Abstract
We study maximum-likelihood-type estimation for diffusion processes when the coefficients are nonrandom and observations occur in nonsynchronous manner. The problem of nonsynchronous observations is important when we consider the analysis of high-frequency data in a financial market. Constructing a quasi-likelihood function to define the estimator, we adaptively estimate the parameter for the diffusion part and the drift part. We consider the asymptotic theory when the terminal time point \( T_n \) and the observation frequency goes to infinity, and show the consistency and the asymptotic normality of the estimator. Moreover, we show local asymptotic normality for the statistical model, and asymptotic efficiency of the estimator as a consequence.

To show the asymptotic properties of the maximum-likelihood-type estimator, we need to control the asymptotic behaviors of some functionals of the sampling scheme. Though it is difficult to directly control those in general, we study tractable sufficient conditions when the sampling scheme is generated by mixing processes.

Keywords Asymptotic efficiency · Diffusion processes · Local asymptotic normality · Maximum-likelihood-type estimation · Nonsynchronous observations

1 Introduction
Given a probability space \((\Omega, \mathcal{F}, P)\) with a right-continuous filtration \( \mathbf{F} = \{\mathcal{F}_t\}_{t \geq 0} \), let \( X^{(1)}(\alpha) = \{X^{(1)}(\alpha), X^{(2)}(\alpha)\}_{t \geq 0} \) be a two-dimensional \( \mathbf{F} \)-adapted process satisfying the following stochastic differential equation:
where $x_0 \in \mathbb{R}^2$, $\{W_t\}_{0 \leq t \leq T}$ is a two-dimensional standard $\mathcal{F}$-Wiener process, $\{\mu_t(\theta)\}_{t \geq 0}$ and $\{b_t(\sigma)\}_{t \geq 0}$ are deterministic functions with values in $\mathbb{R}^2$ and $\mathbb{R}^2 \times \mathbb{R}^2$, respectively, $\alpha = (\sigma, \theta)$, $\sigma \in \Theta_1$, $\theta \in \Theta_2$, and $\Theta_1$ and $\Theta_2$ are bounded open subsets of $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, respectively. Let $\alpha_0 = (\sigma_0, \theta_0) \in \Theta_1 \times \Theta_2$ be the true value, and let $X_t = (X^1_t, X^2_t) = X^{(\alpha_0)}_t$. We consider estimation of $\alpha_0$ when $X$ is observed with nonsynchronous manner, that is, observation times of $X^1$ and $X^2$ are different to each other.

The problem of nonsynchronous observations appears in the analysis of high-frequency financial data. If we analyze the intra-day stock price data, we observe stock prices when a new transaction or a new order arrives. Then, the observation times are different for different stocks, and hence, we cannot avoid the problem of nonsynchronous observations. Statistical analysis with such data is much more complicated compared to the analysis with synchronous data. Parametric estimation for diffusion processes with synchronous and equidistant observations has been analyzed through quasi-maximum-likelihood methods in Florens-Zmirou (1989), Yoshida (1992, 2011), Kessler (1997), and Uchida and Yoshida (2012). Related to the estimation problem for nonsynchronously observed diffusion processes, estimators for the quadratic covariation have been actively studied. Hayashi and Yoshida (2005, 2008, 2009) and Malliavin and Mancino (2002, 2009) have independently constructed consistent estimators under nonsynchronous observations. There are also studies of covariation estimation under the simultaneous presence of microstructure noise and nonsynchronous observations (Barndorff-Nielsen et al., 2011; Bibinger et al., 2014; Christensen et al., 2010, and so on). For parametric estimation with nonsynchronous observations, Ogihara and Yoshida (2014) have constructed maximum-likelihood-type and Bayes-type estimators and have shown the consistency and the asymptotic mixed normality of the estimators when the terminal time point $T_n$ is fixed and the observation frequency goes to infinity. Ogihara (2015) have shown local asymptotic mixed normality for the model in Ogihara and Yoshida (2014), and the maximum-likelihood-type and Bayes-type estimators have been shown to be asymptotically efficient. On the other hand, we need to consider asymptotic theory that the terminal time point $T_n$ goes to infinity to consistently estimate the parameter $\theta$ in the drift term. To the best of the author’s knowledge, there are no studies of the asymptotic theory of parametric estimation for nonsynchronously observed diffusion processes when $T_n \to \infty$.

In this work, we consider the asymptotic theory for nonsynchronously observed diffusion processes when $T_n \to \infty$, and construct maximum-likelihood-type estimators for the parameter $\sigma$ in the diffusion part and the parameter $\theta$ in the drift part. We show the consistency and the asymptotic normality of the estimators. Moreover, we show local asymptotic normality of the statistical model, and we obtain asymptotic efficiency of our estimator as a consequence. Our estimator is constructed based on the quasi-likelihood function that is similarly defined to the one in Ogihara and Yoshida (2014), though we need some modification to deal with the drift part. To investigate asymptotic theory for the maximum-likelihood-type estimator, we need to specify the limit of the quasi-likelihood function. Then, we need to assume some conditions for
the asymptotic behavior of the sampling scheme. In Ogihara and Yoshida (2014), for a matrix

\[ G = \left\{ \left( S_{ni}^{n,1} \wedge S_{nj}^{n,2} - S_{j-1}^{n,2} \lor S_{i-1}^{n,1} \lor 0 \right) \right\}_{i,j} \]

generated by the sampling scheme, the existence of the probability limit of

\[ n^{-1} \text{tr}((GG^T)^p) \quad (p \in \mathbb{Z}_+) \]

is required, where \((S_i^{n,\ell})_i\) are observation times of \(X^\ell\) and \(\top\) denotes transpose of a vector or a matrix. Since we consider the different asymptotics, the asymptotic behavior of the quasi-likelihood function is different from that in Ogihara and Yoshida (2014). We also need to consider estimation for the drift parameter \(\theta\). Then, we need other assumptions for the asymptotic behavior of the sampling scheme [Assumption (A5)]. Though these conditions for the sampling scheme are difficult to check directly, we study tractable sufficient conditions in Sect. 2.4.

The rest of this paper is organized as follows. In Sect. 2, we introduce our model settings and the assumptions for main results. Our estimator is constructed in Sect. 2.1, and the asymptotic normality of the estimator is given in Sect. 2.2. Section 2.3 deals with local asymptotic normality of our model and asymptotic efficiency of the estimator. Tractable sufficient conditions for the assumptions of the sampling scheme are given in Sect. 2.4. Section 3 contains the proofs of main results. Preliminary results are collected in Sect. 3.1. Section 3.2 is for the consistency of the estimator for \(\sigma\), Sect. 3.3 is for the asymptotic normality of the estimator for \(\sigma\), Sect. 3.4 is for the consistency of the estimator for \(\theta\), and Sect. 3.5 is for the asymptotic normality of the estimator for \(\theta\). Other proofs are collected in Sect. 3.6.

