ESTIMATION OF THE LEAD-LAG PARAMETER BETWEEN TWO STOCHASTIC PROCESSES DRIVEN BY FRACTIONAL BROWNIAN MOTIONS

KOHEI CHIBA

ABSTRACT. In this paper, we consider the problem of estimating the lead-lag parameter between two stochastic processes driven by fractional Brownian motions (fBMs) of the Hurst parameter greater than 1/2. First we propose a lead-lag model between two stochastic processes involving fBMs, and then construct a consistent estimator of the lead-lag parameter with possible convergence rate. Our estimator has the following two features. Firstly, we can construct the lead-lag estimator without using the Hurst parameters of the underlying fBMs. Secondly, our estimator can deal with some non-synchronous and irregular observations. We explicitly calculate possible convergence rate when the observation times are (1) synchronous and equidistant, and (2) given by the Poisson sampling scheme. We also present numerical simulations of our results using the R package YUIMA.

1. Introduction

Lead-lag effect is phenomenon that some asset prices follow the fluctuation of others with a small time lag. Recently, milli-second level high frequency trading data are available, and the lead-lag effect is observed at such a fine time scale. For a phenomenological perspective, see Huth and Abergel (2014) and references therein. If we can effectively use the lead-lag effect, it ultimately leads to the prediction of future behavior of stock prices. Hence it is important to analyse the lead-lag effect for developing a trading strategy.

In order to analyse the lead-lag effect statistically, Hoffmann et al. (2013) introduced a regular semimartingale with the lead-lag parameter $\theta \in (-\delta, \delta)$ for $\delta > 0$, and constructed a consistent estimator for the lead-lag parameter θ . Here we briefly review their framework. Before describing the model, we introduce some notations that are used throughout this paper. We follow the notations used in Hoffmann et al. (2013).

Notation 1.1. Let T > 0 be a terminal time and $\delta > 0$ the maximum temporal lead-lag allowed for the model. We set $\Theta = (-\delta, \delta)$. The underlying probability space is denoted by $(\Omega, \mathcal{F}, \mathbb{P})$, and we consider a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [-\delta, T+\delta]}$ satisfying the usual conditions on $(\Omega, \mathcal{F}, \mathbb{P})$. We set $\tau_{-s}(X)_t = X_{t+s}$ for a stochastic process X and a real number $s \in \mathbb{R}$. For a subinterval [a, b] of $[-\delta, T + \delta]$, we set $\mathbb{F}_{[a,b]} = (\mathcal{F}_t)_{t \in [a,b]}$. Finally, for a lead-lag parameter $\theta \in (-\delta, \delta)$, the shifted filtration $(\mathcal{F}_{t-\theta})_{t \in [-\delta+\theta, T+\delta+\theta]}$ is denoted by \mathbb{F}^{θ} .

They defined a regular semimartingale with the lead-lag parameter θ as follows.

Definition 1.2. The two-dimensional stochastic process $\{(X_t, Y_t)\}_{t \in [0, T+\delta]}$ is called a regular semimartingale with the lead-lag parameter $\theta \in [0, \delta)$ if X and Y admit the decomposition $X = X^c + A$ and $Y = Y^c + B$ respectively, where

- (1) the process X^c is a continuous \mathbb{F} -local martingale, and the process Y^c is a continuous $\mathbb{F}^{\theta}_{[0,T+\delta]}$ -local martingale,
- (2) the quadratic variation processes $\langle X^c \rangle$ and $\langle Y^c \rangle$ are absolutely continuous with respect to the Lebesgue measure on $[0, T + \delta]$, and their Radon-Nikodym derivatives have a locally bounded version, and
- (3) the processes A and B are of finite variation on $[0, T + \delta]$.

²⁰¹⁰ Mathematics Subject Classification. 62M09 and 60G22.

Key words and phrases. fractional Brownian motion; lead-lag effect; non-synchronous observations; contrast estimation.

Definition 1.3. If $\theta \in (-\delta, 0]$ and the two-dimensional process (Y, X) is a regular semimartingale with the lead-lag parameter $-\theta$, then the process (X, Y) is called a regular semimartingale with the lead-lag parameter $\theta \in (-\delta, 0]$.

Let us set $\mathcal{T}_n^X = \{0 = s_0^n < s_1^n < \dots < s_{N_X(n)}^n = T + \delta\}$ and $\mathcal{T}_n^Y = \{0 = t_0^n < t_1^n < \dots < t_{N_Y(n)}^n = T + \delta\}$ for some positive integers $N_X(n)$ and $N_Y(n)$. The process X is assumed to be observed at the times in \mathcal{T}_n^X and the process Y at the times in \mathcal{T}_n^Y . Note that the observation times may be non-synchronous, irregular and random. Since we keep in mind high-frequency data, we consider the asymptotics such that the maximum mesh size of the observation times tend to 0 as $n \to \infty$:

$$\lim_{n \to \infty} \max\{|s_i^n - s_{i-1}^n|, |t_j^n - t_{j-1}^n| \mid i = 1, \dots, N_X(n), j = 1, \dots, N_Y(n)\} = 0$$

in probability. They constructed an estimator $\hat{\theta}_n$ for θ by maximizing the *shifted Hayashi-Yoshida covariation contrast* as follows. We define the *shifted Hayashi-Yoshida covariation contrast* $\mathcal{U}_n(\tilde{\theta})$ by

$$\mathcal{U}_{n}(\tilde{\theta}) = \mathbf{1}_{\tilde{\theta} \geq 0} \sum_{i,j:s_{i}^{n} \leq T} (X_{s_{i}^{n}} - X_{s_{i-1}^{n}}) (Y_{t_{j}^{n}} - Y_{t_{j-1}^{n}}) \mathbf{1}_{(s_{i-1}^{n}, s_{i}^{n}] \cap (t_{j-1}^{n} - \tilde{\theta}, t_{j}^{n} - \tilde{\theta}] \neq \emptyset}$$

$$+ \mathbf{1}_{\tilde{\theta} < 0} \sum_{i,j:t_{j}^{n} \leq T} (X_{s_{i}^{n}} - X_{s_{i-1}^{n}}) (Y_{t_{j}^{n}} - Y_{t_{j-1}^{n}}) \mathbf{1}_{(s_{i-1}^{n}, s_{i}^{n}] \cap (t_{j-1}^{n} + \tilde{\theta}, t_{j}^{n} + \tilde{\theta}] \neq \emptyset}.$$

The lead-lag estimator is constructed by maximizing the contrast $\tilde{\theta} \mapsto |\mathcal{U}_n(\tilde{\theta})|$ over a finite subset \mathcal{G}^n of Θ . More precisely, the estimator $\hat{\theta}_n$ is defined by a solution of the equation

$$\hat{\theta}_n = \max_{\tilde{\theta} \in \mathcal{G}^n} |\mathcal{U}_n(\tilde{\theta})|.$$

They proved that the estimator $\hat{\theta}_n$ is consistent under appropriate assumptions on \mathcal{T}_n^X , \mathcal{T}_n^Y and \mathcal{G}^n (see Hoffmann et al. (2013) for detail).

In standard financial theory, it is assumed that there is no arbitrage in the market. Hence semimartingales, which satisfy this no-arbitrage assumption, are regarded as reasonable class for modeling stock prices. However, non-semimartingales, especially ones involving fractional Brownian motion (fBM), also attract attention as a model of stock prices recently. For example, many researchers investigate how arbitrage opportunities arise and how can we exclude them when we use non-semimartingales as a model of stock prices (Bender et al. (2008, 2011); Cheridito (2003); Guasoni (2006); Jarrow et al. (2009) and Rogers (1997) to name but a few). Using fBM in modeling stock prices reflects possible long range dependence property empirically obeserved in some financial time series (for example Cutland et al. (1995); Greene and Fielitz (1977); Hall et al. (2000); Henry (2002); Lo (1991); Teverovsky et al. (1999) and Willinger et al. (1999)). An example of such models is the fractional Black-Scholes model. This model assumes that the stock price dynamics are given by

$$dS_t = \mu S_t \ dt + \sigma S_t dB_t, \ S_0 > 0, \ t \ge 0$$
 (1.1)

where μ and σ are constants and the process $B = (B_t)_{t \geq 0}$ is an fBM with Hurst parameter $H \in (1/2, 1)$. Here the stochastic integral $\sigma \int_0^t S_s dB_s$ can be understood in various ways: if we use the Riemann-Stieltjes integral, then the solution $S^{(1)}$ of the equation (1.1) is

$$S_t^{(1)} = S_0 \exp(\mu t + \sigma B_t),$$

and on the other hand, if we use the Wick-Itô-Skorokhod integral, then the solution $S^{(2)}$ of the equation (1.1) becomes

$$S_t^{(2)} = S_0 \exp\left(\mu t - \frac{\sigma^2}{2}t^{2H} + \sigma B_t\right)$$

(for reference, see Biagini et al. (2008); Mishura (2008) or Sottinen and Valkeila (2003)).

Now a natural question arises: if we consider estimation of the lead-lag parameter between non-semimartingales, especially ones involving fBM, then can we construct a consistent estimator for the lead-lag parameter θ ?

Motivated by this quastion, we consider the problem of estimating the lead-lag parameter θ between two stochastic processes driven by fBMs in this paper. The main aim of this paper is to construct a consistent estimator of the lead-lag parameter θ with possible convergence rate.

Although our analysis is based on the framework by Hoffmann et al. (2013), there are different types of frameworks for lead-lag estimation. For example, Robert and Rosenbaum (2010) studied the lead-lag phenomenon using random matrix theory, and Koike (2016a) investigated the asymptotic structure of the likelihood ratio process when the observed processes are correlated Brownian motions with the microstructure noise.

The rest of this paper is organized as follows. In section 2, we give preliminary results about fractional calculus and stochastic calculus involving fractional Brownian motion, which will be used in subsequent sections of this paper. We also define a covariance structure between two fBMs in Section 2. In section 3, we describe our model. In Section 4, we construct an estimator of the lead-lag parameter θ . We prove its consistency in Section 5. In Section 6, we give an example of our model and a numerical simulation.

2. Preliminaries

2.1. Tools from fractional calculus and stochastic calculus. In this section, we collect some results about fBM that are used in this paper. Then we introduce a correlation between two fBMs. Let us define a fBM at first.

Definition 2.1. Let T > 0 be a positive number. A centered Gaussian process $\{B_t\}_{0 \le t \le T}$ is called fractional Brownian motion (abbreviated fBM) of Hurst parameter $H \in (0,1)$ if its covariance function $R(s,t) = \mathbb{E}\{B(t)B(s)\}, \ s,t \in [0,T]$ satisfies

$$R(t,s) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H})$$
(2.1)

for $s, t \in [0, T]$.

Remark 2.2. In the sequel, we always assume that the Hurst parameter H is greater than 1/2.

When we consider lead-lag relationship between time series, it is reasonable to assume that time series are dependent on each other in appropriate sense. In order to introduce a dependent structure between fBMs, we first define correlated BMs and then construct fBMs from them. Therefore we need a representation of fBM via BM.

Proposition 2.3. Let $W = (W_t)_{t \in [0,T]}$ be a standard BM. Consider the square integrable kernel

$$K_H(t,s) = c_H \mathbf{1}_{(0,t)}(s) s^{1/2-H} \int_s^t (u-s)^{H-3/2} u^{H-1/2} du, \ 0 \le t \le T$$

where $c_H = \sqrt{\frac{H(2H-1)}{B(2-2H,H-1/2)}}$. Then the process $B = (B_t)_{t \in [0,T]}$ defined by

$$B_t = \int_0^T K_H(t,s) \ dW_s \tag{2.2}$$

is an fBM of Hurst parameter H. Here the integral in (2.2) is interpreted as a Wiener integral.

Proof. Since the process B is clearly a centered Gaussian process, it suffices to show (2.1). Later we calculate R(t,s) in a more general setting and (2.1) is obtained as a corollary, see Proposition 2.9 and Corollary 2.10.

Let \mathcal{E} be the set of step functions on [0,T]. Let us consider the linear operator $K_H^* \colon \mathcal{E} \to L^2([0,T])$ defined by

$$(K_H^*\varphi)(s) = \int_0^T \varphi(t) \frac{\partial}{\partial t} K_H(t,s) dt.$$

Note that

$$(K_H^* \mathbf{1}_{[0,t]})(s) = K_H(t,s) \mathbf{1}_{[0,t]}(s).$$

Let \mathcal{H}_H be the completion of \mathcal{E} , with respect to the inner product

$$\langle f, g \rangle_{\mathcal{H}_H} = H(2H - 1) \int_0^T dr \int_0^T du \ f(r)g(u)|r - u|^{2H - 2}$$

= $\langle K_H^* f, K_H^* g \rangle_{L^2([0,T])}$.

The operator K_H^* can be extended to an isometry between \mathcal{H}_H and $L^2([0,T])$, and the extension is also denoted by K_H^* . We set $||f||_{\mathcal{H}}^2 = \langle f, f \rangle_{\mathcal{H}}$ for $f \in \mathcal{H}$. The space \mathcal{H}_H is isomorphic to a Gaussian Hilbert space

$$\mathfrak{H}_H := \overline{\operatorname{span}\left\{B_t; t \in [0, T]\right\}}^{L^2(\mathbb{P})}.$$

We denote the isomorphism between \mathcal{H}_H and \mathfrak{H}_H by B. For proofs of these results, see Nualart (2006). In order to prove Theorem 4.6, we need the notion of multiple Wiener integral and its properties. Let us collect some basic facts concerning multiple Wiener integral which will be used in subsequent sections of this paper.

Definition 2.4. Let \mathcal{H} be a real separable Hilbert space. A centered Gaussian family of random variables $(W(h))_{h\in\mathcal{H}}$ indexed by \mathcal{H} is said to be an isonormal Gaussian process over \mathcal{H} if the covariance function $\mathrm{E}\{W(g)W(h)\} = \langle g,h\rangle_{\mathcal{H}}$ for $g,h\in\mathcal{H}$.

Example 2.5. Let $\mathcal{H} = L^2([0,T];\mathbb{R}^2)$ with an inner product

$$\langle (f_1, g_1), (f_2, g_2) \rangle_{\mathcal{H}} = \langle f_1, g_1 \rangle_{L^2([0,T])} + \langle f_2, g_2 \rangle_{L^2([0,T])}.$$

Let $\tilde{W} = (\tilde{W}^1, \tilde{W}^2)$ be a two-dimensional BM. Then

$$W((f,g)) := \int_0^T f(t) \ d\tilde{W}_t^1 + \int_0^T g(t) \ d\tilde{W}_t^2$$

is an isonormal Gaussian process over \mathcal{H} .

Let p be a positive integer and $h \in \mathcal{H}^{\tilde{\otimes}p}$, where $\mathcal{H}^{\tilde{\otimes}p}$ is the pth symmetric tensor power of \mathcal{H} . After defining an isonormal Gaussian process over \mathcal{H} , we can consider the pth multiple Wiener integral $\mathbb{I}_p(h)$. Note that $\mathbb{I}_1(f)$ coincides with W(f) for $f \in \mathcal{H}$. A detailed discussion about multiple Wiener integral can be found in, for example, Chapter 2 of Nourdin and Peccati (2012). The following properties of multiple Wiener integral are useful.

Proposition 2.6 (Isometry property). Let p and q be positive integers with $1 \leq p \leq q$. For $f \in \mathcal{H}^{\tilde{\otimes}p}$ and $g \in \mathcal{H}^{\tilde{\otimes}p}$, we have

$$\mathbb{E}\{\mathbb{I}_p(f)\mathbb{I}_q(g)\} = \begin{cases} p! \langle f, g \rangle_{\mathcal{H}^{\otimes p}} & \text{if } p = q, \\ 0 & \text{else.} \end{cases}$$

Proof. For example, see Proposition 2.7.5 of Nourdin and Peccati (2012).

Theorem 2.7 (Product formula). Let $p, q \ge 1$ be positive integers. For $f \in \mathcal{H}^{\tilde{\otimes}p}$ and $g \in \mathcal{H}^{\tilde{\otimes}p}$, we have

$$\mathbb{I}_p(f)\mathbb{I}_q(g) = \sum_{r=0}^{p \wedge q} r! \begin{pmatrix} p \\ r \end{pmatrix} \begin{pmatrix} q \\ r \end{pmatrix} \mathbb{I}_{p+q-2r}(f \tilde{\otimes}_r g),$$

where $f \tilde{\otimes}_r g$ denotes the symmetrization of the rth contraction of f and g (for a detail, see Appendix B of Nourdin and Peccati (2012)). In particular, if p = q = 1, then

$$\mathbb{I}_1(f)\mathbb{I}_1(g) = \langle f, g \rangle_{\mathcal{H}} + \mathbb{I}_2(f \tilde{\otimes} g),$$

where $f \tilde{\otimes} g = (1/2)(f \otimes g + g \otimes f)$.

Proof. For example, see Theorem 2.7.10 of Nourdin and Peccati (2012).

Theorem 2.8 (Hypercontractivity). Let q > 0 be a positive number, p be a positive integer, and $h \in \mathcal{H}^{\tilde{\otimes}p}$. Then there exists a constant $C(p,q) \in (0,\infty)$ depending only on p and q such that

$$E\{|\mathbb{I}_p(h)|^q\}^{1/q} \le C(p,q)E\{|\mathbb{I}_p(h)|^2\}^{1/2}.$$

Proof. For example, see Theorem 2.7.2 of Nourdin and Peccati (2012).

