Bernoulli **24**(4B), 2018, 3318–3383
https://doi.org/10.3150/17-BEJ962

Parametric inference for nonsynchronously observed diffusion processes in the presence of market microstructure noise

TEPPEI OGIHARA

The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190–8562, Japan.
E-mail: ogihara@ism.ac.jp
PRESTO, Japan Science and Technology Agency
School of Multidisciplinary Sciences, SOKENDAI (The Graduate University for Advanced Studies)

We study parametric inference for diffusion processes when observations occur nonsynchronously and are contaminated by market microstructure noise. We construct a quasi-likelihood function and study asymptotic mixed normality of maximum-likelihood- and Bayes-type estimators based on it. We also prove the local asymptotic normality of the model and asymptotic efficiency of our estimator when the diffusion coefficients are deterministic and noise follows a normal distribution. We conjecture that our estimator is asymptotically efficient even when the latent process is a general diffusion process. An estimator for the quadratic covariation of the latent process is also constructed. Some numerical examples show that this estimator performs better compared to existing estimators of the quadratic covariation.

Keywords: asymptotic efficiency; Bayes-type estimation; diffusion processes; local asymptotic normality; market microstructure noise; maximum-likelihood-type estimation; nonsynchronous observations; parametric estimation

Contents

1350-7265

1. Introduction

Analysis of volatility and covariation is one of the most important subjects in the study of risk management of financial assets. Studies of high-frequency financial data are increasingly significant as high-frequency financial data become increasingly available and computing technology develops. While realized volatility has been studied as a consistent estimator of integrated volatility at high-frequency limits, estimators of covariation of two securities are also important. The realized covariance, a natural extension of the realized volatility, is a consistent estimator of integrated covariation in ideal settings.

However, there are two significant problems in empirical analysis, one of which is the existence of observation noise. When we model stock price data by a continuous stochastic process, we should assume that the observations are contaminated by additional noise as a way to explain empirical evidence. Consistent estimators of volatility under the presence of microstructure noise are investigated – for example, in Zhang, Mykland, and Aït-Sahalia [31], Barndorff-Nielsen et al. [3], and Podolskij and Vetter [26]—by using various data-averaging or resampling methods to reduce the influence of noise. The other significant problem is that of nonsynchronous observation, namely, that we observe prices of different securities at different time points. The realized covariance has serious bias under models of nonsynchronous observations, though we can calculate the estimator by using some simple "*synchronization*" methods such as linear interpolation or the "previous tick" methods. Hayashi and Yoshida [15–17] and Malliavin and Mancino [22, 23] independently constructed consistent estimators for statistical models of diffusion processes with nonsynchronous observations. There are also studies of covariation estimation under the simultaneous presence of microstructure noise and nonsynchronous observations. We refer interested readers to Barndorff-Nielsen et al. [4] for a kernel based method; Christensen, Kinnebrock, and Podolskij [7], Christensen, Podolskij, and Vetter [8] for the modulated realised covariance and a pre-averaged Hayashi–Yoshida estimator; Aït-Sahalia, Fan, and Xiu [2] for a method using the maximum likelihood estimator of a model with constant diffusion coefficients; and Bibinger et al. [5] for a technique employing the local method of moments.

While the above studies concern estimators under non- or semi-parametric settings, there are also studies about parametric inference of diffusion processes with high-frequency observations. Genon-Catalot and Jacod [11] constructed a quasi-likelihood function and studied an estimator that maximizes it. Gloter and Jacod [13] studied an estimator based on a quasi-likelihood function with noisy observations. Ogihara and Yoshida [25] studied a maximum-likelihood-type estimator and a Bayes-type estimator on nonsynchronous observations without market microstructure noise.

One advantage of maximum-likelihood- and Bayes-type estimators is that they are asymptotically efficient in many models. If a statistical model has the local asymptotic mixed normality

(LAMN) property, then the results in Jeganathan [20,21] ensure that asymptotic variance of estimators cannot be smaller than a certain lower bound. When some estimator attains this bound, it is called asymptotically efficient. For parametric estimation of diffusion processes on fixed intervals, Gobet [14] proved the LAMN property of the statistical model having equidistant observations, and an estimator in [11] is asymptotically efficient. Ogihara [24] proved the LAMN property and asymptotic efficiency of estimators for the setting of [25]. Gloter and Jacod [12] proved the local asymptotic normality (LAN) property for a statistical model with market microstructure noise when diffusion coefficients are deterministic, and the estimator by Gloter and Jacod [13] is asymptotically efficient. There are few studies about the efficiency of estimators that assume the presence of market microstructure noise and nonsynchronous observations. One exception is Bibinger et al. [5], who showed a lower bound of asymptotic variance of estimators in semi-parametric Cramér-Rao sense. We need the LAN or LAMN property of the statistical model to obtain asymptotic efficiency of a parametric model. To the best of our knowledge, this has not been studied for statistical models of noisy, nonsynchronous observations.

This paper examines consistency and asymptotic mixed normality of a maximum-likelihood-type estimator and a Bayes-type estimator based on a quasi-likelihood function, under the simultaneous presence of market microstructure noise and nonsynchronous observations. We also study the LAN property of this model when diffusion coefficients are deterministic, as well as the asymptotic efficiency of our estimators. We expect that our estimators are asymptotically efficient in the general cases. However, it is further difficult to obtain LAMN properties for models of general diffusions. This does not seem to have been obtained even for noisy, equidistant observations, and is left as future work. We will see by simulation that sample variance of the estimation error of our estimator is better than that of existing estimators for some examples in Section 3. These results ensure that our estimator not only is the theoretical best for asymptotic behavior, but also works well in practical finite samplings.

Our study has several advantages in addition to the above arguments regarding asymptotic efficiency.

(i) Our model also allows observation noise that follows a non-Gaussian distribution. We use a quasi-likelihood function for Gaussian noise, but our method is robust enough to allow misspecification of the noise distribution.

(ii) Since we obtain the results regarding asymptotic behaviors of the quasi-likelihood function as a byproduct, many applications become available from the theory of maximum-likelihood-type estimation. For example, we can construct a theory of the likelihood ratio test and one-step estimators as an immediate application. Further, the theory of information criteria is expected to follow from our results of quasi-likelihood functions.

(iii) Our settings contain random sampling schemes where the maximum length of observation intervals is not bounded by any constant multiplication of the minimum length. This is the case for some significant random sampling schemes, such as samplings based on Poisson or Cox processes. Our model encompasses such natural sampling schemes.

To obtain asymptotic mixed normality of our estimator, we investigate asymptotic behaviors of a quasi-likelihood function of noisy, nonsynchronous observations. To this end, we need to specify the limit of some matrix trace related to a ratio of covariance matrices for two different values of parameters, as appearing in (5.2). The inverse of the covariance matrix of observation noise has nontrivial off-diagonal elements, and so the inverse of the covariance matrix of observations is far from a diagonal matrix. This phenomenon is essentially different from the case of *synchronous* observations without noise (where the covariance matrix of observations is diagonal), and the case of *nonsynchronous* observations without noise (where the inverse of the covariance matrix is not a diagonal matrix but is "*close*" to being one).

In a model of noisy, *synchronous* observations, the covariance matrix of a latent process is asymptotically equivalent to a unit matrix of the appropriate size, and is therefore simultaneously diagonalizable with the noise covariance. Gloter and Jacod [12,13] used these facts and closed expressions for the eigenvalues of the noise covariance to identify the limit of the quasi-likelihood function, but we cannot apply their idea because our sampling scheme is irregular and so not well approximated by a unit matrix. Further, the sizes of the covariance matrices are different for different components of the process, which follows from nonsynchronousness. In this paper, we deduce an asymptotically equivalent transform of the trace of the ratio of covariance matrices. This transform changes sizes of matrices and matrix elements into local averages, and arises from specific properties of the noise covariance matrix. We will see these results in Sections 4 and 5.

The remainder of this paper is organized as follows. In Section 2, we describe our detailed settings and main results. We propose a quasi-likelihood function for models with noisy, nonsynchronous observations, and construct a maximum-likelihood-type estimator based on it. We introduce asymptotic mixed normality of our estimator and results about asymptotic efficiency in Section 2.2. Section 2.3 contains results about the LAN property of our model and the asymptotic efficiency of our estimator, and Section 2.4 is devoted to results about Bayes-type estimators and convergence of moments of estimators. Polynomial-type large deviation inequalities, introduced in Yoshida [29,30], are key to deducing these results. In Section 3 we will examine simulation results of our estimator for a simple example where the latent process is a Wiener process. We also construct an estimator of the quadratic covariation and compare the performance of our estimator with that of other estimators. The remaining sections are devoted to a proof of the main results. Section 4 introduces an asymptotically equivalent expression of the quasi-likelihood function. This expression is useful for deducing asymptotic properties of the quasi-likelihood function in Section 5. We also need some results on identifiability of the model to obtain consistency of the maximum-likelihood-type estimator. These are discussed in Section 6. Section 7 shows asymptotic mixed normality of our estimator. The LAN property of the model for deterministic diffusion coefficients is obtained in Section 8. Section 9 contains a proof of results regarding the Bayes-type estimator and the convergence of moments of estimators.

2. Main results

2.1. Settings and construction of the estimator

Let $(\Omega^{(0)}, \mathcal{F}^{(0)}, P^{(0)})$ be a probability space with a filtration $\mathbf{F}^{(0)} = \{\mathcal{F}_t^{(0)}\}_{0\leq t\leq T}$. We consider a two-dimensional $\mathbf{F}^{(0)}$-adapted process $Y = \{Y_t\}_{0\leq t\leq T}$ satisfying the stochastic integral equation:

$$Y_t = Y_0 + \int_0^t \mu_s\, ds + \int_0^t b(s, X_s, \sigma_*)\, dW_s, \qquad t \in [0, T], \tag{2.1}$$

where $\{W_t\}_{0\leq t\leq T}$ is a d_1-dimensional standard $\mathbf{F}^{(0)}$-Wiener process, $b=(b^{ij})_{1\leq i\leq 2,1\leq j\leq d_1}$ is a Borel function, $\mu=\{\mu_t\}_{0\leq t\leq T}$ is a locally bounded $\mathbf{F}^{(0)}$-adapted process with values in $\mathbb{R}^2$, and $X=\{X_t\}_{0\leq t\leq T}$ is a continuous $\mathbf{F}^{(0)}$-adapted processes with values in O, an open subset of $\mathbb{R}^{d_2}$ with $d_2\in\mathbb{N}$. We consider market microstructure noise $\{\varepsilon_i^{n,k}\}_{n\in\mathbb{N},i\in\mathbb{Z}_+,k=1,2}$ as an independent sequence of random variables on another probability space $(\Omega^{(1)},\mathcal{F}^{(1)},P^{(1)})$. We assume that $\mathcal{F}^{(1)}=\mathfrak{B}((\varepsilon_i^{n,k})_{n,k,i})$ and that the distribution of $\varepsilon_j^{n,k}$ does not depend on j, where $\mathfrak{B}(S)$ denotes the minimal σ-field such that any element of S is $\mathfrak{B}(S)$-measurable for a set S of random variables. We use the same notation $\mathfrak{B}(S)$ for a similarly defined σ-field for a set S of measurable sets. We consider a product probability space $(\Omega,\mathcal{F},P)$, where $\Omega=\Omega^{(0)}\times\Omega^{(1)}$, $\mathcal{F}=\mathcal{F}^{(0)}\otimes\mathcal{F}^{(1)}$, and $P=P^{(0)}\otimes P^{(1)}$.

We assume that the observations of processes occur in a nonsynchronous manner and are contaminated by market microstructure noise, that is, we observe the vectors $\{\tilde{Y}_i^k\}_{0\leq i\leq \mathbf{J}_{k,n},k=1,2}$ and $\{\tilde{X}_j^k\}_{0\leq j\leq \mathbf{J}'_{k,n},1\leq k\leq d_2}$, where $\{S_i^{n,k}\}_{i=0}^{\mathbf{J}_{k,n}}$ and $\{T_j^{n,k}\}_{j=0}^{\mathbf{J}'_{k,n}}$ are random times in $(\Omega^{(0)},\mathcal{F}^{(0)})$, $\{\eta_j^{n,k}\}_{j\in\mathbb{Z}_+,1\leq k\leq d_2}$ is a random sequence on $(\Omega,\mathcal{F})$, and

$$\tilde{Y}_i^k=Y^k_{S_i^{n,k}}+\varepsilon_i^{n,k},\qquad \tilde{X}_j^k=X^k_{T_j^{n,k}}+\eta_j^{n,k}. \tag{2.2}$$

Our goal is to estimate the true value σ_* of the parameter from nonsynchronous, noisy observations $\{S_i^{n,k}\}_{0\leq i\leq \mathbf{J}_{k,n},k=1,2}$, $\{T_j^{n,k}\}_{0\leq j\leq \mathbf{J}'_{k,n},1\leq k\leq d_2}$, $\{\tilde{Y}_i^k\}_{0\leq i\leq \mathbf{J}_{k,n},k=1,2}$, and $\{\tilde{X}_j^k\}_{0\leq j\leq \mathbf{J}'_{k,n},1\leq k\leq d_2}$.

By setting $d_2=2$, $X_t\equiv Y_t$, $\mu_t=\mu(t,Y_t)$, $S_i^{n,k}\equiv T_j^{n,k}$, and $\eta_j^{n,k}\equiv\varepsilon_i^{n,k}$, our model contains the case where the latent process Y is a diffusion process satisfying a stochastic differential equation

$$dY_t=\mu(t,Y_t)\,dt+b(t,Y_t,\sigma_*)\,dW_t,\qquad t\in[0,T], \tag{2.3}$$

and Y is observed in a nonsynchronous manner with noise. This model is of particular interest, but our results are also be applied to more general models (2.1).

Remark 2.1. Stochastic volatility models are significant models for modeling stock prices. Unfortunately, our settings are not applied to hidden Markov models including stochastic volatility models because we require (possibly noisy) observations of process X. However, we hope that our results give an essential idea to deal with noisy, nonsynchronous observations, and therefore we can construct an estimator for stochastic volatility models by replacing our quasi-likelihood function. We have left it for future works.

For a vector $x=(x_1,\dots,x_k)$, we denote $\partial_x^l=(\frac{\partial^l}{\partial_{x_{i_1}}\cdots\partial_{x_{i_l}}})_{i_1,\dots,i_l=1}^k$. We assume the true value σ_* of the parameter is contained in a bounded open set $\Lambda\subset\mathbb{R}^d$ that satisfies Sobolev's inequality; that is, for any $p>d$, there exists $C>0$ such that $\sup_{\sigma\in\Lambda}|u(\sigma)|\leq C\sum_{k=0,1}(\int_\Lambda|\partial_\sigma^k u(\sigma)|^p\,d\sigma)^{1/p}$ for any $u\in C^1(\Lambda)$. This is the case when Λ has a Lipshitz boundary. See Adams and Fournier [1] for more details.

Let $\Pi_n=(\{S_i^{n,k}\}_{k,i},\{T_j^{n,k}\}_{k,j})$ and $\{\mathcal{G}_t\}_{0\leq t\leq T}$ be a filtration of $(\Omega,\mathcal{F},P)$ given by

$$\mathcal{G}_t=\mathcal{F}_t^{(0)}\vee\mathfrak{B}(\{\Pi_n\}_n)\vee\mathfrak{B}\big(A\cap\{S_i^{n,k}\leq t\};A\in\mathfrak{B}(\varepsilon_i^{n,k}),k\in\{1,2\},i\in\mathbb{Z}_+,n\in\mathbb{N}\big),$$

where $\mathcal{H}_1 \vee \mathcal{H}_2$ denotes the minimal σ-field which contains σ-fields $\mathcal{H}_1$ and $\mathcal{H}_2$. Let $E_\Pi[\mathbf{X}] = E[\mathbf{X}|\{\Pi_n\}_n]$ for a random variable $\mathbf{X}$. We assume that $\mathcal{F}_T^{(0)}$ and $\mathfrak{B}(\{\Pi_n\}_n)$ are independent. Moreover, we assume that there exist positive constants $v_{1,*}$ and $v_{2,*}$ such that $\eta_j^{n,k} 1_{\{T_j^{n,k} \le t\}}$ is $\mathcal{G}_t$-measurable, $E[\varepsilon_i^{n,k}] = 0$, and $E[(\varepsilon_i^{n,k})^2] = v_{k,*}$ for any n, k, i, j, t, where 1_A is the indicator function for a set A, δ_{ij} is Kronecker's delta. We also assume that the distribution of Y_0 does not depend on σ_*, $v_{1,*}$, nor $v_{2,*}$.

Now we construct the quasi-likelihood function. We apply the idea of Gloter and Jacod [13] to our construction of a quasi-likelihood function; that is, we divide the whole observation interval $[0, T]$ into equidistant subdivisions and construct quasi-likelihood functions for each interval as follows. Let $\{b_n\}_{n\in\mathbb{N}}$ and $\{k_n\}_{n\in\mathbb{N}}$ be sequences of positive numbers satisfying $b_n \ge 1$, $k_n \le b_n$, $b_n \to \infty$, $k_n b_n^{-1/2-\varepsilon} \to \infty$, and $k_n b_n^{-2/3+\varepsilon} \to 0$ as $n \to \infty$ for some $\varepsilon > 0$. We will assume in Assumption (A2) a relation between b_n and our sampling scheme, which implies that b_n represents the order of observation frequency. Let $\ell_n = [b_n k_n^{-1}]$, $s_0 = 0$, $s_m = T\ell_n^{-1} m$, $b^k(t, x, \sigma) = (b^{kj}(t, x, \sigma))_{j=1}^{d_1}$, $K_0^k = -1$, and $K_m^k = \#\{i \in \mathbb{N}; S_i^{n,k} < s_m\}$ for $k \in \{1, 2\}$ and $1 \le m \le \ell_n$. Moreover, let $k_m^j = K_m^j - K_{m-1}^j - 1$, $\bar{k}_n = \max_{m,j} k_m^j$, $\underline{k}_n = \min_{m,j} k_m^j$,

$$I_{i,m}^k = \big[S^{n,k}_{i+K^k_{m-1}}, S^{n,k}_{i+1+K^k_{m-1}}\big), \qquad \tilde{Y}^k\big(I^k_{i,m}\big) = \tilde{Y}^k_{i+1+K^k_{m-1}} - \tilde{Y}^k_{i+K^k_{m-1}},$$

$$\hat{X}_m = \left(\#\big\{j; T_j^{n,k} \in [s_{m-1}, s_m)\big\}^{-1} \sum_{j; T_j^{n,k} \in [s_{m-1}, s_m)} \tilde{X}_j^k\right)_{1\le k\le d_2},$$

and $b^j_{m'}(\sigma) = b^j(s_{m'-1}, \hat{X}_{m'-1}, \sigma)$ for $1 \le m \le \ell_n$, $2 \le m' \le \ell_n$, $j \in \{1, 2\}$ and $1 \le i \le k_m^j$. Then, roughly speaking, we have the following approximations of conditional covariance of observations:

$$\begin{aligned} E\big[\tilde{Y}^k\big(I^k_{i,m}\big)\tilde{Y}^k\big(I^k_{i',m}\big)|\mathcal{G}_{s_{m-1}}\big] &\approx \big(\big|b^k_m\big|^2\big|I^k_{i,m}\big| + 2v_{k,*}\big)\delta_{ii'} - v_{k,*}1_{\{|i-i'|=1\}}, \\ E\big[\tilde{Y}^1\big(I^1_{i'',m}\big)\tilde{Y}^2\big(I^2_{i''',m}\big)|\mathcal{G}_{s_{m-1}}\big] &\approx b^1_m \cdot b^2_m\big|I^1_{i'',m} \cap I^2_{i''',m}\big| \end{aligned} \tag{2.4}$$

for any intervals $I^k_{i,m}, I^k_{i',m}, I^1_{i'',m}, I^2_{i''',m}$.

Let $\top$ denotes the transpose operator for matrices (and vectors), $M(l) = \{2\delta_{i_1 i_2} - 1_{\{|i_1-i_2|=1\}}\}_{i_1,i_2=1}^l$ for $l \in \mathbb{N}$, $M_{j,m} = M(k_m^j)$ for $1 \le j \le 2$. Based on the relation (2.4), we define a quasi-log-likelihood function $H_n(\sigma, v)$ by

$$H_n(\sigma, v) = -\frac{1}{2}\sum_{m=2}^{\ell_n} Z_m^\top S_m^{-1}(\sigma, v) Z_m - \frac{1}{2}\sum_{m=2}^{\ell_n} \log\det S_m(\sigma, v), \tag{2.5}$$

where $Z_m = ((\tilde{Y}^1(I^1_{i,m}))^\top_{1\le i\le k_m^1}, (\tilde{Y}^2(I^2_{i,m}))^\top_{1\le i\le k_m^2})^\top$ and

$$S_m(\sigma, v) = \begin{pmatrix} \{|b^1_m|^2|I^1_{i,m}|\delta_{ii'}\}_{ii'} & \{b^1_m \cdot b^2_m|I^1_{i,m} \cap I^2_{j,m}|\}_{ij} \\ \{b^1_m \cdot b^2_m|I^1_{i,m} \cap I^2_{j,m}|\}_{ji} & \{|b^2_m|^2|I^2_{j,m}|\delta_{jj'}\}_{jj'} \end{pmatrix} + \begin{pmatrix} v_1 M_{1,m} & 0 \\ 0 & v_2 M_{2,m} \end{pmatrix} \tag{2.6}$$

for $v = (v_1, v_2)$.

Remark 2.2. Though such a local Gaussian quasi-log-likelihood function seems valid only when observation noise $\varepsilon_i^{n,k}$ follows a Gaussian distribution, asymptotic properties of the maximum likelihood estimator are robust enough to allow non-Gaussian noise. Hence, we can use the same quasi-likelihood function for general noise.

Remark 2.3. We used subdivisions of $[0, T]$ for the construction of H_n because of technical issues related to deducing the limit of H_n. Since the diffusion coefficient b in S_m is fixed, matrix properties of $M_{j,m}$ introduced in Section 4.2 can be used to deduce the limit of H_n. On the other hand, such a construction of H_n also contributes to reducing the calculation time of the maximum-likelihood-type estimator because the size of S_m is $O(k_n)$ while the size of the covariance matrix of all observations is $O(b_n)$.

Remark 2.4. In [13], k_n is taken so that $n^{1/2}k_n^{-1} \to 0$ and $k_n n^{-3/4} \to 0$. Our rate $b_n^{2/3}$ for the upper bound of k_n is a little bit worse because of some technical issue (for equidistance observations, we have $b_n \equiv n$). When we investigate asymptotic behaviors of the maximum-likelihood-type estimator, we deal with some supremum estimates for the σ of quasi-likelihood ratios. Unlike the one-dimensional settings of [13], our multidimensional setting requires some properties to deal with the supremum. We use Sobolev's inequality here for this purpose. Then we need an additional moment estimate for quasi-likelihood ratios, which causes a worse rate of k_n. See the proofs of Lemmas 4.3 and 4.4 for details.

To construct the maximum-likelihood-type estimator $\hat{\sigma}_n$ for the parameter σ, we need estimators for the unknown noise variance $v_* = (v_{1,*}, v_{2,*})$. We assume the following condition.

Assumption (V). There exist estimators $\{\hat{v}_n\}_{n\in\mathbb{N}}$ of v_* such that $\hat{v}_n \geq 0$ almost surely and $\{b_n^{1/2}(\hat{v}_n - v_*)\}_{n\in\mathbb{N}}$ is tight.

For example, $\hat{v}_n = (\hat{v}_{n,k})_{k=1}^2$ with $\hat{v}_{n,k} = (2\mathbf{J}_{k,n})^{-1}\sum_i(\tilde{Y}_i^k - \tilde{Y}_{i-1}^k)^2$ satisfies (V) if $\{b_n\mathbf{J}_{k,n}^{-1}\}_n$ is tight for $k = 1, 2$ and $\sup_{n,k,i} E[(\varepsilon_i^{n,k})^4] < \infty$.

Let $\mathrm{clos}(A)$ be the closure of a set A. A maximum-likelihood-type estimator $\hat{\sigma}_n$ is a random variable satisfying

$$H_n(\hat{\sigma}_n, \hat{v}_n) = \max_{\sigma\in\mathrm{clos}(\Lambda)} H_n(\sigma, \hat{v}_n).$$

We study asymptotic mixed normality and asymptotic efficiency of the estimator in the following subsections.

Remark 2.5. We can also construct a simultaneous maximum-likelihood-type estimator $(\bar{\sigma}_n, \bar{v}_n)$ satisfying $H_n(\bar{\sigma}_n, \bar{v}_n) = \max_{\sigma,v} H_n(\sigma, v)$. However, it is valid only when the observation noise $\varepsilon_i^{n,k}$ follows a normal distribution. Our interest is on estimating the parameter σ of the latent process, and so the assumptions for observation noise should be reduced as much as possible. Therefore, the nonparametric estimator $\hat{v}_n$ is more suitable for our purpose.

2.2. Asymptotic mixed normality of the maximum-likelihood-type estimator

In the rest of this section, we state our main theorems. Proofs of these results are left to Sections 4–9. In this subsection, we describe the asymptotic mixed normality of the maximum-likelihood-type estimator $\hat{\sigma}_n$.

We first describe assumptions for the theorem. Assumption (A1) is a sequence of assumptions on the latent processes Y and X and observation noise $\varepsilon_i^{n,k}$ and $\eta_j^{n,k}$. We denote by $\mathcal{E}_l$ the unit matrix of size l.

Assumption (A1). 1. For $0 \leq 2i + j \leq 4$ and $0 \leq k \leq 4$, the derivatives $\partial_t^i \partial_x^j \partial_\sigma^k b(t, x, \sigma)$ exist on $[0, T] \times O \times \Lambda$ and have continuous extensions on $[0, T] \times O \times \mathrm{clos}(\Lambda)$.

2. $bb^\top(t, x, \sigma)$ is positive definite for $(t, x, \sigma) \in [0, T] \times O \times \mathrm{clos}(\Lambda)$.

3. $\sup_{n,k,i} E[(\varepsilon_i^{n,k})^q] < \infty$ for any $q > 0$.

4. μ_t is optional and locally bounded (locally in time), that is, there exists an increasing sequence $\{T_l\}_l$ of stopping times such that $\lim_{l\to\infty} T_l = T$ a.s. and $\{\mu_{t\wedge T_l}\}_{0\leq t\leq T}$ is bounded for each l.

5. $b_n^{-\varepsilon} \max_{m,k}(E_\Pi[|\ell_n^{1/2}\mathbf{T}_{m,k}|^q \vee |\ell_n E[\mathbf{T}_{m,k}|\mathcal{G}_{s_{m-1}}]|^q]) \to^p 0$ as $n \to \infty$ for any $q > 0$ and $\varepsilon > 0$, where $\mathbf{T}_{m,k} = \#\{j; T_j^{n,k} \in [s_{m-1}, s_m)\}^{-1} \sum_{j;T_j^{n,k}\in[s_{m-1},s_m)} \eta_j^{n,k}$.

6. There exist progressively measurable processes $\{b_t^{(j)}\}_{0\leq t\leq T, 0\leq j\leq 1}$ and $\{\hat{b}_t^{(j)}\}_{0\leq t\leq T, 0\leq j\leq 1}$ such that $\sup_t E[|b_t^{(j)}|^q \vee |\hat{b}_t^{(j)}|^q] < \infty$, $\sup_{s<t} E[|b_t^{(j)} - b_s^{(j)}|^q \vee |\hat{b}_t^{(j)} - \hat{b}_s^{(j)}|^q]^{1/q}(t-s)^{-1/2} < \infty$ for $0 \leq j \leq 1$ and $q > 0$, and

$$X_t = X_0 + \int_0^t b_s^{(0)}\,ds + \int_0^t b_s^{(1)}\,dW_s, \qquad b_t^{(1)} = b_0^{(1)} + \int_0^t \hat{b}_s^{(0)}\,ds + \int_0^t \hat{b}_s^{(1)}\,dW_s$$

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

Assumption (A1) captures somewhat standard assumptions and whether it holds can easily verified in practical settings. Roughly speaking, point 5 of (A1) is satisfied if the summation of $\eta_j^{n,k}$ is of an order equivalent to the square root of the number of $\eta_j^{n,k}$. This is satisfied under certain independency, martingale conditions or mixing conditions of $\eta_j^{n,k}$. If $\{\eta_j^{n,k}\}_j$ is a sequence of centered, independent and identically distributed random variables and the sequence has finite moments, then $E_\Pi[|\mathbf{T}_{m,k}|^q] = O_p(\#\{j; T_j^{n,k} \in [s_{m-1}, s_m)\}^{-q/2})$. Then, point 5 of (A1) is satisfied if sampling frequency of $\{T_j^{n,k}\}$ is of order b_n and $E[\eta_j^{n,k} 1_{\{T_j^{n,k}>s_{m-1}\}}|\mathcal{G}_{s_{m-1}}] = 0$. Decomposition of X in point 6 of (A1) is used to deduce asymptotically equivalent representation of H_n where the diffusion coefficient $b(t, X_t, \sigma_*)$ is replaced by $b(s_{m-1}, X_{s_{m-1}}, \sigma_*)$. Detailed semimartingale decomposition is required to estimate the difference $b(t, X_t, \sigma_*) - b(s_{m-1}, X_{s_{m-1}}, \sigma_*)$.

In the following, we assume some conditions about our sampling scheme. For $\eta \in (0, 1/2)$, let $\mathcal{S}_\eta$ be the set of all sequences $\{[s'_{n,l}, s''_{n,l})\}_{n\in\mathbb{N}, 1\leq l\leq L_n}$ of intervals on $[0, T]$ satisfying $\{L_n\}_n \subset \mathbb{N}$, $[s'_{n,l_1}, s''_{n,l_2}) \cap [s'_{n,l_2}, s''_{n,l_2}) = \varnothing$ for $n, l_1 \neq l_2$, $\inf_{n,l}(b_n^{1-\eta}(s''_{n,l} - s'_{n,l})) > 0$, and $\sup_{n,l}(b_n^{1-\eta}(s''_{n,l} - s'_{n,l})) < \infty$. Let $r_n = \max_{i,k} |S_i^{n,k} - S_{i-1}^{n,k}|$ and $\underline{r}_n = \min_{i,k} |S_i^{n,k} - S_{i-1}^{n,k}|$.

Assumption (A2). There exist $\eta \in (0, 1/2)$, $\kappa > 0$, $\dot{\eta} \in (0, 1]$ and positive-valued stochastic processes $\{a_t^j\}_{t\in[0,T],j=1,2}$ such that $\sup_{t\neq s}(|a_t^j - a_s^j|/|t-s|^{\dot{\eta}}) < \infty$ almost surely, $b_n^{-1/2+\kappa} \times k_n(b_n^{-1}k_n)^{\dot{\eta}} \to 0$ and

$$k_n b_n^{-1/2+\kappa} \max_{1\leq l\leq L_n} \left|b_n^{-1}\left(s''_{n,l} - s'_{n,l}\right)^{-1}\#\left\{i; \left[S_{i-1}^{n,j}, S_i^{n,j}\right) \subset \left[s'_{n,l}, s''_{n,l}\right)\right\} - a^j_{s'_{n,l}}\right| \xrightarrow{p} 0 \tag{2.7}$$

as $n \to \infty$ for $j = 1, 2$ and $\{[s'_{n,l}, s''_{n,l})\}_{1\leq l\leq L_n, n\in\mathbb{N}} \in \mathcal{S}_\eta$. Moreover, $(r_n b_n^{1-\varepsilon}) \vee (b_n^{-1-\varepsilon}\underline{r}_n^{-1}) \to^p 0$ for any $\varepsilon > 0$.

In particular, Assumption (A2) implies $b_n^{-1}\mathbf{J}_{j,m} \to^p \int_0^T a_t^j\,dt$ and $\max_m |T^{-1}k_n^{-1}k_m^j - a^j_{s_{m-1}}| \to^p 0$ as $n \to \infty$. Roughly speaking, (A2) shows the law of large numbers for sampling schemes in any local time intervals. In the proof of Lemma 5.2, we will see that some properties of $M_{j,m}$ enable us to replace $|I_k^j|$ in $S_m(\sigma, v)$ by the local average in asymptotics. Then (A2) leads to the limit of H_n.

Example 2.1. Let $\{N_t^k\}_{t\geq 0}$ be an exponential α-mixing point process with stationary increments for $k = 1, 2$. Assume that $E[|N_1^k|^q] < \infty$ for any $q > 0$ and $k = 1, 2$. Set $S_i^{m,k} = \inf\{t \geq 0; N^k_{b_n t} \geq i\}$. Then Rosenthal-type inequalities (Theorem 3 and Lemma 7 in Doukhan and Louhichi [10], or Theorem 4 in [25]) and a similar argument to the proof of Proposition 6 in [25] ensure (A2) with $a_t^j \equiv E[N_1^j]$ (constants). Also, (B2) (defined later) is satisfied if further $k_n b_n^{-4/7+\gamma} \to 0$ for some $\gamma > 0$.

Under the above conditions, we can show convergence of the quasi-likelihood ratio $H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)$. Let $b_t = b(t, X_t, \sigma)$, $b_{t,*} = b(t, X_t, \sigma_*)$, $\tilde{a}_t^j = a_t^j / v_{j,*}$ for $j = 1, 2$ and

$$\begin{aligned}
\mathcal{Y}_1(\sigma) = \int_0^T \Bigg\{ & \frac{\sum_{j=1}^2 (|b_t^j|^2 - |b_{t,*}^j|^2)(|b_t^{3-j}|^2\sqrt{\tilde{a}_t^1\tilde{a}_t^2} + \tilde{a}_t^j\sqrt{\det(b_t b_t^\top)})}{4\sqrt{\det(b_t b_t^\top)}(\tilde{a}_t^1|b_t^1|^2 + \tilde{a}_t^2|b_t^2|^2 + 2\sqrt{\tilde{a}_t^1\tilde{a}_t^2\det(b_t b_t^\top)})^{1/2}} \\
& - \frac{(b_t^1\cdot b_t^2 - b_{t,*}^1\cdot b_{t,*}^2)b_t^1\cdot b_t^2\sqrt{\tilde{a}_t^1\tilde{a}_t^2}}{2\sqrt{\det(b_t b_t^\top)}(\tilde{a}_t^1|b_t^1|^2 + \tilde{a}_t^2|b_t^2|^2 + 2\sqrt{\tilde{a}_t^1\tilde{a}_t^2\det(b_t b_t^\top)})^{1/2}} \\
& - \frac{(\tilde{a}_t^1|b_t^1|^2 + \tilde{a}_t^2|b_t^2|^2 + 2\sqrt{\tilde{a}_t^1\tilde{a}_t^2\det(b_t b_t^\top)})^{1/2}}{2} \\
& + \frac{(\tilde{a}_t^1|b_{t,*}^1|^2 + \tilde{a}_t^2|b_{t,*}^2|^2 + 2\sqrt{\tilde{a}_t^1\tilde{a}_t^2\det(b_{t,*} b_{t,*}^\top)})^{1/2}}{2}\Bigg\}\,dt.
\end{aligned} \tag{2.8}$$

Proposition 2.1. *Assume (A1), (A2) and (V). Then* $\sup_{\sigma\in\Lambda}|b_n^{-1/2}\partial_\sigma^k(H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)) - \partial_\sigma^k\mathcal{Y}_1(\sigma)| \to^p 0$ *as* $n \to \infty$ *for* $0 \leq k \leq 3$.

To show consistency and asymptotic normality of $\hat{\sigma}_n$, the limit function $\mathcal{Y}_1(\sigma)$ of the quasi-likelihood ratio should have the unique maximum point at $\sigma = \sigma_*$. More precisely, we use the following as a kind of identifiability condition: $\inf_{\sigma \neq \sigma_*}(-\mathcal{Y}_1(\sigma))/|\sigma - \sigma_*|^2 > 0$ almost surely. Though it is difficult to directly check this condition in general, we can check it under a more tractable sufficient condition. Let

$$\mathcal{Y}_0(\sigma) = -\frac{1}{2}\int_0^T \left\{ \operatorname{tr}\bigl((b_t b_t^\top)^{-1}(b_{t,*}b_{t,*}^\top) - \mathcal{E}_2\bigr) + \log \frac{\det(b_t b_t^\top)}{\det(b_{t,*}b_{t,*}^\top)} \right\} dt.$$

Then $\mathcal{Y}_0$ is the probability limit $n^{-1/2}(H_n^0(\sigma) - H_n^0(\sigma_*))$, where H_n^0 represents a quasi-likelihood function for a statistical model of equidistant observations without noise. See Uchida and Yoshida [28].

Assumption (A3). $\inf_{\sigma \neq \sigma_*}((-\mathcal{Y}_0(\sigma))/|\sigma - \sigma_*|^2) > 0$ almost surely.

We will show in Proposition 6.1 that (A3) is sufficient for the identifiability condition of our model. Moreover, the following condition is a simple sufficient condition for (A3) (see Remark 4 in Ogihara and Yoshida [25] for the details).

Assumption (A3′). $\inf_{\sigma_1 \neq \sigma_2}(|bb^\top(t,x,\sigma_1) - bb^\top(t,x,\sigma_2)|/|\sigma_1 - \sigma_2|) > 0$ for any $t \in [0,T]$ and $x \in O$.

We denote by $\to^{s\text{-}\mathcal{L}}$ the stable convergence of random variables. Let

$$\hat{\Gamma}_{1,n} = -b_n^{-1/2}\partial_\sigma^2 H_n(\hat{\sigma}_n, \hat{v}_n), \qquad \Gamma_1 = -\partial_\sigma^2 \mathcal{Y}_1(\sigma_*). \tag{2.9}$$

Let $\mathcal{N}$ be a d-dimensional random variable on some extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ of $(\Omega, \mathcal{F}, P)$ satisfying the condition that $\mathcal{N}$ is independent of $\mathcal{F}$ and $\mathcal{N}$ follows the d-dimensional standard normal distribution. We denote the expectation with respect to $\tilde{P}$ by the same notation E.

The following theorem is one of our main results.

Theorem 2.1. *Assume (A1)–(A3) and (V). Then Γ_1 is positive definite almost surely and $b_n^{1/4}(\hat{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \Gamma_1^{-1/2}\mathcal{N}$ as $n \to \infty$. Moreover, $\hat{\Gamma}_{1,n} \to^p \Gamma_1$, and therefore $b_n^{1/4} \times \hat{\Gamma}_{1,n}^{1/2} 1_{\{\hat{\Gamma}_{1,n} \text{ is p.d.}\}}(\hat{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \mathcal{N}$ as $n \to \infty$.*

Corollary 2.1. *Assume (A1), (A2), (A3′) and (V). Then the results in Theorem 2.1 hold true.*

Let $\Upsilon_t^{(k)} = \partial_\sigma^k(\tilde{a}_t^1|b_t^1|^2 + \tilde{a}_t^2|b_t^2|^2 + 2\sqrt{\tilde{a}_t^1\tilde{a}_t^2\det(b_t b_t^\top)})^{1/2}|_{\sigma=\sigma_*}$, $B_{j,t}^{(k)} = \partial_\sigma^k|b_t^j|^2|_{\sigma=\sigma_*}$, $B_{3,t}^{(k)} = \partial_\sigma^k b_t^1 \cdot b_t^2|_{\sigma=\sigma_*}$, and $B_{4,t}^{(k)} = \partial_\sigma^k\sqrt{\det(b_t b_t^\top)}|_{\sigma=\sigma_*}$ for $k = 0, 1$ and $j = 1, 2$. Then we can rewrite

Γ_1 as

$$\Gamma_1 = \int_0^T \Bigg\{ \frac{\Upsilon_t^{(1)}(\Upsilon_t^{(1)})^\top}{2\Upsilon_t^{(0)}} + \frac{\sqrt{\tilde{a}_t^1 \tilde{a}_t^2}}{2B_{4,t}^{(0)}\Upsilon_t^{(0)}} \Bigg(B_{4,t}^{(1)}\big(B_{4,t}^{(1)}\big)^\top - \frac{B_{1,t}^{(1)}(B_{2,t}^{(1)})^\top + B_{2,t}^{(1)}(B_{1,t}^{(1)})^\top}{2} + B_{3,t}^{(1)}\big(B_{3,t}^{(1)}\big)^\top \Bigg) \Bigg\} dt. \tag{2.10}$$

The derivation of (2.10) is left to the appendix.