2 Main results

2.1 Setting and parameter estimation

Let \(\mathbb{N}\) be the set of all positive integers. For \(\ell \in \{1, 2\}\), let the observation times \(\{S^n_{ni,\ell}\}_{i=0}^{M_{\ell}}\) be strictly increasing random times with respect to \(i\), and satisfy \(S^n_{0,\ell} = 0\) and \(S^n_{M_{\ell}} = nh_n\), where \(M_{\ell}\) is a random positive integer depending on \(n\) and \((h_n)_{n=1}^\infty\) is a sequence of positive numbers satisfying

\[ h_n \to 0, \quad n^{1-\epsilon_0}h_n \to \infty, \quad nh_n^2 \to 0 \quad (2.1) \]

as \(n \to \infty\) for some \(\epsilon_0 > 0\). Intuitively, \(n\) is of the order of the number of observations and \(h_n\) is of the order of the length of the observation intervals. More precise assumptions of observation times are given in (A2), (A4), and (A5) later. We assume that \(\{S^n_{ni,\ell}\}_{0 \leq i \leq M_{\ell}, \ell = 1, 2}\) is independent of \(\mathcal{F}_T\), and its distribution does not depend on \(\alpha\). We consider nonsynchronous observations of \(X\), that is, we observe \(\{S^n_{ni,\ell}\}_{0 \leq i \leq M_{\ell}, \ell = 1, 2}\) and \(\{X^n_{ni,\ell}\}_{0 \leq i \leq M_{\ell}, \ell = 1, 2}\). In particular, we consider the nonendogenous observation times.
We denote by $\| \cdot \|$ the operator norm with respect to the Euclidean norm for a matrix. We often regard a $p$-dimensional vector $v$ as a $p \times 1$ matrix. For $j \in \mathbb{N}$, we denote $\partial_z = \frac{\partial}{\partial z}$ for a variable $z \in \mathbb{R}^j$, and denote $\partial_z^l = (\partial_{z_{i_1}} \cdots \partial_{z_{i_l}})^T_{i_1, \ldots, i_l=1}$ for $l \in \mathbb{N}$. For functions $f$ and $g$, we often use shorthand notation $\partial_z f \partial_z g = (\partial_z f (\partial_z g)^T + \partial_z g (\partial_z f)^T)/2$.

For a set $A$ in a topological space, let $\text{clos}(A)$ denote the closure of $A$. For a matrix $A$, $[A]_{ij}$ denotes its $(i, j)$ element. For a vector $v = (v_j)^{J}_{j=1}$, we denote $[v]_j = v_j$, and $\text{diag}(v)$ denotes a $K \times K$ diagonal matrix with elements $\{\text{diag}(v)\}_{jj} = v_j$.

Let $M = M_1 + M_2$. For $1 \leq i \leq M$, let

$$\varphi(i) = \begin{cases} i, & \text{if } i \leq M_1, \\ i - M_1, & \text{if } i > M_1, \end{cases} \quad \psi(i) = \begin{cases} 1, & \text{if } i \leq M_1, \\ 2, & \text{if } i > M_1. \end{cases}$$

For a two-dimensional stochastic process $(U_t)_{t \geq 0} = (U^1_t, U^2_t)_{t \geq 0}$, let $\Delta^l_1 U = U^l_{S^n_{p,l}} - U^l_{S^{n,l-1}}$ and let $\Delta^l U = (\Delta^l_1 U)_{1 \leq i \leq M}$ and $\Delta_i U = \Delta^l_{\psi(i)} U$ for $1 \leq i \leq M$.

Let $\Delta U = (\Delta_1 U)^T, (\Delta^2 U)^T$. Let $|K| = b - a$ for an interval $K = (a, b]$. Let $I^l_i = (S^n_{i-1}, S^n_{i,l})$ for $1 \leq i \leq M_1$, and let $I_i = I^l_{\psi(i)}$ for $1 \leq i \leq M$. We denote a unit matrix of size $k$ by $E_k$.

Let $\tilde{\Sigma}_i^l(\sigma) = \int_{I^l_i} [b_i b_i^T(\sigma)]_1 dr$ and $\tilde{\Sigma}_{i,j}^{1,2}(\sigma) = \int_{I^l_i \cap I^l_j} [b_i b_i^T(\sigma)]_2 dr$, and let $\tilde{\Sigma}_i = \tilde{\Sigma}_{\psi(i)}$ for $1 \leq i \leq M$. By setting $\tilde{D} = \text{diag}((\tilde{\Sigma}_i)_{1 \leq i \leq M})$

$$\tilde{G}(\sigma) = \left\{ \frac{\tilde{\Sigma}_{i,j}^{1,2}}{\sqrt{\tilde{\Sigma}_i^{1} \tilde{\Sigma}_j^{2}}} \right\}_{1 \leq i \leq M_1, 1 \leq j \leq M_2},$$

we can calculate the covariance matrix of $\Delta X$ as

$$S_n(\sigma) = \tilde{D}^{1/2} \left( \begin{array}{c} \mathcal{E}_{M_1} \\ \tilde{G}(\sigma) \\ \mathcal{E}_{M_2} \end{array} \right) \tilde{D}^{1/2}.$$  \hspace{1cm} (2.2)

As we will see later, we can ignore the term related to $\mu_t(\theta)$ (drift term) when we consider estimation of $\sigma$, because this term converges to zero very fast. Therefore, we first construct an estimator for $\sigma$, and then construct an estimator for $\theta$. Such adaptive estimation can speed up the calculation.

We define the quasi-likelihood function $H^1_n(\sigma)$ for $\sigma$ as follows:

$$H^1_n(\sigma) = -\frac{1}{2} \Delta X^T S_n^{-1}(\sigma) \Delta X - \frac{1}{2} \log \det S_n(\sigma).$$

Then, the maximum-likelihood-type estimator for $\sigma$ is defined by

$$\hat{\sigma}_n \in \argmax_{\sigma} H^1_n(\sigma).$$

We consider estimation for $\theta$ next. Let $V(\theta) = (V_t(\theta))_{t \geq 0}$ be a two-dimensional stochastic process defined by $V_t(\theta) = (\int_0^t \mu^1_s(\theta)^T ds, \int_0^t \mu^2_s(\theta)^T ds)^T$. Let $\tilde{X}(\theta) = \tilde{X}(\hat{\theta}).$
\( \Delta X - \Delta V(\theta) \). We define the quasi-likelihood function \( H^2_n(\theta) \) for \( \theta \) as follows:

\[
H^2_n(\theta) = -\frac{1}{2} \bar{X}(\theta)^\top S_n^{-1}(\hat{\sigma}_n) \bar{X}(\theta).
\]

Then, the maximum-likelihood-type estimator for \( \theta \) is defined by

\[
\hat{\theta}_n \in \arg\max_{\theta \in \text{clos}(\Theta)} H^2_n(\theta).
\]

The quasi-(log-)likelihood function \( H^1_n(\sigma) \) is defined in the same way as that in Ogihara and Yoshida (2014). Since \( \Delta X \) follows normal distribution, we can construct such a Gaussian quasi-likelihood function even for the nonsynchronous data. When the coefficients are random, though the distribution of \( \Delta X \) is not Gaussian, such Gaussian-type quasi-likelihood function is still valid due to the local Gaussian property of diffusion processes. The Gaussian mean that comes from the drift part is ignored when we construct the quasi-likelihood \( H^1_n \). When we estimate the parameter \( \theta \) for the drift part, we subtract the mean in \( \bar{X}(\theta) \) to construct the quasi-likelihood function \( H^2_n \). Since the effect of the drift term on the estimation of \( \sigma \) is small, it works well to estimate \( \sigma \) in this way and then plug in \( \hat{\sigma}_n \) to \( S_n \) to construct the estimator for \( \theta \). Thus, we can speed up the calculation by separating the estimation for \( \sigma \) and \( \theta \).