2.2. Correlated fBMs. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a stochastic basis supporting a two dimensional standard Brownian motion $W = (\tilde{W}^1, \tilde{W}^2)$. Here we assume that the σ -field \mathcal{F} is complete and the filtration \mathbb{G} satisfies the usual conditions. For $\rho \in [-1, 1]$, we set

$$\begin{cases} W_t^1 = \tilde{W}_t^1, \\ W_t^2 = \rho \tilde{W}_t^1 + \sqrt{1 - \rho^2} \tilde{W}_t^2, \end{cases}$$

for $t \geq 0$. Then the processes W^1 and W^2 are Brownian motions satisfying

$$\langle W^1, W^2 \rangle_t = \rho t, \ t \ge 0.$$

We define correlated fractional Brownian motions on an interval [0, T] by

$$\begin{cases} \tilde{B}_t^1 = \int_0^T K_{H_1}(t,s) \ dW_s^1, \\ \tilde{B}_t^2 = \int_0^T K_{H_2}(t,s) \ dW_s^2, \end{cases}$$
 (2.3)

for $H_1, H_2 \in (1/2, 1)$. Here the integrals in (2.3) are understood in the Wiener sense. In terms of an isonormal Gaussian process, we can write (2.3) as

$$\begin{cases} \tilde{B}_t^1 = W((K_{H_1}^* \mathbf{1}_{[0,t]}, 0)), \\ \tilde{B}_t^2 = W((\rho K_{H_2}^* \mathbf{1}_{[0,t]}, \sqrt{1 - \rho^2} K_{H_2}^* \mathbf{1}_{[0,t]})) \end{cases}$$
(2.4)

(see Example 2.5). For given subinterval $I \subset [0,T]$, we set

$$\begin{cases}
h^{1}(I) = (K_{H_{1}}^{*} \mathbf{1}_{I}, 0), \\
h^{2}(I) = (\rho K_{H_{2}}^{*} \mathbf{1}_{I}, \sqrt{1 - \rho^{2}} K_{H_{2}}^{*} \mathbf{1}_{I}).
\end{cases}$$
(2.5)

Note that since the kernel $K_{H_i}(t,s)$ is a Volterra kernel, i.e., $K_{H_i}(t,s) = 0$ if $s \in [t,T]$, the process \tilde{B}^i is \mathbb{G} -adapted. We can take a modification of the process \tilde{B}^i which has $H_i - \epsilon$ -Hölder continuous trajectory for each $\epsilon \in (0, H_i)$ and is denoted by B^i . The covariance between B^1 and B^2 are given by (2.6) (see Proposition 2.9 below).

Proposition 2.9. Let $B_t^1 = (B_t^1)_{t \in [0,T]}$ and $B^2 = (B_t^2)_{t \in [0,T]}$ be fractional Brownian motions on [0,T] defined as above. Then we have

$$\mathbb{E}\{B_t^1 B_s^2\} = \rho c_{H_1} c_{H_2} \int_0^t du \int_0^s dv \, \beta(u, v) |u - v|^{H_1 + H_2 - 2} u^{H_1 - H_2} v^{H_2 - H_1}, \tag{2.6}$$

where

$$\beta(u,v) = \begin{cases} B(H_1 - 1/2, 2 - H_1 - H_2) & \text{if } u \le v, \\ B(H_2 - 1/2, 2 - H_1 - H_2) & \text{if } v < u. \end{cases}$$

Here B(a,b) denotes the Beta function.

Proof. We have

$$\mathbb{E}\{B_t^1 B_s^2\} = \rho \int_0^T K_{H_1}(t, r) K_{H_2}(s, r) dr
= \rho c_{H_1} c_{H_2} \int_0^{t \wedge s} dr \ r^{1 - H_1 - H_2}
\times \int_r^t du \ (u - r)^{H_1 - 3/2} u^{H_1 - 1/2} \int_r^s dv \ (v - r)^{H_2 - 2/3} v^{H_2 - 1/2}
= \rho c_{H_1} c_{H_2} \int_0^t du \ \int_0^s dv \ u^{H_1 - 1/2} v^{H_2 - 1/2}
\times \int_0^{u \wedge v} dr \ r^{1 - H_1 - H_2} (u - r)^{H_1 - 3/2} (v - r)^{H_2 - 3/2}.$$
(2.7)

Suppose that $u \leq v$. By the change of variables $z = \frac{v-r}{u-r}$ and $x = \frac{v}{uz}$, we obtain

$$\int_{0}^{u} dr \ r^{1-H_{1}-H_{2}} (u-r)^{H_{1}-3/2} (v-r)^{H_{2}-3/2}
= (v-u)^{H_{1}+H_{2}-2} \int_{v/u}^{\infty} dz \ (zu-v)^{1-H_{1}-H_{2}} z^{H_{2}-3/2}
= (v-u)^{H_{1}+H_{2}-2} u^{1/2-H_{2}} v^{1/2-H_{1}} B(H_{1}-1/2, 2-H_{1}-H_{2}).$$
(2.8)

The case where v < u can be obtained similarly:

$$\int_0^v dr \ r^{1-H_1-H_2} (u-r)^{H_1-3/2} (v-r)^{H_2-3/2} = (u-v)^{H_1+H_2-2} u^{1/2-H_2} v^{1/2-H_1} \times B(H_2-1/2, 2-H_1-H_2).$$
(2.9)

Plugging (2.7) and (2.8) into (2.9), we complete the proof.

Corollary 2.10. Assume that $H_1 = H_2 = H$ and $\rho = 1$. Then we have

$$\begin{split} \mathbb{E}\{B_t^1 B_s^1\} &= c_H^2 B(H - 1/2, 2 - 2H) \int_0^t du \int_0^s dv \ |u - v|^{2H - 2} \\ &= \frac{1}{2} (t^{2H} + s^{2H} - |t - s|^{2H}). \end{split}$$

We will use the representation (2.4) in the proof of Theorem 4.6. The next proposition shows that we can always construct an isonormal Gaussian process W over $L^2([0,T];\mathbb{R}^2)$ satisfying (2.4) if we start with fBMs satisfying (2.6).

Proposition 2.11. Let $B^1 = (B_t^1)_{t \in [0,T]}$ and $B^2 = (B_t^2)_{t \in [0,T]}$ be fBMs satisfying (2.6). Then there exists an isonormal Gaussian process W over $L^2([0,T];\mathbb{R}^2)$ such that

$$\begin{cases}
B_t^1 = W((K_{H_1}^* \mathbf{1}_{[0,t]}, 0)), \\
B_t^2 = W((\rho K_{H_2}^* \mathbf{1}_{[0,t]}, \sqrt{1 - \rho^2} K_{H_2}^* \mathbf{1}_{[0,t]}))
\end{cases}$$
(2.10)

holds.

Proof. We define an isonormal Gaussian process over $L^2([0,T])$ by $W^l(\varphi) = B^l((K_{H_l}^*)^{-1}\varphi)$ for $\varphi \in L^2([0,T])$ and l=1,2. Then $B_t^l = W^l(K_{H_l}^* \mathbf{1}_{[0,t]})$ by definition. For step functions f_1, f_2 , it holds that

$$\mathbb{E}\{B^{1}(f_{1})B^{2}(f_{2})\} = \rho \langle K_{H_{1}}^{*}f_{1}, K_{H_{2}}^{*}f_{2} \rangle_{L^{2}([0,T])}. \tag{2.11}$$

We can verify (2.11) holds for any $f_1 \in \mathcal{H}_{H_1}$ and $f_2 \in \mathcal{H}_{H_2}$ by approximating f_l (with respect to the norm $\|\cdot\|_{\mathcal{H}_{H_l}}$) by step functions. Therefore we have

$$\mathbb{E}\{W^1(\varphi)W^2(\psi)\} = \rho\langle\varphi,\psi\rangle_{L^2([0,T])}$$
(2.12)

for $\phi, \psi \in L^2([0,T])$. We set

$$\begin{cases} \tilde{W}^1(\varphi) = W^1(\varphi) \\ \tilde{W}^2(\psi) = (W^2(\psi) - \rho W^1(\psi))/\sqrt{1-\rho^2} \end{cases}$$

and

$$W((\varphi,\psi)) = \tilde{W}^1(\varphi) + \tilde{W}^2(\psi) \tag{2.13}$$

for $\phi, \psi \in L^2([0,T])$. It is easy to check that the isonormal Gaussian process W defined by (2.13) satisfies (2.10).

3. Model assumptions

Now we describe our model in this section. We give precise assumptions on the observed processes and the observation times.

3.1. Lead-lag between stochastic processes involving fBM. Using the correlated fBMs introduced in Section 2.2, we consider the following model for a lead-lag relationship between stochastic processes involving fBM.

Assumption 3.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an underlying probability space. If $\theta \in \Theta_{\geq 0} := [0, \delta)$, then we assume that the two-dimensional process $X=(X_t)_{t\in[0,T+\delta]}=((X_t^1,X_t^2))_{t\in[0,T+\delta]}$ is given by

$$\begin{cases} X_t^1 = X_0^1 + \sigma_1 B_{t+\theta}^1 + A_t^1, \\ X_t^2 = X_0^2 + \sigma_2 B_t^2 + A_t^2, \end{cases}$$

for $t \in [0, T + \delta]$, where

- (A1) X_0^1 and X_0^2 are real-valued random variables, (A2) the processes B^1 and B^2 are fBMs satisfying (2.6), and
- (A3) the drift processes A^1 and A^2 are Lipschitz continuous \mathbb{P} -almost surely.

Moreover, if $\theta \in \Theta_{<0} := (-\delta, 0)$, then we assume that the process $X^* = (X^2, X^1)$ has a representation

$$\begin{cases} X_t^2 = X_0^2 + \sigma_2 B_{t-\theta}^2 + A_t^2, \\ X_t^1 = X_0^1 + \sigma_1 B_t^1 + A_t^1, \end{cases}$$

for $t \in [0, T + \delta]$ where the conditions (A1)-(A3) are in force.

A lead-lag fractional Black-Scholes model, in which we are interested, can be described as follows.

Example 3.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space which supports fBMs B^1 and B^2 satisfying (2.6). Here we assume that the σ -field \mathcal{F} is complete.

Let us consider a system of stochastic differential equations

$$\begin{cases} S_t^1 = S_0^1 + \mu^1 \int_0^t S_s^1 \, ds + \sigma^1 \int_0^t S_s^1 \, dB_s^1, \\ S_t^2 = S_0^2 + \mu^2 \int_0^t S_s^2 \, ds + \sigma^2 \int_0^t S_s^2 \, dB_s^2, \end{cases}$$
(3.1)

where $t \in [T+2\delta]$, $S_0^i > 0$, $\mu^i \in \mathbb{R}$ and $\sigma^i \in \mathbb{R} \setminus \{0\}$ (i=1,2). Then, as we noted in Section 1, the solution of system (3.1) can be written as

$$S_t^i = S_0 \exp(A_t^i + \sigma^i B_t^i),$$

where $A_t^i = \mu^i t$ if we use the Riemann-Stieltjes integral, and $A_t^i = \mu^i t - \frac{(\sigma^i)^2}{2} t^{2H}$ if we consider the Wick-Itô-Skorokhod integral. Note that the function A^i is Lipschitz continuous in either case.

Let $\theta \in [0, \delta)$. We define $X = (X^1, X^2)$ by

$$\begin{cases} X_t^1 = \log S_{t+\theta}^1 = \log S_0^1 + \sigma^1 B_{t+\theta}^1 + A_{t+\theta}^1, \\ X_t^2 = \log S_t^2 = \log S_0^2 + \sigma^2 B_t^2 + A_t^2, \end{cases}$$

for $t \in [0, T + \delta]$. Then the processes X^1 and X^2 satisfy Assumption 3.1.

We give another example of the process that satisfies Assumption 3.1 in Section 6.1.

3.2. **Observation.** Now we give the assumptions on the observation. We consider the problem of estimating the lead-lag parameter $\theta \in \Theta$ from discretely observed X^1 and X^2 . Since we keep in mind that the processes X^1 and X^2 are prices of stocks traded at high frequency, it is natural to consider asymptotics where the number of observations tend to infinity as $n \to \infty$. Therefore we assume that the process X^1 is observed at

$$\mathcal{T}^{1,n} = \{0 = t_0^{1,n} < t_1^{1,n} < \dots < t_{N_1(n)}^{1,n} = T + \delta\},\$$

and the process X^2 at

$$\mathcal{T}^{2,n} = \{0 = t_0^{2,n} < t_1^{2,n} < \dots < t_{N_2(n)}^{2,n} = T + \delta\},$$

where $n, N_1(n)$, and $N_2(n)$ are positive integers. Note that they are in general unevenly spaced, nonsynchronous and may be random and depend on X. Now we introduce some notations that will be useful for describing the assumptions on the observation times $\mathcal{T}^{1,n}$ and $\mathcal{T}^{2,n}$.

Notation 3.3. We set

- $I_i^{k,n} = (t_{i-1}^{k,n}, 1_i^{k,n}], i = 1, \dots, N_k(n) \text{ for } k = 1, 2,$ $|I_i^{k,n}| = t_i^{k,n} t_{i-1}^{k,n}, i = 1, \dots, N_k(n) \text{ for } k = 1, 2,$
- $\mathcal{I}^{k,n} = \{I_i^{k,n} \mid i = 1, \dots, N_k(n)\}$ for k = 1, 2,
- $r_n = \max\{|I_i^{1,n}|, |I_i^{2,n}| \mid i = 1, ..., N_1(n), j = 1, ..., N_2(n)\}$ (the maximum mesh size of the observation times at the step n),
- $I = \inf\{x \mid x \in I\}$ for an interval I,
- $\overline{I} = \sup\{x \mid x \in I\}$ for an interval I,
- $I_{\theta} = \{x + \theta \mid x \in I\}$ for an interval I and $\theta \in \mathbb{R}$,
- $\mathcal{I}_{\theta}^{k,n} = \{(I_i^{k,n})_{\theta} \mid i = 1, \dots, N_k(n)\} \text{ for } k = 1, 2 \text{ and } \theta \in \mathbb{R},$
- $\mathcal{K}(I) = \bigcup_{k \in K_k \cap I \neq \emptyset} K_k$ for given subintervals $\mathcal{K} := \{K_k \mid k\}$ and I (if there are no k such that $K_k \cap I \neq \emptyset$, then we set $\mathcal{K}(I) = \emptyset$),
- $f(I) = f(\overline{I}) f(\underline{I})$ for a function f and an interval I, and
- $V_{>0} = V \cap [0, \infty)$ and $V_{<0} = V \cap (-\infty, 0)$ for a subset V of \mathbb{R} .

Notation 3.4. Let $\{a_i\}_{i\in I}$ be a real sequence indexed by $I\subset \mathbb{Z}_{\geq 0}$. If we omit the range of summation, then we mean the summation over I, that is, $\sum_i a_i = \sum_{i \in I} a_i$. If the index I is empty, then we set $\sum_{i \in I} a_i = 0$. We also denote $\bigcup_{i \in I} A_i$ (resp. $\bigcap_{i \in I} A_i$) as $\bigcup_i A_i$ (resp. $\bigcap_i A_i$) for sets $\{A_i\}_{i \in I}$.

Now we give precise assumptions on the observation times. We give some examples that satisfy Assumption 3.5 in Section 6.2.

Assumption 3.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an underlying probability space. Suppose that there are random variables $\{N_l(n) \in \mathbb{N} \mid l = 1, 2; n \in \mathbb{N}\}$ and $\{t_i^{l,n} \mid i = 0, 1, ..., N_l(n); l = 1, 2; n \in \mathbb{N}\}$ such that, for l = 1, 2and $n \in \mathbb{N}$,

$$0 = t_0^{l,n} < t_1^{l,n} < \dots < t_{N_l(n)}^{l,n} = T + \delta,$$

hold \mathbb{P} -almost surely. We also assume that $\bigcap_{l=1,2}\{i\mid \overline{I_i^{l,n}}\leq T\}\neq\emptyset$ for all $n\in\mathbb{N}$. We set $\mathcal{T}^{l,n}=\{t_i^{l,n}\mid t_i^{l,n}\leq T\}$ $i = 1, \ldots, N_l(n)$.

The observation times $\mathcal{T}^{1,n}$ and $\mathcal{T}^{2,n}$ satisfy following conditions.

- (B1) The σ -algebra $\sigma(\bigcup_{n\in\mathbb{N}}\mathcal{T}^{1,n})$ generated by the observation times $\mathcal{T}^{1,n},\ n\in\mathbb{N}$ is independent of the σ -algebra $\sigma(\bigcup_{n\in\mathbb{N}}\mathcal{T}^{2,n})$ generated by $\mathcal{T}_{2,n},\ n\in\mathbb{N}$. Moreover the σ -algebra $\sigma(\mathcal{T}):=\sigma(\bigcup_{n\in\mathbb{N}}(\mathcal{T}^{1,n}\cup\mathcal{T}^{2,n}))$ $\mathcal{T}^{2,n}$) is independent of the fBMs B^1 and B^2 .
- (B2) There exist positive constants $c_* > 0$ and $\epsilon > 0$ such that

$$\mathbb{P}\left\{\frac{\sum_{i:\overline{I_{i}^{1,n}} \leq T,\underline{I_{i}^{1,n}} \geq \epsilon} |I_{i}^{1,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}}} \geq c_{*}\right\} \to 1$$
(3.2)

and

$$\mathbb{P}\left\{\frac{\sum_{j:\overline{I_{j}^{2,n}} \leq T, I_{\underline{j}}^{2,n} \geq \epsilon} |I_{i}^{2,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}}} \geq c_{*}\right\} \to 1$$
(3.3)

hold as $n \to \infty$.

