Nonsynchronous covariance estimator and limit theorem II

*Takaki Hayashi**, *Keio University*   *Nakahiro Yoshida, University of Tokyo*

ISM Research Memorandum No. 1067, 2008

Abstract: An asymptotic distribution theory of the *nonsynchronous covariation* process for continuous semimartingales (Hayashi and Yoshida (2005b), Hayashi and Yoshida (2006)) is studied. In the setup, two continuous semimartingales are sampled at stopping times, in a *nonsynchronous* manner. The nonsynchronous covariation is a consistent estimator for the 'true' quadratic covariation of the semimartingales, as the mesh size of the sampling intervals shrinks to zero. In particular, we deal with the case when the limiting variation process of the normalized "approximation error" is random, which leads to the convergence to mixed normality, or "conditional" Gaussian martingale. A class of consistent estimators for the asymptotic variation process is proposed based on kernels, which will be useful from a viewpoint of statistical inferences. An example is presented.
Key words: discrete sampling; high-frequency data; martingale central limit theorem; nonsynchronicity; quadratic variation; realized volatility; stable convergence; semimartingale

1 Introduction

1.1 Background

For the last decade, intraday financial time-series, so-called *high-frequency data*, have been becoming increasingly available both in coverage and information contents. The use of high-frequency data has been expected to improve dramatically financial risk managements; one of such applications includes the estimation of variance-covariance structure of the financial markets, which is an essential routine operation for all the financial institutions.

In the literature, it is standard to use *realized volatility* (or *realized variance*) for estimating 'integrated' variance when asset returns are regarded to be sampled from diffusion-type processes. Likewise, when 'integrated' covariance is of interest, the use of *realized covariance* is fairly common. Nevertheless, the standard realized covariance estimator has a serious flaw in its structure when it is applied to multivariate *tick-by-tick* ('raw') data, where time-series are recorded irregularly, in a *nonsynchronous* manner.

The authors addressed the issue in Hayashi and Yoshida (2005b) and showed that realized covariance tends to be downward biased when its defining regular interval size is small relative to the frequency of observations. In the same paper, the authors proposed how to circumvent such biasedness by proposing a new estimator and showed that the estimator is unbiased when the underlying diffusion-type processes have no drift and is consistent as the mesh size of observation intervals tends to zero (or the frequency of observations tends to infinity) for a fixed terminal time of observation.

The authors then proved that the proposed estimator is asymptotically normally distributed (Hayashi and Yoshida (2004), Hayashi and Yoshida (2006)). This paper is in particular a sequel to

[0]*Corresponding author: Keio University, Graduate School of Business Administration, 2-1-1 Hiyoshi-honcho, Yokohama 223-8523, JAPAN. Email: takaki@kbs.keio.ac.jp

Hayashi and Yoshida (2006). The previous paper had a limitation in its results; its main theorem assumed that the asymptotic variance for the proposed estimator is *non-random*, which is certainly restrictive. That is, typically the asymptotic variance can depend on the variance and covariance of the underlying processes, which are not only of our interest but intrinsically *random* in our setting.

This paper removes the restriction, i.e., by allowing the asymptotic variance to be random. We derive asymptotic 'mixed' normality, or convergence to 'conditional' Gaussian martingales, in the space of *càdlàg* functions on $[0, \infty)$. Besides, a class of consistent estimators for the asymptotic variance is suggested based on kernels. Having such an estimator at hand enables us to conduct statistical inferences utilizing the proposed covariance estimator for the 'true' integrated covariance.

The statistics literature has a long list of studies on estimation problems of the diffusion parameter for diffusion processes based on discrete-time samples; see , for instance, the references in Hayashi and Yoshida (2006). Nevertheless, nonsynchronicity seems to have been almost neglected.

1.2 Setup

We fix a stochastic basis $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$. Let $X = (X_t)_{t\in\mathbb{R}_+}$ and $Y = (Y_t)_{t\in\mathbb{R}_+}$ be continuous local martingales, $\left(S^i\right)_{i\in\mathbb{Z}_+}$ and $\left(T^j\right)_{j\in\mathbb{Z}_+}$ be a.s. strictly increasing sequences of stopping times, $S^i \uparrow \infty$ and $T^j \uparrow \infty$ with $S^0 = 0$, $T^0 = 0$.

We will use the following symbols throughout the paper:

$$\begin{aligned} I^i &= \left[S^{i-1}, S^i\right),\ J^j = \left[T^{j-1}, T^j\right), \\ I^i(t) &= \left[S^{i-1}\wedge t, S^i \wedge t\right),\ J^j(t) = \left[T^{j-1}\wedge t, T^j\wedge t\right), \\ r_n(t) &= \max_{i\in\mathbb{N}} |I^i(t)| \vee \max_{j\in\mathbb{N}} |J^j(t)|. \end{aligned}$$

Here, $|\cdot|$ denotes the Lebesgue measure, and $\mathbb{N} = \{1, 2, ...\}$. In the preceding paper, the processes X and Y were implicitly assumed to be observable at some fixed terminal time T, which was non-essential but slightly annoying. The implicit assumption is abandoned in the current paper.

We will refer to $\left(I^i\right)_{i\in\mathbb{N}}$ and $\left(J^j\right)_{j\in\mathbb{N}}$, or equivalently to $\left(S^i\right)_{i\in\mathbb{Z}_+}$ and $\left(T^j\right)_{j\in\mathbb{Z}_+}$, as the *sampling designs* (or simply the *designs*) for X and Y. Also, the sampling designs truncated by time t, $\left(I^i(t)\right)_{i\in\mathbb{N}}$ and $\left(J^j(t)\right)_{j\in\mathbb{N}}$, may be referred to as the *random partitions* of $[0, t)$.

For simplicity, when we say "pair (i, j) overlaps" it will mean either $I^i(t) \cap J^j(t) \neq \varnothing$ (i.e., the two intervals I^i and J^j overlap by time t), or $I^i \cap J^j \neq \varnothing$ (i.e., by any time), depending on the situation.

For processes V and W, $V \cdot W$ denotes the integral (either stochastic or ordinary) of V with respect to W so far as it exists and well-defined. When the integrator W is continuous, it is always true that $V_- \cdot W = V \cdot W$, where $V_{t-} = \lim_{s\uparrow t} V_s$, the left limit at t.

For a stochastic process V and an interval I, $V(I) = V_b - V_{a-}$ with $a = \inf I$ and $b = \sup I$, the increment of V over an interval I; moreover $V(I)_t = V_{b\wedge t} - V_{(a\wedge t)-}$, which defines the increment as a process. For an interval I, we put $I(t) = I \cap [0, t)$. Note that, according to the notation, $V(I)_t = V(I(t))$ for any t provided that V is a continuous process.

Corresponding to every sampling intervals I^i and J^j, we define the point processes

$$I^i_t = 1_{[S^{i-1}, S^i)}(t),\ J^j_t = 1_{[T^{j-1}, T^j)}(t).$$

The quantity of interest is the quadratic covariation of the two processes X and Y, denoted as $[X, Y]$. As its 'sample counterpart,' in this paper we investigate the following quantity:

Definition 1.1 (Hayashi and Yoshida (2005b), Hayashi and Yoshida (2006)) *The* ***non-synchronous covariation process*** *of local martingales X and Y, associated with sampling designs $\mathcal{I} = \left(I^i\right)_{i\in\mathbb{N}}$ and $\mathcal{J} = \left(J^j\right)_{j\in\mathbb{N}}$ is defined by*

$$\{X, Y; \mathcal{I}, \mathcal{J}\}_t = \sum_{i,j=1}^{\infty} X(I^i)_t Y(J^j)_t 1_{\{I^i(t)\cap J^j(t)\neq\varnothing\}},\ t \in \mathbb{R}_+.$$

We may write it simply as $\{X, Y\}_t$ if there is no fear of confusion about the sampling designs $\mathcal{I}$ and $\mathcal{J}$. It has been shown that $\{X, Y\}_t$ is a "consistent estimator" for $[X, Y]_t$.

Theorem 1.1 (Hayashi and Yoshida (2005b), Hayashi and Kusuoka (2004)) *For each t, as $n \to \infty$,*

$$\{X, Y\}_t \xrightarrow{P} [X, Y]_t\,,$$

provided $r_n(t) \xrightarrow{P} 0$.