**Remark 2.1** \( H^1_n(\sigma) \) and \( H^2_n(\theta) \) are well defined only if \( \det S_n(\sigma) > 0 \) and \( \det S_n(\hat{\sigma}_n) > 0 \), respectively. For the covariance matrix \( S_n \) of nonsynchronous observations \( \Delta X \), it is not trivial to check these conditions. Proposition 1 in Section 2 of Ogihara and Yoshida (2014) shows that these conditions are satisfied if \( b_t(\sigma) \) is continuous on \([0, \infty) \times \text{clos}(\Theta_1)\) and \( \inf_{t,\sigma} \det(b_t b_t^\top(\sigma)) > 0 \). We assume such conditions in our setting (Assumption (A1) in Sect. 2.2).

**Remark 2.2** As seen in Ogihara and Yoshida (2014), the quasi-likelihood analysis for nonsynchronously observed diffusion processes becomes much more complicated compared to synchronous observations. In this work, estimation for the drift parameter \( \theta \) is added, and hence, we consider nonrandom drift and diffusion coefficients to avoid overcomplication. For general diffusion processes with the random drift and diffusion coefficients, we need to set predictable coefficients to use the martingale theory. However, the quasi-likelihood function loses a Markov property with nonsynchronous observations and the coefficients in the quasi-likelihood function contain randomness of future time. Then, we need to approximate the coefficients by predictable functions. This operation is particularly complicated. Moreover, approximating the true likelihood function by the quasi-likelihood function is much more difficult problem when we show local asymptotic normality and asymptotic efficiency of the estimators. Therefore, we left asymptotic theory under general random drift and diffusion coefficients as a future work.
2.2 Asymptotic normality of the estimator

In this section, we state the assumptions of our main results, and state the asymptotic normality of the estimator.

For $m \in \mathbb{N}$, an open subset $U \subset \mathbb{R}^m$ is said to admit Sobolev’s inequality if, for any $p > m$, there exists a positive constant $C$ depending on $U$ and $p$, such that $\sup_{x \in U} |u(x)| \leq C \sum_{k=0}^{m} \left( \int |\partial_x^k u(x)|^p dx \right)^{1/p}$ for any $u \in C^1(U)$. This is the case when $U$ has a Lipschitz boundary. We assume that $\Theta, \Theta_1$, and $\Theta_2$ admit Sobolev’s inequality.

Let $\Sigma_t(\sigma) = b_t b_t^T (\sigma)$, and let

$$\rho_t(\sigma) = \frac{[\Sigma_t]_{12}}{[\Sigma_t]_{11}^{1/2} [\Sigma_t]_{22}^{1/2}(\sigma)}, \quad B_{l,t}(\sigma) = \frac{[\Sigma_t(\sigma_0)]_{ll}}{[\Sigma_t(\sigma)]_{ll}}.$$ 

Let $\rho_{t,0} = \rho_t(\sigma_0)$.

**Assumption (A1).** There exists a positive constant $c_1$, such that $c_1 \mathcal{E}_2 \leq \Sigma_t(\sigma)$ for any $t \in [0, \infty)$ and $\sigma \in \Theta_1$. For $k \in \{0, 1, 2, 3, 4\}$, $\partial^k_\sigma \mu_t(\theta)$ and $\partial^k_\sigma b_t(\sigma)$ exist and are continuous with respect to $(t, \sigma, \theta)$ on $[0, \infty) \times \text{clos}(\Theta_1) \times \text{clos}(\Theta_2)$. For any $\epsilon > 0$, there exist $\delta > 0$ and $K > 0$, such that

$$|\partial^k_\sigma \mu_t(\theta)| + |\partial^k_\sigma b_t(\sigma)| \leq K,$$

$$|\partial^k_\sigma \mu_t(\theta) - \partial^k_\sigma \mu_s(\theta)| + |\partial^k_\sigma b_t(\sigma) - \partial^k_\sigma b_s(\sigma)| \leq \epsilon$$

for any $k \in \{0, 1, 2, 3, 4\}$, $\sigma \in \Theta_1$, $\theta \in \Theta_2$, and $t, s \geq 0$ satisfying $|t - s| < \delta$. Let $r_n = \max_{i,l} |I_l^i|$.

**Assumption (A2).** $r_n \to 0$ as $n \to \infty$.

**Assumption (A3).** For any $l \in \{1, 2\}$, $i_1 \in \mathbb{Z}_+$, $i_2 \in \{0, 1\}$, $i_3 \in \{0, 1, 2, 3, 4\}$, $k_1, k_2 \in \{0, 1, 2\}$ satisfying $k_1 + k_2 = 2$, and any polynomial function $F(x_1, \ldots, x_{14})$ of degree equal to or less than 6, there exist continuous functions $\Phi_{i_1,i_2}^{1,F}(\sigma)$, $\Phi_{i_3}^{2}(\sigma)$ and $\Phi_{i_1,i_3}^{3,k_1,k_2}(\theta)$ on $\text{clos}(\Theta_1)$ and $\text{clos}(\Theta_2)$, such that

$$\frac{1}{T} \int_0^T \frac{F((\partial^k_\sigma B_{l,t}(\sigma))_{0 \leq k \leq 4, l = 1, 2})}{(\partial^k_\sigma \rho_t(\sigma))_{k'=1}^4 (\partial^k_\sigma \rho_t(\sigma))_{k=1}^4} \rho_{t,0}^2 \, dt \to \Phi_{i_1,i_2}^{1,F}(\sigma),$$

$$\frac{1}{T} \int_0^T \partial^k_\sigma \log B_{l,t}(\sigma) \, dt \to \Phi_{i_3}^{2}(\sigma),$$

$$\frac{1}{T} \int_0^T \partial^k_\theta (\phi_{i_1,t}^{k_1} \phi_{2,t}^{k_2})(\theta) \rho_{t,0}^2 \, dt \to \Phi_{i_1,i_3}^{3,k_1,k_2}(\theta)$$

as $T \to \infty$ for $\sigma \in \text{clos}(\Theta_1)$, $\theta \in \text{clos}(\Theta_2)$, where $\phi_{l,t}(\theta) = [\Sigma_t(\sigma_0)]_{ll}^{1/2} (\mu^l_t(\theta) - \mu^l_t(\theta_0))$.