2.3. On the LAMN property and asymptotic efficiency of the estimator

In this subsection, we state some results on the so-called LAMN (LAN) property for our model and asymptotic efficiency of our estimator. We also comment on some further studies.

Throughout this subsection, we assume that $X_t \equiv Y_t$, $T_j^{n,k} \equiv S_i^{n,k}$, $\eta_j^{n,k} \equiv \varepsilon_i^{n,k}$, $\mu_t = \mu(t, \sigma_*)$ and $Y_0 = \gamma$ for some Borel function μ and some known $\gamma \in \mathbb{R}^2$. Then the latent process Y is a diffusion process satisfying the stochastic differential equation (2.3) with $\mu = \mu(t, \sigma_*)$. Let $P_{\sigma'_*, v'_*, n}$ be the distribution of $((S_i^{n,k})_{k,i}, (\tilde{Y}_i^k)_{k,i})$ with true values (σ'_*, v'_*) of the parameters. We denote

$$\operatorname{diag}(A, B) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

for square matrices A and B. Let $\mathcal{Y}_2(v) = -\int_0^T \sum_{j=1}^2 a_t^j \{(v_{j,*}/v_j) - 1 + \log(v_j/v_{j,*})\}\, dt/2$,

$$\Gamma_2 = -\partial_v^2 \mathcal{Y}_2(v_*) \quad \text{and} \quad \Gamma = \operatorname{diag}(\Gamma_1, \Gamma_2). \tag{2.11}$$

We adopt the following definition of the LAMN property from Jeganathan [21].

Definition 2.1. Let $P_{\theta,n}$ be a probability measure on some measurable space $(\mathcal{X}_n, \mathcal{A}_n)$ for each $\theta \in \mathbf{\Theta}$ and $n \in \mathbb{N}$, where Θ is a bounded open subset of $\mathbb{R}^d$. Then the family $\{P_{\theta,n}\}_{\theta,n}$ satisfies the local asymptotic mixed normality (LAMN) property at $\theta = \theta_*$ if there exist a sequence $\{\delta_n\}_{n\in\mathbb{N}}$ of $d \times d$ positive definite matrices, $d \times d$ symmetric random matrices Γ_n, Γ and d-dimensional random vectors $\mathcal{N}_n$, $\mathcal{N}$ such that Γ is positive definite a.s., $P_{\theta_*,n}[\Gamma_n$ is positive definite$] = 1$ ($n \in \mathbb{N}$), $\|\delta_n\| \to 0$, and

$$\log \frac{dP_{\theta_* + \delta_n u, n}}{dP_{\theta_*,n}} - \left(u^\top \sqrt{\Gamma_n} \mathcal{N}_n - \frac{1}{2} u^\top \Gamma_n u \right) \to 0$$

in $P_{\theta_*,n}$-probability as $n \to \infty$ for any $u \in \mathbb{R}^d$. Moreover, $\mathcal{N}$ follows the d-dimensional standard normal distribution, $\mathcal{N}$ is independent of Γ and $\mathcal{L}(\mathcal{N}_n, \Gamma_n | P_{\theta_*,n}) \to \mathcal{L}(\mathcal{N}, \Gamma)$ as $n \to \infty$.

If further the limit matrix Γ is non-random, we say $\{P_{\theta,n}\}_{\theta,n}$ has the local asymptotic normality (LAN) property.

To prove the LAMN property of our model, we assume the following additional condition.

Assumption (A1$''$). (A1) is satisfied, $\mu_t = \mu(t, \sigma_*)$, $\sup_{t,\sigma} \mu(t, \sigma) < \infty$, $b(t, x, \sigma)$ does not depend on x and $\varepsilon_i^{n,k}$ follows a normal distribution for any n, k, i.

Theorem 2.2. *Assume (A1$''$), (A2) and (A3). Then the family of distributions $\{P_{\sigma_*,v_*,n}\}_{\sigma_*,v_*,n}$ has the LAN property with Γ in (2.11) and $\delta_n = \mathrm{diag}(b_n^{-1/4}\mathcal{E}_d, b_n^{-1/2}\mathcal{E}_2)$.*

Remark 2.6. Jeganathan [20] studied lower bounds of estimation errors for any estimator of parameters. They showed a version of Hájek's convolution theorem (Corollary 1) and that the optimal asymptotic variance of errors for regular estimators is Γ^{-1}, where Γ is in Definition 2.1. Therefore, Theorems 2.1 and 2.2 ensures that our estimator $\hat{\sigma}_n$ of the parameter σ is asymptotically efficient in this sense under the assumptions of both theorems.

Remark 2.7. The assumptions of Theorem 2.2 are rather strong conditions. We are also interested in the LAMN property in more general settings. In particular, we are interested in the case that $\mu_t = \mu(t, X_t, \sigma_*)$ and μ and b are general functions with suitable conditions. However, we need further analysis using Malliavin calculus to deal with the LAMN property of general diffusion processes, as seen in Gobet [14] and Ogihara [24]. Moreover, when we deal with nonsynchronousness or noise, the true likelihood function is obtained by integrating a likelihood function of ideal observations with respect to unobservable variables. Such integral is not easy to handle, and consequently, it makes the proof of LAMN much more complicated as seen in [24]. To the best of author's knowledge, the LAMN property for a general diffusion model has not been shown even for models with noisy, synchronous observations. We have left this for future works. On the other hand, asymptotic results of the quasi-likelihood function shown in this paper (e.g., Propositions 2.1 and 7.2) are expected to be useful when we try to show LAMN for general cases. Indeed, LAMN for a model with nonsynchronous observations is proved in [24] by showing asymptotic equivalence of the true likelihood ratio and a quasi-likelihood ratio, and using asymptotic results of the quasi-likelihood ratio in Ogihara and Yoshida [25].

2.4. A Bayes-type estimator and convergence of moments of estimation errors

Polynomial-type large deviation theory by Yoshida [29,30] enables us to address the asymptotic properties of a Bayes-type estimator and the convergence of moments of estimation errors, which is a stronger result than asymptotic mixed normality. Convergence of moments is useful when we investigate the theory of information criteria, minimax inequality and asymptotic expansion of estimators. See Uchida [27] for a theory of contrast-based information criteria for ergodic diffusion processes with equidistant observations. We also see asymptotic efficiency of our estimator in the sense of minimax inequality.

We first assume following stronger conditions than (A1)–(A3) and (V).

Assumption (B1). 1. (A1) holds true with $O = \mathbb{R}^{d_2}$.

2. There exists a positive constant C such that $\sup_{t\in[0,T],\sigma\in\Lambda} |\partial_t^i \partial_x^j \partial_\sigma^k b(t,x,\sigma)| \leq C(1+|x|)^C$ for $0 \leq 2i + j \leq 4$, $0 \leq k \leq 4$ and $x \in \mathbb{R}^{d_2}$.

3. $\inf_{t,x,\sigma} \det bb^\top(t,x,\sigma) > 0$.

4. $E[|Y_0|^q] < \infty$ for any $q > 0$.

5. $\sup_t E[|\mu_t|^q] < \infty$, $\sup_{s<t}(E[|\mu_t - \mu_s|^q]^{1/q}(t-s)^{-1/2}) < \infty$ and $\sup_{s<t} E[(E[\mu_t - \mu_s|\mathcal{G}_s]/(t-s))^q] < \infty$ for any $q > 0$.

6. For any $q > 0$, $\max_j \sup_t E[|b_t^{(j)}|^q \vee |\hat{b}_t^{(j)}|^q] < \infty$ and $\max_j \sup_{s<t}(E[|b_t^{(j)} - b_s^{(j)}|^q \vee |\hat{b}_t^{(j)} - \hat{b}_s^{(j)}|^q]^{1/q}(t-s)^{-1/2}) < \infty$.

Assumption (B2). There exist $\eta \in (0,1/2)$, $\dot{\eta} \in (0,1]$, $\delta > 0$ and positive-valued stochastic processes $\{a_t^j\}_{t\in[0,T],j=1,2}$ such that $b_n^{-1/2} k_n (b_n^{-1} k_n)^{\dot{\eta}} \to 0$ as $n \to \infty$, $E[\sup_{j,t>s}(|a_t^j - a_s^j|^q |t-s|^{-q\dot{\eta}})] < \infty$, $E[\sup_{j,t} |a_t^j|^q] \vee E[\sup_{j,t}(|a_t^j|^{-q})] < \infty$, and

$$\sup_n \sup_{\{[s'_{n,l},s''_{n,l})\}\in\mathcal{S}_\eta} E\Big[\Big(k_n b_n^{-1/2+\delta} \max_{1\leq l\leq L_n} \Big| b_n^{-1}\big(s''_{n,l} - s'_{n,l}\big)^{-1} \#\big\{i; \big[S_{i-1}^{n,j}, S_i^{n,j}\big) \subset \big[s'_{n,l}, s''_{n,l}\big)\big\} - a_{s'_{n,l}}^j \Big|\Big)^q\Big]$$

is finite for any $q > 0$. Moreover, there exists a positive constant γ such that $k_n b_n^{-4/7+\gamma} \to 0$ and $E[((r_n b_n^{1-\varepsilon}) \vee (\underline{r}_n^{-1} b_n^{-1-\varepsilon}))^q] \to 0$ as $n \to \infty$ for any $q > 0$ and $\varepsilon > 0$.

Assumption (B3). For any $q > 0$, there exists a positive constant c_q such that $P[\inf_{\sigma\neq\sigma_*}((-\mathcal{Y}_0(\sigma))/|\sigma - \sigma_*|^2) \leq r^{-1}] \leq c_q/r^q$ for any $r > 0$.

Assumption (B4). There exist estimators $\{\hat{v}_n\}_{n\in\mathbb{N}}$ of v_* such that $\hat{v}_n > 0$ almost surely, $\limsup_n E[\hat{v}_n^{-q}] < \infty$, $\sup_n E[|b_n^{1/2}(\hat{v}_n - v_*)|^q] < \infty$, and $\sup_n E[b_n^{-\varepsilon} \max_{m,k}(E_\Pi[|\ell_n^{1/2} \times \mathbf{T}_{m,k}|^q \vee |\ell_n E[\mathbf{T}_{m,k}|\mathcal{G}_{s_{m-1}}]|^q])] < \infty$ for any $q > 0$ and $\varepsilon > 0$.

Though Assumption (B3) is rather difficult to check in a practical setting, Uchida and Yoshida [28] investigated sufficient conditions for (B3). The simplest condition is that (B3) is satisfied if there exists $\varepsilon > 0$ such that $|bb^\top(t,x,\sigma_1) - bb^\top(t,x,\sigma_2)| \geq \varepsilon|\sigma_1 - \sigma_2|$ for any $t \in [0,T]$, $x \in O$ and $\sigma_1, \sigma_2 \in \Lambda$. See Remark 4 in [25] for details.

Let $U_n = \{u \in \mathbb{R}^d; \sigma_* + b_n^{-1/4} u \in \Lambda\}$, $V_n(r) = \{|u| \geq r\} \cap U_n$, and $\mathbf{Z}_n(u) = \exp(H_n(\sigma_* + b_n^{-1/4} u, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n))$ for $u \in U_n$.

Proposition 2.2 (Polynomial-type large deviation inequalities). *Assume (B1)–(B4). Then for any $L > 0$, there exists a positive constant c_L such that $P[\sup_{u\in V_n(r)} \mathbf{Z}_n(u) \geq e^{-r/2}] \leq c_L/r^L$ for any $n \in \mathbb{N}$ and $r > 0$.*

Since $\mathbf{Z}_n(0) = 1$, Proposition 2.2 immediately yields

$$\begin{aligned} E\big[\big|b_n^{1/4}(\hat{\sigma}_n - \sigma_*)\big|^p\big] &= \int_0^\infty pt^{p-1} P\big[\big|b_n^{1/4}(\hat{\sigma}_n - \sigma_*)\big| \geq t\big]\, dt \\ &\leq \int_0^\infty pt^{p-1} P\Big[\sup_{u \in V_n(t)} \mathbf{Z}_n(u) \geq e^{-t/2}\Big]\, dt < \infty \end{aligned} \tag{2.12}$$

for any $p > 0$. Moreover, we obtain the following convergence of moments of the estimation error.

Theorem 2.3. *Assume (B1)–(B4). Then* $E[\mathbf{Y} f(b_n^{1/4}(\hat{\sigma}_n - \sigma_*))] \to E[\mathbf{Y} f(\Gamma_1^{-1/2}\mathcal{N})]$ *as* $n \to \infty$ *for any bounded random variable* $\mathbf{Y}$ *on* $(\Omega, \mathcal{F})$ *and any continuous function* f *of at most polynomial growth.*

In particular, we obtain convergence of moments where $E[|b_n^{1/4}(\hat{\sigma}_n - \sigma_*)|^q] \to E[|\Gamma_1^{-1/2}\mathcal{N}|^q]$ for any $q > 0$. This property is used when we study the theory of information criteria and asymptotic expansion of estimators.

We also obtain results for a Bayes type estimator. Let a prior density $\pi : \Lambda \to (0, \infty)$ be a continuous function satisfying $0 < \inf_\sigma \pi(\sigma) \leq \sup_\sigma \pi(\sigma) < \infty$. Then a Bayes-type estimator $\tilde{\sigma}_n$ for the quadratic loss function is defined by

$$\tilde{\sigma}_n = \left(\int_\Lambda \exp\big(H_n(\sigma, \hat{v}_n)\big)\pi(\sigma)\, d\sigma\right)^{-1} \int_\Lambda \sigma \exp\big(H_n(\sigma, \hat{v}_n)\big)\pi(\sigma)\, d\sigma.$$

Since the Bayes-type estimator $\tilde{\sigma}_n$ contains integrals with respect to σ, we need to deal with tail behaviors of likelihood ratio $H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)$. Hence, Proposition 2.2 is essential to deduce asymptotic properties of a Bayes-type estimator. Since the Bayes-type estimator can be calculated using Markov-Chain Monte Carlo methods, it is often easier to calculate than the maximum-likelihood-type estimator. For the Bayes-type estimator $\tilde{\sigma}_n$, we obtain similar results to the ones for the maximum-likelihood-type estimator.

Theorem 2.4. *Assume (B1)–(B4). Then* $E[\mathbf{Y} f(b_n^{1/4}(\tilde{\sigma}_n - \sigma_*))] \to E[\mathbf{Y} f(\Gamma_1^{-1/2}\mathcal{N})]$ *as* $n \to \infty$ *for any bounded random variable* $\mathbf{Y}$ *on* $(\Omega, \mathcal{F})$ *and any continuous function* f *of at most polynomial growth.*

Remark 2.8. If the assumptions of Theorem 2.2 are satisfied, the asymptotic minimax theorem (Theorem 4 in [21]) holds for our model, so

$$\lim_{\alpha \to \infty} \liminf_{n \to \infty} \sup_{|u| \leq \alpha} E_{\sigma_* + b_n^{-1/4}u}\big[l\big(\big|b_n^{1/4}\big(V_n - \sigma_* - b_n^{-1/4}u\big)\big|\big)\big] \geq E\big[l\big(|\Gamma_1\mathcal{N}|\big)\big]$$

for any estimators $\{V_n\}_n$ of the parameter and any function $l : [0, \infty) \to [0, \infty)$ which is nondecreasing and $l(0) = 0$, where E_σ denotes expectation with respect to $P_{\sigma, v_*, n}$. Using Theorems 2.3 and 2.4 and a similar argument in Theorem 2.2 of Ogihara [24], we can see that $\hat{\sigma}_n$ and $\tilde{\sigma}_n$ attain the lower bound of the above inequality for continuous l of at most polynomial growth, if further (B2) and uniform versions of (B3) and (B4) with respect to the true value (σ_*, v_*) are satisfied. Hence our estimators are asymptotically efficient in this sense as well.

3. Simulation results

In this section, we examine some simulation results of our estimator.

First, we consider the case where the latent process Y is a Brownian motion, that is, Y satisfies the following stochastic differential equation:

$$\begin{cases} dY_t^1 = \sigma_{1,*}\, dW_t^1, \\ dY_t^2 = \sigma_{3,*}\, dW_t^1 + \sigma_{2,*}\, dW_t^2, \end{cases}$$

where $\sigma_* = (\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*}) \in (\varepsilon, R) \times (-R, R) \times (\varepsilon, R)$ for some $0 < \varepsilon < R$. Moreover, let $\{N_t^1\}_{0 \le t \le T}$ and $\{N_t^2\}_{0 \le t \le T}$ be two independent Poisson processes with parameters λ_1 and λ_2, respectively. We give sampling times by $S_i^{n,j} = \inf\{N_{nt}^j \ge j\} \wedge T$ for $j = 1, 2$. Let $\{\varepsilon_i^{n,j}\}_{i \in \mathbb{Z}_+, j=1,2}$ be independent normal random variables with $E[\varepsilon_i^{n,j}] = 0$ and $E[(\varepsilon_i^{n,j})^2] = v_{j,*}$.

Then we can see that this example satisfies (A1″), (A2) and (A3′). So the maximum-likelihood-type estimator $\hat{\sigma}_n$ is asymptotically mixed normal and asymptotically efficient with the asymptotic variance Γ_1^{-1}. For the estimator $\hat{v}_n$ of v_* we first use a simple estimator $\hat{v}_n = (2\mathbf{J}_{k,n})^{-1} \sum_i (\tilde{Y}_i^k - \tilde{Y}_{i-1}^k)^2$, which means that our estimator is calculated by $\hat{\sigma}_n = \operatorname{argmax}_\sigma H_n(\sigma, \hat{v}_n)$. We also consider a plug-in estimator $\hat{v}'_{k,n} = (\hat{v}_{k,n} - |b^k(\hat{\sigma}_n)|^2 T/(2\mathbf{J}_{k,n})) \vee 0$ of $v_{k,*}$, and $\hat{\sigma}'_n = \operatorname{argmax}_\sigma H_n(\sigma, \hat{v}'_n)$. Let $\hat{\sigma}''_n = \operatorname{argmax}_\sigma H_n(\sigma, v_*)$. Then $\hat{\sigma}''_n$ cannot be calculated by observed data, but we can use it for comparison. Though these estimators have the same asymptotic variance, their performances for finite samples are different. In particular, we cannot ignore the bias of $\hat{v}_n$ since v is relatively small compared with σ in practical data.

Table 1 shows results of 1000 independent estimations. Each cell represents the average of estimators, with sample standard deviations given in parentheses. We set the values of parameters as $k_n = [n^{5/8}]$, $T = 1$, $(\lambda_1, \lambda_2) = (1, 1)$, $(\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*}) = (1, \sqrt{1 - 0.5^2}, 0.5)$, and consider two cases of the noise variances : $v_* = (0.001, 0.001)$ and $v_* = (0.005, 0.005)$. In both cases, we can see that $\hat{v}_n$ has an upper bias for $n = 1000$, and causes a lower bias of $\hat{\sigma}_n$ because $\hat{v}_n$ contains variance of the latent process, which is always positive. These biases can be moderated by using the plug-in estimator. For $n = 5000$, the plug-in estimator $\hat{\sigma}'_n$ performs as well as $\hat{\sigma}''_n$. In the case of $v_* = (0.005, 0.005)$, the biases of $\hat{v}_n$ and $\hat{v}'_n$ are relatively small, so the performance of $\hat{\sigma}_n$ and $\hat{\sigma}'_n$ are better.

We can also construct an estimator $\hat{\sigma}'_{1,n}\hat{\sigma}'_{3,n}T$ of the quadratic covariation $\langle Y^1, Y^2 \rangle_T = \sigma_{1,*}\sigma_{3,*}T$. We see that

$$\begin{aligned} & n^{1/4}\big(\hat{\sigma}'_{1,n}\hat{\sigma}'_{3,n}T - \langle Y^1, Y^2 \rangle_T\big) \\ & \quad \xrightarrow{d} N\big(0, T^2\big(\sigma_{3,*}^2\big(\Gamma_1^{-1}\big)_{11} + 2\sigma_{1,*}\sigma_{3,*}\big(\Gamma_1^{-1}\big)_{13} + \sigma_{1,*}^2\big(\Gamma_1^{-1}\big)_{33}\big)\big) \end{aligned} \tag{3.1}$$

Table 1. Simulation results for estimators of parameters

Results with $v_* = (0.001, 0.001)$

n		σ_1	σ_2	σ_3	v_1	v_2
1000	$(\hat{\sigma}_n, \hat{v}_n)$	0.897	0.776	0.451	0.001504	0.001500
		(0.040)	(0.042)	(0.062)	(0.000079)	(0.000080)
	$(\hat{\sigma}'_n, \hat{v}'_n)$	0.971	0.840	0.487	0.001100	0.001094
		(0.046)	(0.047)	(0.067)	(0.000075)	(0.000078)
	$\hat{\sigma}''_n$	0.999	0.863	0.501	–	–
		(0.045)	(0.046)	(0.068)	–	–
5000	$(\hat{\sigma}_n, \hat{v}_n)$	0.964	0.833	0.481	0.001099	0.001099
		(0.028)	(0.029)	(0.040)	(0.000026)	(0.000026)
	$(\hat{\sigma}'_n, \hat{v}'_n)$	0.997	0.862	0.498	0.001006	0.001006
		(0.031)	(0.031)	(0.041)	(0.000027)	(0.000027)
	$\hat{\sigma}''_n$	0.999	0.864	0.499	–	–
		(0.029)	(0.030)	(0.041)	–	–
True values		1	0.866	0.5	0.001	0.001

Results with $v_* = (0.005, 0.005)$

n		σ_1	σ_2	σ_3	v_1	v_2
1000	$(\hat{\sigma}_n, \hat{v}_n)$	0.957	0.818	0.481	0.005515	0.005501
		(0.086)	(0.143)	(0.094)	(0.000293)	(0.000296)
	$(\hat{\sigma}'_n, \hat{v}'_n)$	0.991	0.850	0.498	0.005053	0.005035
		(0.092)	(0.139)	(0.098)	(0.000298)	(0.000306)
	$\hat{\sigma}''_n$	0.997	0.861	0.499	–	–
		(0.069)	(0.070)	(0.096)	–	–
5000	$(\hat{\sigma}_n, \hat{v}_n)$	0.990	0.854	0.495	0.005095	0.005096
		(0.044)	(0.044)	(0.061)	(0.000121)	(0.000123)
	$(\hat{\sigma}'_n, \hat{v}'_n)$	0.999	0.862	0.499	0.004996	0.004998
		(0.045)	(0.045)	(0.062)	(0.000123)	(0.000125)
	$\hat{\sigma}''_n$	0.998	0.862	0.499	–	–
		(0.043)	(0.044)	(0.062)	–	–
True values		1	0.866	0.5	0.005	0.005

as $n \to \infty$ by the delta method, and the estimator is asymptotically efficient since we can reparameterize the model using $\sigma_{1,*}\sigma_{3,*}$. We therefore compared the performance of the estimator (MLE) with existing estimators of the quadratic covariation. We used the pre-averaged Hayashi–Yoshida estimator (PHY) and modulated realised covariance (MRC) by Christensen, Kinnebrock, and Podolskij [7], the local method of moments (LMM) by Bibinger et al. [5], and an estimator based on maximum likelihood estimator of a model of constant diffusion coefficients (QMLE) by Aït-Sahalia, Fan, and Xiu [2] for comparison. Except LMM these estimators

Table 2. Comparison of estimators of $\langle Y^1, Y^2\rangle_T$

Resuls with $v_* = (0.001, 0.001)$							
n	MLE	PHY	MRC_1	MRC_2	QMLE	LMM	minimum
1000	0.474	0.499	0.508	0.501	0.501	0.463	
	(0.073)	(0.121)	(0.182)	(0.110)	(0.095)	(0.082)	(0.066)
5000	0.496	0.497	0.504	0.499	0.498	0.497	
	(0.046)	(0.081)	(0.124)	(0.073)	(0.056)	(0.069)	(0.044)
Results with $v_* = (0.005, 0.005)$							
n	MLE	PHY	MRC_1	MRC_2	QMLE	LMM	minimum
1000	0.496	0.497	0.508	0.5000	0.5000	0.518	
	(0.109)	(0.148)	(0.185)	(0.124)	(0.120)	(0.112)	(0.099)
5000	0.499	0.497	0.505	0.499	0.499	0.514	
	(0.069)	(0.098)	(0.126)	(0.083)	(0.079)	(0.083)	(0.066)

can be calculated using the "cce" function in the "yuima" R package (http://r-forge.r-project.org/projects/yuima). We used the default values of the "cce" function or values used in corresponding papers for parameters of estimators ($\theta = 0.15$ for PHY, $\theta = 1$ for MRC_1, $J = 30$, $h^{-1} = 10$ for LMM). Here we use *the oracle estimator* defined in [5] for LMM to avoid a complicated calculation. For the modulated realised covariance, we also examine an estimator MRC_2 with $\theta = 1/3$ which is used in Jacod et al. [19]. Table 2 shows the results of 1000 estimations. We used the same parameter values as above. Then the true value of the quadratic covariation becomes $\langle Y^1, Y^2\rangle_T = 0.5$. For both cases of observation noise variance, we can see that sample standard deviations of our estimator are the best in large samples. The theoretical (asymptotic) minimum $Tn^{-1/4}(\sigma_{3,*}^2(\Gamma_1^{-1})_{11} + 2\sigma_{1,*}\sigma_{3,*}(\Gamma_1^{-1})_{13} + \sigma_{1,*}^2(\Gamma_1^{-1})_{33})^{1/2}$ of standard deviations for all estimators is calculated in the last column of Table 2. We can see that the sample standard deviations of MLE are close to the minima in large samples.

The asymptotic variance in (3.1) is always the best in this parametric setting. On the other hand, Bibinger et al. [5] gave a Cramér–Rao lower bound for one-dimensional perturbation and the asymptotic variance of LMM becomes optimal in that sense. The derivation of their bound is quite different from that of (3.1), and therefore, it is difficult to judge when these variances are the same even in this simple Brownian model. However, there is an interesting numerical result on this point. Let V_{MLE} be the asymptotic variance of MLE for $\langle Y^1, Y^2\rangle_T$ and let V_{LMM} be that of LMM, and we set parameters as $T = 1$, $(\lambda_1, \lambda_2) = (1, 1)$, $(\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*}) = (1, \sqrt{1 - 0.5^2}, 0.5)$ and $v_* = (0.001, 0.001)$. Then we obtain $V_{\mathrm{MLE}} = T^2(\sigma_{3,*}^2(\Gamma_1^{-1})_{11} + 2\sigma_{1,*}\sigma_{3,*}(\Gamma_1^{-1})_{13} + \sigma_{1,*}^2(\Gamma_1^{-1})_{33}) = 0.1385502\cdots$ and $V_{\mathrm{LMM}} = 0.1385502\cdots$, that is, we see an exact numerical match between these variances. We have this agreement even if we set other values of parameters, and hence we expect that $V_{\mathrm{MLE}} = V_{\mathrm{LMM}}$ always holds for this model.

However, the performances of MLE and LMM are different in above simulation because it is not easy to set parameters of LMM suitably. If we set $J > 30$ for LMM, then the sample standard deviation is reduced, while the bias increase; for example, the average value of LMM is 0.51678 and the sample standard deviation is 0.05017 with $J = 100$, $n = 5000$ and $v_* = (0.001, 0.001)$ in Table 2.

On the other hand, the asymptotic variances of LMM and MLE are different each other in a model with a time-dependent diffusion coefficient. Let the diffusion coefficient b be defined by

$$b(t,\sigma) = f(t)\begin{pmatrix} \sigma_1 & 0 \\ \sigma_3 & \sigma_2 \end{pmatrix}$$

for some smooth, positive function f on [0, 1] with $T = 1$, and let $\{S_i^{n,j}\}_{n,j,i}$ and $\{\varepsilon_i^{n,j}\}_{n,j,i}$ be the same as the above example. Then we can calculate asymptotic variances $\tilde{V}_{\mathrm{MLE}}$ and $\tilde{V}_{\mathrm{LMM}}$ of MLE and LMM, respectively, as

$$\tilde{V}_{\mathrm{MLE}} = V_{\mathrm{MLE}} \times \left(\int_0^1 f(t)^2\,dt\right)^2 \Big/ \int_0^1 f(t)\,dt, \qquad \tilde{V}_{\mathrm{LMM}} = V_{\mathrm{LMM}} \times \int_0^1 f(t)^3\,dt.$$

The Cauchy–Schwarz inequality yields $\int_0^1 f(t)^3\,dt \geq (\int_0^1 f(t)^2\,dt)^2/\int_0^1 f(t)\,dt$, and the equality is attained if and only if f is a constant function. These results imply $\tilde{V}_{\mathrm{MLE}} < \tilde{V}_{\mathrm{LMM}}$ for any non-constant function f if $V_{\mathrm{MLE}} \leq V_{\mathrm{LMM}}$. For example, let us set $f(t) = 2 - \sin(\pi t)$ to capture the U-shape of intra-day activities of a stock market. Then we obtain

$$\tilde{V}_{\mathrm{MLE}}/\tilde{V}_{\mathrm{LMM}} = \frac{(\int_0^1 f(t)^2\,dt)^2}{\int_0^1 f(t)\,dt \int_0^1 f(t)^3\,dt} = 0.9533\cdots$$

if $V_{\mathrm{MLE}} = V_{\mathrm{LMM}}$. Thus, we can see the improvement of efficiency in a model with a time-dependent b.

In the next, we consider the model with random diffusion coefficients and non-Gaussian noise. As mentioned in Remark 2.1, we cannot directly apply our results to stochastic volatility models. Here we consider the Cox-Ingersoll-Ross (CIR) process derived in [9] as a latent process with random diffusion coefficients. Let the latent process Y satisfy

$$dY_t = \begin{pmatrix} \alpha_1 - \beta_1 Y_t^1 \\ \alpha_2 - \beta_2 Y_t^2 \end{pmatrix} dt + \begin{pmatrix} \sigma_{1,*}\sqrt{Y_t^1} & 0 \\ \sigma_{3,*}\sqrt{Y_t^2} & \sigma_{2,*}\sqrt{Y_t^2} \end{pmatrix} dW_t,$$

where $\sigma_* = (\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*}) \in (\varepsilon', R') \times (-R', R') \times (\varepsilon', R')$. We assume Conditions $2\alpha_1 > \sigma_{1,*}^2$ and $2\alpha_2 > \sigma_{2,*}^2 + \sigma_{3,*}^2$ which ensure $Y_t^1 > 0$ and $Y_t^2 > 0$ for $t \in [0,T]$ almost surely. Let $\{\varepsilon_i^{n,j}\}_{i\in\mathbb{Z}}$ be i.i.d. random variables following a centered Gamma distribution with a shape parameter k_j and a scale parameter θ_j for $j = 1, 2$. We define $\{N_t^j\}$, $\hat{v}_n$, $\hat{v}_n'$, $\hat{\sigma}_n$, and $\hat{\sigma}_n'$ similarly to the first example. We set the values of parameters as $k_n = [n^{5/8}]$, $T = 1$, $(\lambda_1, \lambda_2) = (1, 1)$, $(\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*}) = (1, \sqrt{1 - 0.5^2}, 0.5)$, $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (1, 1, 1, 1)$, and

Table 3. Estimation errors of estimators of $\langle Y^1, Y^2\rangle_T$ for the CIR process

n	MLE	PHY	MRC_1	MRC_2	QMLE	LMM
1000	−0.0267	−0.0063	−0.0058	−0.0036	−0.0008	−0.0348
	(0.0733)	(0.1286)	(0.1867)	(0.1162)	(0.1013)	(0.0844)
5000	−0.0023	−0.0036	−0.0022	−0.0016	−0.0005	−0.0033
	(0.0456)	(0.0858)	(0.1305)	(0.0768)	(0.0580)	(0.0719)

$(k_1, k_2, \theta_1, \theta_2) = (2, 2, \sqrt{0.0005}, \sqrt{0.0005})$ which implies $v_* = (0.001, 0.001)$. Table 3 shows averages and sample standard deviations of $T_n - \langle Y^1, Y^2\rangle_T$ for each estimator T_n of the quadratic covariation $\langle Y^1, Y^2\rangle_T$ in 1000 simulations. $\langle Y^1, Y^2\rangle_T$ is random in this model since the diffusion coefficients are random. So we use extra-high-frequency observations $\{Y^l_{k/100\,000}\}_{k=0}^{100\,000}$ of Y to calculate the approximated true value of $\langle Y^1, Y^2\rangle_T$. In this model, we have not obtained the LAMN property nor asymptotic efficiency of our estimator though we expect to obtain them. However, we still see that our estimator achieves the best error variance in large samples.

4. Asymptotically equivalent representation of the quasi-likelihood function

We will prove our main results in the rest of this paper. In this section, we introduce an asymptotically equivalent representation $\tilde{H}_n(\sigma, v)$ of the quasi-likelihood function $H_n(\sigma, v)$, and prove the equivalence. $\tilde{H}_n$ is a useful function for deducing the limit of H_n.

4.1. Some notations

We denote E_m as the $\mathcal{G}_{s_{m-1}}$-conditional expectation and $\bar{E}_m[\mathbf{X}] = \mathbf{X} - E_m[\mathbf{X}]$ for a random variable $\mathbf{X}$. We use the symbol C for a generic positive constant that can vary from line to line.

For a sequence c_n of positive-valued $\mathfrak{B}(\Pi_n)$-measurable random variables, let us denote by $\{\bar{R}_n(c_n)\}_{n\in\mathbb{N}}$, $\{\underline{\mathrm{R}}_n(c_n)\}_{n\in\mathbb{N}}$ and $\{\dot{R}_n(c_n)\}_{n\in\mathbb{N}}$ sequences of random variables (which may depend on $1 \le m \le \ell_n$ and σ) satisfying

$$E\Big[\Big(c_n^{-1}(r_n/b_n)^{-p_1}(b_n/\underline{\mathrm{r}}_n)^{-p_2}(\bar{k}_n/k_n)^{-p_3}(k_n/\underline{\mathrm{k}}_n)^{-p_4}b_n^{-\delta}\sup_{\sigma,m}E_\Pi\big[\big|\bar{R}_n(c_n)\big|^q\big]^{1/q}\Big)^{q'}\Big] \to 0,$$

$$E\Big[\Big(c_n^{-1}(r_n/b_n)^{q_1}(b_n/\underline{\mathrm{r}}_n)^{q_2}(\bar{k}_n/k_n)^{q_3}(k_n/\underline{\mathrm{k}}_n)^{q_4}b_n^{\delta'}\sup_{\sigma,m}E_\Pi\big[\big|\underline{\mathrm{R}}_n(c_n)\big|^q\big]^{1/q}\Big)^{q'}\Big] \to 0,$$

and

$$c_n^{-1}(r_n/b_n)^{q_1}(b_n/\underline{\mathrm{r}}_n)^{q_2}(\bar{k}_n/k_n)^{q_3}(k_n/\underline{\mathrm{k}}_n)^{q_4}\sup_{\sigma,m}\big|\dot{R}_n(c_n)\big| \to^p 0,$$

respectively, as $n \to \infty$ for any $\delta, q, q', q_1, \dots, q_4 > 0$ with some constants $\delta', p_1, \dots, p_4 \ge 0$.

Let $M_m(v) = \mathrm{diag}(v_1 M_{1,m}, v_2 M_{2,m})$ for $v = (v_1, v_2)$, $\tilde{b}^k_m = b^k(s_{m-1}, X_{s_{m-1}}, \sigma)$, $\tilde{b}^k_{m,*} = b^k(s_{m-1}, X_{s_{m-1}}, \sigma_*)$,

$$\tilde{Z}_m = \Big(\big(\big(\tilde{b}^1_{m,*} \cdot (W_{S^{n,1}_i} - W_{S^{n,1}_{i-1}}) + \varepsilon^{n,1}_i - \varepsilon^{n,1}_{i-1}\big)_{i=K^1_{m-1}+2}^{K^1_m} \big)^\top,$$

$$\big(\big(\tilde{b}^2_{m,*} \cdot (W_{S^{n,2}_j} - W_{S^{n,2}_{j-1}}) + \varepsilon^{n,2}_j - \varepsilon^{n,2}_{j-1}\big)_{j=K^2_{m-1}+2}^{K^2_m} \big)^\top \Big)^\top,$$

$$\tilde{S}_m(\sigma, v) = \begin{pmatrix} \mathrm{diag}\big(\big(|\tilde{b}^1_m|^2 |I^1_{i,m}|\big)_i\big) & \big\{\tilde{b}^1_m \cdot \tilde{b}^2_m |I^1_{i,m} \cap I^2_{j,m}|\big\}_{ij} \\ \big\{\tilde{b}^1_m \cdot \tilde{b}^2_m |I^1_{i,m} \cap I^2_{j,m}|\big\}_{ji} & \mathrm{diag}\big(\big(|\tilde{b}^2_m|^2 |I^2_{j,m}|\big)_j\big) \end{pmatrix} + M_m(v), \tag{4.1}$$

and

$$\tilde{H}_n(\sigma, v) = -\frac{1}{2} \sum_{m=2}^{\ell_n} \tilde{Z}_m^\top \tilde{S}_m^{-1}(\sigma, v) \tilde{Z}_m - \frac{1}{2} \sum_{m=2}^{\ell_n} \log\det \tilde{S}_m(\sigma, v).$$

The diffusion coefficient b in $\tilde{Z}_m$ and $\tilde{S}_m$ are either $b(s_{m-1}, X_{s_{m-1}}, \sigma)$ or $b(s_{m-1}, X_{s_{m-1}}, \sigma_*)$. Hence, we do not need to consider the time-dependent structure of b when we study asymptotics of the summands in $\tilde{H}_n$. In particular, we obtain $E_m[\tilde{Z}_m^\top \partial_\sigma \tilde{S}_m(\sigma_*, v_*)^{-1} \tilde{Z}_m + \partial_\sigma \log\det \tilde{S}_m(\sigma_*, v_*)] = 0$ by $\partial_\sigma \log\det \tilde{S}_m(\sigma, v) = \mathrm{tr}(\partial_\sigma \tilde{S}_m \tilde{S}_m^{-1})(\sigma, v)$. We will prove the asymptotic equivalence of H_n and $\tilde{H}_n$ and then investigate asymptotic properties of $\tilde{H}_n$ instead of H_n.

Similarly to the approach of Gloter and Jacod [13], we can use the following stronger condition (A1′) instead of (A1) in several part of the proof. Then localization techniques and Girsanov's theorem enable us to replace (A1′) with (A1).

Assumption (A1′). Assumption (A1) is satisfied, $O = \mathbb{R}^{d_2}$, $\sup_{t,x,\sigma} \|(bb^\top)^{-1}\|(t, x, \sigma) < \infty$, $\mu_t \equiv 0$ and Y_0, $\sup_t(|b^{(l)}_t| \vee |\hat{b}^{(l)}_t|)$, and $\partial_t^i \partial_x^j \partial_\sigma^k b$ are all bounded for $l = 0, 1$, $0 \le 2i + j \le 4$ and $0 \le k \le 4$.

We can also see that (A1′) implies (B1).

