{R}_+}$ is a bounded process.

We need a modification of sampling times as follows. We write $\bar{r}_n = n^{-\xi'}$. Next, let $\upsilon_n = \inf\{t | r_n(t) > \bar{r}_n\}$, and define a sequence $(\widetilde{S}^i)_{i\in\mathbb{Z}^+}$ sequentially by $\widetilde{S}^i = S^i$ if $S^i < \upsilon_n$, otherwise $\widetilde{S}^i = \widetilde{S}^{i-1} + \bar{r}_n$. Then, $(\widetilde{S}^i)$ is obviously a sequence of $\mathbf{F}^{(0)}$-stopping times satisfying (4) and $\sup_{i\in\mathbb{N}}(\widetilde{S}^i - \widetilde{S}^{i-1}) \leq \bar{r}_n$. Furthermore, for any $t > 0$ we have $P\left(\bigcup_i \{\widetilde{S}^i \wedge t \neq S^i \wedge t\}\right) \leq P(\upsilon_n < t) \to 0$ as $n \to \infty$ by (5). By replacing (S^i) with (T^j), we can construct a sequence $(\widetilde{T}^j)$ in a similar manner. This argument implies that we may also assume that

$$\sup_{i\in\mathbb{N}}\left(S^i - S^{i-1}\right) \vee \sup_{j\in\mathbb{N}}\left(T^j - T^{j-1}\right) \leq \bar{r}_n \tag{A.1}$$

by an appropriate localization procedure.

Now we start the main body of the proof. Throughout the discussions, for (random) sequences (x_n) and (y_n), $x_n \lesssim y_n$ means that there exists a (nonrandom) constant $C \in [0,\infty)$ such that $x_n \leq C y_n$ for large n. At the beginning, we remark some elementary properties of refresh times.

PROPOSITION A.1.1. *The following statements are true.*

(a) $\widehat{S}^k \vee \widehat{T}^k = R^k$ *for every* k.

(b) $(\widehat{S}^i < \widehat{T}^j) \Rightarrow (i \leq j)$ *and* $(\widehat{S}^i > \widehat{T}^j) \Rightarrow (i \geq j)$ *for all* i, j.

Proof.

(a) Obvious.

(b) Since $\widehat{T}^j \leq R^j < \widehat{S}^{j+1}$, $(\widehat{S}^i < \widehat{T}^j)$ implies $\widehat{S}^i < \widehat{S}^{j+1}$, hence $i \leq j$. Consecutively, we obtain the former statement. By symmetry we also obtain the latter statement.

■

Set $\bar{I}^i = [\widehat{S}^i, \widehat{S}^{i+k_n})$ and $\bar{J}^j = [\widehat{T}^j, \widehat{T}^{j+k_n})$. Since $(\bar{I}^i \cap \bar{J}^j \neq \emptyset) \Rightarrow (|i-j| \leq k_n)$ by Proposition A.1.1(b), for any $i, j \in \mathbb{Z}_+$ it holds that

$$\sum_{j=0}^{\infty} \bar{K}^{ij} \leq 2k_n + 1, \qquad \sum_{i=0}^{\infty} \bar{K}^{ij} \leq 2k_n + 1 \tag{A.2}$$

Next we introduce some notation. For processes V and W, $V \bullet W$ denotes the integral (either stochastic or ordinary) of V with respect to W. For any semimartingale V and any (random) interval I, we define the processes $V(I)_t$ and I_t by $V(I)_t = \int_0^t 1_I(s-)\mathrm{d}V_s$ and $I_t = 1_I(t)$ respectively. Moreover, for $u \in \{g, g'\}$ define the processes $\bar{V}_u(\widehat{\mathcal{I}})^i$ and $\bar{V}_v(\widehat{\mathcal{J}})^j$ by

$$\bar{V}_u(\widehat{\mathcal{I}})^i_t = \sum_{p=1}^{k_n-1} u_p^n V(\widehat{I}^{i+p})_t, \qquad \bar{V}_u(\widehat{\mathcal{J}})^j_t = \sum_{q=1}^{k_n-1} u_q^n V(\widehat{J}^{j+q})_t,$$

where $u_p^n = u(p/k_n)$.

Define the processes $\mathfrak{E}^1$ and $\mathfrak{E}^2$ by $\mathfrak{E}^1_t = -\frac{1}{k_n}\sum_{p=1}^{\infty} \zeta^1_{\widehat{S}^p} 1_{\{\widehat{S}^p \leq t\}}$ and $\mathfrak{E}^2_t = -\frac{1}{k_n}\sum_{q=1}^{\infty} \zeta^2_{\widehat{T}^q} 1_{\{\widehat{T}^q \leq t\}}$ respectively. Then $\mathfrak{E}^2$ and $\mathfrak{E}^2$ are obviously purely discontinuous locally square-integrable martingales on $\mathcal{B}$ if [C2] holds (note that both $(\widehat{S}^i)$ and $(\widehat{T}^j)$ are $\mathbf{F}^{(0)}$-stopping times). We also define the processes $\mathfrak{M}^1$, $\mathfrak{M}^2$, $\mathfrak{A}^1$, and $\mathfrak{A}^2$ by $\mathfrak{M}^1 = -\frac{\sqrt{n}}{k_n}(\sum_{p=1}^{\infty} \check{I}^p_-) \bullet \underline{M}^1$, $\mathfrak{M}^2 = -\frac{\sqrt{n}}{k_n}(\sum_{q=1}^{\infty} \check{J}^q_-) \bullet \underline{M}^2$, $\mathfrak{A}^1 = -\frac{\sqrt{n}}{k_n}(\sum_{p=1}^{\infty} \check{I}^p_-) \bullet \underline{A}^1$ and $\mathfrak{A}^2 = -\frac{\sqrt{n}}{k_n}(\sum_{q=1}^{\infty} \check{J}^q_-) \bullet \underline{A}^2$. Then, set $\mathfrak{U}^l = \mathfrak{E}^l + \mathfrak{M}^l + \mathfrak{A}^l$ for $l = 1, 2$.

For any semimartingales V, W and any $u, v \in \{g, g'\}$, set

$$\mathbb{H}_{u,v}(V, W)^n_t = \frac{1}{(\psi_{HY} k_n)^2} \sum_{i,j \geq 1, \bar{R}^{\vee}(i,j) \leq t} \bar{V}_u(\widehat{\mathcal{I}})^i_t \bar{W}_v(\widehat{\mathcal{J}})^j_t \bar{K}^{ij}_t,$$

where $\bar{R}^{\vee}(i,j) = \widehat{S}^{i+k_n} \vee \widehat{T}^{j+k_n}$ and $\bar{K}^{ij}_t = 1_{\{\bar{I}^i(t) \cap \bar{J}^j(t) \neq \emptyset\}}$ (recall that $I(t)$ denotes $I \cap [0,t)$ for a random interval I and a time t). Then we can prove the following lemma:

LEMMA A.1.1. *Suppose that* [SC1]-[SC2], [SN$_2^\flat$] *and (A.1) are satisfied. Then we have*

$$\sup_{0 \leq s \leq t} \left| \widehat{PHY}(\mathsf{X}^1, \mathsf{X}^2)^n - \mathbb{H}[g]^n \right| = O_p(n^{-\gamma})$$

for any $t > 0$ as $n \to \infty$, where $\mathbb{H}[g]^n = \mathbb{H}_{g,g}(X^1, X^2)^n + \mathbb{H}_{g,g'}(X^1, \mathfrak{U}^2)^n + \mathbb{H}_{g',g}(\mathfrak{U}^1, X^2)^n + \mathbb{H}_{g',g'}(\mathfrak{U}^1, \mathfrak{U}^2)^n$ and $\gamma = \xi' - 1/2$.

Proof. We decompose the target quantity as

$$\widehat{PHY}\left(\mathsf{X}^1,\mathsf{X}^2\right)_s^n - \mathbb{H}[g]_s^n$$
$$= \frac{1}{(\psi_{HY}k_n)^2}\left[\sum_{\substack{i=0 \text{ or } j=0\\ \bar{R}^\vee(i,j)\le s}} \overline{\mathsf{X}}^1\left(\widehat{\mathcal{I}}\right)^i \overline{\mathsf{X}}^2\left(\widehat{\mathcal{J}}\right)^j \bar{K}^{ij}\right.$$
$$+\sum_{i,j\ge 1\,\bar{R}^\vee(i,j)\le s} \overline{\mathsf{X}}^1\left(\widehat{\mathcal{I}}\right)^i \left\{\sum_{q=0}^{k_n-1}\left\{k_n\Delta(g)_q^n-(g')_q^n\right\}\mathfrak{U}^2\left(\hat{J}^{j+q}\right)_s\right\}\bar{K}_s^{ij}$$
$$\left.+\sum_{i,j\ge 1,\bar{R}^\vee(i,j)\le s}\left\{\sum_{p=0}^{k_n-1}\left\{k_n\Delta(g)_p^n-(g')_p^n\right\}\mathfrak{U}^1\left(\widehat{I}^{i+p}\right)_s\right\}\overline{\mathsf{X}}^2\left(\widehat{\mathcal{J}}\right)^j \bar{K}_s^{ij}\right]$$
$$=: \mathbb{A}_s^1+\mathbb{A}_s^2+\mathbb{A}_s^3,$$

where $\Delta(g)_p^n = g_{p+1}^n - g_p^n$. First, Proposition A.1.1, the Schwarz inequality, [SC2], [$\mathrm{SN}_2^\flat$] and (A.1) imply that $E[\sup_{0\le s\le t}|\mathbb{A}_s^1|]\lesssim \bar{r}_n = O(n^{-\gamma})$. Next, the Lipschitz continuity of g', [SC2], [$\mathrm{SN}_2^\flat$], and (A.1) yield

$$E\left[\left|\sum_{q=0}^{k_n-1}\left\{k_n\Delta(g)_q^n-(g')_q^n\right\}\mathfrak{U}^2\left(\hat{J}^{j+q}\right)_t\right|^2 |\mathcal{F}_{\widehat{T}^{j-1}\wedge t}\right]\lesssim k_n^{-1}\bar{r}_n$$

uniformly in j. On the other hand, [SC2], $\left[\mathrm{SN}_2^\flat\right]$, and (A.1) imply that $E\left[|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 |\mathcal{F}_{S^{i-1}}\right]\lesssim k_n\bar{r}_n$ uniformly in i. Therefore, by Proposition A.1.1, the Schwarz inequality and [SC1] we obtain

$$E\left[\sup_{0\le s\le t}\left|\mathbb{A}_s^2\right|\right]\lesssim k_n^{-2}\bar{r}_n\sum_{|i-j|\le k_n}E\left[1_{\left\{\widehat{S}^{i-1}\wedge\widehat{T}^{j-1}\le t\right\}}\right]\lesssim k_n\bar{r}_n = O\left(b_n^\gamma\right).$$

Similarly, it can also be shown that $E\left[\sup_{0\le s\le t}\left|\mathbb{A}_s^3\right|\right] = O(n^{-\gamma})$, and thus we complete the proof. ■

According to the above lemma, we need to analyze the asymptotic property of the quantity $\mathbb{H}[g]^n$. For this purpose, we define the processes $\bar{L}_{u,v}(V,W)^n$ and $\mathbb{M}_{u,v}(V,W)^n$ by

$$\bar{L}_{u,v}(V,W)^{ij} = \bar{V}_u\left(\widehat{\mathcal{I}}^i\right)_-\bullet\bar{W}_v\left(\widehat{\mathcal{J}}^j\right)+\bar{W}_v\left(\widehat{\mathcal{J}}^j\right)_-\bullet\bar{V}_u\left(\widehat{\mathcal{I}}^i\right),$$
$$\mathbb{M}_{u,v}(V,W)_t^n = \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j\ge 1,\bar{R}^\vee(i,j)\le t}\bar{K}_t^{ij}\bar{L}_{u,v}(V,W)_t^{ij},$$

and set $\mathbb{M}(1)^n = \mathbb{M}_{g,g}\left(X^1,X^2\right)^n$, $\mathbb{M}(2)^n = \mathbb{M}_{g',g'}\left(\mathfrak{U}^1,\mathfrak{U}^2\right)^n$, $\mathbb{M}(3)^n = \mathbb{M}_{g,g'}\left(X^1,\mathfrak{U}^2\right)^n$, and $\mathbb{M}(4)^n = \mathbb{M}_{g',g}\left(\mathfrak{U}^1,X^2\right)^n$. Then we obtain the following result:

LEMMA A.1.2. *Suppose that (A.1) and* [SC1]–[SC2], [$\mathrm{SN}_2^\flat$] *are satisfied. Then*

(a) $\sup_{0\le s\le t}\left|\mathbb{H}_{g,g}\left(X^1,X^2\right)_s^n-\mathbb{M}(1)_s^n-\left[X^1,X^2\right]_s\right|=O_p\left(n^{-\gamma}\right)$,

(b) $\sup_{0\le s\le t}\left|\mathbb{H}_{g',g'}\left(\mathfrak{U}^1,\mathfrak{U}^2\right)_s^n-\mathbb{M}(2)_s^n\right|=O_p\left(n^{-\gamma}\right)$,

(c) $\sup_{0\le s\le t}\left|\mathbb{H}_{g,g'}\left(X^1,\mathfrak{U}^2\right)_s^n-\mathbb{M}(3)_s^n\right|=O_p\left(n^{-\gamma}\right)$,

(d) $\sup_{0\le s\le t}\left|\mathbb{H}_{g',g}\left(\mathfrak{U}^1,X^2\right)_s^n-\mathbb{M}(4)_s^n\right|=O_p\left(n^{-\gamma}\right)$

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

Proof. Let $(V,u)\in\left\{(X^1,g),(\mathfrak{U}^1,g')\right\}$ and $(W,v)\in\left\{(X^2,g),(\mathfrak{U}^2,g')\right\}$. Since integration by parts yields $\bar V_u\left(\widehat{\mathcal{I}}\right)^i\bar W_v\left(\widehat{\mathcal{J}}\right)^j=\bar L_{u,v}(V,W)^{ij}+[\bar V_u(\widehat{\mathcal{I}})^i,\bar W_v(\widehat{\mathcal{J}})^j]$, we have

$$\mathbb{H}_{u,v}(V,W)_s^n=\mathbb{M}_{u,v}(V,W)_s^n+\frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j\ge1,\bar R^\vee(i,j)\le s}[\bar V_u(\widehat{\mathcal{I}})^i,\bar W_v(\widehat{\mathcal{J}})^j]_s\bar K_s^{ij}. \tag{A.3}$$

Since

$$\begin{aligned}[\bar V_u(\widehat{\mathcal{I}})^i,\bar W_v(\widehat{\mathcal{J}})^j]_s&=\sum_{p,q=0}^{k_n-1}u_p^n v_q^n\left(\widehat{I}_-^{i+p}\widehat{J}_-^{j+q}\right)\bullet[V,W]_s\\&=\sum_{p=i}^{i+k_n-1}\sum_{q=i}^{j+k_n-1}u_{p-i}^n v_{q-j}^n(\widehat{I}_-^p\widehat{J}_-^q)\bullet[V,W]_s,\end{aligned}$$

the second term of the right hand of (A.3) is equal to

$$\mathbb{B}_s:=\frac{1}{(\psi_{HY}k_n)^2}\sum_{\substack{p,q=1\\ \widehat{S}^{p+1}\vee\widehat{T}^{q+1}\le s}}^{\infty}\left(\sum_{i=(p-k_n+1)\vee1}^{p}\sum_{j=(q-k_n+1)\vee1}^{q}u_{p-i}^n v_{q-j}^n\bar K_s^{ij}\right)(\widehat{I}_-^p\widehat{J}_-^q)\bullet[V,W]_s.$$

Since $\bar K_s^{ij}=1$ on $\{\widehat{I}^p(s)\cap\widehat{J}^q(s)\neq\emptyset\}$, we have

$$\frac{1}{k_n^2}\sum_{i=(p-k_n+1)\vee1}^{p}\sum_{j=(q-k_n+1)\vee1}^{q}u_{p-i}^n v_{q-j}^n\bar K_s^{ij}=\left(\int_0^1u(x)\mathrm{d}x\right)\left(\int_0^1v(x)\mathrm{d}x\right)+O\left(k_n^{-1}\right)$$

on $\{\widehat{I}^p(s)\cap\widehat{J}^q(s)\neq\emptyset\}$ uniformly in $p,q\ge k_n$. Since $(\widehat{I}_-^p\widehat{J}_-^q)\bullet[V,W]=1_{\{\widehat{I}^p(t)\cap\widehat{J}^q(t)\neq\emptyset\}}(\widehat{I}_-^p\widehat{J}_-^q)\bullet[V,W]$ and $\int_0^1g'(x)\mathrm{d}x=g(1)-g(0)=0$, (A.3) implies that (b)–(d) holds true. Moreover, by using [C2], (A.1) and the fact that $\psi_{HY}=\int_0^1g(x)\mathrm{d}x$, (a) also follows from (A.3). ∎

We also remark the following result:

LEMMA A.1.3. *Let V,W be two semimartingales and $u,v\in\{g,g'\}$. Then, $\bar K_t^{ij}\bar L_{u,v}(V,W)_t=\bar K_-^{ij}\bullet\bar L_{u,v}(V,W)_t$ for any $i,j\in\mathbb{Z}_+$ and any $t\in\mathbb{R}_+$.*

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Proof. By integration by parts we have

$$\bar{K}_t^{ij}\bar{L}_{u,v}(V,W)_t = \bar{K}_-^{ij}\bullet\bar{L}_{u,v}(V,W)_t + \bar{L}_{u,v}(V,W)_-\bullet\bar{K}_t^{ij} + [\bar{K}^{ij},\bar{L}_{u,v}(V,W)]_t,$$

hence it is sufficient to show that $\bar{L}_{u,v}(V,W)_-\bullet\bar{K}_t^{ij} = [\bar{K}^{ij},\bar{L}_{u,v}(V,W)]_t = 0$. $\bar{K}_t^{ij}$ is a step function starting from 0 at $t=0$ and jumps to $+1$ at $t = R^\vee(i,j)$ when $\bar{I}^i\cap\bar{J}^j\neq\emptyset$, where $R^\vee(i,j) = \widehat{S}^i\vee\widehat{T}^j$, for $\bar{K}_t^{ij}$ can be rewritten as $\bar{K}_t^{ij} = \bar{K}^{ij}1_{(R^\vee(i,j),\infty)}(t)$. So, $\bar{L}_{u,v}(V,W)_-\bullet\bar{K}_t^{ij} = \bar{L}_{u,v}(V,W)_{R^\vee(i,j)\wedge t-}\bar{K}_t^{ij}$ and $[\bar{K}^{ij},\bar{L}_{u,v}(V,W)]_t = \bar{K}_t^{ij}\Delta\bar{L}_{u,v}(V,W)_{R^\vee(i,j)\wedge t}$. However, $\bar{L}_{u,v}(V,W)_t = 0$ for $t\leq R^\vee(i,j)$ by its definition. This implies that $\bar{L}_{u,v}(V,W)$ does not jump at $R^\vee(i,j)$. ■

Proof of Theorem 2.1. By a localization procedure, we may assume that [SC1]–[SC2], [$\mathrm{SN}_2^\flat$] and (A.1) hold true.

According to Lemma A.1.1–A.1.2, it suffices to prove $\mathbb{M}(k)^n \xrightarrow{ucp} 0$ as $n\to\infty$ for $k=1,2,3,4$. Moreover, in the light of Lemma A.1.3 it suffices to show that

$$\sup_{0\leq s\leq t}\left|\sum_{i,j\geq 1,\bar{R}^\vee(i,j)\leq s}\bar{K}_-^{ij}\bullet\left\{\bar{V}_u(\widehat{\mathcal{I}})_-^i\bullet\bar{W}_v(\widehat{\mathcal{J}})^j\right\}_s\right| = o_p(k_n^2), \tag{A.4}$$

$$\sup_{0\leq s\leq t}\left|\sum_{i,j\geq 1,\bar{R}^\vee(i,j)\leq s}\bar{K}_-^{ij}\bullet\left\{\bar{W}_v(\widehat{\mathcal{J}})_-^j\bullet\bar{V}_u(\widehat{\mathcal{I}})^i\right\}_s\right| = o_p(k_n^2) \tag{A.5}$$

as $n\to\infty$ for any $(V,u)\in\{(M^1,g),(\mathfrak{E}^1,g'),(\mathfrak{M}^1,g'),(A^1,g),(\mathfrak{A}^1,g')\}$, $(W,v)\in\{(M^2,g),(\mathfrak{E}^2,g'),(\mathfrak{M}^2,g'),(A^2,g),(\mathfrak{A}^2,g')\}$ and $t>0$.

Consider (A.4). First we assume that $(W,v)\in\{(M^2,g),(\mathfrak{M}^2,g'),(\mathfrak{E}^2,g')\}$. We decompose the target quantity as

$$\sum_{i,j\geq 1,\bar{R}^\vee(i,j)\leq s}\bar{K}_-^{ij}\bullet\left\{\bar{V}_u(\widehat{\mathcal{I}})_-^i\bullet\bar{W}_v(\widehat{\mathcal{J}})^j\right\}_s$$
$$=\left\{\sum_{i,j=1}^{\infty} - \sum_{i,j\geq 1,\bar{R}^\vee(i,j)>s}\right\}\bar{K}_-^{ij}\bullet\left\{\bar{V}_u(\widehat{\mathcal{I}})_-^i\bullet\bar{W}_v(\widehat{\mathcal{J}})^j\right\}_s =: \mathbb{B}_s^1+\mathbb{B}_s^2.$$

First we prove $\sup_{0\leq s\leq t}|\mathbb{B}_s^1| = o_p(k_n^2)$. Suppose that $W\in\{(M^2,g),(\mathfrak{M}^2,g')\}$. Then, W is a locally square-integrable martingale with the predictable quadratic variation $[W]$. Therefore, $\mathbb{B}^1$ is also a locally square-integrable martingale and thus it suffices to prove $\langle\mathbb{B}^1\rangle_t = o_p(k_n^4)$. Since $\langle\bar{W}_v(\widehat{\mathcal{J}})^j,\bar{W}_v(\widehat{\mathcal{J}})^{j'}\rangle_t = 0$ if $|j-j'|>k_n$, we have

$$\langle\mathbb{B}^1\rangle_s = \sum_{i,j,i',j':|j-j'|\leq k_n}\bar{K}_-^{ij}\bar{K}_-^{i'j'}\bar{V}_u(\widehat{\mathcal{I}})_-^i\bar{V}_u(\widehat{\mathcal{I}})_-^{i'}\bullet\langle\bar{W}_v(\widehat{\mathcal{J}})^j,\bar{W}_v(\widehat{\mathcal{J}})^{j'}\rangle_s.$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Therefore, the Kunita–Watanabe inequality, Proposition A.1.1, [SC1]–[SC2], [$\mathrm{SN}_2^\flat$], and (A.1) yield

$$E\left[\langle \mathbb{B}^1\rangle_t\right] \leq \sum_{j,j':|j-j'|\leq k_n} E\left[\sup_{0\leq s\leq t}\left(\sum_i \bar{K}_s^{ij}\bar{V}_u(\widehat{\mathcal{I}})_s^i\right)^2 \langle \bar{W}_v(\widehat{\mathcal{J}})^j\rangle_t\right]$$
$$\lesssim k_n^2\cdot k_n\bar{r}_n \sum_{i,j:|i-j|\leq k_n} E\left[\sup_{0\leq s\leq t}\left|\bar{V}_u(\widehat{\mathcal{I}})_s^i\right|^2\right] \lesssim k_n^4\cdot k_n\bar{r}_n$$

and thus $\langle \mathbb{B}^1\rangle_t = o_p(k_n^4)$ because $\xi' > 1/2$. On the other hand, if $W = \mathfrak{E}^2$, then for each $\omega^{(0)} \in \Omega^{(0)}$ $W(\omega^{(0)},\cdot)$ is a locally square-integrable martingale with respect to the filtration $\mathbf{F}^{(1)}$ under the probability measure $Q(\omega^{(0)}, \mathrm{d}\omega^{(1)})$ with the predictable quadratic variation $k_n^{-2}\sum_{j:\widehat{T}^j(\omega^{(0)})\leq\cdot}\Psi^{22}_{\widehat{T}^j(\omega^{(0)})}(\omega^{(0)})$. Therefore, we can adopt an argument similar to the above and prove $\langle \mathbb{B}^1\rangle_t = o_p(k_n^4)$.