Assumption (A1) and the Ascoli–Arzelà theorem yield that the convergences in (A3) can be replaced by uniform convergence with respect to $\sigma$ and $\theta$ (the left-hand
sides of the above equations become relatively compact, and then, any uniformly convergent subsequence converges to the right-hand sides due to the pointwise convergence assumptions. Assumption (A3) is satisfied if \( \mu \) convergent subsequence converges to the right-hand sides due to the pointwise convergence of \( t \), or are periodic functions with respect to \( t \) having a common period (when the period does not depend on \( \sigma \) nor \( \theta \)). Let \( \mathcal{G} \) be the set of all partitions \( (s_k)_{k=0}^\infty \) of \([0, \infty)\) satisfying \( \sup_{k \geq 1} |s_k - s_{k-1}| \leq 1 \) and \( \inf_{k \geq 1} |s_k - s_{k-1}| > 0 \). For \((s_k)_{k=0}^\infty \in \mathcal{G} \), let \( M_l, k = \# \{ i; \sup I_l^i \in (s_{k-1}, s_k) \} \) and \( q_n = \max \{ k; s_k \leq nh_n \} \), and let \( E^{(k)}_{(l)} \) be an \( M_l \times M_l \) matrix satisfying \([E^{(k)}_{(l)}]_{ij} = 1\) if \( i = j \) and \( \sup I_l^i \in (s_{k-1}, s_k) \), and otherwise, \([E^{(k)}_{(l)}]_{ij} = 0 \).

**Assumption (A4).** There exist positive constants \( a_0^1 \) and \( a_0^2 \), such that \( \{h_n M_l, q_n+1\}_{n=1}^\infty \) is \( P \)-tight and

\[
\max_{1 \leq k \leq q_n} |h_n M_l, k - a_0^1(s_k - s_{k-1})| \xrightarrow{P} 0
\]

for \( l \in \{1, 2\} \) and any partition \((s_k)_{k=0}^\infty \in \mathcal{G} \). Moreover, for any \( p \in \mathbb{N} \), there exists a nonnegative constant \( a_p^1 \), such that

\[
\max_{1 \leq k \leq q_n} |h_n \text{tr}(E^{(k)}_{(l)}(GG^\top)^p) - a_p^1(s_k - s_{k-1})| \xrightarrow{P} 0
\]

as \( n \to \infty \) for any partition \((s_k)_{k=0}^\infty \in \mathcal{G} \). Let \( J_l = (|I_l^i|^{1/2})_{i=1}^{M_l} \).

**Assumption (A5).** For \( p \in \mathbb{Z}_+ \), there exist nonnegative constants \( f_p^{1,1} \), \( f_p^{1,2} \), and \( f_p^{2,2} \), such that \( \{|E^{(k)}_{(l)}| J_l\}_{n=1}^\infty \) is \( P \)-tight for \( l \in \{1, 2\} \), and

\[
\max_{1 \leq k \leq q_n} |J_1 E^{(k)}_{(l)}(GG^\top)^p J_1 - f_p^{1,1}(s_k - s_{k-1})| \xrightarrow{P} 0,
\]

\[
\max_{1 \leq k \leq q_n} |J_1 E^{(k)}_{(l)}(GG^\top)^p G J_2 - f_p^{1,2}(s_k - s_{k-1})| \xrightarrow{P} 0,
\]

\[
\max_{1 \leq k \leq q_n} |J_2 E^{(k)}_{(l)}(G^\top G)^p J_2 - f_p^{2,2}(s_k - s_{k-1})| \xrightarrow{P} 0
\]

as \( n \to \infty \) for any partition \((s_k)_{k=0}^\infty \in \mathcal{G} \).

Assumption (A4) corresponds to [A3’] in Ogihara and Yoshida (2014). The functionals in (A4) and (A5) appear in \( H_n^1 \) and \( H_n^2 \), and hence, we cannot specify the limits of \( H_n^1 \) and \( H_n^2 \) unless we assume existence of the limits of these functionals. It is difficult to directly check (A4) and (A5) for concrete statistical experiments with general sampling schemes. We study sufficient conditions for these conditions in Sect. 2.4.
**Assumption (A6).** The constant $a_1^1$ in (A4) is positive, and there exist positive constants $c_2$ and $c_3$, such that

$$\limsup_{T \to \infty} \left( \frac{1}{T} \int_0^T \| \Sigma_t(\sigma) - \Sigma_t(\sigma_0) \|^2 dt \right) \geq c_2 |\sigma - \sigma_0|^2,$$

$$\limsup_{T \to \infty} \left( \frac{1}{T} \int_0^T |\mu_t(\theta) - \mu_t(\theta_0)|^2 dt \right) \geq c_3 |\theta - \theta_0|^2$$

for any $\sigma \in \text{clos}(\Theta_1)$ and $\theta \in \text{clos}(\Theta_2)$.

Assumption (A6) is necessary to identify the parameter $\sigma$ and $\theta$ from the data. For $p < q$

$$\text{tr}(\mathcal{E}^1(\gamma)(GG^\top)^q) \leq \text{tr}(\mathcal{E}^1(\gamma)(GG^\top)^p) \| (GG^\top)^{q-p} \| \leq \text{tr}(\mathcal{E}^1(\gamma)(GG^\top)^p) \quad (2.3)$$

by Lemma 3.3 later and Lemma A.1 in Ogihara (2018). Then, $a_1^p$ is monotone non-increasing with respect to $p$. This implies that $a_1^1 = 0$ for any $p \in \mathbb{N}$ if $a_1^1 = 0$. In this case, the non-diagonal components of the covariance matrix $S_n$ are negligible in the limit. Then, we cannot consistently estimate the parameter in $\rho_t(\sigma)$. This is why, we need the assumption $a_1^1 > 0$ (see Proposition 3.9 and the following discussion to obtain the consistency).

Let $\mathcal{A}(\rho) = \sum_{p=1}^{\infty} a_1^p \rho^{2p}$ for $\rho \in (-1, 1)$. Then, (2.3) implies that $\mathcal{A}(\rho)$ is finite. Moreover, (A5) yields

$$f_p^{1,1} = (nh_n)^{-1} \sum_{k=1}^{q_n} \mathcal{E}_1^1 \mathcal{E}_1^1 (GG^\top)^p \mathcal{J}_1 + o_p(1)$$

which implies $f_p^{1,1} \leq 1$. Similarly, we have $f_p^{1,2} \leq 1$ and $f_p^{2,2} \leq 1$. Let $\partial_{\sigma} B_{1,t,0} = \partial_{\sigma} B_{1,t}(\sigma_0)$, and let

$$\gamma_{1,t} = \mathcal{A}(\rho_t,0) \left( \frac{\partial_{\sigma} \rho_{t,0}}{\rho_{t,0}} - \partial_{\sigma} B_{1,t,0} - \partial_{\sigma} B_{2,t,0} \right)^2 - \partial_{\rho} \mathcal{A}(\rho_t,0) \left( \frac{\partial_{\sigma} \rho_{t,0}}{\rho_{t,0}} \right)^2 - 2 \sum_{l=1}^{2} (a_0^l)$$