4.2. Fundamental properties of the noise covariance matrix

In the following subsection, we will show the asymptotic equivalence of $\tilde{H}_n$ and H_n, namely that

$$b_n^{-1/2} \sup_{\sigma \in \Lambda} \big| \partial_\sigma^j \big(H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)\big) - \partial_\sigma^j \big(\tilde{H}_n(\sigma, v_*) - \tilde{H}_n(\sigma_*, v_*)\big) \big| \xrightarrow{p} 0$$

as $n \to \infty$ for $0 \le j \le 3$. To that end, we first show fundamental properties of S_m and $\tilde{S}_m$. These matrices inherit some properties of $M_{j,m}$, that are necessary to deduce the limit of H_n and $\tilde{H}_n$. The first property (4.2) concerns the trace of a matrix related to $M_{j,m}$ investigated by [13]. In the one-dimensional model with noisy, equidistance observations, this property can be directly applied to the quasi-likelihood function because the covariance matrix of the latent process is the unit matrix. However, this is insufficient for our purpose because our covariance matrix $S_m(\sigma, v) - M_m(v)$ of the latent process is rather complicated. Therefore, we investigate further matrix properties related to $M_{j,m}$.

First, we consider the results in [13]. For any positive constants p, q, a and b, eigenvalues of $(a\mathcal{E} + M_{j,m})^{-1}$ are $\{(a + 2(1 - \cos(i\pi(k_m^j + 1)^{-1}))\}_{i=1}^{k_m^j}$ and we obtain

$$\pi^{-1}\big(k_m^j + 1\big)I_p(a) - a^{-p} \le \mathrm{tr}\big((a\mathcal{E} + M_{j,m})^{-p}\big) \le \pi^{-1}\big(k_m^j + 1\big)I_p(a), \tag{4.2}$$

$$\begin{aligned}\pi^{-1}\big(k_m^j + 1\big)I_{p,q}(a, b) - a^{-p}b^{-q} &\le \mathrm{tr}\big((a\mathcal{E} + M_{j,m})^{-p}(b\mathcal{E} + M_{j,m})^{-q}\big) \\ &\le \pi^{-1}\big(k_m^j + 1\big)I_{p,q}(a, b),\end{aligned} \tag{4.3}$$

where $I_p(a) = \int_0^\pi (a + 2(1 - \cos x))^{-p}\,dx$ and $I_{p,q}(a, b) = \int_0^\pi (a + 2(1 - \cos x))^{-p}(b + 2(1 - \cos x))^{-q}\,dx$. Simple calculations show that $I_1(a) = \pi/\sqrt{a(4 + a)}$, $I_2(a) = \pi(2 + a)a^{-3/2}(4 + a)^{-3/2}$ and $\int_0^\pi \{\log(a + 2(1 - \cos x)) - \log(b + 2(1 - \cos x))\}\,dx = 2\pi(\log(\sqrt{a} + \sqrt{4 + a}) - \log(\sqrt{b} + \sqrt{4 + b}))$. See Section 4.1 in [13] for the details. Moreover, differentiation with respect to a yields

$$I_p(a) = \frac{(-1)^{p-1}}{(p-1)!}\left(\frac{d}{da}\right)^{p-1}\left(\frac{\pi}{\sqrt{a(4+a)}}\right).$$

In particular, if $a = \mathbf{X}_n b_n^{-1}$ for some tight random variables $\{\mathbf{X}_n\}_n$, then we have

$$I_p\big(\mathbf{X}_n b_n^{-1}\big) = \frac{\pi(2p-3)!!}{2^p(p-1)!}\big(\mathbf{X}_n b_n^{-1}\big)^{-p+1/2} + O_p\big(b_n^{p-3/2}\big).$$

For $\varepsilon \ge 0$, let $\{p_j(\varepsilon)\}_{j\in\mathbb{N}}$ and $\{p'_j(\varepsilon)\}_{j\in\mathbb{N}}$ be sequences of positive numbers satisfying $p_1(\varepsilon) = 2 + \varepsilon$, $p'_1(\varepsilon) = 1 + \varepsilon$, $p_{j+1}(\varepsilon) = 2 + \varepsilon - 1/p_j(\varepsilon)$, and $p'_{j+1}(\varepsilon) = 2 + \varepsilon - 1/p'_j(\varepsilon)$ for $j \in \mathbb{N}$. Let $E_{i,j}(a)$ be a $k_m^j \times k_m^j$ matrix satisfying $(E_{i,j}(a))_{k,l} = \delta_{kl} + a\delta_{ik}\delta_{jl}$ for $a \in \mathbb{R}$. Then we have that

$$E_{k_m^j, k_m^j - 1}\big(p_{k_m^j - 1}(\varepsilon)^{-1}\big)\cdots E_{2,1}\big(p_1(\varepsilon)^{-1}\big)(\varepsilon\mathcal{E} + M_{j,m})E_{1,2}\big(p_1(\varepsilon)^{-1}\big)\cdots E_{k_m^j - 1, k_m^j}\big(p_{k_m^j - 1}(\varepsilon)^{-1}\big)$$

is equal to $\mathrm{diag}((p_j(\varepsilon))_{j=1}^{k_m^j})$, and hence

$$\begin{aligned}(\varepsilon\mathcal{E} + M_{j,m})^{-1} &= \left\{\prod_{k+1\le i\le l} p_{i-1}(\varepsilon)^{-1}1_{\{k\le l\}}\right\}_{k,l} \\ &\quad\times \mathrm{diag}\Big(\big(p_j(\varepsilon)^{-1}\big)_{j=1}^{k_m^j}\Big)\left\{\prod_{l+1\le i\le k} p_{i-1}(\varepsilon)^{-1}1_{\{l\le k\}}\right\}_{k,l}.\end{aligned} \tag{4.4}$$

Moreover, we have the following lemma.

Lemma 4.1. *Let* $\varepsilon \in [0, 1)$ *and* $p_+(\varepsilon) = 1 + \varepsilon/2 + \sqrt{\varepsilon + \varepsilon^2/4}$. *Then*

1. $1 \le p'_j(\varepsilon) \le p_+(\varepsilon) < p_j(\varepsilon) \le 1 + 1/j + j\varepsilon$ *for* $j \in \mathbb{N}$, $\{p_j(\varepsilon)\}_j$ *is monotone decreasing, and* $\{p'_j(\varepsilon)\}_j$ *is monotone nondecreasing.*
2. $\{((\varepsilon\mathcal{E} + M_{j,m})^{-1})_{kk}\}_{k=1}^{[k_m^j/2]}$ *is monotone increasing.*
3. $p_j - p_+ \le (1+\sqrt{\varepsilon})^{-(j-2)}$ *and* $p_+ - p'_j \le \sqrt{\varepsilon}(1+\sqrt{\varepsilon})^{-(j-1)}$ *for* $j \ge 2$.
4. $\prod_{j=1}^k p'_j(\varepsilon) = (p_k(\varepsilon) - 1)\prod_{j=1}^{k-1} p_j(\varepsilon)$ *for any* $k \ge 2$.

Proof. 1. We simply denote $p_j = p_j(\varepsilon)$ and $p_+ = p_+(\varepsilon)$. We will prove $p_+ < p_j \le 1 + 1/j + j\varepsilon$ for $j \in \mathbb{N}$ by induction. The results obviously hold for $j = 1$. Assume the results hold for all values in $\mathbb{N}$ up to j. Then since $p_+ = 2 + \varepsilon - 1/p_+$, we obtain $p_{j+1} - p_+ = 1/p_+ - 1/p_j > 0$, and

$$\begin{aligned} p_{j+1} &\le 2 + \varepsilon - \bigl(j/(j+1)\bigr)\bigl(1 + j^2\varepsilon/(j+1)\bigr)^{-1} \le 2 + \varepsilon - \bigl(j/(j+1)\bigr)\bigl(1 - j^2\varepsilon/(j+1)\bigr) \\ &\le 1 + 1/(j+1) + (j+1)\varepsilon. \end{aligned}$$

Hence, we have $p_+ < p_j \le 1 + 1/j + j\varepsilon$ for $j \in \mathbb{N}$. Moreover, we can inductively deduce $p_{j+1} - p_j = 1/p_{j-1} - 1/p_j < 0$. The results for $\{p'_j(\varepsilon)\}_j$ are obtained similarly.

2. By considering the cofactor matrix and (4.4), we have

$$\begin{aligned} \bigl((\varepsilon\mathcal{E} + M_{j,m})^{-1}\bigr)_{kk} &= \frac{\det(\varepsilon\mathcal{E}_{k-1} + M(k-1))\det(\varepsilon\mathcal{E}_{k_m^j - k} + M(k_m^j - k))}{\det(\varepsilon\mathcal{E} + M_{j,m})} \\ &= \frac{\prod_{l=1}^{k-1} p_l \prod_{l=1}^{k_m^j - k} p_l}{\prod_{l=1}^{k_m^j} p_l}. \end{aligned} \tag{4.5}$$

Therefore, we obtain the result by monotonicity of p_j.

3. This is easy since $p_j - p_+ = (p_{j-1} - p_+)/p_+ p_{j-1} \le (p_1 - p_+)/p_+^{j-1} \le p_+^{-j+2}$.

4.

$$\begin{aligned} \prod_{j=1}^k p'_j(\varepsilon) &= \det\bigl(\varepsilon\mathcal{E} + M(k) - \bigl(E_{11}(1) - \mathcal{E}\bigr)\bigr) = \det\bigl(\varepsilon\mathcal{E} + M(k) - \bigl(E_{kk}(1) - \mathcal{E}\bigr)\bigr) \\ &= \bigl(p_k(\varepsilon) - 1\bigr)\prod_{j=1}^{k-1} p_j(\varepsilon). \end{aligned}$$

□

Let $\bar{\rho} = \sup_{t,\sigma}(|b^1 \cdot b^2||b^1|^{-1}|b^2|^{-1})(t, X_t, \sigma)$, $\tilde{D}_m = (\tilde{D}_{1,m}, \tilde{D}_{2,m})$, $\tilde{D}_{j,m} = \mathrm{diag}((|\tilde{b}_m^j|^2 \times |I_{i,m}^j|)_i) + v_{j,*}M_{j,m}$, $D'_m = (D'_{1,m}, D'_{2,m})$, $D'_{j,m} = \mathrm{diag}((|I_{i,m}^j|)_i)$, and $\check{D}_{j,m} = |\tilde{b}_m^j|^2 r_n\mathcal{E} + v_{j,*}M_{j,m}$.

Lemma 4.2. *Assume* (B1). *Then* $\mathrm{tr}(\tilde{S}_m^{-1}(\sigma, v_*)) = \bar{R}_n(b_n^{1/2}k_n)$.

Proof. Let $D''_m = \mathrm{diag}(|\tilde{b}^1_m|^2 D'_{1,m}, |\tilde{b}^2_m|^2 D'_{2,m})$ and $D'''_m = (D''_m)^{-1/2}\tilde{D}_m(D''_m)^{-1/2}$, then we have

$$\tilde{S}_m = (D''_m)^{1/2}(D'''_m)^{1/2}\big(\mathcal{E} + (D'''_m)^{-1/2}(D''_m)^{-1/2}(\tilde{S}_m - \tilde{D}_m)(D''_m)^{-1/2} \\ \times (D'''_m)^{-1/2}\big)(D'''_m)^{1/2}(D''_m)^{1/2}.$$

Moreover, Lemma A.4, (B1), and Lemma 2 in [25] yield

$$\begin{aligned}
&\big\|(D'''_m)^{-1/2}(D''_m)^{-1/2}(\tilde{S}_m - \tilde{D}_m)(D''_m)^{-1/2}(D'''_m)^{-1/2}\big\| \\
&\le \big\|(D''_m)^{-1/2}(\tilde{S}_m - \tilde{D}_m)(D''_m)^{-1/2}\big\| \\
&\le \bar{\rho}\left\{\left\|\left\{\frac{|I^1_{i,m}\cap I^2_{j,m}|}{|I^1_{i,m}|^{1/2}|I^2_{j,m}|^{1/2}}\right\}_{i,j}\right\| \vee \left\|\left\{\frac{|I^1_{i,m}\cap I^2_{j,m}|}{|I^1_{i,m}|^{1/2}|I^2_{j,m}|^{1/2}}\right\}_{j,i}\right\|\right\} \le \bar{\rho} < 1.
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
\mathrm{tr}(\tilde{S}_m^{-1}) &\le \mathrm{tr}\big((D'''_m)^{-1/2}(D''_m)^{-1}(D'''_m)^{-1/2}\big) \\
&\quad \times \big\|\big(\mathcal{E} + (D'''_m)^{-1/2}(D''_m)^{-1/2}(\tilde{S}_m - \tilde{D}_m)(D''_m)^{-1/2}(D'''_m)^{-1/2}\big)^{-1}\big\| \\
&\le \mathrm{tr}(\tilde{D}_m^{-1})/(1-\bar{\rho}) \le \sum_{j=1}^2 \mathrm{tr}(\check{D}_{j,m}^{-1}) r_n \underline{r}_n^{-1}(1-\bar{\rho})^{-1},
\end{aligned}$$

by Lemma A.1, the equation $\tilde{D}_{j,m} = \check{D}^{1/2}_{j,m}(\mathcal{E} - \check{D}^{-1/2}_{j,m}(\check{D}_{j,m} - \tilde{D}_{j,m})\check{D}^{-1/2}_{j,m})\check{D}^{1/2}_{j,m}$, and that

$$\big\|\check{D}^{-1/2}_{j,m}(\check{D}_{j,m} - \tilde{D}_{j,m})\check{D}^{-1/2}_{j,m}\big\| \le \big(|\tilde{b}^j_m|^2 r_n\big)^{-1}|\tilde{b}^j_m|^2(r_n - \underline{r}_n) = 1 - \underline{r}_n/r_n. \tag{4.6}$$

We thus obtain the desired results by (4.2). □

4.3. Asymptotic equivalence of H_n and $\tilde{H}_n$

In this section, we prove the asymptotic equivalence of H_n and $\tilde{H}_n$. We provide the following lemma about estimates of moments of the quantities related to H_n and $\tilde{H}_n$. The proof is given in the appendix; it is obtained based on the properties of $M_{j,m}$ in Section 4.2, standard Itô calculus, and some results from linear algebra.

Let $\tilde{S}_{m,*} = \tilde{S}_m(\sigma_*, v_*)$.

Lemma 4.3. *Assume (B1). Let $\mathbf{S} = \mathbf{S}_{n,m}$ be a symmetric, $\mathcal{G}_{s_{m-1}}$-measurable random matrix of size $k^1_m + k^2_m$ satisfying $\|\tilde{S}_{m,*}\mathbf{S}\tilde{S}_{m,*}\| \le b_n^{-1}$ for $n \in \mathbb{N}$ and $1 \le m \le \ell_n$. Then there exists a positive random sequence $\{Q_{n,q}\}_{q\ge 2}$ which does not depend on $\mathbf{S}$ such that $Q_{n,q} = \bar{R}_n(1)$ for $q \ge 2$ and*

1. $|E_m[(\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m)^2] - 2\,\mathrm{tr}((\mathbf{S}\tilde{S}_{m,*})^2) - \mathrm{tr}(\mathbf{S}\tilde{S}_{m,*})^2| \le Q_{n,2}$, $E_m[(\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m)^4] \le ((b_n^{-4}k_n^7) \vee (b_n^{-2}k_n^4))Q_{n,4}$ *and* $E_m[|\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m|^q] \le b_n^{-q}k_n^{2q}Q_{n,q}$ *for* $q > 4$.

2. $E_\Pi[|\sum_m (Z_m - \tilde{Z}_m)^\top \mathbf{S}(Z_m + \tilde{Z}_m)|^q] \leq (b_n^{-3}k_n^7)^{q/4} Q_{n,q}$ *for* $q \geq 4$.
3. $E_\Pi[|\sum_m (Z_m - \tilde{Z}_m)^\top \mathbf{S}(Z_m + \tilde{Z}_m)|^2] \leq b_n^{-1}k_n^2 Q_{n,2}$.

Proof. See the appendix. □

Now we obtain the asymptotic equivalence of H_n and $\tilde{H}_n$.

Lemma 4.4. *Assume* (B1), (A2) *and* (V). *Then*

$$b_n^{-1/2}\sup_{\sigma\in\Lambda}\bigl|\partial_\sigma^j\bigl(H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n)\bigr)-\partial_\sigma^j\bigl(\tilde{H}_n(\sigma,v_*)-\tilde{H}_n(\sigma_*,v_*)\bigr)\bigr|\to^p 0,$$

and $b_n^{-1/4}(\partial_\sigma H_n(\sigma_*,\hat{v}_n)-\partial_\sigma \tilde{H}_n(\sigma_*,v_*))\to^p 0$ *as* $n\to\infty$ *for* $0\leq j\leq 3$. *If* (B4) *holds as well, then*

$$\begin{aligned}&E_\Pi\Bigl[\Bigl(b_n^{-1/2}\sup_{\sigma\in\Lambda}\bigl|\partial_\sigma^j\bigl(H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n)\bigr)-\partial_\sigma^j\bigl(\tilde{H}_n(\sigma,v_*)-\tilde{H}_n(\sigma_*,v_*)\bigr)\bigr|\Bigr)^q\Bigr]\\&\quad=\bar{R}_n\bigl((b_n^{-5}k_n^7)^{q/4}\bigr)\end{aligned}$$

for any $0\leq j\leq 3$ *and* $q>0$.

Proof. We first obtain

$$\begin{aligned}&H_n(\sigma,v_*)-H_n(\sigma_*,v_*)-\bigl(\tilde{H}_n(\sigma,v_*)-\tilde{H}_n(\sigma_*,v_*)\bigr)\\&\quad=-\frac{1}{2}\sum_m\biggl\{(Z_m-\tilde{Z}_m)^\top\bigl(S_m^{-1}(\sigma,v_*)-S_m^{-1}(\sigma_*,v_*)\bigr)(Z_m+\tilde{Z}_m)\\&\qquad+\tilde{Z}_m^\top\bigl(S_m^{-1}(\sigma,v_*)-S_m^{-1}(\sigma_*,v_*)-\tilde{S}_m^{-1}(\sigma,v_*)+\tilde{S}_m^{-1}(\sigma_*,v_*)\bigr)\tilde{Z}_m\\&\qquad+\biggl(\log\frac{\det S_m(\sigma,v_*)}{\det\tilde{S}_m(\sigma,v_*)}-\log\frac{\det S_m(\sigma_*,v_*)}{\det\tilde{S}_m(\sigma_*,v_*)}\biggr)\biggr\}\\&\quad=:\hat{\Psi}_{1,n}(\sigma)+\hat{\Psi}_{2,n}(\sigma)+\hat{\Psi}_{3,n}(\sigma).\end{aligned}$$

We will give estimates for these quantities. Point 2 of Lemma 4.3 yields $\sup_\sigma E_\Pi[|b_n^{-1/2}\times\partial_\sigma^j\hat{\Psi}_{1,n}|^q]=\bar{R}_n((b_n^{-5}k_n^7)^{q/4})$ for $0\leq j\leq 4$ and $q>0$, and consequently by Sobolev's inequality $E_\Pi[\sup_\sigma|b_n^{-1/2}\partial_\sigma^j\hat{\Psi}_{1,n}|^q]=\bar{R}_n((b_n^{-5}k_n^7)^{q/4})$ as $n\to\infty$ for $0\leq j\leq 3$ and $q>d$.

Let

$$\begin{aligned}&\mathbf{C}(x,\sigma,v)\\&\quad=\begin{pmatrix}\{|b^1(s_{m-1},x,\sigma)|^2|I_{i,m}^1|\delta_{ii'}\}_{ii'}+v_1M_{1,m} & \{b^1\cdot b^2(s_{m-1},x,\sigma)|I_{i,m}^1\cap I_{j,m}^2|\}_{ij}\\ \{b^1\cdot b^2(s_{m-1},x,\sigma)|I_{i,m}^1\cap I_{j,m}^2|\}_{ji} & \{|b^2(s_{m-1},x,\sigma)|^2|I_{j,m}^2|\delta_{jj'}\}_{jj'}+v_2M_{2,m}\end{pmatrix}\end{aligned}$$

for $v=(v_1,v_2)$, $\tilde{\mathbf{C}}(x,v)=\mathbf{C}(x,\sigma,v)^{-1}-\mathbf{C}(x,\sigma_*,v)^{-1}$, $\mathbf{C}^{(1)}=\int_0^1\partial_x\tilde{\mathbf{C}}(t\hat{X}_{m-1}+(1-t)X_{s_{m-1}},v_*)\,dt$, $\mathbf{C}^{(2)}=\partial_x\tilde{\mathbf{C}}(X_{s_{m-2}},v_*)$ and $\mathbf{C}^{(3)}=\mathbf{C}(X_{s_{m-2}},\sigma_*,v_*)$. Then we obtain

$$\begin{aligned}
E_\Pi\big[|\hat{\Psi}_{2,n}|^q\big]&=2^{-q}E_\Pi\bigg[\bigg|\sum_m\tilde{Z}_m^\top\mathbf{C}^{(1)}\tilde{Z}_m(\hat{X}_{m-1}-X_{s_{m-1}})\bigg|^q\bigg]\\
&\le CE_\Pi\bigg[\bigg|\sum_m\operatorname{tr}\big(\mathbf{C}^{(1)}E_m\big[\tilde{Z}_m\tilde{Z}_m^\top\big]\big)(\hat{X}_{m-1}-X_{s_{m-1}})\bigg|^q\bigg]\\
&\quad+CE_\Pi\bigg[\bigg|\sum_m\operatorname{tr}\big(\mathbf{C}^{(1)}\bar{E}_m\big[\tilde{Z}_m\tilde{Z}_m^\top\big]\big)^2(\hat{X}_{m-1}-X_{s_{m-1}})^2\bigg|^{q/2}\bigg]\\
&\le CE_\Pi\bigg[\bigg|\sum_m\operatorname{tr}\big(\mathbf{C}^{(2)}\mathbf{C}^{(3)}\big)(\hat{X}_{m-1}-X_{s_{m-1}})\bigg|^q\bigg]+O_p\big(\big(\ell_nb_n^{-1/2}k_n\ell_n^{-1}\big)^qb_n^\varepsilon\big)\\
&\quad+O_p\big(\ell_n^{q/2}b_n^{-q}k_n^{2q}\ell_n^{-q/2}b_n^\varepsilon\big)\\
&=O_p\big(b_n^\varepsilon\ell_n^{q/2}b_n^{-q/2}k_n^q\ell_n^{-q/2}\big)+O_p\big(b_n^\varepsilon b_n^{-q/2}k_n^q\big)+O_p\big(b_n^\varepsilon b_n^{-q}k_n^{2q}\big)\\
&=O_p\big(b_n^{-q+\varepsilon}k_n^{2q}\big)
\end{aligned}$$

for any $q>0$ and $\varepsilon>0$, by Lemma 4.3, the Burkholder–Davis–Gundy inequality, $\sup_m E_\Pi[|E_{m-1}[\hat{X}_{m-1}-X_{s_{m-1}}]|^p]^{1/p}=O_p(\ell_n^{-1}b_n^\varepsilon)$ and $\sup_m E_\Pi[|\hat{X}_{m-1}-X_{s_{m-1}}|^q]^{1/q}=O_p(\ell_n^{-1/2}b_n^\varepsilon)$. Similar estimates for $\partial_\sigma^j\hat{\Psi}_{2,n}$ and Sobolev's inequality yield $E_\Pi[\sup_\sigma|b_n^{-1/2}\times\partial_\sigma^j\hat{\Psi}_{2,n}|^q]=O_p(b_n^{-3q/2+\varepsilon}k_n^{2q})$ for $0\le j\le 3$, $q>0$ and any $\varepsilon>0$. If (B4) is satisfied, we obtain $E_\Pi[\sup_\sigma|b_n^{-1/2}\partial_\sigma^j\hat{\Psi}_{2,n}|^q]=\bar{R}_n(b_n^{-3q/2}k_n^{2q})$.

Similarly, we have $E_\Pi[\sup_\sigma|b_n^{-1/2}\partial_\sigma^j\hat{\Psi}_{3,n}|^q]=O_p(b_n^{-q+\varepsilon}k_n^q)$, and therefore we obtain $b_n^{-1/2}\sup_\sigma|\partial_\sigma^j(H_n(\sigma,v_*)-H_n(\sigma_*,v_*))-\partial_\sigma^j(\tilde{H}_n(\sigma,v_*)-\tilde{H}_n(\sigma_*,v_*))|\to^p 0$ for $0\le j\le 3$. If further (B4) is satisfied, we have

$$\begin{aligned}
&E_\Pi\Big[\Big(b_n^{-1/2}\sup_\sigma\big|\partial_\sigma^j\big(H_n(\sigma,v_*)-H_n(\sigma_*,v_*)\big)-\partial_\sigma^j\big(\tilde{H}_n(\sigma,v_*)-\tilde{H}_n(\sigma_*,v_*)\big)\big|\Big)^q\Big]\\
&\quad=\bar{R}_n\big(\big(b_n^{-5}k_n^7\big)^{q/4}\big)
\end{aligned}$$

for $0\le j\le 3$.

Taylor's formula yields

$$\begin{aligned}
&H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n)-\big(H_n(\sigma,v_*)-H_n(\sigma_*,v_*)\big)\\
&\quad=-\frac{1}{2}\sum_m\bigg\{Z_m^\top\big(\tilde{\mathbf{C}}(\hat{X}_{m-1},\hat{v}_n)-\tilde{\mathbf{C}}(\hat{X}_{m-1},v_*)\big)Z_m+\log\frac{\det S_m(\sigma,\hat{v}_n)}{\det S_m(\sigma,v_*)}-\log\frac{\det S_m(\sigma_*,\hat{v}_n)}{\det S_m(\sigma_*,v_*)}\bigg\}\\
&\quad=-\frac{1}{2}\sum_m\sum_{j=1}^3\bigg\{Z_m^\top\partial_v^j\tilde{\mathbf{C}}(\hat{X}_{m-1},v_*)Z_m+\partial_v^j\log\frac{\det S_m(\sigma,v_*)}{\det S_m(\sigma_*,v_*)}\bigg\}\frac{(\hat{v}_n-v_*)^j}{j!}\\
&\qquad-\frac{1}{2}\sum_m\int_0^1\frac{(1-s)^3}{6}\bigg\{Z_m^\top\partial_v^4\tilde{\mathbf{C}}(\hat{X}_{m-1},v_s)Z_m+\partial_v^4\log\frac{\det S_m(\sigma,v_s)}{\det S_m(\sigma_*,v_s)}\bigg\}ds(\hat{v}_n-v_*)^4,
\end{aligned}\tag{4.7}$$

where $v_s=s\hat{v}_n+(1-s)v_*$.

Then we obtain $b_n^{-1/4}|H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n)-(H_n(\sigma,v_*)-H_n(\sigma_*,v_*))| \to^p 0$, by Lemma 4.3, (V), and the equation $\partial_v \log\det S_m = \mathrm{tr}(\partial_v S_m S_m^{-1})$. Similarly, we obtain $b_n^{-1/4} \times |\partial_\sigma^j H_n(\sigma,\hat{v}_n)-\partial_\sigma^j H_n(\sigma,v_*)| \to^p 0$ for $1\le j\le 4$. Sobolev's inequality yields $\sup_\sigma(b_n^{-1/4} \times |\partial_\sigma^j(H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n))-\partial_\sigma^j(H_n(\sigma,v_*)-H_n(\sigma_*,v_*))|) \to^p 0$ for $0\le j\le 3$ and consequently we obtain $\sup_\sigma(b_n^{-1/2}|\partial_\sigma^j(H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n))-\partial_\sigma^j(\tilde{H}_n(\sigma,v_*)-\tilde{H}_n(\sigma_*,v_*))|) \to^p 0$ for $0\le j\le 3$.

Moreover, points 1 and 3 of Lemma 4.3 yield $E_\Pi[|b_n^{-1/4}\partial_\sigma\hat{\Psi}_{1,n}(\sigma_*)|^2]=\bar{R}_n(b_n^{-3/2}k_n^2)\to^p 0$, $E_\Pi[|b_n^{-1/4}\partial_\sigma\hat{\Psi}_{2,n}(\sigma_*)|^2]=O_p(b_n^{-3/4}k_n b_n^\varepsilon)\to^p 0$ for sufficiently small $\varepsilon>0$, and consequently $b_n^{-1/4}(\partial_\sigma H_n(\sigma_*,\hat{v}_n)-\partial_\sigma\tilde{H}_n(\sigma_*,v_*))\to^p 0$.

If further (B4) is satisfied, then for any $q>0$, (4.7) yields

$$\begin{aligned}&\sup_\sigma E_\Pi\big[b_n^{-q/2}\big|H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n)-\big(H_n(\sigma,v_*)-H_n(\sigma_*,v_*)\big)\big|^q\big]\\&\quad=\bar{R}_n\big(b_n^{-q/2}b_n^{-q}k_n^{2q}\ell_n^q b_n^{-q/2}\big)+\bar{R}_n\big(b_n^{-q/2}(b_nk_n\ell_n)^q b_n^{-2q}\big)=\bar{R}_n\big(b_n^{-q}k_n^q\big),\end{aligned}$$

by Lemma 4.3, (V), and the equation $\partial_v\log\det S_m=\mathrm{tr}(\partial_v S_m S_m^{-1})$. Similarly, we obtain $\sup_\sigma E_\Pi[b_n^{-q/2}|\partial_\sigma^j H_n(\sigma,\hat{v}_n)-\partial_\sigma^j H_n(\sigma,v_*)|^q]=\bar{R}_n(b_n^{-q}k_n^q)$ for $1\le j\le 4$. Sobolev's inequality yields $E_\Pi[\sup_\sigma(b_n^{-1/2}|\partial_\sigma^j(H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n))-\partial_\sigma^j(H_n(\sigma,v_*)-H_n(\sigma_*,v_*))|)^q]=\bar{R}_n(b_n^{-q}k_n^q)$ for $0\le j\le 3$, which completes the proof. □

5. The limit of the quasi-likelihood function

We complete the proof of Proposition 2.1 in this section. To do so, it is essential to specify the asymptotic behavior of some functions of approximate covariance matrix $\tilde{S}_m$, as seen in (5.2). Unlike previous studies by Gloter and Jacod [12,13], the eigenvalues of the diagonal blocks $\tilde{D}_{1,m}$ and $\tilde{D}_{2,m}$ of $\tilde{S}_m$ are not identified because of the irregular sampling, and even the sizes of $\tilde{D}_{1,m}$ and $\tilde{D}_{2,m}$ are different. These problems make it difficult to deduce asymptotic behaviors of the right-hand side of (5.2). To solve these problems, in Lemma 5.1, we approximate $\tilde{D}_{j,m}$ by $\dot{D}_{j,m}$, which is a kind of local averaged versions of $\tilde{D}_{j,m}$ and has similar properties to the covariance matrix of equidistant sampling scheme. Moreover, we can also change the sizes of $\tilde{D}_{j,m}$ using some specific properties of $\dot{D}_{j,m}$. We deal with this in Lemma 5.2, and show convergence of some trace functions that appear in a decomposition of H_n. The decomposition (4.4) and the nice properties of p_i in Lemma 4.1 are essential in the proofs.

Lemma 4.4 yields

$$\begin{aligned}&b_n^{-1/2}\partial_\sigma^j\big(H_n(\sigma,\hat{v}_n)-H_n(\sigma_*,\hat{v}_n)\big)\\&\quad=b_n^{-1/2}\partial_\sigma^j\big(\tilde{H}_n(\sigma,v_*)-\tilde{H}_n(\sigma_*,v_*)\big)+o_p(1)\\&\quad=-\frac{1}{2}b_n^{-\frac{1}{2}}\sum_m\bigg(E_m\big[\tilde{Z}_m^\top\partial_\sigma^j\big(\tilde{S}_m^{-1}-\tilde{S}_{m,*}^{-1}\big)\tilde{Z}_m\big]+\partial_\sigma^j\log\frac{\det\tilde{S}_m}{\det\tilde{S}_{m,*}}\bigg)\\&\qquad-\frac{1}{2}b_n^{-\frac{1}{2}}\sum_m\bar{E}_m\big[\tilde{Z}_m^\top\partial_\sigma^j\big(\tilde{S}_m^{-1}-\tilde{S}_{m,*}^{-1}\big)\tilde{Z}_m\big]+o_p(1)\end{aligned}\tag{5.1}$$

for $0 \le j \le 3$. Together with the relation $E_m[\tilde{Z}_m^\top \partial_\sigma^j (\tilde{S}_m^{-1} - \tilde{S}_{m,*}^{-1})\tilde{Z}_m] = \mathrm{tr}(\partial_\sigma^j (\tilde{S}_m^{-1} - \tilde{S}_{m,*}^{-1})\tilde{S}_{m,*})$, we obtain

$$\begin{aligned} & b_n^{-\frac{1}{2}} \partial_\sigma^j \bigl(H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)\bigr) \\ & \quad = -\frac{1}{2} b_n^{-\frac{1}{2}} \sum_m \partial_\sigma^j \left(\mathrm{tr}\bigl(\tilde{S}_m^{-1}\tilde{S}_{m,*} - \mathcal{E}\bigr) + \log \frac{\det \tilde{S}_m}{\det \tilde{S}_{m,*}} \right) + o_p(1), \end{aligned} \tag{5.2}$$

since the residual terms are $o_p(1)$ by Lemma 4.3.

We first investigate asymptotics of $\mathrm{tr}(\tilde{S}_m^{-1}\tilde{S}_{m,*} - \mathcal{E})$. Let $\tilde{L} = \{\tilde{b}_m^1 \cdot \tilde{b}_m^2 |I_{i,m}^1 \cap I_{j,m}^2|\}_{i,j}$ and $\tilde{G} = \{|I_{i,m}^1 \cap I_{j,m}^2|\}_{i,j}$. Then since

$$\begin{aligned} \tilde{S}_m^{-1} &= \tilde{D}_m^{-1/2} \sum_{p=0}^\infty (-1)^p \begin{pmatrix} 0 & \tilde{D}_{1,m}^{-1/2}\tilde{L}\tilde{D}_{2,m}^{-1/2} \\ \tilde{D}_{2,m}^{-1/2}\tilde{L}^\top \tilde{D}_{1,m}^{-1/2} & 0 \end{pmatrix}^p \tilde{D}_m^{-1/2} \\ &= \sum_{p=0}^\infty \begin{pmatrix} (\tilde{D}_{1,m}^{-1}\tilde{L}\tilde{D}_{2,m}^{-1}\tilde{L}^\top)^p \tilde{D}_{1,m}^{-1} & -\tilde{D}_{1,m}^{-1}\tilde{L}(\tilde{D}_{2,m}^{-1}\tilde{L}^\top \tilde{D}_{1,m}^{-1}\tilde{L})^p \tilde{D}_{2,m}^{-1} \\ -\tilde{D}_{2,m}^{-1}\tilde{L}^\top (\tilde{D}_{1,m}^{-1}\tilde{L}\tilde{D}_{2,m}^{-1}\tilde{L}^\top)^p \tilde{D}_{1,m}^{-1} & (\tilde{D}_{2,m}^{-1}\tilde{L}^\top \tilde{D}_{1,m}^{-1}\tilde{L})^p \tilde{D}_{2,m}^{-1} \end{pmatrix}, \end{aligned}$$

we have

$$\begin{aligned} \mathrm{tr}\bigl(\tilde{S}_m^{-1}\tilde{S}_{m,*} - \mathcal{E}\bigr) = \sum_{p=0}^\infty \bigl\{ & \bigl(|\tilde{b}_{m,*}^1|^2 - |\tilde{b}_m^1|^2\bigr) \mathrm{tr}\bigl((\tilde{D}_{1,m}^{-1}\tilde{L}\tilde{D}_{2,m}^{-1}\tilde{L}^\top)^p \tilde{D}_{1,m}^{-1} D'_{1,m}\bigr) \\ & + \bigl(|\tilde{b}_{m,*}^2|^2 - |\tilde{b}_m^2|^2\bigr) \mathrm{tr}\bigl((\tilde{D}_{2,m}^{-1}\tilde{L}^\top \tilde{D}_{1,m}^{-1}\tilde{L})^p \tilde{D}_{2,m}^{-1} D'_{2,m}\bigr) \\ & - 2\bigl(\tilde{b}_{m,*}^1 \cdot \tilde{b}_{m,*}^2 - \tilde{b}_m^1 \cdot \tilde{b}_m^2\bigr) \mathrm{tr}\bigl(\tilde{D}_{1,m}^{-1}\tilde{L}\tilde{D}_{2,m}^{-1}(\tilde{L}^\top \tilde{D}_{1,m}^{-1}\tilde{L}\tilde{D}_{2,m}^{-1})^p \tilde{G}^\top\bigr)\bigr\}. \end{aligned} \tag{5.3}$$

Note that $\|\{|I_{i,m}^1 \cap I_{j,m}^2||I_{i,m}^1|^{-1/2}|I_{j,m}^2|^{-1/2}\}_{ij}\| \le 1$ by Lemma 2 in [25].

We will see the limit of each term on the right-hand side of (5.3). Let $\hat{a}_m^j = a_{s_{m-1}}^j$ and $\dot{D}_{j,m} = |\tilde{b}_m^j|^2 b_n^{-1}(\hat{a}_m^j)^{-1}\mathcal{E} + v_{j,*}M_{j,m}$. It is difficult to calculate the eigenvalues of $\tilde{D}_{j,m}$. However, we can easily calculate the eigenvalues of $\dot{D}_{j,m}$ as seen in Section 4.2, and it has nice properties which are useful when we deduce the limit. Moreover, we can replace $\tilde{D}_{j,m}^{-1}$ by $\dot{D}_{j,m}^{-1}$ using the following lemma.

Lemma 5.1. *Let $j \in \{1,2\}$ and $A_{n,m}$ be a $k_m^j \times k_m^j$ matrix for $1 \le m \le \ell_n$. Assume (B1), (A2) and that all elements of $A_{n,m}$ are nonnegative and $\|A_{n,m}\| \le 1$ for any m. Then*

$$\mathrm{tr}\bigl(\partial_\sigma^k \tilde{D}_{j,m}^{-1} A_{n,m}\bigr) = \mathrm{tr}\bigl(\partial_\sigma^k \dot{D}_{j,m}^{-1} A_{n,m}\bigr) + \dot{R}_n\bigl(b_n^{3/2}\ell_n^{-1}\bigr)$$

for $0 \le k \le 3$. If further (B2) is satisfied, then

$$\sup_\sigma \bigl|\mathrm{tr}\bigl(\partial_\sigma^k \tilde{D}_{j,m}^{-1} A_{n,m}\bigr) - \mathrm{tr}\bigl(\partial_\sigma^k \dot{D}_{j,m}^{-1} A_{n,m}\bigr)\bigr| = \underline{R}_n\bigl(b_n^{3/2}\ell_n^{-1}\bigr)$$

for $0 \le k \le 3$.