Next we prove $\sup_{0\leq s\leq t}|\mathbb{B}_s^2| = o_p(k_n^2)$. Since $\bar{K}_s^{ij} = 0$ if $\widehat{S}^i\vee\widehat{T}^j \geq s$ or $|i-j|>k_n$, the Schwarz inequality yields

$$\sup_{0\leq s\leq t}|\mathbb{B}_s^2| \lesssim k_n\sqrt{\sum_{i,j}\left|\bar{K}_-^{ij}\bullet\left\{\bar{V}_u(\widehat{\mathcal{I}})_-^i\bullet\bar{W}_v(\widehat{\mathcal{J}})^j\right\}_s\right|^2}.$$

Moreover, an argument similar to the above implies that $\sup_{0\leq s\leq t}\sum_{i,j}\left|\bar{K}_-^{ij}\bullet\left\{\bar{V}_u(\widehat{\mathcal{I}})_-^i\bullet\bar{W}_v(\widehat{\mathcal{J}})^j\right\}_s\right|^2 = O_p(k_n^2\cdot k_n\bar{r}_n)$, hence we obtain $\sup_{0\leq s\leq t}|\mathbb{B}_s^2| = O_p(k_n^2\cdot\sqrt{k_n\bar{r}_n}) = o_p(k_n^2)$ because $\xi' > 1/2$. Consequently, we have shown (A.4) in the case that $(W,v) \in \{(M^2,g),(\mathfrak{M}^2,g'),(\mathfrak{E}^2,g')\}$.

Next we assume that $(W,v) \in \{(A^2,g),(\mathfrak{A}^2,g')\}$. Then, [SC1]–[SC2], [$\mathrm{SN}_2^\flat$], (A.1), and the Schwarz inequality yield

$$E\left[\sup_{0\leq s\leq t}\left|\sum_{i,j\geq 1,\bar{R}^\vee(i,j)\leq s}\bar{K}_-^{ij}\bullet\left\{\bar{V}_u(\widehat{\mathcal{I}})_-^i\bullet\bar{W}_v(\widehat{\mathcal{J}})^j\right\}_s\right|\right]$$
$$\lesssim k_n\left\{E\left[\sum_i\sup_{0\leq s\leq t}\left|\bar{V}_u(\widehat{\mathcal{I}})_s^i\right|^2\right]\right\}^{1/2}\left\{E\left[\sum_j\left(\sum_{q=0}^{k_n-1}|\widehat{J}^{j+q}(t)|\right)^2\right]\right\}^{1/2}$$
$$\lesssim k_n\cdot\sqrt{k_n}\cdot\sqrt{k_n^2\bar{r}_n} = k_n^2\sqrt{k_n\bar{r}_n},$$

hence we obtain (A.4) because $\xi' > 1/2$.

By symmetry we can also prove (A.5), and thus we complete the proof. ∎

A.2. Proof of Theorem 2.2

First, we note that the condition [C2] is implied by [A3] and [A5]. Moreover, [C1] holds true under the assumptions of Theorem 2.2 (see Remark 2.7(i)). In addition, we may assume (A.1) by the same argument as in the previous section. Consequently, the lemmas proved in the previous section can be used in this section.

Then, in the light of Lemma A.1.1–A.1.2 we focus on the process $\mathbb{M}_{u,v}(V,W)^n$. We replace it by a more tractable one. Define the process $\mathbf{M}_{u,v}(V,W)^n$ by

$$\mathbf{M}_{u,v}(V,W)^n_t = \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j=1}^{\infty}\bar{K}^{ij}_{-}\bullet\bar{L}_{u,v}(V,W)^{ij}_t.$$

LEMMA A.2.1. *Under the assumptions of Theorem 2.2, $n^{1/4}\{\mathbb{M}_{u,v}(V,W)^n - \mathbf{M}_{u,v}(V,W)^n\} \xrightarrow{ucp} 0$ as $n\to\infty$ for any $V\in\{X^1,\mathfrak{U}^1\}$, $W\in\{X^2,\mathfrak{U}^2\}$ and any $u,v\in\{g,g'\}$.*

Proof. Fix a $t>0$. Since $\bar{K}^{ij}_s = 0$ if $\widehat{S}^i\vee\widehat{T}^j\geq s$ or $|i-j|>k_n$, Lemma A.1.3 and the Hölder inequality yield

$$\sup_{0\leq s\leq t}\left|\mathbb{M}_{u,v}(V,W)^n_s - \mathbf{M}_{u,v}(V,W)^n_s\right| \lesssim k_n^{-2}\cdot k_n^{3/2}\left\{\sum_{i,j:|i-j|\leq k_n}\sup_{0\leq s\leq t}\left|\bar{L}_{u,v}(V,W)^{ij}_s\right|^4\right\}^{1/4}.$$

By [A3], [A5], [N$_8$], and (A.1) we obtain $\sum_{i,j:|i-j|\leq k_n}\sup_{0\leq s\leq t}\left|\bar{L}_{u,v}(V,W)^{ij}_s\right|^4 = O_p(k_n^2(k_n\bar{r}_n)^3)$, hence

$$\sup_{0\leq s\leq t}\left|\mathbb{M}_{u,v}(V,W)^n_s - \mathbf{M}_{u,v}(V,W)^n_s\right| = O_p((k_n\bar{r}_n)^{3/4}) = O_p\left(n^{-\frac{3}{4}\left(\xi'-\frac{1}{2}\right)}\right).$$

Since $\xi' > \frac{9}{10}$ by [A4], the proof is completed. ∎

Since $\mathbf{M}_{u,v}(V,W)^n$ has a similar structure to the estimation error process of the Hayashi–Yoshida estimator (see equation (3.2) of Hayashi and Yoshida, 2011), we can apply arguments mimicking those in Hayashi and Yoshida (2011) for the proof. For example, the following result can be obtained in such a manner:

LEMMA A.2.2. *Let $u,v\in\{g,g'\}$. Under the assumptions of Theorem 2.2, it holds that*

(a) $n^{1/4}\mathbf{M}_{u,v}(A,W)^n \xrightarrow{ucp} 0$ *as $n\to\infty$ for $A\in\{A^1,\mathfrak{A}^1\}$ and $W\in\{X^2,\mathfrak{U}^2\}$.*

(b) $n^{1/4}\mathbf{M}_{u,v}(V,A)^n \xrightarrow{ucp} 0$ *as $n\to\infty$ for $V\in\{X^1,\mathfrak{U}^1\}$ and $A\in\{A^2,\mathfrak{A}^2\}$.*

Proof. Since $\mathbf{M}_{u,v}(A,W)^n$ and $\mathbf{M}_{u,v}(V,A)^n$ have similar structures to the estimation error process of the Hayashi–Yoshida estimator, we can adopt an argument mimicking the proof of Lemma 13.1–13.2 in Hayashi and Yoshida (2011). This completes the proof. ∎

The above lemma tells us that we can replace $\sum_{k=1}^4\mathbf{M}(k)^n$ by $\widetilde{\mathbf{M}}^n := \sum_{k=1}^4\widetilde{\mathbf{M}}(k)^n$, where $\widetilde{\mathbf{M}}(1)^n = \mathbf{M}_{g,g}(M^1,M^2)^n$, $\widetilde{\mathbf{M}}(2)^n = \mathbf{M}_{g',g'}(\mathfrak{E}^1+\mathfrak{M}^1,\mathfrak{E}^2+\mathfrak{M}^2)^n$, $\widetilde{\mathbf{M}}(3)^n = \mathbf{M}_{g,g'}(M^1,\mathfrak{E}^2+\mathfrak{M}^2)^n$, and $\widetilde{\mathbf{M}}(4)^n = \mathbf{M}_{g',g}(\mathfrak{E}^1+\mathfrak{M}^1,M^2)^n$. Since $\widetilde{\mathbf{M}}^n$ is a locally square-integrable martingale, it is more tractable. In fact, we can apply a simplified martingale version of the stable central limit theorem from Jacod (1997) to it as follows:

Proof of Theorem 2.2.

(a) According to Lemma A.1.1–A.1.2 and A.2.1–A.2.2, it suffices to show that $n^{1/4}\widetilde{\mathbf{M}}^n \to^{d_s} \int_0^\cdot w_s\mathrm{d}\widetilde{W}_s$ as $n\to\infty$.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

We apply Theorem 2–2 from Jacod (1997). The above stable convergence is implied by the following three conditions:

$$\langle n^{1/4}\widetilde{\mathbf{M}}^n\rangle_t \to^p \int_0^t w_s^2 \mathrm{d}s, \tag{A.6}$$

$$\langle n^{1/4}\widetilde{\mathbf{M}}^n, N\rangle_t \to^p 0, \tag{A.7}$$

$$\sum_{s:0\le s\le t} |n^{1/4}\Delta\widetilde{\mathbf{M}}_s^n|^4 \to^p 0 \tag{A.8}$$

for any $t \in \mathbb{R}_+$ and any $\mathbf{F}$-square-integrable martingale N.

First we prove (A.6). The bi-linearity of predictable quadratic variations yields

$$\langle n^{1/4}\widetilde{\mathbf{M}}^n\rangle_t = \sqrt{n}\sum_{k,l=1}^{4} \langle \widetilde{\mathbf{M}}(k)^n, \widetilde{\mathbf{M}}(l)^n\rangle_t.$$

Consider $\langle \widetilde{\mathbf{M}}(1)^n, \widetilde{\mathbf{M}}(1)^n\rangle_t$. Since $\widetilde{\mathbf{M}}(1)^n$ has a similar structure to the estimation error process of the Hayashi–Yoshida estimator, we can adopt an argument mimicking the proof of Proposition 5.1 in Hayashi and Yoshida (2011). Consequently, we obtain

$$\begin{aligned}\langle &\widetilde{\mathbf{M}}(1)^n, \widetilde{\mathbf{M}}(1)^n\rangle_t \\ &= \frac{1}{(\psi_{HY}k_n)^4}\sum_{i,j,i',j'} (\bar{K}_-^{ij}\bar{K}_-^{i'j'})\bullet\Big\{\Big\langle \bar{M}_g^1(\widehat{\mathcal{I}})^i, \bar{M}_g^1(\widehat{\mathcal{I}})^{i'}\rangle_t\langle \bar{M}_g^2(\widehat{\mathcal{J}})^j, \bar{M}_g^2(\widehat{\mathcal{J}})^{j'}\Big\rangle_t \\ &\quad + \Big\langle \bar{M}_g^1(\widehat{\mathcal{I}})^i, \bar{M}_g^2(\widehat{\mathcal{J}})^{j'}\rangle_t\langle \bar{M}_g^1(\widehat{\mathcal{I}})^{i'}, \bar{M}_g^2(\widehat{\mathcal{J}})^{j}\Big\rangle_t\Big\} + o_p(n^{-1/2}).\end{aligned}$$

Moreover, noting that it can be shown that

$$\sup_{p,q:p,q\ge k_n}\left| k_n^{-2}\sum_{i=(p-k_n+1)\vee 1}^{p}\ \sum_{j=(q-k_n+1)\vee 1}^{q} g_{p-i}^n g_{q-j}^n \bar{K}^{ij} - \psi_{g,g}\left(\frac{q-p}{k_n}\right)\right| = O_p\left(n^{-1/2}\right)$$

due to Proposition A.1.1, a simple computation yields

$$\begin{aligned}\langle \widetilde{\mathbf{M}}(1)^n, \widetilde{\mathbf{M}}(1)^n\rangle_t &= \frac{1}{\psi_{HY}^4}\sum_{p,q=1}^{\infty}\psi_{g,g}\left(\frac{q-p}{k_n}\right)^2 \\ &\quad\times\Big\{\langle M^1\rangle(\widehat{I}^p)_t\langle M^2\rangle(\widehat{J}^q)_t + \langle M^1, M^2\rangle(\widehat{I}^p)_t\langle M^1, M^2\rangle(\widehat{J}^q)_t\Big\} \\ &\quad+ o_p\left(n^{-1/2}\right).\end{aligned}$$

Then, by using [A2]–[A4] and (9) it can be shown that

$$\begin{aligned}\langle &\widetilde{\mathbf{M}}(1)^n, \widetilde{\mathbf{M}}(1)^n\rangle_t \\ &= \frac{1}{\psi_{HY}^4}\sum_{p,q=1}^{\infty}\psi_{g,g}\left(\frac{q-p}{k_n}\right)^2\Big\{[X^1](\Gamma^p)_t[X^2](\Gamma^q)_t + [X^1, X^2](\Gamma^p)_t[X^1, X^2](\Gamma^q)_t\Big\} \\ &\quad + o_p(n^{-1/2})\end{aligned}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

$$= \frac{n^{-2}}{\psi_{HY}^4} \sum_{p=1}^{\infty} \left\{ \sum_q \psi_{g,g} \left(\frac{q-p}{k_n} \right)^2 \right\} \left\{ [X^1]'_{R^{p-1}} [X^2]'_{R^{p-1}} + \left([X^1, X^2]'_{R^{p-1}} \right)^2 \right\}$$
$$\times \left(G^n_{R^{p-1}} \right)^2 1_{\{R^{p-1} \le t\}} + o_p(n^{-1/2})$$
$$= n^{-1/2} \theta \kappa \int_0^t \left\{ [X^1]'_s [X^2]'_s + \left([X^1, X^2] \right)^2 \right\} G_s \mathrm{d}s + o_p(n^{-1/2}),$$

hence we obtain $\sqrt{n}\langle \widetilde{\mathbf{M}}(1)^n, \widetilde{\mathbf{M}}(1)^n \rangle_t \to^p \theta\kappa \int_0^t \left\{ [X^1]'_s [X^2]'_s + \left([X^1, X^2]\right)^2 \right\} G_s \mathrm{d}s$. In a similar manner we can also show that $\sqrt{n}\langle \widetilde{\mathbf{M}}(1)^n, \widetilde{\mathbf{M}}(k)^n \rangle_t \to^p 0$ $(k = 2, 3, 4)$, $\sqrt{n}\langle \widetilde{\mathbf{M}}(2)^n, \widetilde{\mathbf{M}}(l)^n \rangle_t \to^p 0$ $(l = 3, 4)$ and

$$\sqrt{n}\langle \widetilde{\mathbf{M}}(2)^n, \widetilde{\mathbf{M}}(2)^n \rangle_t \to^p \theta^{-3} \widetilde{\kappa} \int_0^t \left\{ \Psi_s^{11} \Psi_s^{22} + \left(\Psi_s^{12} \chi_s \right)^2 \right\} \frac{1}{G_s} \mathrm{d}s,$$
$$\sqrt{n}\langle \widetilde{\mathbf{M}}(3)^n, \widetilde{\mathbf{M}}(3)^n \rangle_t \to^p \theta^{-1} \overline{\kappa} \int_0^t [X^1]'_s \Psi_s^{11} \mathrm{d}s,$$
$$\sqrt{n}\langle \widetilde{\mathbf{M}}(4)^n, \widetilde{\mathbf{M}}(4)^n \rangle_t \to^p \theta^{-1} \overline{\kappa} \int_0^t \Psi_s^{22} [X^2]'_s \mathrm{d}s,$$
$$\sqrt{n}\langle \widetilde{\mathbf{M}}(3)^n, \widetilde{\mathbf{M}}(4)^n \rangle_t \to^p \theta^{-1} \overline{\kappa} \int_0^t [X^1, X^2]'_s \Psi_s^{12} \chi_s \mathrm{d}s$$

as $n \to \infty$. Consequently, (A.6) holds true.

Next we prove (A.7). Let $\mathcal{N}$ be the set of all $\mathbf{F}$-square-integrable martingales satisfying (A.7). Then, $\mathcal{N}$ is a closed subspace of the Hilbert space $\mathcal{M}_2$ of all $\mathbf{F}$-square-integrable martingales due to (A.6) and the Kunita–Watanabe inequality, hence it suffices to show that $\mathcal{N}$ is total in $\mathcal{M}_2$.

Since (11) is satisfied for $f = [X^1]', [X^2]'$ and $[X^1, X^2]'$ by [A3], Lemma A.2.2 yields $M^1, M^2 \in \mathcal{N}$. On the other hand, if N is an $\mathbf{F}^{(0)}$-square integrable martingale orthogonal to (M^1, M^2), then obviously $\langle \widetilde{\mathbf{M}}^n, N \rangle_t = 0$, so that $N \in \mathcal{N}$. Consequently, $\mathcal{N}$ includes the set $\mathcal{N}^0$ of all $\mathbf{F}^{(0)}$-square integrable martingales. Moreover, it can also be shown that $\mathcal{N}$ includes the set $\mathcal{N}^1$ of all $\mathbf{F}^{(1)}$-square integrable martingales by an argument similar to the Step (5) of the proof of Lemma 5.7 in Jacod et al. (2009). Consequently, $\mathcal{N}$ is total in $\mathcal{M}_2$ because so is $\mathcal{N}^0 \cup \mathcal{N}^1$.

Finally we prove (A.8). Since equation I-4.36 in Jacod and Shiryaev (2003) yields

$$\Delta \widetilde{\mathbf{M}}^n_s = \frac{1}{(\psi_{HY} k_n)^2} \sum_{i,j} \left\{ \bar{K}_s^{ij} \bar{\mathsf{X}}^1(\widehat{\mathcal{I}})^i_s \Delta \bar{\mathfrak{E}}^2_{g'}(\widehat{\mathcal{J}})^j_s + \bar{K}_s^{ij} \bar{\mathsf{X}}^2(\widehat{\mathcal{J}})^j_s \Delta \bar{\mathfrak{E}}^1_{g'}(\widehat{\mathcal{I}})^i_s \right\},$$

where $\bar{\mathsf{X}}^1(\widehat{\mathcal{I}})^i_s = \bar{X}^1_g(\widehat{\mathcal{I}})^i_s + \bar{\mathfrak{E}}^1_{g'}(\widehat{\mathcal{I}})^i_s$ and $\bar{\mathsf{X}}^2(\widehat{\mathcal{J}})^j_s = \bar{X}^2_g(\widehat{\mathcal{J}})^j_s + \bar{\mathfrak{E}}^2_{g'}(\widehat{\mathcal{J}})^j_s$, it suffices to prove that

$$\frac{n}{k_n^8} \sum_{0 \le s \le t} \left| \sum_{i,j} \bar{K}_s^{ij} \bar{\mathsf{X}}^1(\widehat{\mathcal{I}})^i_s \Delta \bar{\mathfrak{E}}^2_{g'}(\widehat{\mathcal{J}})^j_s \right|^4 \to^p 0,$$
$$\frac{n}{k_n^8} \sum_{0 \le s \le t} \left| \sum_{i,j} \bar{K}_s^{ij} \bar{\mathsf{X}}^2(\widehat{\mathcal{J}})^j_s \Delta \bar{\mathfrak{E}}^1_{g'}(\widehat{\mathcal{I}})^i_s \right|^4 \to^p 0 \qquad \textbf{(A.9)}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

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

$$\sum_{i,j} \bar{K}_s^{ij} \mathsf{X}^1(\widehat{\mathcal{I}})_s^i \Delta \bar{\mathfrak{E}}_{g'}^2(\widehat{\mathcal{J}})_s^j = -\frac{1}{k_n} \sum_{q=1}^{\infty} \zeta_{\widehat{T}^q}^2 1_{\{\widehat{T}^q = s\}} \sum_{i=1}^{\infty} \sum_{j=(q-k_n+1)\vee 1}^{q} (g')_{q-j}^n \bar{K}_{\widehat{T}^q}^{ij} \bar{\mathsf{X}}^1(\widehat{\mathcal{I}})_{\widehat{T}^q}^i,$$

we have

$$\begin{aligned}
&\frac{n}{k_n^8} \sum_{0 \le s \le t} \left| \sum_{i,j} \bar{K}_s^{ij} \bar{\mathsf{X}}^1(\widehat{\mathcal{I}})_s^i \Delta \bar{\mathfrak{E}}_{g'}^2(\widehat{\mathcal{J}})_s^j \right|^4 \\
&\quad \le \frac{n}{k_n^8} \cdot \frac{1}{k_n^4} \sum_{q=1}^{\infty} \left(\zeta_{\widehat{T}^q}^2 \right)^4 \left| \sum_{i=1}^{\infty} \sum_{j=(q-k_n+1)\vee 1}^{q} (g')_{q-j}^n \bar{K}_{\widehat{T}^q}^{ij} \bar{\mathsf{X}}^1(\widehat{\mathcal{I}})_{\widehat{T}^q}^i \right|^4 1_{\{\widehat{T}^q \le t\}} \\
&\quad = O_p(n k_n^{-8} \cdot k_n^{-4} \cdot n k_n^8 (k_n \bar{r}_n)) = o_p(1)
\end{aligned}$$

by [C1], [C2], [N_8], and [A4]. Consequently, we have proved the first equation of (A.9). By symmetry we also obtain the second equation of (A.9), hence we complete the proof.

(b) Similar to the proof of (a).

■

A.3. Proof of Theorem 3.1

Exactly as in Section A.1, we may assume that [SC1]–[SC2], [$\mathrm{SN}_r^\flat$], and (A.1) are satisfied. Similarly, we may strengthen the condition [F] as follows:

[SF] We have [F] and there is a positive constant B such that

$$B^{-1} < \inf_{k \in \mathbb{N}} |\gamma_k^l| \le \sup_{k \in \mathbb{N}} |\gamma_k^l| < B \tag{A.10}$$

for each $l = 1, 2$.