$$+ \mathcal{A}(\rho_t,0)) (\partial_{\sigma} B_{1,t,0})^2,$$

and let $\Gamma_1 = \lim_{T \to \infty} T^{-1} \int_0^T \gamma_{1,t} dt$, which exists under (A1), (A3), and (A4). Let

$$\Gamma_2 = \lim_{T \to \infty} \frac{1}{T} \int_0^T \sum_{p=0}^{\infty} \rho_{t,0}^{2p} \left\{ \sum_{l=1}^{2} f_p^{2,l} (\partial_{\theta} \phi_{l,t})^2 (\theta_0) \right\}.$$
\[-2 \rho_1 \rho_2 f \left( \frac{1}{p} \right) \phi_{1,t} \phi_{2,t} (\theta_0) \right] \, \text{d}t,

which exists under (A1), (A3), and (A5). Let \( T_n = nh_n \) and

\[
\Gamma = \begin{pmatrix}
\Gamma_1 & 0 \\
0 & \Gamma_2
\end{pmatrix}.
\]

**Theorem 2.3** Assume (A1)–(A6). Then, \( \Gamma \) is positive definite, and

\[
(\sqrt{n}(\hat{\sigma}_n - \sigma_0), \sqrt{T_n}(\hat{\theta}_n - \theta_0)) \xrightarrow{d} N(0, \Gamma^{-1})
\]
as \( n \to \infty \).

### 2.3 Local asymptotic normality

Let \( \alpha_0 \in \Theta, \Theta \subset \mathbb{R}^d \), and \( \{P_{\alpha,n}\}_{\alpha \in \Theta} \) be a family of probability measures defined on a measurable space \((X_n, A_n)\) for \( n \in \mathbb{N} \), where \( \Theta \) is an open subset of \( \mathbb{R}^d \). As usual, we shall refer to \( dP_{\alpha_2,n}/dP_{\alpha_1,n} \) the derivative of the absolutely continuous component of the measure \( P_{\alpha_2,n} \) with respect to measure \( P_{\alpha_1,n} \) at the observation \( x \) as the likelihood ratio. The following definition of local asymptotic normality is Definition 2.1 in Chapter II of Ibragimov and Has’minski˘ı (1981).

**Definition 2.4** A family \( P_{\alpha,n} \) is called locally asymptotically normal (LAN) at point \( \alpha_0 \in \Theta \) as \( n \to \infty \) if for some nondegenerate \( d \times d \) matrix \( \epsilon_n \) and any \( u \in \mathbb{R}^d \), the representation

\[
\log \frac{dP_{\alpha_0 + \epsilon_n u,n}}{dP_{\alpha_0,n}} - (u^\top \Delta_n - |u|^2/2) \to 0
\]
in \( P_{\alpha_0,n} \)-probability as \( n \to \infty \), where

\[
\mathcal{L}(\Delta_n | P_{\alpha_0,n}) \to N(0, \mathcal{E}_d)
\]
as \( n \to \infty \), and \( \mathcal{L}(\cdot | P_{\alpha,n}) \) denotes the distribution with respect to \( P_{\alpha,n} \).

Let \( \Theta = \Theta_1 \times \Theta_2 \). For \( \alpha \in \Theta \), let \( P_{\alpha,n} \) be the probability measure generated by the observations \( \{S_i^{n,l}\}_{i,l} \) and \( \{X_{S_i}^{(\alpha,l)}\}_{i,l} \).

**Theorem 2.5** Assume (A1)–(A6). Then, \( \{P_{\alpha,n}\}_{\alpha,n} \) satisfies the LAN property at \( \alpha = \alpha_0 \) with

\[
\epsilon_n = \begin{pmatrix}
n^{-1/2} \Gamma_1^{-1/2} & 0 \\
0 & T_n^{-1/2} \Gamma_2^{-1/2}
\end{pmatrix}.
\]
The proof is left to Sect. 3.6. Theorem 11.2 in Chapter II of Ibragimov and Has’minskii (1981) gives lower bounds of estimation errors for any regular estimator of parameters under the LAN property. Then, the optimal asymptotic variance of \( \varepsilon_n^{-1}(U_n - \alpha_0) \) for regular estimator \( U_n \) is \( E_d \). We will show that \((\hat{\sigma}_n, \hat{\theta}_n)\) is regular in Remark 3.18. Therefore, Theorem 2.5 ensures that our estimator \((\hat{\sigma}_n, \hat{\theta}_n)\) is asymptotically efficient in this sense under the assumptions of the theorem.

### 2.4 Sufficient conditions for the assumptions

It is not easy to directly check Assumptions (A4) and (A5) for general random sampling schemes (even for a sampling scheme generated by simple Poisson processes given in Example 2.6). In this section, we study tractable sufficient conditions for these assumptions. The proofs of the results in this section are left to Sect. 3.6.

Let \( q > 0 \) and \( \mathcal{N}_{t}^{n,l} = \sum_{i=1}^{M_l} 1_{\{s_i^{n,l} \leq t\}} \). We consider the following conditions for the point process \( \mathcal{N}_{t}^{n,l} \).

**Assumption (B1-\( q \)).**

\[
\sup_{n \geq 1} \max_{l \in \{1, 2\}} \sup_{0 \leq t \leq (n-1)h_n} E[(\mathcal{N}_{t+h_n}^{n,l} - \mathcal{N}_{t}^{n,l})^q] < \infty.
\]

**Assumption (B2-\( q \)).**

\[
\limsup_{u \to \infty} \sup_{n \geq 1} \max_{l \in \{1, 2\}} \sup_{0 \leq t \leq nh_n - uh_n} u^q P(\mathcal{N}_{t+uh_n}^{n,l} - \mathcal{N}_{t}^{n,l} = 0) < \infty.
\]

**Example 2.6** Let \((\hat{\mathcal{N}}_t^1, \hat{\mathcal{N}}_t^2)\) be two independent homogeneous Poisson processes with positive intensities \( \lambda_1 \) and \( \lambda_2 \), respectively, and \( \mathcal{N}_{t}^{n,l} = \hat{\mathcal{N}}_{h_n-1}^{l}, \) that is, \( S_i^{n,l} = \inf\{t \geq 0; \hat{\mathcal{N}}_{h_n-1}^{l} \geq i\}. \) Even in this simple case, it is not trivial to directly check (A4) and (A5). On the other hand, (B1-\( q \)) obviously holds for any \( q > 0 \). Moreover, (B2-\( q \)) holds for any \( q > 0 \), since

\[
\limsup_{u \to \infty} \sup_{n \geq 1} \max_{l \in \{1, 2\}} \sup_{0 \leq t \leq nh_n - uh_n} u^q P(\mathcal{N}_{t+uh_n}^{n,l} - \mathcal{N}_{t}^{n,l} = 0) = \lim_{u \to \infty} u^q e^{-(\lambda_1 \wedge \lambda_2)u} = 0.
\]

Then, by Corollary 2.12, we can check Assumptions (A2), (A4), and (A5) for this sampling scheme.