Proof. We first consider the case where $k=0$. (4.4), (4.6), and the equation above it yield

$$\begin{aligned}
&\mathrm{tr}\big(\tilde{D}_{j,m}^{-1}A_{n,m}\big)\\
&\quad=\sum_{p=0}^{\infty}\mathrm{tr}\big(\check{D}_{j,m}^{-1/2}\big(\check{D}_{j,m}^{-1/2}(\check{D}_{j,m}-\tilde{D}_{j,m})\check{D}_{j,m}^{-1/2}\big)^p\check{D}_{j,m}^{-1/2}A_{n,m}\big)\\
&\quad=\sum_{p=0}^{\infty}\frac{1}{v_{j,*}^{p+1}}\sum_{i_1,\dots,i_{p+1}}\frac{1}{p_{i_1}\cdots p_{i_{p+1}}(|\tilde{b}_m^j|^2 r_n v_{j,*}^{-1})}\\
&\qquad\times\sum_{\substack{l_q\le i_q\wedge i_{q+1}\\ l'\le i_{p+1},l''\le i_1}}\frac{(A_{n,m})_{l',l''}}{P_{l',l'',i_{p+1},i_1}}\prod_{q=1}^{p}\frac{(\check{D}_{j,m}-\tilde{D}_{j,m})_{l_q,l_q}}{P_{l_q,l_q,i_q,i_{q+1}}},
\end{aligned}\tag{5.4}$$

where $P_{k_1,k_2,l_1,l_2}=\prod_{m_1;k_1\le m_1\le l_1-1}p_{m_1}\prod_{m_2;k_2\le m_2\le l_2-1}p_{m_2}(|\tilde{b}_m^j|^2 r_n v_{j,*}^{-1})$.

Then the nice properties of p_i in Lemma 4.1 will lead us to the desired results. Roughly speaking, we have $1\le p_i(|\tilde{b}_m^j|^2 r_n v_{j,*}^{-1})\sim 1+Cb_n^{-1/2}$ for sufficiently large i. This means that P_{k_1,k_2,l_1,l_2} and P_{k_1',k_2',l_1,l_2} are asymptotically equivalent if $|k_1-k_1'|$ and $|k_2-k_2'|$ are of order less than $b_n^{1/2}$. Then we can replace $\tilde{D}_{j,m}$ in the right-hand side of (5.4) by $\dot{D}_{j,m}$ since the diagonal elements of both matrices have the same local average. We will verify these rough sketches by the following.

We first see that terms containing small l_q can be ignored. Let η be the one in (A2), $\eta'\in(\eta,1/2)$, $t_{\tilde{l},m}=s_{m-1}+T\ell_n^{-1}b_n^{\eta'}/k_n+T\ell_n^{-1}(k_n-b_n^{\eta'})[(k_n-b_n^{\eta'})b_n^{-\eta}]^{-1}\tilde{l}/k_n$ for $0\le\tilde{l}\le[(k_n-b_n^{\eta'})b_n^{-\eta}]$, $\mathcal{I}_m(\tilde{l})=\{l;I_{l,m}^j\subset[t_{\tilde{l}-1,m},t_{\tilde{l},m})\}$, and $\mathcal{E}'=\{\delta_{i_1i_2}1_{\{\inf I_{i_1,m}^j<t_0\}}\}_{1\le i_1,i_2\le k_m^j}$. Then the absolute value of summation involving terms with l_q satisfying $\inf I_{l_q,m}^j<t_0$ for some $1\le q\le p$ on the right-hand side of (5.4) is less than

$$\begin{aligned}
&\|A_{n,m}\|\sum_{p=1}^{\infty}p\big\|\check{D}_{j,m}^{-1/2}(\check{D}_{j,m}-\tilde{D}_{j,m})\check{D}_{j,m}^{-1/2}\big\|^{p-1}\big\|\check{D}_{j,m}^{-1/2}\big\|^2\big|\tilde{b}_m^j\big|^2\,\mathrm{tr}\big(\check{D}_{j,m}^{-1/2}(r_n-\underline{r}_n)\mathcal{E}'\check{D}_{j,m}^{-1/2}\big)\\
&\quad\le r_n^2\underline{r}_n^{-2}\,\mathrm{tr}\big(\check{D}_{j,m}^{-1}\mathcal{E}'\big),
\end{aligned}\tag{5.5}$$

by (4.6), Lemma A.1, and the assumptions. Let $J_1=[T\ell_n^{-1}b_n^{\eta'}k_n^{-1}r_n^{-1}]$ and $J_2=[T\ell_n^{-1}b_n^{\eta'}\times k_n^{-1}\underline{r}_n^{-1}]+1$, then we obtain $J_1\le\#\{i;\inf I_{i,m}^j<t_0\}\le J_2$. Therefore, point 2 of Lemma 4.1 ensures that $\mathrm{tr}(\check{D}_{j,m}^{-1}\mathcal{E}')$ is less than

$$\frac{J_2}{[k_m^j/2]-J_2}\,\mathrm{tr}\big(\check{D}_{j,m}^{-1}\big)\le\frac{2J_2}{[k_m^j/2]}\,\mathrm{tr}\big(\check{D}_{j,m}^{-1}\big)$$

if $[k^j_m/2] \geq 2J_2$, and hence the right-hand side of (5.5) is $\bar{R}_n(\mathcal{V}_n)$ by $\mathrm{tr}(\check{D}^{-1}_{j,m}) = \bar{R}_n(r_n^{-1/2}k_n)$, where $\mathcal{V}_n = b_n^{1/2+\eta'}1_{\{[\underline{k}_n/2]\geq 2J_2\}} + b_n^{1/2}k_n 1_{\{[\underline{k}_n/2]<2J_2\}}$.

On the other hand, for $\tilde{l}$, i_q, i_{q+1} and l_q satisfying $l_q \in \mathcal{I}_m(\tilde{l})$ and $\max \mathcal{I}_m(\tilde{l}) \leq i_q \wedge i_{q+1}$, Lemma 4.1 yields that $P_{\max \mathcal{I}_m(\tilde{l}),\max \mathcal{I}_m(\tilde{l}),i_q,i_{q+1}}/P_{l_q,l_q,i_q,i_{q+1}}$ is less than 1 and greater than

$$\left(1 + 1/J_1 + \left|\tilde{b}^j_m\right|^2 r_n v^{-1}_{j,*} J_1\right)^{-Cb_n^{-1+\eta}\underline{r}_n^{-1}} \geq 1 - \bar{R}_n\left(b_n^{\eta-\eta'} + b_n^{-1+\eta'+\eta}\right)$$

for sufficiently large n. Moreover, $\max_{\tilde{l}} |\sum_{l\in\mathcal{I}_m(\tilde{l})}(\tilde{D}_{j,m} - \dot{D}_{j,m})_{l,l}|$ is $\dot{R}_n(b_n^{-1+\eta})$ by (A2). We also have $\sup_{\sigma,m}\max_{\tilde{l}} |\sum_{l\in\mathcal{I}_m(\tilde{l})}(\tilde{D}_{j,m} - \dot{D}_{j,m})_{l,l}| = \underline{\mathrm{R}}_n(b_n^{-1+\eta})$ if (B2) is satisfied.

Therefore, we obtain

$$\begin{aligned}
&\mathrm{tr}\left(\tilde{D}^{-1}_{j,m}A_{n,m}\right) \\
&\quad= \sum_{p=0}^{\infty}\frac{1}{v^{p+1}_{j,*}}\sum_{i_1,\ldots,i_{p+1}}\frac{1}{p_{i_1}\cdots p_{i_{p+1}}(|\tilde{b}^j_m|^2 r_n v^{-1}_{j,*})} \\
&\qquad\times \sum_{\substack{l_q\leq i_q\wedge i_{q+1},\inf I^j_{l_q,m}\geq t_0\\ l'\leq i_{p+1},l''\leq i_1}}\frac{(A_{n,m})_{l',l''}}{P_{l',l'',i_{p+1},i_1}}\prod_{q=1}^{p}\frac{(\check{D}_{j,m} - \tilde{D}_{j,m})_{l_q,l_q}}{P_{l_q,l_q,i_q,i_{q+1}}} \\
&\qquad+ \bar{R}_n(\mathcal{V}_n) \\
&\quad= \sum_{p=0}^{\infty}\frac{\mathcal{T}^{n,p}_{m,1}}{v^{p+1}_{j,*}}\sum_{i_1,\ldots,i_{p+1}}\frac{1}{p_{i_1}\cdots p_{i_{p+1}}(|\tilde{b}^j_m|^2 r_n v^{-1}_{j,*})}\sum_{l'\leq i_{p+1},l''\leq i_1}\frac{(A_{n,m})_{l',l''}}{P_{l',l'',i_{p+1},i_1}} \\
&\qquad\times\prod_{q=1}^{p}\sum_{1\leq\tilde{l}_q\leq[(k_n-b_n^{\eta'})b_n^{-\eta}]}\frac{\sum_{l_q\in\mathcal{I}_m(\tilde{l}_q)}(\check{D}_{j,m} - \dot{D}_{j,m} + \mathcal{T}^{n,p}_{m,3}\mathcal{E})_{l_q,l_q}}{P_{\max\mathcal{I}_m(\tilde{l}_q),\max\mathcal{I}_m(\tilde{l}_q),i_q,i_{q+1}}} + \bar{R}_n(\mathcal{V}_n) \\
&\quad= \sum_{p=0}^{\infty}\mathcal{T}^{n,p}_{m,1}\left(\mathcal{T}^{n,p}_{m,2}\right)^{-1}\mathrm{tr}\left(\check{D}^{-1/2}_{j,m}\left(\check{D}^{-1/2}_{j,m}\left(\check{D}_{j,m} - \dot{D}_{j,m} + \mathcal{T}^{n,p}_{m,3}\mathcal{E}\right)\check{D}^{-1/2}_{j,m}\right)^p\check{D}^{-1/2}_{j,m}A_{n,m}\right) \\
&\qquad+ \dot{R}_n\left(b_n^{3/2}\ell_n^{-1}\right),
\end{aligned} \tag{5.6}$$

where $\mathcal{T}^{n,p}_{m,i}$ is a random variable which does not depend on $l_q, \tilde{l}_q, i_q, i_{q+1}$ and satisfies

$$\left(1 - \bar{R}_n\left(b_n^{\eta-\eta'} + b_n^{-1+\eta'+\eta}\right)\right)^p \leq \mathcal{T}^{n,p}_{m,i} \leq 1$$

for $i = 1, 2$ and $\sup_p |\mathcal{T}^{n,p}_{m,3}| = \dot{R}_n(b_n^{-1})$.

Let $F_p(t) = \mathrm{tr}(\check{D}_{j,m}^{-1/2}(\check{D}_{j,m}^{-1/2}(\check{D}_{j,m} - \dot{D}_{j,m} + t\mathcal{T}_{m,3}^{n,p}\mathcal{E})\check{D}_{j,m}^{-1/2})^p \check{D}_{j,m}^{-1/2} A_{n,m})$. Then

$$\begin{aligned}|F_p(1) - F_p(0)| &\leq \int_0^1 |F_p'(t)|\,dt \\ &\leq p|\tilde{b}_m^j|^{-2}|\mathcal{T}_{m,3}^{n,p}|(1 - b_n^{-1}(\hat{a}_m^j)^{-1}r_n^{-1} + |\tilde{b}_m^j|^{-2}\mathcal{T}_{m,3}^{n,p}r_n^{-1})^{p-1}\bar{R}_n(r_n^{-1/2}b_n^2\ell_n^{-1}),\end{aligned}$$

and hence $\sum_{p=0}^\infty |\mathcal{T}_{m,1}^{n,p}(\mathcal{T}_{m,2}^{n,p})^{-1}||F_p(1) - F_p(0)| \leq \sup_p |\mathcal{T}_{m,3}^{n,p}| \cdot \dot{R}_n(b_n^{3/2}k_n) = \dot{R}_n(b_n^{3/2}\ell_n^{-1})$. Therefore, we obtain the desired conclusion by

$$\begin{aligned}&\sum_{p=0}^\infty |\mathcal{T}_{m,1}^{n,p}(\mathcal{T}_{m,2}^{n,p})^{-1} - 1|F_p(0) \\ &\quad\leq C\sum_p p\bar{R}_n(b_n^{\eta-\eta'} + b_n^{-1+\eta+\eta'})(1 + \bar{R}_n(b_n^{\eta-\eta'} + b_n^{-1+\eta+\eta'}))^{p-1} \\ &\qquad\times (1 - b_n^{-1}(\hat{a}_m^j)^{-1}r_n^{-1})^p r_n^{-1/2}k_m^j \\ &\quad= \dot{R}_n(b_n^{3/2}\ell_n^{-1}).\end{aligned}$$

For the case $k=1$, we have $\mathrm{tr}(\partial_\sigma \tilde{D}_{j,m}^{-1}A_{n,m}) = -\mathrm{tr}(\dot{D}_{j,m}^{-1}\partial_\sigma \tilde{D}_{j,m}\dot{D}_{j,m}^{-1}A_{n,m}) + \dot{R}_n(b_n^{3/2}\ell_n^{-1})$ by using the result for $k=0$, $\partial_\sigma \tilde{D}_{j,m}^{-1} = -\tilde{D}_{j,m}^{-1}\partial_\sigma \tilde{D}_{j,m}\tilde{D}_{j,m}^{-1}$, $\|\tilde{D}_{j,m}^{-1}\partial_\sigma \tilde{D}_{j,m}\| = \bar{R}_n(1)$ and all elements of $\tilde{D}_{j,m}^{-1}$ are nonnegative by a similar argument to (4.4). Then a similar argument to (5.6) enables us to replace $\partial_\sigma \tilde{D}_{j,m}$ by $\partial_\sigma \dot{D}_{j,m}$. Similarly, we obtain $\mathrm{tr}(\partial_\sigma^k \tilde{D}_{j,m}^{-1}A_{n,m}) = \mathrm{tr}(\partial_\sigma^k \dot{D}_{j,m}^{-1}A_{n,m}) + \dot{R}_n(b_n^{3/2}\ell_n^{-1})$ for $k = 2, 3$.

If further (B2) is satisfied, then similarly we have $\sup_\sigma |\mathrm{tr}(\partial_\sigma^k \tilde{D}_{j,m}^{-1}A_{n,m}) - \mathrm{tr}(\partial_\sigma^k \dot{D}_{j,m}^{-1}A_{n,m})| = \underline{\mathrm{R}}_n(b_n^{3/2}\ell_n^{-1})$. □

Remark 5.1. The proof shows that there are upperbounds of the absolute values of residual terms in the statement of Lemma 5.1 which do not depend on $A_{n,m}$.

Let $c_j = |\tilde{b}_m^j|^2 b_n^{-1}(\hat{a}_m^j)^{-1}/v_{j,*}$ and $c_j' = c_j(\hat{a}_m^j/\hat{a}_m^{3-j})^2$ for $j = 1, 2$.

Lemma 5.2. *Assume (B1) and (A2). Then*

$$\begin{aligned}&\sup_{\sigma,m}\left|\partial_\sigma^k \mathrm{tr}((\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top)^p\tilde{D}_{1,m}^{-1}D_{1,m}') - \frac{T\hat{a}_m^1 k_n}{\pi}\frac{(\hat{a}_m^2)^p\partial_\sigma^k I_{p+1,p}(c_1,c_2')}{b_n^{2p+1}(\hat{a}_m^1)^{3p+1}v_{1,*}^{p+1}v_{2,*}^p}\right| \\ &\quad= o_p(b_n^{1/2}\ell_n^{-1}),\end{aligned} \tag{5.7}$$

$$\begin{aligned}&\sup_{\sigma,m}\left|\partial_\sigma^k \mathrm{tr}((\tilde{D}_{2,m}^{-1}\tilde{G}^\top\tilde{D}_{1,m}^{-1}\tilde{G})^p\tilde{D}_{2,m}^{-1}D_{2,m}') - \frac{T\hat{a}_m^1 k_n}{\pi}\frac{(\hat{a}_m^2)^{p+1}\partial_\sigma^k I_{p,p+1}(c_1,c_2')}{b_n^{2p+1}(\hat{a}_m^1)^{3p+2}v_{1,*}^p v_{2,*}^{p+1}}\right| \\ &\quad= o_p(b_n^{1/2}\ell_n^{-1}),\end{aligned} \tag{5.8}$$

$$\sup_{\sigma,m}\left|\partial_\sigma^k \operatorname{tr}\big(\big(\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top\big)^{p+1}\big) - \frac{T\hat{a}_m^1 k_n}{\pi}\frac{(\hat{a}_m^2)^{p+1}\partial_\sigma^k I_{p+1,p+1}(c_1,c_2')}{(b_n^2(\hat{a}_m^1)^3 v_{1,*}v_{2,*})^{p+1}}\right| \\ = o_p\big(b_n^{1/2}\ell_n^{-1}\big) \tag{5.9}$$

for $0 \le k \le 3$ *and* $p \in \mathbb{Z}_+$. *If further* (B2) *is satisfied, then* $o_p(b_n^{1/2}\ell_n^{-1})$ *in* (5.7)–(5.9) *can be replaced by* $\underline{R}_n(b_n^{1/2}\ell_n^{-1})$.

Proof. For any $p \in \mathbb{N}$, Lemma 5.1 yields

$$b_n^{-1/2}\partial_\sigma^k \operatorname{tr}\big(\big(\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top\big)^p\big) = b_n^{-1/2}\partial_\sigma^k \operatorname{tr}\big(\big(\dot{D}_{1,m}^{-1}\tilde{G}\dot{D}_{2,m}^{-1}\tilde{G}^\top\big)^p\big) + \dot{R}_n\big(\ell_n^{-1}\big). \tag{5.10}$$

Moreover, we have

$$b_n^{-1/2}\operatorname{tr}\big(\big(\dot{D}_{1,m}^{-1}\tilde{G}\dot{D}_{2,m}^{-1}\tilde{G}^\top\big)^p\big) \\ = b_n^{-1/2}\frac{1}{v_{1,*}^p v_{2,*}^p}\sum_{\substack{i_1,\dots,i_p \\ j_1,\dots,j_p}}\ \sum_{\substack{\alpha_{2q-1}\le i_q,\beta_{2q-1}\le j_q \\ \alpha_{2q}\le i_{q+1},\beta_{2q}\le j_q (1\le q\le p)}}\ \prod_{q=1}^p \frac{\tilde{G}_{\alpha_{2q-1},\beta_{2q-1}}}{\tilde{P}_{\alpha_{2q-1},\beta_{2q-1},i_q+1,j_q+1}}\frac{\tilde{G}_{\alpha_{2q},\beta_{2q}}}{\tilde{P}_{\alpha_{2q},\beta_{2q},i_{q+1},j_q}} \tag{5.11}$$

by (4.4), where $i_{p+1} = i_1$ and $\tilde{P}_{\alpha,\beta,i,j} = \prod_{k_1=\alpha}^{i-1} p_{k_1}(c_1)\prod_{k_2=\beta}^{j-1} p_{k_2}(c_2)$.

We will apply (4.3) to obtain the limit of the traces. To do so, we need to change the size of matrices $\tilde{G}$ and $\dot{D}_{2,m}^{-1}$. This is again achieved by the nice properties of p_i. The essential idea is that point 3 of Lemma 4.1 ensures $p_i \approx p_+$ for sufficiently large i, and therefore $\tilde{P}_{\alpha,\beta,i,j} \approx p_+(c_1)^{i-\alpha}p_+(c_2)^{j-\beta} \approx \exp(\sqrt{c_1}(i-\alpha)+\sqrt{c_2}(j-\beta)) \approx \acute{P}_{k\alpha,k\beta,ki,kj}$, where $k \in \mathbb{N}$ and $\acute{P}_{\alpha',\beta',i',j'} = \prod_{k_1=\alpha'}^{i'-1} p_{k_1}(c_1/k^2)\prod_{k_2=\beta'}^{j'-1} p_{k_2}(c_2/k^2)$. The size of $\dot{D}_{2,m}^{-1}$ decides the ranges of summation of $j_1,\dots,j_p$ in (5.11). By changing these ranges using the above relation on $\tilde{P}_{\alpha,\beta,i,j}$ and $\acute{P}_{k\alpha,k\beta,ki,kj}$, we can change the size of matrices $\tilde{G}$ and $\dot{D}_{2,m}^{-1}$.

Now we verify the above idea. First, we see that the terms involving small α_q or β_q in (5.11) can be ignored. Let $\eta \in (0,1/2)$ be the one in (A2), $\delta \in (1/2,1)$ such that $b_n^{\delta+\varepsilon}k_n^{-1} \to 0$ for some $\varepsilon > 0$, $\tilde{s}_{l'} = s_{m-1} + T\ell_n^{-1}[k_n b_n^{-\eta}]^{-1}((l' + [b_n^{\delta-\eta}]) \wedge [k_n b_n^{-\eta}])$ for $0 \le l' \le ([k_n b_n^{-\eta}] - [b_n^{\delta-\eta}]) \vee 0$, $\dot{D}_{3,m} = (c_1 \wedge c_2)(v_{1,*}\wedge v_{2,*})\mathcal{E}_{k_m^1\vee k_m^2} + (v_{1,*}\wedge v_{2,*})M(k_m^1 \vee k_m^2)$, $G' = \{|I_{i,m}^1 \cap I_{j,m}^2|1_{\{\inf I_{i,m}^1\wedge \inf I_{j,m}^2 < \tilde{s}_0\}}\}_{1\le i,j\le k_m^1\vee k_m^2}$, $\hat{G} = \{|I_{i,m}^1 \cap I_{j,m}^2|1_{\{i\le k_m^1 \text{ and } j\le k_m^2\}}\}_{1\le i,j\le k_m^1\vee k_m^2}$, and $\mathcal{E}'' = \{\delta_{ij}1_{\{\inf I_{i,m}^1\wedge \inf I_{i,m}^2 < \tilde{s}_0\}}\}_{1\le i,j\le k_m^1\vee k_m^2}$. Similarly to the proof of Lemma 5.1, the absolute value Λ_1 of a summation involving the terms with (α_q,β_q) satisfying $\inf I_{\alpha_q,m}^1 \wedge \inf I_{\beta_q,m}^2 < \tilde{s}_0$ is less than $pb_n^{-1/2}\operatorname{tr}(\dot{D}_{3,m}^{-1}(G'\dot{D}_{3,m}^{-1}\hat{G}^\top + \hat{G}\dot{D}_{3,m}^{-1}(G')^\top)(\dot{D}_{3,m}^{-1}\hat{G}\dot{D}_{3,m}^{-1}\hat{G}^\top)^{p-1})$. Lemma 3 in [25] implies $\|G' + (G')^\top\| \le 2r_n$ and hence all the eigenvalues of $G' + (G')^\top$ are greater than or equal to $-2r_n$. Therefore, $G' + (G')^\top + 2r_n\mathcal{E}''$ is nonnegative definite, and hence Lemma A.1 yields

$$\Lambda_1 \le 2p(r_n/\underline{r}_n)^{2p-1}b_n^{-1/2}\operatorname{tr}\big(\dot{D}_{3,m}^{-1}\big(G' + \big(G'\big)^\top + 2r_n\mathcal{E}''\big)\big) \\ \le 4p(r_n/\underline{r}_n)^{2p-1}b_n^{-1/2}\big(\operatorname{tr}\big(\dot{D}_{3,m}^{-1}G'\big) + r_n\operatorname{tr}\big(\dot{D}_{3,m}^{-1}\mathcal{E}''\big)\big).$$

Let $(\dot{G})_{i,j}=(\sum_{l\le i}G'_{l,i}+\sum_{m<i}G'_{i,m})\delta_{ij}$ and $\dot{k}=\max\{i;\dot{G}_{ii}>0\}$. Then Lemma 4.1 yields

$$\begin{aligned}\operatorname{tr}\big(\dot{D}_{3,m}^{-1}G'\big)&=\frac{1}{v_{1,*}\wedge v_{2,*}}\sum_i\sum_{\alpha,\beta\le i}\frac{G'_{\alpha,\beta}}{p_\alpha\cdots p_i p_\beta\cdots p_{i-1}}\\&\le\frac{1}{v_{1,*}\wedge v_{2,*}}\sum_i\sum_{\alpha\le i}\frac{\dot{G}_{\alpha,\alpha}}{p_\alpha\cdots p_i p_\alpha\cdots p_{i-1}}\\&=\operatorname{tr}\big(\dot{D}_{3,m}^{-1}\dot{G}\big)\le\big(\dot{D}_{3,m}^{-1}\big)_{\dot{k},\dot{k}}(\tilde{s}_0-s_{m-1}+r_n)\le\frac{Cb_n^{-1+\delta}+r_n}{[(k_m^1\vee k_m^2)/2]-\dot{k}}\operatorname{tr}\big(\dot{D}_{3,m}^{-1}\big).\end{aligned}$$

Therefore, we obtain $\Lambda_1=\dot{R}_n(\ell_n^{-1})$, and $\Lambda_1=\underline{\mathrm{R}}_n(\ell_n^{-1})$ if (B2) is satisfied.

Let $\ddot{D}_{2,m}=v_{2,*}c_2(\hat{a}_m^2/\hat{a}_m^1)^2\mathcal{E}+v_{2,*}M_{1,m}$, $i(\alpha')=\min\{i;S_i^{n,1}\ge\tilde{s}_{\alpha'-1}\}$, and $j(\alpha')=\min\{j;S_j^{n,2}\ge\tilde{s}_{\alpha'-1}\}$. We will show that $b_n^{-1/2}\operatorname{tr}((\dot{D}_{1,m}^{-1}\tilde{G}\dot{D}_{2,m}^{-1}\tilde{G}^\top)^p)$ is approximated by $b_n^{-1/2-2p}(\hat{a}_m^2)^p(\hat{a}_m^1)^{-3p}\operatorname{tr}((\dot{D}_{1,m}^{-1}\ddot{D}_{2,m})^p)$. First, a similar argument to the proof of Lemma 5.1 yields $|\tilde{P}_{\alpha,\beta,i,j}/\tilde{P}_{i(\alpha'),j(\alpha'),i(i'),j(j')}-1|=\dot{R}_n(1)$ for $i(\alpha')\le\alpha<i(\alpha'+1)$, $j(\alpha')\le\beta<j(\alpha'+1)$, $i(i')\le i<i(i'+1)$, and $j(j')\le j<j(j'+1)$. Therefore, repeated use of (A2) yields

$$\begin{aligned}&b_n^{-1/2}\operatorname{tr}\big(\big(\dot{D}_{1,m}^{-1}\tilde{G}\dot{D}_{2,m}^{-1}\tilde{G}^\top\big)^p\big)\\&\quad=b_n^{-1/2}\frac{\mathcal{T}_{m,4}^{n,p}}{v_{1,*}^p v_{2,*}^p}\\&\qquad\times\sum_{\substack{i'_1,\dots,i'_p\\ j'_1,\dots,j'_p}}\sum_{\alpha'_1,\dots,\alpha'_{2p}}\prod_{q=1}^p\frac{(Tb_n^{-1+\eta})^2\#\{i_q;I^1_{i_q,m}\subset[\tilde{s}_{i'_q-1},\tilde{s}_{i'_q})\}\#\{j_q;I^2_{j_q,m}\subset[\tilde{s}_{j'_q-1},\tilde{s}_{j'_q})\}}{\tilde{P}_{i(\alpha'_{2q-1}),j(\alpha'_{2q-1}),i(i'_q),j(j'_q)}\tilde{P}_{i(\alpha'_{2q}),j(\alpha'_{2q}),i(i'_{q+1}),j(j'_q)}}\\&\qquad+\dot{R}_n\big(l_n^{-1}\big)\qquad(5.12)\\&\quad=b_n^{-1/2}\frac{\mathcal{T}_{m,5}^{n,p}(Tb_n^{-1+\eta})^{2p}(\hat{a}_m^2)^p}{v_{1,*}^p v_{2,*}^p(\hat{a}_m^1)^p}\\&\qquad\times\sum_{\substack{i'_1,\dots,i'_p\\ j'_1,\dots,j'_p}}\sum_{\alpha'_1,\dots,\alpha'_{2p}}\prod_{q=1}^p\frac{\#\{\alpha;I^1_{\alpha,m}\subset[\tilde{s}_{\alpha'_{2q-1}-1},\tilde{s}_{\alpha'_{2q-1}})\}\#\{\alpha;I^1_{\alpha,m}\subset[\tilde{s}_{\alpha'_{2q}-1},\tilde{s}_{\alpha'_{2q}})\}}{\tilde{P}_{i(\alpha'_{2q-1}),j(\alpha'_{2q-1}),i(i'_q),j(j'_q)}\tilde{P}_{i(\alpha'_{2q}),j(\alpha'_{2q}),i(i'_{q+1}),j(j'_q)}}\\&\qquad+\dot{R}_n\big(l_n^{-1}\big),\end{aligned}$$

where the summations of $\alpha'_1,\dots,\alpha'_{2p}$ are over $1\le\alpha'_{2q-1}\le i'_q\wedge j'_q$ and $1\le\alpha'_{2q}\le i'_{q+1}\wedge j'_q$ for $1\le q\le p$, $i'_{p+1}=i'_1$ and $\mathcal{T}_{m,i}^{n,p}$ is a random variable satisfying $\sup_{\sigma,m}|\mathcal{T}_{m,i}^{n,p}-1|=\dot{R}_n(1)$ for $i=4,5$.

Since Lemma 4.1 and (A2) yield

$$\begin{aligned}
&p_{j(\alpha)}\cdots p_{j(\beta)-1}(c_2)\\
&\quad=\big(p_+(c_2)\big)^{j(\beta)-j(\alpha)}\big(1+\dot{R}_n(1)\big)\\
&\quad=\exp\big(\big(b_n\hat{a}_m^2(\tilde{s}_\beta-\tilde{s}_\alpha)+\dot{R}_n\big(b_n^{1/2}\big)\big)\log p_+(c_2)\big)\big(1+\dot{R}_n(1)\big)\\
&\quad=\exp\big(\hat{a}_m^2\big(\hat{a}_m^1\big)^{-1}\big(i(\beta)-i(\alpha)\big)\big)\log p_+(c_2)\big)\big(1+\dot{R}_n(1)\big)=p_{i(\alpha)}\cdots p_{i(\beta)-1}\big(c_2'\big)\big(1+\dot{R}_n(1)\big),
\end{aligned}$$

we may replace $\tilde{P}_{i(\alpha'_{2q-1}),j(\alpha'_{2q-1}),i(i'_q),j(j'_q)}$ and $\tilde{P}_{i(\alpha'_{2q}),j(\alpha'_{2q}),i(i'_{q+1}),j(j'_q)}$ in the right-hand side of (5.12) by $\hat{P}_{i(\alpha'_{2q-1}),i(\alpha'_{2q-1}),i(i'_q),i(j'_q)}$ and $\hat{P}_{i(\alpha'_{2q}),i(\alpha'_{2q}),i(i'_{q+1}),i(j'_q)}$, respectively, where $\hat{P}_{\alpha,\beta,i,j}=\prod_{k_1=\alpha}^{i-1}p_{k_1}(c_1)\prod_{k_2=\beta}^{j-1}p_{k_2}(c_2')$. Therefore, we obtain

$$\sup_{\sigma,m}\big|b_n^{-1/2}\operatorname{tr}\big(\big(\dot{D}_{1,m}^{-1}\tilde{G}\dot{D}_{2,m}^{-1}\tilde{G}^\top\big)^p\big)-b_n^{-1/2-2p}\big(\hat{a}_m^2\big)^p\big(\hat{a}_m^1\big)^{-3p}\operatorname{tr}\big(\big(\dot{D}_{1,m}^{-1}\ddot{D}_{2,m}^{-1}\big)^p\big)\big|=o_p\big(\ell_n^{-1}\big),$$

by a similar argument to (5.12). Since $\partial_\sigma^l\dot{D}_{j,m}=\partial_\sigma^l c_j v_{j,*}\mathcal{E}$ for $1\le l\le 3$, we similarly obtain

$$\begin{aligned}
&\sup_{\sigma,m}\big|b_n^{-1/2}\partial_\sigma^k\operatorname{tr}\big(\big(\dot{D}_{1,m}^{-1}\tilde{G}\dot{D}_{2,m}^{-1}\tilde{G}^\top\big)^p\big)-b_n^{-1/2-2p}\big(\hat{a}_m^2\big)^p\big(\hat{a}_m^1\big)^{-3p}\partial_\sigma^k\operatorname{tr}\big(\big(\dot{D}_{1,m}^{-1}\ddot{D}_{2,m}^{-1}\big)^p\big)\big|\\
&\quad=o_p\big(\ell_n^{-1}\big).
\end{aligned}\tag{5.13}$$

Then (4.3), (5.10) and (5.13) yield (5.9).

We also have (5.7) and (5.8) by a similar argument.

Similar arguments enable us to replace $o_p(b_n^{1/2}\ell_n^{-1})$ by $\underline{\mathrm{R}}_n(b_n^{1/2}\ell_n^{-1})$ in (5.7)–(5.9) if (B2) is satisfied. □

Remark 5.2. We can also show that the summations of the left-hand side of (5.7)–(5.9) in $1\le p<\infty$ are bounded by $\underline{\mathrm{R}}_n(b_n^{1/2}\ell_n^{-1})\sum_{p=1}^\infty p(\tilde{b}_m^1\cdot\tilde{b}_m^2|\tilde{b}_m^1|^{-1}|\tilde{b}_m^2|^{-1})^{2(p-l)_+}$ for some $l\in\mathbb{N}$ if (B2) is satisfied. To show this, it is sufficient to replace $\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top$ in the traces with $b_n^{-1}\hat{a}_m^2(\hat{a}_m^1)^{-3}\dot{D}_{1,m}^{-1}\ddot{D}_{2,m}^{-1}$ one by one, and estimate residual terms separately on the event $\{b_n^{-1-\varepsilon'}\le\underline{r}_n\le r_n\le b_n^{-1+\varepsilon'}\}$ and on its complement for some small $\varepsilon'>0$. This fact is used in the proof of Proposition 2.2.

Proof of Proposition 2.1. We first prove the results under the additional condition (A1′).

Since $\|\tilde{D}_{1,m}^{-1/2}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top\tilde{D}_{1,m}^{-1/2}\|\le|\tilde{b}_m^1|^{-2}|\tilde{b}_m^2|^{-2}$ by Lemma 3 in [25], for any $\varepsilon,\delta>0$, there exists $P_1\in\mathbb{N}$ such that

$$\sup_n P\left[\sup_\sigma b_n^{-\frac{1}{2}}\sum_m\sum_{p=P+1}^\infty\big|\partial_\sigma^k\big(\big(\tilde{b}_m^1\cdot\tilde{b}_m^2\big)^{2p-1}\operatorname{tr}\big(\big(\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top\big)^p\big)\big)\big|\ge\delta\right]<\varepsilon,\tag{5.14}$$

$$\sup_n P\left[\sup_\sigma b_n^{-\frac{1}{2}}\sum_m\sum_{p=P+1}^\infty\left|\partial_\sigma^k\left(\frac{(\tilde{b}_m^1\cdot\tilde{b}_m^2)^{2p-1}\hat{a}_m^1 k_n(\hat{a}_m^2)^p I_{p,p}(c_1,c_2')}{(b_n^2(\hat{a}_m^1)^3 v_{1,*}v_{2,*})^p}\right)\right|\geq\delta\right]<\varepsilon \quad (5.15)$$

for $P \geq P_1$. Together with Lemma 5.2, we obtain that

$$\sup_\sigma\left|b_n^{-1/2}\partial_\sigma^k\sum_m\sum_{p=1}^\infty(\tilde{b}_m^1\cdot\tilde{b}_m^2)^{2p-1}\operatorname{tr}\left(\left(\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top\right)^p\right)\right.$$
$$\left.-\frac{T\hat{a}_m^1 k_n}{\pi b_n^{1/2}}\partial_\sigma^k\sum_m\sum_{p=1}^\infty\frac{(\tilde{b}_m^1\cdot\tilde{b}_m^2)^{2p-1}(\hat{a}_m^2)^p I_{p,p}(c_1,c_2')}{(b_n^2(\hat{a}_m^1)^3 v_{1,*}v_{2,*})^p}\right|\xrightarrow{p}0.$$

Let $\dot{a}_m^j=\tilde{a}_{s_{m-1}}^j$, $\mathfrak{C}_m=|\tilde{b}_m^1|^2|\tilde{b}_m^2|^2-(\tilde{b}_m^1\cdot\tilde{b}_m^2)^2$,

$$\mathfrak{A}_t=\left(\tilde{a}_t^1\left|b_t^1\right|^2+\tilde{a}_t^2\left|b_t^2\right|^2+2\sqrt{\tilde{a}_t^1\tilde{a}_t^2\det(b_t b_t^\top)}\right)^{1/2},$$

and

$$P_n=\left(c_1+c_2'+2\sqrt{b_n^{-2}\hat{a}_m^2\left(\hat{a}_m^1\right)^{-3}v_{1,*}^{-1}v_{2,*}^{-1}\mathfrak{C}_m}\right)^{1/2}$$
$$=b_n^{-1/2}\left(\hat{a}_m^1\right)^{-1}\left(\dot{a}_m^1\left|\tilde{b}_m^1\right|^2+\dot{a}_m^2\left|\tilde{b}_m^2\right|^2+2\sqrt{\dot{a}_m^1\dot{a}_m^2\mathfrak{C}_m}\right)^{1/2}.$$

Then Lemma A.9 yields

$$\frac{T\hat{a}_m^1 k_n}{\pi b_n^{1/2}}\partial_\sigma^k\sum_m\sum_{p=1}^\infty\frac{(\tilde{b}_m^1\cdot\tilde{b}_m^2)^{2p-1}(\hat{a}_m^2)^p I_{p,p}(c_1,c_2')}{(b_n^2(\hat{a}_m^1)^3 v_{1,*}v_{2,*})^p}$$
$$=\partial_\sigma^k\sum_m\frac{Tb_n^{1/2}\hat{a}_m^1\ell_n^{-1}\frac{\hat{a}_m^2\tilde{b}_m^1\cdot\tilde{b}_m^2}{(\hat{a}_m^1)^3 b_n^2 v_{1,*}v_{2,*}}}{b_n^{-1/2}(\hat{a}_m^1)^{-1}\cdot 2(\dot{a}_m^1|\tilde{b}_m^1|^2+\dot{a}_m^2|\tilde{b}_m^2|^2+2\sqrt{\dot{a}_m^1\dot{a}_m^2\mathfrak{C}_m})^{1/2}\sqrt{\dot{a}_m^1\dot{a}_m^2}\sqrt{\mathfrak{C}_m}b_n^{-1}(\hat{a}_m^1)^{-2}}$$
$$+o_p(1)$$
$$=\partial_\sigma^k\sum_m\frac{T\ell_n^{-1}\sqrt{\dot{a}_m^1\dot{a}_m^2}\tilde{b}_m^1\cdot\tilde{b}_m^2}{2\sqrt{\mathfrak{C}_m}(\dot{a}_m^1|\tilde{b}_m^1|^2+\dot{a}_m^2|\tilde{b}_m^2|^2+2\sqrt{\dot{a}_m^1\dot{a}_m^2\mathfrak{C}_m})^{1/2}}+o_p(1)$$
$$=\partial_\sigma^k\int_0^T\frac{\sqrt{\tilde{a}_t^1\tilde{a}_t^2}b_t^1\cdot b_t^2}{2\sqrt{\det(b_t b_t^\top)}\mathfrak{A}_t}\,dt+o_p(1).$$

Therefore, we have

$$\begin{aligned}\sup_\sigma\bigg|&b_n^{-1/2}\partial_\sigma^k\sum_m\sum_{p=1}^\infty(\tilde{b}_m^1\cdot\tilde{b}_m^2)^{2p-1}\operatorname{tr}\big((\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top)^p\big)\\&-\partial_\sigma^k\int_0^T\frac{\sqrt{\tilde{a}_t^1\tilde{a}_t^2}b_t^1\cdot b_t^2}{2\sqrt{\det(b_tb_t^\top)}\mathfrak{A}_t}\,dt\bigg|\xrightarrow{p}0.\end{aligned}\tag{5.16}$$