Moreover, we assume that there exists a deterministic sequence $\{v_n\}_{n\in\mathbb{N}}$ such that $v_n\in(0,\delta)$ for all $n \in \mathbb{N}, v_n \to 0$ as $n \to \infty$ and the following properties (B3) and (B4) hold.

(B3) For any $\mu > 0$, there exists $\gamma(\mu) > \mu$ such that

$$\frac{r_n^{2H_1-1+\mu}}{\sum_{i:\overline{I_i^{1,n}} \le T} |I_i^{1,n}|^{2H_1}} = O_p(v_n^{\gamma(\mu)})$$

and

$$\frac{r_n^{2H_2-1+\mu}}{\sum_{j:\overline{I_j^{2,n}} \le T} |I_j^{2,n}|^{2H_2}} = O_p(v_n^{\gamma(\mu)})$$

hold.

(B4) It holds that

$$\lim_{n \to \infty} v_n^{2 - (H_1 \vee H_2)} (\mathbb{E} \{ \# \mathcal{I}^{1,n} \} + \mathbb{E} \{ \# \mathcal{I}^{2,n} \}) = 0.$$

Remark 3.6. The assumption $\bigcap_{l=1,2}\{i\mid \overline{I_i^{l,n}}\leq T\}\neq\emptyset$ for all $n\in\mathbb{N}$ is in fact not a constraint. Since

$$\bigcup_{l=1,2} \bigcap_{i} \left\{ \overline{I_{i}^{l,n}} > T \right\} \subset \bigcup_{l=1,2} \left\{ \frac{\sum_{i:\overline{I_{i}^{l,n}} \leq T, \underline{I_{i}^{l,n}} \geq \epsilon} |I_{i}^{l,n}|^{H_{1} + H_{2}}}{\sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}}} = 0 \right\}$$

holds, we have $\mathbb{P}\{\bigcap_{l=1,2}\bigcup_i\{\overline{I_i^{l,n}}\leq T\}\}\to 1$ as $n\to\infty$ by (B2). Hence we can assume $\bigcap_{l=1,2}\{i\mid \overline{I_i^{l,n}}\leq T\}\neq\emptyset$ outside an asymptotically negligible set.

Assumption 3.5 is for consistency of the lead-lag estimator defined later in Definition 4.4. These conditions may seem strange in comparison with the assumptions in Hoffmann et al. (2013). In Section 4, we compare Assumption 3.5 with the assumptions in Hoffmann et al. (2013).

4. Construction of an esimator for the lead-lag parameter

Let us now turn to construct an estimator for the lead-lag parameter $\theta \in \Theta$. First we explain the idea for our estimator. Then we define our estimator in Definition 4.4 and state the main theorem of this paper (Theorem 4.6). Finally we compare our consistency result with that of Hoffmann et al. (2013). In particular, we consider differences of the assumptions on the sampling scheme and the grid between our case and the semimartingale case.

4.1. The idea for the lead-lag estimator. To explain the idea for our estimator, let us outline the proof of consistency of the lead-lag estimator in semimartingale case. For now, we assume that the observed process $X = (X^1, X^2)$ is a regular semimartingale with lead-lag parameter $\theta \in [0, \delta)$. We also assume that the drift terms A^1 and A^2 are zero for simplicity. Let $\tilde{\theta}_n \in [0, \delta) \cap \mathcal{G}^n$.

Suppose first that $|\tilde{\theta}_n - \theta|$ converges to 0 rapidly as $n \to \infty$. Then the Hayashi-Yoshida covariance contrast $\mathcal{U}_n(\tilde{\theta}_n)$ is "nearly" the Hayashi-Yoshida estimator:

$$\mathcal{U}_{n}(\tilde{\theta}_{n}) = \sum_{i,j:\overline{I}_{i}^{1,n} \leq T} X^{1}(I_{i}^{1,n}) X^{2}(I_{j}^{2,n}) \mathbf{1}_{I_{i}^{1,n} \cap (I_{j}^{2,n})_{-\tilde{\theta}_{n}} \neq \emptyset}$$

$$= \sum_{i,j:\overline{I}_{i}^{1,n} \leq T} X^{1}(I_{i}^{1,n}) \tau_{-\tilde{\theta}_{n}}(X^{2}) ((I_{j}^{2,n})_{-\tilde{\theta}_{n}}) \mathbf{1}_{I_{i}^{1,n} \cap (I_{j}^{2,n})_{-\tilde{\theta}_{n}} \neq \emptyset}.$$

Therefore, the limit

$$\mathcal{U}_n(\tilde{\theta}_n) \to^p \langle X, \tau_{-\theta}(Y) \rangle_T$$
 (4.1)

holds as $n \to \infty$. This is Proposition 4 of Hoffmann et al. (2013). For the properties of the Hayashi-Yoshida estimator, we refer to Hayashi and Kusuoka (2008); Hayashi and Yoshida (2005, 2008) for example.

On the other hand, we can prove

$$\sup_{\tilde{\theta} \in \mathcal{G}^n \cap \{\tilde{\theta} \in \Theta | |\tilde{\theta} - \theta| \ge 2v_n\}} |\mathcal{U}_n(\tilde{\theta})| \to^p 0$$
(4.2)

as $n \to \infty$. This is a consequence of Proposition 3 of Hoffmann et al. (2013). If $|\tilde{\theta} - \theta| \ge 2v_n$, then the contrast function $\mathcal{U}_n(\tilde{\theta})$ can be rewritten as the terminal value of a martingale outside of an asymptotically negligible set. To handle the probability of the supremum, the Bürkholder-Davis-Gundy inequality is used in Hoffmann et al. (2013).

By (4.1) and (4.2), the contrast function $\mathcal{U}_n(\tilde{\theta})$ has a peak at $\tilde{\theta} = \theta$ asymptotically on the set $\{\langle X, \tau_{-\theta}(Y) \rangle_T \neq 0\}$. Hence we can conclude that the maximum contrast estimator is consistent.

Now we go back to the original problem. Let us assume that the observed process X satisfies Assumption 3.1. We essentially follow the strategy above. However, we can not use the Hayashi-Yoshida covariance contrast function directly. This is because the Hayashi-Yoshida estimator may converge to zero if the observed process is not a semimartingale. For example, let us set $A^1 = A^2 = 0$, $\sigma_1 = \sigma_2 = 0$, $H_1 = H_2 = H$ and $\rho = 1$ in Assumption 3.1 and $\mathcal{T}^{l,n} = \{(1+\delta)i/n \mid i=1,\ldots,n\}$ for l=1,2 in Assumption 3.5. We also assume $\theta = 0$ for simplicity. Then we obtain

$$\mathcal{U}_n(0) = \frac{1}{n^{2H-1}} n^{2H-1} \sum_{i=1}^{\lfloor n/(1+\delta) \rfloor} (B_{i/n}^1 - B_{(i-1)/n}^1)^2$$

$$\to^p 0$$

as $n \to \infty$ by ergodicity and self-similarity of fractional Gaussian noise. To ensure that the contrast function is away from zero when $|\tilde{\theta}_n - \theta|$ converges to 0 rapidly as $n \to \infty$, we consider the "correlation version" of the contrast $\mathcal{U}_n(\tilde{\theta})$. We consider the *shifted Hayashi-Yoshida correlation contrast function* $\mathcal{U}_n^{\text{cor}} \colon \Theta \to \mathbb{R}$ defined by

$$\begin{split} \mathcal{U}_{n}^{\text{cor}}(\tilde{\theta}) &= \mathbf{1}_{\Theta_{\geq 0}}(\tilde{\theta}) \frac{\sum_{i,j:\overline{I_{i}^{1,n}} \leq T} X^{1}(I_{i}^{1,n}) X^{2}(I_{j}^{2,n}) \mathbf{1}_{I_{i}^{1,n} \cap (I_{j}^{2,n})_{-\tilde{\theta}} \neq \emptyset}}{\sqrt{\sum_{i:\overline{I_{i}^{1,n}} \leq T} X^{1}(I_{i}^{1,n})^{2}} \sqrt{\sum_{j} X^{2}(I_{j}^{2,n})^{2}}} \\ &+ \mathbf{1}_{\Theta_{<0}}(\tilde{\theta}) \frac{\sum_{i,j:\overline{I_{j}^{2,n}} \leq T} X^{1}(I_{i}^{1,n}) X^{2}(I_{j}^{2,n}) \mathbf{1}_{(I_{i}^{1,n})_{\tilde{\theta}} \cap I_{j}^{2,n} \neq \emptyset}}{\sqrt{\sum_{i} X^{1}(I_{i}^{1,n})^{2}} \sqrt{\sum_{j:\overline{I_{j}^{2,n}} \leq T} X^{2}(I_{j}^{2,n})^{2}}}. \end{split}$$

Then we can ensure that the contrast $\mathcal{U}_n^{\text{cor}}(\tilde{\theta}_n)$ is away from zero when $|\tilde{\theta}_n - \theta|$ converges to 0 rapidly as $n \to \infty$.

We have another problem. In the semimartingale case, the proof of (4.2) relies on the martingale property of the observed process. However, we cannot expect it in our case. We overcome this problem by exploiting the Gaussian property of the observed process: the Wiener chaos decomposition and the hypercontractivity.

4.2. **Definition of the lead-lag estimator.** Following the discussion above, we now define the lead-lag estimator in our case.

Definition 4.1. We define the shifted Hayashi-Yoshida correlation contrast function $\mathcal{U}_n^{\text{cor}} : \Theta \to \mathbb{R}$ by

$$\mathcal{U}_{n}^{\text{cor}}(\tilde{\theta}) = \mathbf{1}_{\Theta \geq 0}(\tilde{\theta}) \frac{\sum_{i,j:\overline{I_{i}^{1,n}} \leq T} X^{1}(I_{i}^{1,n}) X^{2}(I_{j}^{2,n}) \mathbf{1}_{I_{i}^{1,n} \cap (I_{j}^{2,n})_{-\tilde{\theta}} \neq \emptyset}}{\sqrt{\sum_{i:\overline{I_{i}^{1,n}} \leq T} X^{1}(I_{i}^{1,n})^{2}} \sqrt{\sum_{j} X^{2}(I_{j}^{2,n})^{2}}} + \mathbf{1}_{\Theta < 0}(\tilde{\theta}) \frac{\sum_{i,j:\overline{I_{j}^{2,n}} \leq T} X^{1}(I_{i}^{1,n}) X^{2}(I_{j}^{2,n}) \mathbf{1}_{(I_{i}^{1,n})_{\tilde{\theta}} \cap I_{j}^{2,n} \neq \emptyset}}{\sqrt{\sum_{i} X^{1}(I_{i}^{1,n})^{2}} \sqrt{\sum_{j:\overline{I_{j}^{2,n}} \leq T} X^{2}(I_{j}^{2,n})^{2}}}.$$
(4.3)

if $\sum_{i:\overline{I_i^{l,n}}\leq T}X^l(I_i^{l,n})^2\neq 0$ for l=1,2. When $\sum_{i:\overline{I_i^{l,n}}\leq T}X^l(I_i^{l,n})^2=0$ for some l, we set $\mathcal{U}_n^{\mathrm{cor}}(\tilde{\theta})=0$. Hereafter, we denote the contrast function $\mathcal{U}_n^{\mathrm{cor}}(\tilde{\theta})$ in (4.3) by $\mathcal{U}_n(\tilde{\theta})$ for notational simplicity.

Our estimator $\hat{\theta}_n$ is obtained by maximizing the contrast function $|\mathcal{U}_n(\tilde{\theta})|$ over a finite grid \mathcal{G}^n in the parameter space Θ . The assumptions on the grid \mathcal{G}^n are as follows.

Assumption 4.2. Let \mathcal{G}^n be a finite subset of the parameter space Θ such that (C1) $0 \in \mathcal{G}^n$,

- (C2) $\#\mathcal{G}^n = O(v_n^{-\beta})$ for some $\beta > 0$, and
- (C3) for some deterministic sequence $\rho_n > 0$ and positive constant $\varsigma \in (0, 1 (H_1 \vee H_2))$, we have

$$\Theta \subset \bigcup_{\tilde{\theta} \in \mathcal{G}^n} [\tilde{\theta} - \rho_n, \tilde{\theta} + \rho_n]$$

and

$$\lim_{n \to \infty} \rho_n^{1-\varsigma}(\mathbb{E}\{\#\mathcal{I}^{1,n}\} + \mathbb{E}\{\#\mathcal{I}^{2,n}\}) = 0.$$

Here the sequence $(v_n)_{n\in\mathbb{N}}$ is from Assumption 3.5.

Example 4.3. An example of the grid \mathcal{G}^n satisfying Assumption 4.2 is

$$\mathcal{G}^n = \{ v_n^{2 - (H_1 \vee H_2) + \epsilon} m \mid m \in \mathbb{Z} \} \cap \Theta$$

for any $\epsilon > 0$ by (B4).

Now we are ready to define our lead-lag estimator.

Definition 4.4. Let $\mathcal{U}_n(\hat{\theta}_n)$ be the shifted Hayashi-Yoshida correlation contrast function defined in Definition 4.1. We define the lead-lag estimator $\hat{\theta}_n$ as a solution of

$$|\mathcal{U}_n(\hat{\theta}_n)| = \max_{\tilde{\theta} \in \mathcal{G}^n} |\mathcal{U}_n(\tilde{\theta})|.$$

Remark 4.5. We can assume that $\hat{\theta}_n$ is chosen to be measurable. For example, let $\hat{\theta}_n^*$ denote the largest element in the set $\arg\max_{\tilde{\theta}\in\mathcal{G}^n}|\mathcal{U}_n(\tilde{\theta})|$. Then we have

$$\{\hat{\theta}_n^* \le r\} = \left\{ \max_{\tilde{\theta} \in \mathcal{G}^n \cap (r, \delta)} |\mathcal{U}_n(\tilde{\theta})| < \max_{\tilde{\theta} \in \mathcal{G}^n \cap [0, r]} |\mathcal{U}_n(\tilde{\theta})| \right\}$$

for any $r \in [0, \delta)$, and hence $\hat{\theta}_n^*$ is measurable.

Let us state our main theorem.

Theorem 4.6. Suppose that

- (a) the processes X^1 and X^2 are defined as in Assumption 3.1 with $\rho \neq 0$,
- (b) the observation times $\mathcal{T}^{1,n}$ and $\mathcal{T}^{2,n}$ satisfy Assumption 3.5, and
- (c) the grid \mathcal{G}^n satisfies Assumption 4.2.

Then the estimation $\hat{\theta}_n$ satisfies

$$v_n^{-1}(\hat{\theta}_n - \theta) \to^p 0 \tag{4.4}$$

as $n \to \infty$. Here the symbol \to^p denotes convergence in probability. Note that the sequence $(v_n)_{n \in \mathbb{N}}$ is from Assumption 3.5.

- Remark 4.7. (1) It is interesting to consider removing the assumption $H \in (1/2, 1)$. By constructing correlated fBMs from correlated BMs as in Section 2.2, we could consider a lead-lag model between two processes involving fBMs with any Hurst parameter. However, we use the representation (2.6) of the covariance when we prove the consistency of the lead-lag estimator.
- (2) As we shall see in Section 6, the convergence rate v_n essentially equals n^{-1} when the observation times are synchronous and equispaced with n points or given by the Poisson sampling scheme with frequency proportional to n. However, as mentioned in Hoffmann et al. (2013), the convergence rate v_n could be improved in our case. In fact, Koike (2016a) analyzed the lead-lag model based on Gaussian likelihood and suggested the possibility such that the convergence rate becomes $n^{-3/2}$ when the observed processes are correlated Brownian motions with the microstructure noise and the observation is synchronous and evenly spaced with n points.

(3) For financial high-frequency data, the true price process are considered to be contaminated by market microstructure noise, although we do not take into account the presence of it in this paper. In the semimartingale case, it is suggested in Hoffmann et al. (2013) that using a contrast function based on the pre-averaged Hayashi-Yoshida estimator gives a consistent lead-lag estimator even if microstructure noise is present. However, it is unclear whether the same approach works or not in our non-semimartingale case. For the properties of the pre-averaged Hayashi-Yoshida estimator, see Koike (2014, 2016b) and references therein.

Before proceeding the proof of Theorem 4.6, let us compare our consistency result with the one of Hoffmann et al. (2013).

The condition (B1) in Assumption 3.5 is a counterpart of the condition [B2] of Hoffmann et al. (2013). Although they allowed the observation times to be dependent on the process, we content ourselves with the assumption that the observation times are independent of the process.

The condition (B2) has no counterpart in Hoffmann et al. (2013). This is because (B2) is naturally satisfied in semimartingale case. Let us set $H_1 = H_2 = 1/2$ in (B2). Then (B2) becomes as follows.

(B2') There exist positive constants $c_* > 0$ and $\epsilon > 0$ such that

$$\mathbb{P}\left\{\sum_{i:\overline{I_i^{l,n}} \leq T, I_i^{l,n} \geq \epsilon} |I_i^{l,n}| \geq c_*\right\} \to^p 1$$

holds as $n \to \infty$ for l = 1, 2.

Let us assume $r_n = o_p(1)$. Since we consider high-frequency asymptotics, this is a minimum requirement. Then (B2') is derived from $r_n = o_p(1)$ since the inequality

$$\mathbb{P}\left\{ \sum_{i: \overline{I_{i}^{l,n}} \leq T, I_{i}^{l,n} \geq T/2} |I_{i}^{l,n}| \geq T/4 \right\} \geq \mathbb{P}\{r_{n} < T/8\}$$

holds.