In light of the stated theorem, the authors emphasize that $\{X, Y\}_t$ *is regarded as a generalization in the context of nonsynchronous sampling schemes of the standard definition of the quadratic covariation process for semimartingales in stochastic analysis.*

In this paper, we are going to investigate the limiting distribution of the process $\{X, Y\}_\cdot$.

We denote by $\mathbb{C}\left(\mathbb{R}_+\right)$ the space of continuous functions on $\mathbb{R}_+$ equipped with the locally uniform topology, and by $\mathbb{D}\left(\mathbb{R}_+\right)$ 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 $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ of $(\Omega, \mathcal{F}, P)$ if $E\left[Y g\left(X^n\right)\right] \to E\left[Y g\left(X\right)\right]$ for any $\mathcal{F}$-measurable and bounded random variable Y and any bounded and continuous function g. We then write $X^n \to^{d_s} X$. See Rényi (1972), Aldous and Eagleson (1978), and Rootzén (1980). Clearly, stable convergence in law implies (ordinary) *convergence in law*. A sequence of (X^n) of stochastic processes is said to converge to a process X *uniformly on compacts in probability* (abbreviated *ucp*) if, for each $t > 0$, $\sup_{0\le s\le t}|X_s^n - X_s| \xrightarrow{P} 0$ as $n \to \infty$.

2 Asymptotic distribution theory

2.1 Stable convergence of the estimation error

The estimation error of $\{X, Y\}$ is given by

$$M_t^n = \{X, Y\}_t - [X, Y]_t = \sum_{i,j} L_t^{ij} K_t^{ij} \tag{2.1}$$

where

$$K_t^{ij} = 1_{\{I^i(t)\cap J^j(t)\neq\varnothing\}},$$
$$L_t^{ij} = \left(I_-^i \cdot X\right)_- \cdot \left(J_-^j \cdot Y\right)_t + \left(J_-^j \cdot Y\right)_- \cdot \left(I_-^i \cdot X\right)_t.$$

Our interest lies on the quadratic variation of M^n so as to explore asymptotics. According to Lemma 3.1 of Hayashi and Yoshida (2006),

$$M_t^n = \sum_{i,j} L_-^{ij} \cdot K_t^{ij},$$

in particular, M_t^n is a continuous local martingale with

$$[M^n, M^n]_t = \sum_{i,j,i',j'} \left(K_-^{ij} K_-^{i'j'}\right) \cdot \left[L^{ij}, L^{i'j'}\right]_t =: V^n (M, M)_t . \tag{2.2}$$

Let

$$V^n (M, X, Y)_t = \sum_{i,j} K_-^{ij} \cdot \left\{X \left(I^i\right)_- \cdot [X, Y] \left(J^j\right)\right\}_t + \sum_{i,j} K_-^{ij} \cdot \left\{Y \left(J^j\right) \cdot [X, X] \left(I^i\right)\right\}_t$$

and

$$V^n (M, Y, X)_t = \sum_{i,j} K_-^{ij} \cdot \left\{X \left(I^i\right)_- \cdot [Y, Y] \left(J^j\right)\right\}_t + \sum_{i,j} K_-^{ij} \cdot \left\{Y \left(J^j\right) \cdot [X, Y] \left(I^i\right)\right\}_t .$$

In view of the standard martingale central limit theorem, we formally state the following condition. (b_n) denotes a sequence of positive numbers tending to 0 as $n \to \infty$.

[A1*] *There exist an* $(\mathcal{F}_t)$-adapted*, nondecreasing, continuous process* $(V_t)_{t\in\mathbb{R}_+}$ *such that* $b_n^{-1} V^n (M, M)_t \xrightarrow{P} V_t$ as $n \to \infty$ *for every* t.

[B1] $b_n^{-1} V^n (M, X, Y)_t \xrightarrow{P} 0$ and $b_n^{-1} V^n (M, Y, X)_t \xrightarrow{P} 0$ as $n \to \infty$ *for every* t.

We consider the condition:

[W] *There exist an* $(\mathcal{F}_t)$-predictable *process* w such that $V_\cdot = \int_0^\cdot w_s^2 ds$.

An aim of this paper is to prove the following statement:

[SC] $b_n^{-\frac{1}{2}} M^n \to^{d_s} M$ in $\mathbb{C}(\mathbb{R}_+)$ as $n \to \infty$, where $M_\cdot = \int_0^\cdot w_s d\widetilde{W}_s$ and $\widetilde{W}$ is a one-dimensional Wiener process (defined on an extension of $\mathcal{B}$) which is independent of $\mathcal{F}$.

Proposition 2.1 *Suppose that [A1*], [B1] and [W] are fulfilled. Then [SC] holds.*

Proof. Notice that $[M^n, X] = V^n (M, X, Y)$ and $[M^n, Y] = V^n (M, Y, X)$. Since $[M^n, N] = 0$ for any bounded martingale N on $\mathcal{B}$, we obtain the stable convergence of $b_n^{-\frac{1}{2}} M^n$ from Jacod (1997) (or Jacod and Shiryaev (2000), Theorem IX.7.3, p.584). ∎

Each expression of $V^n (M, M)$, $V^n (M, X, Y)$ and $V^n (M, Y, X)$ is rather abstract; it may be of little help for explicitly calculating the quadratic variation/covariation and identifying the limiting distribution of M^n. From this regard, the preceding paper Hayashi and Yoshida (2006) introduce the following quantity so as to pursue more concrete appearance particularly of $V^n (M, M)$. We remark that the argument there is general enough so that it can indeed apply in the current setup as well.

Let

$$V_t^{iji'j'} = \left\{\left(I_-^i I_-^{i'}\right) \cdot [X, X]\right\}_t \left\{\left(J_-^j J_-^{j'}\right) \cdot [Y, Y]\right\}_t + \left\{\left(I_-^i J_-^{j'}\right) \cdot [X, Y]\right\}_t \left\{\left(I_-^{i'} J_-^j\right) \cdot [X, Y]\right\}_t \tag{2.3}$$

and set $V^{ij} = V^{ijij}$. $V^{iji'j'}$ is designed to approximate $\left[L^{ij}, L^{i'j'}\right]$ when the interval lengths $\left|I^i\right|$, $\left|I^{i'}\right|$, $\left|J^j\right|$, and $\left|J^{j'}\right|$ are sufficiently small.

We postulate the following hypothesis:

[B2] *For every $t \in \mathbb{R}_+$,*

$$b_n^{-1} \sum_{i,j,i',j'} \left(K_-^{ij} K_-^{i'j'}\right) \cdot \left[L^{ij}, L^{i'j'}\right]_t = b_n^{-1} \sum_{i,j,i',j'} \left(K_-^{ij} K_-^{i'j'}\right) \cdot V_t^{iji'j'} + o_P(1) \tag{2.4}$$

as $n \to \infty$.

Denote $[X] = [X, X]$ and $[Y] = [Y, Y]$ as usual. Let

$$\begin{aligned}\overline{V}_t^n &= \sum_{i,j} [X]\left(I^i(t)\right)[Y]\left(J^j(t)\right) K_t^{ij} + \sum_i [X, Y]\left(I^i(t)\right)^2 \\ &\quad + \sum_j [X, Y]\left(J^j(t)\right)^2 - \sum_{i,j} [X, Y]\left(\left(I^i(t)\right) \cap \left(J^j(t)\right)\right)^2,\end{aligned}$$

The following proposition will be used for identifying the limiting form of the quadratic variation:

Proposition 2.2 *Suppose [B2] holds. Then, $M^n = (M_t^n)_{0 \leq t \leq T}$ is a continuous local martingale with*

$$[M^n, M^n]_t = \overline{V}_t^n + o_P(b_n),$$

as $n \to \infty$.

For a proof, see Hayashi and Yoshida (2006).

We modify [A1*] as follows:

[A1] *There exist an $(\mathcal{F}_t)$-*adapted*, nondecreasing, continuous process $(V_t)_{0 \in \mathbb{R}_+}$ such that $b_n^{-1}\overline{V}_t^n \xrightarrow{P} V_t$* as *$n \to \infty$ for every t.*

Accordingly, we can rephrase Proposition 2.1 as follows:

Proposition 2.3 *Suppose that [A1], [B1], [B2] and [W] are fulfilled. Then [SC] holds.*

Condition [B2] is too technical to check in practice. A sufficient condition will be given in Section 2.3. Condition [B2] will also disappear in the main theorems that hypothesize amiable conditions.