Next we introduce the following strengthened version of the condition [T]:

[ST] For each $l = 1, 2$ we have $\varrho_n^l(t) = \alpha_n^l(t)\rho_n$, where $(\rho_n)_{n \in \mathbb{N}}$ is the same one as in [T] and $(\alpha_n^l(t))_{n \in \mathbb{N}}$ is a sequence of (not necessarily adapted) positive-valued stochastic processes such that there exists a positive constant K_0 satisfying

$$\frac{1}{K_0} < \inf_{t \in \mathbb{R}_+} \alpha_n^l(t) \le \sup_{t \in \mathbb{R}_+} \alpha_n^l(t) < K_0, \qquad n = 1, 2, \ldots.$$

LEMMA A.3.1. *Let (c_n) be a sequence of positive numbers. If we have*

$$c_n^{-1}\{\widehat{PTHY}(\mathsf{Z}^1, \mathsf{Z}^2)^n - \widehat{PHY}(\mathsf{X}^1, \mathsf{X}^2)^n\} \xrightarrow{ucp} 0 \quad \text{as} \quad n \to \infty \tag{A.11}$$

under the condition [ST], *then we have also (A.11) under the condition* [T].

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Proof. Let $t > 0$ and $k \in \mathbb{N}$. Suppose that [T] holds. Then, for an arbitrary $\varepsilon > 0$, there exists a positive number K such that

$$\sup_{n\in\mathbb{N}} P\left(\sup_{0\le s<R_k^l} \alpha_n^l(s) \ge K\right) < \varepsilon \quad \text{and} \quad \sup_{n\in\mathbb{N}} P\left(\sup_{0\le s<R_k^l} [1/\alpha_n^l(s)] \ge K\right) < \varepsilon, \quad l=1,2.$$

Hence for any $\eta > 0$ we have

$$P\left(\Psi^n(t) > \eta\right) \le P(R_k^1 \wedge R_k^2 \le t)$$
$$+4\varepsilon + P\left(\Psi^n(t\wedge R_k^1 \wedge R_k^2) > \eta, \max_{l\in\{1,2\}}\left[\sup_{0\le s<R_k^l} \alpha_n^l(s) \vee \sup_{0\le s<R_k^l} \frac{1}{\alpha_n^l(s)}\right] < K\right),$$

where $\Psi^n(t) := \sup_{0\le s\le t} c_n^{-1}|\widehat{PTHY}(\mathsf{Z}^1,\mathsf{Z}^2)_s^n - \widehat{PHY}(\mathsf{X}^1,\mathsf{X}^2)_s^n|$. Therefore, by the assumption we obtain

$$\limsup_{n\to\infty} P\left(\Psi^n(t) > \eta\right) \le P(R_k^1 \wedge R_k^2 \le t) + 4\varepsilon.$$

Since ε is arbitrary, we can replace ε in the above inequality with 0. Finally, with k tending to 0, we obtain the desired result. ■

Now we introduce some notation and prove some lemmas which we will also use later. Set

$$\widetilde{V}(\widehat{\mathcal{I}})_t^i = -\sqrt{n}\sum_{p=0}^{k_n-1} \Delta(g)_p^n V(\check{I}^{i+p})_t, \qquad \widetilde{V}(\widehat{\mathcal{J}})_t^j = -\sqrt{n}\sum_{q=0}^{k_n-1} \Delta(g)_q^n V(\check{J}^{j+q})_t$$

for each $t \in \mathbb{R}_+$ and $i, j \in \mathbb{Z}_+$. The following lemma is an analog to Lemma 2.1 of Mancini and Gobbi (2012):

LEMMA A.3.2. *Suppose that* [SC1]–[SC2] *and (A.1) are satisfied. Then there exists a positive constant L such that*

$$\limsup_{n\to\infty}\sup_{i\in\mathbb{N}} \frac{|\bar{X}^1(\widehat{\mathcal{I}})_t^i|}{\sqrt{2k_n\bar{r}_n\log\frac{1}{\bar{r}_n}}} \le L, \qquad \limsup_{n\to\infty}\sup_{i\in\mathbb{N}} \frac{|\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})_t^i|}{\sqrt{2k_n\bar{r}_n\log\frac{1}{\bar{r}_n}}} \le L, \tag{A.12}$$

$$\limsup_{n\to\infty}\sup_{i\in\mathbb{N}} \frac{|\bar{X}^2(\widehat{\mathcal{J}})_t^i|}{\sqrt{2k_n\bar{r}_n\log\frac{1}{\bar{r}_n}}} \le L, \qquad \limsup_{n\to\infty}\sup_{i\in\mathbb{N}} \frac{|\widetilde{\underline{X}^2}(\widehat{\mathcal{J}})_t^i|}{\sqrt{2k_n\bar{r}_n\log\frac{1}{\bar{r}_n}}} \le L \tag{A.13}$$

a.s. for any $t > 0$.

Proof. First, note that by [SC2] there exists a constant $c_0 > 0$ such that $\sup_{0\le s\le t}|[X^1]_s'| + \sup_{0\le s\le t}|[\underline{X}^1]_s'| \le c_0$. We also remark that (A.1) implies that $|\widehat{I}^p|, |\check{I}^p| \le 2\bar{r}_n$.

Next, combining a representation of a continuous local martingale with Brownian motion and Lévy's theorem on the uniform modulus of continuity of Brownian motion, we obtain a.s.

$$\limsup_{\delta\to+0}\sup_{\substack{s,u\in[0,t]\\|s-u|\le\delta}} \frac{|M_s^1 - M_u^1|}{\sqrt{2\delta\log\frac{1}{\delta}}} \le \sup_{0\le s\le t}|[X^1]_s'|,$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

$$\limsup_{\delta\to+0}\sup_{\substack{s,u\in[0,t]\\|s-u|\leq\delta}}\frac{|\underline{M}^1_s-\underline{M}^1_u|}{\sqrt{2\delta\log\frac{1}{\delta}}}\leq\sup_{0\leq s\leq t}|[\underline{X}^1]'_s|.$$

Since $\bar{X}^1(\widehat{\mathcal{I}})^i_t=-\sum_{p=0}^{k_n-1}\Delta(g)^n_p(X^1_{\widehat{S}^{i+p}\wedge t}-X^1_{\widehat{S}^i\wedge t})$, $|\Delta(g)^n_p|\leq\frac{1}{k_n}\|g\|_\infty$ and $|A^1_{\widehat{S}^{i+p}\wedge t}-A^1_{\widehat{S}^i\wedge t}|\lesssim\bar{r}_n$, we obtain the first inequality in (A.12). Moreover, by the Freedman inequality (Thm. 1.6 of Freedman, 1975) we have

$$\begin{aligned}&P\left(|\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t|>\sqrt{12c_0\|g\|_\infty k_n\bar{r}_n\log n}\right)\\&\quad\leq 2\exp\left(-\frac{12c_0\|g\|_\infty k_n\bar{r}_n\log n}{2(2c_0\|g\|_\infty k_n\bar{r}_n+\sqrt{12k_n}4c_0\|g\|_\infty\bar{r}_n\log n)}\right)\lesssim n^{-2}.\end{aligned}$$

Therefore, [SC1] and the Borel–Cantelli lemma as well as $|\underline{A}^1(\check{I}^p)_t|\lesssim\bar{r}_n$ imply that the second inequality in (A.12).

By symmetry we also obtain (A.13). ∎

We can strengthen Lemma A.3.2 by a localization if we assume that (A.1) and [SC2] hold, so that in the remainder of this section we always assume that we have a positive constant K and a positive integer n_0 such that

$$\sup_{i\in\mathbb{N}}\frac{|\bar{X}^1(\widehat{\mathcal{I}})^i_t(\omega)|+|\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t(\omega)|}{\sqrt{2k_n\bar{r}_n\log n}}+\sup_{j\in\mathbb{N}}\frac{|\bar{X}^2(\widehat{\mathcal{J}})^j_t(\omega)|+|\widetilde{\underline{X}^2}(\widehat{\mathcal{J}})^j_t(\omega)|}{\sqrt{2k_n\bar{r}_n\log n}}\leq K \tag{A.14}$$

for all $t>0$ and $\omega\in\Omega$ if $n\geq n_0$. Moreover, we only consider sufficiently large n such that $n\geq n_0$.

Next, set

$$\bar{\zeta}^1(\widehat{\mathcal{I}})^i=\sum_{p=0}^{k_n-1}\Delta(g)^n_p\zeta^1_{\widehat{S}^{i+p}},\qquad\bar{\zeta}^2(\widehat{\mathcal{J}})^j=\sum_{q=0}^{k_n-1}\Delta(g)^n_q\zeta^2_{\widehat{T}^{j+q}}$$

for each $i,j\in\mathbb{Z}_+$. We denote by E_0 a conditional expectation given $\mathcal{F}^{(0)}$, i.e. $E_0[\cdot]:=E[\cdot|\mathcal{F}^{(0)}]$.

LEMMA A.3.3. *Suppose* $[\mathrm{SN}^\flat_r]$ *hold for some* $r\in[2,\infty)$. *Then there exists a some positive constant* K_r *independent of n such that*

$$E_0[|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^r]\leq K_rk_n^{-r/2},\qquad E_0[|\bar{\zeta}^2(\widehat{\mathcal{J}})^j|^r]\leq K_rk_n^{-r/2} \tag{A.15}$$

for all $i,j\in\mathbb{Z}_+$.

Proof. The Burkholder–Davis–Gundy inequality, Jensen's inequality and the Lipschitz continuity of g yield

$$\begin{aligned}E_0[|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^r]&\lesssim E_0\left[\left\{\sum_{p=0}^{k_n-1}|\Delta(g)^n_p\zeta^1_{\widehat{S}^{i+p}}|^2\right\}^{r/2}\right]\\&\leq k_n^{r/2-1}\sum_{p=0}^{k_n-1}E_0\left[|\Delta(g)^n_p\zeta^1_{\widehat{S}^{i+p}}|^r\right]\lesssim k_n^{-r/2},\end{aligned}$$

hence we obtain the first inequality of (A.15). By symmetry we also obtain the second one. ∎

Recall that $\bar{I}^i = [\widehat{S}^i, \widehat{S}^{i+k_n})$, $\bar{J}^j = [\widehat{T}^j, \widehat{T}^{j+k_n})$, and $\bar{R}^\vee(i,j) = \widehat{S}^{i+k_n} \vee \widehat{T}^{j+k_n}$.

LEMMA A.3.4. *Let c be a positive number. Suppose* [SC2] *and* [SF] *hold. Suppose also* [SN$^\flat_r$] *holds for some* $r \in (2,\infty)$. *Then for all* $t > 0$ *we have*

$$\sum_{i=1}^{\infty} P\left(|\overline{\zeta}^1(\widehat{\mathcal{I}})^i| \geq c, N^1(\bar{I}^i)_t \neq 0, \widehat{S}^{i+k_n} \leq t\right) \to 0,$$
$$\sum_{j=1}^{\infty} P\left(|\overline{\zeta}^2(\widehat{\mathcal{J}})^j| \geq c, N^2(\bar{J}^j)_t \neq 0, \widehat{T}^{j+k_n} \leq t\right) \to 0 \tag{A.16}$$

as $n \to \infty$.

Proof. Lemma A.3.3 yields $E_0\big[|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r 1_{\{N^1(\bar{I}^i)_t \neq 0\}}\big] \lesssim k_n^{-r/2} 1_{\{N^1(\bar{I}^i)_t \neq 0\}}$ uniformly in i. Since N^1 is a point process, we obtain

$$\sum_{i=1}^{\infty} P\left(|\overline{\zeta}^1(\widehat{\mathcal{I}})^i| \geq c, N^1(\bar{I}^i)_t \neq 0, \widehat{S}^{i+k_n} \leq t \big| \mathcal{F}^{(0)}\right)$$
$$\leq \frac{1}{c^r}\sum_{i=1}^{\infty} E_0[|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r 1_{\{N^1(\bar{I}^i)_t \neq 0\}}] \lesssim k_n^{1-r/2} N_t^1,$$

and thus we obtain the first equation of (A.16) since $r > 2$ and $k_n \to \infty$ as $n \to \infty$. By symmetry we also obtain the second equation of (A.16), and thus we complete the proof of lemma. ∎

Proof of Theorem 3.1. By a localization procedure, we may replace the conditions [F], [C1]–[C2], and [N$^\flat_r$] with [SF], [SC1]–[SC2] and [SN$^\flat_r$] respectively. Moreover, we can also replace the condition [T] with [ST] by Lemma A.3.1, while (5) can be replaced with (A.1) due to the same argument as in Section A.1.

We decompose the target quantity as

$$\begin{aligned}
&PTHY(\mathsf{Z}^1,\mathsf{Z}^2)^n_t - PHY(\mathsf{X}^1,\mathsf{X}^2)^n_t \\
&= \frac{1}{(\psi_{HY}k_n)^2}\Bigg(-\sum_{i,j:\bar{R}^\vee(i,j)\leq t} \overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i \overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j \bar{K}^{ij} 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 > \varrho_n^1[i]\}\cup\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 > \varrho_n^2[j]\}} \\
&\quad + \sum_{i,j:\bar{R}^\vee(i,j)\leq t}\left\{\bar{D}^1(\widehat{\mathcal{I}})^i_t \overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j + \overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i \bar{D}^2(\widehat{\mathcal{J}})^j_t + \bar{D}^1(\widehat{\mathcal{I}})^i_t \bar{D}^2(\widehat{\mathcal{J}})^j_t\right\} \\
&\quad \times \bar{K}^{ij} 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \leq \varrho_n^1[i], |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \leq \varrho_n^2[j]\}}\Bigg) \\
&=: \mathbb{I}_t + \mathbb{II}_t + \mathbb{III}_t + \mathbb{IV}_t,
\end{aligned}$$

where $D^l_t := \sum_{k=1}^{N^l_t} \gamma^l_k$ for each $l = 1, 2$.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

First consider $\mathbb{I}$. By the Schwarz inequality and (A.2), we have

$$\sup_{0\le s\le t}|\mathbb{I}_s| \le \frac{1}{(\psi_{HY}k_n)^2}\left\{\sum_{i,j:\bar{R}^\vee(i,j)\le t}|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2\bar{K}^{ij}1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]\}}\right\}^{1/2}$$
$$\times\left\{\sum_{i,j:\bar{R}^\vee(i,j)\le t}|\overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j|^2\bar{K}^{ij}1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j]\}}\right\}^{1/2}$$
$$\lesssim \frac{1}{k_n}\left\{\sum_{i:\widehat{S}^{i+k_n}\le t}|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]\}}\right\}^{1/2}$$
$$\times\left\{\sum_{j:\widehat{T}^{j+k_n}\le t}|\overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j|^2\, 1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j]\}}\right\}^{1/2}.$$

Consider $\sum_{i:\widehat{S}^{i+k_n}\le t}|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]\}}$. We decompose it as

$$\sum_{i:\widehat{S}^{i+k_n}\le t}|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]\}}$$
$$=\sum_{i:\widehat{S}^{i+k_n}\le t}|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2\left(1_{\{|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i],N^1(\bar{I}^i)_t=0\}}+1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i],N^1(\bar{I}^i)_t\neq 0\}}\right)$$
$$=: A_{1,t}+A_{2,t}.$$

On $\{|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i],\widehat{S}^{i+k_n}\le t\}$ we have

$$|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|\ge|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|-|\bar{X}^1(\widehat{\mathcal{I}})^i_t|-|\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t|>\sqrt{\rho_n}\left(\frac{1}{\sqrt{K_0}}-2K\sqrt{\frac{2k_n\bar{r}_n\log n}{\rho_n}}\right)$$

by [ST] and (A.14). Hence by (18) we have

$$A_{1,t}\le\sum_{i:\widehat{S}^{i+k_n}\le t}|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2>\rho_n/4K_0\}}$$
$$\le 2\left(\frac{4K_0}{\rho_n}\right)^{\frac{r}{2}}\sum_{i:\widehat{S}^{i+k_n}\le t}|\bar{X}^1(\widehat{\mathcal{I}})^i_t|^2|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r+2\left(\frac{4K_0}{\rho_n}\right)^{\frac{r-2}{2}}\sum_{i:\widehat{S}^{i+k_n}\le t}|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r,$$

and thus (18) and (A.14) imply that $A_{1,t}\lesssim(\rho_n)^{-\frac{r-2}{2}}\sum_{i:\widehat{S}^{i+k_n}\le t}|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r$. Hence Lemma A.3.3 and [SC1] yield

$$E[A_{1,t}]\lesssim nk_n^{-1}(k_n\rho_n)^{-\frac{r-2}{2}}. \tag{A.17}$$

On the other hand, since $|\bar{I}^i(t)|\le k_n\bar{r}_n\to 0$ and N^1 is a point process, pathwise for sufficiently large n there exists a some index $k(i)\in\mathbb{N}$ for each i such that $\bar{D}^1(\widehat{\mathcal{I}})^i_t=\gamma_k(i)N^1$

$(\bar{I}^i)_t$. Hence by (A.10) we have $|\bar{D}^1(\widehat{\mathcal{I}})^i_t| \geq B^{-1}$ on $\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \leq \varrho^1_n[i], N^1(\bar{I}^i)_t \neq 0, \widehat{S}^{i+k_n} \leq t\}$ for each i pathwise for sufficiently large n. Moreover, on $\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \leq \varrho^1_n[i], |\bar{D}^1(\widehat{\mathcal{I}})^i_t| \geq B^{-1}, \widehat{S}^{i+k_n} \leq t\}$ we have

$$
\begin{aligned}
|\overline{\zeta}^1(\widehat{\mathcal{I}})^i| &\geq |\bar{D}^1(\widehat{\mathcal{I}})^i_t| - |\mathsf{Z}^1(\widehat{\mathcal{I}})^i| - |\bar{X}^1(\widehat{\mathcal{I}})^i_t| - |\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t| \\
&\geq B^{-1} - \sqrt{\varrho^1_n[i]} - |\bar{X}^1(\widehat{\mathcal{I}})^i_t| - |\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t|,
\end{aligned}
$$

hence by (A.14) a.s. for sufficiently large n we have $A_{2,t} \leq \sum_{i:\widehat{S}^{i+k_n} \leq t} |\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\overline{\zeta}^1(\widehat{\mathcal{I}})^i| > 1/2B, N^1(\bar{I}^i)_t \neq 0\}}$. Therefore Lemma A.3.4 yields

$$
A_{2,t} = o_p\left(nk_n^{-1}(k_n\rho_n)^{-\frac{r-2}{2}}\right). \tag{A.18}
$$

By (A.17) and (A.18) we obtain $\sum_{i:\widehat{S}^{i+k_n} \leq t} |\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 > \varrho^1_n[i]\}} = O_p\big(nk_n^{-1} (k_n\rho_n)^{-\frac{r-2}{2}}\big)$, and by symmetry we also obtain $\sum_{i:\widehat{T}^{j+k_n} \leq t} |\overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j|^2 1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 > \varrho^2_n[j]\}} = O_p\left(nk_n^{-1}(k_n\rho_n)^{-\frac{r-2}{2}}\right)$. Consequently, by (7) we have $\sup_{0\leq s\leq t} |\mathbb{I}_s| = O_p\left(\left(n^{-1/2} \rho_n^{-1}\right)^{\frac{r-2}{2}}\right)$.

Next consider $\mathbb{III}$. Since $\bar{D}^1(\widehat{\mathcal{I}})^i_t = 0$ on $\{N^1(\bar{I}^i)_t = 0\}$, we have

$$
\mathbb{III}_t = \frac{1}{(\psi_{HY}k_n)^2} \sum_{i,j\in\mathbb{Z}_+, \bar{R}^\vee(i,j)\leq t} \bar{D}^1(\widehat{\mathcal{I}})^i_t \overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j \bar{K}^{ij} 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \leq \varrho^1_n[i], |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \leq \varrho^2_n[j], N^1(\bar{I}^i)_t \neq 0\}},
$$

and thus an argument similar to the proof of (A.18) yield $\sup_{0\leq s\leq t} |\mathbb{III}_s| = o_p(n^{-1/4})$. Similarly we can show $\sup_{0\leq s\leq t} |\mathbb{IIII}_s| = o_p(n^{-1/4})$ and $\sup_{0\leq s\leq t} |\mathbb{IV}_s| = o_p(n^{-1/4})$. Consequently, we complete the proof of Theorem 3.1. ■

A.4. Proof of Theorem 3.4

Exactly as in the previous section, we can use a localization procedure for the proof, and which allows us to systematically replace the conditions [A4], [A6], and [K$_\beta$] by the following strengthened versions:

[SA4] $\xi \vee \frac{9}{10} < \xi'$ and (A.1) holds.

[SA6] There exists a positive constant C such that $nH_n(t) \leq C$ for every t.

[SK$_\beta$] We have [K$_\beta$] with $E^1 = E^2 =: E$ and $(A^l)'$, $[M^l]'$, $(\underline{A}^l)'$, and $[\underline{M}^l]'$ $(l = 1, 2)$ are bounded. Moreover, there is a nonnegative bounded measurable function ψ on E such that

$$
\sup_{\omega\in\Omega, t\in\mathbb{R}_+} |\delta^l(\omega, t, x)| \leq \psi(x) \text{ and } \int_E \psi(x)^\beta F^l(\mathrm{d}x) < \infty,\ l = 1, 2.
$$

Next, an argument similar to the one in the first part of Section 12 of Hayashi and Yoshida (2011) allows us to assume that $\frac{9}{10} < \xi < \xi' < 1$ under [A1]. Furthermore, in the following we only consider sufficiently large n such that

$$
k_n\bar{r}_n < n^{-\xi+1/2}. \tag{A.19}
$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Now we prove some auxiliary results. Let

$$N^l := 1_{\{|\delta^l|>1\}} \star \mu^l, \qquad L^l := \kappa(\delta^l) \star \left(\mu^l - \nu^l\right).$$

First we need the preaveraged versions of some lemmas in Section 7 of Koike (2013). For processes V and W, $V \bullet W$ denotes the integral (either stochastic or ordinary) of V with respect to W.

LEMMA A.4.1. *Suppose* [A1], [SA4], [SA6], *and* [SK$_2$] *hold. Suppose also that* [SN$_r^\flat$] *holds for some* $r \in (2,\infty)$. *Then for any* $t > 0$ *and any* $q \in [0,2]$ *we have*

$$\sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^q 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], N^1(\bar{I}^i)_t \ne 0\}} = O_p(1),$$
$$\sum_{j:\widehat{T}^{j+k_n}\le t} |\bar{L}^2(\widehat{\mathcal{J}})^j_t|^q 1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \le \varrho^2_n[j], N^2(\bar{J}^j)_t \ne 0\}} = O_p(1) \tag{A.20}$$

as $n \to \infty$.