The condition (B3) is a counterpart of the condition [B1] of Hoffmann et al. (2013), that is, $r_n = o_p(v_n)$. If $H_1 = H_2 = 1/2$, then the condition (B3) becomes as follows.

(B3') For any $\mu > 0$ there exists $\gamma(\mu) > \mu$ such that $r_n^{\mu} = O_p(v_n^{\gamma(\mu)})$ holds as $n \to \infty$.

The condition (B3') clearly implies $r_n = o_p(v_n)$. Hence (B3) is a bit stronger than the condition [B1] of Hoffmann et al. (2013).

The condition (B4) also has no counterpart in the semimartingale case. Unfortunately, the author has no explanation for the condition (B4). We use the condition (B4) only when we show that the contrast $\mathcal{U}_n(\tilde{\theta}_n)$ is away from zero if $|\tilde{\theta}_n - \theta|$ converges to 0 rapidly as $n \to \infty$. In semimartingale case, the consistency of the Hayashi-Yoshida estimator is exploited here.

Finally, Assumption 4.2 is a counterpart of [B3] in Hoffmann et al. (2013).

5. Proof of Theorem 4.6

In this section, we give the proof of Theorem 4.6. Without loss of generality, we can assume that $\theta \in \Theta_{\geq 0}$. By Assumption 3.1, the case where $\theta \in \Theta_{< 0}$ is equivalent to the case of $-\theta \in \Theta_{\geq 0}$ with indices 1 and 2 interchanged. Therefore the proof when $\theta \in \Theta_{\geq 0}$ can be applied in the case of $\theta \in \Theta_{< 0}$. First we prove Theorem 4.6 assuming that $A^1 = A^2 = 0$, and then we remove this assumption.

Notation 5.1. Let $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ be sequences of real numbers. If there exists a positive constant c>0 independent of n such that $a_n\leq cb_n$ for all n, then we write $a_n\lesssim b_n$.

5.1. **Proof of Theorem 4.6 when** $A^1 = A^2 = 0$. Under the additional assumption $A^1 = A^2 = 0$, the contrast function $\mathcal{U}_n(\tilde{\theta})$ becomes

$$\mathcal{U}_{n}(\tilde{\theta}) = \mathbf{1}_{\Theta \geq 0}(\tilde{\theta}) \frac{\sum_{i,j:\overline{I_{i}^{1,n}} \leq T} B^{1}((I_{i}^{1,n})_{\theta}) B^{2}(I_{j}^{2,n}) \mathbf{1}_{I_{i}^{1,n} \cap (I_{j}^{2,n})_{-\tilde{\theta}} \neq \emptyset}}{\sqrt{\sum_{i:\overline{I_{i}^{1,n}} \leq T} B^{1}((I_{i}^{n})_{\theta})^{2}} \sqrt{\sum_{j} B^{2}(I_{j}^{2,n})^{2}}} + \mathbf{1}_{\Theta < 0}(\tilde{\theta}) \frac{\sum_{i,j:\overline{I_{j}^{2,n}} \leq T} B^{1}((I_{i}^{1,n})_{\theta}) B^{2}(I_{j}^{2,n}) \mathbf{1}_{(I_{i}^{1,n})_{\tilde{\theta}} \cap I_{j}^{2,n} \neq \emptyset}}{\sqrt{\sum_{i} B^{1}((I_{i}^{n})_{\theta})^{2}} \sqrt{\sum_{j:\overline{I_{j}^{2,n}} \leq T} B^{2}(I_{j}^{2,n})^{2}}} \tag{5.1}$$

if the denominators appearing in (5.1) are not equal zero. Note that they are not equal zero \mathbb{P} -almost surely. We start with the following lemma.

Lemma 5.2. Let J denote [0,T] or $[0,T+\delta]$. Then we have, for all $a \in [0,\delta)$,

$$\frac{\sum_{i:\overline{I_{i}^{l,n}}\in J} B^{l}((I_{i}^{l,n})_{a})^{2}}{\sum_{i:\overline{I_{i}^{l,n}}\in J} |I_{i}^{l,n}|^{2H_{l}}} \to^{p} 1$$
(5.2)

as $n \to \infty$ for l = 1, 2.

Proof. First we define an auxiliary set $A_n(K)$ by

$$A_n(K) = \bigcap_{l=1,2} \left\{ \frac{r_n^{2H_l - 1 + \mu} v_n^{-\gamma(\mu)}}{\sum_{i: \overline{I_i^{l,n}} \le T} |I_i^{l,n}|^{2H_l}} \le K \right\}.$$

Since $\sum_{i:\overline{I_i^{l,n}} \leq T} |I_i^{l,n}|^{2H_l} \leq Tr_n^{2H_l-1}$ holds, we have

$$r_n \leq (TK)^{1/\mu} v_n^{\gamma(\mu)/\mu},$$

on the set $A_n(K)$. In particular, it holds that $(r_n/v_n)1_{A_n(K)} \to 0$ as $n \to \infty$ (recall that $\gamma(\mu) > \mu$). Thanks to (B3), it holds that $\sup_n \mathbb{P}(A_n(K)^c) < \epsilon$ for each $\epsilon > 0$ if $K = K(\epsilon)$ is sufficiently large. Let $R_1^{l,n}$ denote $(\sum_{i:\overline{I_i^{l,n}}\in J} B^l((I_i^{l,n})_a)^2)/(\sum_{i:\overline{I_i^{l,n}}\in J} |I_i^{l,n}|^{2H_l})$ for simplicity. For any positive number $\epsilon_1 > 0$, we have

$$\mathbb{P}\left\{\left|R_1^{l,n} - 1\right| \ge \epsilon_1\right\} \le \mathbb{P}\left\{\left|R_1^{l,n} - 1\right| \ge \epsilon_1, \ A_n(K)\right\} + \mathbb{P}\left\{A_n(K)^c\right\}.$$

Hence we obtain

$$\mathbb{P}\left\{\left|R_{1}^{l,n}-1\right| \geq \epsilon_{1}\right\} \leq \mathbb{P}\left\{\left|R_{1}^{l,n}-1\right| \geq \epsilon_{1}, \ A_{n}(K(\epsilon_{2}))\right\} + \epsilon_{2}$$

$$\leq \epsilon_{1}^{-2} \mathbb{E}\left\{\left|R_{1}^{l,n}-1\right|^{2} \mathbf{1}_{A_{n}(K(\epsilon_{2}))}\right\} + \epsilon_{2}.$$
(5.3)

By using [B1] and conditioning, we can calculate the expectation in (5.3) as

$$\begin{split} & \mathbb{E}\left\{\left|R_{1}^{l,n}-1\right|^{2}\mathbf{1}_{A_{n}(K(\epsilon_{2}))}\right\} \\ & = \mathbb{E}\left\{\mathbb{E}\left\{\left(\frac{\sum_{i:\overline{I_{i}^{l,n}}\in J}(B^{l}((I_{i}^{l,n})_{a})^{2}-|I_{i}^{l,n}|^{2H_{l}})}{\sum_{i:\overline{I_{i}^{l,n}}\in J}|I_{i}^{l,n}|^{2H_{l}}}\right)^{2}\mathbf{1}_{A_{n}(K(\epsilon_{2}))}\middle|\sigma(\mathcal{T})\right\}\right\} \\ & = \mathbb{E}\left\{\frac{\mathbf{1}_{A_{n}(K(\epsilon_{2}))}\mathbb{E}\left\{\left(\sum_{i:\overline{I_{i}^{l,n}}\in J}\left(B^{l}((I_{i}^{l,n})_{a})^{2}-|I_{i}^{l,n}|^{2H_{l}}\right)\right)^{2}\middle|\sigma(\mathcal{T})\right\}}{\left(\sum_{i:\overline{I_{i}^{l,n}}\in J}|I_{i}^{l,n}|^{2H_{l}}\right)^{2}}\right\}. \end{split}$$

Since $B^l((I_i^{l,n})_a)^2 - |(I_i^{l,n})|^{2H_l} = B^l((I_i^{l,n})_a)^2 - |(I_i^{l,n})_a|^{2H_l}$ coincides with the multiple Wiener integral $\mathbb{I}_2(h^l((I_i^{l,n})_a)^{\otimes 2})$ (recall that the definition of $h^l(I)$ is given in (2.5)) by Theorem 2.7, we have

$$\mathbb{E}\left\{\left(\sum_{i:\overline{I_{i}^{l,n}}\in J}\left(B^{l}((I_{i}^{l,n})_{a})^{2}-|I_{i}^{l,n}|^{2H_{l}}\right)\right)^{2}\middle|\sigma(\mathcal{T})\right\} = \mathbb{E}\left\{\mathbb{I}_{2}^{2}\left(\sum_{i:\overline{I_{i}^{l,n}}\in J}h^{l}((I_{i}^{l,n})_{a})^{\otimes 2}\right)\middle|\sigma(\mathcal{T})\right\}$$

$$=\sum_{i,j}\langle h^{l}((I_{i}^{l,n})_{a}), h^{l}((I_{j}^{l,n})_{a})\rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})}^{2}$$

$$=\sum_{i,j}\langle \mathbf{1}_{(I_{i}^{l,n})_{a}}, \mathbf{1}_{(I_{j}^{l,n})_{a}}\rangle_{\mathcal{H}_{H_{l}}}$$

Since $|\langle \mathbf{1}_{(I_i^{l,n})_a}, \mathbf{1}_{(I_j^{l,n})_a} \rangle_{\mathcal{H}_{H_l}}| \leq r_n^{2H_l}$ because of Cauchy-Schwarz inequality, we have

$$\mathbb{E}\left\{\left(\sum_{i:\overline{I_i^{l,n}}\in J}\left(B^l((I_i^{l,n})_a)^2-|I_i^{l,n}|^{2H_l}\right)\right)^2\middle|\sigma(\mathcal{T})\right\} \leq r_n^{2H_l}\sum_{i,j}\langle\mathbf{1}_{(I_i^{l,n})_a},\mathbf{1}_{(I_j^{l,n})_a}\rangle_{\mathcal{H}_{H_l}}$$
$$\lesssim r_n^{2H_l}.$$

Note that $\langle \mathbf{1}_{(I_i^{l,n})_a}, \mathbf{1}_{(I_i^{l,n})_a} \rangle_{\mathcal{H}_{H_l}}$ is non-negative. Plugging this into (5.3), we obtain

$$\mathbb{P}\left\{\left|R_{1}^{l,n}-1\right| \geq \epsilon_{1}\right\} \lesssim \epsilon_{1}^{-2} \mathbb{E}\left\{\left(\frac{r_{n}^{H_{l}}}{\sum_{i:\overline{I_{i}^{l,n}} \in J} |I_{i}^{l,n}|^{2H_{l}}}\right)^{2} \mathbf{1}_{A_{n}(K(\epsilon_{2}))}\right\} + \epsilon_{2}$$

$$= \epsilon_{1}^{-2} \mathbb{E}\left\{\left(\frac{r_{n}^{2H_{l}-1+\mu} v_{n}^{-\gamma(\mu)}}{\sum_{i:\overline{I_{i}^{l,n}} \in J} |I_{i}^{l,n}|^{2H_{l}}} r_{n}^{1-H_{l}-\mu} v_{n}^{\gamma(\mu)}\right)^{2} \mathbf{1}_{A_{n}(K(\epsilon_{2}))}\right\} + \epsilon_{2}$$

$$\lesssim \epsilon_{1}^{-2} (TK(\epsilon_{2}))^{2/\mu} v_{n}^{2(1-H_{l})} + \epsilon_{2},$$

if we choose sufficiently small $\mu > 0$ such that $1 - H - \mu > 0$ holds. Since ϵ_1 and ϵ_2 are arbitrary, we obtain (5.2) by letting $n \to \infty$.

Let us set

$$\begin{split} R_2^n(\tilde{\theta}) &= \mathbf{1}_{\Theta \geq 0}(\tilde{\theta}) \sum_{i,j: \overline{I_i^{1,n}} \leq T} B^1((I_i^{1,n})_{\theta}) B^2(I_j^{2,n}) \mathbf{1}_{I_i^{1,n} \cap (I_j^{2,n})_{-\tilde{\theta}} \neq \emptyset} \\ &+ \mathbf{1}_{\Theta < 0}(\tilde{\theta}) \sum_{i,j: \overline{I_j^{2,n}} \leq T} B^1((I_i^{1,n})_{\theta}) B^2(I_j^{2,n}) \mathbf{1}_{(I_i^{1,n})_{\tilde{\theta}} \cap (I_j^{2,n}) \neq \emptyset}, \end{split}$$

and

$$\begin{split} D^{n}(\tilde{\theta}) &= \mathbf{1}_{\Theta_{\geq 0}}(\tilde{\theta}) \sqrt{\sum_{i: \overline{I_{i}^{1,n}} \leq T} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}} \\ &+ \mathbf{1}_{\Theta_{< 0}}(\tilde{\theta}) \sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j: \overline{I_{j}^{2,n}} \leq T} |I_{j}^{2,n}|^{2H_{2}}}. \end{split}$$

Lemma 5.2 reduces calculation of $\mathcal{U}_n(\tilde{\theta})$ to that of $R_2^n(\tilde{\theta})/D^n(\tilde{\theta})$. Next we show that $R_2^n(\tilde{\theta})/D^n(\tilde{\theta})$ goes to zero as $n \to \infty$ if $\tilde{\theta}$ is distant from θ .

Proposition 5.3. We have, for any $\epsilon > 0$,

$$\sup_{\mathcal{G}^n \cap \{\tilde{\theta} \in \Theta \mid |\tilde{\theta} - \theta| \ge \epsilon v_n\}} \left| \frac{R_2^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \to^p 0 \tag{5.4}$$

as $n \to \infty$.

Proof. The set $\mathcal{G}^n \cap \{\tilde{\theta} \in \Theta \mid |\tilde{\theta} - \theta| \ge \epsilon v_n\}$ can be decomposed as

$$\mathcal{G}^{n} \cap \{\tilde{\theta} \in \Theta \mid |\tilde{\theta} - \theta| \ge \epsilon v_{n}\} = (\mathcal{G}_{\ge 0}^{n} \cap \{\tilde{\theta} \in \Theta \mid \tilde{\theta} \ge \theta + \epsilon v_{n}\}) \cup (\mathcal{G}_{\ge 0}^{n} \cap \{\tilde{\theta} \in \Theta \mid \tilde{\theta} \le \theta - \epsilon v_{n}\})$$
$$\cup (\mathcal{G}_{< 0}^{n} \cap \{\tilde{\theta} \in \Theta \mid \tilde{\theta} \le \theta - \epsilon v_{n}\})$$
$$=: \mathcal{G}_{1}^{n} \cup \mathcal{G}_{2}^{n} \cup \mathcal{G}_{3}^{n}.$$

Clearly (5.4) is equivalent to

$$\sup_{\mathcal{G}_i^n} \left| \frac{R_2^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \to^p 0 \tag{5.5}$$

for i = 1, 2, 3. We only prove (5.5) when i = 1. The other cases are similar.

Let us assume $H_1 \leq H_2$ for the moment. Since $\tilde{\theta} \in \mathcal{G}_1^n$, we have

$$\begin{split} R_2^n(\tilde{\theta}) &= \sum_{i,j: \overline{I_i^{1,n}} \leq T} B^1((I_i^{1,n})_{\theta}) B^2((I_j^{2,n})) \mathbf{1}_{(I_i^{1,n})_{\theta} \cap (I_j^{2,n})_{\theta - \tilde{\theta}} \neq \emptyset} \\ &= \sum_{i: \overline{I_i^{1,n}} \leq T} B^1((I_i^{1,n})_{\theta}) B^2(\mathcal{I}_{\theta - \tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta} - \theta}) \\ &= \sum_{i: \overline{I_i^{1,n}} \leq T} \langle h^1((I_i^{1,n})_{\theta}), h^2(\mathcal{I}_{\theta - \tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta} - \theta}) \rangle_{L^2([0,T + 2\delta]; \mathbb{R}^2)} \\ &+ \sum_{i: \overline{I_i^{1,n}} \leq T} \left(B^1((I_i^{1,n})_{\theta}) B^2(\mathcal{I}_{\theta - \tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta} - \theta}) \rangle_{L^2([0,T + 2\delta]; \mathbb{R}^2)} \right) \\ &- \langle h^1((I_i^{1,n})_{\theta}), h^2(\mathcal{I}_{\theta - \tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta} - \theta}) \rangle_{L^2([0,T + 2\delta]; \mathbb{R}^2)} \right) \\ &=: R_3^n(\tilde{\theta}) + R_4^n(\tilde{\theta}). \end{split}$$

Hence it suffices to show

$$\sup_{\tilde{\theta} \in \mathcal{G}_1^n} \left| \frac{R_3^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \to^p 0 \tag{5.6}$$

and

$$\sup_{\tilde{\theta} \in \mathcal{G}_{i}^{n}} \left| \frac{R_{4}^{n}(\tilde{\theta})}{D^{n}(\tilde{\theta})} \right| \to^{p} 0 \tag{5.7}$$

as $n \to \infty$.