2.2 Convergence of the sampling measures and a representation of V_t

In Hayashi and Yoshida (2005a), the authors introduced empirical distribution functions of the sampling times given by

$$\begin{aligned} H_n^1(t) &= \sum_i |I^i(t)|^2,\ H_n^2(t) = \sum_j |J^j(t)|^2, \\ H_n^{1\cap 2}(t) &= \sum_{i,j} |\left(I^i \cap J^j\right)(t)|^2,\ H_n^{1*2}(t) = \sum_{i,j} |I^i(t)||J^j(t)|K_t^{ij}, \end{aligned} \tag{2.5}$$

where $|\cdot|$ is the Lebesgue measure. Clearly, the four functions are $(\mathcal{F}_t)$-adapted, non-decreasing, piecewise-quadratic continuous functions, whose graphs contain 'kinks' at the observation stopping times. Note that they altogether preserve the information regarding the (random) positions of the nonsynchronous sampling times $\left(S^i\right)$ and $\left(T^j\right)$.

[A1']: *There exists a possibly random, nondecreasing, continuous functions* $H^1, H^2, H^{1\cap 2}$, *and* H^{1*2} on $[0,T]$, *such that each* $H^k = \int_0^t h_s^k ds$ for some density h^k, and that $b_n^{-1} H_n^k(t) \xrightarrow{P} H^k(t)$ *as* $n \to \infty$ *for every* $t \in \mathbb{R}_+$ *and* $k = 1, 2, 1\cap 2, 1*2$.

Then, an extension of Theorem 2.2 of Hayashi and Yoshida (2005a) is given as follows.

Proposition 2.4 *Suppose [A1'], [B1] and [B2] are fulfilled, and that each* $[X]$, $[Y]$, *and* $[X,Y]$ *is absolutely continuous with a locally bounded derivative. Then [SC] is valid with* w_s *given by*

$$w_s = \sqrt{[X]'_s\,[Y]'_s\,h_s^{1*2} + \left([X,Y]'_s\right)^2 (h_s^1 + h_s^2 - h_s^{1\cap 2})}. \tag{2.6}$$

Proof. Lemma 2.1 below identifies the limiting variation process V in the condition [A1] by (2.6). Thus Proposition 2.3 implies the assertion. ∎

Lemma 2.1 *Suppose* $[X]$, $[Y]$, *and* $[X,Y]$ *are continuously differentiable. Then, [A1'] implies that*

$$\text{(i) } b_n^{-1} \sum_i [X,Y]\left(I^i(t)\right)^2 \xrightarrow{P} \int_0^t \left([X,Y]'_s\right)^2 H^1(ds),$$

$$\text{(ii) } b_n^{-1} \sum_j [X,Y]\left(J^j(t)\right)^2 \xrightarrow{P} \int_0^t \left([X,Y]'_s\right)^2 H^2(ds),$$

$$\text{(iii) } b_n^{-1} \sum_j [X,Y]\left(\left(I^i \cap J^j\right)(t)\right)^2 \xrightarrow{P} \int_0^t \left([X,Y]'_s\right)^2 H^{1\cap 2}(ds),$$

$$\text{(iv) } b_n^{-1} \sum_{i,j} [X]\left(I^i(t)\right)[Y]\left(J^j(t)\right) K_t^{ij} \xrightarrow{P} \int_0^t [X]'_s\,[Y]'_s\, H^{1*2}(ds).$$

as $n \to \infty$ *for every* t.

For a proof, see Hayashi and Yoshida (2006).

2.3 Sufficient condition for [B2]: strong predictability

We have presented a basic version of the limit theorems as Propositions 2.1 and 2.3, in the latter of which Condition [B2] is assumed. Under more general sampling, however, [B2] is still technical. A more tractable condition on the sampling scheme to ensure [B2] is desirable for practical perspectives. The "strong predictability condition" introduced by Hayashi and Yoshida (2006) serves for such purpose.

Let ξ and ξ' be constants satisfying $\frac{4}{5} \vee \xi < \xi' < 1$.

[A2]: *For every* $n, i \in \mathbb{N}$, S^i *and* T^i *are* $\left(\mathcal{G}_t^{(n)}\right)$*-stopping times, where* $\left\{\mathcal{G}_t^{(n)}\right\}_{t\in\mathbb{R}_+}$ *is the filtration given by* $\mathcal{G}_t^{(n)} = \mathcal{F}_{\left(t - b_n^{\xi}\right)\vee 0}$ *for* $t \in \mathbb{R}_+$.

For real-valued function x on $\mathbb{R}_+$, the *modulus of continuity* on $[0,T]$ is denoted as

$$w(x;\delta,T) = \sup_{\substack{s,t;|s-t|\le\delta \\ 0\le s,t\le T}} |x(t)-x(s)| \tag{2.7}$$

for $T,\delta \ge 0$. Write $H_t^* = \sup_{0\le s\le t}|H_s|$ for a process H.

[A3]: (i) $[X]$, $[Y]$, *and* $[X,Y]$ *are absolutely continuous. (ii) Moreover, for the density processes* $f=[X]'$, $[Y]'$ *and* $[X,Y]'$, $w(f;h,t) = O_P\left(h^{\frac{1}{2}-\lambda}\right)$ *as* $h\to 0$ *for every* $t,\lambda\in(0,\infty)$, and $|f_0|<\infty$, a.s.

Remark: It should also be noted that (ii) will imply that, for any t,

$$\left([X]'\right)_t^* < \infty,\ \left([Y]'\right)_t^* < \infty,\ \left([X,Y]'\right)_t^* < \infty,\ \textit{almost surely}. \tag{2.8}$$

[A4]: $r_n(t) = o_P\left(b_n^{\xi'}\right)$ *for every* $t\in\mathbb{R}_+$.

Remark: In Hayashi and Yoshida (2006), the authors used a different parametrization for the exponents of b_n in [A2] and [A4]. That is, α and β adopted there can be rewritten here in terms of ξ and ξ' as

$$\beta = \xi - \frac{2}{3}, \text{ and } \alpha = \xi' - \frac{2}{3}.$$

As mentioned earlier, these conditions altogether form a sufficient condition for [B2].

Proposition 2.5 *[B2] holds true under [A2], [A3] and [A4].*

For a proof, see Hayashi and Yoshida (2006).

2.4 Limit theorems for semimartingales: main results

Up to the previous sections we focused on the case without drifts, i.e., when X and Y are martingales. In this section, we relax the restriction; we consider the case when X and Y are continuous *semimartingales.*

Suppose that $X\equiv A^X+M^X$ and $Y\equiv A^Y+M^Y$ are continuous semimartingales, where A^X and A^Y are finite variation parts, M^X and M^Y are continuos martingale parts. Like the previous sections, we treat X and Y square-integrable, because the usual localization argument can apply due to the path continuity. In this case the nonsynchronous covariation process of X and Y associated with $\left(I^i\right)$ and $\left(J^j\right)$ can be defined exactly by the same way as Definition 1.1.

The following are used for the case when X and Y have drifts A^X *and* A^Y.

[A5]: A^X *and* A^Y *are absolutely continuous, and* $w(f;h,t) = O_P\left(h^{\frac{1}{2}-\lambda}\right)$ as $h\to 0$ for every $t\in\mathbb{R}_+$ *and some* $\lambda\in(0,1/4)$, for the density processes $f=\left(A^X\right)'$ *and* $\left(A^Y\right)'$.

Remark: [A5] implies that *for any* $t>0$,

$$\left(A^{l'}\right)_t^* < \infty, \text{ almost surely, } l = X,Y. \tag{2.9}$$

[A6]: *For every t, as $n \to \infty$,*

$$b_n^{-1} \sum_i \left|I^i(t)\right|^2 + b_n^{-1} \sum_j \left|J^j(t)\right|^2 = O_p(1).$$

Remark: Condition [A5] is slightly stronger than (C4') of Hayashi and Yoshida (2004). Condition [A1] does not imply [A6]. Indeed, it is possible to make a sampling scheme that includes $[n^{7/10}]$ intervals of $n^{-4/5}$ in length and $[X]$ and $[Y]$ do not increase on the union of those intervals, and also [A1] holds. However [A6] breaks in this case. On the other hand, [A1'] implies [A6]. The Poisson sampling scheme considered in Hayashi and Yoshida (2004) is an example.