Proof. For sufficiently large n we have $|\bar{D}(\widehat{\mathcal{I}})^i_t| \ge 1$ on $\{N^1(\bar{I}^i)_t \ne 0\}$, hence

$$|\bar{L}^1(\widehat{\mathcal{I}})^i_t| + |\bar{\zeta}^1(\widehat{\mathcal{I}})^i| \ge |\bar{D}(\widehat{\mathcal{I}})^i_t| - |\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i| - |\bar{X}^1(\widehat{\mathcal{I}})^i_t| - |\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t| > 1/2$$

on $\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], N^1(\bar{I}^i)_t \ne 0\}$. Therefore, we obtain

$$\begin{aligned}
&\sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^q 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], N^1(\bar{I}^i)_t \ne 0\}} \\
&\quad\le \sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^q 1_{\{|\bar{\zeta}^1(\widehat{\mathcal{I}})^i| \ge 1/4, N^1(\bar{I}^i)_t \ne 0\}} + 4^{2-q}\sum_{i=0}^\infty N^1(\bar{I}^i)_t |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 \\
&\quad= \mathbf{I} + \mathbf{II}.
\end{aligned}$$

Lemma A.3.4 yields $\mathbf{I} = O_p(1)$. On the other hand, since N^1 and L^1 have no common jump, Itô's formula yields

$$N^1(\bar{I}^i)|\bar{L}^1(\widehat{\mathcal{I}})^i|^2 = |\bar{L}^1(\widehat{\mathcal{I}})^i|^2_- \bar{I}^i_- \bullet N^1 + 2N^1(\bar{I}^i)_- \bar{L}^1(\widehat{\mathcal{I}})^i_- \bullet \bar{L}^1(\widehat{\mathcal{I}})^i + N^1(\bar{I}^i)_- \bullet [\bar{L}^1(\widehat{\mathcal{I}})^i],$$

and thus we obtain $E[N^1(\bar{I}^i)_t|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2] = E[|\bar{L}^1(\widehat{\mathcal{I}})^i|^2_- \bar{I}^i_- \bullet \Lambda^1_t] + E[N^1(\bar{I}^i)_- \bullet \langle \bar{L}^1(\widehat{\mathcal{I}})^i\rangle_t]$ by the optional sampling theorem, where Λ^1 is the compensator of N^1. Since $\Lambda^1 = 1_{\{|\delta^1|>1\}}\star\nu^1$ and $\langle \bar{L}^1(\widehat{\mathcal{I}})^i\rangle = \sum_{p=1}^{k_n-1} g^n_p(\widehat{I}^{i+p}_- \bullet \langle L^1\rangle) = \sum_{p=1}^{k_n-1} g^n_p[\widehat{I}^{i+p}_- \kappa(\delta^1)]\star\nu^1$, by [SK$_2$], [A1] and the optional sampling theorem we have

$$\begin{aligned}
E[N^1(\bar{I}^i)_t|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2] &\lesssim \int_0^t E[|\bar{L}^1(\widehat{\mathcal{I}})^i_s|^2 \bar{I}^i_s]\mathrm{d}s + \sum_{p=1}^{k_n-1}\int_0^t E[N^1(\bar{I}^i)_s \widehat{I}^{i+p}_s]\mathrm{d}s \\
&= \int_0^t E[\langle \bar{L}^1(\widehat{\mathcal{I}})^i\rangle_s \bar{I}^i_s]\mathrm{d}s + \sum_{p=1}^{k_n-1}\int_0^t E[\Lambda^1(\bar{I}^i)_s \widehat{I}^{i+p}_s]\mathrm{d}s,
\end{aligned}$$

and thus again [SK$_2$] and the representations of Λ^1 and $\langle \bar{L}^1(\widehat{\mathcal{I}})^i\rangle$ yield

$$E[N^1(\bar{I}^i)_t|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2] \lesssim \sum_{p=1}^{k_n-1}\int_0^t E[\widehat{I}^{i+p}_s|\bar{I}^i(t)|]\mathrm{d}s = \sum_{p=1}^{k_n-1} E[|\widehat{I}^{i+p}(t)||\bar{I}^i(t)|],$$

Since [SA6] implies $\sum_{i=1}^{\infty}\sum_{p=1}^{k_n-1}|\widehat{I}^{i+p}(t)||\bar{I}^i(t)| \lesssim k_n^2 n^{-1} \lesssim 1$, we obtain $\mathbf{II} = O_p(1)$, which completes the proof of the first equation of (A.20). Similarly we can prove the second equation of (A.20). ■

Let $\varphi_p(\varepsilon) = \sum_{l=1}^2 \int_{\{\psi\le\varepsilon\}} \psi(x)^p F^l(\mathrm{d}x)$ for each $p \in [\beta,\infty)$. The following lemma is the same one as Lemma 9 of Koike (2013), and will be useful to prove the lemmas below.

LEMMA A.4.2. *Suppose* [SK$_\beta$] *for some* $\beta \in [0,2]$. *Let* p *be a positive number and* (ρ_n) *be a sequence of positive numbers which tends to 0. Then there exists a sequence of numbers* $\varepsilon_n \in [0,1]$ *such that*

$$\limsup_{n\to\infty}\left(\rho_n^{-1}\varepsilon_n^2\right)^p \varphi_\beta(\varepsilon_n) \le 1 \tag{A.21}$$

and

$$\varphi_\beta(\varepsilon_n) \to 0, \qquad \sqrt{\rho_n}/\varepsilon_n \to 0 \tag{A.22}$$

as $n \to \infty$.

Proof. The strategy of the proof is the same as the one in the proof of Lemma 7.4 of Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006). Let

$$a_n' := \sup\left\{y \in (0,\infty) | y^p \varphi_\beta(y\sqrt{\rho_n}) \le 1\right\}, \qquad a_n := 1 \vee \left(a_n' - n^{-1}\right).$$

Since $\varphi_\beta(\varepsilon) \to 0$ as $\varepsilon \to 0$, for any $C > 0$ there exists a positive number ε_0 such that $\varepsilon \le \varepsilon_0$ implies $\varphi_\beta(\varepsilon) \le C^{-p}$. Moreover, since $\rho_n \to 0$ as $n \to \infty$, there exists a positive integer n_0 such that $n \ge n_0$ implies $\rho_n < \varepsilon_0^2/C^2$, hence $C^p\varphi_\beta(C\sqrt{\rho_n}) \le 1$. Therefore we have $a_n' \to \infty$, hence $a_n \to \infty$. Furthermore, for sufficiently large n $a_n < a_n'$, hence $a_n^p\varphi_\beta(a_n\sqrt{\rho_n}) \le 1$. Therefore, if we put $\varepsilon_n := a_n\sqrt{\rho_n} \wedge 1$, we obtain (A.21) and $\sqrt{\rho_n}/\varepsilon_n \to 0$. Moreover, since $\varphi_\beta(\varepsilon_n) \le \varphi_\beta(a_n\sqrt{\rho_n}) \le a_n^{-p} \to 0$, we complete the proof. ■

We introduce some auxiliary notation. We introduce an auxiliary sequence (ε_n) of numbers in [0, 1] such that

$$\limsup_{n\to 0}\frac{\sqrt{\rho_n}}{\varepsilon_n} < \infty, \tag{A.23}$$

and we set $E_n := \left\{x \in E \middle| \psi(x) > \varepsilon_n\right\}$. We will more specify the sequence (ε_n) later. Furthermore, we put

$$D^l := \kappa'(\delta^l)\star\mu^l, \qquad X^l := A^l + M^l, \qquad Y^l := X^l + D^l, \qquad \tilde{N}^l := 1_{E_n}\star\mu^l,$$
$$\mathfrak{L}^l := \kappa(\delta^l)1_{E_n^c}\star(\mu^l-\nu^l) - \kappa(\delta^l)1_{E_n}\star\nu^l, \qquad \Xi^l := \kappa(\delta^l)1_{E_n^c}\star(\mu^l-\nu^l)$$
$$\Theta^l := \kappa(\delta^l)1_{E_n}\star\nu^l$$

for each $l = 1, 2$ and

$$\mathsf{X}^1_{S^i} = X^1_{S^i} + U^1_{S^i}, \quad \mathsf{X}^2_{T^j} = X^2_{T^j} + U^2_{T^j}, \quad \mathsf{Y}^1_{S^i} = Y^1_{S^i} + U^1_{S^i}, \quad \mathsf{Y}^2_{T^j} = Y^2_{T^j} + U^2_{T^j}$$

for each $i, j \in \mathbb{Z}_+$.

LEMMA A.4.3. *Suppose* [SK$_\beta$] *holds for some* $\beta \in [0,2]$. *Then for any* $t > 0$ *we have*

$$\sum_{i=1}^{\infty} \tilde{N}^1(\bar{I}^i)_t = O_p(k_n \varepsilon_n^{-\beta}), \qquad \sum_{j=1}^{\infty} \tilde{N}^2(\bar{J}^j)_t = O_p(k_n \varepsilon_n^{-\beta}). \tag{A.24}$$

Proof. Since $E\left[\sum_{i=1}^{\infty} \tilde{N}^1(\bar{I}^i)_t\right] = \sum_{i=1}^{\infty} E[\hat{\Lambda}^1(\bar{I}^i)_t] \lesssim \varepsilon_n^{-\beta} \sum_{i=1}^{\infty} E\left[|\bar{I}^i(t)|\right] \leq \varepsilon_n^{-\beta} k_n t$, we obtain the first equation of (A.4.3). By symmetry we also obtain the second one. ■

LEMMA A.4.4. *Suppose* [SK$_\beta$] *holds for some* $\beta \in [0,2]$. *Then for any* $t > 0$ *we have*

$$E\left[\sum_{i=1}^{\infty} |\bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})^i_t|^2\right] \lesssim \varepsilon_n^{2-\beta} \varphi_\beta(\varepsilon_n) k_n, \qquad E\left[\sum_{j=1}^{\infty} |\bar{\mathfrak{L}}^2(\widehat{\mathcal{J}})^j_t|^2\right] \lesssim \varepsilon_n^{2-\beta} \varphi_\beta(\varepsilon_n) k_n. \tag{A.25}$$

Proof. Since $E[|\bar{\Xi}^{(1)}(\widehat{\mathcal{I}})^i_t|^2] = E[\langle \bar{\Xi}^{(1)}(\widehat{\mathcal{I}})^i\rangle_t] \lesssim \varphi_2(\varepsilon_n) E\left[\sum_{p=1}^{k_n-1} |\widehat{I}^{i+p}(t)|\right]$ and $E[|\bar{\Theta}(\widehat{\mathcal{I}})^i_t|^2] \lesssim \varepsilon_n^{-2(\beta-1)_+} E\left[|\bar{I}^i(t)|^2\right]$, we have

$$E\left[\sum_{i=1}^{\infty} |\bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})^i_t|^2\right] \lesssim \varphi_2(\varepsilon_n) k_n t + \varepsilon_n^{-2(\beta-1)_+} k_n \bar{r}_n \cdot k_n t.$$

Since $\varphi_2(\varepsilon_n) \leq (\varepsilon_n)^{2-\beta} \varphi_\beta(\varepsilon_n)$, $(\beta-1)_+ \leq \beta/2$ and $k_n \bar{r}_n = o((\varepsilon_n)^2)$ by (18) and (A.23), we obtain the first equation of (A.25). By symmetry we obtain the second equation of (A.25). ■

LEMMA A.4.5. *Suppose* [ST] *and* [SK$_\beta$] *hold for some* $\beta \in [0,2]$. *Then for any* $t > 0$ *we have*

$$\begin{aligned} &\sum_{i=1}^{\infty} 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > c\varrho^1_n[i]\}} = o_p\left(k_n \rho_n^{-\beta/2}\right), \\ &\sum_{j=1}^{\infty} 1_{\{|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2 > c\varrho^{(2)}_n[j]\}} = o_p\left(k_n \rho_n^{-\beta/2}\right). \end{aligned} \tag{A.26}$$

Proof. Combining Lemma A.4.3 with Lemma A.4.4, we obtain

$$\sum_{i=1}^{\infty} 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > c\varrho^1_n[i]\}} \leq \frac{K_0}{\rho_n} \sum_{i=1}^{\infty} |\bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})^i_t|^2 + \sum_{i=1}^{\infty} \tilde{N}^1(\bar{I}^i)_t \lesssim \frac{\varepsilon_n^2}{\rho_n} \varphi_\beta(\varepsilon_n) \varepsilon_n^{-\beta} k_n + \varepsilon_n^{-\beta} k_n,$$

hence Lemma A.4.2 with $p = 1$ yields the first equation of (A.26). By symmetry we also obtain the second one. ■

LEMMA A.4.6. *Suppose* [ST] *and* [SK$_\beta$] *hold for some* $\beta \in [0,2]$. *Then for any* $t > 0$ *we have*

$$\sum_{i=1}^{\infty} |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 \le c\varrho_n^1[i]\}} = o_p\left(k_n \rho_n^{1-\beta/2}\right),$$
$$\sum_{j=1}^{\infty} |\bar{L}^2(\widehat{\mathcal{J}})_t^j|^2 1_{\{|\bar{L}^2(\widehat{\mathcal{J}})_t^j|^2 \le c\varrho_n^2[j]\}} = o_p\left(k_n \rho_n^{1-\beta/2}\right). \tag{A.27}$$

Proof. Since

$$\begin{aligned}
&\sum_{i=1}^{\infty} |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 \le c\varrho_n^1[i]\}} \\
&\quad \le \sum_{i=1}^{\infty} |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 1_{\{N^1(\bar{I}^i)_t = 0\}} + \sum_{i=1}^{\infty} |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 \le c\varrho_n^1[i], N^1(\bar{I}^i)_t \neq 0\}} \\
&\quad \lesssim \sum_{i=1}^{\infty} |\bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})_t^i|^2 + \rho_n \sum_{i=1}^{\infty} 1_{\{N^1(\bar{I}^i)_t \neq 0\}},
\end{aligned}$$

Lemma A.4.3 and Lemma A.4.4 yield $\sum_{i=1}^{\infty} |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2 \le c\varrho_n^1[i]\}} = O_p\left(\{(\varepsilon_n)^2 \varphi_\beta (\varepsilon_n) + \rho_n\} k_n \varepsilon_n^{-\beta}\right)$, hence by Lemma A.4.2 with $p = 2$ we obtain the first equation of (A.27). By symmetry we also obtain the second equation of (A.27). ∎

Next we prove some lemmas which deal with the events that the noise part corrects the effect of small jumps. Let $\eta_n := (k_n \rho_n)^{-1}$. Set

$$\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i = \sum_{p=1}^{k_n-1} g_p^n \left(\mathsf{Y}^1_{\widehat{S}^{i+p}} - \mathsf{Y}^1_{\widehat{S}^{i+p-1}}\right), \qquad \overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j = \sum_{q=1}^{k_n-1} g_q^n \left(\mathsf{Y}^2_{\widehat{T}^{j+q}} - \mathsf{Y}^2_{\widehat{T}^{j+q-1}}\right)$$

for each $i, j \in \mathbb{Z}_+$.

LEMMA A.4.7. *Let* c_1 *and* c_2 *be two positive numbers. Suppose* [ST], [A1], [SA4], [SA6], *and* [SK$_\beta$] *hold for some* $\beta \in [0,2]$. *Suppose also that* [SN$_r^\flat$] *holds for some* $r \in (2,\infty)$. *Then for any* $t > 0$ *we have*

$$\begin{aligned}
&\sum_{i:\widehat{S}^{i+k_n} \le t} |\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2 1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le c_1\varrho_n^1[i], |\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})_t^i|^2 > c_2\varrho_n^1[i]\right\}} \\
&\quad = O_p(1) + o_p\left(\eta_n^{\frac{r-2}{2}} \rho_n^{-\beta/2}\right),
\end{aligned} \tag{A.28}$$

$$\begin{aligned}
&\sum_{j:\widehat{T}^{j+k_n} \le t} |\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2 1_{\left\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \le c_1\varrho_n^2[j], |\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})_t^j|^2 > c_2\varrho_n^2[j]\right\}} \\
&\quad = O_p(1) + o_p\left(\eta_n^{\frac{r-2}{2}} \rho_n^{-\beta/2}\right).
\end{aligned} \tag{A.29}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Proof. Consider (A.28). We decompose the target quantity as

$$\begin{aligned}
&\sum_{i:\widehat{S}^{i+k_n}\leq t} |\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2 1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq c_1\varrho_n^1[i],|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i_t|^2>c_2\varrho_n^1[i]\right\}}\\
&\quad=\sum_{i:\widehat{S}^{i+k_n}\leq t} |\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2 1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq c_1\varrho_n^1[i],|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2>c_2\varrho_n^1[i],N^1(\bar{I}^i)_t\neq 0\right\}}\\
&\qquad+\sum_{i:\widehat{S}^{i+k_n}\leq t} |\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq c_1\varrho_n^1[i],|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2>c_2\varrho_n^1[i],N^1(\bar{I}^i)_t=0\right\}}\\
&\quad=: A_{1,t}+A_{2,t}.
\end{aligned}$$

First, since $(E_0[|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2])_{i=0}^\infty$ is bounded by (A.14), [SK$_\beta$], and [SN$_r$], Lemma A.4.1 yields $A_{1,t}=O_p(1)$. Next, on $\{|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2>c_1\varrho_n^1[i]\}$ we have

$$|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|\geq|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|-|\bar{X}^1(\widehat{\mathcal{I}})^i_t|-|\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t|>\sqrt{\rho_n}\left(\sqrt{\frac{c_1}{K_0}}-2K\sqrt{\frac{2k_n\bar{r}_n\log n}{\rho_n}}\right)$$

by [SA4], [ST], and (A.14). Hence by (18) we have

$$\begin{aligned}
A_{2,t}&\leq\sum_{i:\widehat{S}^{i+k_n}\leq t}|\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>\varrho_n^1[i],|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2>c_1\rho_n/4K_0\}}\\
&\leq 2\left(\frac{2K_0}{c_1\rho_n}\right)^{\frac{r}{2}}\sum_{i:\widehat{S}^{i+k_n}\leq t}|\bar{X}^1(\widehat{\mathcal{I}})^i_t|^2|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>c_2\varrho_n^1[i]\}}\\
&\quad+2\left(\frac{2K_0}{c_1\rho_n}\right)^{\frac{r-2}{2}}\sum_{i:\widehat{S}^{i+k_n}\leq t}|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>c_2\varrho_n^1[i]\}},
\end{aligned}$$

and thus (18) and (A.14) imply that $A_{2,t}\lesssim(\rho_n)^{-\frac{r-2}{2}}\sum_{i:\widehat{S}^{i+k_n}\leq t}|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>c_2\varrho_n^1[i]\}}$. By Lemma A.3.3 we obtain $E_0\left[A_{2,t}\right]\lesssim\eta_n^{\frac{r-2}{2}}k_n^{-1}\sum_{i=1}^\infty 1_{\{\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>\varrho_n^1[i]\}}$, hence Lemma A.4.5 yield $A_{2,t}=o_p\left(\eta_n^{\frac{r-2}{2}}\rho_n^{-\beta/2}\right)$. Consequently, we obtain (A.28). By symmetry we also obtain (A.29), and thus we complete the proof. ∎

LEMMA A.4.8. *Under the assumptions of Lemma A.4.7, for any $t>0$ we have*

$$\begin{aligned}
&\sum_{i:\widehat{S}^{i+k_n}\leq t}|\bar{L}^1(\widehat{\mathcal{I}})^i_t|1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho_n^1[i],|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho_n^1[i]\}}\\
&\quad=O_p(1)+o_p\left(k_n\eta_n^{r/4}\rho_n^{-\beta/4}\right),
\end{aligned}\tag{A.30}$$

$$\begin{aligned}
&\sum_{j:\widehat{T}^{j+k_n}\leq t}|\bar{L}^2(\widehat{\mathcal{J}})^j_t|1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\leq\varrho_n^2[j],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho_n^2[j]\}}\\
&\quad=O_p(1)+o_p\left(k_n\eta_n^{r/4}\rho_n^{-\beta/4}\right).
\end{aligned}\tag{A.31}$$

Proof. Consider (A.30). We decompose the target quantity as

$$\sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t| 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i]\}}$$

$$= \sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t| \Bigg(1_{\left\{|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^2 > \varrho^1_n[i]/4, |\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i]\right\}}$$

$$+ 1_{\left\{|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i]/4, |\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i]\right\}} \Bigg)$$

$$=: \mathbb{B}_{1,t} + \mathbb{B}_{2,t}.$$

By the Schwarz inequality we have

$$\mathbb{B}_{1,t} \le \sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t| 1_{\left\{|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^2 > \varrho^1_n[i]/4, |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i]\right\}}$$

$$\le \left\{ \sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 1_{\{|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^2 > \varrho^1_n[i]/4\}} \right\}^{1/2} \left\{ \sum_{i:\widehat{S}^{i+k_n}\le t} 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i]\}} \right\}^{1/2}.$$

[ST] and Lemma A.3.3 yield

$$E\left[\sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 1_{\{|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^2 > \varrho^1_n[i]/4\}} \right] \lesssim \rho_n^{-r/2} \sum_{i=1}^{\infty} E\left[|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 E_0[|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^r] \right]$$

$$\lesssim \eta_n^{r/2} \sum_{i=1}^{\infty} E[|\bar{I}^i(t)|] \le \eta_n^{r/2} k_n.$$

Combining this with Lemma A.4.5, we obtain $\mathbb{B}_{1,t} = o_p\left(k_n \eta_n^{r/4} \rho_n^{-\beta/4}\right)$. On the other hand, on $\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i]\}$ we have $|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i| \ge |\bar{L}^1(\widehat{\mathcal{I}})^i_t| - |\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i| > \sqrt{\varrho^1_n[i]}$. Moreover, by (A.14) we have

$$|\bar{D}^1(\widehat{\mathcal{I}})^i_t| \ge |\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i| - |\bar{X}^1(\widehat{\mathcal{I}})^i_t| - |\bar{\zeta}^1(\widehat{\mathcal{I}})^i| - |\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})^i_t| > \sqrt{\rho_n}\left(\frac{1}{2\sqrt{K_0}} - 2\sqrt{\frac{2k_n \bar{r}_n \log n}{\rho_n}} \right),$$

on $\{|\bar{\zeta}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i]/4, |\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2 > \varrho^1_n[i], \widehat{S}^{i+k_n} \le t\}$, and thus (18) yields $|\bar{D}^1(\widehat{\mathcal{I}})^i_t| > 0$, hence $N^1(\bar{I}^i) \ne 0$. Therefore, we obtain

$$\mathbb{B}_{2,t} \le \sum_{i:\widehat{S}^{i+k_n}\le t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t| 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], N^1(\bar{I}^i)_t \ne 0\}}, \tag{A.32}$$

and thus Lemma A.4.1 yields $\mathbb{B}_{2,t} = O_p(1)$. Consequently, we obtain (A.30). By symmetry we also obtain (A.31), and thus we complete the proof. ∎

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

LEMMA A.4.9. *Under the assumptions of Lemma A.4.7, for any $t > 0$ we have*

$$\sum_{i:\widehat{S}^{i+k_n}\leq t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 1_{\{|\overline{Z}^1(\widehat{\mathcal{I}})^i|^2\leq \varrho_n^1[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho_n^1[i]\}} = O_p(1) + O_p\left(k_n \eta_n^{r/2}\right), \tag{A.33}$$

$$\sum_{j:\widehat{T}^{j+k_n}\leq t} |\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2 1_{\{|\overline{Z}^2(\widehat{\mathcal{J}})^j|^2\leq \varrho_n^2[j], |\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho_n^2[j]\}} = O_p(1) + O_p\left(k_n \eta_n^{r/2}\right). \tag{A.34}$$

Proof. Consider (A.33). We decompose the target quantity as

$$\begin{aligned}
&\sum_{i:\widehat{S}^{i+k_n}\leq t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 1_{\{|\overline{Z}^1(\widehat{\mathcal{I}})^i|^2\leq \varrho_n^1[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho_n^1[i]\}} \\
&\quad = \sum_{i:\widehat{S}^{i+k_n}\leq t} |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 \Big(1_{\{|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]/4, |\overline{Z}^1(\widehat{\mathcal{I}})^i|^2\leq \varrho_n^1[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho_n^1[i]\}} \\
&\qquad + 1_{\{|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho_n^1[i]/4, |\overline{Z}^1(\widehat{\mathcal{I}})^i|^2\leq \varrho_n^1[i], |\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho_n^1[i]\}}\Big) \\
&\quad =: \Gamma_{1,t} + \Gamma_{2,t}.
\end{aligned}$$

An argument similar to that in the proof of (A.32) yields $\Gamma_{2,t} = O_p(1)$. On the other hand, Lemma A.3.3 yields

$$\begin{aligned}
E[\Gamma_{1,t}] &\lesssim \rho_n^{-r/2} \sum_{i=1}^{\infty} E\left[|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 E_0[|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^r]\right] \\
&\lesssim \eta_n^{r/2} \sum_{i=1}^{\infty} E\left[|\bar{I}^i(t)|\right] \leq \eta_n^{r/2} k_n t,
\end{aligned}$$

hence we obtain (A.33). By symmetry we also obtain (A.34), and thus we complete the proof. ∎

Let $\bar{K}_t^{ij} = 1_{\{\bar{I}^i(t)\cap \bar{J}^j(t)\neq\emptyset\}}$ for each $i, j \in \mathbb{Z}_+$ and $t \in \mathbb{R}_+$.