The limit (5.6). It suffices to show

$$\sup_{\tilde{\theta} \in \mathcal{G}_1^n} \left| \frac{R_3^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \mathbf{1}_{A_n(K)} \to 0 \tag{5.8}$$

as $n \to \infty$ for each K > 0. By (2.6) in Proposition 2.9, it holds that

$$|R_3^n(\tilde{\theta})| = |\rho| c_{H_1} c_{H_2} \sum_{i: \overline{I_i^{1,n}} \le T} \int_{(I_i^{1,n})_{\theta}} du \int_{\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta}-\theta}} dv \ \beta(u,v)$$
$$\times |u-v|^{H_1+H_2-2} u^{H_1-H_2} v^{H_2-H_1}.$$

Note that for each K > 0 and $\epsilon > 0$ we can choose $n_0 = n_0(K, \epsilon)$ such that $n \ge n_0$ implies

$$(TK)^{1/\mu}v_n^{(\gamma(\mu)/\mu)-1} \le \frac{\epsilon}{3}.$$

Hence if $n \ge n_0$ then $r_n \mathbf{1}_{A_n(K)} \le (\epsilon v_n/3)$ holds. We can always assume $n \ge n_0$ in this proof. Since $\tilde{\theta} \in \mathcal{G}_1^n$ and $r_n \le (\epsilon v_n)/3$ on $A_n(K)$, we have $|u - v| \ge (\epsilon v_n)/3$. Therefore

$$|R_3^n(\tilde{\theta})| \lesssim (\epsilon v_n)^{H_1 + H_2 - 2} \sum_{i: \overline{I_i^{1,n}} \le T} \int_{(I_i^{1,n})_{\theta}} du \ u^{H_1 - H_2} \int_{\mathcal{I}_{\theta - \tilde{\theta}}^{2,n} ((I_i^{1,n})_{\theta})_{\tilde{\theta} - \theta}} dv \ v^{H_2 - H_1}$$

$$(5.9)$$

on $A_n(K)$. Since $H_1 \leq H_2$, we have

$$\int_{\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta}-\theta}} dv \ v^{H_2-H_1} \lesssim r_n. \tag{5.10}$$

Since the integral $\int_{(I_i^{1,n})_{\theta}} du \ u^{H_1-H_2}$ is summable (note that $H_1 - H_2 \in (-1/2, 1/2)$), we have

$$|R_3^n(\tilde{\theta})| \lesssim (\epsilon v_n)^{H_1 + H_2 - 2} r_n. \tag{5.11}$$

By (5.11) and (B3), we have

$$\left| \frac{R_3^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \mathbf{1}_{A_n(K)} \lesssim v_n^{\gamma(\mu) - \mu} \mathbf{1}_{A_n(K)}. \tag{5.12}$$

Since the right-hand-side of (5.12) does not depend on $\tilde{\theta}$, we obtain (5.8).

The limit (5.7). We define $\tilde{h}_i^n(\tilde{\theta}) \in L^2([0,T+2\delta];\mathbb{R}^2) \otimes L^2([0,T+2\delta];\mathbb{R}^2)$ by

$$\tilde{h}_i^n(\tilde{\theta})(\omega) = h^1((I_i^{1,n})_{\theta}(\omega)) \tilde{\otimes} h^2(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta}-\theta}(\omega)).$$

Let ϵ_1 and ϵ_2 be positive numbers. By the same conditioning technique as in the proof of Lemma 5.2, we have

$$\mathbb{P}\left\{\sup_{\tilde{\theta}\in\mathcal{G}_{1}^{n}}\left|\frac{R_{4}^{n}(\tilde{\theta})}{D^{n}(\tilde{\theta})}\right| \geq \epsilon_{1}\right\} \leq \epsilon_{2} + \epsilon_{1}^{-2p}\sum_{\tilde{\theta}\in\mathcal{G}_{1}^{n}}\mathbb{E}\left\{\frac{\mathbf{1}_{A_{n}(K(\epsilon_{2}))}\mathbb{E}\left\{\left|\mathbb{I}_{2}\left(\sum_{i:\overline{I_{i}^{1,n}}\leq T}\tilde{h}_{i}^{n}(\tilde{\theta})\right)\right|^{2p}\right|\sigma(\mathcal{T})\right\}}{|D^{n}(\tilde{\theta})|^{2p}}\right\} \tag{5.13}$$

for any $p \ge 1$. Using Theorem 2.8, we obtain

$$\mathbb{E}\left\{\left|\mathbb{I}_{2}\left(\sum_{i:\overline{I_{i}^{1,n}}\leq T}\tilde{h}_{i}^{n}(\tilde{\theta})\right)\right|^{2p}\middle|\sigma(\mathcal{T})\right\}\leq C_{p}\mathbb{E}\left\{\left|\mathbb{I}_{2}\left(\sum_{i:\overline{I_{i}^{1,n}}\leq T}\tilde{h}_{i}^{n}(\tilde{\theta})\right)\right|^{2}\middle|\sigma(\mathcal{T})\right\}^{p}$$
(5.14)

where C_p is a positive constant depending only on $p \geq 1$. Let $R_5^n(\tilde{\theta})$ denote the expectation in the right-hand-side of (5.14). A simple calculation using Proposition 2.6 yields

$$\begin{split} R_5^n(\tilde{\theta}) &= \sum_{i,k: \overline{I_i^{1,n}} \leq T, \ \overline{I_k^{2,n}} \leq T} \langle h^1((I_i^{1,n})_{\theta}), h^1((I_k^{1,n})_{\theta}) \rangle_{L^2([0,T+2\delta];\mathbb{R}^2)} \\ &\times \langle h^2(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta}-\theta}, h^2(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_k^{1,n})_{\theta})_{\tilde{\theta}-\theta} \rangle_{L^2([0,T+2\delta];\mathbb{R}^2)} \\ &+ \sum_{i,k: \overline{I_i^{1,n}} \leq T, \ \overline{I_k^{2,n}} \leq T} \langle h^1((I_i^{1,n})_{\theta}), h^2(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_k^{1,n})_{\theta})_{\tilde{\theta}-\theta} \rangle_{L^2([0,T+2\delta];\mathbb{R}^2)} \\ &\times \langle h^1((I_k^{1,n})_{\theta}), h^2(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_i^{1,n})_{\theta})_{\tilde{\theta}-\theta} \rangle_{L^2([0,T+2\delta];\mathbb{R}^2)} \\ &=: R_{5,1}^n(\tilde{\theta}) + R_{5,2}^n(\tilde{\theta}). \end{split}$$

(1) The first term can be estimated as follows. Since

$$\langle h^2(\mathcal{I}^{2,n}_{\theta-\tilde{\theta}}((I_i^{1,n})_\theta)_{\tilde{\theta}-\theta}, h^2(\mathcal{I}^{2,n}_{\theta-\tilde{\theta}}((I_k^{1,n})_\theta)_{\tilde{\theta}-\theta}\rangle_{L^2([0,T+2\delta];\mathbb{R}^2)} \lesssim r_n^{2H_2},$$

we have

$$|R_{5,1}^n(\tilde{\theta})| \lesssim r_n^{2H_2}$$
 (5.15)

on $A_n(K(\epsilon_2))$. Note that $I \mapsto h^1(I)$ is linear so that

$$\sum_{i,k: \overline{I_i^{1,n}} \leq T, \ \overline{I_k^{1,n}} \leq T} \langle h^1((I_i^{1,n})_\theta), h^1((I_k^{1,n})_\theta) \rangle_{L^2([0,T+2\delta];\mathbb{R}^2)} < \infty.$$

(2) The second term can be estimated as follows. We first recall that

$$|\langle h^{1}((I_{i}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_{k}^{1,n})_{\theta})_{\tilde{\theta}-\theta}\rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})}|$$

$$= |\rho|c_{H_{1}}c_{H_{2}}\int_{(I_{i}^{1,n})_{\theta}} du \int_{\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_{k}^{1,n})_{\theta})_{\tilde{\theta}-\theta}} dv \beta(u,v)$$

$$\times |u-v|^{H_{1}+H_{2}-2}u^{H_{1}-H_{2}}v^{H_{2}-H_{1}}.$$

By the same reasoning as (5.9), we have

$$\langle h^{1}((I_{i}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_{k}^{1,n})_{\theta})_{\tilde{\theta}-\theta}\rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})} \lesssim \left(\frac{\epsilon v_{n}}{3}\right)^{H_{1}+H_{2}-2} r_{n}|I_{i}^{1,n}|$$

if $i \leq k$, and

$$\langle h^{1}((I_{k}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_{i}^{1,n})_{\theta})_{\tilde{\theta}-\theta}\rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})} \lesssim \left(\frac{\epsilon v_{n}}{3}\right)^{H_{1}+H_{2}-2} r_{n}|I_{k}^{1,n}|$$

if i > k, on $A_n(K(\epsilon_2))$. Therefore, it holds that

$$|R_{5,2}^{n}(\tilde{\theta})| \lesssim r_{n} v_{n}^{H_{1}+H_{2}-2} \left(\sum_{i \leq k} |I_{i}^{1,n}| |\langle h^{1}((I_{k}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_{i}^{1,n})_{\theta})_{\tilde{\theta}-\theta} \rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})} | \right)$$

$$+ \sum_{k>i} |I_{k}^{1,n}| |\langle h^{1}((I_{i}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}}^{2,n}((I_{k}^{1,n})_{\theta})_{\tilde{\theta}-\theta} \rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})} | \right)$$

$$\lesssim r_{n}^{1+H_{2}} v_{n}^{H_{1}+H_{2}-2}$$

$$(5.16)$$

on $A_n(K(\epsilon_2))$. Plugging (5.15) and (5.16) into (5.13), we obtain

$$\mathbb{P}\left\{ \sup_{\tilde{\theta} \in \mathcal{G}_{1}^{n}} \left| \frac{R_{4}^{n}(\tilde{\theta})}{D^{n}(\tilde{\theta})} \right| \geq \epsilon_{1} \right\} \lesssim \epsilon_{2} + \epsilon_{1}^{-2p} \sum_{\tilde{\theta} \in \mathcal{G}_{1}^{n}} \mathbb{E}\left\{ \frac{\mathbf{1}_{A_{n}(K(\epsilon_{2}))} | r_{n}^{2H_{2}} + r_{n}^{1+H_{2}} v_{n}^{H_{1}+H_{2}-2} |^{p}}{|D^{n}(\tilde{\theta})|^{2p}} \right\} \\
\lesssim \epsilon_{2} + \epsilon_{1}^{-2p} (\#\mathcal{G}^{n}) v_{n}^{p(1-H_{1})}$$

by (B3). Thanks to (C2), we have $\lim_{n\to\infty} (\#\mathcal{G}^n) v_n^{p(1-H_1)} = 0$ if we choose sufficiently large $p \geq 1$. This gives (5.7).

Thanks to (5.6) and (5.7), we obtain (5.5) when $H_1 \leq H_2$. Now let us consider the case where $H_1 > H_2$. In the case where $H_1 > H_2$, we use an alternative expression of $R_2^n(\tilde{\theta})$:

$$R_{2}^{n}(\tilde{\theta}) = \sum_{i,j: I_{i}^{1,n} \leq T} B^{1}((I_{i}^{1,n})_{\theta}) B^{2}(I_{j}^{2,n}) \mathbf{1}_{(I_{i}^{1,n})_{\tilde{\theta}} \cap I_{j}^{2,n} \neq \emptyset}$$

$$= \sum_{j} B^{2}(I_{j}^{2,n}) B^{1}(\mathcal{I}_{\tilde{\theta}}^{1,n,\leq T}(I_{j}^{2,n})_{\theta-\tilde{\theta}}), \tag{5.17}$$

where the symbol $\mathcal{I}_{\hat{\theta}}^{1,n,\leq T}$ denotes the family of shifted intervals

$$\mathcal{I}_{\tilde{\theta}}^{1,n,\leq T} = \{ (I_i^{1,n})_{\tilde{\theta}} \mid \overline{I_i^{1,n}} \leq T \}.$$

Using the expression (5.17), we can prove (5.5) when $H_1 > H_2$ in the same manner as we did in the case where $H_1 \leq H_2$.

Let us show that $R_2^n(\tilde{\theta}_n)/D^n(\tilde{\theta}_n)$ is away from zero if $|\tilde{\theta}_n - \theta|$ decreases rapidly as $n \to \infty$.

Proposition 5.4. Suppose that $\tilde{\theta}_n$ satisfies $|\theta - \tilde{\theta}_n| \le \rho_n$ for all $n \in \mathbb{N}$. Then there exists $c'_* > 0$ such that

$$\mathbb{P}\left\{ \left| \frac{R_2^n(\tilde{\theta}_n)}{D^n(\tilde{\theta}_n)} \right| \ge c_*' \right\} \to 1 \tag{5.18}$$

as $n \to \infty$.

Proof. We formally extend B^1 and B^2 by setting $B_t^1 = B_t^2 = 0$ for t < 0. Let us first suppose that $\tilde{\theta}_n \in \Theta_{\geq 0}$. This is the case for sufficiently large n if $\theta \in (0, \delta)$. Then we decompose $R_2^n(\tilde{\theta}_n)$ into three terms:

$$R_{2}^{n}(\tilde{\theta}_{n}) = \sum_{i:\overline{I_{i}^{1,n}} \leq T} B^{1}((I_{i}^{1,n})_{\theta}) \left(B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})_{\tilde{\theta}_{n}-\theta}) - B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta}) \cap [0, T+2\delta]) \right)$$

$$+ \sum_{i:\overline{I_{i}^{1,n}} \leq T} \left(B^{1}((I_{i}^{1,n})_{\theta}) B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta}) \cap [0, T+2\delta]) \right)$$

$$- \langle h^{1}((I_{i}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta}) \cap [0, T+2\delta]) \rangle_{L^{2}([0, T+2\delta]; \mathbb{R}^{2})} \right)$$

$$+ \sum_{i:\overline{I_{i}^{1,n}} \leq T} \langle h^{1}((I_{i}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta}) \cap [0, T+2\delta]) \rangle_{L^{2}([0, T+2\delta]; \mathbb{R}^{2})}$$

$$=: R_{6}^{n}(\tilde{\theta}_{n}) + R_{7}^{n}(\tilde{\theta}_{n}) + R_{8}^{n}(\tilde{\theta}_{n}).$$

$$(5.19)$$

To simplify the notation, we set $\mathcal{I}^{2,n}_{\theta-\tilde{\theta}_n}((I_i^{1,n})_{\theta})^+ = \mathcal{I}^{2,n}_{\theta-\tilde{\theta}_n}((I_i^{1,n})_{\theta}) \cap [0,T+2\delta]$. Let us show

$$\left| \frac{R_6^n(\tilde{\theta}_n)}{D^n(\tilde{\theta}_n)} \right| \to^p 0, \tag{5.20}$$

$$\left| \frac{R_7^n(\tilde{\theta}_n)}{D^n(\tilde{\theta}_n)} \right| \to^p 0 \tag{5.21}$$

and that there exists $c_* > 0$ such that

$$\mathbb{P}\left\{ \left| \frac{R_8^n(\tilde{\theta}_n)}{D^n(\tilde{\theta}_n)} \right| \ge c_*' \right\} \to 1$$
(5.22)

as $n \to \infty$.

The limit (5.20). Thanks to (B3), it suffices to show

$$\left| \frac{R_6^n(\tilde{\theta}_n)}{D^n(\tilde{\theta}_n)} \right| \mathbf{1}_{A_n(K)} \to^p 0 \tag{5.23}$$

as $n \to \infty$. By the conditioning argument as in the proof of Lemma 5.2, we have

$$\mathbb{E}\left\{\frac{|R_{6}^{n}(\tilde{\theta_{n}})|\mathbf{1}_{A_{n}(K)}}{|D^{n}(\tilde{\theta}_{n})|}\right\} \leq \mathbb{E}\left\{\frac{\mathbf{1}_{A_{n}(K)}}{|D^{n}|}\sum_{i}\mathbb{E}\{|B^{1}((I_{i}^{1,n})_{\theta})|\right. \\
\left. \times |B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})_{\tilde{\theta}_{n}-\theta}) - B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})^{+})| \mid \sigma(\mathcal{T})\}\right\}.$$
(5.24)

Let p > 1 and q > 1 be conjugate indices: 1/p + 1/q = 1. Using (conditional) Hölder's inequality and Theorem 2.8, we have

$$\sum_{i} \mathbb{E}\{|B^{1}((I_{i}^{1,n})_{\theta})||B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})_{\tilde{\theta}_{n}-\theta}) - B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})^{+})| \mid \sigma(\mathcal{T})\} \\
\leq \left(\sum_{i} \mathbb{E}\{|B((I_{i}^{1,n})_{\theta})|^{p} \mid \sigma(\mathcal{T})\}\right)^{1/p} \left(\sum_{i} \mathbb{E}\{|B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})_{\tilde{\theta}_{n}-\theta}) - B^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})^{+})|^{q} \mid \sigma(\mathcal{T})\}\right)^{1/q} \\
\lesssim r_{n}^{H_{1}-1/p}((\#\mathcal{I}^{1,n})|\theta-\tilde{\theta}_{n}|^{qH_{2}})^{1/q}. \tag{5.25}$$

Plugging (5.25) into (5.24) with $p = \left(1 - \frac{H_2}{1-\varsigma}\right)^{-1}$ and $q = \left(\frac{H_2}{1-\varsigma}\right)^{-1}$, we obtain

$$\mathbb{E}\left\{\frac{|R_{6}^{n}(\tilde{\theta_{n}})|\mathbf{1}_{A_{n}(K)}}{|D^{n}(\tilde{\theta}_{n})|}\right\} \leq \mathbb{E}\left\{\frac{r_{n}^{H_{1}+H_{2}-1+\mu+\frac{H_{2}\varsigma}{1-\varsigma}-\mu}\left((\#\mathcal{I}^{1,n})|\theta-\tilde{\theta}_{n}|^{1-\varsigma}\right)^{\frac{H_{2}}{1-\varsigma}}}{|D^{n}(\tilde{\theta}_{n})|}\mathbf{1}_{A_{n}(K)}\right\}.$$
 (5.26)

Now we can choose sufficiently small $\mu > 0$ satisfying $\frac{H_{2\varsigma}}{1-\varsigma} - \mu > 0$. Applying (B3), (C3) and Jensen's inequality to (5.26), we complete the proof of (5.20).