Similarly, we obtain

$$\begin{aligned}\sup_\sigma\bigg|&b_n^{-1/2}\partial_\sigma^k\sum_m\sum_{p=0}^\infty\operatorname{tr}\big((\tilde{D}_{1,m}^{-1}\tilde{L}\tilde{D}_{2,m}^{-1}\tilde{L}^\top)^p\tilde{D}_{1,m}^{-1}D'_{1,m}\big)\\&-\partial_\sigma^k\int_0^T\frac{|b_t^2|^2\sqrt{\tilde{a}_t^1\tilde{a}_t^2}+\tilde{a}_t^1\sqrt{\det(b_tb_t^\top)}}{2\sqrt{\det(b_tb_t^\top)}\mathfrak{A}_t}\,dt\bigg|\xrightarrow{p}0,\end{aligned}\tag{5.17}$$

$$\begin{aligned}\sup_\sigma\bigg|&b_n^{-1/2}\partial_\sigma^k\sum_m\sum_{p=0}^\infty\operatorname{tr}\big((\tilde{D}_{2,m}^{-1}\tilde{L}^\top\tilde{D}_{1,m}^{-1}\tilde{L})^p\tilde{D}_{2,m}^{-1}D'_{2,m}\big)\\&-\partial_\sigma^k\int_0^T\frac{|b_t^1|^2\sqrt{\tilde{a}_t^1\tilde{a}_t^2}+\tilde{a}_t^2\sqrt{\det(b_tb_t^\top)}}{2\sqrt{\det(b_tb_t^\top)}\mathfrak{A}_t}\,dt\bigg|\xrightarrow{p}0.\end{aligned}\tag{5.18}$$

Lemmas A.3 and 5.2 and a similar argument yield

$$\begin{aligned}\partial_\sigma^k\log\det\big(\tilde{S}_m\tilde{D}_m^{-1}\big)&=\partial_\sigma^k\log\det\left(\mathcal{E}+\begin{pmatrix}0&\tilde{D}_{1,m}^{-1/2}\tilde{L}\tilde{D}_{2,m}^{-1/2}\\\tilde{D}_{2,m}^{-1/2}\tilde{L}^\top\tilde{D}_{1,m}^{-1/2}&0\end{pmatrix}\right)\\&=-\sum_{p=1}^\infty\frac{1}{p}\partial_\sigma^k\operatorname{tr}\big((\tilde{D}_{1,m}^{-1}\tilde{L}\tilde{D}_{2,m}^{-1}\tilde{L}^\top)^p\big)\\&=-\frac{T\hat{a}_m^1k_n}{\pi}\sum_{p=1}^\infty\frac{(\hat{a}_m^2)^p(\tilde{b}_m^1\cdot\tilde{b}_m^2)^{2p}I_{p,p}(c_1,c_2')}{p(\hat{a}_m^1)^{3p}b_n^{2p}v_{1,*}^pv_{2,*}^p}+o_p\big(b_n^{1/2}\ell_n^{-1}\big).\end{aligned}$$

Then Lemma A.9 yields

$$\begin{aligned}&\partial_\sigma^k\log\det\big(\tilde{S}_m\tilde{D}_m^{-1}\big)\\&=-T\hat{a}_m^1k_n\partial_\sigma^k\Bigg(\sqrt{\frac{|\tilde{b}_m^1|^2}{b_n\hat{a}_m^1v_{1,*}}}+\sqrt{\frac{\hat{a}_m^2|\tilde{b}_m^2|^2}{b_n(\hat{a}_m^1)^2v_{2,*}}}\end{aligned}\tag{5.19}$$

$$
\begin{aligned}
&- \frac{b_n^{-1/2}}{\hat{a}_m^1}\Big(\dot{a}_m^1|\tilde{b}_m^1|^2 + \dot{a}_m^2|\tilde{b}_m^2|^2 + 2\sqrt{\dot{a}_m^1\dot{a}_m^2 \det(\tilde{b}_m\tilde{b}_m^\top)}\Big)^{1/2}\Big) \\
&+ o_p\big(b_n^{1/2}\ell_n^{-1}\big) \\
={}& Tb_n^{1/2}\ell_n^{-1}\partial_\sigma^k\Big(\Big(\dot{a}_m^1|\tilde{b}_m^1|^2 + \dot{a}_m^2|\tilde{b}_m^2|^2 + 2\sqrt{\dot{a}_m^1\dot{a}_m^2 \det(\tilde{b}_m\tilde{b}_m^\top)}\Big)^{1/2} - \sqrt{\dot{a}_m^1|\tilde{b}_m^1|^2} - \sqrt{\dot{a}_m^2|\tilde{b}_m^2|^2}\Big) \\
&+ o_p\big(b_n^{1/2}\ell_n^{-1}\big).
\end{aligned}
$$

Moreover, Lemmas A.3 and 5.1 yield

$$
\begin{aligned}
&\partial_\sigma^k \log\det\big(\tilde{D}_{j,m}\tilde{D}_{j,m,*}^{-1}\big) \\
&\quad = \partial_\sigma^k \log\det\big(\mathcal{E} + \tilde{D}_{j,m,*}^{-1/2}(\tilde{D}_{j,m} - \tilde{D}_{j,m,*})\tilde{D}_{j,m,*}^{-1/2}\big) \\
&\quad = \sum_{p=1}^\infty \frac{(-1)^{p-1}}{p}\partial_\sigma^k \operatorname{tr}\big(\big(\tilde{D}_{j,m,*}^{-1}(\tilde{D}_{j,m} - \tilde{D}_{j,m,*})\big)^p\big) \qquad (5.20)\\
&\quad = \sum_{p=1}^\infty \frac{(-1)^{p-1}}{p}\partial_\sigma^k \operatorname{tr}\big(\big(\dot{D}_{j,m,*}^{-1}(\dot{D}_{j,m} - \dot{D}_{j,m,*})\big)^p\big) + o_p\big(b_n^{1/2}\ell_n^{-1}\big) \\
&\quad = \partial_\sigma^k \log\det\big(\dot{D}_{j,m}\dot{D}_{j,m,*}^{-1}\big) + o_p\big(b_n^{1/2}\ell_n^{-1}\big)
\end{aligned}
$$

when $|\tilde{b}_{m,*}^j| \geq |\tilde{b}_m^j|$, where $\tilde{D}_{j,m,*}$ and $\dot{D}_{j,m,*}$ are obtained by substituting $\sigma = \sigma_*$ in $\tilde{D}_{j,m}$ and $\dot{D}_{j,m}$, respectively. Similarly, we have $\partial_\sigma^k \log\det(\tilde{D}_{j,m}\tilde{D}_{j,m,*}^{-1}) = \partial_\sigma^k \log\det(\dot{D}_{j,m}\dot{D}_{j,m,*}^{-1}) + o_p(b_n^{1/2}\ell_n^{-1})$ when $|\tilde{b}_{m,*}^j| < |\tilde{b}_m^j|$.

On the other hand, results in Section 4.2 yield

$$
\begin{aligned}
\partial_\sigma^k \log\frac{\det \dot{D}_{j,m}}{\det \dot{D}_{j,m,*}} &= \frac{k_m^j + 1}{\pi}\partial_\sigma^k \int_0^\pi \log\frac{c_j + 2(1-\cos x)}{c_{j,*} + 2(1-\cos x)}\,dx + o_p\big(b_n^{1/2}\ell_n^{-1}\big) \\
&= 2\big(k_m^j + 1\big)\partial_\sigma^k \log\frac{\sqrt{c_j} + \sqrt{4 + c_j}}{\sqrt{c_{j,*}} + \sqrt{4 + c_{j,*}}} + o_p\big(b_n^{1/2}\ell_n^{-1}\big) \\
&= k_m^j\partial_\sigma^k(\sqrt{c_j} - \sqrt{c_{j,*}}) + o_p\big(b_n^{1/2}\ell_n^{-1}\big) \\
&= Tb_n^{1/2}\ell_n^{-1}\sqrt{\dot{a}_m^j}\partial_\sigma^k\big(|\tilde{b}_m^j| - |\tilde{b}_{m,*}^j|\big) + o_p\big(b_n^{1/2}\ell_n^{-1}\big).
\end{aligned} \qquad (5.21)
$$

The residuals are bounded uniformly with respect to σ and m. Then we obtain $\sup_\sigma |b_n^{-1/2} \times \partial_\sigma^k(H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)) - \partial_\sigma^k\mathcal{Y}_1(\sigma)| \to^p 0$ as $n \to \infty$ for any $\sigma \in \Lambda$ and $0 \leq k \leq 3$ by (5.2), (5.3) and (5.16)–(5.21).

Finally, we obtain the desired results without (A1′) by using the arguments in Proposition 3.1 of Gloter and Jacod [13]. □

6. Identifiability of the model

In this section, we check the identifiability condition, $\inf_{\sigma\neq\sigma_*}((-\mathcal{Y}_1(\sigma))/|\sigma-\sigma_*|^2)>0$ almost surely. This condition is necessary to deduce consistency of the maximum-likelihood-type estimator, as seen in Proposition 7.1. In general, it is not easy to check this condition directly because $\mathcal{Y}_1(\sigma)$ is a complicated function of b_t and a_t^j. On the other hand, Ogihara and Yoshida [25] proved that the identifiability condition (A3) of a model for equidistant observations without noise is sufficient for the identifiability of a model for nonsynchronous observations. This is also the case for our model. Recall that $\bar{\rho}$ has been defined just before Lemma 4.2.

Proposition 6.1. *Assume (A1), (A2) and (V). Then there exists a positive constant c such that*

$$-\mathcal{Y}_1(\sigma)\geq \chi\int_0^T\bigl\{\bigl(|b_t^1|^2-|b_{t,*}^1|^2\bigr)^2+\bigl(|b_t^2|^2-|b_{t,*}^2|^2\bigr)^2+\bigl(b_t^1\cdot b_t^2-b_{t,*}^1\cdot b_{t,*}^2\bigr)^2\bigr\}\,dt \tag{6.1}$$

for any σ, where

$$\chi=c\bigl(1-\bar{\rho}^2\bigr)\frac{\inf_{j,t}(a_t^j/v_{j,*})}{\sup_{j,t}(a_t^j/v_{j,*})^{1/2}}\Bigl(\sup_{j,t,\sigma}\bigl(\bigl|b^j(t,X_t,\sigma)\bigr|\vee\bigl|b^j(t,X_t,\sigma)\bigr|^{-1}\bigr)\Bigr)^{-9}.$$

In particular, $\inf_{\sigma\neq\sigma_*}((-\mathcal{Y}_1(\sigma))/|\sigma-\sigma_*|^2)>0$ *almost surely under (A1)–(A3) and (V).*

Proof. It is sufficient to show the results under the additional condition (A1′) by localization techniques similar to the proof of Proposition 2.1.

Let $\tilde{\rho}_m=|\tilde{b}_m^1|^{-1}|\tilde{b}_m^2|^{-1}\tilde{b}_m^1\cdot\tilde{b}_m^2$, $\hat{D}_m=\tilde{D}_m-M_{m,*}$ and $\mathbf{B}=\sup_{j,t,\sigma}(|b^j(t,X_t,\sigma)|\vee|b^j(t,X_t,\sigma)|^{-1})$, then since

$$\begin{aligned}&u^\top\hat{D}_m^{-1/2}\tilde{S}_m\hat{D}_m^{-1/2}u\\&\quad\geq u^\top\begin{pmatrix}\mathcal{E} & \tilde{\rho}_m\bigl\{\bigl|I_{i,m}^1\cap I_{j,m}^2\bigr|\bigl|I_{i,m}^1\bigr|^{-1/2}\bigl|I_{j,m}^2\bigr|^{-1/2}\bigr\}_{i,j}\\ \tilde{\rho}_m\bigl\{\bigl|I_{i,m}^1\cap I_{j,m}^2\bigr|\bigl|I_{i,m}^1\bigr|^{-1/2}\bigl|I_{j,m}^2\bigr|^{-1/2}\bigr\}_{j,i} & \mathcal{E}\end{pmatrix}u\end{aligned}$$

for any $u\in\mathbb{R}^{k_m^1+k_m^2}$, we have $\|(\hat{D}_m^{1/2}\tilde{S}_m^{-1}\hat{D}_m^{1/2})^{1/2}\|\leq C\mathbf{B}(1-\bar{\rho}^2)^{-1/2}$ by Lemma A.4, and hence we obtain

$$\begin{aligned}&\bigl\|\bigl(\hat{D}_m^{1/2}\tilde{S}_m^{-1}\hat{D}_m^{1/2}\bigr)^{1/2}\bigl(\hat{D}_m^{-1/2}\tilde{S}_{m,*}\hat{D}_m^{-1/2}\bigr)\bigl(\hat{D}_m^{1/2}\tilde{S}_m^{-1}\hat{D}_m^{1/2}\bigr)^{1/2}\bigr\|\\&\quad=\bigl\|\mathcal{E}+\bigl(\hat{D}_m^{1/2}\tilde{S}_m^{-1}\hat{D}_m^{1/2}\bigr)^{1/2}\bigl(\hat{D}_m^{-1/2}(\tilde{S}_{m,*}-\tilde{S}_m)\hat{D}_m^{-1/2}\bigr)\bigl(\hat{D}_m^{1/2}\tilde{S}_m^{-1}\hat{D}_m^{1/2}\bigr)^{1/2}\bigr\|\\&\quad\leq 1+C\mathbf{B}^6\bigl(1-\bar{\rho}^2\bigr)^{-1}.\end{aligned}$$

Then Lemma A.6 yields

$$\begin{aligned}&\operatorname{tr}\bigl(\tilde{S}_m^{-1}\tilde{S}_{m,*}-\mathcal{E}\bigr)+\log\det\tilde{S}_m-\log\det\tilde{S}_{m,*}\\&\quad\geq C\mathbf{B}^{-6}\bigl(1-\bar{\rho}^2\bigr)\operatorname{tr}\bigl(\tilde{S}_m^{-1}(\tilde{S}_{m,*}-\tilde{S}_m)\tilde{S}_m^{-1}(\tilde{S}_{m,*}-\tilde{S}_m)\bigr).\end{aligned} \tag{6.2}$$

Therefore, repeated use of Lemma A.5 yields

$$
\begin{aligned}
&\mathrm{tr}\big(\tilde{S}_m^{-1}\tilde{S}_{m,*}-\mathcal{E}\big)+\log\det\tilde{S}_m-\log\det\tilde{S}_{m,*}\\
&\quad\geq C\mathbf{B}^{-6}\big(1-\bar{\rho}^2\big)\,\mathrm{tr}\big(\tilde{D}_m^{-1}(\tilde{S}_{m,*}-\tilde{S}_m)\tilde{S}_m^{-1}(\tilde{S}_{m,*}-\tilde{S}_m)\big)\\
&\quad\geq C\mathbf{B}^{-6}\big(1-\bar{\rho}^2\big)\,\mathrm{tr}\big(\tilde{D}_m^{-1}(\tilde{S}_{m,*}-\tilde{S}_m)\tilde{D}_m^{-1}(\tilde{S}_{m,*}-\tilde{S}_m)\big)\\
&\quad= C\mathbf{B}^{-6}\big(1-\bar{\rho}^2\big)\Bigg\{\sum_{j=1}^{2}\big(\big|\tilde{b}_{m,*}^j\big|^2-\big|\tilde{b}_m^j\big|^2\big)^2\,\mathrm{tr}\big(\tilde{D}_{j,m}^{-1}D'_{j,m}\tilde{D}_{j,m}^{-1}D'_{j,m}\big)\\
&\qquad+2\big(\tilde{b}_{m,*}^1\cdot\tilde{b}_{m,*}^2-\tilde{b}_m^1\cdot\tilde{b}_m^2\big)^2\,\mathrm{tr}\big(\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top\big)\Bigg\}.
\end{aligned}
$$

Hence, it is sufficient to show that lim sup of $\mathrm{tr}(\tilde{D}_{j,m}^{-1}D'_{j,m}\tilde{D}_{j,m}^{-1}D'_{j,m})$ for $j=1,2$ and $\mathrm{tr}(\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top)$ are estimated from below by positive random variables.

By Lemma 5.1, (4.2) and (5.13) with a sampling scheme $\{S_i^{n,1}\}_i\equiv\{S_j^{n,2}\}_j$, we obtain

$$
\begin{aligned}
&b_n^{-1/2}\,\mathrm{tr}\big(\tilde{D}_{j,m}^{-1}D'_{j,m}\tilde{D}_{j,m}^{-1}D'_{j,m}\big)\\
&\quad=b_n^{-1/2}\,\mathrm{tr}\big(\dot{D}_{j,m}^{-1}D'_{j,m}\dot{D}_{j,m}^{-1}D'_{j,m}\big)+\dot{R}_n\big(\ell_n^{-1}\big)=b_n^{-5/2}\big(\hat{a}_m^j\big)^{-2}\,\mathrm{tr}\big(\dot{D}_{j,m}^{-2}\big)+\dot{R}_n\big(\ell_n^{-1}\big)\\
&\quad=\frac{b_n^{-5/2}(k_m^j+1)}{(\hat{a}_m^j)^2v_{j,*}^2\pi}I_2\bigg(\frac{b_n^{-1}|\tilde{b}_m^j|^2}{\hat{a}_m^jv_{j,*}}\bigg)+\dot{R}_n\big(\ell_n^{-1}\big)=\frac{T(\hat{a}_m^j)^{1/2}\ell_n^{-1}}{4v_{j,*}^{1/2}|\tilde{b}_m^j|^3}+\dot{R}_n\big(\ell_n^{-1}\big).
\end{aligned}
$$

Moreover, Lemma 5.1, (4.2) and (5.13) yield

$$
\begin{aligned}
&b_n^{-1/2}\,\mathrm{tr}\big(\tilde{D}_{1,m}^{-1}\tilde{G}\tilde{D}_{2,m}^{-1}\tilde{G}^\top\big)\\
&\quad=b_n^{-5/2}\frac{\hat{a}_m^2}{(\hat{a}_m^1)^3}\,\mathrm{tr}\big(\dot{D}_{1,m}^{-1}\ddot{D}_{2,m}^{-1}\big)+\dot{R}_n\big(\ell_n^{-1}\big)\\
&\quad\geq\frac{b_n^{-5/2}\hat{a}_m^2}{(\hat{a}_m^1)^3v_{1,*}v_{2,*}}\,\mathrm{tr}\Bigg(\Bigg(\Bigg(\bigg(\frac{|\tilde{b}_m^2|^2b_n^{-1}\hat{a}_m^2}{v_{2,*}(\hat{a}_m^1)^2}\bigg)\vee\frac{|\tilde{b}_m^1|^2b_n^{-1}}{\hat{a}_m^1v_{1,*}}\Bigg)\mathcal{E}+M_{1,m}\Bigg)^{-2}\Bigg)+\dot{R}_n\big(\ell_n^{-1}\big)\\
&\quad=\frac{T\ell_n^{-1}\dot{a}_m^1\dot{a}_m^2}{4((\dot{a}_m^1|\tilde{b}_m^1|^2)\vee(\dot{a}_m^2|\tilde{b}_m^2|^2))^{3/2}}+\dot{R}_n\big(\ell_n^{-1}\big).
\end{aligned}
$$

Therefore, we obtain (6.1).

In particular, by Lemma 6 and Remark 4 in [25], there exists a positive-valued random variable $\mathcal{R}$ such that

$$
-\mathcal{Y}_1(\sigma)\geq\chi\mathcal{R}\big(-\mathcal{Y}_0(\sigma)\big)
$$

for any σ. Therefore, we have $\inf_{\sigma\neq\sigma_*}((-\mathcal{Y}_1(\sigma))/|\sigma-\sigma_*|^2)>0$ almost surely under (A1)–(A3) and (V). □

7. Asymptotic mixed normality of the estimator

In this section, we prove the consistency and asymptotic mixed normality of $\hat{\sigma}_n$. To obtain asymptotic mixed normality, we prove stable convergence of the score function $b_n^{-1/4}\partial_\sigma H_n(\sigma_*, v_*)$ by means of the martingale limit theorem for a mixed normal limit in Jacod [18]. We also use the idea by Jacod et al. [19] to adapt the limit theorem to models containing observation noise.

Consistency is an immediate consequence of Proposition 2.1 and the identifiability condition.

Proposition 7.1. *Assume (A1)–(A3) and (V). Then $\hat{\sigma}_n \to^p \sigma_*$ as $n \to \infty$.*

Proof. Let ε, δ be arbitrary positive constants. By Proposition 2.1, we have $\sup_\sigma |b_n^{-1/2} \times (H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)) - \mathcal{Y}_1(\sigma)| \to^p 0$ as $n \to \infty$. Moreover, Proposition 6.1 ensures that there exists $\eta > 0$ such that $P[\inf_{\sigma \neq \sigma_*}((-\mathcal{Y}_1(\sigma))/|\sigma - \sigma_*|^2) \leq \eta] < \varepsilon$. Since $H_n(\hat{\sigma}_n, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n) \geq 0$ by the definition of $\hat{\sigma}_n$, we obtain

$$\begin{aligned}
P\big[|\hat{\sigma}_n - \sigma_*| \geq \delta\big] &< P\big[\mathcal{Y}_1(\hat{\sigma}_n) \leq -\eta\delta^2\big] + \varepsilon \\
&\leq P\Big[\sup_\sigma \big|b_n^{-1/2}\big(H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)\big) - \mathcal{Y}_1(\sigma)\big| \geq \eta\delta^2\Big] + \varepsilon \\
&< 2\varepsilon
\end{aligned}$$

for sufficiently large n. □

Proposition 7.2. *Assume (A1), (A2) and (V). Then $b_n^{-1/4}\partial_\sigma H_n(\sigma_*, \hat{v}_n) \to^{s\text{-}\mathcal{L}} \Gamma_1^{1/2}\mathcal{N}$ as $n \to \infty$.*

Proof. It is sufficient to prove the results assuming the additional condition (A1′).

Since $b_n^{-1/4}\partial_\sigma \tilde{H}_n(\sigma_*, v_*) = -2^{-1}b_n^{-1/4}\sum_m \bar{E}_m[\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m]$, we only need to check assumptions of Theorem 3.2 in Jacod [18] for $\mathcal{X}_m^n = -2^{-1}b_n^{-1/4}\bar{E}_m[\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m]$. For any $\varepsilon > 0$, Lemma 4.3 yields

$$\sum_{m=2}^{[\ell_n t/T]} E_m\big[|\mathcal{X}_m^n|^2 1_{\{|\mathcal{X}_m^n| > \varepsilon\}}\big] \leq \frac{Cb_n^{-1}}{\varepsilon^2}\sum_{m=2}^{[\ell_n t/T]} E_m\big[\big(\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big)^4\big] \xrightarrow{p} 0. \tag{7.1}$$

It is easy to see that $\sum_{m=2}^{[\ell_n t/T]} E_m[\mathcal{X}_m^n(W_{s_m} - W_{s_{m-1}})] = 0$.

Lemma 4.3 yields

$$\begin{aligned}
E_m\big[(\mathcal{X}_m^n)^2\big] &= \frac{b_n^{-1/2}}{4}\big\{E_m\big[\big(\tilde{Z}_m \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big)^2\big] - E_m\big[\tilde{Z}_m \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big]^2\big\} \\
&= \frac{b_n^{-1/2}}{2}\operatorname{tr}\big(\tilde{S}_{m,*}\partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{S}_{m,*}\partial_\sigma \tilde{S}_{m,*}^{-1}\big) + \bar{R}_n\big(b_n^{-1/2}\big).
\end{aligned}$$

On the other hand, since $\partial_\sigma \log\det \tilde{S}_m(x, \sigma) = \operatorname{tr}(\partial_\sigma \tilde{S}_m \tilde{S}_m^{-1})$, we have

$$\begin{aligned}
&E_m\big[\tilde{Z}_m^\top \partial_\sigma^2 \tilde{S}_m^{-1}\tilde{Z}_m + \partial_\sigma^2 \log\det \tilde{S}_m\big]|_{\sigma=\sigma_*} \\
&\quad = \operatorname{tr}\big(\partial_\sigma^2 \tilde{S}_{m,*}^{-1}\tilde{S}_{m,*}\big) - \operatorname{tr}\big(\partial_\sigma^2 \tilde{S}_{m,*}^{-1}\tilde{S}_{m,*}\big) + \operatorname{tr}\big(\tilde{S}_{m,*}^{-1}\partial_\sigma \tilde{S}_{m,*}\tilde{S}_{m,*}^{-1}\partial_\sigma \tilde{S}_{m,*}\big).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\sum_{m=2}^{[\ell_n t/T]} E_m\big[(\mathcal{X}_m^n)^2\big] \\
&\quad = \frac{b_n^{-1/2}}{2} \sum_{m=2}^{[\ell_n t/T]} E_m\big[\tilde{Z}_m^\top \partial_\sigma^2 \tilde{S}_m^{-1} \tilde{Z}_m + \partial_\sigma^2 \log\det \tilde{S}_m)\big]|_{\sigma=\sigma_*} \xrightarrow{p} -\partial_\sigma^2 \mathcal{Y}_1(\sigma_*, t),
\end{aligned} \tag{7.2}$$

where

$$\begin{aligned}
&\mathcal{Y}_1(\sigma, t) \\
&= \int_0^t \bigg\{ \frac{\sum_{j=1}^2 (|b_s^j|^2 - |b_{s,*}^j|^2)(|b_s^{3-j}|^2 \sqrt{\tilde{a}_s^1 \tilde{a}_s^2} + \tilde{a}_s^j \sqrt{\det(b_s b_s^\top)}) - 2(b_s^1 \cdot b_s^2 - b_{s,*}^1 \cdot b_{s,*}^2) b_s^1 \cdot b_s^2 \sqrt{\tilde{a}_s^1 \tilde{a}_s^2}}{4\sqrt{\det(b_s b_s^\top)}(\tilde{a}_s^1 |b_s^1|^2 + \tilde{a}_s^2 |b_s^2|^2 + 2\sqrt{\tilde{a}_s^1 \tilde{a}_s^2 \det(b_s b_s^\top)})^{1/2}} \\
&\quad - \frac{(\tilde{a}_s^1 |b_s^1|^2 + \tilde{a}_s^2 |b_s^2|^2 + 2\sqrt{\tilde{a}_s^1 \tilde{a}_s^2 \det(b_s b_s^\top)})^{1/2}}{2} \\
&\quad + \frac{(\tilde{a}_s^1 |b_{s,*}^1|^2 + \tilde{a}_s^2 |b_{s,*}^2|^2 + 2\sqrt{\tilde{a}_s^1 \tilde{a}_s^2 \det(b_{s,*} b_{s,*}^\top)})^{1/2}}{2} \bigg\} ds.
\end{aligned}$$

Then by Theorem 3.2 in Jacod [18], it is sufficient to show that $\sum_{m=2}^{[\ell_n t/T]} E_m[\mathcal{X}_m^n (N_{s_m} - N_{s_{m-1}})] \to^p 0$ for any bounded $\mathcal{G}_t$-martingale $N = (N_t)_{0 \le t \le T}$ orthogonal to W.

Orthogonality of N and W yields

$$E_m\big[\mathcal{X}_m^n (N_{s_m} - N_{s_{m-1}})\big] = E_m\big[\dot{\mathcal{X}}_m^n (N_{s_m} - N_{s_{m-1}})\big],$$

where $\dot{\mathcal{X}}_m^n = -2^{-1} b_n^{-1/4} \bar{E}_m[\tilde{Z}_{2,m}^\top \partial_\sigma S_{m,*}^{-1} \tilde{Z}_{2,m} + 2\tilde{Z}_{1,m}^\top \partial_\sigma S_{m,*}^{-1} \tilde{Z}_{2,m}]$, and $\tilde{Z}_{1,m}$ and $\tilde{Z}_{2,m}$ are defined in (A.4).

Let $N' = (N_t')_{0 \le t \le T}$ be another bounded $\mathcal{G}_t$-martingale. Then we obtain

$$\begin{aligned}
\sum_{m=2}^{[\ell_n t/T]} E_m\big[\dot{\mathcal{X}}_m^n (N_{s_m} - N_{s_{m-1}})\big] &= \sum_{m=2}^{[\ell_n t/T]} E_m\big[\dot{\mathcal{X}}_m^n \big(N'_{s_m} - N'_{s_{m-1}}\big)\big] \\
&\quad + \sum_{m=2}^{[\ell_n t/T]} E_m\big[\dot{\mathcal{X}}_m^n \big(N_{s_m} - N'_{s_m} - N_{s_{m-1}} + N'_{s_{m-1}}\big)\big],
\end{aligned}$$

and

$$\begin{aligned}
&E\Bigg[\Bigg|\sum_{m=2}^{[\ell_n t/T]} E_m\big[\dot{\mathcal{X}}_m^n \big(N_{s_m} - N'_{s_m} - N_{s_{m-1}} + N'_{s_{m-1}}\big)\big]\Bigg| 1_{\{\sum_m E_m[(\dot{\mathcal{X}}_m^n)^2] \le K\}}\Bigg] \\
&\quad \le E\Bigg[\sum_{m=2}^{\ell_n} E_m\big[(\dot{\mathcal{X}}_m^n)^2\big]^{1/2} E_m\big[\big(N_{s_m} - N'_{s_m} - N_{s_{m-1}} + N'_{s_{m-1}}\big)^2\big]^{1/2} 1_{\{\sum_m E_m[(\dot{\mathcal{X}}_m^n)^2] \le K\}}\Bigg]
\end{aligned}$$

$$\leq E\left[\sum_{m=2}^{\ell_n} E_m\big[(\dot{\mathcal{X}}_m^n)^2\big]1_{\{\sum_m E_m[(\dot{\mathcal{X}}_m^n)^2]\leq K\}}\right]^{1/2} E\left[\sum_{m=2}^{\ell_n}\big(N_{s_m}-N'_{s_m}-N_{s_{m-1}}+N'_{s_{m-1}}\big)^2\right]^{1/2}$$

$$\leq \sqrt{K}E\big[\big(N_T-N'_T\big)^2\big]^{1/2}$$

for any positive constant K.

Since we have the tightness of $\sum_{m=2}^{\ell_n} E_m[(\dot{X}_m^n)^2]$ similarly to (7.2), we obtain the desired result if for any $\varepsilon > 0$, we can find a bounded $\mathcal{G}_t$-martingale N' which satisfies $E[(N_T - N'_T)^2] \leq \varepsilon$ and $E_m[\dot{\mathcal{X}}_m^n(N'_{s_m} - N'_{s_{m-1}})] = 0$ for any m. We obtain such N' by setting $N'_t = E[\tilde{N}|\mathcal{G}_t]$ for suitable $\tilde{N} \in \mathbf{N}$, where $\mathbf{N}$ is the set of finite sums of random variables $\mathbf{X}\prod_{j=1}^{l} g_j(\varepsilon_{i_j}^{n_j,k_j})$ with bounded Borel functions $(g_j)_{j=1}^{l}$, an $\mathcal{F}^{(0)}$-measurable bounded random variable $\mathbf{X}$, $n_1, \dots, n_l \in \mathbb{N}$, $1 \leq k_1, \dots, k_l \leq 2$, and $i_1, \dots, i_l \in \mathbb{Z}_+$. Then $\mathbf{N}$ is dense in $L^2(\Omega, \mathcal{G}_T, P)$ and we obtain $E_m[\dot{\mathcal{X}}_m^n(N'_{s_m} - N'_{s_{m-1}})] = 0$ for any m if $n > \max_j n_j$. □

Proof of Theorem 2.1. Since the parameter space Λ is open, there exists $\varepsilon > 0$ such that $O(\varepsilon, \sigma_*) = \{\sigma; |\sigma - \sigma_*| < \varepsilon\} \subset \Lambda$. Then we have

$$-\partial_\sigma H_n(\sigma_*, \hat{\upsilon}_n) = \int_0^1 \partial_\sigma^2 H_n(\sigma_*, \hat{\upsilon}_n)\big(\sigma_* + t(\hat{\sigma}_n - \sigma_*)\big)(\hat{\sigma}_n - \sigma_*)\,dt$$

for $\hat{\sigma}_n \in O(\varepsilon, \sigma_*)$, by $\partial_\sigma H_n(\hat{\sigma}_n, \hat{\upsilon}_n) = 0$.

Hence, we obtain $b_n^{1/4}(\hat{\sigma}_n - \sigma_*) = \tilde{\Gamma}_{1,n}^{-1} b_n^{-1/4}\partial_\sigma H_n(\sigma_*, \hat{\upsilon}_n)$ on $\{\det\tilde{\Gamma}_{1,n} \neq 0$ and $\hat{\sigma}_n \in O(\varepsilon, \sigma_*)\}$, where $\tilde{\Gamma}_{1,n} = -b_n^{-1/2}\int_0^1 \partial_\sigma^2 H_n(\sigma_* + t(\hat{\sigma}_n - \sigma_*))\,dt$. Then since Propositions 2.1, 6.1 and 7.1 yield $P[\det\tilde{\Gamma}_{1,n} = 0] \to 0$, $P[\hat{\sigma}_n \notin O(\varepsilon, \sigma_*)] \to 0$ and $\tilde{\Gamma}_{1,n}^{-1}1_{\{\det\tilde{\Gamma}_{1,n}\neq 0\}} \to^p \Gamma_1^{-1}$, we have $b_n^{1/4}(\hat{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \Gamma_1^{-1/2}\mathcal{N}$ as $n \to \infty$ by Proposition 7.2.

Moreover, Propositions 2.1 and 7.1 ensure that $\hat{\Gamma}_{1,n} \to^p \Gamma_1$, which completes the proof. □

8. Proof of the LAN property

To obtain the LAN property of our model, the arguments in the proof of Theorem 2.1 are essential. Indeed, by using Propositions 2.1 and 7.2, we obtain a LAMN-type property of the quasi-log-likelihood function H_n with respect to σ: $H_n(\sigma_* + b_n^{-1/4}u_1, \upsilon_*) - H_n(\sigma_*, \upsilon_*) - u_1 \cdot b_n^{-1/4}\partial_\sigma H_n(\sigma_*, \upsilon_*) - u_1^\top b_n^{-1/2}\partial_\sigma^2 H_n(\sigma_*, \upsilon_*)u_1/2 \to^p 0$ as $n \to \infty$ for any $u_1 \in \mathbb{R}^d$, and $(b_n^{-1/4}\partial_\sigma H_n(\sigma_*, \upsilon_*), -b_n^{-1/2}\partial_\sigma^2 H_n(\sigma_*, \upsilon_*)) \to^{s\text{-}\mathcal{L}} (\Gamma_1^{1/2}\mathcal{N}, \Gamma_1)$, where $\mathcal{N}$ is a d-dimensional standard normal random variable independent of $\mathcal{F}$. On the other hand, under the assumptions of Theorem 2.2, the *true* log-likelihood ratio $\log(dP_{\sigma_*+b_n^{-1/4}u_1, \upsilon_*+b_n^{-1/2}u_2, n}/dP_{\sigma_*, \upsilon_*, n})$ for $u_1 \in \mathbb{R}^d$ and $u_2 \in \mathbb{R}^2$ is obtained as $-(Z_1^\top S_1^{-1}Z_1 + \log\det S_1)/2$ if we set $k_n = b_n$. We cannot apply the argument of Section 5 to this quantity because the estimate $\ell_n \to \infty$ is essential there. Therefore, we follow the approaches by Gloter and Jacod [12] to show the LAN property. We set a "subexperiment" and a "superexperiment", which are obtained by respectively removing and adding observations from the original experiment. The likelihood functions of these experiments have similar properties to H_n, and therefore we can prove the LAN properties for these experiments with the same limit distribution. We can prove that these results lead us to the LAN property of the original one.

Let $\mathcal{Z} = (\mathbb{R}^8)^{\mathbb{Z}_+}$, $\pi_i(z) = (x_i^{k,j}, t_i^k, e_i^k)_{j,k=1,2}$ for $i \in \mathbb{Z}_+$ and $z = (x_{i'}^{k,j}, t_{i'}^k, e_{i'}^k)_{i' \in \mathbb{Z}_+, j,k=1,2} \in \mathcal{Z}$. Let $\mathcal{H} = \mathfrak{B}(\{\pi_i^{-1}(A); i \in \mathbb{Z}_+, A \in \mathcal{B}(\mathbb{R}^8)\})$, $P'_{\sigma'_*, v'_*}$ be the induced probability measure on $(\mathcal{Z}, \mathcal{H})$ by $((Y^j_{S_i^{n,k}} 1_{\{i \le \mathbf{J}_{k,n}\}}, S_i^{n,k} 1_{\{i \le \mathbf{J}_{k,n}\}}, \varepsilon_i^{n,k} 1_{\{i \le \mathbf{J}_{k,n}\}})_{i \in \mathbb{Z}_+, j,k=1,2})$ with a true value (σ'_*, v'_*). We can ignore the event $\min_{j,m} k_m^j \le 0$.

Let $\mathcal{H}' = \mathfrak{B}(t_i^k; i \in \mathbb{Z}_+, k = 1, 2)$, $j_0^k = -1$, $j_m^k = \max\{i; t_i^k < s_m\} \vee 0$ $(1 \le m \le \ell_n)$, $\mathbf{l}(0) = 1$, $\mathbf{l}(m) = \min\{k; t_i^k = \max_{i',k'}\{t_{i'}^{k'} < s_m\}$ for some $i\}$ for $1 \le m \le \ell_n$,

$$
\begin{aligned}
\mathcal{H}^{n,0} &= \mathfrak{B}\big((x_{i+1}^{k,k} + e_{i+1}^k - x_i^{k,k} - e_i^k) 1_{\{i \notin \{j_m^k\}_m\}}; i \in \mathbb{Z}_+, k = 1, 2\big) \vee \mathcal{H}', \\
\mathcal{H}^{n,1} &= \mathfrak{B}\big(x_0^{k,k}, x_i^{k,k} + e_i^k; i \in \mathbb{Z}_+, k = 1, 2\big) \vee \mathcal{H}', \\
\mathcal{H}^{n,2} &= \mathcal{H}^{n,1} \vee \mathfrak{B}\big(x^{\mathbf{l}(m),j}_{j_m^{\mathbf{l}(m)}}; 1 \le m \le \ell_n - 1, j = 1, 2\big) \vee \mathfrak{B}\big(x^{1,1}_{j^1_{\ell_n}+1}, x^{1,2}_{j^1_{\ell_n}+1}\big).
\end{aligned}
$$

Then we can see $\mathcal{H}^{n,0} \subset \mathcal{H}^{n,1} \subset \mathcal{H}^{n,2}$ and

$$
\log(dP_{\sigma_u,v_u}/dP_{\sigma_*,v_*}) = \log \frac{d(P'_{\sigma_u,v_u}|_{\mathcal{H}^{n,1}})}{d(P'_{\sigma_*,v_*}|_{\mathcal{H}^{n,1}})}. \tag{8.1}
$$

Moreover, we obtain

$$
\begin{aligned}
&\log \frac{d(P'_{\sigma_u,v_u}|_{\mathcal{H}^{n,l}})}{d(P'_{\sigma_*,v_*}|_{\mathcal{H}^{n,l}})} \big((Y^j_{S_i^{n,k}} 1_{\{i \le \mathbf{J}_{k,n}\}}, S_i^{n,k} 1_{\{i \le \mathbf{J}_{k,n}\}}, \varepsilon_i^{n,k} 1_{\{i \le \mathbf{J}_{k,n}\}})_{i \in \mathbb{Z}_+, j,k=1,2}\big) \\
&\quad = H_n^{(l)}(\sigma_u, v_u) - H_n^{(l)}(\sigma_*, v_*)
\end{aligned} \tag{8.2}
$$

for $l = 0$, where $\mu(t, \sigma) = (\mu^1(t, \sigma), \mu^2(t, \sigma))^\top$, $Z_m^{(0)}(\sigma) = (\tilde{Y}^k(I_{i,m}^k) - \int_{I_{i,m}^k} \mu^k(t, \sigma)\, dt)_{i,k}$, and

$$
S_m^{(0)}(\sigma, v) = \begin{pmatrix} \operatorname{diag}\Big(\Big(\int_{I_{i,m}^1} |b_t^1|^2\, dt\Big)_i\Big) & \Big(\int_{I_{i,m}^1 \cap I_{j,m}^2} b_t^1 \cdot b_t^2\, dt\Big)_{i,j} \\ \Big(\int_{I_{i,m}^1 \cap I_{j,m}^2} b_t^1 \cdot b_t^2\, dt\Big)_{j,i} & \operatorname{diag}\Big(\Big(\int_{I_{j,m}^2} |b_t^2|^2\, dt\Big)_j\Big) \end{pmatrix} + \begin{pmatrix} v_1 M_{1,m} & 0 \\ 0 & v_2 M_{2,m} \end{pmatrix}
$$

for $1 \le m \le \ell_n$, $H_n^{(0)}(\sigma, v) = -\sum_{m=1}^{\ell_n} \{(Z_m^{(0)}(\sigma))^\top (S_m^{(0)})^{-1}(\sigma, v) Z_m^{(0)}(\sigma) + \log\det S_m^{(0)}(\sigma, v)\}/2$, $\sigma_u = \sigma_* + b_n^{-1/4} u_1$ and $v_u = v_* + b_n^{-1/2} u_2$ for $u = (u_1, u_2) \in \mathbb{R}^d \times \mathbb{R}^2$. $\mathcal{H}^{n,0}$ and $\mathcal{H}^{n,2}$ are σ-fields for "subexperiment" and "superexperiment", respectively, while $\mathcal{H}^{n,1}$ is the one for the original one.