LEMMA A.4.10. *Suppose that* [A1] *and* [SA4] *are satisfied. Let τ be a $\mathbf{G}^{(n)}$-stopping time. Then $H\bar{K}_\tau^{ij}$ is $\mathcal{F}_{\widehat{S}^i\wedge\widehat{T}^j}$-measurable for any $i, j \in \mathbb{Z}_+$, provided that H is a $\mathcal{G}^{(n)}_{\bar{R}^\vee(i,j)}$-measurable random variable.*

Proof. Let $B = \left\{\bar{I}^i(\tau)\cap \bar{J}^j(\tau)\neq\emptyset\right\}$. It is sufficient to show that $A\cap B\cap C \in \mathcal{F}_u$ for any $u \in \mathbb{R}_+$, where $A \in \mathcal{G}^{(n)}_{\bar{R}^\vee(i,j)}$ and $C = \{\widehat{S}^i \wedge \widehat{T}^j \leq u\}$. On B we have $\bar{R}^\vee(i,j) - \widehat{S}^i\wedge\widehat{T}^j \leq |\bar{I}^i| \vee |\bar{J}^j| \vee (\hat{S}^{i+k_n} - \hat{T}^j) \vee (\hat{T}^{j+k_n} - \hat{S}^i) \leq k_n\bar{r}_n$, hence $\bar{R}^\vee(i,j) = \{\bar{R}^\vee(i,j) - \widehat{S}^i \wedge \widehat{T}^j\} + \widehat{S}^i\wedge\widehat{T}^j \leq \widehat{S}^i\wedge\widehat{T}^j + k_n\bar{r}_n$, and thus we have $B\cap C = B\cap C\cap\{\bar{R}^\vee(i,j)\leq u + k_n\bar{r}_n\}$. Since $A, B \in \mathcal{G}^{(n)}_{\bar{R}^\vee(i,j)}$, we have $A\cap B\cap\{\bar{R}^\vee(i,j)\leq u+k_n\bar{r}_n\} \in \mathcal{G}^{(n)}_{u+k_n\bar{r}_n}$, however, $\mathcal{G}^{(n)}_{u+k_n\bar{r}_n} = \mathcal{F}_{(u+k_n\bar{r}_n - n^{-\xi+\frac{1}{2}})_+} \subset \mathcal{F}_u$ by (A.19). This together with the fact that $C \in \mathcal{F}_u$ implies $A\cap B\cap C\in\mathcal{F}_u$. ∎

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

LEMMA A.4.11. *Suppose* [SC1]–[SC2], [A1], [SA4], [SA6], [SK$_2$], *and* [SN$_2^\flat$] *hold. Then for any $t > 0$, there exists a positive constant K independent of both n and (ε_n) such that*

$$E\left[\sup_{0\le s\le t}\left|\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Xi}^1\left(\widehat{\mathcal{I}}\right)^i_s\bar{M}^2\left(\widehat{\mathcal{J}}\right)^j_s\bar{K}^{ij}\right|\right]\le Kk_n^2\sqrt{\varphi_2(\varepsilon_n)}n^{-\frac{1}{4}}, \tag{A.35}$$

$$E\left[\sup_{0\le s\le t}\left|\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Xi}^1\left(\widehat{\mathcal{I}}\right)^i_s\widetilde{\underline{M}^2}\left(\widehat{\mathcal{J}}\right)^j_s\bar{K}^{ij}\right|\right]\le Kk_n^2\sqrt{\varphi_2(\varepsilon_n)}n^{-\frac{1}{4}}, \tag{A.36}$$

$$E\left[\sup_{0\le s\le t}\left|\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Xi}^1\left(\widehat{\mathcal{I}}\right)^i_s\bar{\zeta}^2\left(\widehat{\mathcal{J}}\right)^j\bar{K}^{ij}\right|\right]\le Kk_n^2\sqrt{\varphi_2(\varepsilon_n)}n^{-\frac{1}{4}}. \tag{A.37}$$

Proof. First consider (A.35). Since integration by parts and Lemma A.1.3 yield

$$\bar{\Xi}^1(\widehat{\mathcal{I}})^i_t\bar{M}^2(\widehat{\mathcal{J}})^j_t\bar{K}^{ij}_s=\left\{\bar{K}^{ij}_-\bar{\Xi}^1(\widehat{\mathcal{I}})^i_-\right\}\bullet\bar{M}^2(\widehat{\mathcal{J}})^j_s+\left\{\bar{K}^{ij}_-\bar{M}^2(\widehat{\mathcal{J}})^j_-\right\}\bullet\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s, \tag{A.38}$$

we can decompose the target quantity as

$$\begin{aligned}&\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s\bar{M}^2(\widehat{\mathcal{J}})^j_s\bar{K}^{ij}\\&\quad=\sum_{i,j=0}^{\infty}\left[\left\{\bar{K}^{ij}_-\bar{\Xi}^1(\widehat{\mathcal{I}})^i_-\right\}\bullet\bar{M}^2(\widehat{\mathcal{J}})^j_s+\left\{\bar{K}^{ij}_-\bar{M}^2(\widehat{\mathcal{J}})^j_-\right\}\bullet\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s\right]\\&\qquad+\sum_{i,j:\bar{R}^\vee(i,j)>s}\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s\bar{M}^2(\widehat{\mathcal{J}})^j_s\bar{K}^{ij}_s\\&\quad=:\Delta_{1,s}+\Delta_{2,s}+\Delta_{3,s}.\end{aligned}$$

First we estimate $\Delta_{1,s}$. Since $\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s=\sum_{p=1}^{k_n-1}g^n_p\Xi^1(\widehat{I}^{i+p})_s$ and $\bar{M}^2(\widehat{\mathcal{J}})^j_s=\sum_{q=1}^{k_n-1}g^n_q M^2(\widehat{J}^{j+q})_s$, we have

$$\begin{aligned}\Delta_{1,s}&=\sum_{i,j=0}^{\infty}\sum_{p,q=1}^{k_n-1}g^n_pg^n_q\bar{K}^{ij}_-\Xi^1(\widehat{I}^{i+p})_-\widehat{J}^{j+q}_-\bullet M^2_s\\&=\sum_{i,j=0}^{\infty}\sum_{p=i+1}^{i+k_n-1}\sum_{q=j+1}^{j+k_n-1}g^n_{p-i}g^n_{q-j}\bar{K}^{ij}_-\Xi^1(\widehat{I}^{i+p})_-\widehat{J}^{j+q}_-\bullet M^2_s\\&=\sum_{p,q=1}^{\infty}\upsilon(p,q)_-\Xi^1(\widehat{I}^p)_-\widehat{J}^q_-\bullet M^2_s,\end{aligned}$$

where $\upsilon(p,q)_s=\sum_{i=(p-k_n+1)_+}^{p-1}\sum_{j=(q-k_n+1)_+}^{q-1}g^n_{p-i}g^n_{q-j}\bar{K}^{ij}_s$. Hence we have

$$\langle\Delta_{1,\cdot}\rangle_s=\sum_{q=1}^{\infty}\left[\sum_{p=1}^{\infty}\upsilon(p,q)_-\Xi^1(\widehat{I}^p)_-\right]^2\widehat{J}^q_-\bullet\langle M^2\rangle_s,$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

and thus we obtain

$$
\begin{aligned}
E[\langle \Delta_{1,\cdot}\rangle_t] &\lesssim \sum_{q=1}^{\infty}\int_0^t E\left[\left\{\sum_{p=1}^{\infty} \upsilon(p,q)_s\,\Xi^1(\widehat{I}^p)_s\right\}^2 \widehat{J}^q_s\right]\mathrm{d}s \\
&= \sum_{q=1}^{\infty}\int_0^t \sum_{p=1}^{\infty} E\left[\upsilon(p,q)_s^2\widehat{J}^q_s\, E\left[|\Xi^1(\widehat{I}^p)_s|^2|\mathcal{F}^{(0)}_{\widehat{S}^{p-1}}\right]\right]\mathrm{d}s \\
&\lesssim k_n^2\varphi_2(\varepsilon_n)\sum_{p,q=1}^{\infty}\sum_{i=(p-k_n+1)_+}^{p-1}\sum_{j=(q-k_n+1)_+}^{q-1} E\left[\bar{K}^{ij}_t|\widehat{I}^p(t)||\widehat{J}^q(t)|\right] \\
&\lesssim k_n^4\varphi_2(\varepsilon_n)k_n n^{-1} \lesssim k_n^4\varphi_2(\varepsilon_n)n^{-1/2}
\end{aligned}
$$

by the representation of $\langle M^2\rangle$ and $\langle \Xi^1\rangle$, [A1], (A.2), [SA6], and (7). Combining this with the Schwarz and Doob inequalities, we conclude that

$$
E\left[\sup_{0\leq s\leq t}|\Delta_{1,s}|\right] \lesssim k_n^2\sqrt{\varphi_2(\varepsilon_n)}n^{-1/4}. \tag{A.39}
$$

Similarly we can also show that $E\left[\sup_{0\leq s\leq t}|\Delta_{2,s}|\right] \lesssim k_n^2\sqrt{\varphi_2(\varepsilon_n)}n^{-1/4}$. Now we estimate $\Delta_{3,s}$. (A.38), the Doob inequality, [A1] and the optional sampling theorem, (A.2) and [SA6] imply that

$$
\begin{aligned}
E\left[\sup_{0\leq s\leq t}\sum_{i,j=0}^{\infty}|\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s\bar{M}^2(\widehat{\mathcal{J}})^j_s\bar{K}^{ij}_s|^2\right] &\lesssim \varphi_2(\varepsilon_n)E\left[\sum_{i,j=1}^{\infty}\bar{K}^{ij}|\bar{I}^i(t)||\bar{J}^j(t)|\right] \\
&\lesssim \varphi_2(\varepsilon_n)k_n.
\end{aligned} \tag{A.40}
$$

On the other hand, (A.2) also implies

$$
\sum_{i,j:\bar{R}^\vee(i,j)>s}\bar{K}^{ij}_s \leq (2k_n+1)\left[\sum_{i:\widehat{S}^{i+k_n}>s}1+\sum_{j:\widehat{T}^{j+k_n}>s}1\right] \leq (2k_n+1)k_n. \tag{A.41}
$$

Therefore, the Schwarz inequality and (7) yield $E\left[\sup_{0\leq s\leq t}|\Delta_{3,s}|\right] \lesssim \sqrt{\varphi_2(\varepsilon_n)k_n}\cdot k_n \lesssim k_n^2\sqrt{\varphi_2(\varepsilon_n)}n^{-1/4}$. Consequently, we conclude that (A.35) holds. (A.36) can be shown in a similar manner.

Finally consider (A.37). Define the process $\mathfrak{z}^2_t$ by $\mathfrak{z}^2_t = n^{-1/2}\sum_{j=1}^{\infty}\zeta^2_{\widehat{T}^j}1_{\{\widehat{T}^j\leq t\}}$. Then obviously $\mathfrak{z}^2_t$ is a purely discontinuous locally square-integrable martingale $\mathcal{B}^{(0)}$ and $\overline{\zeta}^2(\widehat{\mathcal{J}})^j = \widetilde{\mathfrak{z}}^2(\widehat{\mathcal{J}})^j$ on $\{\widehat{T}^{j+k_n}\leq t\}$. On the other hand, since Ξ^1 is quasi-left continuous by Theorem I-4.2 of Jacod and Shiryaev (2003) and for every j $\widehat{T}^j$ is $\mathbf{F}^{(0)}$-predictable time by [A1], we have $\Delta\Xi^1_{\widehat{T}^j} = 0$ for every j. Therefore, we have $[\Xi^1,\mathfrak{z}^2] = 0$, and thus we can decompose the target quantity as

$$
\sum_{i,j:\bar{R}^\vee(i,j)\leq s}\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s\overline{\zeta}^2(\widehat{\mathcal{J}})^j_s\bar{K}^{ij} = \sum_{i,j=0}^{\infty}\left[\left\{\bar{K}^{ij}_-\bar{\Xi}^1(\widehat{\mathcal{I}})^i_-\right\}\bullet\widetilde{\mathfrak{z}}^2(\widehat{\mathcal{J}})^j_s+\left\{\bar{K}^{ij}_-\widetilde{\mathfrak{z}}^2(\widehat{\mathcal{J}})^j_-\right\}\bullet\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s\right]
$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

$$
\begin{aligned}
&+ \sum_{i,j:\bar{R}^{\vee}(i,j)>s} \bar{\Xi}^1(\widehat{\mathcal{I}})^i_s \widetilde{\mathfrak{Z}}^2(\widehat{\mathcal{J}})^j_s \bar{K}^{ij}_s \\
&=: \Upsilon_{1,s} + \Upsilon_{2,s} + \Upsilon_{3,s}
\end{aligned}
$$

due to integration by parts and Lemma A.1.3. First we estimate $\Upsilon_{1,s}$. Since

$$
\begin{aligned}
\Upsilon_{1,s} &= \sqrt{n} \sum_{i,j=0}^{\infty} \sum_{p=1}^{k_n-1} \sum_{q=0}^{k_n-1} g^n_p \Delta(g)^n_q \bar{K}^{ij}_- \Xi^1(\widehat{I}^{i+p})_- \widehat{J}^{j+q}_- \bullet \mathfrak{Z}^2_s \\
&= \sqrt{n} \sum_{i,j=0}^{\infty} \sum_{p=i+1}^{i+k_n-1} \sum_{q=j}^{j+k_n-1} g^n_{p-i} \Delta(g)^n_{q-j} \bar{K}^{ij}_- \Xi^1(\widehat{I}^{i+p})_- \widehat{J}^{j+q}_- \bullet \mathfrak{Z}^2_s \\
&= \sqrt{n} \sum_{p=1}^{\infty} \sum_{q=0}^{\infty} v'(p,q)_- \Xi^1(\widehat{I}^p)_- \widehat{J}^q_- \bullet \mathfrak{Z}^2_s,
\end{aligned}
$$

where $v'(p,q)_s = \sum_{i=(p-k_n+1)_+}^{p-1} \sum_{j=(q-k_n+1)_+}^{q} g^n_{p-i} \Delta(g)^n_{q-j} \bar{K}^{ij}_s$. We have

$$
[\Upsilon_{1,\cdot}]_s = \sum_{q=0}^{\infty} \left[\sum_{p=1}^{\infty} v'(p,q)_{\widehat{T}^q} \Xi^1(\widehat{I}^p)_{\widehat{T}^q} \right]^2 \left(\zeta^2_{\widehat{T}^q} \right)^2 1_{\{\widehat{T}^q \le s\}},
$$

hence we obtain

$$
\begin{aligned}
E\left[[\Upsilon_{1,\cdot}]_t\right] &\lesssim \sum_{q=0}^{\infty} E\left[\left\{ \sum_{p=1}^{\infty} v'(p,q)_{\widehat{T}^q} \Xi^1(\widehat{I}^p)_{\widehat{T}^q} \right\}^2 1_{\{\widehat{T}^q \le t\}} \right] \\
&= \sum_{q=0}^{\infty} \sum_{p=1}^{\infty} E\left[v'(p,q)^2_{\widehat{T}^q \wedge t} E\left[|\Xi^1(\widehat{I}^p)_{\widehat{T}^q \wedge t}|^2 | \mathcal{F}^{(0)}_{\widehat{S}^{p-1} \wedge \widehat{T}^q \wedge t} \right] \right] \\
&\lesssim \varphi_2(\varepsilon_n) \sum_{q=0}^{\infty} \sum_{p=1}^{\infty} \sum_{i=(p-k_n+1)_+}^{p-1} \sum_{j=(q-k_n+1)_+}^{q} E\left[\bar{K}^{ij}_t |\widehat{I}^p(t)| \right] \\
&\lesssim \varphi_2(\varepsilon_n) k_n^3 \lesssim k_n^4 \varphi_2(\varepsilon_n) n^{-1/2}
\end{aligned}
$$

by the optional sampling theorem, the representation of $\langle \Xi^1 \rangle$, Lemma A.4.10, the Lipschitz continuity of g, (A.2) and (7). Since $\langle \Upsilon_{1,\cdot} \rangle$ is the predictable compensator of $[\Upsilon_{1,\cdot}]$, the above result and the Schwarz and Doob inequalities yield $E\left[\sup_{0 \le s \le t} |\Upsilon_{1,s}|\right] \lesssim k_n^2 \sqrt{\varphi_2(\varepsilon_n)} n^{-1/4}$. On the other hand, we can show that $E\left[\sup_{0 \le s \le t} |\Upsilon_{2,s}|\right] \lesssim k_n^2 \sqrt{\varphi_2(\varepsilon_n)} n^{-1/4}$ in a similar manner to the proof of (A.39). Finally, since

$$
\begin{aligned}
E\left[\sup_{0 \le s \le t} \sum_{i,j=0}^{\infty} |\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s \bar{\zeta}^2(\widehat{\mathcal{J}})^j_s \bar{K}^{ij}_s|^2 \right] &\lesssim \sum_{i=0}^{\infty} E\left[\sup_{0 \le s \le t} |\bar{\Xi}^1(\widehat{\mathcal{I}})^i_s|^2 \right] \\
&\lesssim \varphi_2(\varepsilon_n) E\left[\sum_{i=0}^{\infty} |\bar{I}^i(t)| \right] \lesssim \varphi_2(\varepsilon_n) k_n
\end{aligned}
$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

by (A.2) and the Doob inequality, the Schwarz inequality and (A.41) yield $E\left[\sup_{0\le s\le t}|\Upsilon_{3,s}|\right]\lesssim k_n^2\sqrt{\varphi_2(\varepsilon_n)}n^{-1/4}$. Consequently, we obtain (A.37). ∎

LEMMA A.4.12. *Suppose* [ST], [A1], [SA4], *and* [SK$_\beta$] *hold for some* $\beta\in[0,2]$. *Suppose also that* [SN$_r^\flat$] *holds for some* $r\in(2,\infty)$. *Then for any* $t>0$ *we have*

$$k_n^{-2}\sup_{0\le s\le t}\left|\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Theta}^1(\widehat{\mathcal{I}})^i_s\bar{M}^2(\widehat{\mathcal{J}})^j_s\bar{K}^{ij}\right|=o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right),\tag{A.42}$$

$$k_n^{-2}\sup_{0\le s\le t}\left|\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Theta}^1(\widehat{\mathcal{I}})^i_s\underline{\widetilde{M^2}}(\widehat{\mathcal{J}})^j_s\bar{K}^{ij}\right|=o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right),\tag{A.43}$$

$$k_n^{-2}\sup_{0\le s\le t}\left|\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Theta}^1(\widehat{\mathcal{I}})^i_s\bar{\zeta}^2(\widehat{\mathcal{J}})^j\bar{K}^{ij}\right|=o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right).\tag{A.44}$$

Proof. Define the process Υ_s by $\Upsilon_s=\sum_{i,j:\bar{R}^\vee(i,j)\le s}\bar{\Theta}^1(\widehat{\mathcal{I}})^i_s\bar{M}^2(\widehat{\mathcal{J}})^j_s\bar{K}^{ij}$. If $\beta\ge 1$, the Schwarz inequality, [SK$_\beta$], and [SA4] yield

$$\begin{aligned}E\left[\sup_{0\le s\le t}|\Upsilon_s|\right]&\le E\left[\sum_{i,j=1}^\infty\bar{K}^{ij}|\bar{\Theta}^1(\widehat{\mathcal{I}})^i_t\bar{M}^2(\widehat{\mathcal{J}})^j_t|\right]\\&\le\left\{E\left[\sum_{i,j=1}^\infty\bar{K}^{ij}|\bar{\Theta}^1(\widehat{\mathcal{I}})^i_t|^2\right]\right\}^{1/2}\left\{E\left[\sum_{i,j=1}^\infty\bar{K}^{ij}|\bar{M}^2(\widehat{\mathcal{J}})^j_t|^2\right]\right\}^{1/2}\\&=o\left((k_n\bar{r}_n)^{1/2}\rho_n^{-(\beta-1)/2}\right),\end{aligned}$$

hence we obtain $\sup_{0\le s\le t}|\Upsilon_s|=o_p\left(\rho_n^{1-\beta/2}\right)$. If $\beta<1$, we decompose the target quantity as

$$\Upsilon_s=\sum_{i,j:\bar{R}^\vee(i,j)\le s}\left\{\bar{\hat{\Theta}}^1(\widehat{\mathcal{I}})^i_s-\bar{\check{\Theta}}^1(\widehat{\mathcal{I}})^i_s\right\}\bar{M}^2(\widehat{\mathcal{J}})^j\bar{K}^{ij}=\Upsilon_{1,s}+\Upsilon_{2,s},$$

where $\hat{\Theta}^1=\kappa(\delta^1)\star\nu^1$ and $\check{\Theta}^1=\kappa(\delta^1)1_{(E_n^1)^c}\star\nu^1$. Then, by [SA6] and an argument similar to the above, we can show $\sup_{0\le s\le t}|\Upsilon_{2,s}|=O_p\left(k_n^2\rho_n^{(1-\beta)/2}k_nn^{-1/2}\right)=o_p(k_n^2n^{-1/4})$. On the other hand, note that $|\bar{\hat{\Theta}}^1(\widehat{\mathcal{I}})^i_s|\lesssim k_n\bar{r}_n$ because $\beta<1$, an argument similar to the proof of Lemma A.2.1 yields

$$\sup_{0\le s\le t}\left|\Upsilon_{1,s}-\sum_{i,j=1}^\infty\bar{\hat{\Theta}}^1(\widehat{\mathcal{I}})^i_s\bar{M}^2(\widehat{\mathcal{J}})^j_s\bar{K}^{ij}_s\right|=o_p\left(k_n^2n^{-1/4}\right),$$

hence Lemma A.2.2 yields $\sup_{0\le s\le t}|\Upsilon_{1,s}|=o_p\left(k_n^2n^{-1/4}\right)$ since we have the conditions [A1], [K$_\beta$](v), and [SA6]. Consequently, we conclude that (A.42). Similarly we can also show (A.43) and (A.44). ∎

Proof of Theorem 3.4. By a localization procedure, we may replace the conditions $[K_\beta]$, [C1]–[C2], and $[N_r^\flat]$ with [SF], [SC1]–[SC2], and $[SN_r^\flat]$ respectively. Moreover, we can also replace the condition [T] with [ST] by Lemma A.3.1, while (5) can be replaced with (A.1) due to the above argument.