The limit (5.21). This can be proved in the same way as (5.7). Note that in this case the term corresponding to $R_{5,2}^n(\tilde{\theta})$ is

$$R_{7,2}^{n}(\tilde{\theta}_{n}) := \sum_{i,k:\overline{I_{i}^{1,n}} \leq T, \ \overline{I_{k}^{2,n}} \leq T} \langle h^{1}((I_{i}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{k}^{1,n})_{\theta})^{+}) \rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})} \times \langle h^{1}((I_{k}^{1,n})_{\theta}), h^{2}(\mathcal{I}_{\theta-\tilde{\theta}_{n}}^{2,n}((I_{i}^{1,n})_{\theta})^{+}) \rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})}.$$

This is estimated as follows:

$$\begin{split} |R_{7,2}^n(\tilde{\theta}_n)| &\lesssim r_n^{H_1+H_2} \sum_{i,k: \overline{I_i^{1,n}} \leq T, \ \overline{I_k^{2,n}} \leq T} \langle h^1((I_i^{1,n})_\theta), h^2(\mathcal{I}_{\theta-\tilde{\theta}_n}^{2,n}((I_k^{1,n})_\theta)^+) \rangle_{L^2([0,T+2\delta];\mathbb{R}^2)} \\ &\lesssim r_n^{H_1+2H_2}(\#\mathcal{I}^{1,n}) \\ &= r_n^{2H_1-1+\mu} r_n^{2H_2-1+\mu} r_n^{2-H_1-2\mu}(\#\mathcal{I}^{1,n}). \end{split}$$

This, combined with (B4), completes the proof of (5.21).

The limit (5.22). Finally we analyze the term $R_8^n(\theta_n)$. Recall that

$$|R_8^n(\tilde{\theta}_n)| = |\rho| c_{H_1} c_{H_2} \sum_{i: \overline{I_i^n} \le T} \int_{(I_i^{1,n})_{\theta}} du \int_{\mathcal{I}_{\theta-\tilde{\theta}_n}^{2,n}((I_i^{1,n})_{\theta})^+} dv \ \beta(u,v)$$
$$\times |u-v|^{H_1+H_2-2} u^{H_1-H_2} v^{H_2-H_1}.$$

Since $(I_i^{1,n})_{\theta} \subset \mathcal{I}_{\theta-\tilde{\theta}_n}^{2,n}((I_i^{1,n})_{\theta})^+$ for all i with $\overline{I_i^{1,n}} \leq T$, it holds that

$$|R_8^n(\tilde{\theta}_n)| \ge |\rho| c_{H_1} c_{H_2} \sum_{i: \overline{I_i^{1,n}} \le T} \int_{(I_i^{1,n})_{\theta}} du \int_{(I_i^{1,n})_{\theta}} dv \ \beta(u,v)$$

$$\times |u-v|^{H_1+H_2-2} u^{H_1-H_2} v^{H_2-H_1}. \tag{5.27}$$

To obtain a lower bound for (5.27), we note the following fact: for any $\epsilon > 0$, it holds that

$$\sum_{i:\overline{I_{i}^{1,n}} \leq T} \int_{(I_{i}^{1,n})_{\theta}} du \int_{(I_{i}^{1,n})_{\theta}} dv \, \beta(u,v) |u-v|^{H_{1}+H_{2}-2} u^{H_{1}-H_{2}} v^{H_{2}-H_{1}}$$

$$\geq \sum_{i:\overline{I_{i}^{1,n}} \leq T,\underline{I_{i}^{1,n}} \geq \epsilon} \int_{(I_{i}^{1,n})_{\theta}} du \int_{(I_{i}^{1,n})_{\theta}} dv \, \beta(u,v) |u-v|^{H_{1}+H_{2}-2} u^{H_{1}-H_{2}} v^{H_{2}-H_{1}}$$

$$\geq \sum_{i:\overline{I_{i}^{1,n}} \leq T,\overline{I_{i}^{1,n}} \geq \epsilon} \int_{(I_{i}^{1,n})_{\theta}} du \int_{(I_{i}^{1,n})_{\theta}} dv \, |u-v|^{H_{1}+H_{2}-2}.$$
(5.28)

Combining (5.27) nad (5.28), we obtain

$$|R_8^n(\tilde{\theta}_n)| \gtrsim \sum_{i:\overline{I_i^{1,n}} \le T, I_i^{1,n} \ge \epsilon} |I_i^{1,n}|^{H_1 + H_2}$$
 (5.29)

for any $\epsilon > 0$. Therefore we obtain

$$\left| \frac{R_8^n(\tilde{\theta}_n)}{D^n(\tilde{\theta}_n)} \right| \gtrsim \frac{\sum_{i: \overline{I_i^{1,n}} \leq T, \underline{I_i^{1,n}} \geq \epsilon} |I_i^{1,n}|^{H_1 + H_2}}{\sqrt{\sum_i |I_i^{1,n}|^{2H_1}} \sqrt{\sum_j |I_j^{2,n}|^{2H_2}}}.$$

We complete the proof of (5.22) by (B2) (especially (3.2)).

As we noted above, if the true parameter value θ is positive, then $\tilde{\theta}_n$ is in $\Theta_{\geq 0}$ for sufficiently large n since $|\tilde{\theta}_n - \theta| \leq \rho_n$. Therefore we can conclude (5.18) by (5.20)-(5.22) if $\theta > 0$. However $\tilde{\theta}_n$ may be negative for any n when $\theta = 0$. In order to obtain (5.18) when $\theta = 0$, it suffices to show (5.20)-(5.22) hold for $\theta = 0$ and $\tilde{\theta}_n \in [-\rho_n, 0)$. This is completely analogous to the case where $\tilde{\theta}_n \in \Theta_{\geq 0}$, so that we only add a few remarks and omit the proof.

If $\theta = 0$ and $\theta_n \in [-\rho_n, 0)$, then (5.19) becomes

$$R_{2}^{n}(\tilde{\theta}_{n}) = \sum_{i,j:\overline{I_{j}^{2,n}} \leq T} B^{1}(I_{i}^{1,n})B^{2}(I_{j}^{2,n})\mathbf{1}_{(I_{i}^{1,n})_{\tilde{\theta}_{n}} \cap I_{j}^{2,n} \neq \emptyset}$$

$$= \sum_{j:\overline{I_{j}^{2,n}} \leq T} \left(B^{1}(\mathcal{I}_{\tilde{\theta}_{n}}^{1,n}(I_{j}^{2,n})_{-\tilde{\theta}_{n}}) - B^{1}(\mathcal{I}_{\tilde{\theta}_{n}}^{1,n}(I_{j}^{2,n})^{+})\right) B^{2}(I_{j}^{2,n})$$

$$+ \sum_{j:\overline{I_{j}^{2,n}} \leq T} \left(B^{1}(\mathcal{I}_{\tilde{\theta}_{n}}^{1,n}(I_{j}^{2,n})^{+})B^{2}(I_{j}^{2,n}) - \langle h^{1}(\mathcal{I}_{\tilde{\theta}_{n}}^{1,n}(I_{j}^{2,n})^{+}), h^{2}(I_{j}^{2,n}) \rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})}\right)$$

$$+ \sum_{j:\overline{I_{j}^{2,n}} \leq T} \langle h^{1}(\mathcal{I}_{\tilde{\theta}_{n}}^{1,n}(I_{j}^{2,n})^{+}), h^{2}(I_{j}^{2,n}) \rangle_{L^{2}([0,T+2\delta];\mathbb{R}^{2})}. \tag{5.30}$$

Here the symbol $\mathcal{I}_{\tilde{\theta}_n}^{1,n}(I_j^{2,n})^+$ denotes $\mathcal{I}_{\tilde{\theta}_n}^{1,n}(I_j^{2,n}) \cap [0,T+2\delta]$ as before. Let $\hat{R}_8^n(\tilde{\theta}_n)$ denote the last term in (5.30). Then the inequality corresponding to (5.29) is

$$|\hat{R}_8^n(\tilde{\theta}_n)| \gtrsim \sum_{j: \overline{I_j^{2,n}} \le T, I_j^{2,n} \ge \epsilon} |I_i^{2,n}|^{H_1 + H_2}.$$

In particular, we use (3.3) of (B2) in this case.

The following corollary is immediate from Propositions 5.3 and 5.4 and Lemma 5.2.

Corollary 5.5. (1) We have

$$\sup_{\tilde{\theta} \in \mathcal{G}^n \cap \{\tilde{\theta} \in \Theta | |\tilde{\theta} - \theta| \ge \epsilon v_n\}} |\mathcal{U}_n(\tilde{\theta})| \to^p 0$$

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

(2) Suppose that $\tilde{\theta}_n \in \mathcal{G}^n$ satisfies $|\theta - \tilde{\theta}_n| \leq \rho_n$ for all $n \in \mathbb{N}$. Then there exists $c_* > 0$ such that

$$\mathbb{P}\left\{|\mathcal{U}_n(\tilde{\theta}_n)| \ge c_*\right\} \to 1$$

as $n \to \infty$.

Now we are ready to prove Theorem 4.6 under the additional assumptions $A^1 = A^2 = 0$.

Proof of Theorem 4.6 when $A^1 = A^2 = 0$. Thanks to (C3), there is a sequence $(\tilde{\theta}_n)_{n \in \mathbb{N}} \subset \mathcal{G}^n$ satisfying $|\tilde{\theta}_n - \theta| \leq \rho_n$. By the definition of $\hat{\theta}_n$, we have $|\mathcal{U}_n(\tilde{\theta}_n)| \leq |\mathcal{U}_n(\hat{\theta}_n)|$. On the other hand, one obtains $|\mathcal{U}_n(\hat{\theta}_n)| \leq \sup_{\tilde{\theta} \in \mathcal{G}^n \cap \{\tilde{\theta} \in \Theta | |\tilde{\theta} - \theta| > \epsilon v_n\}} |\mathcal{U}_n(\tilde{\theta})|$ on the set $\{v_n^{-1} | \hat{\theta}_n - \theta| \geq \epsilon\}$. Therefore, we obtain

$$\mathbb{P}\{v_n^{-1}|\hat{\theta_n} - \theta| \ge \epsilon\} \le \mathbb{P}\left\{ |\mathcal{U}_n(\tilde{\theta}_n)| \le \sup_{\tilde{\theta} \in \mathcal{G}^n \cap \{\tilde{\theta} \in \Theta | |\tilde{\theta} - \theta| \ge \epsilon v_n\}} |\mathcal{U}_n(\tilde{\theta})| \right\}. \tag{5.31}$$

The right-hand-side of (5.31) tends to 0 as $n \to \infty$ by Corollary 5.5.

5.2. **Proof of Theorem 4.6.** Let $\mathcal{U}_n^0(\tilde{\theta})$ denote the right-hand-side of (5.1). Now we consider the general case: we no longer assume $A^1 = A^2 = 0$. We reduce the general case to the previous one. We establish a relation between $\mathcal{U}_n(\tilde{\theta})$ and $\mathcal{U}_n^0(\tilde{\theta})$ in Proposition 5.8 below. We begin with a lemma that relates the denominators of $\mathcal{U}_n(\tilde{\theta})$ and those of $\mathcal{U}_n^0(\tilde{\theta})$.

Lemma 5.6. Let J denote $[0, T + \delta]$ or [0, T]. We set $a_1 = \theta$ and $a_2 = 0$. Then we have

$$\frac{\sum_{i:\overline{I_i^n}\in J} (X^l(I_i^{l,n})^2 - \sigma_l^2 B^l((I_i^{l,n})_{a_l})^2)}{\sum_{i:\overline{I_i^n}\in J} |I_i^{l,n}|^{2H_l}} \to^p 0$$
(5.32)

as $n \to \infty$ for l = 1, 2.

Proof. By the Hölder continuity of B^l , we have, for any $\epsilon > 0$,

$$|X^{l}(I_{i}^{l,n})^{2} - \sigma_{l}^{2}B^{l}((I_{i}^{l,n})_{a_{l}})^{2}| = |A^{l}(I_{i}^{l,n})^{2} + 2\sigma_{l}A^{l}(I_{i}^{l,n})B^{l}((I_{i}^{l,n})_{a_{l}})|$$

$$\lesssim |I_{i}^{l,n}|^{2H_{l}} + |I_{i}^{l,n}|^{1+H_{l}-\epsilon}$$

$$\lesssim r_{n}|I_{i}^{l,n}| + r_{n}^{H_{l}-\epsilon}|I_{i}^{l,n}|.$$

Since $H_l - \epsilon = (2H_l - 1) + 1 - H_l - \epsilon$, we obtain (5.32) by (B3) if we choose $\epsilon > 0$ such that $1 - H_l - \epsilon > 0$.

We set

$$\begin{split} \check{R}_{2}^{n}(\tilde{\theta}) &= \mathbf{1}_{\Theta_{\geq 0}}(\tilde{\theta}) \sum_{i,j: \overline{I_{i}^{1,n}} \leq T} X^{1}(I_{i}^{1,n}) X^{2}(I_{j}^{2,n}) \mathbf{1}_{I_{i}^{1,n} \cap (I_{j}^{2,n})_{-\tilde{\theta}} \neq \emptyset} \\ &+ \mathbf{1}_{\Theta_{<0}}(\tilde{\theta}) \sum_{i,j: \overline{I_{i}^{2,n}} \leq T} X^{1}(I_{i}^{1,n}) X^{2}(I_{j}^{2,n}) \mathbf{1}_{(I_{i}^{1,n})_{\tilde{\theta}} \cap (I_{j}^{2,n}) \neq \emptyset}. \end{split}$$

The next lemma relates the numerators of $\mathcal{U}_n(\tilde{\theta})$ and those of $\mathcal{U}_n^0(\tilde{\theta})$.

Lemma 5.7. We have

$$\sup_{\tilde{\theta} \in \mathcal{G}^n} \left| \frac{\check{R}_2^n(\tilde{\theta}) - \sigma_1 \sigma_2 R_2^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \to^p 0 \tag{5.33}$$

as $n \to \infty$.

Proof. Note that

$$\begin{split} X^1(I_i^{1,n})X^2(I_j^{2,n}) &= A^1(I_i^{1,n})A^2(I_j^{2,n}) + \sigma_2 A^1(I_i^{1,n})B^2(I_j^{2,n}) \\ &+ \sigma_1 B^1((I_i^{1,n})_\theta)A^2(I_j^{2,n}) + \sigma_1 \sigma_2 B^1((I_i^{1,n})_\theta)B^2(I_j^{2,n}). \end{split}$$

Obviously (5.33) is equivalent to

$$\sup_{\tilde{\theta} \in \mathcal{G}^n \cap \Theta_{\geq 0}} \left| \frac{\check{R}_2^n(\tilde{\theta}) - \sigma_1 \sigma_2 R_2^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \to^p 0$$
(5.34)

and

$$\sup_{\tilde{\theta} \in \mathcal{G}^n \cap \Theta_{<0}} \left| \frac{\check{R}_2^n(\tilde{\theta}) - \sigma_1 \sigma_2 R_2^n(\tilde{\theta})}{D^n(\tilde{\theta})} \right| \to^p 0.$$
 (5.35)

We only prove (5.34). The proof of (5.35) is completely analogous. Since $\tilde{\theta} \in \Theta_{\geq 0}$, we have

$$\begin{split} |\check{R}_{2}^{n}(\tilde{\theta}) - R_{2}^{n}(\tilde{\theta})| &\leq \sum_{i: \overline{I_{i}^{1,n}} \leq T} |A^{1}(I_{i}^{1,n})A^{2}(\mathcal{I}_{-\tilde{\theta}}^{2,n}(I_{i}^{1,n})_{\tilde{\theta}})| \\ &+ \sum_{i: \overline{I_{i}^{1,n}} \leq T} |\sigma_{2}A^{1}(I_{i}^{1,n})B^{2}(\mathcal{I}_{-\tilde{\theta}}^{2,n}(I_{i}^{1,n})_{\tilde{\theta}})| \\ &+ \sum_{i: \overline{I_{i}^{1,n}} \leq T} |\sigma_{1}B^{1}(\mathcal{I}_{\theta}^{1,n,\leq T}((I_{j}^{2,n})_{\theta-\tilde{\theta}}))A^{2}(I_{j}^{2,n})| \\ &\lesssim r_{n} + r_{n}^{H_{1}} + r_{2}^{H_{2}} \\ &= r_{n}^{H_{1}+H_{2}-1+\mu}(r_{n}^{2-H_{1}-H_{2}-\mu} + r_{n}^{1-H_{2}-\mu} + r_{n}^{1-H_{1}-\mu}). \end{split}$$

We can choose $\mu > 0$ such that $1 - (H_1 \vee H_2) - \mu > 0$ so that we conclude (5.34) by (B3).

Let us set $S_n = \bigcap_{l=1,2} \{ \sum_{i:\overline{I_i^{l,n}} \leq T} X^l (I_i^{l,n})^2 \neq 0 \}$. Note that $\mathbb{P}\{S_n\} \to 1$ as $n \to \infty$ by Lemmas 5.2 and 5.6. The next proposition shows that we can reduce calculation of $\mathcal{U}_n(\tilde{\theta})$ to that of $\mathcal{U}_n^0(\tilde{\theta})$.