To obtain similar formula to (8.2) for $l=2$, let $\mathbf{R}_m = S^{n,1}_{K^1_m} \vee S^{n,2}_{K^2_m}$, $\mathbf{Q}^k_m = S^{n,k}_{K^k_m}$, $\mathbf{Q}^k_{m,\pm} = S^{n,k}_{K^k_m \pm 1}$,

$$\tilde{\mathbf{Y}}^k_{m,-}(\sigma) = \tilde{Y}^k_{K^k_{m-1}+1} - Y^k_{\mathbf{R}_{m-1}} - \int_{\mathbf{R}_{m-1}}^{\mathbf{Q}^k_{m-1,+}} \mu^k(t,\sigma)\,dt,$$

$$\tilde{\mathbf{Y}}^k_{m,+}(\sigma) = \begin{cases} Y^k_{\mathbf{Q}^k_m} - \tilde{Y}^k_{K^k_m - 1} - \int_{\mathbf{Q}^k_{m,-}}^{\mathbf{Q}^k_m} \mu^k(t,\sigma)\,dt, & \text{if } \mathbf{Q}^k_m = \mathbf{R}_m, \\ \begin{pmatrix} \tilde{Y}^k(I^k_{k^k_m,m}) - \int_{I^k_{k^k_m,m}} \mu^k(t,\sigma)\,dt \\ Y^k_{\mathbf{R}_m} - \tilde{Y}^k_{K^k_m} - \int_{\mathbf{Q}^k_m}^{\mathbf{R}_m} \mu^k(t,\sigma)\,dt \end{pmatrix}, & \text{if } \mathbf{Q}^k_m < \mathbf{R}_m \end{cases}$$

$\mathbf{Y}_{m,0} = \varepsilon^{n,k}_{K^k_m}$ if $\mathbf{Q}^{3-k}_m < \mathbf{R}_m$, $\mathbf{Y}_{m,0} = (\varepsilon^{n,1}_{K^1_m}, \varepsilon^{n,2}_{K^2_m})^\top$ if $\mathbf{Q}^1_m = \mathbf{Q}^2_m$, and

$$Z^{(2)}_m(\sigma) = \left(\left(\left(\tilde{\mathbf{Y}}^k_{m,-}(\sigma)\right)^\top, \left(\tilde{Y}^k\left(I^k_{i,m}\right) - \int_{I^k_{i,m}} \mu^k(t,\sigma)\,dt \right)^\top_{1 \le i < k^k_m}, \left(\tilde{\mathbf{Y}}^k_{m,+}(\sigma)\right)^\top \right)^2_{k=1}, \mathbf{Y}^\top_{m,0} \right)^\top$$

for $2 \le m \le \ell_n - 1$. Then (8.2) holds for $l=2$, where $\bar{I}^k_{1,m} = [\mathbf{R}^k_{m-1}, \mathbf{Q}^k_{m-1,+})$, $\bar{I}^k_{i,m} = I^k_{i-1,m}$ for $2 \le i \le k^k_m + 1$,

$$\begin{aligned}
&\mathbf{E}(v) = v_{3-k}, \qquad k^{(2),k}_m = k^k_m + 2, \qquad k^{(2),3-k}_m = k^{3-k}_m + 1, \quad \text{and} \\
&\bar{I}^k_{k^{(2),k}_m,m} = \left[\mathbf{Q}^k_m, \mathbf{R}_m\right), \quad \text{if } \mathbf{Q}^k_m < \mathbf{R}_m, \\
&\mathbf{E}(v) = \operatorname{diag}(v_1, v_2) \quad \text{and} \quad \left(k^{(2),1}_m, k^{(2),2}_m\right) = \left(k^1_m + 1, k^2_m + 1\right) \quad \text{if } \mathbf{Q}^1_m = \mathbf{Q}^2_m, \\
&\left(M^{(2)}_{j,m}\right)_{ii'} = 2\delta_{ii'} - 1_{\{|i-i'|=1\}} - 1_{\{i=i'=1\}} - 1_{\{i=i'=k^{(2),j}_m\}}, \\
&S^{(2)}_m(\sigma, v) \\
&\quad = \begin{pmatrix} \operatorname{diag}\left(\left(\int_{\bar{I}^1_{i,m}} \left|b^1_t\right|^2 dt\right)_{1 \le i \le k^{(2),1}_m}\right) & \left\{\int_{\bar{I}^1_{i,m} \cap \bar{I}^2_{j,m}} b^1_t \cdot b^2_t\,dt\right\}_{\substack{1 \le i \le k^{(2),1}_m \\ 1 \le j \le k^{(2),2}_m}} & 0 \\ \left\{\int_{\bar{I}^1_{i,m} \cap \bar{I}^2_{j,m}} b^1_t \cdot b^2_t\,dt\right\}_{\substack{1 \le j \le k^{(2),2}_m \\ 1 \le i \le k^{(2),1}_m}} & \operatorname{diag}\left(\left(\int_{\bar{I}^2_{j,m}} \left|b^2_t\right|^2 dt\right)_{1 \le j \le k^{(2),2}_m}\right) & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&\qquad + \begin{pmatrix} v_1 M^{(2)}_{1,m} & 0 & 0 \\ 0 & v_2 M^{(2)}_{2,m} & 0 \\ 0 & 0 & \mathbf{E}(v) \end{pmatrix}
\end{aligned}$$

for $2 \le m \le \ell_n - 1$, $Z_m^{(2)}$, $M_m^{(2)}$, and $S_m^{(2)}$ are similarly defined for $m = 1, \ell_n$, and

$$H_n^{(2)}(\sigma, v) = -\frac{1}{2}\sum_{m=1}^{\ell_n}\bigl\{\bigl(Z_m^{(2)}(\sigma)\bigr)^\top S_m^{(2)}(\sigma, v)^{-1} Z_m^{(2)}(\sigma) + \log\det S_m^{(2)}(\sigma, v)\bigr\}.$$

The log-likelihood functions $H_n^{(0)}$ and $H_n^{(2)}$ of "subexperiment" and "superexperiment", respectively, have similar properties to that of H_n, and hence we can prove convergence of likelihood ratios. Gloter and Jacod [12] showed that convergence of likelihood ratios of "subexperiment" and "superexperiment" imply convergence of that of the original experiment (Theorem 4.1). Here, we use a slight extension of their result. The proof is straightforward. Let $\mathbf{U}_{\sigma,v}^{n,l} = d(P'_{\sigma,v}|_{\mathcal{H}^{n,l}})/d(P'_{\sigma_*,v_*}|_{\mathcal{H}^{n,l}})$, $K \in \mathbb{N}$, and $\{\sigma_n^k\}_{n\in\mathbb{N},1\le k\le K} \subset \Lambda$ and $v_n^k\}_{n\in\mathbb{N},1\le k\le K} \subset (0,\infty)\times(0,\infty)$ be arbitrary sequences.

Theorem 8.1. *Suppose that* $(\mathbf{U}^{n,l}_{\sigma_n^1,v_n^1},\dots,\mathbf{U}^{n,l}_{\sigma_n^K,v_n^K})$ *converges in law under* $P'^n_{\sigma_*,v_*}$ *to a limit* $Y = (Y^1,\dots,Y^K)$ *with* $0 < Y^k < \infty$ *almost surely and* $E[Y^k] = 1$ *for* $l = 0, 2$ *and* $1 \le k \le K$. *Then the same convergence holds for* $l = 1$.

We will show the LAN properties of "subexperiment" and 'superexperiment'. Then Theorem 8.1 leads to the LAN property of the original one.

First, Taylor's formula yields

$$\begin{aligned}
&H_n^{(l)}(\sigma_u, v_u) - H_n^{(l)}(\sigma_*, v_*)\\
&\quad= b_n^{-1/4}\partial_\sigma H_n^{(l)}(\sigma_*, v_*)\cdot u_1 + 2^{-1}b_n^{-1/2}u_1^\top \partial_\sigma^2 H_n^{(l)}(\sigma_*, v_*)u_1 + b_n^{-1/2}\partial_v H_n^{(l)}(\sigma_*, v_*)\cdot u_2\\
&\qquad+ 2^{-1}b_n^{-1}u_2^\top\partial_v^2 H_n^{(l)}(\sigma_*, v_*)u_2 + \int_0^1\int_0^1\sum_{i,j} b_n^{-\frac{3}{4}}\partial_{v_i}\partial_{\sigma_j} H_n^{(l)}(\sigma_{tu}, v_{su})u_{2,i}u_{1,j}\,ds\,dt\\
&\qquad+ \int_0^1\frac{(1-t)^2}{2}\sum_{i,j,k} b_n^{-\frac{3}{4}}\partial_{\sigma_i}\partial_{\sigma_j}\partial_{\sigma_k} H_n^{(l)}(\sigma_{tu}, v_*)u_{1,i}u_{1,j}u_{1,k}\,dt\\
&\qquad+ \int_0^1\frac{(1-t)^2}{2}\sum_{i,j,k} b_n^{-\frac{3}{2}}\partial_{v_i}\partial_{v_j}\partial_{v_k} H_n^{(l)}(\sigma_*, v_{tu})u_{2,i}u_{2,j}u_{2,k}\,dt.
\end{aligned}$$

We examine the limit of each term on the right-hand side. First, we prepare an auxiliary lemma. Let $\mathfrak{P}_q = \{\alpha = (\alpha_1,\dots,\alpha_L); L\in\mathbb{N}, 1\le\alpha_l\le q, \sum_{l=1}^L \alpha_l = q\}$, $M_{m,*}^{(0)} = M_{m,*}$, $M_{m,*}^{(2)} = \mathrm{diag}(v_{1,*}M_{1,m}^{(2)}, v_{2,*}M_{2,m}^{(2)}, \mathbf{E}(v_*))$,

$$\mathbf{Y}_{2,m,+}^k = \begin{cases} -\varepsilon^{n,k}_{K_m^k-1}, & \text{if } \mathbf{Q}_m^k \ge \mathbf{Q}_m^{3-k},\\ \bigl(\varepsilon^{n,k}_{K_m^k} - \varepsilon^{n,k}_{K_m^k-1}, -\varepsilon^{n,k}_{K_m^k}\bigr), & \text{if } \mathbf{Q}_m^k < \mathbf{Q}_m^{3-k}\end{cases}$$

$Z_{2,m}^{(0)}=(\varepsilon^{n,k}_{K^k_{m-1}+i+1}-\varepsilon^{n,k}_{K^k_{m-1}+i})_{1\leq i\leq k^k_m,k=1,2}$, and

$$Z_{2,m}^{(2)}=\Big(\big(\varepsilon^{n,k}_{K^k_{m-1}+1},\big(\varepsilon^{n,k}_{K^k_{m-1}+i+1}-\varepsilon^{n,k}_{K^k_{m-1}+i}\big)^\top_{1\leq i<k^k_m},\big(\mathbf{Y}^k_{2,m,+}\big)^\top\big)_{k=1,2},\mathbf{Y}^\top_{m,0}\Big)^\top \tag{8.3}$$

for $2\leq m\leq \ell_n-1$. Though we can similarly define $Z_{2,m}^{(0)}$ and $Z_{2,m}^{(2)}$ for $m=1,\ell_n$ so that they satisfy the following lemma, we omit it to avoid redundancy. We also omit discussions for $m=1,\ell_n$ throughout the rest of this section.

Lemma 8.1. *Assume (A1″). Let $p\in\mathbb{N}$, $l\in\{0,2\}$, $1\leq m\leq\ell_n$ and let $\dot{\mathbf{S}}_m$ be a $\mathcal{G}_{s_{m-1}}$-measurable random matrix of suitable size. Then there exists a positive constant C_p depending only on p such that*

(i) $|E_m[((Z_{2,m}^{(l)})^\top\dot{\mathbf{S}}_mZ_{2,m}^{(l)})^p]|\leq C_p\sum_{\alpha=(\alpha_1,\dots,\alpha_L)\in\mathfrak{P}_p}\prod_{l=1}^L|\mathrm{tr}((\dot{\mathbf{S}}_mM_{m,*}^{(l)})^{\alpha_l})|$.

(ii) $E_m[\bar{E}_m[(Z_{2,m}^{(l)})^\top\dot{\mathbf{S}}_mZ_{2,m}^{(l)}]^4]=12\,\mathrm{tr}((M_{m,*}^{(l)}\dot{\mathbf{S}}_m)^2)^2+48\,\mathrm{tr}((M_{m,*}^{(l)}\dot{\mathbf{S}}_m)^4)$.

Proof. (i) Let $U_m^{(l)}$ be an orthogonal matrix and $\Lambda_m^{(l)}$ be a diagonal matrix satisfying $U_m^{(l)}M_{m,*}^{(l)}(U_m^{(l)})^\top=\Lambda_m^{(l)}$. Then since $Z_{2,m}^{(l)}\sim N(0,M_{m,*}^{(l)})$, a similar estimate to (A.5) yields

$$E_m\big[\big((Z_{2,m}^{(l)})^\top\dot{\mathbf{S}}_mZ_{2,m}^{(l)}\big)^p\big]=\sum_{i_1,\dots,i_{2p}}\prod_{q=1}^p\big(U_m^{(l)}\dot{\mathbf{S}}_m(U_m^{(l)})^\top\big)_{i_{2q-1}i_{2q}}\sum_{(l_{2q'-1},l_{2q'})_{q'=1}^p}\prod_{q'=1}^p\big(\Lambda_m^{(l)}\big)_{l_{2q'-1},l_{2q'}},$$

where the second summation on the right-hand side is over all disjoint pairs $(l_{2q'-1},l_{2q'})_{q'=1}^p$ of variables $i_1,\dots,i_{2p}$.

(ii)

$$\begin{aligned}
&E_m\big[\big(\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)-\big(Z_{2,m}^{(l)}\big)^\top\dot{\mathbf{S}}_mZ_{2,m}^{(l)}\big)^4\big]\\
&=\sum_{r=0}^4\frac{(-1)^r4!}{r!(4-r)!}\big(\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)\big)^rE_m\big[\big(\big(Z_{2,m}^{(l)}\big)^\top\dot{\mathbf{S}}_mZ_{2,m}^{(l)}\big)^{4-r}\big]\\
&=\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^4-4\,\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^4+6\,\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2\big(\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2+2\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2\big)\big)\\
&\quad-4\,\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)\big(\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^3+6\,\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2\big)+8\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^3\big)\big)\\
&\quad+\big\{\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^4+12\,\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2\big)+12\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2\big)^2\\
&\quad+32\,\mathrm{tr}\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^3\big)+48\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^4\big)\big\}\\
&=12\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^2\big)^2+48\,\mathrm{tr}\big(\big(M_{m,*}^{(l)}\dot{\mathbf{S}}_m\big)^4\big).
\end{aligned}$$

□

Lemma 8.2. *Assume (A1″) and (A2). Then*

1. $\sup_\sigma|b_n^{-1/2}\partial_\sigma^k(H_n^{(l)}(\sigma,v_*)-H_n^{(l)}(\sigma_*,v_*))-\partial_\sigma^k\mathcal{Y}_1(\sigma)|\to^p0$,

2. $\sup_v |b_n^{-1}\partial_v^k(H_n^{(l)}(\sigma_*, v) - H_n^{(l)}(\sigma_*, v_*)) - \partial_v^k \mathcal{Y}_2(v)| \to^p 0$,
3. $\sup_{\sigma,v} |b_n^{-3/4}\partial_\sigma \partial_v H_n^{(l)}(\sigma, v)| \to^p 0$

as $n \to \infty$ *for* $0 \le k \le 3$ *and* $l = 0, 2$.

Proof. 1. For any $\varepsilon > 0$, $(\varepsilon\mathcal{E} + M_{j,m}^{(2)})^{-1}$ has a similar decomposition to (4.4) by replacing $p_1, \dots, p_{k_m^j-1}, p_{k_m^j}$ by $p'_1, \dots, p'_{k_m^j-1}, p'_{k_m^j} - 1$. Then, We obtain the desired results by a similar argument to the proof of Proposition 2.1 together with Lemmas 4.1 and 8.1, the results in Section 8 of [12], and similar estimates to Lemmas 5.1, 4.2 and 4.3. Estimate for the quantity corresponding to Λ_1 in Lemma 5.2 is obtained since $\{((\varepsilon\mathcal{E} + M_{j,m}^{(2)})^{-1})_{kk}\}_{k=1}^{[k_m^j/2]}$ is nonincreasing similarly to Lemma 4.2, and

$$\big((c_j v_{j,*}^{-1}\mathcal{E} + M_{j,m}^{(2)})^{-1}\big)_{11} = \frac{\prod_{l=1}^{k_m^j-1} p'_l(c_j v_{j,*}^{-1})}{(p'_{k_m^j}(c_j v_{j,*}^{-1}) - 1)\prod_{l=1}^{k_m^j-1} p'_l(c_j v_{j,*}^{-1})} = O_p\big(b_n^{1/2}\big).$$

2. Similarly to (5.1), we obtain

$$b_n^{-1}\partial_v^l H_n^{(l)}(\sigma, v) = -\frac{1}{2}b_n^{-1}\sum_m \big\{\mathrm{tr}\big(\partial_v^l\big(S_m^{(l)}\big)^{-1} S_{m,*}^{(l)}\big) + \partial_v^l \log\det S_m^{(l)}\big\} + o_p(1).$$

Let $k_m^{(0),j} = k_m^j$ for $j = 1, 2$, $\tilde{D}_{1,m}^{(l)} = ((S_m^{(l)})_{i,i'})_{1\le i,i'\le k_m^{(l),1}}$,

$$\begin{aligned}
\tilde{D}_{2,m}^{(0)} &= \big(\big(S_m^{(0)}\big)_{j,j'}\big)_{k_m^{(0),1}<j,j'\le k_m^{(0),1}+k_m^{(0),2}}, \\
\tilde{D}_{2,m}^{(2)} &= \mathrm{diag}\big(\big(\big(S_m^{(2)}\big)_{j,j'}\big)_{k_m^{(2),1}<j,j'\le k_m^{(2),1}+k_m^{(2),2}}, \mathbf{E}(v)\big), \\
\hat{G}_m^{(0)} &= \big(\tilde{D}_{1,m}^{(0)}\big)^{-1/2}\bigg\{\int_{I_{i,m}^1\cap I_{j,m}^2} b_t^1 \cdot b_t^2\,dt\bigg\}_{1\le i\le k_m^{(0),1}, 1\le j\le k_m^{(0),2}}\big(\tilde{D}_{2,m}^{(0)}\big)^{-1/2}, \\
\hat{G}_m^{(2)} &= \big(\tilde{D}_{1,m}^{(2)}\big)^{-1/2}\bigg\{\int_{\bar{I}_{i,m}^1\cap \bar{I}_{j,m}^2} b_t^1 \cdot b_t^2\,dt 1_{\{j\le k_m^{(2),2}\}}\bigg\}_{1\le i\le k_m^{(2),1}, 1\le j\le \tilde{k}_m^{(2),2}}\big(\tilde{D}_{2,m}^{(2)}\big)^{-1/2},
\end{aligned}$$

and $\tilde{D}_m^{(l)} = \mathrm{diag}(\tilde{D}_{1,m}^{(l)}, \tilde{D}_{2,m}^{(l)})$ for $l = 0, 2$, where $\tilde{k}_m^{(2),2}$ is the size of $\tilde{D}_{2,m}^{(2)}$. Then we obtain

$$\begin{aligned}
&\mathrm{tr}\big(\big(S_m^{(l)}\big)^{-1} S_{m,*}^{(l)}\big) \\
&= \mathrm{tr}\left(\big(\tilde{D}_m^{(l)}\big)^{-1/2}\begin{pmatrix}\mathcal{E} & \hat{G}_m^{(l)} \\ \big(\hat{G}_m^{(l)}\big)^\top & \mathcal{E}\end{pmatrix}^{-1}\big(\tilde{D}_m^{(l)}\big)^{-1/2}\big(\tilde{D}_{m,*}^{(l)}\big)^{1/2}\begin{pmatrix}\mathcal{E} & \hat{G}_{m,*}^{(l)} \\ \big(\hat{G}_{m,*}^{(l)}\big)^\top & \mathcal{E}\end{pmatrix}\big(\tilde{D}_{m,*}^{(l)}\big)^{1/2}\right) \\
&= \sum_{p=0}^\infty \big\{\mathrm{tr}\big(\big(\tilde{D}_{1,m}^{(l)}\big)^{-1/2}\big(\hat{G}_m^{(l)}\big(\hat{G}_m^{(l)}\big)^\top\big)^p\big(\tilde{D}_{1,m}^{(l)}\big)^{-1/2}\tilde{D}_{1,m,*}^{(l)}
\end{aligned}$$

$$
\begin{aligned}
&- (\tilde{D}_{1,m}^{(l)})^{-1/2}(\hat{G}_m^{(l)}(\hat{G}_m^{(l)})^\top)^p \hat{G}_m^{(l)}(\tilde{D}_{2,m}^{(l)})^{-1/2}(\tilde{D}_{2,m,*}^{(l)})^{1/2}(\hat{G}_{m,*}^{(l)})^\top(\tilde{D}_{1,m,*}^{(l)})^{1/2}) \\
&+ \mathrm{tr}((\tilde{D}_{2,m}^{(l)})^{-1/2}((\hat{G}_m^{(l)})^\top \hat{G}_m^{(l)})^p(\tilde{D}_{2,m}^{(l)})^{-1/2}\tilde{D}_{2,m,*}^{(l)} \\
&- (\tilde{D}_{2,m}^{(l)})^{-1/2}(\hat{G}_m^{(l)})^\top(\hat{G}_m^{(l)}(\hat{G}_m^{(l)})^\top)^p(\tilde{D}_{1,m}^{(l)})^{-1/2}((\tilde{D}_{1,m,*}^{(l)}))^{1/2}\hat{G}_{m,*}^{(l)}(\tilde{D}_{2,m,*}^{(l)})^{1/2})\}.
\end{aligned}
$$

Since $\|(\tilde{D}_{j,m}^{(l)})^{-1/2}\tilde{D}_{j,m,*}^{(l)}(\tilde{D}_{j,m}^{(l)})^{-1/2}\| = O_p(1)$, terms involving $\hat{G}_m^{(l)}$ on the right-hand side are $O_p(b_n^{1/2}\ell_n^{-1})$. Therefore, we have

$$
\begin{aligned}
&b_n^{-1}\mathrm{tr}((S_m^{(l)})^{-1}S_{m,*}^{(l)}) \\
&\quad = b_n^{-1}\sum_{j=1}^2 \mathrm{tr}((\tilde{D}_{j,m}^{(l)})^{-1}\tilde{D}_{j,m,*}^{(l)}) + o_p(\ell_n^{-1}) \\
&\quad = b_n^{-1}\sum_{j=1}^2 \frac{v_{j,*}}{v_j}\mathrm{tr}(\mathcal{E}_{k_m^j} - (\tilde{D}_{j,m}^{(l)})^{-1}(\tilde{D}_{j,m}^{(l)} - v_j v_{j,*}^{-1}\tilde{D}_{j,m,*}^{(l)})) + o_p(\ell_n^{-1}) \\
&\quad = T\ell_n^{-1}\sum_{j=1}^2 \hat{a}_m^j \frac{v_{j,*}}{v_j} + o_p(\ell_n^{-1}).
\end{aligned}
$$

Similarly, we have $b_n^{-1}\mathrm{tr}(\partial_v^k(S_m^{(l)})^{-1}S_{m,*}^{(l)}) = T\ell_n^{-1}\sum_{j=1}^2 \hat{a}_m^j \partial_v^k \frac{v_{j,*}}{v_j} + o_p(\ell_n^{-1})$. Moreover, we obtain

$$
\begin{aligned}
&b_n^{-1}\partial_v^k \log\frac{\det S_m^{(l)}}{\det S_{m,*}^{(l)}} \\
&\quad = \sum_{j=1}^2 b_n^{-1}\partial_v^k \log\det((\tilde{D}_{j,m,*}^{(l)})^{-1}\tilde{D}_{j,m}^{(l)}) + b_n^{-1}\partial_v^k\log\det(\mathcal{E} - \hat{G}_m^{(l)}(\hat{G}_m^{(l)})^\top) \\
&\qquad - b_n^{-1}\partial_v^k\log\det(\mathcal{E} - \hat{G}_{m,*}^{(l)}(\hat{G}_{m,*}^{(l)})^\top) \\
&\quad = b_n^{-1}\sum_{j=1}^2 \partial_v^k\log\det(v_j v_{j,*}^{-1}\mathcal{E}_{k_m^j} + (\tilde{D}_{j,m,*}^{(l)})^{-1}(\tilde{D}_{j,m}^{(l)} - v_j v_{j,*}^{-1}\tilde{D}_{j,m,*}^{(l)})) + o_p(\ell_n^{-1}) \\
&\quad = T\ell_n^{-1}\sum_{j=1}^2 \hat{a}_m^j \partial_v^k \log(v_j v_{j,*}^{-1}) + o_p(\ell_n^{-1}).
\end{aligned}
$$

3. Since $\partial_v \log\det S_m^{(l)} = \mathrm{tr}(\partial_v S_m^{(l)}(S_m^{(l)})^{-1})$, we have

$$
b_n^{-3/4}\partial_\sigma\partial_v H_n^{(l)}(\sigma, v) = -\frac{1}{2}b_n^{-3/4}\sum_m\{\mathrm{tr}(\partial_\sigma\partial_v(S_m^{(l)})^{-1}S_{m,*}^{(l)}) + \partial_\sigma\partial_v\log\det S_m^{(l)}\} + o_p(1) = o_p(1).
$$

Sobolev's inequality and similar estimates for $\partial_\sigma \partial_v^2$ and $\partial_\sigma^2 \partial_v$ yield the results. □

The following lemma completes the proof of the LAN properties of "subexperiment" and "superexperiment".

Lemma 8.3. *Assume* (A1″) *and* (A2). *Then* $(b_n^{-1/4}\partial_\sigma H_n^{(l)}(\sigma_*, v_*)^\top, b_n^{-1/2} \times \partial_v H_n^{(l)}(\sigma_*, v_*)^\top)^\top \to^{s\text{-}\mathcal{L}} \operatorname{diag}(\Gamma_1^{1/2}, \Gamma_2^{1/2})\tilde{\mathcal{N}}$ *for* $l = 0, 2$, *where* $\tilde{\mathcal{N}}$ *is a* $(d+2)$*-dimensional normal random variable independent of* $\mathcal{F}$.

Proof. Let $Z_{m,*}^{(l)} = Z_m^{(l)}(\sigma_*)$, $\tilde{\mathcal{X}}_m^{n,1} = -b_n^{-1/4}\bar{E}_m[(Z_{m,*}^{(l)})^\top \partial_\sigma (S_{m,*}^{(l)})^{-1} Z_{m,*}^{(l)}]/2$, $\tilde{\mathcal{X}}_m^{n,2} = -b_n^{-1/2} \times \bar{E}_m[(Z_{m,*}^{(l)})^\top \partial_v (S_{m,*}^{(l)})^{-1} Z_{m,*}^{(l)}]/2$, and $\tilde{\mathcal{X}}_m^n = ((\tilde{\mathcal{X}}_m^{n,1})^\top, (\tilde{\mathcal{X}}_m^{n,2})^\top)^\top$, then we have

$$\begin{aligned}\sum_{m=1}^{[\ell_n t/T]} E_m\bigl[|\tilde{\mathcal{X}}_m^n|^2 1_{\{|\tilde{\mathcal{X}}_m^n|>\varepsilon\}}\bigr] &\le \frac{C}{\varepsilon^2} b_n^{-1} \sum_m E_m\bigl[\bigl|(Z_{m,*}^{(l)})^\top \partial_\sigma (S_{m,*}^{(l)})^{-1} Z_{m,*}^{(l)}\bigr|^4\bigr] \\ &\quad + \frac{C}{\varepsilon^2} b_n^{-2} \sum_m E_m\bigl[\bigl|\bar{E}_m\bigl[(Z_{m,*}^{(l)})^\top \partial_v (S_{m,*}^{(l)})^{-1} Z_{m,*}^{(l)}\bigr]\bigr|^4\bigr].\end{aligned}$$

We can see that the first term on the right-hand side converges to 0 by Lemma 8.1 and a similar estimate to Point 1 of Lemma 4.3. We show the second term on the right-hand side converges to 0 in probability.

Let $Z_{2,m}^{(0)}$ be similarly defined to $\tilde{Z}_{2,m}$ in (A.4). Let $Z_{1,m,*}^{(l)} = Z_m^{(l)}(\sigma_*) - Z_{2,m}^{(l)}$ and $S_{1,m,*}^{(l)} = E_m[Z_{1,m,*}^{(l)}(Z_{1,m,*}^{(l)})^\top]$. Then Lemmas 8.1 and a similar estimate to (A.5) yield

$$\begin{aligned}E_m\bigl[\bigl|(Z_{1,m,*}^{(l)})^\top \partial_v (S_{m,*}^{(l)})^{-1} Z_{2,m}^{(l)}\bigr|^4\bigr] &\le C\bigl|E_m\bigl[\bigl((Z_{2,m}^{(l)})^\top \partial_v (S_{m,*}^{(l)})^{-1} S_{1,m,*}^{(l)} \partial_v (S_{m,*}^{(l)})^{-1} Z_{2,m}^{(l)}\bigr)^2\bigr]\bigr| \\ &= \bar{R}_n\bigl(b_n^{-1}k_n^2\bigr),\end{aligned}$$

and $E_m[|(Z_{1,m,*}^{(l)})^\top \partial_v (S_{m,*}^{(l)})^{-1} Z_{1,m,*}^{(l)}|^4] \le \bar{R}_n(b_n^{-4})\sum_{j=1}^4 |\operatorname{tr}((S_{m,*}^{(l)})^{-1})^j| = \bar{R}_n(b_n^{-2}k_n^4)$. Moreover, Lemma 8.1 yields

$$b_n^{-2}\sum_m E_m\bigl[\bar{E}_m\bigl[(Z_{2,m}^{(l)})^\top \partial_v (S_{m,*}^{(l)})^{-1} Z_{2,m}^{(l)}\bigr]^4\bigr] = \bar{R}_n\bigl(b_n^{-2}\ell_n k_n^2\bigr) \xrightarrow{p} 0,$$

and therefore, we obtain

$$\sum_{m=1}^{[\ell_n t/T]} E_m\bigl[|\tilde{\mathcal{X}}_m^n|^2 1_{\{|\tilde{\mathcal{X}}_m^n|>\varepsilon\}}\bigr] \xrightarrow{p} 0.$$

Similarly to the proof of Proposition 7.2, we have

$$\sum_{m=1}^{[\ell_n t/T]} E_m\bigl[\tilde{\mathcal{X}}_m^n (N_{s_m} - N_{s_{m-1}})\bigr] \to^p 0 \quad \text{and} \quad \sum_{m=1}^{[\ell_n t/T]} E_m\bigl[\tilde{\mathcal{X}}_m^n (W_{s_m} - W_{s_{m-1}})\bigr] = 0$$

for any bounded $\mathcal{G}_t$-martingale N orthogonal to $(W_t)_t$.

Therefore, by Theorem 3.2 in Jacod [18], it is sufficient to show that

$$\sum_{m=1}^{[\ell_n t/T]} E_m\big[\tilde{\mathcal{X}}_m^n\big(\tilde{\mathcal{X}}_m^n\big)^\top\big] \xrightarrow{p} \operatorname{diag}\big(-\partial_\sigma^2\mathcal{Y}_1(\sigma_*,t), -\partial_v^2\mathcal{Y}_2(v_*,t)\big),$$

where $\mathcal{Y}_2(v,t) = -\int_0^t \sum_{j=1}^2 a_s^j\{(v_{j,*}/v_j) - 1 + \log(v_j/v_{j,*})\}\,ds/2$.

Then we obtain the desired results by

$$\sum_{m=1}^{[\ell_n t/T]} E_m\big[\tilde{\mathcal{X}}_m^{n,1}\big(\tilde{\mathcal{X}}_m^{n,2}\big)^\top\big] = \frac{b_n^{-3/4}}{2}\sum_{m=1}^{[\ell_n t/T]} \operatorname{tr}\big(\partial_\sigma S_{m,*}^{(l)}\big(S_{m,*}^{(l)}\big)^{-1}\partial_v S_{m,*}^{(l)}\big(S_{m,*}^{(l)}\big)^{-1}\big) \xrightarrow{p} 0,$$

$$\begin{aligned}
&\sum_{m=1}^{[\ell_n t/T]} E_m\big[\tilde{\mathcal{X}}_m^{n,1}\big(\tilde{\mathcal{X}}_m^{n,1}\big)^\top\big] \\
&= \frac{b_n^{-1/2}}{2}\sum_{m=1}^{[\ell_n t/T]} \operatorname{tr}\big(\big(\partial_\sigma S_{m,*}^{(l)}\big(S_{m,*}^{(l)}\big)^{-1}\big)^2\big) \\
&= \frac{b_n^{-1/2}}{2}\sum_{m=1}^{[\ell_n t/T]} E_m\big[\partial_\sigma^2\big\{\big(Z_{m,*}^{(l)}\big)^\top\big(S_m^{(l)}\big)^{-1}Z_{m,*}^{(l)} + \log\det S_m^{(l)}\big\}\big]\Big|_{(\sigma,v)=(\sigma_*,v_*)} \\
&\xrightarrow{p} -\partial_\sigma^2\mathcal{Y}_1(\sigma_*,t),
\end{aligned}$$

and similarly

$$\begin{aligned}
&\sum_{m=1}^{[\ell_n t/T]} E_m\big[\tilde{\mathcal{X}}_m^{n,2}\big(\tilde{\mathcal{X}}_m^{n,2}\big)^\top\big] \\
&= \frac{b_n^{-1}}{2}\sum_{m=1}^{[\ell_n t/T]} E_m\big[\partial_v^2\big\{\big(Z_{m,*}^{(l)}\big)^\top\big(S_m^{(l)}\big)^{-1}Z_{m,*}^{(l)} + \log\det S_m^{(l)}\big\}\big]\Big|_{(\sigma,v)=(\sigma_*,v_*)} \\
&\xrightarrow{p} -\partial_v^2\mathcal{Y}_2(v_*,t).
\end{aligned}$$

□

Proof of Theorem 2.2. Let $\mathbf{U}(u) = \exp(u^\top\Gamma^{1/2}\tilde{\mathcal{N}} - u^\top\Gamma u/2)$ for $u \in \mathbb{R}^{d+2}$. Let

$$Z^{(1)}(\sigma) = \left(\varepsilon_0^{n,k}, \left(\tilde{Y}_i^k - \tilde{Y}_{i-1}^k - \int_{S_{i-1}^{n,k}}^{S_i^{n,k}} \mu^k(t,\sigma)\,dt\right)_{i=1}^{\mathbf{J}_{k,n}}\right)_{k=1,2},$$

and $S^{(1)}(\sigma,v)$ be a symmetric matrix of size $\mathbf{J}_{1,n} + \mathbf{J}_{2,n} + 2$ defined by $(S^{(1)}(\sigma,v))_{11} = v_1$, $(S^{(1)}(\sigma,v))_{\mathbf{J}_{1,n}+2,\mathbf{J}_{1,n}+2} = v_2$,

$$\big(S^{(1)}(\sigma,v)\big)_{ij} = \operatorname{diag}\big(v_1 M(\mathbf{J}_{1,n}+1), v_2 M(\mathbf{J}_{2,n}+1)\big)_{ij} \qquad \text{if } i \neq j \text{ and } \{i,j\} \cap \{1, \mathbf{J}_{1,n}+2\} \neq \varnothing,$$

$$\big(S^{(1)}(\sigma,v)\big)_{ij} = \int_{S^{n,1}_{i-2}}^{S^{n,1}_{i-1}} \big|b^1_t\big|^2\,dt\,\delta_{ij} + v_1 M(\mathbf{J}_{1,n}+1)_{ij} \qquad \text{if } 2\le i,\, j\le \mathbf{J}_{1,n}+1,$$

$$\big(S^{(1)}(\sigma,v)\big)_{ij} = \int_{S^{n,2}_{i'-2}}^{S^{n,2}_{i'-1}} \big|b^2_t\big|^2\,dt\,\delta_{ij} + v_2 M(\mathbf{J}_{2,n}+1)_{i'j'} \qquad \text{if } 2\le i',\, j'\le \mathbf{J}_{2,n}+1,$$

$$\big(S^{(1)}(\sigma,v)\big)_{ij} = \int_{[S^{n,1}_{i-2},S^{n,1}_{i-1})\cap[S^{n,2}_{j'-2},S^{n,2}_{j'-1})} b^1_t\cdot b^2_t\,dt \qquad \text{if } 2\le i\le \mathbf{J}_{1,n}+1 \text{ and } 2\le j'\le \mathbf{J}_{2,n}+1,$$

where $i'=i-\mathbf{J}_{1,n}-1$ and $j'=j-\mathbf{J}_{1,n}-1$. Then we have (8.2) for $l=1$ with

$$H^{(1)}_n(\sigma,v) = -\frac{1}{2}\big(Z^{(1)}(\sigma)\big)^\top\big(S^{(1)}(\sigma,v)\big)^{-1}Z^{(1)}(\sigma) - \frac{1}{2}\log\det S^{(1)}(\sigma,v).$$

Moreover, Theorem 8.1 and Lemmas 8.2 and 8.3 yield

$$\begin{aligned}&\big(H^{(1)}_n(\sigma_{u^{(1)}},v_{u^{(1)}}) - H^{(1)}_n(\sigma_*,v_*),\dots,H^{(1)}_n(\sigma_{u^{(k)}},v_{u^{(k)}}) - H^{(1)}_n(\sigma_*,v_*)\big)\\ &\quad\xrightarrow{d}\big(\log\mathbf{U}\big(u^{(1)}\big),\dots,\log\mathbf{U}\big(u^{(k)}\big)\big)\end{aligned} \tag{8.4}$$

as $n\to\infty$ for $u^{(1)},\dots,u^{(k)}\in\mathbb{R}^{d+2}$.