We decompose the target quantity as

$$\begin{aligned}
&PTHY(\mathsf{Z}^1,\mathsf{Z}^2)^n_t - PHY(\mathsf{X}^1,\mathsf{X}^2)^n_t \\
&= \frac{1}{(\psi_{HY}k_n)^2}\Bigg[\sum_{i,j:\bar{R}^\vee(i,j)\le t}\Big\{\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}}^i)\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}}^j)1_{\{|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n[i],|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2\le 4\varrho_n[j]\}} \\
&\quad -\overline{\mathsf{X}}^1(\widehat{\mathcal{I}})^i\overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j\Big\}\bar{K}^{ij} + \sum_{i,j:\bar{R}^\vee(i,j)\le t}\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j\bar{K}^{ij} \\
&\quad \times\left(1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i],|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho_n^2[j]\}} - 1_{\{|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n^1[i],|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2\le 4\varrho_n^2[j]\}}\right) \\
&\quad + \sum_{i,j:\bar{R}^\vee(i,j)\le t}\Big\{\bar{L}^1(\widehat{\mathcal{I}})^i_t\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j + \overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i\bar{L}^2(\widehat{\mathcal{J}})^j_t + \bar{L}^1(\widehat{\mathcal{I}})^i_t\bar{L}^2(\widehat{\mathcal{J}})^j_t\Big\} \\
&\quad \times\bar{K}^{ij}1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i],|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho_n^2[j]\}}\Bigg] \\
&=: \mathbb{I}_t + \mathbb{II}_t + \mathbb{III}_t + \mathbb{IV}_t + \mathbb{V}_t. \qquad \textbf{(A.45)}
\end{aligned}$$

(a) By Theorem 3.1, we have

$$\sup_{0\le s\le t}|\mathbb{I}_s| = o_p(n^{-1/4}) + O_p\left(\eta_n^{\frac{r-2}{2}}\right). \qquad \textbf{(A.46)}$$

(b) Next consider $\mathbb{II}$. We decompose it as

$$\begin{aligned}
\mathbb{II}_t &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\le t}\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j\bar{K}^{ij} \\
&\quad\times\Big(1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i],|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho_n^2[j]\}}1_{\{|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2>4\varrho_n^1[i]\}\cup\{|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2>4\varrho_n^2[j]\}} \\
&\qquad -1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]\}\cup\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j]\}}1_{\{\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n^1[i],\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2\le 4\varrho_n^2[j]\}}\Big) \\
&=: \mathbb{II}_{1,t} + \mathbb{II}_{2,t}.
\end{aligned}$$

First estimate $\mathbb{II}_{1,t}$. We decompose it as

$$\begin{aligned}
\mathbb{II}_{1,t} &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\le t}\overline{\mathsf{Y}}^{(1)}(\widehat{\mathcal{I}})^i\overline{\mathsf{Y}}^{(2)}(\widehat{\mathcal{J}})^j\bar{K}^{ij}1_{\{|\overline{\mathsf{Z}}^{(1)}(\widehat{\mathcal{I}})^i|^2\le\varrho_n^{(1)}[i],|\overline{\mathsf{Z}}^{(2)}(\widehat{\mathcal{J}})^j|^2\le\varrho_n^{(2)}[j]\}} \\
&\quad\times\Big(1_{\{|\overline{\mathsf{Y}}^{(1)}(\widehat{\mathcal{I}})^i|^2>4\varrho_n^{(1)}[i],|\overline{\mathsf{Y}}^{(2)}(\widehat{\mathcal{J}})^j|^2\le 4\varrho_n^{(2)}[j]\}} + 1_{\{|\overline{\mathsf{Y}}^{(1)}(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n^{(1)}[i],|\overline{\mathsf{Y}}^{(2)}(\widehat{\mathcal{J}})^j|^2>4\varrho_n^{(2)}[j]\}} \\
&\qquad +1_{\{|\overline{\mathsf{Y}}^{(1)}(\widehat{\mathcal{I}})^i|^2>4\varrho_n^{(1)}[i],|\overline{\mathsf{Y}}^{(2)}(\widehat{\mathcal{J}})^j|^2>4\varrho_n^{(2)}[j]\}}\Big) \\
&=: \mathbb{II}_{1,t}^{(1)} + \mathbb{II}_{1,t}^{(2)} + \mathbb{II}_{1,t}^{(3)}.
\end{aligned}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Consider $\mathbb{III}_{1,t}^{(1)}$. Since (A.2) yields

$$\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}^{(1)}|\lesssim\frac{1}{k_n}\sum_{i:\widehat{S}^{i+k_n}\leq t}|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2 1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho_n^1[i],|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2>4\varrho_n^1[i]\right\}},$$

Lemma A.4.7 implies that $\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}^{(1)}|=O_p(k_n^{-1})+o_p(k_n^{-1}\eta_n^{\frac{r-2}{2}}\rho_n^{-\beta/2})$. Since $k_n^{-1}=o(n^{-1/4})$ and $k_n^{-1}\rho_n^{-\beta/2}=o(1)$, we conclude that

$$\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}^{(1)}|=o_p(n^{-1/4})+o_p\left(\eta_n^{\frac{r-2}{2}}\right). \tag{A.47}$$

By symmetry we also obtain

$$\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}^{(2)}|=o_p(n^{-1/4})+o_p\left(\eta_n^{\frac{r-2}{2}}\right). \tag{A.48}$$

On the other hand, the Schwarz inequality and (A.2) yield

$$\begin{aligned}\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}^{(3)}|\leq{}&\frac{1}{\psi_{HY}^2k_n}\left\{\sum_{i:\widehat{S}^{i+k_n}\leq t}|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho_n^1[i],|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2>4\varrho_n^1[i]\}}\right\}^{1/2}\\&\times\left\{\sum_{j:\widehat{T}^{j+k_n}\leq t}|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2 1_{\{|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2>4\varrho_n^2[j],|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\leq\varrho_n^2[j]\}}\right\}^{1/2},\end{aligned}$$

hence Lemma A.4.7 yields $\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}^{(3)}|=O_p(k_n^{-1})+O_p\left(k_n^{-1}\eta_n^{\frac{r-2}{2}}\rho_n^{-\beta/2}\right)=o_p(n^{-1/4})+o_p\left(\eta_n^{\frac{r-2}{2}}\right)$, and thus, we obtain

$$\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}^{(3)}|=o_p(n^{-1/4})+o_p\left(\eta_n^{\frac{r-2}{2}}\right). \tag{A.49}$$

(A.47), (A.48) and (A.49) yield

$$\sup_{0\leq s\leq t}|\mathbb{III}_{1,s}|=o_p(n^{-1/4})+o_p\left(\eta_n^{\frac{r-2}{2}}\right)+o_p\left(\rho_n^{1-\beta/2}\right). \tag{A.50}$$

Next estimate $\mathbb{III}_{2,t}$. Note that (A.2), we obtain

$$\begin{aligned}\sup_{0\leq s\leq t}|\mathbb{III}_{2,t}|&\lesssim\frac{\rho_n}{k_n}\left\{\sum_{i:\widehat{S}^{i+k_n}\leq t}1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i],|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2\leq4\varrho_n^1[i]\right\}}\right.\\&\qquad\left.+\sum_{j:\widehat{T}^{j+k_n}\leq t}1_{\left\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j],|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2\leq4\varrho_n^2[j]\right\}}\right\}\\&\leq\frac{\rho_n}{k_n}\left\{\sum_{i:\widehat{S}^{i+k_n}\leq t}1_{\left\{|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]/4\right\}}+\sum_{j:\widehat{T}^{j+k_n}\leq t}1_{\left\{|\overline{\zeta}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j]/4\right\}}\right.\end{aligned}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

$$+\sum_{i:\widehat{S}^{i+k_n}\le t}1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i],|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n^1[i],|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i]/4\right\}}$$
$$\left.+\sum_{j:\widehat{T}^{j+k_n}\le t}1_{\left\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j],|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2\le 4\varrho_n^2[j],|\overline{\zeta}^2(\widehat{\mathcal{J}})^j|^2\le\varrho_n^2[j]/4\right\}}\right\}.$$

[SC1] and Lemma A.3.3 yield

$$\sum_{i:\widehat{S}^{i+k_n}\le t}1_{\left\{|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]/4\right\}}\lesssim n\eta_n^{r/2},\qquad \sum_{j:\widehat{T}^{j+k_n}\le t}1_{\left\{|\overline{\zeta}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j]/4\right\}}\lesssim n\eta_n^{r/2}. \quad \textbf{(A.51)}$$

On the other hand, on $\left\{|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n^{(1)}[i],|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i]/4,\widehat{S}^{i+k_n}\le t\right\}$ we have

$$\begin{aligned}|\bar{D}^1(\widehat{\mathcal{I}})_t^i|&\le|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|+|\bar{X}^1(\widehat{\mathcal{I}})_t^i|+|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|+|\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})_t^i|\\&\le 5\sqrt{\varrho_n^1[i]}/2+2K\sqrt{2k_n\bar{r}_n\log n}\to 0,\end{aligned}$$

hence a.s. for sufficiently large n we have $\bar{D}^1(\widehat{\mathcal{I}})_t^i=0$. Moreover, on $\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]$, $|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i]/4,\bar{D}^1(\widehat{\mathcal{I}})_t^i=0,\widehat{S}^{i+k_n}\le t\}$, by (A.14) we have

$$\begin{aligned}|\bar{L}^1(\widehat{\mathcal{I}})_t^i|&\ge|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|-|\bar{X}^1(\widehat{\mathcal{I}})_t^i|-|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|-|\widetilde{\underline{X}^1}(\widehat{\mathcal{I}})_t^i|\\&>\sqrt{\rho_n}\left(\frac{1}{2\sqrt{K_0}}-2K\sqrt{\frac{2k_n\bar{r}_n\log n}{\rho_n}}\right),\end{aligned}$$

hence (18) yields $|\bar{L}^1(\widehat{\mathcal{I}})_t^i|>\sqrt{\rho_n/9K_0}$. Therefore we obtain

$$\sum_{i:\widehat{S}^{i+k_n}\le t}1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i],|\overline{\mathsf{Y}}^1(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n^1[i],|\overline{\zeta}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i]/4\}}\le\sum_{i=1}^{\infty}1_{\{|\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2>\rho_n/9K_0\}}.$$

By symmetry we also obtain

$$\begin{aligned}&\sum_{j:\widehat{T}^{j+k_n}\le t}1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j],|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2\le 4\varrho_n^2[j],|\overline{\zeta}^2(\widehat{\mathcal{J}})^j|^2\le\varrho_n^2[j]/4\}}\\&\quad\le\sum_{j=1}^{\infty}1_{\{|\bar{L}^2(\widehat{\mathcal{J}})_t^j|^2>\rho_n/9K_0\}}.\end{aligned}$$

Combining these results with Lemma A.4.5 and (7), we obtain

$$\sup_{0\le s\le t}|\mathbb{III}_{2,t}|=O_p\left(\eta_n^{\frac{r-2}{2}}\right)+o_p\left(\rho_n^{1-\beta/2}\right). \quad \textbf{(A.52)}$$

By (A.50) and (A.52), we conclude

$$\sup_{0\le s\le t}|\mathbb{III}_t|=o_p(n^{-1/4})+O_p\left(\eta_n^{\frac{r-2}{2}}\right)+o_p\left(\rho_n^{1-\beta/2}\right). \quad \textbf{(A.53)}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

(c) Next consider $\mathbb{IIII}$. We decompose it as

$$\mathbb{IIII}_t = \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\le t}\bar{L}^1(\widehat{\mathcal{I}})^i_t\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j\bar{K}^{ij}1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\le\varrho^1_n[i],|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho^2_n[j]\}}$$
$$\times\Big(1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho^1_n[i],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho^2_n[j]\}}+1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho^1_n[i],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2\le4\varrho^2_n[j]\}}$$
$$+1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2\le4\varrho^1_n[i],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho^2_n[j]\}}+1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2\le4\varrho^1_n[i],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2\le4\varrho^2_n[j]\}}\Big)$$
$$=: \mathbb{IIII}_{1,t}+\mathbb{IIII}_{2,t}+\mathbb{IIII}_{3,t}+\mathbb{IIII}_{4,t}.$$

First we estimate $\mathbb{IIII}_{1,t}$. The Schwarz inequality and (A.2) yield

$$\sup_{0\le s\le t}|\mathbb{IIII}_{1,s}|\le\frac{1}{\psi^2_{HY}k_n}\left\{\sum_{i:\widehat{S}^{i+k_n}\le t}|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\le\varrho^1_n[i],|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho^1_n[i]\}}\right\}^{1/2}$$
$$\times\left\{\sum_{j:\widehat{T}^{j+k_n}\le t}|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2 1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho^2_n[j],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho^2_n[j]\}}\right\}^{1/2}.$$

On $\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho^2_n[j],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho^2_n[j]\}$ we have $|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|\ge|\bar{L}^2(\widehat{\mathcal{J}})^j_t|-|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|>\sqrt{\varrho^2_n[j]}$, hence Lemma A.4.7 and Lemma A.4.9 imply that

$$\sup_{0\le s\le t}|\mathbb{IIII}_{1,s}|=O_p\left(k_n^{-1}\right)+o_p\left(k_n^{-1}\eta_n^{\frac{r-2}{4}}\rho_n^{-\beta/4}\right)+o_p\left(k_n^{-1/2}\eta_n^{r/4}\right)+o_p\left(k_n^{-1/2}\eta_n^{\frac{r-1}{2}}\rho_n^{-\beta/4}\right)$$
$$=o_p\left(n^{-1/4}\right)+o_p\left(\eta_n^{\frac{r-2}{4}}\rho_n^{1/2-\beta/4}\right). \quad \textbf{(A.54)}$$

Next we estimate $\mathbb{IIII}_{2,t}$. On $\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho^2_n[j],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2\le4\varrho^2_n[j]\}$ we have $|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|\le|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|+|\bar{L}^2(\widehat{\mathcal{J}})^j_t|\le3\sqrt{\varrho^2_n[j]}$, hence by (A.2) we obtain

$$\sup_{0\le s\le t}|\mathbb{IIII}_{2,s}|\lesssim\frac{\sqrt{\rho_n}}{k_n}\sum_{i:\widehat{S}^{i+k_n}\le t}|\bar{L}^1(\widehat{\mathcal{I}})^i|1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\le\varrho^1_n[i],|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho^1_n[i]\}},$$

and thus Lemma A.4.8 yields

$$\sup_{0\le s\le t}|\mathbb{IIII}_{2,s}|=o_p(n^{-1/4})+o_p\left(\eta_n^{r/4}\rho_n^{1/2-\beta/4}\right). \quad \textbf{(A.55)}$$

Next we estimate $\mathbb{IIII}_{3,t}$. By the Schwarz inequality and (A.2) we have

$$\sup_{0\le s\le t}|\mathbb{IIII}_{3,s}|\lesssim\frac{1}{k_n}\left\{\sum_{i:\widehat{S}^{i+k_n}\le t}|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2\le4\varrho^1_n[i]\}}\right\}^{1/2}$$
$$\times\left\{\sum_{j:\widehat{T}^{j+k_n}\le t}|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j|^2 1_{\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho^2_n[j],|\bar{L}^2(\widehat{\mathcal{I}})^i_t|^2>4\varrho^2_n[j]\}}\right\}^{1/2}.$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Note that on $\left\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \leq \varrho_n^2[j], |\bar{L}^2(\widehat{\mathcal{J}})_t^j|^2 > 4\varrho_n^2[j]\right\}$ we have $|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j| > |\bar{L}^2(\widehat{\mathcal{J}})_t^j| - |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j| > \sqrt{\varrho_n^2[j]}$, Lemma A.4.6, Lemma A.4.7, $k_n\bar{r}_n = o(\rho_n)$ and $k_n^{-1} = o(\rho_n)$ yield

$$\sup_{0\leq s\leq t} |\mathbb{III}_{3,s}| = O_p\left(k_n^{-1/2}\rho_n^{1/2-\beta/4}\right) + o_p\left(k_n^{-1/2}\rho_n^{1/2-\beta/2}\eta_n^{\frac{r-2}{4}}\right)$$
$$= o_p(n^{-1/4}) + o_p\left(\rho_n^{1-\beta/2}\right). \tag{A.56}$$

Finally we estimate $\mathbb{III}_{4,t}$. First we specify (ε_n). By Lemma A.4.2 we can choose the sequence (ε_n) satisfying (A.21) for $p = 2$ and (A.22). Next we decompose the target quantity as

$$\begin{aligned}\mathbb{III}_{4,t} =& \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\leq t} \bar{L}^1(\widehat{\mathcal{I}})_t^i \overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j \bar{K}^{ij}\\
&\times 1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho_n^1[i], |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\leq\varrho_n^2[j], |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2\leq 4\varrho_n^1[i], |\bar{L}^2(\widehat{\mathcal{J}})_t^j|^2\leq 4\varrho_n^2[j]\right\}}\\
&\times\left(1_{\left\{N^2(\bar{J}^j)_t\neq 0\right\}} + 1_{\left\{N^2(\bar{J}^j)_t=0, \tilde{N}^1(\bar{I}^i)_t\neq 0\right\}} + 1_{\left\{N^2(\bar{J}^j)_t=\tilde{N}^1(\bar{I}^i)_t=0\right\}}\right)\\
=:& \mathbb{III}_{4,t}^{(1)} + \mathbb{III}_{4,t}^{(2)} + \mathbb{III}_{4,t}^{(3)}.\end{aligned}$$

By Lemma A.3.4, we obtain $\sup_{0\leq s\leq t}|\mathbb{III}_{4,s}^{(1)}| = o_p(n^{-1/4})$. Moreover, on $\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \leq \varrho_n^2[j], |\bar{L}^2(\widehat{\mathcal{J}})^j|^2 \leq 4\varrho_n^2[j]\}$ we have $|\overline{\mathsf{Y}}^2(\widehat{\mathcal{J}})^j| \leq |\bar{L}^2(\widehat{\mathcal{J}})^j| + |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j| \leq 3\sqrt{\varrho_n^2[j]}$, and thus by (A.2), Lemma A.4.3, and (A.22) we have

$$\sup_{0\leq s\leq t}|\mathbb{III}_{4,s}^{(2)}| \lesssim k_n^{-1}\rho_n\sum_{i=1}^{\infty} 1_{\{\tilde{N}^1(\bar{I}^i)\neq 0\}} = o_p\left(\rho_n^{1-\beta/2}\right).$$

On the other hand, since

$$\begin{aligned}\mathbb{III}_{4,t}^{(3)} =& \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\leq t} \bar{L}^1(\widehat{\mathcal{I}})_t^i \overline{\mathsf{X}}^2(\widehat{\mathcal{J}})^j \bar{K}^{ij} 1_{\left\{N^2(\bar{J}^j)_t=\tilde{N}^1(\bar{I}^i)_t=0\right\}}\\
&\times 1_{\left\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho_n^1[i], |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\leq\varrho_n^2[j], |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2\leq 4\varrho_n^1[i], |\bar{L}^2(\widehat{\mathcal{J}})_t^j|^2\leq 4\varrho_n^2[j]\right\}},\end{aligned}$$

we can decompose it as

$$\begin{aligned}\mathbb{III}_{4,t}^{(1)} =& \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\leq t} \bar{L}^1(\widehat{\mathcal{I}})_t^i\\
&\times\left[\left\{\bar{A}^2(\widehat{\mathcal{J}})_t^j + \widetilde{\underline{A}^2}(\widehat{\mathcal{J}})_t^j\right\} + \bar{M}^2(\widehat{\mathcal{J}})_t^j + \widetilde{\underline{M}^2}(\widehat{\mathcal{J}})_t^j + \bar{\zeta}^2(\widehat{\mathcal{J}})^j\right]\bar{K}^{ij}\\
&\times 1_{\left\{N^2(\bar{J}^j)_t=\tilde{N}^1(\bar{I}^i)_t=0, |\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\leq\varrho_n^1[i], |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\leq\varrho_n^2[j], |\bar{L}^1(\widehat{\mathcal{I}})_t^i|^2\leq 4\varrho_n^1[i], |\bar{L}^2(\widehat{\mathcal{J}})_t^j|^2\leq 4\varrho_n^2[j]\right\}}\\
=:& \Gamma_{1,t} + \Gamma_{2,t} + \Gamma_{3,t} + \Gamma_{4,t}.\end{aligned}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

First consider $\Gamma_{1,t}$. By the Schwarz inequality and (A.2), we have

$$\sup_{0\le s\le t}|\Gamma_{1,s}| \lesssim \frac{1}{k_n}\left\{\sum_{i:\widehat{S}^{i+k_n}\le t}|\bar{L}^1(\widehat{\mathcal{I}})^i|^2 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i|^2\le 4\varrho_n^1[i]\}}\right\}^{1/2} \times\left\{\sum_{j:\widehat{T}^{j+k_n}\le t}|\bar{A}^2(\widehat{\mathcal{J}})^j+\widetilde{\underline{A}^2}(\widehat{\mathcal{J}})^j|^2\right\}^{1/2},$$

hence Lemma A.4.6, the boundedness of $(A^2)'$ and $(\underline{A}^2)'$, the Lipschitz continuity of g and [SA6] yield

$$\sup_{0\le s\le t}|\Gamma_{1,s}| = o_p\left(k_n^{-1/2}\rho_n^{1/2-\beta/4}\right) = o_p\left(n^{-1/4}\right). \tag{A.57}$$