Let

$$B_{1,t} = \sum_{i,j=1}^{\infty} A^X(I^i)_t M^Y\left(J^j\right)_t K_t^{ij},\ B_{2,t} = \sum_{i,j=1}^{\infty} A^Y(J^j)_t M^X\left(X^i\right)_t K_t^{ij},$$
$$B_{3,t} = \sum_{i,j=1}^{\infty} A^X(I^i)_t A^Y\left(J^j\right)_t K_t^{ij}.$$

Then,

$$\{X,Y\}_t \equiv \sum_{i,j=1}^{\infty} M^X(I^i)_t M^Y(J^j)_t K_t^{ij} + B_{1,t} + B_{2,t} + B_{3,t}.$$

Lemma 2.2 *Suppose that [A2]–[A6] are satisfied. Then, for any $t > 0$, as $n \to \infty$,*

$$b_n^{-1/2} B_{l,t}^* = o_P(1), \quad l = 1,2,3.$$

The negligibility of $B_{l,t}^*$ is proved by Lemma 5.1 in Hayashi and Yoshida (2006).

Here is our first main result:

Theorem 2.1 *Suppose that X and Y are continuous semimartingales.*
(a) If [A1]-[A6] and [W] are satisfied, then [SC] holds.
(b) If [A1'], [A2]-[A5] are satisfied, then [SC] holds for w given by (2.6).

Proof. In both cases (a) and (b), [A2]–[A6] holds, hence Lemma 2.2 ensures the behavior of $\{X,Y\}_t$ is the same as that of $\left\{M^X, M^Y\right\}$ in the first order; that $[X,Y] = \left[M^X, M^Y\right]$ is trivial. Eventually, we will consider

$$M_t^n = \left\{M^X, M^Y\right\}_t - \left[M^X, M^Y\right]_t = \sum_{i,j} L_t^{ij} K_t^{ij}$$

in place of (2.1), but in the present situation with

$$L_t^{ij} = \left(I_-^i \cdot M^X\right)_- \cdot \left(J_-^j \cdot M^Y\right)_t + \left(J_-^j \cdot M^Y\right)_- \cdot \left(I_-^i \cdot M^X\right)_t.$$

Condition [B2] holds under the assumptions according to Proposition 2.5; note that $V_t^{iji'j'}$ is unchanged and that [A3] still holds even if $\left(M^X, M^Y\right)$ replaces (X,Y). Therefore, once condition [B1] is verified for $\left(M^X, M^Y\right)$, (a) follows from Proposition 2.2 and (b) from Proposition 2.4.

After all, what we have to show is that $b_n^{-\frac{1}{2}}V^n(M, M^X, M^Y) = o_P(1)$ and $b_n^{-\frac{1}{2}}V^n(M, M^Y, M^X) = o_P(1)$ as $n \to \infty$ for every t. In the current context, for instance, $V^n(M, M^X, M^Y)$ is now given by

$$V^n(M, M^X, M^Y) = \sum_{i,j} K_-^{ij} \cdot \Big\{ M^X\left(I^i\right)_- \cdot \left[M^X, M^Y\right]\left(J^j\right)\Big\}_t + \sum_{i,j} K_-^{ij} \cdot \left\{ M^Y\left(J^j\right) \cdot \left[M^X, M^X\right]\left(I^i\right)\right\}_t.$$

Since $\left[M^X, M^X\right]$ and $\left[M^X, M^Y\right]$ possessing [A3] also satisfy the property described in [A5] (which is originally postulated for A^X and A^Y), exactly the same argument made for $\mathbb{III}_t$ in the proof for Lemma 5.1 of Hayashi and Yoshida (2006) to give $b_n^{-\frac{1}{2}}V^n(M, M^X, M^Y) = o_P(1)$. The convergence of $b_n^{-\frac{1}{2}}V^n(M, M^Y, M^X) = o_P(1)$ is verified in the same fashion. ■

2.5 Empirical nonsynchronous covariation process

A keen reader may be aware that $\{X, Y\}_t$ *may not always be observable*. Although this fact is not crucial in the development of the asymptotic distribution theory, from a viewpoint of practical applications it is certainly a distracting feature. The argument here is the way how to amend such a minor flaw pertaining to the previous construction (1.1).

Definition 2.1 *The empirical nonsynchronous covariation process of X and Y associated with sampling designs $\mathcal{I} = \left(I^i\right)_{i\in\mathbb{N}}$ and $\mathcal{J} = \left(J^j\right)_{j\in\mathbb{N}}$ is the process*

$$\overline{\{X, Y\}}_t = \sum_{\substack{i,j=1 \\ S^i \vee T^j \le t}}^{\infty} X(I^i)Y(J^j)1_{\{I^i \cap J^j \neq \varnothing\}}, \quad t \in \mathbb{R}_+.$$

Obviously,

$$\overline{\{X, Y\}}_t = \sum_{\substack{i,j=1 \\ S^i \vee T^j \le t}}^{\infty} X(I^i)_t Y(J^j)_t 1_{\{I^i(t) \cap J^j(t) \neq \varnothing\}}.$$

It is the piecewise constant, càdlàg version of the nonsynchronous covariation process and

$$\overline{\{X, Y\}}_t = \{X, Y\}_t$$

at $t \in \Pi = \left(S^i\right) \cap \left(T^j\right)$ (i.e., 'synchronous' points). Otherwise they do not coincide in general; however the difference is negligible as shown next.

The empirical version $\overline{\{X, Y\}}_t$ is observable at *any* t, hence is a *statistic*, as intended.

Lemma 2.3 *As $n \to \infty$,*

$$b_n^{-\frac{1}{2}} \left(\overline{\{M^X, M^Y\}} - \left\{M^X, M^Y\right\} \right) \overset{ucp}{\to} 0.$$

Proof. To proceed, we introduce the following symbols:

$$I\left(J^j\right)_t = \sum_i K_t^{ij} I_t^i, \; j \ge 1; \qquad J\left(I^i\right)_t = \sum_j K_t^{ij} J_t^j, \; i \ge 1.$$

We can interpret for instance that $I\left(J^j\right)_t$ is the $\{0, 1\}$-valued process corresponding to the smallest (aggregate) interval consisting of I^is that covers a given J^j. As per the notation system introduced earlier, $I\left(J^j\right)(t)$ denotes the aggregate interval *truncated by time* t, hence $X\left(I\left(J^j\right)(t)\right)$ will mean the increment of X over it.

Clearly, for any t, there exists one (and only one) pair (i,j) such that $t \in \left[S^{i-1}, S^i\right)$ and $t \in \left[T^{j-1}, T^j\right)$. Call such indices i_t and j_t in what follows. Notice by construction, for arbitrary fixed time s, $i_s - 1 = \max\left\{i : S^i \leq s\right\}$ and $j_s - 1 = \max\left\{j : T^j \leq s\right\}$.

We remark that $\left|I\left(J^{j_s}\right)(s)\right| \leq 2r_n(s)$ and $\left|J\left(I^{i_s}\right)(s)\right| \leq 2r_n(s)$ for any s. In fact, for the former, because

$$I\left(J^{j_s}\right)(s) \equiv \left[S^{i_{Tj_s-1}-1}, s\right) \equiv I^{i_{Tj_s-1}} \cup \left[S^{i_{Tj_s-1}}, s\right),$$

we have

$$\left|I\left(J^{j_s}\right)(s)\right| \equiv \left|I^{i_{Tj_s-1}}\right| + \left(s - S^{i_{Tj_s-1}}\right) \leq \left|I^{i_{Tj_s-1}}\right| + \left(s - T^{j_s-1}\right) \leq 2r_n(s)$$

since $T^{j_s-1} \leq S^{i_{Tj_s-1}}$. The latter can be shown by the same way.