To give sufficient conditions for (A4) and (A5), we consider mixing properties of \( \mathcal{N}_{t}^{n,l} \). That is, we assume conditions for the following mixing coefficient \( \alpha^n_k \). Let

\[
\mathcal{G}_{k,i,j}^n = \sigma(\mathcal{N}_{t}^{n,l} - \mathcal{N}_{s}^{n,l}; i h_n \leq s < t \leq j h_n, l = 1, 2) \quad (0 \leq i, j \leq n),
\]

and let

\[
\alpha^n_k = 0 \lor \sup_{1 \leq i, j \leq n-1, j-i \geq k} \sup_{A \in \mathcal{G}_{0,i}^n, B \in \mathcal{G}_{j,n}^n} |P(A \cap B) - P(A) P(B)|.
\]
Proposition 2.7 Assume that (B1-q) and (B2-q) hold and that
\[ \sup_{n \in \mathbb{N}} \sum_{k=0}^{\infty} (k + 1)^q \alpha_k^n < \infty \] (2.4)
for any \( q > 0 \). Moreover, assume that there exist positive constants \( a_0^1 \) and \( a_0^2 \), and a nonnegative constant \( a_p^1 \) for \( p \in \mathbb{N} \), such that \( \{E[h_n M_{l,qn+1}]\}_{n=1}^{\infty} \) is bounded and
\[
\max_{1 \leq k \leq q_n} |h_n E[M_{l,k}] - a_0^1 (s_k - s_{k-1})| \to 0,
\]
\[
\max_{1 \leq k \leq q_n} |h_n E[\text{tr}(e_{(k)}^1 (GG^\top)^p)] - a_p^1 (s_k - s_{k-1})| \to 0
\] (2.5)
as \( n \to \infty \) for \( p \in \mathbb{Z}_+ \), \( l \in \{1, 2\} \) and any partition \( (s_k)_{k=0}^{\infty} \in \mathcal{S} \). Then, (A4) holds.

Proposition 2.8 Assume that (B1-q) and (B2-q) hold and that (2.4) is satisfied for any \( q > 0 \). Moreover, assume that there exist nonnegative constants \( f_p^{1.1}, f_p^{1.2}, \) and \( f_p^{2.2} \) for \( p \in \mathbb{Z}_+ \), such that \( \{E[|\mathcal{J}_l^{(q_n+1)}|]\}_{n=1}^{\infty} \) is bounded and
\[
\max_{1 \leq k \leq q_n} |E[\mathcal{J}_1 e_{(k)}^1 (GG^\top)^p \mathcal{J}_1] - f_p^{1.1} (s_k - s_{k-1})| \to 0,
\]
\[
\max_{1 \leq k \leq q_n} |E[\mathcal{J}_1 e_{(k)}^1 (GG^\top)^p G \mathcal{J}_2] - f_p^{1.2} (s_k - s_{k-1})| \to 0,
\]
\[
\max_{1 \leq k \leq q_n} |E[\mathcal{J}_2 e_{(k)}^2 (G^\top G)^p \mathcal{J}_2] - f_p^{2.2} (s_k - s_{k-1})| \to 0
\] (2.6)
as \( n \to \infty \) for \( l \in \{1, 2\}, p \in \mathbb{Z}_+ \) and any partition \( (s_k)_{k=0}^{\infty} \in \mathcal{S} \). Then, (A5) holds.

Proposition 2.9 Assume that there exists \( q > 0 \), such that (A4) and (B2-q) hold, \( \{N_{l+h_n} - N_{l}^{n,l}\}_{0 \leq l \leq h_n, l \in \{1,2\}, n \in \mathbb{N}} \) is \( P \)-tight, and \( \sum_{k=1}^{\infty} k \alpha_k^n < \infty \). Then, \( a_1^l > 0 \).

In the following, let \( \mathcal{N}_l^{q,l} \) be an exponential \( \alpha \)-mixing point process for \( l \in \{1, 2\} \). Assume that the distribution of \( (\mathcal{N}_{l+h_n}^{q,l} - \mathcal{N}_{l}^{q,l})_{1 \leq k \leq l, l = 1,2} \) does not depend on \( t \geq 0 \) for any \( K \in \mathbb{N} \) and \( 0 \leq t_0 < t_1 < \cdots < t_K \).

Lemma 2.10 Let \( N_{l}^{q,l} = \mathcal{N}_{l+h_n/l}^{q,l} \) for \( 0 \leq t \leq nh_n \) and \( l \in \{1, 2\} \). Then, (2.4) is satisfied for any \( q > 2 \), and there exist constants \( a_0^1, a_0^2, \) and \( a_p^1 = a_p^2 \) for \( p \in \mathbb{N} \), such that (2.5) holds and \( \{E[h_n M_{l,qn+1}]\}_{n=1}^{\infty} \) is bounded for any \( (s_k)_{k=0}^{\infty} \in \mathcal{S} \). Moreover, there exist nonnegative constants \( f_p^{1.1}, f_p^{1.2}, \) and \( f_p^{2.2} \) for \( p \in \mathbb{Z}_+ \), such that (2.6) holds and \( \{E[|\mathcal{J}_l^{(q_n+1)}|]\}_{n=1}^{\infty} \) is bounded for \( l \in \{1, 2\} \) and any \( (s_k)_{k=0}^{\infty} \in \mathcal{S} \).

Proposition 2.11 (Proposition 8 in Ogihara & Yoshida, 2014) Let \( q \in \mathbb{N} \). Assume (B2-(q + 1)). Then, \( \sup_n E[h_n q^{1+q} r_n^{q}] < \infty \). In particular, (A2) holds under (B2-2).

By the above results, we obtain simple tractable sufficient conditions for the assumptions of the sampling scheme when the observation times are generated by the exponential \( \alpha \)-mixing point process \( \mathcal{N}_l^{q,l} \) defined above.
Corollary 2.12 Let $\mathcal{N}_{t,n}^{l,d} = \hat{N}_{h_n^{-1} t}^{d}$ for $0 \leq t \leq T_n$ and $l \in \{1, 2\}$. Assume that (B1-q) and (B2-q) hold for any $q > 0$. Then, (A2), (A4), and (A5) hold, and $a_1^\dagger > 0$.

3 Proofs

3.1 Preliminary results

For a real number $a$, $[a]$ denotes the maximum integer which is not greater than $a$. Let $\Pi = \Pi_n = \{S^n_{1,t}\}_{1 \leq i \leq M_t, t \in \{1,2\}}$. We denote $|x|^2 = \sum_{i_1, \ldots, i_k} |x_{i_1, \ldots, i_k}|^2$ for $x = \{x_{i_1, \ldots, i_k}\}_{i_1, \ldots, i_k}$ with $k \in \mathbb{N}$ and $x_{i_1, \ldots, i_k} \in \mathbb{R}$. For a matrix $A = (A_{ij})_{i,j}$, $\text{Abs}(A)$ denotes the matrix $(|A_{ij}|)_{i,j}$. $C$ denotes generic positive constant whose value may vary depending on context. We often omit the parameters $\sigma$ and $\theta$ in general functions $f(\sigma)$ and $g(\theta)$.