Proposition 5.8. It holds that

$$\mathcal{U}_n(\tilde{\theta}) = \left(\bar{o}_p(1)(\tilde{\theta}) + \frac{1 + o_p(1)}{1 + o_p(1)} \mathcal{U}_n^0(\tilde{\theta})\right) \mathbf{1}_{S_n}.$$
 (5.36)

Here $\bar{o}_p(1)(\tilde{\theta})$ denotes a quantity such that

$$\sup_{\tilde{\theta} \in \mathcal{G}^n} \bar{o}_p(1)(\tilde{\theta}) \to^p 0.$$

In particular,

(1) we have

$$\sup_{\tilde{\theta} \in \mathcal{G}^n \cap \{\tilde{\theta} \in \Theta \mid |\tilde{\theta} - \theta| \ge \epsilon v_n\}} |\mathcal{U}_n(\tilde{\theta})| \to^p 0$$

as $n \to \infty$ for any $\epsilon > 0$, and

(2) if $\tilde{\theta}_n \in \mathcal{G}^n$ satisfies $|\tilde{\theta}_n - \theta| \le v_n^{1+\alpha}$, then there exists $c_* > 0$ such that

$$\mathbb{P}\left\{|\mathcal{U}_n(\tilde{\theta}_n)| \ge c_*\right\} \to 1$$

as $n \to \infty$.

Proof. We have (5.36) by Lemmas 5.2, 5.6 and 5.7.

Now we are ready to prove Theorem 4.6.

Proof of Theorem 4.6. Thanks to Proposition 5.8, (4.4) follows by the same line of argument as in the proof of Theorem 4.6 when $A^1 = A^2 = 0$.

6. Examples

In this section, we give some examples that satisfy the assumptions of Theorem 4.6.

6.1. Processes.

6.1.1. The solution of a stochastic differential equation driven by fBM. Let $(\Omega, \mathcal{F}, \mathbb{G} = (\mathcal{G}_t)_{t \in [0, T+2\delta]}, \mathbb{P})$ be a stochastic basis which supports fBMs B^1 and B^2 satisfying (2.6). Here we assume that the σ -field \mathcal{F} is complete and the filtration \mathbb{G} satisfies the usual conditions. Let us consider a stochastic differential equation

$$X_t^l = X_0^l + \int_0^t b(s, X_s^l) \, ds + \sigma^l B_t^l \tag{6.1}$$

for $t \in [0, T + 2\delta]$ and l = 1, 2. We assume that X_0 is a real-valued random variable, $\sigma^l \in \mathbb{R} \setminus \{0\}$ and the function b satisfies the following conditions (D1) and (D2).

Assumption 6.1. The function $b: [0, T + 2\delta] \times \mathbb{R} \to \mathbb{R}$ satisfies that

(D1) for each positive integer N there exists a positive constant $L_N > 0$ such that

$$|b(t,x) - b(t,y)| \le L_N|x - y|$$

for all $x \in [-N, N]$, $y \in [-N, N]$ and $t \in [0, T + 2\delta]$, and

(D2) there exists a positive constant $L_0 > 0$ such that

$$|b(t,x)| \le L_0(1+|x|)$$

for all $x \in \mathbb{R}$ and $t \in [0, T + 2\delta]$.

Then there exists a unique solution X^l to the equation (6.1) which has 1/2-Hölder continuous sample paths almost surely. This is a special case of Theorem 2.1 of Nualart and Răsçanu (2002).

Let us check the processes $\tau_{-\theta}(X^1)$ and X^2 satisfy Assumption 3.1. Since X^l satisfies the equation (6.1), it suffices to verify (A3). We have

$$|b(s, X_s^l)| \le L_0(|X_s^l| + 1) \le L_0(|X_s^l - X_0^l| + |X_0^l| + 1) \le L_0(C(T + 2\delta)^{1/2} + |X_0^l| + 1)$$

for some Hölder constant C > 0. Hence it holds that

$$|A_t^l - A_s^l| \le L_0(C(T+2\delta)^{1/2} + |X_0^l| + 1)|t-s|$$

for $0 \le s \le t \le T + 2\delta$.

6.2. Sampling scheme.

6.2.1. Synchronous and evenly spaced observations. Let us check that the synchronous and evenly spaced observations satisfy Assumption 3.5. Let $i_*^{l,n}(a)$ denote

$$i_*^{l,n}(a) = \inf\{i \mid \overline{I_i^{l,n}} \ge a\}$$

for a > 0. Note that, using $i_*^{l,n}(a)$, we obtain the following relation:

$$\frac{\sum_{i:\overline{I_{i}^{l,n}} \leq T,\underline{I_{i}^{l,n}} \geq \epsilon} |I_{i}^{l,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}}} = \frac{\sum_{i=i_{*}^{l,n}(\tau)-1}^{i,n} |I_{i}^{l,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}}}$$

Proposition 6.2. Let us set

$$\mathcal{T}^{1,n} = \mathcal{T}^{2,n} = \left\{ \frac{i}{n} (T + \delta) \mid i = 0, \dots, n \right\}.$$

Then $\mathcal{T}^{1,n}$ and $\mathcal{T}^{2,n}$ satisfy Assumption 3.5 with $v_n = (1/n)^{1-\xi}$ for any $\xi \in \left(0, \frac{1-(H_1 \vee H_2)}{2-(H_1 \vee H_2)}\right)$.

Proof. First (B1) is clear. Note that $i_*^{l,n}(\epsilon) = \left\lceil \frac{\epsilon n}{T+\delta} \right\rceil$ in this case. Let us check (B2):

$$\begin{split} \frac{\sum_{i:\overline{I_{i}^{l,n}} \leq T,\underline{I_{i}^{l,n}} \geq \epsilon} |I_{i}^{l,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i}|I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j}|I_{j}^{2,n}|^{2H_{2}}}} &= \frac{\sum_{i=i_{*}^{l,n}(T)-1}^{i_{*}(T)-1} |I_{i}^{l,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i}|I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j}|I_{j}^{2,n}|^{2H_{2}}}} \\ &= \frac{1}{n} \left(\left\lceil \frac{Tn}{T+\delta} \right\rceil - 1 - \left\lceil \frac{\epsilon n}{T+\delta} \right\rceil \right) \\ &\to \frac{T-\epsilon}{T+\delta} \end{split}$$

as $n \to \infty$. Next we verify (B3):

$$\frac{r_n^{2H_l-1+\mu}}{\sum_{i:\overline{I_i^{l,n}}\leq T}|I_i^{l,n}|^{2H_l}} = \left(\frac{1}{n}\right)^{\mu}\left(\frac{1}{n}\left\lceil\frac{Tn}{T+\delta}\right\rceil\right)^{-1} = v_n^{\frac{\mu}{1-\xi}}\left(\frac{1}{n}\left\lceil\frac{Tn}{T+\delta}\right\rceil\right)^{-1}.$$

We can take $\gamma(\mu) = \mu/(1-\xi)$. Finally we verify (B4):

$$v_n^{2-(H_1\vee H_2)}\mathbb{E}\{\#\mathcal{I}^{1,n}\} = v_n^{2-(H_1\vee H_2)}\mathbb{E}\{\#\mathcal{I}^{2,n}\} = \left(\frac{1}{n}\right)^{(1-\xi)(2-(H_1\vee H_2))-1}.$$

If
$$\xi \in \left(0, \frac{1 - (H_1 \vee H_2)}{2 - (H_1 \vee H_2)}\right)$$
, then $(1 - \xi)(2 - (H_1 \vee H_2)) - 1 > 0$ so that (B4) holds.

6.2.2. Poisson Sampling. We check that the Poisson sampling scheme satisfies Assumption 3.5. Let $p_1>0$ and $p_2>0$ be positive numbers, and $N^{1,n}$ and $N^{2,n}$ be mutually independent Poisson processes with intensity p_1n and p_2n , respectively. We also assume $N^{1,n}$ and $N^{2,n}$ are independent of the underlying fractional Brownian motions B^1 and B^2 . We set $T_i^{l,n}=\inf\{t\geq 0: N_t^{l,n}=i\} \land (T+\delta), \ \tilde{T}_i^{l,n}=\inf\{t\geq 0: N_t^{l,n}=i\}, \ I_i^{l,n}=(T_{i-1}^{l,n},T_i^{l,n}], \ \tilde{I}_i^{l,n}=(\tilde{T}_{i-1}^{l,n},\tilde{T}_i^{l,n}] \ \text{and} \ \lambda^{l,n}(t)=\lceil p_lnt\rceil.$ Note that $i_*^{l,n}(t)=N_{t-}^{l,n}+1$ in this case.

Proposition 6.3. We set

$$\mathcal{T}^{1,n} = \left\{ T_i^{1,n} \mid i = 0, \dots, i_*^{1,n} (T + \delta) \right\}$$

and

$$\mathcal{T}^{2,n} = \left\{ T_j^{2,n} \mid j = 0, \dots, i_*^{2,n} (T + \delta) \right\}.$$

Then $\mathcal{T}^{1,n}$ and $\mathcal{T}^{2,n}$ satisfy Assumption 3.5 with $v_n = (1/n)^{1-\eta}$ for any $\eta \in \left(0, \frac{1-(H_1 \vee H_2)}{2-(H_1 \vee H_2)}\right)$.

In order to prove Proposition 6.3, we use the following lemmas.

Lemma 6.4. Let

$$\tilde{r}_n = \left(\max_{1 \le i \le \lambda^{1,n}(T+\delta) \lor i_*^{1,n}(T+\delta)} |\tilde{I}_i^{1,n}|\right) \lor \left(\max_{1 \le j \le \lambda^{2,n}(T+\delta) \lor i_*^{2,n}(T+\delta)} |\tilde{I}_j^{2,n}|\right).$$

Then for any $q \geq 1$, we have

$$E\{\tilde{r}_n^q\} = o(n^\alpha),$$

for any $0 < \alpha < q$.

Proof. See Lemma 8 of Hayashi and Yoshida (2008).

Lemma 6.5. Let $\epsilon \in (1/2,1)$. Then, for each T>0, it holds that

$$P\{|\lambda^{l,n}(T) - i_*^{l,n}(T)| \le n^{\epsilon}\} \to 1,$$

as $n \to \infty$ for l = 1, 2.

Proof. See Lemma 9 of Hayashi and Yoshida (2008).

Before starting the proof of Proposition 6.3, we introduce some notation: for $a \in (0, T + \delta]$ and $H \in (1/2, 1)$, we set

$$R_{9}^{l,n}(a) = \frac{\sum_{i=1}^{i_{i}^{l,n}}(a)^{-1} |I_{i}^{l,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}}},$$

$$R_{10}^{l,n,\pm}(a) = \frac{\sum_{i=1}^{\lambda^{l,n}}(a)^{\pm \lceil 2n^{2/3} \rceil} |\tilde{I}_{i}^{l,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i=1}^{\lambda^{l,n}}(T+\delta)^{\mp \lceil 2n^{2/3} \rceil} |\tilde{I}_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j=1}^{\lambda^{l,n}}(T+\delta)^{\mp \lceil 2n^{2/3} \rceil} |\tilde{I}_{j}^{2,n}|^{2H_{2}}},$$

$$R_{11}^{l,n,\pm}(H,a) = \sum_{i=1}^{\lambda^{l,n}} \left(|\tilde{I}_{i}^{l,n}|^{2H} - \frac{\Gamma(2H+1)}{(np_{l})^{2H}} \right),$$

and

$$R_{12}^{l,n,\pm}(H,a) = \frac{\Gamma(2H+1)(\lambda^{l,n}(a) \pm \lceil 2n^{2/3} \rceil)}{(np_l)^{2H}}.$$

Note that we have the following relation:

$$R_{10}^{l,n,\pm}(a) = \left(R_{11}^{l,n,\pm}\left(\frac{H_1 + H_2}{2}, a\right) + R_{12}^{l,n,\pm}\left(\frac{H_1 + H_2}{2}, a\right)\right) \times \left(R_{11}^{l,n,\mp}(H_1, T + \delta) + R_{12}^{l,n,\mp}(H_1, T + \delta)\right)^{-1/2} \times \left(R_{11}^{2,n,\mp}(H_2, T + \delta) + R_{12}^{2,n,\mp}(H_2, T + \delta)\right)^{-1/2}.$$
(6.2)

Proof of Proposition 6.3. By assumption, (B1) holds. Let us show that Poisson sampling scheme satisfies condition (B2).

Claim. Let $a \in (0, T + \delta]$. Then it holds that

$$R_9^n(a) \to^p \frac{a}{T+\delta} c_l$$
 (6.3)

as $n \to \infty$. Here c_l is a constant depending on l, H_1 , H_2 , p_1 and p_2 :

$$c_l = \frac{p_1^{H_1 - 1/2} p_2^{H_2 - 1/2} \Gamma(H_1 + H_2 + 1)}{p_l^{H_1 + H_2 - 1} \Gamma(2H_1 + 1) \Gamma(2H_2 + 1)}.$$

In particular, (B2) holds.

Let us show (6.3). First note that a straightforward calculation yields $\mathrm{E}\{|\tilde{I}_i^{l,n}|^{2H}\} = \Gamma(2H+1)/(np_l)^{2H}$ and $\mathrm{Var}\{|\tilde{I}_i^{l,n}|^{2H}\} = (\Gamma(4H+1) - \Gamma(2H+1)^2)/(np_l)^{4H}$ for $H \in (1/2,1)$. By Lemma 6.5, we can assume $|\lambda^{l,n}(a) - i_*^{l,n}(a)| \leq n^{\epsilon}$ and $|\lambda^{l,n}(T+\delta) - i_*^{l,n}(T+\delta)| \leq n^{\epsilon}$ for $\epsilon \in (1/2,1)$. Since $|\lambda^{l,n}(a) - i_*^{l,n}(a)| \leq n^{\epsilon}$, we have $\lambda^{l,n}(a) - \lceil 2n^{\epsilon} \rceil < i_*^{l,n}(a)$ and hence $I_i^{l,n} = \tilde{I}_i^{l,n}$ if $i \leq \lambda^{l,n}(a) - \lceil 2n^{\epsilon} \rceil$. Note that clearly $|I_i^{l,n}| \leq |\tilde{I}_i^{l,n}|$ holds for all i. Hence we have

$$R_{10}^{l,n,-}(a) \le R_9^{l,n}(a) \le R_{10}^{l,n,+}(a)$$

In order to prove (6.3), it suffices to show

$$R_{10}^{l,n,\pm}(a) \to^p \frac{a}{T+\delta} c_l \tag{6.4}$$

as $n \to \infty$. Thanks to (6.2), we can further reduce (6.4) to the following limits:

$$n^{2H-1}R_{12}^{l,n,\pm}(H,a) \to \frac{a\Gamma(2H+1)}{(T+\delta)p_l^{2H-1}}$$
 (6.5)

and

$$R_{11}^{l,n,\pm}(H,a) = o_p(n^{1-2H})$$
(6.6)

as $n \to \infty$ for each $H \in (1/2, 1)$. First (6.5) is straightforward. Next, Kolmogorov's inequality shows that

$$\mathbb{P}\left\{n^{2H-1}\left|R_{11}^{l,n,\pm}\right| \ge n^{-\xi}\right\} \le \frac{\lambda^{l,n}(a) \pm \lceil 2n^{2/3} \rceil}{n^{2(1-\xi)}p_t^{4H}} (\Gamma(4H+1) - \Gamma(2H+1)^2) \tag{6.7}$$

holds for $H \in (1/2, 1)$, and the right-hand-side of (6.7) goes to 0 as $n \to \infty$ if $\xi < 1/2$. Hence (6.6) holds. As a consequence of (6.3), we can conclude that (B2) holds: we have

$$\frac{\sum_{i:\overline{I_{i}^{l,n}} \leq T,\underline{I_{i}^{l,n}} \geq \epsilon} |I_{i}^{l,n}|^{H_{1}+H_{2}}}{\sqrt{\sum_{i} |I_{i}^{1,n}|^{2H_{1}}} \sqrt{\sum_{j} |I_{j}^{2,n}|^{2H_{2}}}} \to^{p} \frac{T-\epsilon}{T+\delta} c_{l}$$

as $n \to \infty$.

Next we show that (B3) holds. Note that for each $\eta > 0$, we can choose $\zeta > 0$ satisfying $\eta > \zeta/(1+\zeta)$. Claim. If $\zeta > 0$ satisfies $\eta > \zeta/(1+\zeta)$, then, for each $\mu > 0$ and $\gamma(\mu) := (1+\zeta)\mu > \mu$, it holds that

$$\frac{r_n^{2H_l-1+\mu}}{\sum_{i:\overline{I_i^{l,n}} \le T} |I_i^{l,n}|^{2H_l}} = o_p(v_n^{\gamma(\mu)}).$$

In particular, (B3) holds.