Put $\Delta_t = \overline{\{M^X, M^Y\}}_t - \left\{M^X, M^Y\right\}_t$, the gap between the two quantities of interest.
Case: $S^{i_s-1} < T^{j_s-1}$: For $i \leq i_s - 1$,

$$\left[I^i \cap J^j \neq \varnothing\right] \quad \Longrightarrow \quad \left[\sup J^j \leq T^{j_s-1} \leq s\right] \quad \Longrightarrow \quad \left[I^i(s) \cap J^j(s) \neq \varnothing\right],$$

i.e., any "overlapping" pair (i,j) with $i \leq i_s - 1$ must be one "completed" by the fixed time s. Such pairs are included in the summation in both $\overline{\{M^X, M^Y\}}$ and $\left\{M^X, M^Y\right\}$. Consequently, when the gap between the two quantities has to be evaluated, only the remaining overlapping pairs (i,j) with $i = i_s$ are to be taken into account. Thus, for any s, t with $s \leq t$,

$$\begin{aligned}
|\Delta_s| &\leq \left|M^X\left(I^{i_s}(s)\right)\right| \left|M^Y\left(J\left(I^{i_s}\right)(s)\right)\right| \\
&\leq w\left(M^X; r_n(t), t\right) \cdot w\left(M^Y; 2r_n(t), t\right).
\end{aligned}$$

Case: $S^{i_s-1} > T^{j_s-1}$: By symmetry,

$$\begin{aligned}
|\Delta_s| &\leq \left|M^X\left(I\left(J^{j_s}\right)(s)\right)\right| \left|M^Y\left(J^{j_s}(s)\right)\right| \\
&\leq w\left(M^X; 2r_n(t), t\right) \cdot w\left(M^Y; r_n(t), t\right).
\end{aligned}$$

Case: $S^{i_s-1} = T^{j_s-1}$: By the same token,

$$\begin{aligned}
|\Delta_s| &\leq \left|M^X\left(I^{i_s}(s)\right)\right| \left|M^Y\left(J^{j_s}(s)\right)\right| \\
&\leq w\left(M^X; r_n(t), t\right) \cdot w\left(M^Y; r_n(t), t\right).
\end{aligned}$$

Therefore, putting all together, we have

$$|\Delta_s| \leq w\left(M^X; 2r_n(t), t\right) \cdot w\left(M^Y; 2r_n(t), t\right).$$

Since M^X and M^Y are continuous local martingales, due to the martingale representation as Brownian motion, for any $t > 0$ and any $\varepsilon > 0$,

$$w\left(M^X; h, t\right) \leq h^{\frac{1}{2}-\varepsilon}$$

as $h \downarrow 0$. The same inequality is true for Y as well. Hence,

$$|\Delta_s| \leq 2r_n(t)^{1-2\varepsilon}.$$

From [A4], one can conclude

$$b_n^{-\frac{1}{2}} \sup_{s\in[0,t]} |\Delta_s| = o_P\left(b_n^{\left(\frac{2}{3}+\alpha\right)(1-2\varepsilon)-\frac{1}{2}}\right) = o_P(1),$$

since $\alpha \in \left(\beta, \frac{1}{3}\right)$ with $\beta \in \left(\frac{2}{15}, \frac{1}{3}\right)$ and $\varepsilon > 0$ can be taken arbitrarily small. ■

Set

$$\overline{B}_{1,t} = \sum_{\substack{i,j=1 \\ S^i\vee T^j\le t}}^{\infty} A^X(I^i)M^Y\left(J^j\right)K^{ij},\ \overline{B}_{2,t} = \sum_{\substack{i,j=1 \\ S^i\vee T^j\le t}}^{\infty} A^Y(J^j)M^X\left(X^i\right)K^{ij},$$

$$\overline{B}_{3,t} = \sum_{\substack{i,j=1 \\ S^i\vee T^j\le t}}^{\infty} A^X(I^i)A^Y\left(J^j\right)K^{ij},$$

where $K^{ij} = 1_{\{I^i\cap J^j\neq\varnothing\}}$, then we obtain the corresponding discrete version of the decomposition

$$\overline{\{X,Y\}}_t \equiv \sum_{\substack{i,j=1 \\ S^i\vee T^j\le t}}^{\infty} M^X(I^i)M^Y(J^j)K^{ij} + \overline{B}_{1,t} + \overline{B}_{2,t} + \overline{B}_{3,t}.$$

Lemma 2.4 *Suppose that [A2]–[A6] are satisfied. Then, for any $t > 0$, as $n \to \infty$,*

$$b_n^{-1/2}\overline{B}_{l,t}^{*} = o_P(1), \quad l = 1,2,3.$$

Outline of Proof. The negligibility of $B_{l,t}^*$ is already stated in Lemma 2.2. The (uniform) difference between $\overline{B}_l$ and B_1, after rescaled $b_n^{-1/2}$, can also be shown to be negligible, by applying a similar argument to the proof of Lemma 2.3. ■

For $\overline{M}^n = \overline{\{X,Y\}} - [X,Y]$, we will show:

[SCE] $b_n^{-\frac{1}{2}}\overline{M}^n \to^{d_s} M$ in $\mathbb{D}(\mathbb{R}_+)$ as $n \to \infty$, where $M = \int_0^{\cdot} w_s d\widetilde{W}_s$ and $\widetilde{W}$ is a one-dimensional Wiener process (defined on an extension of $\mathcal{B}$) which is independent of $\mathcal{F}$.

Here is our second main result:

Theorem 2.2 *Suppose that X and Y are continuous semimartingales.*
(a) If [A1]-[A4] and [W] are satisfied, then [SCE] holds.
(b) If [A1'], [A2]-[A5] are satisfied, then [SCE] holds for w given by (2.6).

3 Statistical applications and example

3.1 Stochastic differential equations

Suppose that X^1 and X^2 are continuous Itô semimartingales,

(3.1) $$dX_t^k = \mu_t^k dt + \sigma_t^k dW_t^k,\ k = 1,2,$$

where μ_t^k are $(\mathcal{F}_t)$-adapted, continuous processes, σ_t^k are strictly positive, $(\mathcal{F}_t)$-adapted, continuous processes, $k = 1,2$. The $(\mathcal{F}_t)$-adapted Brownian motions W_t^k are correlated with $d\left[W^1, W^2\right]_t =$

$\rho_t dt$, where ρ_t is an $(\mathcal{F}_t)$-adapted, continuous process. This model is a *stochastic volatility model*, a standard model in the finance literature. In the meantime, we continue to use the same symbols $\mathcal{I} = \left(I^i\right)_{i\in\mathbb{N}}$ and $\mathcal{J} = \left(J^j\right)_{j\in\mathbb{N}}$ for the sampling designs associated with X^1 and X^2, respectively.

Under the assumption the quadratic variations / covariation are expressible as

$$\left[X^k\right]_t = \int_0^t \left(\sigma_s^k\right)^2 ds, \; k = 1, 2,$$
$$\left[X^1, X^2\right]_t = \int_0^t \rho_s \sigma_s^1 \sigma_s^2 ds.$$

Besides, we assume the following smooth condition for ρ, σ^k, $k = 1, 2$, together with μ^k, $k = 1, 2$:

[A3']: *For every* $\lambda > 0$ and $t \in \mathbb{R}_+$, $w\left(f; h, t\right) = O_P\left(h^{\frac{1}{2}-\lambda}\right)$ as $h \downarrow 0$ for $f = \sigma^1, \sigma^2$ and ρ.

[A5']: *For some* $\lambda \in (0, 1/4)$ and any $t \in \mathbb{R}_+$, $w\left(f; h, t\right) = O_P\left(h^{\frac{1}{2}-\lambda}\right)$ as $h \downarrow 0$ for $f = \mu^k$, $k = 1, 2$.

Notice that, if 'X^1' is read as 'X' and 'X^2' as 'Y' as in the previous section, it can be seen that [A3'] implies [A3](ii) while [A5'] implies [A5](ii), respectively.

Remark: Together with path continuity, [A3'] and [A5'] imply that the paths of ρ, σ^k, μ^k, $k = 1, 2$ are all Hölder continuous with exponent $\left(\frac{1}{2} - \lambda\right)$.

Now, define the distribution functions associated with the sampling designs $\mathcal{I}$ and $\mathcal{J}$ by

$$\overline{H}_n^1(t) = \sum_{i:\, S^i \le t} |I^i|^2, \quad \overline{H}_n^2(t) = \sum_{j:\, T^j \le t} |J^j|^2,$$
$$\overline{H}_n^{1\cap 2}(t) = \sum_{\substack{i,j:\\ S^i \vee T^j \le t}} |I^i \cap J^j|^2, \quad \overline{H}_n^{1*2}(t) = \sum_{\substack{i,j:\\ S^i \vee T^j \le t}} |I^i||J^j| K^{ij},$$

where $K^{ij} = 1_{\{I^i \cap J^j \neq \varnothing\}}$. This time, the four functions are $(\mathcal{F}_t)$-adapted, with non-decreasing, *piecewise constant, càdlàg* paths, which 'jump' at the observation stopping times. They are all observable at any t.