Furthermore, similar estimates to the proof of Lemma 8.2 yield $\sup_{\sigma,v}|b_n^{-3/4}\partial_\sigma\partial_v H^{(1)}_n(\sigma,v)|\to^p 0$, $\sup_\sigma|b_n^{-3/4}\partial_\sigma^3 H^{(1)}_n(\sigma,v_*)|\to^p 0$, and $\sup_v|b_n^{-3/2}\partial_v^3 H^{(1)}_n(\sigma_*,v)|\to^p 0$. Therefore we obtain

$$H^{(1)}_n(\sigma_u,v_u) - H^{(1)}_n(\sigma_*,v_*) - \big(u\cdot\mathbf{V}_{1,n} - u^\top\mathbf{V}_{2,n}u/2\big)\xrightarrow{p}0 \tag{8.5}$$

as $n\to\infty$ for any $u\in\mathbb{R}^{d+2}$, where $\mathbf{V}_{1,n}=(b_n^{-1/4}\partial_\sigma H^{(1)}_n(\sigma_*,v_*),b_n^{-1/2}\partial_v H^{(1)}_n(\sigma_*,v_*))$ and $\mathbf{V}_{2,n}=-\mathrm{diag}(b_n^{-1/2}\partial_\sigma^2 H^{(1)}_n(\sigma_*,v_*),b_n^{-1}\partial_v^2 H^{(1)}_n(\sigma_*,v_*))$. (8.4) and (8.5) yield $\mathbf{V}_{1,n}\to^d\Gamma^{1/2}\tilde{\mathcal{N}}$ and $\mathbf{V}_{2,n}\to^p\Gamma$, and therefore we obtain the LAN property of the original experiment with $\Gamma_n=\mathbf{V}_{2,n}$ and $\mathcal{N}_n=\Gamma^{-1/2}\mathbf{V}_{1,n}$ by (8.1). □

9. Proof of the results in Section 2.4

In this final section, we complete the proof of remaining results in Section 2. Proposition 2.2 is proven by the scheme of Yoshida [29,30]. Proposition 6.1 and moment estimates in Lemmas 4.4 and 5.2 enable us to check the assumptions of Theorem 2 in [30]. Then the results on convergence of moments and the Bayes-type estimator are obtained by Proposition 2.2.

Outline of the proof of Proposition 2.2. We apply Theorem 2 in Yoshida [30]. It is sufficient to prove the following five conditions for any $L>0$ with some positive constant δ_1 and δ_2:

1. There exists $C_L>0$ such that $P[\inf_{\sigma\neq\sigma_*}(-\mathcal{Y}_1(\sigma)/|\sigma-\sigma_*|^2)\le r^{-1}]\le C_L/r^L$ and $P[\{r^{-1}|u|^2\le u^\top\Gamma_1 u/4 \text{ for any } u\in\mathbb{R}^d\}^c]\le C_L/r^L$ for any $r>0$.

2. $\sup_n E[(b_n^{-1/4}|\partial_\sigma H_n(\sigma_*, \hat{v}_n)|)^L] < \infty$.
3. $\sup_n E[(b_n^{\delta_1} \sup_\sigma |b_n^{-1/2}(H_n(\sigma, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)) - \mathcal{Y}_1(\sigma)|)^L] < \infty$.
4. $\sup_n E[(b_n^{-1/2} \sup_\sigma |\partial_\sigma^3 H_n(\sigma, \hat{v}_n)|)^L] < \infty$.
5. $\sup_n E[(b_n^{\delta_2}|b_n^{-1/2}\partial_\sigma^2 H_n(\sigma_*, \hat{v}_n) + \Gamma_1|)^L] < \infty$.

By Taylor's formula for $\mathcal{Y}_1(\sigma)$ and relations $\mathcal{Y}_1(\sigma_*) = \partial_\sigma \mathcal{Y}_1(\sigma_*) = 0$, we obtain $\inf_{\sigma \neq \sigma_*}(-\mathcal{Y}_1(\sigma)/|\sigma - \sigma_*|^2) \leq \inf_{u \in \mathbb{R}^d \setminus \{0\}} u^\top \Gamma_1 u/(2|u|^2)$. Then (B3) and the proof of Proposition 6.1 yield point 1. By Lemmas 4.4 and 5.2, Remark 5.2, and a similar argument to the proof of Proposition 2.1, we obtain 3–5 and $\sup_n E[(b_n^{-1/4}|\partial_\sigma H_n(\sigma_*, \hat{v}_n) - \partial_\sigma \tilde{H}_n(\sigma_*, v_*)|)^L] < \infty$. Moreover, by the Burkholder–Davis–Gundy inequality, we obtain

$$\begin{aligned}
&E\big[\big|b_n^{-1/4}\partial_\sigma \tilde{H}_n(\sigma_*, v_*)\big|^L\big] \\
&\quad = E\bigg[\bigg|\frac{b_n^{-1/4}}{2}\sum_m \bar{E}_m\big[\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big]\bigg|^L\bigg] \leq CE\bigg[\bigg|b_n^{-\frac{1}{2}}\sum_m \bar{E}_m\big[\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big]^2\bigg|^{L/2}\bigg] \\
&\quad \leq CE\bigg[\bigg|b_n^{-\frac{1}{2}}\sum_m E_m\big[\bar{E}_m\big[\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big]^2\big]\bigg|^{L/2}\bigg] \\
&\qquad + CE\bigg[\bigg|b_n^{-1}\sum_m \bar{E}_m\big[\bar{E}_m\big[\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big]^2\big]^2\bigg|^{L/4}\bigg] \\
&\quad \leq CE\bigg[\bigg|b_n^{-\frac{1}{2}}\sum_m \mathrm{tr}\big(\big(\partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{S}_{m,*}\big)^2\big)\bigg|^{L/2}\bigg] \\
&\qquad + CE\bigg[\bigg|b_n^{-1}\sum_m E_m\big[\big(\tilde{Z}_m^\top \partial_\sigma \tilde{S}_{m,*}^{-1}\tilde{Z}_m\big)^L\big]^{4/L}\bigg|^{L/4}\bigg] + O(1) \\
&\quad = O\big(\big(b_n^{-1/2}\ell_n b_n^{1/2}\ell_n^{-1}\big)^{L/2}\big) + E\big[\bar{R}_n\big(\big(b_n^{-5}k_n^8\ell_n\big)^{L/4}\big)\big] + O(1) \\
&\quad = O(1)
\end{aligned}$$

for $L \geq 4$, which implies point 2. □

Proof of Theorem 2.3. We extend $\mathbf{Z}_n(u)$ to a continuous function on $\mathbb{R}^d$ satisfying $\lim_{|u| \to \infty} \mathbf{Z}_n(u) = 0$ with the supremum norm of the extended function the same as for the original one. Then by Theorem 5 and Remark 5 in Yoshida [30], it is sufficient to show $\limsup_{n \to \infty} E[|b_n^{1/4}(\hat{\sigma}_n - \sigma_*)|^p] < \infty$ for any $p > 0$ and $\mathbf{Z}_n \to^{s\text{-}\mathcal{L}} \mathbf{Z}$ in $C(B(R))$ as $n \to \infty$ for any $R > 0$, where $\mathbf{Z}(u) = \exp(\mathcal{N} \cdot u - u^\top \Gamma_1 u/2)$ and $B(R) = \{u; |u| \leq R\}$.

By Lemma 4.4 and a similar argument to the proof of Proposition 2.1, we have

$$\lim_{N \to \infty} \sup_{n \geq N} E\Big[\sup_{u \in B(R)} \big|\partial_u \log \mathbf{Z}_n(u)\big|\Big] < \infty.$$

Then Propositions 2.1 and 7.2 and tightness criterion in C space in Billingsley [6] yield $\log \mathbf{Z}_n \to^{s\text{-}\mathcal{L}} \log \mathbf{Z}$ in $C(B(R))$. Then (2.12) completes the proof. □

Proof of Theorem 2.4. By Theorem 10 in Yoshida [30], it is sufficient to show

$$\sup_n E\biggl[\biggl(\int_{U_n} \mathbf{Z}_n(u)\pi\bigl(\sigma_* + b_n^{-1/4}u\bigr)\,du\biggr)^{-1}\biggr] < \infty. \tag{9.1}$$

Similarly to the proof of Proposition 2.2, there exists $\delta > 0$ such that $\sup_n E[|H_n(\sigma_* + b_n^{-1/4}u, \hat{v}_n) - H_n(\sigma_*, \hat{v}_n)|^p] \le C_p|u|^p$ for any $u \in U(\delta)$, where $U(\delta) = \{u \in \mathbb{R}^d; |u_i| \le \delta (i = 1, \dots, d)\}$. Then we have (9.1) by Lemma 2 in [30]. □

Appendix

A.1. Results from linear algebra

Lemma A.1. *Let A and B be matrices, with A nonnegative definite and symmetric. Then*

$$\bigl|\mathrm{tr}(AB)\bigr| \le \mathrm{tr}(A)\|B\|.$$

Lemma A.2. *Let $l \in \mathbb{N}$, A^j and B^j be real-valued matrices and $\{\lambda_k^j\}_k$ be eigenvalues of A^j for $1 \le j \le l$. Assume that A^j is symmetric and all the elements of B^j are nonnegative for $1 \le j \le l$. Then*

$$\sum_{i_1,\dots,i_{2l}} \prod_{j=1}^{l} \bigl(|A^j_{i_{2j-1},i_{2j}}|B^j_{i_{2j},i_{2j+1}}\bigr) \le \prod_{j=1}^{l}\biggl(\|B^j\|\sum_k |\lambda_k^j|\biggr),$$

where $i_{2l+1} = i_1$.

Proof. Let U^j be an orthogonal matrix such that $A^j = (U^j)^\top \operatorname{diag}((\lambda_k^j)_k)U^j$. Then

$$\begin{aligned}
\sum_{i_1,\dots,i_{2l}} \prod_{j=1}^{l} \bigl(|A^j_{i_{2j-1},i_{2j}}|B^j_{i_{2j},i_{2j+1}}\bigr) &\le \sum_{k_1,\dots,k_l}\sum_{i_1,\dots,i_{2l}} \prod_{j=1}^{l}\bigl(|\lambda^j_{k_j}||U^j_{k_j,i_{2j-1}}||U^j_{k_j,i_{2j}}|B^j_{i_{2j},i_{2j+1}}\bigr) \\
&\le \sum_{k_1,\dots,k_l}\prod_{j=1}^{l}\biggl\{|\lambda^j_{k_j}|\|B^j\|\sum_i (U^j_{k_j,i})^2\biggr\} \\
&= \prod_{j=1}^{l}\biggl(\|B^j\|\sum_k|\lambda_k^j|\biggr).
\end{aligned}$$

□

Lemma A.3. *Let A be a symmetric matrix with $\|A\| < 1$. Then $\log\det(\mathcal{E} + A) = \sum_{p=1}^{\infty}(-1)^{p-1}p^{-1}\,\mathrm{tr}(A^p)$.*

Proof. Let $\{\lambda_j\}_{j=1}^k$ be eigenvalues of A. Then $\sup_j |\lambda_j| = \|A\| < 1$, and hence

$$\log\det(\mathcal{E} + A) = \sum_j \log(1 + \lambda_j) = \sum_j \sum_{p=1}^{\infty} (-1)^{p-1} p^{-1} \lambda_j^p = \sum_{p=1}^{\infty} (-1)^{p-1} p^{-1} \operatorname{tr}(A^p).$$

□

Lemma A.4. *Let A and B be symmetric, positive definite matrices. Assume that $v^\top A v \geq v^\top B v$ for any vector v. Then $\|A^{-1}\| \leq \|B^{-1}\|$ and $\|A^{-1/2}\| \leq \|B^{-1/2}\|$.*

Proof. Let $(\lambda_j^A)_j$ and $(\lambda_j^B)_j$ be eigenvalues of A and B, respectively. Then for any unit vector v, there exists an orthogonal matrix U such that

$$\sum_j \lambda_j^A v_j^2 \geq \sum_j \lambda_j^B (Uv)_j^2 \geq \inf_j \lambda_j^B.$$

Therefore, we obtain $\|A^{-1}\|^{-1} = \inf_j \lambda_j^A \geq \inf_j \lambda_j^B = \|B^{-1}\|^{-1}$ and $\|A^{-1/2}\|^{-1} = \inf_j (\lambda_j^A)^{1/2} \geq \inf_j (\lambda_j^B)^{1/2} = \|B^{-1/2}\|^{-1}$. □

Lemma A.5. *Let B be a symmetric, positive definite matrix and A be a symmetric, nonnegative definite matrix. Then* $\operatorname{tr}(AB) \geq \operatorname{tr}(A)\|B^{-1}\|^{-1}$.

Proof. Let $\{\lambda_j^A\}_j$ and $\{\lambda_j^B\}_j$ be eigenvalues of A and B, respectively, and U be an orthogonal matrix satisfying $UAU^\top = \operatorname{diag}((\lambda_j^A)_j)$. Then since $(UBU^\top)_{jj} \geq \inf_j \lambda_j^B = \|B^{-1}\|^{-1}$, we obtain

$$\operatorname{tr}(AB) = \sum_j \lambda_j^A \left(UBU^\top\right)_{jj} \geq \sum_j \lambda_j^A \left\|B^{-1}\right\|^{-1} = \operatorname{tr}(A)\left\|B^{-1}\right\|^{-1}.$$

□

Lemma A.6. *Let $\eta > 0$ and A be a symmetric matrix. Assume that $\mathcal{E} + A$ is positive definite and $\|\mathcal{E} + A\| \leq \eta$. Then* $\operatorname{tr}(A) - \log\det(\mathcal{E} + A) \geq \operatorname{tr}(A^2)/(2(\eta \vee 1))$.

Proof. We easily obtain the desired results by using the fact that $\log\det(\mathcal{E}+A) = \sum_k \log(1+\lambda_k)$ and that $x - x^2/(2(\eta \vee 1)) \geq \log(1+x)$ for $-1 < x \leq \eta - 1$, where $(\lambda_j)_j$ are eigenvalues of A. □

Lemma A.7. *Let A be a symmetric matrix, B a matrix of suitable size and $(\lambda_j)_j$ eigenvalues of $B^\top AB$. Then*

1. $|(B^\top AB)_{ii}| \leq \|A\|(B^\top B)_{ii}$ *for any i.*
2. $\sum_j |\lambda_j| \leq \|A\| \operatorname{tr}(B^\top B)$.

Proof. 1. Let U be an orthogonal matrix and let $\{\lambda_j'\}_j$ be eigenvalues of A such that $U^\top AU = \operatorname{diag}((\lambda_j')_j)$. Then we obtain

$$\left|\left(B^\top AB\right)_{ii}\right| = \left|\sum_j \lambda_j' \left(\left(U^\top B\right)_{ji}\right)^2\right| \leq \|A\| \sum_j \left(\left(U^\top B\right)_{ji}\right)^2 = \|A\|\left(B^\top B\right)_{ii}.$$

2. There exists an orthogonal matrix V such that $\lambda_j = (V^\top B^\top ABV)_{jj}$ for any j. Then

$$\sum_j |\lambda_j| = \sum_j \big|\big(V^\top B^\top ABV\big)_{jj}\big| \le \|A\| \sum_j \big(V^\top B^\top BV\big)_{jj} = \|A\| \operatorname{tr}\big(B^\top B\big)$$

by 1. $\square$

A.2. Proof of (2.10)

Proof of (2.10). Let $\tilde{a}_t^3 = \sqrt{\tilde{a}_t^1 \tilde{a}_t^2}$, $B_t = \sqrt{\det(b_t b_t^\top)}$, $\Upsilon_t = (\tilde{a}_t^1 |b_t^1|^2 + \tilde{a}_t^2 |b_t^2| + 2\tilde{a}_t^3 B_t)^{1/2}$, and $\psi(x,y) = (xy^\top + yx^\top)/2$ for vectors x, y. Then (2.8) yields

$$\mathcal{Y}_1(\sigma) = \int_0^T \left\{ \frac{\sum_{j=1}^2 (|b_t^j|^2 - |b_{t,*}^j|^2)(\tilde{a}_t^3 |b_t^{3-j}|^2 + \tilde{a}_t^j B_t) - 2\tilde{a}_t^3 (b_t^1 \cdot b_t^2 - b_{t,*}^1 \cdot b_{t,*}^2) b_t^1 \cdot b_t^2}{4 B_t \Upsilon_t} \right.$$
$$\left. - \frac{\Upsilon_t - \Upsilon_t^{(0)}}{2} \right\} dt.$$

Hence we obtain

$$\partial_\sigma^2 \mathcal{Y}_1(\sigma_*) = \int_0^T \left\{ \frac{\sum_{j=1}^2 \partial_\sigma^2 |b_t^j|^2 (\tilde{a}_t^3 |b_t^{3-j}|^2 + \tilde{a}_t^j B_t) - 2\tilde{a}_t^3 \partial_\sigma^2 b_t^1 \cdot b_t^2 b_t^1 \cdot b_t^2}{4 B_t \Upsilon_t} \Big|_{\sigma=\sigma_*} \right.$$
$$+ \sum_{j=1}^2 \psi\left(\partial_\sigma |b_t^j|^2, \partial_\sigma \frac{\tilde{a}_t^3 |b_t^{3-j}|^2 + \tilde{a}_t^j B_t}{2 B_t \Upsilon_t} \right)\Big|_{\sigma=\sigma_*} - \psi\left(\partial_\sigma b_t^1 \cdot b_t^2, \tilde{a}_t^3 \partial_\sigma \frac{b_t^1 \cdot b_t^2}{B_t \Upsilon_t} \right)\Big|_{\sigma=\sigma_*}$$
$$\left. - \frac{\partial_\sigma^2 \Upsilon_t^2}{4 \Upsilon_t}\Big|_{\sigma=\sigma_*} + \frac{\Upsilon_t^{(1)} (\Upsilon_t^{(1)})^\top}{2 \Upsilon_t^{(0)}} \right\} dt.$$

Terms involving ∂_σ^2 in the integrand of right-hand side are rewritten as

$$\left(\tilde{a}_t^3 \frac{\sum_{j=1}^2 \partial_\sigma^2 |b_t^j|^2 |b_t^{3-j}|^2 - 2\partial_\sigma^2 b_t^1 \cdot b_t^2 b_t^1 \cdot b_t^2}{4 B_t \Upsilon_t} - \tilde{a}_t^3 \frac{2\partial_\sigma^2 B_t}{4 \Upsilon_t} \right)\Big|_{\sigma=\sigma_*}$$
$$= \frac{\tilde{a}_t^3}{4 \Upsilon_t^{(0)}} \left\{ \frac{\partial_\sigma^2 |b_t^1|^2 |b_t^2|^2 + |b_t^1|^2 \partial_\sigma^2 |b_t^2|^2 - 2 b_t^1 \cdot b_t^2 \partial_\sigma^2 b_t^1 \cdot b_t^2}{B_t} - \frac{\partial_\sigma^2 B_t^2}{B_t} + \frac{2 B_{4,t}^{(1)} (B_{4,t}^{(1)})^\top}{B_{4,t}^{(0)}} \right\}\Big|_{\sigma=\sigma_*}$$
$$= \frac{\tilde{a}_t^3}{4 \Upsilon_t^{(0)}} \left\{ -\frac{2\psi(B_{1,t}^{(1)}, B_{2,t}^{(1)}) - 2 B_{3,t}^{(1)} (B_{3,t}^{(1)})^\top}{B_{4,t}^{(0)}} + \frac{2 B_{4,t}^{(1)} (B_{4,t}^{(1)})^\top}{B_{4,t}^{(0)}} \right\}.$$

Other terms are rewritten as

$$\sum_{j=1}^{2}\psi\left(\partial_\sigma|b_t^j|^2,\partial_\sigma\frac{\tilde{a}_t^3|b_t^{3-j}|^2+\tilde{a}_t^jB_t}{2B_t\Upsilon_t}\right)\Big|_{\sigma=\sigma_*}-\psi\left(\partial_\sigma b_t^1\cdot b_t^2,\tilde{a}_t^3\partial_\sigma\frac{b_t^1\cdot b_t^2}{B_t\Upsilon_t}\right)\Big|_{\sigma=\sigma_*}$$
$$+\frac{\Upsilon_t^{(1)}(\Upsilon_t^{(1)})^\top}{2\Upsilon_t^{(0)}}$$
$$=\sum_{j=1}^{2}\psi\left(B_{j,t}^{(1)},\frac{\tilde{a}_t^3B_{3-j,t}^{(1)}}{2B_{4,t}^{(0)}\Upsilon_t^{(0)}}-\frac{\tilde{a}_t^3B_{3-j,t}^{(0)}B_{4,t}^{(1)}}{2(B_{4,t}^{(0)})^2\Upsilon_t^{(0)}}-\frac{(\tilde{a}_t^3B_{3-j,t}^{(0)}+\tilde{a}_t^jB_{4,t}^{(0)})\Upsilon_t^{(1)}}{2B_{4,t}^{(0)}(\Upsilon_t^{(0)})^2}\right)$$
$$-\tilde{a}_t^3\psi\left(B_{3,t}^{(1)},\frac{B_{3,t}^{(1)}}{B_{4,t}^{(0)}\Upsilon_t^{(0)}}-\frac{B_{3,t}^{(0)}B_{4,t}^{(1)}}{(B_{4,t}^{(0)})^2\Upsilon_t^{(0)}}-\frac{B_{3,t}^{(0)}\Upsilon_t^{(1)}}{B_{4,t}^{(0)}(\Upsilon_t^{(0)})^2}\right)+\frac{\Upsilon_t^{(1)}(\Upsilon_t^{(1)})^\top}{2\Upsilon_t^{(0)}}$$
$$=-\tilde{a}_t^3\psi\left(B_{1,t}^{(1)}B_{2,t}^{(0)}+B_{2,t}^{(1)}B_{1,t}^{(0)}-2B_{3,t}^{(0)}B_{3,t}^{(1)},\frac{B_{4,t}^{(1)}}{2(B_{4,t}^{(0)})^2\Upsilon_t^{(0)}}+\frac{\Upsilon_t^{(1)}}{2B_{4,t}^{(0)}(\Upsilon_t^{(0)})^2}\right)$$
$$+\tilde{a}_t^3\frac{\psi(B_{1,t}^{(1)},B_{2,t}^{(1)})-B_{3,t}^{(1)}(B_{3,t}^{(1)})^\top}{B_{4,t}^{(0)}\Upsilon_t^{(0)}}-\psi\left(\tilde{a}_t^1B_{1,t}^{(1)}+\tilde{a}_t^2B_{2,t}^{(1)},\frac{\Upsilon_t^{(1)}}{2(\Upsilon_t^{(0)})^2}\right)+\frac{\Upsilon_t^{(1)}(\Upsilon_t^{(1)})^\top}{2\Upsilon_t^{(0)}}$$
$$=\tilde{a}_t^3\frac{\psi(B_{1,t}^{(1)},B_{2,t}^{(1)})-B_{3,t}^{(1)}(B_{3,t}^{(1)})^\top}{B_{4,t}^{(0)}\Upsilon_t^{(0)}}-\frac{\tilde{a}_t^3B_{4,t}^{(1)}(B_{4,t}^{(1)})^\top}{B_{4,t}^{(0)}\Upsilon_t^{(0)}}-\frac{\Upsilon_t^{(1)}(\Upsilon_t^{(1)})^\top}{2\Upsilon_t^{(0)}}.$$

Here we have used equations $B_{1,t}^{(1)}B_{2,t}^{(0)}+B_{2,t}^{(1)}B_{1,t}^{(0)}-2B_{3,t}^{(0)}B_{3,t}^{(1)}=2B_{4,t}^{(0)}B_{4,t}^{(1)}$ and $\tilde{a}_t^1B_{1,t}^{(1)}+\tilde{a}_t^2B_{2,t}^{(1)}+2\tilde{a}_t^3B_{4,t}^{(1)}=2\Upsilon_t^{(0)}\Upsilon_t^{(1)}$. Therefore, we obtain (2.10). □

A.3. Proof of Lemma 4.3

Let $\mathbf{A}_m$ be a $(k_m^1+k_m^2)\times(k_m^1+k_m^2)$ matrix with elements $(\mathbf{A}_m)_{ij}=1_{i\geq j}1_{\{i\leq k_m^1 \text{ or } j>k_m^1\}}$, let $\mathbf{1}$ be a matrix with all elements equal to 1 and $M_{m,*}=M_m(v_*)$.

Lemma A.8. *Let $n\in\mathbb{N}$, $1\leq m\leq\ell_n$, $q,q'\in\mathbb{N}$ such that $q'\leq 2q$, $A_m:\{1,\dots,k_m^1+k_m^2\}^{q'}\to\{0,1\}$ be a random map, $\iota:\{1,\dots,q'\}\to\{1,\dots,2q\}$ be an injection, and $\mathbf{S}$ be a symmetric, $\mathcal{G}_{s_{m-1}}$-measurable random matrix of size $k_m^1+k_m^2$ satisfying $\|M_{m,*}\mathbf{S}M_{m,*}\|\leq b_n^{-1}$. Let $\mathcal{K}=\sum_{j_1,\dots,j_{q'}=1}^{k_m^1+k_m^2}A_m(j_1,\dots,j_{q'})$. Then there exists a positive constant C that depends only on v_* and q such that*

$$\left|\sum_{i_1,\dots,i_{2q}}\prod_{j=1}^{q}\left((\mathbf{A}_m^\top)^{-1}\mathbf{S}\mathbf{A}_m^{-1}\right)_{i_{2j-1},i_{2j}}A_m(i_{\iota(1)},\dots,i_{\iota(q')})\right|\leq C\mathcal{K}(\bar{k}_n/\underline{k}_n)^{2q}b_n^{-q}\bar{k}_n^{q+q'-2[q'/2]}.$$

Proof. Let $\hat{\mathbf{S}} = (\mathbf{A}_m^\top)^{-1}\mathbf{S}\mathbf{A}_m^{-1}$ and $\tilde{\mathbf{A}}_m = M_{m,*}\mathbf{A}_m^\top$. Then a simple calculation shows that

$$\big(\tilde{\mathbf{A}}_m^{-1}\big)_{i,j} = \begin{cases} \big(k_m^{\mathbf{k}(i)}+1\big)^{-1}\big(j - k_m^1 1_{\{\mathbf{k}(i)=2\}}\big)v_{\mathbf{k}(i),*}^{-1}, & \mathbf{k}(i)=\mathbf{k}(j) \text{ and } i \geq j, \\ -\big(k_m^{\mathbf{k}(i)}+1\big)^{-1}\big(k_m^{\mathbf{k}(i)} - j + 1 + k_m^1 1_{\{\mathbf{k}(i)=2\}}\big)v_{\mathbf{k}(i),*}^{-1}, & \mathbf{k}(i)=\mathbf{k}(j) \text{ and } i < j, \\ 0, & \text{otherwise,} \end{cases}$$

where $\{\mathbf{k}(i)\}_{i=1}^{k_m^1+k_m^2}$ is defined by $\mathbf{k}(i) = 1$ for $1 \leq i \leq k_m^1$ and $\mathbf{k}(i) = 2$ for $k_m^1 < i \leq k_m^1 + k_m^2$. Therefore, we have

$$\begin{aligned} |\hat{\mathbf{S}}_{ij}| &= \big|\big(\tilde{\mathbf{A}}_m^{-1}M_{m,*}\mathbf{S}M_{m,*}\big(\tilde{\mathbf{A}}_m^\top\big)^{-1}\big)_{ij}\big| \\ &\leq \left(\sum_k\big(\big(\tilde{\mathbf{A}}_m^{-1}\big)_{i,k}\big)^2\right)^{1/2}\|M_{m,*}\mathbf{S}M_{m,*}\|\left(\sum_k\big(\big(\tilde{\mathbf{A}}_m^{-1}\big)_{j,k}\big)^2\right)^{1/2} \\ &\leq \big(v_{1,*}^{-1}+v_{2,*}^{-1}\big)b_n^{-1}\bar{k}_n. \end{aligned} \tag{A.1}$$

Similarly, we have

$$\big|(\hat{\mathbf{S}}\mathbf{1}\hat{\mathbf{S}})_{ij}\big| \leq \left(\max_i\sum_k\big(\tilde{\mathbf{A}}_m^{-1}\big)_{ik}^2\right)\|M_{m,*}\mathbf{S}M_{m,*}\|^2\big\|\big(\tilde{\mathbf{A}}_m^{-1}\big)^\top\mathbf{1}\tilde{\mathbf{A}}_m^{-1}\big\| \leq Cb_n^{-2}\bar{k}_n^2. \tag{A.2}$$

Moreover, since $\sum_i((\mathbf{A}_m M_{m,*}\mathbf{A}_m^\top)^{-1})_{ij} = v_{\mathbf{k}(j),*}^{-1}/(k_m^j+1)$, we obtain

$$\begin{aligned} \big|\mathrm{tr}(\hat{\mathbf{S}}\mathbf{1})\big| &\leq \|M_{m,*}\mathbf{S}M_{m,*}\|\sum_{i,j,k,l}\big(\big(\mathbf{A}_m M_{m,*}\mathbf{A}_m^\top\big)^{-1}\big)_{ik}\big(\mathbf{A}_m\mathbf{A}_m^\top\big)_{kl}\big(\big(\mathbf{A}_m M_{m,*}\mathbf{A}_m^\top\big)^{-1}\big)_{lj} \\ &\leq C\underline{k}_n^{-2}b_n^{-1}\sum_{k,l}\big(\mathbf{A}_m\mathbf{A}_m^\top\big)_{kl} \leq C\underline{k}_n^{-2}b_n^{-1}\bar{k}_n^3. \end{aligned} \tag{A.3}$$

If both i_{2j-1} and i_{2j} are outside the image of ι, we have

$$\sum_{i_{2j-1},i_{2j}}\hat{\mathbf{S}}_{i_{2j-1},i_{2j}}A_m(i_{\iota(1)},\ldots,i_{\iota(q')}) = \mathrm{tr}(\hat{\mathbf{S}}\mathbf{1})A_m(i_{\iota(1)},\ldots,i_{\iota(q')}).$$

Moreover, if both i_{2j-1} and i_{2k-1} are in the image of ι and neither i_{2j} nor i_{2k} is in it, then we have

$$\begin{aligned} &\sum_{i_{2j},i_{2k}}\hat{\mathbf{S}}_{i_{2j-1},i_{2j}}\hat{\mathbf{S}}_{i_{2k-1},i_{2k}}A_m(i_{\iota(1)},\ldots,i_{\iota(q')}) \\ &\quad = (\hat{\mathbf{S}}\mathbf{1}\hat{\mathbf{S}})_{i_{2j-1},i_{2k-1}}A_m(i_{\iota(1)},\ldots,i_{\iota(q')}). \end{aligned}$$

Therefore there exist $\alpha_k \in \{0,1\}$ for $1 \le k \le [q'/2]$, $0 \le s \le [(2q-q')/2]$ and a bijection $\iota' : \{1,\dots,q'\} \to \{1,\dots,q'\}$ such that $\sum_{k=1}^{[q'/2]} \alpha_k + [q'/2] + s = q - (q' - 2[q'/2])$ and

$$
\begin{aligned}
&\left| \sum_{i_1,\dots,i_{2q}} \prod_{j=1}^{q} \hat{\mathbf{S}}_{i_{2j-1},i_{2j}} A_m(i_{\iota(1)},\dots,i_{\iota(q')}) \right| \\
&\quad \le \big(Cb_n^{-1}\bar{k}_n^2\big)^{q'-2[q'/2]} \left| \sum_{j_1,\dots,j_{q'}} \prod_{k=1}^{[q'/2]} \big|\big((\hat{\mathbf{S}}\mathbf{1})^{\alpha_k}\hat{\mathbf{S}}\big)_{j_{\iota'(2k-1)},j_{\iota'(2k)}}\big| \operatorname{tr}(\hat{\mathbf{S}}\mathbf{1})^s A_m(j_1,\dots,j_{q'}) \right| \\
&\quad \le C\big(b_n^{-1}\bar{k}_n^2\big)^{q'-2[q'/2]} (\bar{k}_n/\underline{k}_n)^{2q} \big(b_n^{-1}\bar{k}_n\big)^{q-(q'-2[q'/2])} \mathcal{K} \le C\mathcal{K}(\bar{k}_n/\underline{k}_n)^{2q} b_n^{-q} \bar{k}_n^{q+q'-2[q'/2]},
\end{aligned}
$$

by (A.1)–(A.3). □

Proof of Lemma 4.3. Let $\{\tilde{\varepsilon}_{i,m}\}_{1\le i\le k_m^1+k_m^2}$ and $\{\dot{\varepsilon}_{i,m}\}_{1\le i\le k_m^1+k_m^2}$ be sequences of random variables defined by $\tilde{\varepsilon}_{i,m} = \varepsilon^{n,1}_{i+K^1_{m-1}+1}$ and $\dot{\varepsilon}_{i,m} = \varepsilon^{n,1}_{K^1_{m-1}+1}$ for $i \le k_m^1$ and $\tilde{\varepsilon}_{i,m} = \varepsilon^{n,2}_{i-k_m^1+K^2_{m-1}+1}$ and $\dot{\varepsilon}_{i,m} = \varepsilon^{n,2}_{K^2_{m-1}+1}$ for $i > k_m^1$. Moreover, let

$$
\begin{aligned}
\tilde{Z}_{1,m} &= \Big(\Big(\big(\tilde{b}^1_{m,*}\cdot (W_{S_i^{n,1}} - W_{S_{i-1}^{n,1}})\big)_{i=K^1_{m-1}+2}^{K^1_m}\Big)^\top, \Big(\big(\tilde{b}^2_{m,*}\cdot (W_{S_j^{n,2}} - W_{S_{j-1}^{n,2}})\big)_{j=K^2_{m-1}+2}^{K^2_m}\Big)^\top\Big)^\top, \\
\tilde{Z}_{2,m} &= \Big(\Big(\big(\varepsilon_i^{n,1} - \varepsilon_{i-1}^{n,1}\big)_{i=K^1_{m-1}+2}^{K^1_m}\Big)^\top, \Big(\big(\varepsilon_j^{n,2} - \varepsilon_{j-1}^{n,2}\big)_{j=K^2_{m-1}+2}^{K^2_m}\Big)^\top\Big)^\top
\end{aligned}
\tag{A.4}
$$

and $\tilde{S}_{1,m,*} = \tilde{S}_{m,*} - M_{m,*}$.

Let $\tilde{U}_{1,m,*}$ be an orthogonal matrix and let $\Lambda_{1,m,*}$ be a diagonal matrix satisfying $\tilde{U}_{1,m,*}\tilde{S}_{1,m,*}\tilde{U}_{1,m,*}^\top = \Lambda_{1,m,*}$. Then since $\tilde{Z}_{1,m}|\mathcal{G}_{s_{m-1}} \sim N(0,\tilde{S}_{1,m,*})$, we have $\tilde{U}_{1,m,*}\tilde{Z}_{1,m}|\mathcal{G}_{s_{m-1}} \sim N(0,\Lambda_{1,m,*})$. Therefore, for any $q\in\mathbb{N}$ and $1 \le j_1,\dots,j_{2q} \le k_m^1 + k_m^2$, we obtain

$$
E_m\left[\prod_{k=1}^{2q} (\tilde{U}_{1,m,*}\tilde{Z}_{1,m})_{j_k}\right] = \sum_{(l_{2k-1},l_{2k})_{k=1}^q} \prod_{k=1}^{q} (\Lambda_{1,m,*})_{l_{2k-1},l_{2k}}, \tag{A.5}
$$

where the summation on the right-hand side is over all q-pairs $(l_{2k-1},l_{2k})_{k=1}^q$ of variables $j_1,\dots,j_{2q}$.

1. Let $\phi(A,B)_{i_1,\dots,i_4} = (A_{i_1,i_2}B_{i_3,i_4} + A_{i_1,i_3}B_{i_2,i_4} + A_{i_1,i_4}B_{i_2,i_3} + A_{i_2,i_3}B_{i_1,i_4} + A_{i_2,i_4}B_{i_1,i_3} + A_{i_3,i_4}B_{i_1,i_2})/2$ for square matrices A and B of the same size, and $\delta_{i_1,\dots,i_q}$ be a $\{0,1\}$-valued function that is equal to 1 if and only if $i_1 = \cdots = i_q$. Then we have

$$
\begin{aligned}
&E_m\big[\big(\tilde{Z}_{2,m}^\top \mathbf{S}\tilde{Z}_{2,m}\big)^2\big] \\
&\quad = E_m\big[\big((\mathbf{A}_m\tilde{Z}_{2,m})^\top \hat{\mathbf{S}}(\mathbf{A}_m\tilde{Z}_{2,m})\big)^2\big] = \sum_{i_1,\dots,i_4} \hat{\mathbf{S}}_{i_1,i_2}\hat{\mathbf{S}}_{i_3,i_4} E_m\left[\prod_{j=1}^4 (\tilde{\varepsilon}_{i_j,m} - \dot{\varepsilon}_{i_j,m})\right]
\end{aligned}
$$

$$= \sum_{i_1,\dots,i_4} \hat{\mathbf{S}}_{i_1,i_2}\hat{\mathbf{S}}_{i_3,i_4}\big\{\phi(\mathbf{M}_1,\mathbf{M}_1)_{i_1,\dots,i_4} + \big(E\big[(\dot{\varepsilon}_{i_1,m})^4\big] - 3E\big[(\dot{\varepsilon}_{i_1,m})^2\big]^2\big)$$

$$\times 1_{\{\max_{1\le j\le 4} i_j \le k_m^1 \text{ or } \min_{1\le j\le 4} i_j > k_m^1\}}$$

$$+ 2\phi(\mathbf{M}_1,\mathbf{M}_2)_{i_1,\dots,i_4} + \phi(\mathbf{M}_2,\mathbf{M}_2)_{i_1,\dots,i_4} + \big(E\big[(\tilde{\varepsilon}_{i_1,m})^4\big] - 3E\big[(\tilde{\varepsilon}_{i_1,m})^2\big]^2\big)\delta_{i_1,i_2,i_3,i_4}\big\},$$

where $\mathbf{M}_1 = \mathrm{diag}(v_{1,*}\mathbf{1}, v_{2,*}\mathbf{1})$ and $\mathbf{M}_2 = \mathrm{diag}(v_{1,*}\mathcal{E}, v_{2,*}\mathcal{E})$.

Hence, we obtain

$$\begin{aligned} &E_m\big[\big(\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m\big)^2\big] \\ &\quad = E_m\big[\big(\tilde{Z}_{2,m}^\top \mathbf{S}\tilde{Z}_{2,m}\big)^2\big] + \sum_{i_1,\dots,i_4} \mathbf{S}_{i_1,i_2}\mathbf{S}_{i_3,i_4}\big(\phi(\tilde{S}_{1,m,*},\tilde{S}_{1,m,*}) + 2\phi(\tilde{S}_{1,m,*}, M_{m,*})\big)_{i_1,\dots,i_4} \\ &\quad = 2\,\mathrm{tr}(\tilde{S}_{m,*}\mathbf{S}\tilde{S}_{m,*}\mathbf{S}) + \mathrm{tr}(\tilde{S}_{m,*}\mathbf{S})^2 + \sum_{j=1}^{2}\big(E\big[\big(\varepsilon^{n,j}_{K^j_{m-1}+1}\big)^4\big] - 3v_{j,*}^2\big)\,\mathrm{tr}(\hat{\mathbf{S}}\mathcal{E}_{(j)}\hat{\mathbf{S}}\mathcal{E}_{(j)}) \\ &\qquad + \sum_i C_{n,i}|\hat{\mathbf{S}}_{ii}|^2, \end{aligned} \tag{A.6}$$

by (A.5), where $C_{n,i} = E[(\tilde{\varepsilon}_{i,m})^4] - 3E[(\tilde{\varepsilon}_{i,m})^2]^2$ and $\mathcal{E}_{(j)}$ is a $(k_m^1 + k_m^2)\times(k_m^1+k_m^2)$ matrix with elements $(\mathcal{E}_{(1)})_{kl} = 1_{\{k\le k_m^1\}}1_{\{l\le k_m^1\}}$ and $(\mathcal{E}_{(2)})_{kl} = 1_{\{k>k_m^1\}}1_{\{l>k_m^1\}}$.