Next consider $\Gamma_{2,t}$. Since $\bar{L}^1(\widehat{\mathcal{I}})^i_t = \bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})^i_t = \bar{\Xi}^1(\widehat{\mathcal{I}})^i_t + \bar{\Theta}^1(\widehat{\mathcal{I}})^i_t$ on $\{\tilde{N}^2(\bar{J}^j)_t = 0\}$, we can decompose the target quantity as

$$\begin{aligned}\Gamma_{2,t} &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\le t}\left\{\bar{\Xi}^1(\widehat{\mathcal{I}})^i_t\bar{M}^2(\widehat{\mathcal{J}})^j_t + \bar{\Theta}^1(\widehat{\mathcal{I}})^i_t\bar{M}^2(\widehat{\mathcal{J}})^j_t\right\}\bar{K}^{ij}\\ &\quad\times 1_{\left\{N^2(\bar{J}^j)_t=\tilde{N}^2(\bar{J}^j)_t=0,|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2\le\varrho_n^1[i],|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2\le\varrho_n^2[j],|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2\le 4\varrho_n^1[i],|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2\le 4\varrho_n^2[j]\right\}}\\ &=: \Gamma^{(1)}_{2,t}+\Gamma^{(2)}_{2,t}.\end{aligned}$$

First estimate $\Gamma^{(1)}_{2,t}$. We decompose it further as

$$\begin{aligned}\Gamma^{(1)}_{2,t} &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j:\bar{R}^\vee(i,j)\le t}\bar{\Xi}^1(\widehat{\mathcal{I}})^i_t\bar{M}^2(\widehat{\mathcal{J}})^j_t\bar{K}^{ij}\\ &\quad\times\left(1-1_{\substack{\{N^2(\bar{J}^j)_t\neq 0\}\cup\{\tilde{N}^2(\bar{J}^j)_t\neq 0\}\cup\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]\}\\ \cup\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho_n^1[i]\}\cup\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j]\}\cup\{|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho_n^2[j]\}}}\right)\\ &=: \Gamma^{(1)\prime}_{2,t}+\Gamma^{(1)\prime\prime}_{2,t}.\end{aligned}$$

By Lemma A.4.11 we have $\sup_{0\le s\le t}|\Gamma^{(1)\prime}_{2,s}| = o_p(n^{-1/4})$. On the other hand, (A.40), the Schwarz inequality, (A.2), Lemma A.4.3, Lemma A.4.5, (A.51), and (A.22) yield

$$\sup_{0\le s\le t}|\Gamma^{(1)\prime\prime}_{2,s}| = o_p\left(k_n^{-1}\varphi_2(\varepsilon_n)^{1/2}k_n^{1/2}\rho_n^{-\beta/4}\right)+O_p\left(k_n^{-1}\varphi_2(\varepsilon_n)^{1/2}n^{1/2}\eta_n^{r/4}\right).$$

Since we have

$$\varphi_2(\varepsilon_n)\le\varepsilon_n^{-\beta}\varepsilon_n^2\varphi_\beta(\varepsilon_n)\le\rho_n^2\varepsilon_n^{-2-\beta} = o\left(\rho_n^{1-\beta/2}\right) \tag{A.58}$$

due to (A.21) for $p=2$ and (A.22), note that $n^{-1}=o(\rho_n^2)$, we obtain $\sup_{0\le s\le t}|\Gamma^{(1)\prime\prime}_{2,s}| = o_p\left(n^{-1/4}\rho_n^{1/2-\beta/2}\right)+o_p\left(\rho_n^{1/2-\beta/4}\eta_n^{r/4}\right) = o_p\left(\rho_n^{1-\beta/2}\right)+o_p\left(\eta_n^{r/2}\right)$. Consequently, we conclude that

$$\sup_{0\le s\le t}|\Gamma^{(1)}_{2,s}| = o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right)+o_p\left(\eta_n^{r/2}\right). \tag{A.59}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

Next estimate $\Gamma^{(2)}_{2,t}$. We decompose it further as

$$\Gamma^{(2)}_{2,t} = \frac{1}{(\psi_{HY}k_n)^2} \sum_{i,j:\bar{R}^\vee(i,j)\le t} \bar{\Theta}^1(\widehat{\mathcal{I}})^i_t \bar{M}^2(\widehat{\mathcal{J}})^j_t \bar{K}^{ij}$$

$$\times \left(1 - 1_{\{N^2(\bar{J}^j)_t \neq 0\}\cup\{\tilde{N}^2(\bar{J}^j)_t\neq 0\}\cup\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 > \varrho^1_n[i]\} \cup\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i]\}\cup\{|\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 > \varrho^2_n[j]\}\cup\{|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2 > 4\varrho^2_n[j]\}}\right)$$

$$=: \Gamma^{(2)\prime}_{2,t} + \Gamma^{(2)\prime\prime}_{2,t}.$$

Lemma A.4.12 yields $\sup_{0\le s\le t}|\Gamma^{(3)\prime}_{2,s}| = o_p\left(n^{-1/4}\right) + o_p\left(\rho_n^{1-\beta/2}\right)$. On the other hand, since $\bar{\Theta}^1(\widehat{\mathcal{I}})^i_t \bar{M}^2(\widehat{\mathcal{J}})^j_t \bar{K}^{ij}_t = \bar{K}^{ij}_- \bar{\Theta}^1(\widehat{\mathcal{I}})^i_- \bullet \bar{M}^2(\widehat{\mathcal{J}})^j_t + \bar{K}^{ij}_- \bar{M}^2(\widehat{\mathcal{J}})^j_- \bullet \bar{\Theta}^1(\widehat{\mathcal{I}})^i_t$ by integration by parts and Lemma A.1.3, we have

$$E\left[|\bar{\Theta}^1(\widehat{\mathcal{I}})^i_t \bar{M}^2(\widehat{\mathcal{J}})^j_t \bar{K}^{ij}_t|^2\right] \lesssim k_n^2 \bar{r}_n \varepsilon_n^{-2(\beta-1)_+} E\left[\sum_{p=1}^{k_n-1} |\widehat{I}^{i+p}(t)|^2 \bar{K}^{ij}_t\right]$$

by [A1], the optional sampling theorem and the inequality $|\Theta^1(\widehat{I}^p)_t|^2 \lesssim \varepsilon_n^{-2(\beta-1)_+}|\widehat{I}^p(t)|^2$ and (A.14). Therefore, (A.2) and [SA6] yield

$$E\left[\sum_{i,j:\bar{R}^\vee(i,j)\le t} |\bar{\Theta}^1(\widehat{\mathcal{I}})^i_t \bar{M}^2(\widehat{\mathcal{J}})^j_t \bar{K}^{ij}|^2\right] \lesssim k_n^4 \bar{r}_n \varepsilon_n^{-2(\beta-1)_+} n^{-1} \lesssim k_n^2 \bar{r}_n \varepsilon_n^{-2(\beta-1)_+}$$

and thus an argument similar to the above yields

$$\begin{aligned}\sup_{0\le s\le t}|\Gamma^{(2)\prime\prime}_{2,s}| &= o_p\left(k_n^{-1/2}\bar{r}_n^{1/2}\varepsilon_n^{-(\beta-1)_+}k_n^{1/2}\rho_n^{-\beta/4}\right) + O_p\left(k_n^{-1/2}\bar{r}_n^{1/2}\varepsilon_n^{-(\beta-1)_+}n^{1/2}\eta_n^{r/4}\right)\\ &= o_p\left(\rho_n^{1-\beta/2}\right) + o_p\left(\eta_n^{r/2}\right).\end{aligned}$$

Consequently, we conclude that $\sup_{0\le s\le t}|\Gamma^{(2)}_{2,s}| = o_p\left(n^{-1/4}\right) + o_p\left(\eta_n^{r/2}\right) + o_p\left(\rho_n^{1-\beta/2}\right)$. Combining this result with (A.59), we conclude

$$\sup_{0\le s\le t}|\Gamma_{2,s}| = o_p\left(n^{-1/4}\right) + o_p\left(\eta_n^{r/2}\right) + o_p\left(\rho_n^{1-\beta/2}\right). \tag{A.60}$$

Similarly we can also show that

$$\sup_{0\le s\le t}|\Gamma_{3,s}| = o_p\left(n^{-1/4}\right) + o_p\left(\eta_n^{r/2}\right) + o_p\left(\rho_n^{1-\beta/2}\right). \tag{A.61}$$

Now we deal with $\Gamma_{4,t}$. Since $\bar{L}^1(\widehat{\mathcal{I}})^i_t = \bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})^i_t$ on $\{\tilde{N}^2(\bar{J}^j)_t = 0\}$, we can decompose the target quantity as

$$\Gamma_{4,t} = \frac{1}{(\psi_{HY}k_n)^2} \sum_{i,j:\bar{R}^\vee(i,j)\le t} \bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})^i_t \bar{\zeta}^2(\widehat{\mathcal{J}})^j \bar{K}^{ij}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

$$\times\left(\begin{array}{l}1-1_{\{N^2(\bar{J}^j)_t\neq 0\}\cup\{\tilde{N}^2(\bar{J}^j)_t\neq 0\}\cup\{|\overline{\mathbb{Z}}^1(\widehat{\mathcal{I}})^i|^2>\varrho_n^1[i]\}}\\ \qquad\cup\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2>4\varrho_n^1[i]\}\cup\{|\overline{\mathbb{Z}}^2(\widehat{\mathcal{J}})^j|^2>\varrho_n^2[j]\}\cup\{|\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2>4\varrho_n^2[j]\}\}\end{array}\right)$$
$$=: \Gamma^{(1)}_{4,t}+\Gamma^{(2)}_{4,t}.$$

By Lemma A.4.11 and Lemma A.4.12 we have $\sup_{0\leq s\leq t}|\Gamma^{(1)}_{4,s}|=o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right)$. On the other hand, by the Lipschitz continuity of g, (A.2) and [SN$^\flat_2$], we have

$$E_0\left[\sum_{i,j:\bar{R}^\vee(i,j)\leq t}|\bar{\mathfrak{L}}^1(\widehat{\mathcal{I}})^i_t|^2|\bar{\zeta}^2(\widehat{\mathcal{J}})^j|^2\bar{K}^{ij}\right]\lesssim\sum_{i=1}^\infty|\bar{\mathfrak{L}}^{(1)}(\widehat{\mathcal{I}})^i_t|^2,$$

hence the Schwarz inequality, (A.2), Lemma A.4.3, Lemma A.4.5, (A.51), and (A.22) yield

$$\begin{aligned}\sup_{0\leq s\leq t}|\Gamma^{(2)}_{4,s}|=&o_p\left(k_n^{-1}\varepsilon_n^{1-\beta/2}\varphi_\beta(\varepsilon_n)^{1/2}k_n^{1/2}\rho_n^{-\beta/4}\right)\\&+O_p\left(k_n^{-1}\varepsilon_n^{1-\beta/2}\varphi_\beta(\varepsilon_n)^{1/2}n^{1/2}\eta_n^{r/4}\right).\end{aligned}$$

Since we have (A.58) due to (A.21) for $p=2$ and (A.22), note that $n^{-1}=o(\rho_n^2)$, we obtain

$$\begin{aligned}\sup_{0\leq s\leq t}|\Gamma^{(2)}_{4,s}|=&o_p\left(n^{-1/4}\rho_n^{1/2-\beta/2}\right)+o_p\left(\rho_n^{1/2-\beta/4}\eta_n^{r/4}\right)\\=&o_p\left(\rho_n^{1-\beta/2}\right)+o_p\left(\eta_n^{r/2}\right).\end{aligned}$$

Consequently, we obtain

$$\sup_{0\leq s\leq t}|\Gamma_{4,s}|=o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right)+o_p\left(\eta_n^{r/2}\right).\tag{A.62}$$

By (A.57), (A.60), and (A.61), we obtain $\sup_{0\leq s\leq t}|\mathbb{III}^{(3)}_{4,s}|=o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right)+o_p\left(\eta_n^{r/2}\right)$. Consequently, we obtain

$$\sup_{0\leq s\leq t}|\mathbb{III}_{4,s}|=o_p\left(n^{-1/4}\right)+o_p\left(\rho_n^{1-\beta/2}\right)+o_p\left(\eta_n^{r/2}\right).\tag{A.63}$$

Note that $\eta_n=o(1)$ and $\beta\geq 0$, (A.54), (A.55), (A.56), and (A.63) yield

$$\sup_{0\leq s\leq t}|\mathbb{III}_s|=o_p\left(n^{-1/4}\right)+o_p\left(\eta_n^{\frac{r-2}{2}}\right)+o_p\left(\rho_n^{1-\beta/2}\right).\tag{A.64}$$

(d) By symmetry, we obtain

$$\sup_{0\leq s\leq t}|\mathbb{IV}_s|=o_p\left(n^{-1/4}\right)+o_p\left(\eta_n^{\frac{r-2}{2}}\right)+o_p\left(\rho_n^{1-\beta/2}\right).\tag{A.65}$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

(e) Finally we consider $\mathbb{V}$. We decompose it as

$$
\begin{aligned}
\mathbb{V}_t = \frac{1}{(\psi_{HY} k_n)^2} \sum_{i,j:\bar{R}^\vee(i,j)\le t} & \bar{L}^1(\widehat{\mathcal{I}})^i_t \bar{L}^2(\widehat{\mathcal{J}})^j_t \bar{K}^{ij} 1_{\{|\overline{\mathsf{Z}}^1(\widehat{\mathcal{I}})^i|^2 \le \varrho^1_n[i], |\overline{\mathsf{Z}}^2(\widehat{\mathcal{J}})^j|^2 \le \varrho^2_n[j]\}} \\
\times \Big(& 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i], |\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2 > 4\varrho^2_n[j]\}} + 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 > 4\varrho^1_n[i], |\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2 \le 4\varrho^2_n[j]\}} \\
& + 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 \le 4\varrho^1_n[i], |\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2 > 4\varrho^2_n[j]\}} + 1_{\{|\bar{L}^1(\widehat{\mathcal{I}})^i_t|^2 \le 4\varrho^1_n[i], |\bar{L}^2(\widehat{\mathcal{J}})^j_t|^2 \le 4\varrho^2_n[j]\}} \Big) \\
=: \mathbb{V}_{1,t} + \mathbb{V}_{2,t} + \mathbb{V}_{3,t} + \mathbb{V}_{4,t}. &
\end{aligned}
$$

By the Schwarz inequality, (A.2) and Lemma A.4.9 we have $\sup_{0\le s\le t} |\mathbb{V}_{1,s}| = o_p\left(n^{-1/4}\right) + O_p\left(\eta_n^{\frac{r}{2}}\right)$. Moreover, by (A.2) and Lemma A.4.8 we have $\sup_{0\le s\le t} |\mathbb{V}_{2,s}| = o_p\left(n^{-1/4}\right) + o_p\left(\eta_n^{r/4} \rho_n^{1/2-\beta/4}\right)$ and $\sup_{0\le s\le t} |\mathbb{V}_{3,s}| = o_p\left(n^{-1/4}\right) + o_p\left(\eta_n^{r/4} \rho_n^{1/2-\beta/4}\right)$. Furthermore, the Schwarz inequality, (A.2) and Lemma A.4.6 yield $\sup_{0\le s\le t} |\mathbb{V}_{4,s}| = o_p\left(\rho_n^{1-\beta/2}\right)$. Consequently, we obtain

$$
\sup_{0\le s\le t} |\mathbb{V}_s| = o_p\left(n^{-1/4}\right) + O_p\left(\eta_n^{r/2}\right) + o_p\left(\rho_n^{1-\beta/2}\right). \tag{A.66}
$$

Note that $\eta_n = o(1)$, (A.45), (A.46), (A.53), (A.64), (A.65), and (A.66) yield

$$
\sup_{0\le s\le t} |PTHY(\mathsf{Z}^1,\mathsf{Z}^2)^n_s - PHY(\mathsf{X}^1,\mathsf{X}^2)^n_s| = o_p\left(n^{-1/4}\right) + O_p\left(\eta_n^{\frac{r-2}{2}}\right) + o_p\left(\rho_n^{1-\beta/2}\right).
$$

Since $\eta_n = (k_n\rho_n)^{-1} = O(n^{-\frac{1}{2}}\rho_n^{-1})$, we complete the proof of Theorem 3.4. ■

A.5. Proof of Proposition 4.1

By a localization procedure, we may assume that [SC1]-[SC2], [SA4], [SA6], and [$\mathrm{SN}_2^\flat$] hold. In a similar manner we may also assume that $E^1 = E^2 =: E$ and that there is a nonnegative bounded measurable function ψ on E such that

$$
\sup_{\omega\in\Omega, t\in\mathbb{R}_+} |\delta^l(\omega,t,x)| \le \psi(x) \quad \text{and} \quad \int_E \psi(x)^2 F^l(\mathrm{d}x) < \infty, \qquad l = 1,2.
$$

Under the above assumption we can define the process L'^l by $L'^l = \delta^l \star (\mu^l - \nu^l)$ for each $l = 1,2$. Then, for each $l = 1,2$ we have $Z^l = X'^l + L'^l$, where $X'^l_t = X^l_t - \int_0^t \int_E \kappa'(\delta^l(s,x))\mathrm{d}s\, F(\mathrm{d}x)$. Hence we can decompose the target quantity as

$$
\begin{aligned}
PHY(\mathsf{Z}^1,\mathsf{Z}^2)^n_t = \frac{1}{(\psi_{HY}k_n)^2} \sum_{i,j:\bar{R}^\vee(i,j)\le t} & \Big\{ \overline{\mathsf{X}}'^1(\widehat{\mathcal{I}})^i \overline{\mathsf{X}}'^2(\widehat{\mathcal{J}})^j + \overline{\mathsf{X}}'^1(\widehat{\mathcal{I}})^i \bar{L}'^2(\widehat{\mathcal{J}})^j_t \\
& + \bar{L}'^1(\widehat{\mathcal{I}})^i_t \overline{\mathsf{X}}'^2(\widehat{\mathcal{J}})^j + \bar{L}'^1(\widehat{\mathcal{I}})^i_t \bar{L}'^2(\widehat{\mathcal{J}})^j_t \Big\} \bar{K}^{ij} \\
=: \mathbb{I}_t + \mathbb{II}_t + \mathbb{III}_t + \mathbb{IV}_t, &
\end{aligned}
$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

where $\overline{\mathsf{X}}'^1(\widehat{\mathcal{I}})^i = \sum_{p=1}^{k_n-1} g\left(\frac{p}{k_n}\right)\left(\mathsf{X}'^1_{\widehat{S}^{i+p}} - \mathsf{X}'^1_{\widehat{S}^{i+p-1}}\right)$, $\overline{\mathsf{X}}'^2(\widehat{\mathcal{J}})^j = \sum_{q=1}^{k_n-1} g\left(\frac{q}{k_n}\right)\left(\mathsf{X}'^2_{\widehat{T}^{j+q}} - \mathsf{X}'^2_{\widehat{T}^{j+q-1}}\right)$ and $\mathsf{X}'^1_{\widehat{S}^i} = X'^1_{\widehat{S}^i} + U^1_{\widehat{S}^i}$, $\mathsf{X}'^2_{\widehat{T}^j} = X'^2_{\widehat{T}^j} + U^2_{\widehat{T}^j}$ for any i, j.

First, we can adopt an argument similar to the proof of Lemma A.4.11 for the proof of $\mathbb{II}_t = O_p(n^{-1/4})$ and $\mathbb{III}_t = O_p(n^{-1/4})$. Next, combining Lemma A.1.1–A.1.2 with an argument similar to the proof of Lemma A.4.11 we can show that $\mathbb{I}_t = [X'^1, X'^2]_t + O_p(n^{-1/4})$. Finally we consider $\mathbb{IV}_t$. By integration by parts we can decompose it as

$$\begin{aligned}\mathbb{IV}_t &= \frac{1}{(\psi_{HY}k_n)^2}\\ &\quad\times \sum_{i,j:\bar{R}^\vee(i,j)\leq t}\left\{\bar{L}'^1(\widehat{\mathcal{I}})^i_- \bullet \bar{L}'^2(\widehat{\mathcal{J}})^j_t + \bullet\bar{L}'^2(\widehat{\mathcal{J}})^j_- \bullet \bar{L}'^1(\widehat{\mathcal{I}})^i_t + [\bar{L}'^1(\widehat{\mathcal{I}})^i, \bar{L}'^2(\widehat{\mathcal{J}})^j]_t\right\}\bar{K}^{ij}\\ &=: \mathbb{IV}^{(1)}_t + \mathbb{IV}^{(4)}_t + \mathbb{IV}^{(3)}_t.\end{aligned}$$

By an argument similar to the proof of Lemma A.4.11 we can show that $\mathbb{IV}^{(1)}_t = O_p(n^{-1/4})$ and $\mathbb{IV}^{(2)}_t = O_p(n^{-1/4})$. On the other hand, an argument similar to the proof of Lemma A.1.2 yields

$$\begin{aligned}\mathbb{IV}^{(3)}_t &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{p,q=1}^{\infty}\left(\sum_{i=(p-k_n+1)\vee 1}^{p}\sum_{j=(q-k_n+1)\vee 1}^{q} g^n_{p-i}g^n_{q-j}\bar{K}^{ij}1_{\{\bar{R}^\vee(i,j)\leq t\}}\right)\\ &\quad\times\left(\widehat{I}^p_-\widehat{J}^q_-\right)\bullet[L'^1, L'^2]_t.\end{aligned}$$

and

$$\begin{aligned}[L'^1, L'^2]_t &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{p,q=1}^{\infty}\left(\sum_{i=(p-k_n+1)\vee 1}^{p}\sum_{j=(q-k_n+1)\vee 1}^{q} g^n_{p-i}g^n_{q-j}\bar{K}^{ij}_t\right)\\ &\quad\times\left(\widehat{I}^p_-\widehat{J}^q_-\right)\bullet[L'^1, L'^2]_t + O_p\left(k_n^{-1}\right)\\ &=: \mathbf{IV}^{(3)}_t + O_p\left(k_n^{-1}\right).\end{aligned}$$

Since we have

$$\begin{aligned}\mathbb{IV}^{(3)}_t - \mathbf{IV}^{(3)}_t &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{p,q=1}^{\infty}\left(\sum_{i=(p-k_n+1)\vee 1}^{p}\sum_{j=(q-k_n+1)\vee 1}^{q} g^n_{p-i}g^n_{q-j}\bar{K}^{ij}_t 1_{\{\bar{R}^\vee(i,j)>t\}}\right)\\ &\quad\times\left(\widehat{I}^p_-\widehat{J}^q_-\right)\bullet[L'^1, L'^2]_t + O_p(n^{-1/2})\\ &= \frac{1}{(\psi_{HY}k_n)^2}\sum_{i,j=0}^{\infty}\bar{K}^{ij}_t 1_{\{\bar{R}^\vee(i,j)>t\}}\sum_{p,q=0}^{k_n-1} g^n_p g^n_q\left(\widehat{I}^{i+p}_-\widehat{J}^{j+q}_-\right)\bullet[L'^1, L'^2]_t + O_p(n^{-1/2}),\end{aligned}$$

we obtain $\left|\mathbb{IV}^{(3)}_t - \mathbf{IV}^{(3)}_t\right| \lesssim \frac{1}{k_n^2}\sum_{i,j=0}^{\infty}\bar{K}^{ij}_t 1_{\{\bar{R}^\vee(i,j)>t\}}(\bar{I}^i_-\bar{J}^j_-)\bullet\left|[L'^1, L'^2]\right|_t + O_p(n^{-1/2})$, and thus the Kunita–Watanabe inequality and the inequality of arithmetic and geometric means yield

$$\left|\mathbb{IV}^{(3)}_t - \mathbf{IV}^{(3)}_t\right| \lesssim \frac{1}{k_n^2}\sum_{i,j=0}^{\infty}\bar{K}^{ij}_t 1_{\{\bar{R}^\vee(i,j)>t\}}\left\{[L'^1](\bar{I}^i)_t + [L'^2](\bar{J}^j)_t\right\} + O_p(n^{-1/2}).$$

Since $\bar{K}_t^{ij} 1_{\{\bar{R}^\vee(i,j)>t\}}$ is $\mathcal{F}_{\widehat{S}^i \wedge \widehat{T}^j}$-measurable by Lemma A.4.10, we obtain

$$E\left[\left|\sum_{i,j=0}^{\infty} \bar{K}_t^{ij} 1_{\{\bar{R}^\vee(i,j)>t\}} \left\{[L'^1](\bar{I}^i)_t + [L'^2](\bar{J}^j)_t\right\}\right|\right]$$
$$\lesssim E\left[\sum_{i,j=0}^{\infty} \bar{K}_t^{ij} 1_{\{\bar{R}^\vee(i,j)>t\}} \left\{\langle L'^1\rangle(\bar{I}^i)_t + \langle L'^2\rangle(\bar{J}^j)_t\right\}\right],$$

hence by [SK$_2$], [SA4], and (A.41) we conclude that $\left|\mathbb{IV}_t^{(3)} - \mathbf{IV}_t^{(3)}\right| = O_p(k_n \bar{r}_n) + O_p(n^{-1/2}) = o_p(n^{-1/4})$. Consequently, we obtain $\mathbb{IV}_t^{(3)} = [L'^1, L'^2]_t + o_p(n^{-1/4})$, and thus we complete the proof of the proposition because $[Z^1, Z^2] = [X'^1, X'^2] + [L'^1, L'^2]$. ■

A.6. Proof of Proposition 4.2

By a localization procedure, we may systematically replace the conditions [C1]–[C2], [A4], [A6], [N$_2^\flat$], [T], and [N$_r^\flat$] with [SC1]–[SC2], [SA4], [SA6], [SN$_2^\flat$], [ST], and [SN$_r^\flat$] respectively.