[A1"]: *There exist a possibly random, nondecreasing functions* $H^1, H^2, H^{1\cap 2}$*, and* H^{1*2} on $[0, T]$, such that each $h^1, h^2, h^{1\cap 2}$*, and* h^{1*2} on $[0, \infty)$, such that $H^k(t) = \int_0^\infty h_s^k ds$ for some density h^k, and that $< \infty$ and $b_n^{-1}\overline{H}_n^k(t) \xrightarrow{P} H^k(t)$ as $n \to \infty$ *for every* $t \in \mathbb{R}_+$ *and* $k = 1, 2, 1\cap 2, 1 * 2$.

The equivalence of [A1'] and [A1"] can be verified. In fact,

$$\overline{H}_n^k(s) \le H_n^k(s) \le \overline{H}_n^k(s) + 2r_n(t)^2, \quad \text{for all } s \in [0, t]\,, \, k = 1, 2, 1\cap 2, 1 * 2.$$

We take on the case $k = 1 * 2$ only, for all the others are straightforward. The first inequality is obvious by construction. Moreover, according to a similar argument adopted in the proof for Lemma 2.3, for any s, t with $s \le t$,

$$H_n^{1*2}(s) - \overline{H}_n^{1*2}(s) \le \left|I^{i_t}(t)\right| \left|J\left(I^{i_t}\right)(t)\right| \vee \left|I\left(J^{j_t}\right)(t)\right| \left|J^{j_t}(t)\right| \le 2r_n\left(t\right)^2$$

so that

$$\sup_{s\in[0,t]} \left|H_n^{1*2}(s) - \overline{H}_n^{1*2}(s)\right| \le 2r_n\left(t\right)^2.$$

Hence, the second inequality also holds. Since $b_n^{-1} r_n(t)^2 \xrightarrow{P} 0$ under [A4] for example, we have ascertained that the convergence of $b_n^{-1}\overline{H}_n^{1*2}$ is equivalent to that of $b_n^{-1} H_n^{1*2}$.

Then, by the application of Theorem 3.1 we have the following theorem:

Theorem 3.1 *Suppose that X and Y are continuous semimartingales. Suppose that either [A1'] or [A1"], and [A2], [A3'], [A4] and [A5'] are satisfied. Then, for $M^n = b_n^{-1/2}\left(\{X,Y\} - \int_0^\cdot \rho_s\sigma_s^1\sigma_s^2 ds\right)$ and $\overline{M}^n = b_n^{-1/2}\left(\overline{\{X,Y\}} - \int_0^\cdot \rho_s\sigma_s^1\sigma_s^2 ds\right)$, [SC] and [SCE] hold for w given by*

$$w_s = \int_0^\cdot \sqrt{\left(\sigma_s^1\sigma_s^2\right)^2 h_s^{1*2} + \left(\rho_s\sigma_s^1\sigma_s^2\right)^2 \left(h_s^1 + h_s^2 - h_s^{1\cap 2}\right)}. \tag{3.2}$$

3.2 Studentization

We shall briefly discuss studentization. Consider w_s given in (3.2). In our context, w_t is not observable since it contains unknown quantities such as $\rho_t\sigma_t^1\sigma_t^2$; neither is $\int_0^\cdot w_s^2 ds$. Suppose we find a statistic $\widehat{\int_0^t w_s^2 ds}$ such that

$$\widehat{\int_0^\cdot w_s^2 ds} \quad \xrightarrow{P} \quad \int_0^\cdot w_s^2 ds \qquad (\text{in } \mathbb{D})$$

as $n \to \infty$.

Then, the stable convergence stated in Theorem 3.1 implies that

$$\left(b_n^{-1/2}\overline{M}^n, \widehat{\int_0^\cdot w_s^2 ds}\right) \quad \xrightarrow{\mathcal{L}} \quad \left(M, \int_0^\cdot w_s^2 ds\right),$$

with $[M,M] = \int_0^\cdot w_s^2 ds$; in particular, for every $t > 0$,

$$\frac{b_n^{-\frac{1}{2}}\left(\overline{\{X^1,X^2\}}_t - \int_0^t \rho_s\sigma_s^1\sigma_s^2 ds\right)}{\sqrt{\widehat{\int_0^t w_s^2 ds}}} \quad \xrightarrow{\mathcal{L}} \quad N(0,t)$$

as $n \to \infty$, whenever $\int_0^t w_s^2 ds > 0$ a.s. I.e., the 'studentized' estimation error of the nonsynchronous covariation is asymptotically normally distributed, which can be used for inferences.

3.3 Construction of $\widehat{\int_0^t w_s^2 ds}$: Kernel-based approach

Let $K(u)$ be a kernel function such that $\int_{-\infty}^{\infty} K(u)\,du = 1$ and $K(u) \geq 0$ for all u. K is assumed to be absolutely continuous with $\int_{-\infty}^{\infty} |K'(s)|\,ds < \infty$. Here, K' may mean a 'generalized' derivative. Moreover, for convenience, we introduce the following symbol: For arbitrary window size $h > 0$, let $K_h(u) = \frac{K\left(\frac{u}{h}\right)}{h}$ (so that $\int_{-\infty}^{\infty} K_h(u)\,du = 1$).

For every $t \geq 0$, define

$$\partial\widetilde{\{X^1,X^2\}}_t^h = \int_{-\infty}^{\infty} \overline{\{X^1,X^2\}}_s K_h'(t-s)\,ds,$$

$$\partial\widetilde{\{X^k\}}_t^h = \partial\widetilde{\{X^k,X^k\}}_t^h, \qquad k = 1,2$$

which are observable for any $h > 0$ and $n \geq 1$.

Using these quantities, define, for any $h > 0$, $n \geq 1$,

$$\widehat{\int_0^t w_s^2 ds} = \int_0^t \partial\widetilde{\{X^1\}}_s^h \partial\widetilde{\{X^2\}}_s^h b_n^{-1} \overline{H}_n^{1*2}(ds) + \int_0^t \partial\widetilde{\{X^1, X^2\}}_s^h b_n^{-1} \left(\overline{H}_n^1 + \overline{H}_n^2 - \overline{H}_n^{1\cap 2}\right)(ds).$$

This quantity is also *observable* for any $h > 0$, $n \geq 1$, and $t > 0$. Since the integrators $\overline{H}_n^k(\cdot)$ have piecewise constant, càdlàg paths, the integrations are indeed summations taken at the sampling times up to time t, which are easy to implement.

Proposition 3.1 *Under the assumptions in Theorem 3.1,*

$$\widehat{\int_0^{\cdot} w_s^2 ds} \overset{ucp}{\rightarrow} \int_0^{\cdot} w_s^2 ds$$

as $n \to \infty$, provided that $b_n^{\frac{1}{2}} h^{-1} \to 0$.

Proof. Let $\overline{M}_t^n = \overline{\{X^1, X^2\}}_t - \left[X^1, X^2\right]_t$, then clearly it satisfies $b_n^{-\frac{1}{2}} \sup_{s\in[0,t]} \left|\overline{M}_s^n\right| = O_p(1)$ (for any $t > 0$) under the conditions [A1] through [A6]. Thus, by the integration-by-parts formula,

$$\begin{aligned}
\partial\widetilde{\{X^1, X^2\}}_t^h &= \int_{-\infty}^{\infty} \left(\left[X^1, X^2\right]_s + \overline{M}_s^n\right) K_h'(t-s)\, ds \\
&= \left[X^1, X^2\right]_s K_h(t-s)\Big|_{s=-\infty}^{\infty} + \int_{-\infty}^{\infty} \left[X^1, X^2\right]_s' K_h(t-s)\, ds + O_p(b_n^{\frac{1}{2}}) \int_{-\infty}^{\infty} \left|K_h'(t-s)\right| ds \\
&= \rho_t \sigma_t^1 \sigma_t^2 + w(\rho\sigma^1\sigma^2; h, t) + O_p(b_n^{\frac{1}{2}} h^{-1}),
\end{aligned}$$

uniformly in $t \in [0, t']$ for any $t' > 0$. This can be realized by choosing h such that $b_n^{\frac{1}{2}} h^{-1} \to 0$, or $h = b_n^{\frac{1}{2}-\epsilon}$ for arbitrary $\epsilon \in \left(0, \frac{1}{2}\right)$. The last term is because $\int_{-\infty}^{\infty} |K_h'(t-s)|\, ds = \int_{-\infty}^{\infty} \frac{\left|K'\left(\frac{t-s}{h}\right)\right|}{h^2} ds = O(h^{-1})$.