Lemma A.7 and (A.1) yield

$$\begin{aligned} \sum_i C_{n,i}|\hat{\mathbf{S}}_{ii}|^2 &\le C\sup_i C_{n,i} b_n^{-1}\bar{k}_n \|\tilde{S}_{m,*}\mathbf{S}\tilde{S}_{m,*}\|\,\mathrm{tr}\big(\big(\mathbf{A}_m^\top\big)^{-1}\tilde{S}_{m,*}^{-2}\mathbf{A}_m^{-1}\big) \\ &\le \bar{R}_n\big(b_n^{-2}k_n\big)\,\mathrm{tr}\left(\begin{pmatrix} M_{1,m} & 0 \\ 0 & M_{2,m}\end{pmatrix}\tilde{S}_{m,*}^{-2}\right) \\ &\le \bar{R}_n\big(b_n^{-2}k_n\big)\,\mathrm{tr}\big(\tilde{S}_{m,*}^{-1}\big) = \bar{R}_n\big(b_n^{-3/2}k_n^2\big) = \bar{R}_n(1). \end{aligned} \tag{A.7}$$

Here, we used the fact that diagonal elements $(\tilde{S}_{m,*}^{-2})_{11}$ and $(\tilde{S}_{m,*}^{-2})_{k_m^1+1,k_m^1+1}$ are positive since $\tilde{S}_{m,*}^{-2}$ is positive definite.

Moreover, the assumptions and Lemmas A.1, A.4 and A.7 yield

$$\begin{aligned} \mathrm{tr}(\hat{\mathbf{S}}\mathcal{E}_{(1)}\hat{\mathbf{S}}\mathcal{E}_{(1)}) &\le \mathrm{tr}\left(\hat{\mathbf{S}}\begin{pmatrix}\mathbf{1}+\mathcal{E} & 0 \\ 0 & 0\end{pmatrix}\hat{\mathbf{S}}\mathcal{E}_{(1)}\right) \\ &\le C\,\mathrm{tr}\big(\tilde{S}_{m,*}^{-1/2}\mathbf{A}_m^{-1}\mathcal{E}_{(1)}\big(\mathbf{A}_m^\top\big)^{-1}\tilde{S}_{m,*}^{-1/2}\big)\left\|\tilde{S}_{m,*}^{1/2}\mathbf{S}\begin{pmatrix} M_{1,m} & 0 \\ 0 & 0\end{pmatrix}\mathbf{S}\tilde{S}_{m,*}^{1/2}\right\| \\ &\le Cb_n^{-2}r_n\underline{r}_n^{-3}\big(\tilde{S}_{m,*}^{-1}\big)_{11} \le Cb_n^{-2}r_n\underline{r}_n^{-3}\big(M_{m,*}^{-1}\big)_{11} = \bar{R}_n(1), \end{aligned} \tag{A.8}$$

since $(M_{m,*}^{-1})_{11} \le v_{1,*}^{-1}$ by (4.5).

(A.6)–(A.8) and similar estimates for $\mathrm{tr}(\hat{\mathbf{S}}\mathcal{E}_{(2)}\hat{\mathbf{S}}\mathcal{E}_{(2)})$ yield $E_m[(\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m)^2] = 2\,\mathrm{tr}((\mathbf{S}\tilde{S}_{m,*})^2) + \mathrm{tr}(\mathbf{S}\tilde{S}_{m,*})^2 + \bar{R}_n(1)$.

We next prove the estimate for $E_m[|\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m|^q]$. Let $p \in \mathbb{N}$ satisfy $q \le 2p$. Then it is sufficient to show that $E_m[(\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m)^{2p}] = \bar{R}_n(b_n^{-2p}k_n^{4p})$.

Note that

$$E_m\big[\big(\tilde{Z}_{2,m}^\top \mathbf{S}\tilde{Z}_{2,m}\big)^{2p}\big] = \sum_{i_1,\dots,i_{4p}} \hat{\mathbf{S}}_{i_1,i_2}\cdots\hat{\mathbf{S}}_{i_{4p-1},i_{4p}} E_m\left[\prod_{j=1}^{4p}(\tilde{\varepsilon}_{i_j,m} - \dot{\varepsilon}_{i_j,m})\right],$$

and there exist $\{0,1\}$-valued maps $\{A_{l,m}\}_l$, constants C_l, positive integers $\{q_l'\}$ not greater than $4p$ and injections $\{\iota_l\}$ such that $E_m[\prod_{j=1}^{4p}(\tilde{\varepsilon}_{i_j,m} - \dot{\varepsilon}_{i_j,m})] = \sum_l C_l A_{l,m}(i_{\iota_l(1)},\dots,i_{\iota_l(q_l')})$ and $\sum_{j_1,\dots,j_{q_l'}} A_{l,m}(j_1,\dots,j_{q_l'}) = \bar{R}_n(k_n^{[q_l'/2]})$ for any l. Then Lemma A.8 yields

$$E_m\big[\big(\tilde{Z}_{2,m}^\top \mathbf{S}\tilde{Z}_{2,m}\big)^{2p}\big] = \bar{R}_n\big(b_n^{-2p}k_n^{4p}\big). \tag{A.9}$$

Moreover, (A.5) yields

$$E_m\big[\big(\tilde{Z}_{1,m}^\top \mathbf{S}\tilde{Z}_{1,m}\big)^{2p}\big] \le C_p \sum_{\substack{\gamma=(\gamma_1,\dots,\gamma_L);\\ L\in\mathbb{N},\gamma_k\ge 1,\sum_k\gamma_k=2p}} \prod_k \mathrm{tr}\big((\mathbf{S}\tilde{S}_{1,m,*})^{\gamma_k}\big) = \bar{R}_n\big(b_n^{-p}k_n^{2p}\big). \tag{A.10}$$

Furthermore, by calculating the expectation of $\tilde{Z}_{1,m}$ and using (A.5) and Lemma A.7, we have

$$\begin{aligned}
E_m\big[\big(\tilde{Z}_{1,m}^\top \mathbf{S}\tilde{Z}_{2,m}\big)^{2p}\big] &= \left(\sum_{(l_{2k-1},l_{2k})_{k=1}^q} 1\right) E_m\big[\big(\tilde{Z}_{2,m}^\top \mathbf{S}\tilde{S}_{1,m,*}\mathbf{S}\tilde{Z}_{2,m}\big)^p\big] \\
&\le (2p-1)!!\|M_{m,*}\mathbf{S}\tilde{S}_{1,m,*}\mathbf{S}M_{m,*}\|^p E_m\big[\big(\tilde{Z}_{2,m}^\top M_{m,*}^{-2}\tilde{Z}_{2,m}\big)^p\big] \\
&= \bar{R}_n\big(b_n^{-p}k_n^{2p}\big).
\end{aligned}$$

Then we obtain $E_m[(\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m)^{2p}] = \bar{R}_n(b_n^{-2p}k_n^{4p})$.

For the estimate of $E_m[(\tilde{Z}_m^\top \mathbf{S}\tilde{Z}_m)^4]$, we have $E_m[(\tilde{Z}_{1,m}^\top \mathbf{S}\tilde{Z}_{1,m})^4] = \bar{R}_n(b_n^{-2}k_n^4)$ and $E_m[(\tilde{Z}_{1,m}^\top \mathbf{S}\tilde{Z}_{2,m})^4] = \bar{R}_n(b_n^{-2}k_n^4)$ by the above results. Moreover, we have

$$E_m\big[\big(\tilde{Z}_{2,m}^\top \mathbf{S}\tilde{Z}_{2,m}\big)^4\big] = \sum_{i_1,\dots,i_8}\prod_{k=1}^4\big(\big(\mathbf{A}_m^\top\big)^{-1}\mathbf{S}\mathbf{A}_m^{-1}\big)_{i_{2k-1},i_{2k}} E_m\left[\prod_{k=1}^8(\tilde{\varepsilon}_{i_k,m} - \dot{\varepsilon}_{i_k,m})\right] \tag{A.11}$$

and there exist $\{0,1\}$-valued maps $\{A'_l\}_l$, constants C'_l, positive integers $\{q''_l\}$ not greater than 8 and injections $\{\iota'_l\}$ such that $\sum_{j_1,\dots,j_{q''_l}} A'_l(j_1,\dots,j_{q''_l}) = \bar{R}_n(k_n^{[q''_l/2]\wedge 3})$ and

$$E_m\left[\prod_{j=1}^{8}(\tilde{\varepsilon}_{i_j,m} - \dot{\varepsilon}_{i_j,m})\right] = \sum_{(l_{2k-1},l_{2k})_{k=1}^4}\prod_{k=1}^{4}\delta_{l_{2k-1},l_{2k}} + \sum_l C_l A'_l(i_{\iota_l(1)},\dots,i_{\iota_l(q'_l)}), \qquad \text{(A.12)}$$

where the summation in the first term of the right-hand side is over all 4-pairs $(l_{2k-1},l_{2k})_{k=1}^4$ of variables $i_1,\dots,i_8$.

Then (A.11), (A.12) and Lemma A.8 yield

$$\begin{aligned} E_m\big[(\tilde{Z}_{2,m}^\top \mathbf{S}\tilde{Z}_{2,m})^4\big] &= \sum_{i_1,\dots,i_8}\prod_{k=1}^{4}\big((\mathbf{A}_m^{-1})^\top \mathbf{S}\mathbf{A}_m^{-1}\big)_{i_{2k-1},i_{2k}} \sum_{(l_{2k-1},l_{2k})_{k=1}^4}\prod_{k=1}^{4}\delta_{l_{2k-1},l_{2k}} + \bar{R}_n\big(b_n^{-4}k_n^7\big) \\ &= \bar{R}_n\big(\big(b_n^{-4}k_n^7\big)\vee\big(b_n^{-2}k_n^4\big)\big). \end{aligned}$$

2. Let $\{\tilde{I}_{i,m}\}_{i=1}^{k_m^1+k_m^2}$ be defined by $\tilde{I}_{i,m} = I_{i,m}^1$ for $1\le i\le k_m^1$ and $\tilde{I}_{i,m} = I^2_{i-k_m^1,m}$ for $k_m^1 < i \le k_m^1 + k_m^2$. Since $(Z_m - \tilde{Z}_{2,m} \pm \tilde{Z}_{1,m})_i = \int_{\tilde{I}_{i,m}}(b_{t,*}^{\mathbf{k}(i)} \pm \tilde{b}_{m,*}^{\mathbf{k}(i)})\,dW_t + \mu_{s_{m-1}}^{\mathbf{k}(i)}|\tilde{I}_{i,m}| + \int_{\tilde{I}_{i,m}}(\mu_t^{\mathbf{k}(i)} - \mu_{s_{m-1}}^{\mathbf{k}(i)})\,dt$, we have $(Z_m - \tilde{Z}_m)^\top \mathbf{S}(Z_m + \tilde{Z}_m) = \Psi_{m,1} + \Psi_{m,2} + \Psi_{m,3}$, where

$$\begin{aligned} \Psi_{m,1} &= 2(Z_m - \tilde{Z}_m)^\top \mathbf{S}\tilde{Z}_{2,m} + \sum_{k=1}^{2}\sum_{i,j}\mathbf{S}_{i,j}\mu_{s_{m-1}}^{\mathbf{k}(i)}|\tilde{I}_{i,m}|\int_{\tilde{I}_{j,m}}\big(b_{t,*}^{\mathbf{k}(j)} + (-1)^k\tilde{b}_{m,*}^{\mathbf{k}(j)}\big)\,dW_t \\ &\quad + \sum_{k=1}^{2}\sum_{i,j}\mathbf{S}_{i,j}\int_{\tilde{I}_{i,m}}\big(b_{t,*}^{\mathbf{k}(i)} + (-1)^k\tilde{b}_{m,*}^{\mathbf{k}(i)}\big)\int_{\tilde{I}_{j,m}\cap[0,t)}\big(b_{s,*}^{\mathbf{k}(j)} + (-1)^{k-1}\tilde{b}_{m,*}^{\mathbf{k}(j)}\big)\,dW_s\,dW_t, \end{aligned}$$

$E_m[\Psi_{m,2}] = 0$, $\Psi_{m,2} = \bar{R}_n(b_n^{-1}k_n^{3/2})$ and $\Psi_{m,3} = \bar{R}_n(b_n^{-3/2}k_n^2)$. Here we used the facts that

$$\big|x^\top Ay\big| \le \|A\||x||y| \qquad \text{(A.13)}$$

for vectors x, y and a matrix A,

$$\big|E_m\big[\mathrm{tr}(\mathbf{S}\mathcal{D})\big]\big| = \big|\mathrm{tr}\big(\mathbf{S}E_m[\mathcal{D}]\big)\big| \le \big\|E_m[\mathcal{D}]\big\|\,\|\mathbf{S}\tilde{S}_{m,*}\|\,\mathrm{tr}\big(\tilde{S}_{m,*}^{-1}\big) = \bar{R}_n\big(b_n^{-3/2}k_n^2\big),$$

and

$$\begin{aligned} E_\Pi\big[\bar{E}_m\big[\mathrm{tr}(\mathbf{S}\mathcal{D})\big]^q\big] &\le C E_\Pi\big[E_m\big[\|\mathcal{D}\|^q\big]\|\tilde{S}_{m,*}\mathbf{S}\|^q\,\mathrm{tr}\big(\tilde{S}_{m,*}^{-1}\big)^q\big] \\ &\le C E_\Pi\big[\bar{R}_n\big(b_n^{3/2q}\ell_n^{-q}\big)\bar{R}_n\big(b_n^{-q}\ell_n^{-q/2}\big)\big] \end{aligned}$$

for any $q\ge 1$, where $\mathcal{D} = (\mathcal{D}_{ij})_{ij}$ and $\mathcal{D}_{ij} = \int_{\tilde{I}_{i,m}\cap\tilde{I}_{j,m}}(b_{t,*}^{\mathbf{k}(i)} + \tilde{b}_{m,*}^{\mathbf{k}(i)})(b_{t,*}^{\mathbf{k}(j)} - \tilde{b}_{m,*}^{\mathbf{k}(j)})\,dt$.

Then the Burkholder–Davis–Gundy inequality yields

$$\begin{aligned}
&E_\Pi\biggl[\biggl|\sum_m (Z_m-\tilde{Z}_m)^\top \mathbf{S}(Z_m+\tilde{Z}_m)\biggr|^q\biggr] \\
&\quad\le CE_\Pi\biggl[\biggl(\sum_m (\Psi_{m,1}+\Psi_{m,2})^2\biggr)^{\frac{q}{2}}\biggr]+\bar{R}_n\bigl(b_n^{-\frac{q}{2}}k_n^q\bigr) \\
&\quad\le CE_\Pi\biggl[\biggl(\sum_m E_m\bigl[\Psi_{m,1}^2\bigr]\biggr)^{\frac{q}{2}}\biggr]+CE_\Pi\biggl[\biggl(\sum_m \bar{E}_m\bigl[\Psi_{m,1}^2\bigr]^2\biggr)^{\frac{q}{4}}\biggr]+\bar{R}_n\bigl(b_n^{-\frac{q}{2}}k_n^q\bigr).
\end{aligned} \tag{A.14}$$

Let $\dot{Z}_{m,i}=\int_{\tilde{I}_{i,m}}(b_{t,*}^{\mathbf{k}(i)}-\tilde{b}_{m,*}^{\mathbf{k}(i)})\,dW_t$ and $\dot{Z}_m=(\dot{Z}_{m,i})_i$, then we obtain

$$\begin{aligned}
\Psi_{m,1}&=2(Z_m-\tilde{Z}_m)^\top\mathbf{S}\tilde{Z}_{2,m}+\dot{Z}_m^\top\mathbf{S}\dot{Z}_m+2\tilde{Z}_{1,m}^\top\mathbf{S}\dot{Z}_m \\
&\quad+(2\dot{Z}_m+2\tilde{Z}_{1,m})^\top\mathbf{S}\bigl(\mu_{s_{m-1}}^{\mathbf{k}(i)}|\tilde{I}_{i,m}|\bigr)_i+\bar{R}_n\bigl(b_n^{-1}k_n^{3/2}\bigr).
\end{aligned} \tag{A.15}$$

Moreover, for any $p\in\mathbb{N}$, Lemma A.7, (A.13), (A.9) and the Cauchy–Schwarz inequality yield

$$\begin{aligned}
&E_m\bigl[\bigl(2(Z_m-\tilde{Z}_m)^\top\mathbf{S}\tilde{Z}_{2,m}+(2\dot{Z}_m+2\tilde{Z}_{1,m})^\top\mathbf{S}\bigl(\mu_{s_{m-1}}^{\mathbf{k}(i)}|\tilde{I}_{i,m}|\bigr)_i\bigr)^{2p}\bigr] \\
&\quad\le CE_m\bigl[\bigl\|\tilde{S}_{m,*}\mathbf{S}(Z_m-\tilde{Z}_m)(Z_m-\tilde{Z}_m)^\top\mathbf{S}\tilde{S}_{m,*}\bigr\|^{2p}\bigr]^{1/2}E_m\bigl[\bigl(\tilde{Z}_{2,m}^\top\tilde{S}_{m,*}^{-2}\tilde{Z}_{2,m}\bigr)^{2p}\bigr]^{1/2} \\
&\qquad+\bar{R}_n\bigl(\bigl(b_nk_nb_n^{-1/2}b_n^{-1}\bigr)^{2p}\bigr) \\
&\quad=\bar{R}_n\bigl(\bigl(b_n^{-1}\ell_n^{-1}k_n\bigr)^pk_n^{2p}\bigr)+\bar{R}_n\bigl(b_n^{-p}k_n^{2p}\bigr)=\bar{R}_n\bigl(b_n^{-2p}k_n^{4p}\bigr).
\end{aligned} \tag{A.16}$$

We can rewrite $\dot{Z}_{m,i}=\mathbf{L}_i^1+\mathbf{L}_i^2+\mathbf{L}_i^3+\bar{R}_n(b_n^{-1/2}\ell_n^{-3/2})$, where $\mathbf{L}_i^1=\sum_j\xi_j^1\int_{\tilde{I}_{i,m}}(t-s_{m-1})\,dW_t^j$, $\mathbf{L}_i^2=\sum_{j,k}\xi_{j,k}^2\int_{\tilde{I}_{i,m}}(W_t^k-W_{s_{m-1}}^k)\,dW_t^j$ and $\mathbf{L}_i^3=\sum_{j,k,l}\xi_{j,k,l}^3\int_{\tilde{I}_{i,m}}\int_{s_{m-1}}^t(W_s^l-W_{s_{m-1}}^l)\,dW_s^k\,dW_t^j$ for some $\mathcal{G}_{s_{m-1}}$-measurable random variables ξ_j^1, $\xi_{j,k}^2$, and $\xi_{j,k,l}^3$ with bounded moments.

Let $\mathbf{L}^j=(\mathbf{L}_i^j)_i$. Then, Lemma A.7 and (A.10) yield

$$\begin{aligned}
&E_m\bigl[\bigl(\dot{Z}_m^\top\mathbf{S}\dot{Z}_m+2\tilde{Z}_{1,m}^\top\mathbf{S}\dot{Z}_m\bigr)^{2p}\bigr] \\
&\quad\le C\bigl\|\tilde{S}_{m,*}^{1/2}\mathbf{S}\tilde{S}_{m,*}^{1/2}\bigr\|^{2p}\bigl(E_m\bigl[\bigl((\mathbf{L}^2)^\top\tilde{S}_{m,*}^{-1}\mathbf{L}^2\bigr)^{2p}\bigr]+\bar{R}_n\bigl(\bigl(b_nk_nb_n^{-1}\ell_n^{-2}\bigr)^{2p}\bigr)\bigr) \\
&\qquad+C\sum_{j=1}^3E_m\bigl[\bigl(\tilde{Z}_{1,m}^\top\mathbf{S}\mathbf{L}^j\bigr)^{2p}\bigr] \\
&\qquad+E_m\biggl[\biggl\|\tilde{S}_{m,*}\mathbf{S}\biggl(\dot{Z}_m-\sum_{j=1}^3\mathbf{L}^j\biggr)^\top\biggl(\dot{Z}_m-\sum_{j=1}^3\mathbf{L}^j\biggr)\mathbf{S}\tilde{S}_{m,*}\biggr\|^{2p}\biggr]^{1/2}E_m\bigl[\bigl(\tilde{Z}_{1,m}^\top\tilde{S}_{m,*}^{-2}\tilde{Z}_{1,m}\bigr)^{2p}\bigr]^{1/2} \\
&\quad\le\bar{R}_n(1)E_m\bigl[\bigl((\mathbf{L}^2)^\top\tilde{S}_{m,*}^{-1}\mathbf{L}^2\bigr)^{2p}\bigr]+C\sum_{j=1}^3E_m\bigl[\bigl(\tilde{Z}_{1,m}^\top\mathbf{S}\mathbf{L}^j\bigr)^{2p}\bigr]+\bar{R}_n\bigl(b_n^{-4p}k_n^{6p}\bigr).
\end{aligned} \tag{A.17}$$

Here we also used that $(a+b)^\top A(a+b) \le 2a^\top Aa + 2b^\top Ab$ for a positive definite matrix A and vectors a, b.

Moreover, we can see that there exists a positive constant C_p such that

$$E_m\left[\sum_{\substack{i_1,\dots,i_{2p}\\ j_1,\dots,j_{2p}}} \prod_{k=1}^{2p}\left(W^{p_{1,k}}(\tilde{I}_{i_k,m})\int_{\tilde{I}_{j_k,m}}\left(W_t^{p_{2,k}} - W_{s_{m-1}}^{p_{2,k}}\right) dW_t^{p_{3,k}}\right)\right]$$

$$\le C_p \sum_{\substack{i_1,\dots,i_{2p}\\ j_1,\dots,j_{2p}}} \sum_{(l_{2q-1},l_{2q})_{q=1}^{2p-\alpha},\alpha} \left(\prod_{q=1}^{2p-\alpha} |\tilde{I}_{l_{2q-1},m}\cap \tilde{I}_{l_{2q},m}|\right) r_n^{2\alpha}(s_m - s_{m-1})^{p-\alpha}$$

for any $\{p_{l,k}\}_{1\le l\le 3, 1\le k\le 2p} \subset \{1,2\}$, where the second summation on the right-hand side is taken over $0 \le \alpha \le p$ and $(2p-\alpha)$ disjoint pairs $(l_{2q-1}, l_{2q})_{q=1}^{2p-\alpha}$ in the variables $i_1,\dots,i_{2p}, j_1,\dots,j_{2p}$. Here we used the fact that all $6p$ factors $(W^{p_{1,k}}(\tilde{I}_{i_k,m}), \int_{\tilde{I}_{j_k,m}} \cdot\, dW_t^{p_{3,k}}, (W_t^{p_{2,k}} - W_{s_{m-1}}^{p_{2,k}})_{t\in \tilde{I}_{j_k,m}})_{k=1}^{2p}$ should be separated into $3p$ pairs in the non-zero terms. 2α represents the number of pairs with the form $(W^{p_{1,k}}(\tilde{I}_{i_k,m}), (W_t^{p_{2,k'}} - W_{s_{m-1}}^{p_{2,k'}})_{t\in\tilde{I}_{j_{k'},m}})$ or $(\int_{\tilde{I}_{j_k,m}} \cdot\, dW_t^{p_{3,k}}, (W_t^{p_{2,k'}} - W_{s_{m-1}}^{p_{2,k'}})_{t\in\tilde{I}_{j_{k'},m}})$. Therefore we obtain

$$E_m\left[\left(\tilde{Z}_{1,m}^\top \mathbf{S}\mathbf{L}^2\right)^{2p}\right] = \bar{R}_n\left(\sum_{0\le\alpha\le p} \left(b_n^{1/2}\ell_n^{-1}\right)^{2p-\alpha}\left(b_n^{3/2}\ell_n^{-1}k_n\right)^{\alpha} r_n^{2\alpha}\ell_n^{-p+\alpha}\right) = \bar{R}_n\left(\left(b_n^{-1}k_n^{3/2}\right)^{2p}\right).$$

(A.17) and similar arguments for $E_m[(\tilde{Z}_{1,m}^\top \mathbf{S}\mathbf{L}^1)^{2p}]$, $E_m[(\tilde{Z}_{1,m}^\top \mathbf{S}\mathbf{L}^3)^{2p}]$ and $E_m[((\mathbf{L}^2)^\top \mathbf{S}\mathbf{L}^2)^{2p}]$ yield

$$E_m\left[\left(\dot{Z}_m^\top \mathbf{S}\dot{Z}_m + 2\tilde{Z}_{1,m}^\top \mathbf{S}\dot{Z}_m\right)^{2p}\right] = \bar{R}_n\left(b_n^{-2p}k_n^{3p}\right). \tag{A.18}$$

(A.15), (A.16) and (A.18) yield

$$E_m\left[\Psi_{m,1}^{2p}\right] = \bar{R}_n\left(b_n^{-2p}k_n^{4p}\right), \quad \text{and hence} \quad E_\Pi\left[\left(\sum_m \Psi_{m,1}^4\right)^{q/4}\right] = \bar{R}_n\left(\left(b_n^{-3}k_n^7\right)^{q/4}\right). \tag{A.19}$$

(A.13), (A.15) and (A.18) yield

$$\begin{aligned} E_m\left[\Psi_{m,1}^2\right] &\le C\left\|\tilde{S}_{m,*}^{1/2}\mathbf{S}M_{m,*}\mathbf{S}\tilde{S}_{m,*}^{1/2}\right\| E_m\left[(Z_m - \tilde{Z}_m)^\top \tilde{S}_{m,*}^{-1}(Z_m - \tilde{Z}_m)\right] + \bar{R}_n\left(b_n^{-2}k_n^3\right) \\ &\quad + \left\|\mathbf{S}\left\{E_m\left[(2\dot{Z}_m + 2\tilde{Z}_{1,m})_i(2\dot{Z}_m + 2\tilde{Z}_{1,m})_j\right]\right\}_{i,j}\mathbf{S}\right\| \sum_i \left(\mu_{s_{m-1}}^{\mathbf{k}(i)}|\tilde{I}_{i,m}|\right)^2 \\ &\le \bar{R}_n\left(b_n^{-3/2}k_n^2\right) + \bar{R}_n\left(b_n^{-2}k_n^3\right) + \bar{R}_n\left(b_nk_nb_n^{-2}\right) = \bar{R}_n\left(b_n^{-2}k_n^3\right), \end{aligned}$$

and therefore, we have $E_\Pi[(\sum_m E_m[\Psi_{m,1}^2])^{q/2}] = \bar{R}_n(b_n^{-q/2}k_n^q)$, which completes the proof of point 2.

Point 3 is easily obtained by the proof of point 2 since we only need the estimate for $E_\Pi[(\sum_m[\Psi^2_{m,1}])]$ if $q=2$. □

A.4. An additional lemma

Lemma A.9. *Let e_n be a sequence of positive numbers, $\mathcal{S}$ be an open set in a Euclidean space, $A_n(\lambda)$ and $B_n(\lambda)$ be sequences of positive-valued random variables, and $C_n(\lambda)$ be a sequence of non-negative-valued random variables for $\lambda \in \mathcal{S}$. Assume that $A_n(\lambda)$, $B_n(\lambda)$, and $C_n(\lambda)$ are C^3 with respect to λ, $e_n \to \infty$, $\sup_{0\le k\le 3,\lambda\in\mathcal{S}}(|\partial_\lambda^k A_n| \vee |\partial_\lambda^k B_n|) = O_p(e_n^{-1})$, and $\sup_{0\le k\le 3,\lambda\in\mathcal{S}}|\partial_\lambda^k C_n| = O_p(e_n^{-2})$ as $n\to\infty$, $C_n < A_nB_n$ a.s. for any $n\in\mathbb{N}$, and $\lim_{n\to\infty}(e_n^2(A_nB_n - C_n)) > 0$ a.s. Then*

$$\sup_{\lambda\in\mathcal{S}}\left|\partial_\lambda^k\left(\sum_{p=1}^\infty\int_0^\pi \frac{C_n^p}{f_p(A_n,x)f_p(B_n,x)}dx - \frac{\pi C_n}{2P_n\sqrt{A_nB_n-C_n}}\right)\right| = O_p\left(e_n^{-\frac{3}{2}}\right), \quad \text{(A.20)}$$

$$\sup_{\lambda\in\mathcal{S}}\left|\partial_\lambda^k\left(\sum_{p=1}^\infty\int_0^\pi \frac{C_n^p f_1(A_n,x)}{f_p(A_n,x)f_p(B_n,x)}dx - \frac{\pi C_n(A_n+\sqrt{A_nB_n-C_n})}{2P_n\sqrt{A_nB_n-C_n}}\right)\right| = O_p\left(e_n^{-\frac{5}{2}}\right), \quad \text{(A.21)}$$

$$\sup_{\lambda\in\mathcal{S}}\left|\partial_\lambda^k\left(\sum_{p=1}^\infty\frac{1}{p}\int_0^\pi \frac{C_n^p}{f_p(A_n,x)f_p(B_n,x)}dx - \pi(\sqrt{A_n}+\sqrt{B_n}) + \pi P_n\right)\right| = O_p\left(e_n^{-1}\right) \quad \text{(A.22)}$$

for $0\le k\le 3$, where $f_p(a,x) = (a+2(1-\cos x))^p$ and $P_n = \sqrt{A_n+B_n+2\sqrt{A_nB_n-C_n}}$.

Proof. An elementary calculation yields

$$\begin{aligned}
&\sum_{p=1}^\infty\int_0^\pi \frac{C_n^p}{(A_n+2(1-\cos x))^p(B_n+2(1-\cos x))^p}dx \\
&= \int_{-\infty}^\infty\left(1-\frac{C_n}{(A_n+\frac{4t^2}{1+t^2})(B_n+\frac{4t^2}{1+t^2})}\right)^{-1}\frac{C_n}{(A_n+\frac{4t^2}{1+t^2})(B_n+\frac{4t^2}{1+t^2})}\frac{1}{1+t^2}dt \\
&= \int_{-\infty}^\infty\frac{C_n}{(A_n+\frac{4t^2}{1+t^2})(B_n+\frac{4t^2}{1+t^2})-C_n}\frac{1}{1+t^2}dt \\
&= \int_{-\infty}^\infty\frac{C_n(1+t^2)}{((A_n+4)t^2+A_n)((B_n+4)t^2+B_n)-C_n(1+t^2)^2}dt \\
&= \int_{-\infty}^\infty\frac{C_n(1+t^2)}{(A_nB_n+4A_n+4B_n-C_n+16)t^4+2(A_nB_n+2A_n+2B_n-C_n)t^2+A_nB_n-C_n}dt.
\end{aligned} \quad \text{(A.23)}$$

We only consider the case $(A_n - B_n)^2 + 16C_n > 0$ a.s. We can easily obtain the desired results for the other case with a slight modification.

Let

$$\alpha_1 = \frac{-2A_n - 2B_n - \sqrt{4(A_n - B_n)^2 + 16C_n}}{16} \quad \text{and}$$
$$\alpha_2 = \frac{-2A_n - 2B_n + \sqrt{4(A_n - B_n)^2 + 16C_n}}{16},$$

then we have $\alpha_1 < \alpha_2 < 0$ by the assumptions. Moreover, we can calculate the right-hand side of (A.23) as

$$\begin{aligned}
&\bigl(1 + O_p\bigl(e_n^{-1}\bigr)\bigr) \int_{-\infty}^{\infty} \frac{C_n(1+t^2)}{16(t^2-\alpha_1)(t^2-\alpha_2)}\,dt \\
&= \bigl(1 + O_p\bigl(e_n^{-1}\bigr)\bigr) \frac{2\pi i}{16}\left[\frac{C_n(1+\alpha_1)}{2\sqrt{-\alpha_1}i(\alpha_1-\alpha_2)} + \frac{C_n(1+\alpha_2)}{2\sqrt{-\alpha_2}i(\alpha_2-\alpha_1)}\right] \\
&= \bigl(1 + O_p\bigl(e_n^{-1}\bigr)\bigr) \frac{\pi}{16} \frac{C_n(\sqrt{-\alpha_2} - \sqrt{-\alpha_1})(1+\sqrt{\alpha_1\alpha_2})}{\sqrt{\alpha_1\alpha_2}(\alpha_1-\alpha_2)} \\
&= \frac{\pi}{16} \frac{C_n}{\sqrt{\alpha_1\alpha_2}(\sqrt{-\alpha_1}+\sqrt{-\alpha_2})} + O_p\bigl(e_n^{-3/2}\bigr).
\end{aligned}$$

Therefore, we obtain (A.20) with $k=0$ by noting that $16\alpha_1\alpha_2 = A_nB_n - C_n$ and

$$\sqrt{x+y+\sqrt{(x-y)^2+4z}} + \sqrt{x+y-\sqrt{(x-y)^2+4z}} = \sqrt{2x+2y+4\sqrt{xy-z}}$$

for non-negative numbers x, y, z satisfying $xy \ge z$.

We similarly have (A.20) with $1 \le k \le 3$ and (A.21) with $0 \le k \le 3$.

Furthermore, we obtain $\int_{-\infty}^{\infty}(1+t^2)^{-1}\log(t^2+\alpha^2)\,dt = 2\pi\log(\alpha+1)$ for any $\alpha > 0$ by the residue theorem.

Therefore, we obtain

$$\begin{aligned}
&\partial_\lambda^k \sum_{p=1}^{\infty} \frac{1}{p} \int_0^{\pi} \frac{C_n^p}{f_p(A_n,x)f_p(B_n,x)}\,dx \\
&= -\partial_\lambda^k \int_{-\infty}^{\infty} \frac{1}{1+t^2} \log\left(1 - \frac{C_n}{(A_n + 4t^2/(1+t^2))(B_n + 4t^2/(1+t^2))}\right) dt \\
&= -\partial_\lambda^k \int_{-\infty}^{\infty} \frac{1}{1+t^2} \log\left(\frac{(t^2-\alpha_1)(t^2-\alpha_2)}{(t^2 + A_n/(A_n+4))(t^2 + B_n/(B_n+4))}\right) dt + O_p\bigl(e_n^{-1}\bigr) \\
&= -2\pi\,\partial_\lambda^k \left(\log(1+\sqrt{-\alpha_1}) + \log(1+\sqrt{-\alpha_2}) - \log\left(1 + \sqrt{\frac{A_n}{A_n+4}}\right)\right.
\end{aligned}$$

$$-\log\left(1+\sqrt{\frac{B_n}{B_n+4}}\right)\right)+O_p\left(e_n^{-1}\right)$$

$$=2\pi\partial_\lambda^k\left(\sqrt{A_n}/2+\sqrt{B_n}/2-\sqrt{-\alpha_1}-\sqrt{-\alpha_2}\right)+O_p\left(e_n^{-1}\right),$$

which completes the proof. □

Acknowledgement

The author is grateful to Yuta Koike for very useful discussions and his help on numerical analysis of quadratic covariation estimators. He had a great contribution on constructing the "cce" function in the "yuima" R package as well. This work was partially supported by Japan Society for the Promotion of Science KAKENHI Grant Number 15K21598, and JST PRESTO Grant Number JPMJPR15E2, Japan.

References

[1] Adams, R.A. and Fournier, J.J.F. (2003). *Sobolev Spaces*, 2nd ed. *Pure and Applied Mathematics* (*Amsterdam*) **140**. Amsterdam: Elsevier/Academic Press. MR2424078

[2] Aït-Sahalia, Y., Fan, J. and Xiu, D. (2010). High-frequency covariance estimates with noisy and asynchronous financial data. *J. Amer. Statist. Assoc.* **105** 1504–1517. MR2796567

[3] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica* **76** 1481–1536. MR2468558

[4] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *J. Econometrics* **162** 149–169. MR2795610

[5] Bibinger, M., Hautsch, N., Malec, P. and Reiss, M. (2014). Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency. *Ann. Statist.* **42** 80–114. MR3226158

[6] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. MR1700749

[7] Christensen, K., Kinnebrock, S. and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *J. Econometrics* **159** 116–133. MR2720847

[8] Christensen, K., Podolskij, M. and Vetter, M. (2013). On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. *J. Multivariate Anal.* **120** 59–84. MR3072718

[9] Cox, J.C., Ingersoll, J.E. Jr. and Ross, S.A. (1985). A theory of the term structure of interest rates. *Econometrica* **53** 385–407. MR0785475

[10] Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84** 313–342. MR1719345

[11] Genon-Catalot, V. and Jacod, J. (1994). Estimation of the diffusion coefficient for diffusion processes: Random sampling. *Scand. J. Stat.* **21** 193–221. MR1292636

[12] Gloter, A. and Jacod, J. (2001). Diffusions with measurement errors. I. Local asymptotic normality. *ESAIM Probab. Stat.* **5** 225–242. MR1875672
[13] Gloter, A. and Jacod, J. (2001). Diffusions with measurement errors. II. Optimal estimators. *ESAIM Probab. Stat.* **5** 243–260. MR1875673
[14] Gobet, E. (2001). Local asymptotic mixed normality property for elliptic diffusion: A Malliavin calculus approach. *Bernoulli* **7** 899–912. MR1873834
[15] Hayashi, T. and Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli* **11** 359–379. MR2132731
[16] Hayashi, T. and Yoshida, N. (2008). Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Ann. Inst. Statist. Math.* **60** 367–406. MR2403524
[17] Hayashi, T. and Yoshida, N. (2011). Nonsynchronous covariation process and limit theorems. *Stochastic Process. Appl.* **121** 2416–2454. MR2822782
[18] Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités, XXXI. Lecture Notes in Math.* **1655** 232–246. Berlin: Springer. MR1478732
[19] Jacod, J., Li, Y., Mykland, P.A., Podolskij, M. and Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.* **119** 2249–2276. MR2531091
[20] Jeganathan, P. (1982). On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A* **44** 173–212. MR0688800
[21] Jeganathan, P. (1983). Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. *Sankhyā Ser. A* **45** 66–87. MR0749355
[22] Malliavin, P. and Mancino, M.E. (2002). Fourier series method for measurement of multivariate volatilities. *Finance Stoch.* **6** 49–61. MR1885583
[23] Malliavin, P. and Mancino, M.E. (2009). A Fourier transform method for nonparametric estimation of multivariate volatility. *Ann. Statist.* **37** 1983–2010. MR2533477
[24] Ogihara, T. (2015). Local asymptotic mixed normality property for nonsynchronously observed diffusion processes. *Bernoulli* **21** 2024–2072. MR3378458
[25] Ogihara, T. and Yoshida, N. (2014). Quasi-likelihood analysis for nonsynchronously observed diffusion processes. *Stochastic Process. Appl.* **124** 2954–3008. MR3217430
[26] Podolskij, M. and Vetter, M. (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli* **15** 634–658. MR2555193
[27] Uchida, M. (2010). Contrast-based information criterion for ergodic diffusion processes from discrete observations. *Ann. Inst. Statist. Math.* **62** 161–187. MR2577445
[28] Uchida, M. and Yoshida, N. (2013). Quasi likelihood analysis of volatility and nondegeneracy of statistical random field. *Stochastic Process. Appl.* **123** 2851–2876. MR3054548
[29] Yoshida, N. (2006). Polynomial type large deviation inequalities and convergence of statistical random fields. *ISM Research Memorandum* 1021.
[30] Yoshida, N. (2011). Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. *Ann. Inst. Statist. Math.* **63** 431–479. MR2786943
[31] Zhang, L., Mykland, P.A. and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100** 1394–1411. MR2236450

Received December 2015 and revised April 2017