Set $\tilde{\lambda}_u^l = \sum_{v=u}^{\infty} \lambda_v^l$ for each $u \in \mathbb{Z}_+$ and $l = 1, 2$. We define the random variables $\tilde{\zeta}_i^1$ and $\tilde{\zeta}_j^2$ by $\tilde{\zeta}_i^1 = \sum_{u=0}^{i} \tilde{\lambda}_{u+1}^1 \zeta_{S^{i-u}}^1$ and $\tilde{\zeta}_j^2 = \sum_{u=0}^{i} \tilde{\lambda}_{u+1}^2 \zeta_{T^{j-u}}^2$ for any i, j. Then we have

$$\tilde{\zeta}_i^1 - \tilde{\zeta}_{i-1}^1 = \sum_{u=0}^{i} \tilde{\lambda}_{u+1}^1 \zeta_{S^{i-u}}^1 - \sum_{u=1}^{i} \tilde{\lambda}_u^1 \zeta_{S^{i-u}}^1 = \sum_{u=0}^{i} (\tilde{\lambda}_{u+1}^1 - \tilde{\lambda}_u^1) \zeta_{S^{i-u}}^1 + \tilde{\lambda}_0 \zeta_{S^i}^1$$
$$= -\sum_{u=0}^{i} \lambda_u \zeta_{S^{i-u}}^1 + \tilde{\lambda}_0 \zeta_{S^i}^1,$$

hence we obtain

$$\sum_{u=0}^{i} \lambda_u \zeta_{S^{i-u}}^1 = \tilde{\lambda}_0 \zeta_{S^i}^1 - (\tilde{\zeta}_i^1 - \tilde{\zeta}_{i-1}^1). \tag{A.67}$$

In the time series analysis, this relation is known as the Beveridge–Nelson decomposition. See Beveridge and Nelson (1981) for details. Combining (A.67) with Abel's partial summation formula, we obtain

$$\sum_{p=0}^{k_n-1} \Delta(g)_p^n \left(\sum_{u=0}^{i+p} \lambda_u \zeta_{S^{i+p-u}}^1\right) = \tilde{\lambda}_0 \overline{\zeta}^1 (\widehat{\mathcal{I}})^i - \sum_{p=0}^{k_n-1} \Delta(g)_p^n \left(\tilde{\zeta}_{i+p}^1 - \tilde{\zeta}_{i+p-1}^1\right)$$
$$= \tilde{\lambda}_0 \overline{\zeta}^1 (\widehat{\mathcal{I}})^i + \sum_{p=0}^{k_n-1} \left\{\Delta(g)_{p+1}^n - \Delta(g)_p^n\right\} \left(\tilde{\zeta}_{i+p}^1 - \tilde{\zeta}_{i-1}^1\right).$$

for any i. Similarly we can deduce

$$\sum_{q=0}^{k_n-1} \Delta(g)_q^n \left(\sum_{u=0}^{j+q} \lambda_u^2 \zeta_{T^{j+q-u}}^2\right) = \tilde{\lambda}_0^2 \overline{\zeta}^2 (\widehat{\mathcal{J}})^j + \sum_{q=0}^{k_n-1} \left\{\Delta(g)_{q+1}^n - \Delta(g)_q^n\right\} \left(\tilde{\zeta}_{j+q}^2 - \tilde{\zeta}_{j-1}^2\right),$$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

$$\sum_{p=0}^{k_n-1} \Delta(g)_p^n \left(\sum_{u=0}^{i+p} \mu_u^1 \sqrt{n} \underline{X}^1 (I^{i+p-u})_t \right)$$
$$= \tilde{\mu}_0^1 \widetilde{\underline{X}^1}(\widehat{\mathcal{I}})_t^i + \sqrt{n} \sum_{p=0}^{k_n-1} \{\Delta(g)_{p+1}^n - \Delta(g)_p^n\}\big(\underline{X}^1(I^{i+p})_t - \underline{X}^1(I^{i-1})_t\big),$$
$$\sum_{q=0}^{k_n-1} \Delta(g)_q^n \left(\sum_{u=0}^{j+q} \mu_u^2 \sqrt{n} \underline{X}^2 (J^{j+q-u})_t \right)$$
$$= \tilde{\mu}_0^2 \widetilde{\underline{X}^2}(\widehat{\mathcal{J}})_t^j + \sqrt{n} \sum_{q=0}^{k_n-1} \{\Delta(g)_{q+1}^n - \Delta(g)_q^n\}\big(\underline{X}^2(J^{j+q})_t - \underline{X}^2(J^{j-1})_t\big)$$

for every i, j, where $\tilde{\mu}_u^l = \sum_{v=u}^{\infty} \mu_v^l$ for each $u \in \mathbb{Z}_+$ and $l = 1, 2$. Note that $\sum_{u=1}^{\infty} |\tilde{\lambda}_u^l| \leq \sum_{u=1}^{\infty} \sum_{v=u}^{\infty} |\lambda_v^l| = \sum_{v=1}^{\infty} v|\lambda_v^l| < \infty$ and $\sum_{u=1}^{\infty} |\tilde{\mu}_u^l| \leq \sum_{u=1}^{\infty} \sum_{v=u}^{\infty} |\mu_v^l| = \sum_{v=1}^{\infty} v |\mu_v^l| < \infty$ for $l = 1, 2$, by using the above formulas we can show that

$$\widehat{PHY}(\mathsf{Z}^1, \mathsf{Z}^2)_t^n - \widehat{PHY}(\tilde{\mathsf{Z}}^1, \tilde{\mathsf{Z}}^2)_t^n \to^p 0 \quad \text{and} \quad \widehat{PTHY}(\mathsf{Z}^1, \mathsf{Z}^2)_t^n - \widehat{PTHY}(\tilde{\mathsf{Z}}^1, \tilde{\mathsf{Z}}^2)_t^n \to^p 0$$

as $n \to \infty$ for any $t \in \mathbb{R}_+$, where $\tilde{\mathsf{Z}}_{S^i}^1 = Z_{S^i}^1 + \tilde{\lambda}_0^1 \zeta_{S^i}^1 + \tilde{\mu}_0^1 \sqrt{n}(\underline{X}_{S^i}^1 - \underline{X}_{S^{i-1}}^1)$ and $\tilde{\mathsf{Z}}_{T^j}^2 = Z_{T^j}^2 + \tilde{\lambda}_0^2 \zeta_{T^j}^2 + \tilde{\mu}_0^2 \sqrt{n}(\underline{X}_{T^j}^2 - \underline{X}_{T^{j-1}}^2)$ for any i, j. Consequently, we complete the proof of the proposition due to Proposition 4.1 and Corollary 3.5. □

A.7. List of assumptions

We list below the assumptions used in this paper. They are listed in alphabetical order. Numbers ξ, ξ', β, and r appearing in the following always satisfy $\frac{1}{2} < \xi < 1$, $0 < \xi' < 1$, $0 \leq \beta \leq 2$, and $r > 0$, respectively. Also, $\mathbf{H}^n$ denotes a filtration of $\mathcal{F}$ to which the processes $\sum_{k=1}^{\infty} 1_{\{R^k \leq t\}}$, $\sum_{k=1}^{\infty} 1_{\{\widehat{S}^k \leq t\}}$ and $\sum_{k=1}^{\infty} 1_{\{\widehat{T}^k \leq t\}}$ are adapted.

[A1] For any $n, i \in \mathbb{N}$, S^i and T^i are $\mathbf{G}^{(n)}$-stopping times, where $\mathbf{G}^{(n)} = (\mathcal{G}_t^{(n)})_{t \in \mathbb{R}_+}$ is the filtration given by $\mathcal{G}_t^{(n)} = \mathcal{F}^{(0)}_{(t - n^{-\xi+1/2})_+}$ for $t \in \mathbb{R}_+$.

[A2] (i) For each n, we have a càdlàg $\mathbf{H}^n$-adapted process G^n and a random subset $\mathcal{N}_n^0$ of $\mathbb{N}$ such that $(\#\mathcal{N}_n^0)_{n \in \mathbb{N}}$ is tight, $G(1)_{R^{k-1}}^n = G_{R^{k-1}}^n$ for any $k \in \mathbb{N} - \mathcal{N}_n^0$, and there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process G satisfying that G and G_- do not vanish and that $G^n \xrightarrow{\text{Sk.p.}} G$ as $n \to \infty$.

(ii) There exists a constant $\rho \geq 1/\xi'$ such that $\big(\sup_{0 \leq s \leq t} G(\rho)_s^n\big)_{n \in \mathbb{N}}$ is tight for all $t > 0$.

(iii) For each n, we have a càdlàg $\mathbf{H}^n$-adapted process χ'^n and a random subset $\mathcal{N}_n'$ of $\mathbb{N}$ such that $(\#\mathcal{N}_n')_{n \in \mathbb{N}}$ is tight, $\chi_{R^{k-1}}^n = \chi'^n_{R^{k-1}}$ for any $k \in \mathbb{N} - \mathcal{N}_n'$, and there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process χ such that $\chi'^n \xrightarrow{\text{Sk.p.}} \chi$ as $n \to \infty$.

(iv) For each n and $l = 1, 2, 1*2$, we have a càdlàg $\mathbf{H}^n$-adapted process $F^{n,l}$ and a random subset $\mathcal{N}_n^l$ of $\mathbb{N}$ such that $(\#\mathcal{N}_n^l)_{n \in \mathbb{N}}$ is tight, $F(1)_{R^{k-1}}^{n,l} = F_{R^{k-1}}^{n,l}$

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press

for any $k \in \mathbb{N} - \mathcal{N}_n^l$, and there exists a càdlàg $\mathbf{F}^{(0)}$-adapted processes F^l satisfying $F^{n,l} \xrightarrow{\text{Sk.p.}} F^l$ as $n \to \infty$.

(v) There exists a constant $\rho' \geq 1/\xi'$ such that $\left(\sup_{0 \leq s \leq t} F(\rho')_s^{n,l}\right)_{n \in \mathbb{N}}$ is tight for all $t > 0$ and $l = 1, 2$.

[A2$^\sharp$] (i) For every $\rho \in [0, 1/\xi']$ there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process $G(\rho)$ such that $G(\rho)^n \xrightarrow{\text{Sk.p.}} G(\rho)$ as $n \to \infty$. Furthermore, G and G_- do not vanish, where $G = G(1)$.

(ii) There exists a càdlàg $\mathbf{F}^{(0)}$-adapted process χ such that $\chi^n \xrightarrow{\text{Sk.p.}} \chi$ as $n \to \infty$.

(iii) For every $l = 1, 2$ and every $\rho' \in [0, 1/\xi']$, there exists a càdlàg $\mathbf{F}^{(0)}$-adapted process $F(\rho)^l$ such that $F(\rho)^{n,l} \xrightarrow{\text{Sk.p.}} F(\rho)^l$ as $n \to \infty$.

(iv) There exists a càdlàg $\mathbf{F}^{(0)}$-adapted process $F(1)^{1*2}$ such that $F(1)^{n,1*2} \xrightarrow{\text{Sk.p.}} F(1)^{1*2}$ as $n \to \infty$.

[A3] For each $V, W = X^1, X^2, \underline{X}^1, \underline{X}^2$, $[V, W]$ is absolutely continuous with a càdlàg derivative, and for the density process $f = [V, W]'$ there is a sequence (σ_k) of $\mathbf{F}^{(0)}$-stopping times such that $\sigma_k \uparrow \infty$ as $k \to \infty$ and for every k and any $\lambda > 0$, we have a positive constant $C_{k,\lambda}$ satisfying

$$E\left[|f_{\tau_1}^{\sigma_k} - f_{\tau_2}^{\sigma_k}|^2 \middle| \mathcal{F}_{\tau_1 \wedge \tau_2}\right] \leq C_{k,\lambda} E\left[|\tau_1 - \tau_2|^{1-\lambda} \middle| \mathcal{F}_{\tau_1 \wedge \tau_2}\right]$$

for any bounded $\mathbf{F}^{(0)}$-stopping times τ_1 and τ_2, and f is adapted to $\mathbf{H}^n$.

[A4] $\xi \vee \frac{9}{10} < \xi'$ and

$$\sup_{i \in \mathbb{Z}_+} [(S^i \wedge t) - (S^{i-1} \wedge t)] \vee \sup_{j \in \mathbb{Z}_+} [(T^j \wedge t) - (T^{j-1} \wedge t)] = o_p(n^{-\xi'})$$

as $n \to \infty$ for any $t \in \mathbb{R}_+$.

[A5] A^1, A^2, $\underline{A}^1$, and $\underline{A}^2$ are absolutely continuous with càdlàg derivatives, and there is a sequence (σ_k) of $\mathbf{F}^{(0)}$-stopping times such that $\sigma_k \uparrow \infty$ as $k \to \infty$ and for every k we have a positive constant C_k and $\lambda_k \in (0, 3/4)$ satisfying

$$E\left[|f_t^{\sigma_k} - f_\tau^{\sigma_k}|^2 \middle| \mathcal{F}_{\tau \wedge t}\right] \leq C_k E\left[|t - \tau|^{1-\lambda_k} \middle| \mathcal{F}_{\tau \wedge t}\right] \tag{A.68}$$

for every $t > 0$ and any bounded $\mathbf{F}^{(0)}$-stopping time τ, for the density processes $f = (A^1)'$, $(A^2)'$, $(\underline{A}^1)'$, and $(\underline{A}^2)'$.

[A6] For each $t \in \mathbb{R}_+$, $nH_n(t) = O_p(1)$ as $n \to \infty$, where $H_n(t) = \sum_{k=1}^\infty |\Gamma^k(t)|^2$.

[C1] $n^{-1} N_t^n = O_p(1)$ as $n \to \infty$ for every t.

[C2] $A^1, A^2, \underline{A}^1, \underline{A}^2$, and $[V, W]$ for $V, W = X^1, X^2, \underline{X}^1, \underline{X}^2$ are absolutely continuous with locally bounded derivatives.

[F] For each $l = 1, 2$ we have $Z_t^l = X_t^l + \sum_{k=1}^{N_t^l} \gamma_k^l$, where X^l is a continuous semimartingale on $\mathcal{B}^{(0)}$ given by (3), N^l is a (simple) point process adopted to $\mathbf{F}^{(0)}$, and $(\gamma_k^l)_{k \in \mathbb{N}}$ is a sequence of nonzero random variables.

[K$_\beta$] For each $l=1,2$, we have

$$Z^l = X^l + \kappa(\delta^l)\star(\mu^l - \nu^l) + \kappa'(\delta^l)\star\mu^l,$$

where

(i) X^l is a continuous semimartingale given by (3).
(ii) μ^l is a Poisson random measure on $\mathbb{R}_+ \times E^l$ with intensity measure $\nu^l(\mathrm{d}t,\mathrm{d}x) = \mathrm{d}t F^l(\mathrm{d}x)$, where $(E^l,\mathcal{E}^l)$ is a Polish space and F^l is a σ-finite measure on $(E^l,\mathcal{E}^l)$.
(iii) $\kappa(x) = x1_{\{|x|\leq 1\}}$ and $\kappa'(x) = x - \kappa(x)$ for each $x \in \mathbb{R}$.
(iv) δ^l is a predictable map from $\mathbf{\Omega}^{(0)} \times \mathbb{R}_+ \times E^l$ into $\mathbb{R}$. Moreover, there are a sequence (R^l_k) of stopping times increasing to ∞ and a sequence (ψ^l_k) of nonnegative measurable functions on E^l such that

$$\sup_{\omega^{(0)}\in\Omega^{(0)},t<R^l_k(\omega^{(0)})} |\delta^l(\omega,t,x)| \leq \psi^l_k(x) \text{ and } \int_{E^l} 1\wedge\psi^l_k(x)^\beta F^l(\mathrm{d}x) < \infty.$$

(v) If $\beta < 1$, for the process $f_t = \int_{E^l}\kappa(\delta^l(t,x))F^l(\mathrm{d}x)$, there is a sequence (σ_k) of $\mathbf{F}^{(0)}$-stopping times such that for every k we have a positive constant C_k and $\lambda_k \in (0,3/4)$ satisfying (A.68) for every $t > 0$ and any bounded $\mathbf{F}^{(0)}$-stopping time τ.

[N$^\flat_r$] $(\int |z|^r Q_t(\mathrm{d}z))_{t\in\mathbb{R}_+}$ is a locally bounded process.

[N$_r$] $(\int |z|^r Q_t(\mathrm{d}z))_{t\in\mathbb{R}_+}$ is a locally bounded process, and the covariance matrix process

$$\Psi_t(\omega^{(0)}) = \int zz^* Q_t(\omega^{(0)},\mathrm{d}z).$$

is càdlàg, quasi-left continuous and adapted to $\mathbf{H}^n$ for every n. Furthermore, there is a sequence (σ^k) of $\mathbf{F}^{(0)}$-stopping times such that $\sigma^k \uparrow \infty$ as $k\to\infty$ and for every k and any $\lambda > 0$, we have a positive constant $C_{k,\lambda}$ satisfying

$$E\left[|\Psi^{ij}_{\sigma^k\wedge t} - \Psi^{ij}_{\sigma^k\wedge(t-h)_+}|^2 \middle| \mathcal{F}_{(t-h)_+}\right] \leq C_{k,\lambda}h^{1-\lambda}$$

for any $i,j\in\{1,2\}$ and any $t,h>0$.

[SA4] $\xi\vee\frac{9}{10} < \xi'$ and

$$\sup_{i\in\mathbb{N}}(S^i - S^{i-1}) \vee \sup_{j\in\mathbb{N}}(T^j - T^{j-1}) \leq \bar{r}_n$$

holds true for every n.

[SA6] There exists a positive constant C such that $nH_n(t) \leq C$ for every t.

[SC1] There is a positive constant K such that $n^{-1}N^n_t \leq K$ for all n and t.

[SC2] [C2] holds, and $(A^1)'$, $(A^2)'$, $(\underline{A}^1)'$, $(\underline{A}^2)'$, and $[V,W]'$ for each $V,W = X^1, X^2, \underline{X}^1, \underline{X}^2$ are bounded.

[SF] We have [F] and there is a positive constant B such that

$$B^{-1} < \inf_{k\in\mathbb{N}} |\gamma_k^l| \leq \sup_{k\in\mathbb{N}} |\gamma_k^l| < B$$

for each $l = 1, 2$.

[SK$_\beta$] We have [K$_\beta$] with $E^1 = E^2 =: E$ and $(A^l)'$, $[M^l]'$, $(\underline{A}^l)'$, and $[\underline{M}^l]'$ $(l = 1, 2)$ are bounded. Moreover, there is a nonnegative bounded measurable function ψ on E such that

$$\sup_{\omega\in\Omega, t\in\mathbb{R}_+} |\delta^l(\omega, t, x)| \leq \psi(x) \text{ and } \int_E \psi(x)^\beta F^l(\mathrm{d}x) < \infty, \ l = 1, 2.$$

[SN$_r^\flat$] $(\int |z|^r Q_t(\mathrm{d}z))_{t\in\mathbb{R}_+}$ is a bounded process.

[ST] For each $l = 1, 2$ we have $\varrho_n^l(t) = \alpha_n^l(t)\rho_n$, where $(\rho_n)_{n\in\mathbb{N}}$ is a sequence of (deterministic) positive numbers satisfying $\rho_n \to 0$ and

$$\frac{n^{-\xi'+1/2}\log n}{\rho_n} \to 0$$

as $n \to \infty$, and $(\alpha_n^l(t))_{n\in\mathbb{N}}$ is a sequence of (not necessarily adapted) positive-valued stochastic processes such that there exists a positive constant K_0 satisfying

$$\frac{1}{K_0} < \inf_{t\in\mathbb{R}_+} \alpha_n^l(t) \leq \sup_{t\in\mathbb{R}_+} \alpha_n^l(t) < K_0, \qquad n = 1, 2, \ldots.$$

[T] $\xi' > 1/2$, and for each $l = 1, 2$ we have $\varrho_n^l(t) = \alpha_n^l(t)\rho_n$, where

(i) $(\rho_n)_{n\in\mathbb{N}}$ is a sequence of (deterministic) positive numbers satisfying $\rho_n \to 0$ and

$$\frac{n^{-\xi'+1/2}\log n}{\rho_n} \to 0$$

as $n \to \infty$.

(ii) $(\alpha_n^l(t))_{n\in\mathbb{N}}$ is a sequence of (not necessarily adapted) positive-valued stochastic processes. Moreover, there exists a sequence (R_k^l) of stopping times (with respect to $\mathbf{F}$) such that $R_k^l \uparrow \infty$ and both of the sequences $\left(\sup_{0\leq t<R_k^l} \alpha_n^l(t)\right)_{n\in\mathbb{N}}$ and $\left(\sup_{0\leq t<R_k^l} [1/\alpha_n^l(t)]\right)_{n\in\mathbb{N}}$ are tight for all k.

https://doi.org/10.1017/S0266466614000954 Published online by Cambridge University Press