Similarly,

$$\partial\widetilde{\{X^k\}}_t^h = \left(\sigma_t^k\right)^2 + w(\left(\sigma^k\right)^2; h, t) + O_p(b_n^{\frac{1}{2}} h^{-1}),$$

uniformly in $t \in [0, t']$ for any $t' > 0$.

Therefore, it is immediate to obtain the assertion of the proposition. ∎

Special case:

A special case of kernel-based approach is the following naïve approach ("historical averages"). For any $s > 0$ and for any $h > 0$, define

$$\begin{aligned}
\partial\overline{\{X^1, X^2\}}_s^h &= \frac{1}{h}\left(\overline{\{X^1, X^2\}}_s - \overline{\{X^1, X^2\}}_{s-h}\right), \\
\partial\overline{\{X^k\}}_s^h &= \partial\overline{\{X^k, X^k\}}_s^h, \qquad k = 1, 2.
\end{aligned}$$

Specifically, as $h \to 0$ (with $b_n^{\frac{1}{2}} h^{-1} \to 0$), they can be expressed as

$$\partial\overline{\{X^1, X^2\}}_s^h = \rho_s \sigma_s^1 \sigma_s^2 + w(\rho\sigma^1\sigma^2; h, s) + O_p(b_n^{\frac{1}{2}} h^{-1})$$

and
$$\partial\overline{\{X^k\}}^h_s = \left(\sigma^k_s\right)^2 + w(\left(\sigma^k\right)^2; h, s) + O_p(b_n^{\frac{1}{2}} h^{-1}),$$
uniformly in $s \in [0, t]$ for $t > 0$. The corresponding $\widehat{\int_0^t w_s^2 ds}$ is especially easy to implement in practice.

3.4 Poisson sampling with a random change point

As an illustration, we consider the case when the sampling times are Poisson arrival times but there is a 'change point' in that their arrival intensities will change once the underlying processes hit pre-specified boundaries. More precisely, suppose that $(\mathcal{F}'_t)$-adapted processes $\rho, \mu^k, \sigma^k, W^k, \tau^k, k = 1, 2$ are given on a stochastic basis $(\Omega', \mathcal{F}', (\mathcal{F}'_t), P')$. The processes X^k are defined by (3.1). Let τ^k $(k = 1, 2)$ be $(\mathcal{F}'_t)$-stopping times.

On an auxiliary probability space $(\Omega'', \mathcal{F}'', P'')$, there are random variables $\left(\underline{S}^i\right), \left(\underline{T}^j\right), (\overline{S}^i), (\overline{T}^j)$ that are mutually independent Poisson arrival times with intensity $\underline{\lambda}^k = n\underline{p}^k$, $\underline{p}^k \in (0, \infty)$, $k = 1, 2$, respectively for $\left(\underline{S}^i\right), \left(\underline{T}^j\right)$ and with $\overline{\lambda}^k = n\overline{p}^k$, $\overline{p}^k \in (0, \infty)$, $k = 1, 2$, respectively for $(\overline{S}^i), (\overline{T}^j)$.

We construct the product stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ by
$$\Omega = \Omega' \times \Omega'', \quad \mathcal{F} = \mathcal{F}' \times \mathcal{F}'', \quad \mathcal{F}_t = \mathcal{F}'_t \times \mathcal{F}'' \quad P = P' \times P''.$$
On the new basis the random elements aforementioned can be extended in the usual way; i.e., $W^k(\omega', \omega'') = W^k(\omega')$, $\underline{S}^i(\omega', \omega'') = \underline{S}^i(\omega'')$, $(\omega', \omega'') \in \Omega$, and so forth.

The sampling design $\mathcal{I} = \left(S^i\right)$ for X will be made of $\left(\underline{S}^i\right)$ and $(\overline{S}^i)$ as follows: The sampling for X takes place first at each of $\left(\underline{S}^i\right)$ of the 'first kind'. Sooner or later, the stopping time τ^1 may arrive. Once it arrives, then at $\left(\tau^1 + \frac{1}{\sqrt{n}}\right)$ the stochastic 'alarm clock' of the 'second kind' $(\overline{S}^i)$ will start, at each arrival time of which X will subsequently be sampled. Accordingly, all the arrival times used will form the sampling times $\left(S^i\right)$, from which the sampling design $\mathcal{I}$ will be made. $\left(\underline{T}^j\right)$ and $(\overline{T}^j)$ will form the sampling design $\mathcal{J}$ for Y in the same way. Note that in this construction we control the magnitude of the 'average' interval size $b_n = \frac{1}{n}$ (up to a multiplier).

Formally, $\left(S^i\right)$ is constructed as follows: Put $\tau^1_n = \tau^1 + \frac{1}{\sqrt{n}}$. Define S^i sequentially by
$$\begin{aligned} S^1 &= \inf_{l \in \mathbb{N}} \left\{ \underline{S}^l_{\{\underline{S}^l < \tau^1_n\}}, \tau^1_n + \overline{S}^1 \right\}, \\ S^i &= \inf_{l, m \in \mathbb{N}} \left\{ \underline{S}^l_{\{S^{i-1} < \underline{S}^l < \tau^1_n\}}, \left(\tau^1_n + \overline{S}^m\right)_{\{S^{i-1} < \tau^1_n + \overline{S}^m\}} \right\}, \quad i \geq 2. \end{aligned}$$
$\left(T^j\right)$ is defined by the same way. Here, we have introduced the notation: For a stopping time T (with respect to a filtration $(\mathcal{F}_t)$) and a set $A \in \mathcal{F}_T$, let $T_A(\omega) = T(\omega)$ if $\omega \in A; T_A(\omega) = +\infty$ otherwise. Then, T_A is in fact a stopping time as well (cf., J-S I.1.15, p.4).

In the present situation, the filtration $\left(\mathcal{G}^{(n)}_t\right)$ consists of
$$\mathcal{G}^{(n)}_t \equiv \mathcal{F}'_{(t - n^{-\xi}) \vee 0} \times \mathcal{F}'', \quad t \in \mathbb{R}_+.$$
For a $\left(\mathcal{G}^{(n)}_t\right)$-stopping time T, the 'stopped' σ-field $\mathcal{G}^{(n)}_T$ is defined in the usual way:
$$\mathcal{G}^{(n)}_T = \left\{ A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{G}^{(n)}_t, \text{ all } t \geq 0 \right\}.$$

We claim:

Lemma 3.1 *S^is and T^js are $\left(\mathcal{G}_t^{(n)}\right)$-stopping times.*

Proof. It suffices to consider $\left(S^i\right)$ only. We prove the assertion by induction. Beforehand, first note that $\underline{S}^l$ and $\overline{S}^m$ are $\left(\mathcal{G}_t^{(n)}\right)$-stopping times. Since τ^1 is an $(\mathcal{F}_t)$-stopping time, $\tau_n^1 \equiv \tau^1 + \frac{1}{\sqrt{n}}$ is a $\left(\mathcal{G}_t^{(n)}\right)$-stopping time, hence, $\tau_n^1 + \overline{S}^m$ is a $\left(\mathcal{G}_t^{(n)}\right)$-stopping time as well. Moreover, since $\left\{\underline{S}^l < \tau_n^1\right\} \in \mathcal{G}_{\underline{S}^l}^{(n)}$, $\underline{S}^l_{\left\{\underline{S}^l < \tau_n^1\right\}}$ is a $\left(\mathcal{G}_t^{(n)}\right)$-stopping time. Therefore, S^1 is a $\left(\mathcal{G}_t^{(n)}\right)$-stopping time as well.