For a sequence $p_n$ of positive numbers, let us denote by $(\tilde{R}_n(p_n))_{n \in \mathbb{N}}$ a sequence of random variables (which may also depend on $1 \leq i \leq M$ and $\alpha \in \Theta$) satisfying that $\{\sup_{\alpha,i} E_{\Pi}[[p_n^{-1} \tilde{R}_n(p_n)]]\}_{n \in \mathbb{N}}$ is $P$-tight for any $q \geq 1$, where $E_{\Pi}[X] = E[X|\sigma(\Pi_n)]$ for a random variable $X$.

For a matrix $A$ and vectors $v, w$ with suitable sizes, we repeatedly use the following inequality:

$$|w^\top A v| \leq |w||A v| \leq \|A\||v||w|.$$  

Lemma 3.1 (A special case of Lemma 3.1 in Ogihara and Uehara, 2022) Let $(Z_n)_{n \in \mathbb{N}}$ be nonnegative-valued random variables. Then

1. $E_{\Pi}[Z_n] \xrightarrow{P} 0$ as $n \to \infty$ implies that $Z_n \xrightarrow{P} 0$ as $n \to \infty$.
2. $P$-tightness of $(E_{\Pi}[Z_n])_{n \in \mathbb{N}}$ implies $P$-tightness of $(Z_n)_{n \in \mathbb{N}}$.

Let $\tilde{V} = V(\theta_0)$, and let

$$\rho_{i,j}(\sigma) = \begin{cases} \frac{\hat{E}_{i,j}^{1,2}}{\sqrt{\hat{E}_{i,j}^{1,1} \hat{E}_{j,j}^{2,2}}}, & \text{if } |I_i^1 \cap I_j^2| \neq \emptyset, \\ 0, & \text{otherwise}. \end{cases}$$

Let $\tilde{\rho}_n = \sup_{\sigma} (\max_{i,j} |\rho_{i,j}(\sigma)|) \vee \sup_{i} |\rho_i(\sigma)|$, and let

$$\hat{S}_n = \left( \begin{array}{cc} \hat{E}_{M_1} & -\hat{\rho}_n G \\ -\hat{\rho}_n G^\top & \hat{E}_{M_2} \end{array} \right).$$ (3.1)

Let $\Delta_{i,t}^{l} U = U_{t \wedge S_{i}^{r,t}}^{l} - U_{t \wedge S_{i-1}^{r,t}}^{l}$, and let $\Delta_{i,t}^{l} U = \Delta_{\psi(t),t}^{\psi(i),t} U$ for $t \geq 0$ and a two-dimensional stochastic process $(U_t)_{t \geq 0} = ((U_{t}^{1}, U_{t}^{2}))_{t \geq 0}$. 

\(\circledS\) Springer
Under (A4), we have
\[ h_n M_l = h_n \sum_{k=1}^{q_n+1} M_{l,k} = d_0^l n h_n + o_p(nh_n). \]
Then, we obtain
\[ M_l = d_0^l n + o_p(n). \]  \hfill (3.2)

**Lemma 3.2** Assume (A1). Then, for any \( p \geq 1 \), there exist positive constants \( C_p \) (depending on \( p \)) and \( C \), such that
\[
\sup_\theta |\Delta^l_i V(\theta)| \leq C |I^l_i|, \quad E\Pi[|\Delta^l_i X|^p]^{1/p} \leq C_p (|I^l_i| + \sqrt{|I^l_i|})
\]
for \( l \in \{1, 2\} \) and \( 1 \leq i \leq M_l \).

**Proof** Since \( \mu^l_i(\theta) \) and \( [b_t b_t(\sigma_0)]_{lt} \) are bounded by (A1), the Burkholder–Davis–Gundy inequality yields
\[
\sup_\theta |\Delta^l_i V(\theta)| = \sup_\theta \left| \int_{I^l_i} \mu^l_i(\theta) d\theta \right| \leq C |I^l_i|,
\]
\[
E\Pi[|\Delta^l_i X|^p]^{1/p} = E\Pi \left[ \left| \int_{I^l_i} \mu^l_i(\theta_0) d\theta + \int_{I^l_i} [b_t(\sigma_0) dW_t]_l \right|^p \right]^{1/p} \leq C_p |I^l_i| + C_p E\Pi \left[ \left| \int_{I^l_i} [b_t b_t(\sigma_0)]_{lt} dW_t \right|^p \right]^{1/p} \leq C_p (|I^l_i| + \sqrt{|I^l_i|}).
\]

\( \Box \)

**Lemma 3.3** (Lemma 2 in Ogihara & Yoshida, 2014) \( \|G\| \vee \|G^T\| \leq 1. \)

**Lemma 3.4** \( \|\tilde{G}\| \vee \|\tilde{G}^T\| \leq \tilde{\rho}_n. \)

**Proof** Since all the elements of \( G \) are nonnegative, we have
\[
\|\tilde{G}\|^2 = \sup_{|x|=1} |\tilde{G}x|^2 = \sup_{|x|=1} \sum_{i} \left( \sum_{j} \rho_{ij} G_{ij} x_j \right)^2 \leq \tilde{\rho}_n^2 \sup_{|x|=1} \sum_{i} \left( \sum_{j} G_{ij} |x_j| \right)^2 \leq \tilde{\rho}_n^2 \|G\|^2 \leq \tilde{\rho}_n^2.
\]
Since \( \|\tilde{G}^T\| = \|\tilde{G}\| \), we obtain the conclusion. \( \Box \)
Lemma 3.5 Assume (A1). Then, there exists a positive constant C, such that
\( \|D^{1/2} \partial^k \sigma S_n^{-1}(\sigma)D^{1/2}\| \leq C(1 - \bar{\rho}_n)^{-k-1} \) and \( \|S_n^{-1}(\sigma)\|_{ij} \leq C[D^{-1/2} \hat{S}_n^{-1}D^{-1/2}]_{ij} \) if \( \bar{\rho}_n < 1 \), and
\( \|D^{-1/2} \partial^k \sigma S_n(\sigma)D^{-1/2}\| \leq C \) for any \( \sigma \in \Theta_1 \), \( 1 \leq i, j \leq M \), and \( k \in \{0, 1, 2, 3, 4\} \).

**Proof** By (A1) and Lemma 3.3, we have
\[
\|D^{-1/2} \partial^k \sigma S_n(\sigma)D^{-1/2}\| \leq C \sum_{j=0}^{k} \| \partial^j \{ \mathcal{E}_M + \left( \frac{0}{G^\top} \right) \} \| \leq C.
\]

Moreover, by (A1) and Lemma 3.4, we have
\[
\|D^{1/2} S_n^{-1}D^{1/2}\| \leq C \left\| \left( \mathcal{E}_M + \left( \frac{0}{G^\top} \right) \right)^{-1} \right\| \leq C(1 - \bar{\rho}_n)^{-1}
\]
if \( \bar{\rho}_n < 1 \).