Indeed, let $\epsilon > 0$ be any positive number. Then we have

$$\mathbb{P}\left\{\frac{r_n^{2H_l-1+\mu}v_n^{-\gamma(\mu)}}{\sum_{i:\overline{I_i^{l,n}}\leq T}|I_i^{l,n}|^{2H_l}}\geq \epsilon\right\} \leq \mathbb{P}\left\{\frac{\tilde{r}_n^{2H_l-1+\mu}n^{(1-\eta)(1+\zeta)\mu}}{\sum_{i=1}^{\lambda^{l,n}(T)-\lceil 2n^{2/3}\rceil}|\tilde{I}_i^{l,n}|^{2H_l}}\geq \epsilon, \ |\lambda^{l,n}(T)-i_*^{l,n}(T)|\leq n^{2/3}\right\} \\
+\mathbb{P}\left\{|\lambda^{l,n}(T)-i_*^{l,n}(T)|>n^{2/3}\right\} \\
\leq \mathbb{P}\left\{\frac{\tilde{r}_n^{2H_l-1+\mu}n^{2H_l-1+(1-\eta)(1+\zeta)\mu}}{n^{2H_l-1}R_{11}^{l,n,\pm}(H_l,T)+n^{2H_l-1}R_{12}^{l,n,\pm}(H_l,T)}\geq \epsilon\right\}+o(1).$$

Hence it holds that

$$\mathbb{P}\left\{\frac{r_n^{2H_l-1+\mu}v_n^{-\gamma(\mu)}}{\sum_{i:\overline{I_i^{l,n}} \leq T} |I_i^{l,n}|^{2H_l}} \geq \epsilon\right\} \leq \mathbb{P}\left\{\tilde{r}_n^{2H_l-1+\mu}n^{2H_l-1+(1-\eta)(1+\zeta)\mu} \geq \frac{\epsilon T\Gamma(2H_l+1)}{2(T+\delta)p_l^{2H_l-1}}\right\} + o(1)$$

$$\leq \left(\frac{\epsilon T\Gamma(2H_l+1)}{2(T+\delta)p_l^{2H_l-1}}\right)^{\frac{-1}{2H_l-1+\mu}} \mathbb{E}\{\tilde{r}_n\}n^{\frac{2H_l-1+(1+\zeta)(1-\eta)\mu}{2H_l-1+\mu}} + o(1). \tag{6.8}$$

Since $(1+\zeta)(1-\eta)\mu < \mu$, Lemma 6.4 implies that the right-hand-side of (6.8) goes to zero as $n \to \infty$. Note that $\mathbb{E}\{\#\mathcal{I}^{l,n}\} = p_l n(T+\delta)$. We can verify (B4) easily: since $(1-\eta)(2-(H_1\vee H_2)) > 1$ when $\eta \in \left(0, \frac{1-(H_1\vee H_2)}{2-(H_1\vee H_2)}\right)$, the condition (B4) holds.

6.3. **Simulation study.** In this subsection, we present a numerical simulation. We use the R package YUIMA to generate the data, plot the paths, and calculate the estimated values.

We consider the following setting. We set $dX_t^1 = dB_t^1$ and $dX_t^2 = dB_t^2$ where B^1 and B^2 satisfy (2.6) with $H_1 = 0.6$, $H_2 = 0.7$. We vary the correlation parameter ρ as $\rho = 0.25, 0.50$ and 0.75. The time horizon T equals 1. We assume that the processes X^1 and X^2 are observed by independent Poisson sampling schemes whose intensities are both n. We also vary the intensity n as n = 100, 300 and 500. We set the true lead-lag parameter $\theta = 0.02$, so that the process X^1 has a lead of $\theta = 0.02$ over the process X^2 . To calculate the estimator, we use the grids $\mathcal{G}_1 = \{-1 + 10^{-3}k \mid k \in \mathbb{Z}_{\geq 0}\} \cap [-1,1]$ and $\mathcal{G}_2 = \{-1 + (3 \cdot 10^{-3})k \mid k \in \mathbb{Z}_{\geq 0}\} \cap [-1,1]$. Hence the grid \mathcal{G}_1 contains the true parameter $\theta = 0.02$, and the grid \mathcal{G}_2 does not.

First we fix the grid. Then we repeat 300 simulations and compute the value of the estimator $\hat{\theta}_n$ each time, letting ρ and n vary as we explained above. That is, we consider the following nine cases after we

fix the grid: $(\rho, n) = (0.25, 100), (0.25, 300), (0.25, 500), (0.50, 100), (0.50, 300), (0.50, 500), (0.75, 100), (0.75, 300) and (0.75, 500).$

Figure 1 (resp. 2) shows a plot of estimated values $\hat{\theta}_n$ with the grid \mathcal{G}_1 (resp. \mathcal{G}_2). The red dashed lines show the true parameter $\theta = 0.02$.

Figure 3 shows a simulated observation data. Figure 4 shows a plot of the contrast function $\mathcal{U}_n(\tilde{\theta})$ using the simulated data shown in Figure 3. We use the grid $\mathcal{G} = \mathcal{G}_1$. Again the red dashed line shows the true parameter $\theta = 0.02$. In this simulation, the estimated parameter value is $\hat{\theta}_n = 0.02$.

ACKNOWLEDGEMENT

The author would like to express deepest gratitude to Professor Nakahiro Yoshida for introducing him to this problem and for many valuable suggestions. He is also very grateful to two anonymous referees for their really helpful comments and suggestions. He could never have improved the results and the presentation of this paper without their helpful comments.

This work was supported by JST CREST and the Program for Leading Graduate Schools, MEXT, Japan.

Conflict of interest

The author has no conflicts of interest to declare.

References

- C. Bender, T. Sottinen, and E. Valkeila. Pricing by hedging and no-arbitrage beyond semimartingales. *Finance and Stochastics*, 12(4):441–468, 2008.
- C. Bender, T. Sottinen, and E. Valkeila. Fractional processes as models in stochastic finance. In *Advanced mathematical methods for finance*, pages 75–103. Springer, 2011.
- F. Biagini, Y. Hu, B. Øksendal, and T. Zhang. Stochastic Calculus for Fractional Brownian Motion and Applications. Springer Science & Business Media, 2008.
- P. Cheridito. Arbitrage in fractional brownian motion models. Finance and Stochastics, 7(4):533–553, 2003.
- N. J. Cutland, P. E. Kopp, and W. Willinger. Stock price returns and the joseph effect: a fractional version of the black-scholes model. In *Seminar on stochastic analysis, random fields and applications*, pages 327–351. Springer, 1995.
- M. T. Greene and B. D. Fielitz. Long-term dependence in common stock returns. *Journal of Financial Economics*, 4(3):339–349, 1977.
- P. Guasoni. No arbitrage under transaction costs, with fractional brownian motion and beyond. *Mathematical Finance*, 16(3):569–582, 2006.
- P. Hall, W. Härdle, T. Kleinow, and P. Schmidt. Semiparametric bootstrap approach to hypothesis tests and confidence intervals for the hurst coefficient. *Statistical inference for stochastic processes*, 3(3): 263–276, 2000.
- T. Hayashi and S. Kusuoka. Consistent estimation of covariation under nonsynchronicity. *Statistical Inference for Stochastic Processes*, 11(1):93–106, 2008.
- T. Hayashi and N. Yoshida. On covariance estimation of non-synchronously observed diffusion processes. Bernoulli, 11(2):359–379, 2005.
- T. Hayashi and N. Yoshida. Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Annals of the Institute of Statistical Mathematics*, 60(2):367–406, 2008.
- O. T. Henry. Long memory in stock returns: some international evidence. Applied Financial Economics, 12(10):725–729, 2002.
- M. Hoffmann, M. Rosenbaum, and N. Yoshida. Estimation of the lead-lag parameter from non-synchronous data. *Bernoulli*, 19(2):426–461, 2013.
- N. Huth and F. Abergel. High frequency lead/lag relationshipsempirical facts. *Journal of Empirical Finance*, 26:41–58, 2014.

- R. A. Jarrow, P. Protter, and H. Sayit. No arbitrage without semimartingales. *The annals of applied probability*, pages 596–616, 2009.
- Y. Koike. Limit theorems for the pre-averaged hayashi-yoshida estimator with random sampling. Stochastic processes and their applications, 124(8):2699–2753, 2014.
- Y. Koike. On the asymptotic structure of brownian motions with an infinitesimal lead-lag effect. arXiv preprint arXiv:1601.03614, 2016a.
- Y. Koike. Quadratic covariation estimation of an irregularly observed semimartingale with jumps and noise. Bernoulli, 22(3):1894–1936, 2016b.
- A. W. Lo. Long-term memory in stock market prices. *Econometrica: Journal of the Econometric Society*, pages 1279–1313, 1991.
- Y. Mishura. Stochastic Calculus for Fractional Brownian Motion and Related Processes, volume 1929. Springer Science & Business Media, 2008.
- I. Nourdin and G. Peccati. Normal approximations with Malliavin calculus: from Stein's method to universality, volume 192. Cambridge University Press, 2012.
- D. Nualart. The Malliavin Calculus and Related Topics, volume 1995. Springer, 2006.
- D. Nualart and A. Răsçanu. Differential equations driven by fractional brownian motion. *Collectanea Mathematica*, 53(1):55–81, 2002.
- C. Y. Robert and M. Rosenbaum. On the limiting spectral distribution of the covariance matrices of time-lagged processes. *Journal of multivariate analysis*, 101(10):2434–2451, 2010.
- L. C. G. Rogers. Arbitrage with fractional brownian motion. Mathematical Finance, 7(1):95–105, 1997.
- T. Sottinen and E. Valkeila. On arbitrage and replication in the fractional black-scholes pricing model. Statistics & Decisions/International mathematical Journal for stochastic methods and models, 21(2): 93–108, 2003.
- V. Teverovsky, M. S. Taqqu, and W. Willinger. A critical look at lo's modified r/s statistic. *Journal of statistical Planning and Inference*, 80(1):211–227, 1999.
- W. Willinger, M. S. Taqqu, and V. Teverovsky. Stock market prices and long-range dependence. *Finance and stochastics*, 3(1):1–13, 1999.

GRADUATE SCHOOL OF MATHEMATICAL SCIENCES, THE UNIVERSITY OF TOKYO, KOMABA, TOKYO, 153-8914, JAPAN. *E-mail address*: kchiba@ms.u-tokyo.ac.jp

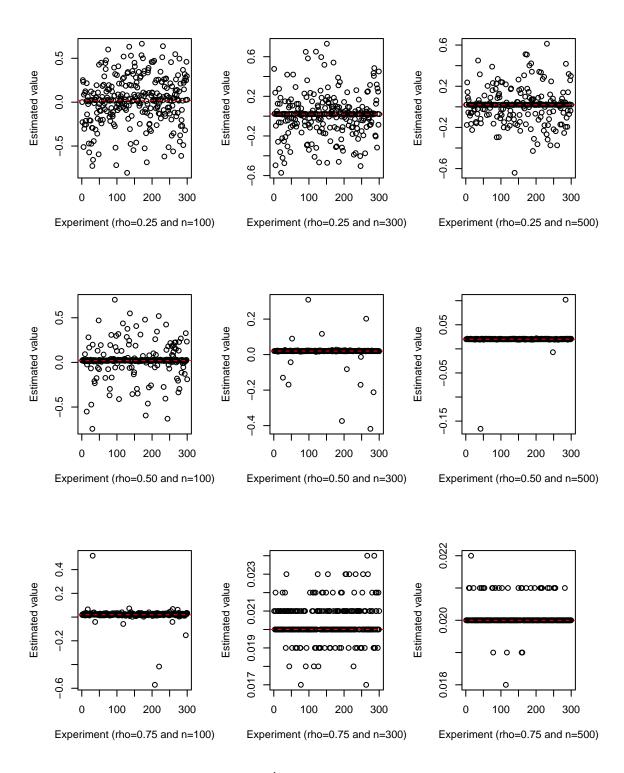


FIGURE 1. A plot of estimated values $\hat{\theta}_n$ (colored online). We fix a grid $\mathcal{G}_1 = \{-1+10^{-3}k \mid k \in \mathbb{Z}_{\geq 0}\} \cap [-1,1]$, repeat 300 simulations and plot the computed values of the estimator $\hat{\theta}_n$. The red dashed lines show the true parameter $\theta = 0.02$.

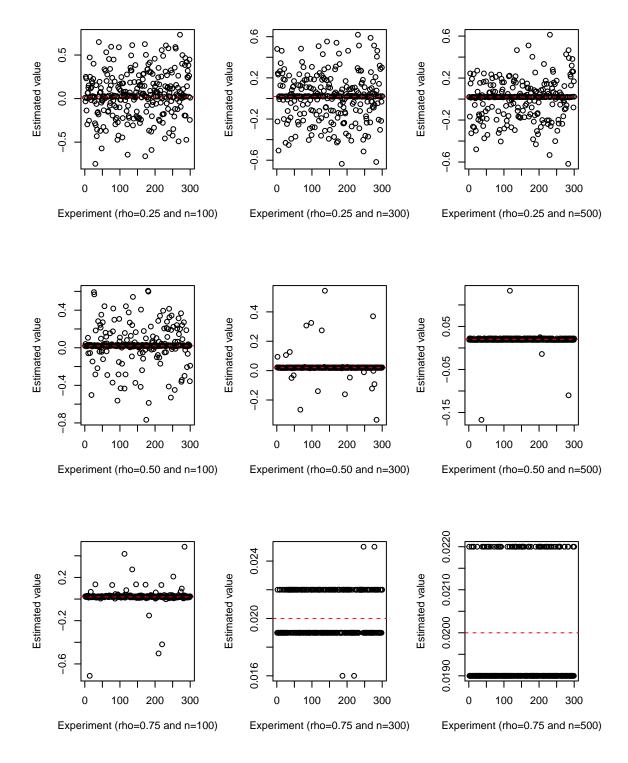


FIGURE 2. A plot of estimated values $\hat{\theta}_n$ (colored online). We fix a grid $\mathcal{G}_2 = \{-1 + (3 \cdot 10^{-3})k \mid k \in \mathbb{Z}_{\geq 0}\} \cap [-1,1]$, repeat 300 simulations and plot the computed values of the estimator $\hat{\theta}_n$. The red dashed lines show the true parameter $\theta = 0.02$. Note that the grid \mathcal{G}_2 does not contain the true parameter $\theta = 0.02$.

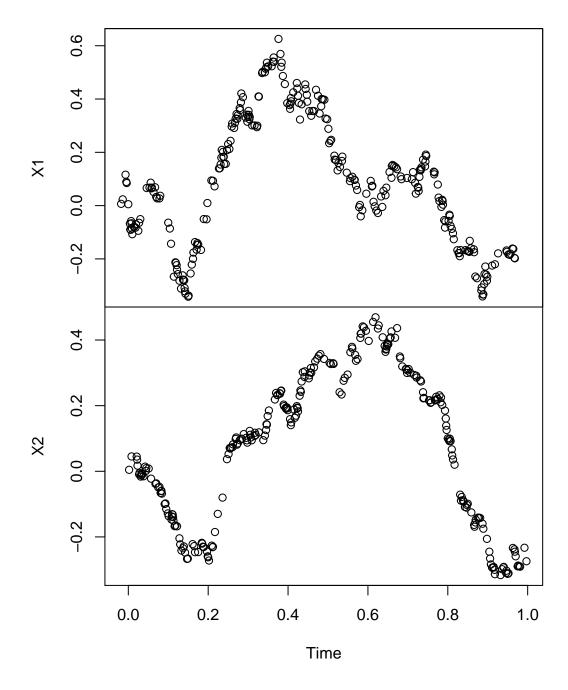


FIGURE 3. A simulated observation data. The underlying processes are $dX_t^1 = dB_t^1$ and $dX_t^2 = dB_t^2$ where B^1 and B^2 satisfy (2.6) with $H_1 = 0.6$, $H_2 = 0.7$ and $\rho = 0.50$. The time horizon is T = 1, and we use independent Poisson sampling schemes both with intensity n = 300. The lead-lag parameter is $\theta = 0.02$, so that the process X^1 has a lead of $\theta = 0.02$ over the process X^2 .

1 vs 2 (positive theta means 1 leads 2)

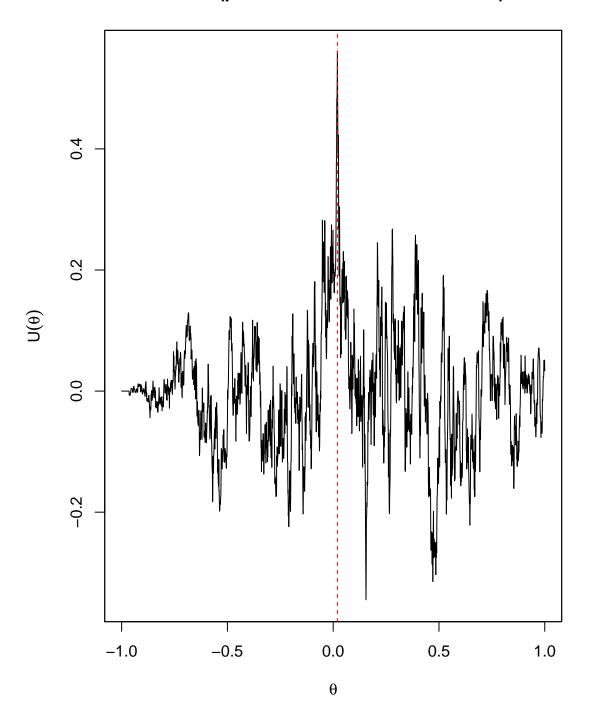


FIGURE 4. A plot of the contrast function $\mathcal{U}_n(\tilde{\theta})$ (colored online). We use the simulated data shown in Figure 3. In order to maximize the contrast, we use the grid $\mathcal{G}_1 = \{-1 + 10^{-3}k \mid k \in \mathbb{Z}_{\geq 0}\} \cap [-1,1]$. The red dashed vertical line shows $\tilde{\theta} = 0.02$. In this simulation, the estimated value is $\hat{\theta}_n = 0.02$.