Suppose for now that S^{i-1} is a $\left(\mathcal{G}_t^{(n)}\right)$-stopping time. Then, it is true that $\left\{S^{i-1} < \tau_n^1 + \overline{S}^m\right\} \in \mathcal{G}_{\tau_n^1+\overline{S}^m}^{(n)}$. At the same time, $\left\{S^{i-1} < \underline{S}^l < \tau_n^1\right\} = \left\{S^{i-1} < \underline{S}^l\right\} \cap \left\{\underline{S}^l < \tau_n^1\right\} \in \mathcal{G}_{\underline{S}^l}^{(n)}$. These facts implies that S^i is in fact a $\left(\mathcal{G}_t^{(n)}\right)$-stopping time, as asserted. ∎

Consequently, by the argument made for Proposition 1 of , we have, as $n \to \infty$,

$$
\begin{aligned}
nH_n^1(t) &\xrightarrow{P} \frac{2}{\underline{p}^1}\left(\tau^1 \wedge t\right) + \frac{2}{\overline{p}^1}\left(t - \tau^1 \wedge t\right) =: \int_0^t h_s^1 ds,\\
nH_n^2(t) &\xrightarrow{P} \frac{2}{\underline{p}^2}\left(\tau^2 \wedge t\right) + \frac{2}{\overline{p}^2}\left(t - \tau^2 \wedge t\right) =: \int_0^t h_s^2 ds,\\
nH_n^{1\cap 2}(t) &\xrightarrow{P} \frac{2}{\underline{p}^1 + \underline{p}^2}\left(\tau^1 \wedge \tau^2 \wedge t\right) + \frac{2}{\underline{p}^1 + \overline{p}^2}\left(\tau^2 \wedge t - \tau^1 \wedge \tau^2 \wedge t\right)\\
&+\frac{2}{\overline{p}^1 + \underline{p}^2}\left(\tau^1 \wedge t - \tau^1 \wedge \tau^2 \wedge t\right) + \frac{2}{\overline{p}^1 + \overline{p}^2}\left(t - \left(\tau^1 \vee \tau^2\right) \wedge t\right) =: \int_0^t h_s^{1\cap 2} ds,\\
nH_n^{1*2}(t) &\xrightarrow{P} \left(\frac{2}{\underline{p}^1} + \frac{2}{\underline{p}^2}\right)\left(\tau^1 \wedge \tau^2 \wedge t\right) + \left(\frac{2}{\underline{p}^1} + \frac{2}{\overline{p}^2}\right)\left(\tau^2 \wedge t - \tau^1 \wedge \tau^2 \wedge t\right)\\
&+\left(\frac{2}{\overline{p}^1} + \frac{2}{\underline{p}^2}\right)\left(\tau^1 \wedge t - \tau^1 \wedge \tau^2 \wedge t\right) + \left(\frac{2}{\overline{p}^1} + \frac{2}{\overline{p}^2}\right)\left(t - \tau \wedge t\right) =: \int_0^t h_s^{1*2} ds,
\end{aligned}
$$

for every t.

Corollary 3.1 *Under the same hypotheses as Theorem **??**, as $n \to \infty$,*

$$
b_n^{-1/2} M^n \xrightarrow{\mathcal{L}} \int_0^\cdot g_s d\widetilde{W}_s \qquad \textit{(stably)},
$$

where

$$
g_s = \begin{cases}
\sqrt{\left(\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\underline{p}^1} + \frac{2}{\underline{p}^2}\right) + \left(\rho_s\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\underline{p}^1} + \frac{2}{\underline{p}^2} - \frac{2}{\underline{p}^1+\underline{p}^2}\right)} & (s \le \tau^1 \wedge \tau^2)\\
\sqrt{\left(\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\underline{p}^1} + \frac{2}{\overline{p}^2}\right) + \left(\rho_s\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\underline{p}^1} + \frac{2}{\overline{p}^2} - \frac{2}{\underline{p}^1+\overline{p}^2}\right)}1_{\{\tau^1 \le \tau^2\}} & \\
+\sqrt{\left(\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\overline{p}^1} + \frac{2}{\underline{p}^2}\right) + \left(\rho_s\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\overline{p}^1} + \frac{2}{\underline{p}^2} - \frac{2}{\overline{p}^1+\underline{p}^2}\right)}1_{\{\tau^1 > \tau^2\}} & (\tau^1 \wedge \tau^2 < s \le \tau^1 \vee \tau^2)\\
\sqrt{\left(\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\overline{p}^1} + \frac{2}{\overline{p}^2}\right) + \left(\rho_s\sigma_s^1\sigma_s^2\right)^2\left(\frac{2}{\overline{p}^1} + \frac{2}{\overline{p}^2} - \frac{2}{\overline{p}^1+\overline{p}^2}\right)} & (\tau^1 \vee \tau^2 < s)
\end{cases}
$$

and $\widetilde{W}$ is an independent Brownian motion.

Remark: Since τ^1 and τ^2 are $(\mathcal{F}_t)$-stopping times the process (g_t) is $(\mathcal{F}_t)$-adapted, as dictated. An example of such τ^ks are boundary hitting times: $\tau^1 = \inf\left\{t > 0 : X_t > K^1\right\}$, $\tau^2 = \inf\left\{t > 0 : Y_t > K^2\right\}$, $K^1, K^2 \in (0, \infty)$.

Special case: For simplicity, suppose there is only one change point; in particular, let $\tau = \tau^2 \equiv \tau^1 = \inf\left\{t > 0 : X_t > K^1\right\}$. Then, the express for g_t reduces to

$$g_s = \begin{cases} \sqrt{(\sigma_s^1\sigma_s^2)^2\left(\frac{2}{\underline{p}^1} + \frac{2}{\underline{p}^2}\right) + (\rho_s\sigma_s^1\sigma_s^2)^2\left(\frac{2}{\underline{p}^1} + \frac{2}{\underline{p}^2} - \frac{2}{\underline{p}^1 + \underline{p}^2}\right)} & (s \leq \tau) \\ \sqrt{(\sigma_s^1\sigma_s^2)^2\left(\frac{2}{\overline{p}^1} + \frac{2}{\overline{p}^2}\right) + (\rho_s\sigma_s^1\sigma_s^2)^2\left(\frac{2}{\overline{p}^1} + \frac{2}{\overline{p}^2} - \frac{2}{\overline{p}^1 + \overline{p}^2}\right)} & (\tau < s) \end{cases}.$$

Moreover, if $X^1 = \sigma^1 W^1$, then the density of the first passage time τ (to K^1) can be written explicitly as

$$f(t) = \frac{K^1}{\sqrt{2\pi\left((\sigma^1)^2 t\right)^3}} e^{-\frac{\left(K^1\right)^2}{2(\sigma^1)^2 t}},\ t > 0$$

(cf. Karatzas-Shreve (91), p.96).

Acknowledgement

This paper is written with the following financial supports: (to Hayashi) the Grant-in-Aid for Scientific Research (C), Japan Society for the Promotion of Science (JSPS), the Grants for the Advancement of Research at the Graduate Schools of Keio University; (to Yoshida) the Grant-in-Aid for Scientific Research (B), JSPS, the Cooperative Research Program of the Institute of Statistical Mathematics; (to both) the 21st Century COE Program at the University of Tokyo.

References

ALDOUS, D. J., AND G. EAGLESON (1978): On Mixing and Stability of Limit Theorems, *Ann. Probab.* 6(2), 325–331.

HAYASHI, T., AND S. KUSUOKA (2004): Nonsynchronous Covariation Measurement for Continuous Semimartingales, Preprint 2004-21, Grad. Sch. of Math. Sci., Univ. of Tokyo.

HAYASHI, T., AND N. YOSHIDA (2004): Asymptotic Normality of Nonsynchronous Covariance Estimators for Diffusion Processes, preprint.

——— (2005a): Estimating Correlations with Missing Observations in Continuous Diffusion Models, preprint.

——— (2005b): On Covariance Estimation of Non-synchronously Observed Diffusion Processes, *Bernoulli* 11(2), 359–379.

——— (2006): Nonsynchronous Covariance Estimator and Limit Theorem, Research Memorandum No.1020, Institute of Statistical Mathematics.

JACOD, J. (1997): On Continuous Conditional Gausian Martingales and Stable Convergence in Law, in *Séminaire de Probabilitiés XXXI, Lecture Notes in Mathematics 1655.*Berlin; New York: Springer, pp. 232–246.

Jacod, J., and A. N. Shiryaev (2000): *Limit Theorems for Stochastic Processes*, second edn. Berlin; New York: Springer-Verlag.

Rényi, A. (1972): On Stable Sequences of Events, *Sankhya Ser. A* 25, 293–302.

Rootzén, H. (1980): Limit Distributions for the Error in Approximations of Stochastic Integrals, *Ann. Probab.* 80(2), 241–251.