Using the equation \( \partial \sigma S_n^{-1} = -S_n^{-1} \partial \sigma S_n S_n^{-1} \), we obtain
\[
\|D^{1/2} \partial \sigma S_n^{-1}D^{1/2}\| = \|D^{1/2} S_n^{-1} \partial \sigma S_n S_n^{-1}D^{1/2}\|
\leq \|D^{1/2} S_n^{-1}D^{1/2}\|^2 \|D^{-1/2} \partial \sigma S_n D^{-1/2}\| \leq C(1 - \bar{\rho}_n)^{-2}
\]
if \( \bar{\rho}_n < 1 \). Similarly, we obtain
\[
\|D^{1/2} \partial \sigma S_n^{-1}D^{1/2}\| \leq C(1 - \bar{\rho}_n)^{-k-1}
\]
if \( \bar{\rho}_n < 1 \) for \( k \in \{0, 1, 2, 3, 4\} \).

If \( \bar{\rho}_n < 1 \), since Lemma 3.4 yields
\[
S_n^{-1} = \hat{D}^{-1/2} \left( \begin{array}{cc} \mathcal{E}_{M_1} & 0 \\ 0 & \mathcal{E}_{M_2} \end{array} \right) \left( \begin{array}{cc} 0 & \tilde{G} \\ \tilde{G}^\top & 0 \end{array} \right) \hat{D}^{-1/2} = \hat{D}^{-1/2} \sum_{p=0}^{\infty} (-1)^p \left( \begin{array}{cc} 0 & \tilde{G} \\ \tilde{G}^\top & 0 \end{array} \right)^p \hat{D}^{-1/2},
\]
we obtain
\[
[S_n^{-1}]_{ij} \leq C \left[ \hat{D}^{-1/2} \sum_{p=0}^{\infty} \bar{\rho}_n^p \left( \begin{array}{cc} 0 & \tilde{G} \\ \tilde{G}^\top & 0 \end{array} \right)^p \hat{D}^{-1/2} \right]_{ij} = [D^{-1/2} \hat{S}_n^{-1}D^{-1/2}]_{ij}.
\]
Under (A1), we have $\Sigma_t(\sigma) \geq c_1 \mathcal{E}_2$, which implies that $\sup_{t,\sigma} |\rho_t(\sigma)| < 1$. Then, by (A2) and uniform continuity of $b_t$, for some fixed $\delta > 0$ and any $\epsilon > 0$, there exists $N \in \mathbb{N}$, such that $P(1 - \tilde{\rho}_n < \delta) < \epsilon$ for $n \geq N$. Therefore, we have

$$P(\tilde{\rho}_n < 1 - \delta) \to 1$$

(3.3) as $n \to \infty$, and we have

$$P((1 - \tilde{\rho}_n)^{-q} > \delta^{-q}) < \epsilon$$

for any $q > 0$ and $n \geq N$, which implies that

$$(1 - \tilde{\rho}_n)^{-q} = O_P(1).$$

(3.4)

Moreover, Lemma 3.4 yields

$$S_n^{-1}(\sigma) = \tilde{D}^{-1/2} \sum_{p=0}^{\infty} (-1)^p \left( \begin{array}{c} 0 \\ \tilde{G}^\top \end{array} \right)^p \tilde{D}^{-1/2}$$

$$= \tilde{D}^{-1/2} \sum_{p=0}^{\infty} \left( \begin{array}{c} (\tilde{G} \tilde{G}^\top)^p \\ -(\tilde{G} \tilde{G}^\top)^p \tilde{G}^\top \\ -(\tilde{G} \tilde{G}^\top)^p \end{array} \right) \tilde{D}^{-1/2}$$

(3.5) if $\tilde{\rho}_n < 1$.

3.2 Consistency of $\hat{\sigma}_n$

In this section, we show consistency: $\hat{\sigma}_n \overset{P}{\to} \sigma_0$ as $n \to \infty$. For this purpose, we specify the limit of $H_n^1(\sigma) - H_n^1(\sigma_0)$.

Lemma 3.6 Assume (A1) and (A2). Then

$$\frac{1}{n} \sup_{\sigma \in \Theta_1} \left| \partial_{\sigma}^k (H_n^1(\sigma) - H_n^1(\sigma_0)) + \frac{1}{2} \partial_{\sigma}^k \text{tr}(S_n^{-1}(\sigma)(S_n(\sigma_0) - S_n(\sigma))) + \frac{1}{2} \partial_{\sigma}^k \log \frac{\det S_n(\sigma)}{\det S_n(\sigma_0)} \right| \overset{P}{\to} 0$$

(3.6)

as $n \to \infty$ for $k \in \{0, 1, 2, 3\}$.

Proof Let $X_t^\gamma = \int_0^t b_s(\sigma_0) dW_s$. By the definition of $H_n^1$, we have

$$H_n^1(\sigma) - H_n^1(\sigma_0) = -\frac{1}{2} \Delta X^\top (S_n^{-1}(\sigma) - S_n^{-1}(\sigma_0)) \Delta X - \frac{1}{2} \log \frac{\det S_n(\sigma)}{\det S_n(\sigma_0)}.$$ 

We first show that
\begin{equation}
H_n^1(\sigma) - H_n^1(\sigma_0) = \frac{1}{2} (\Delta X^c)^\top (S_n^{-1}(\sigma) - S_n^{-1}(\sigma_0)) \Delta X^c - \frac{1}{2} \log \frac{\det S_n(\sigma)}{\det S_n(\sigma_0)} + \sqrt{n}\hat{e}_n(\sigma),
\end{equation}

where \((\hat{e}_n(\sigma))_{n=1}^\infty\) denotes a general sequence of random variables, such that
\[
\sup_\sigma |\hat{e}_n(\sigma)| \xrightarrow{P} 0 \text{ as } n \to \infty.
\]

Since
\[
\Delta X^\top S_n^{-1}(\sigma) \Delta X - (\Delta X^c)^\top S_n^{-1}(\sigma) \Delta X^c = 2(\Delta \bar{V})^\top S_n^{-1}(\sigma) \Delta X^c + (\Delta \bar{V})^\top S_n^{-1}(\sigma) \Delta \bar{V} =: \Psi_1 + \Psi_2,
\]

it suffices to show that \(\Psi_i = \sqrt{n}\hat{e}_n\) for \(i \in \{1, 2\}\).

Lemma 3.5 and (3.4) yield
\[
|\Psi_2| \leq \|D^{1/2}S_n^{-1}(\sigma)D^{1/2}\|D^{-1/2}\Delta \bar{V}\|^2 = O_p(1) \times |D^{-1/2}\Delta \bar{V}|^2.
\]

Moreover, Lemma 3.2 yields
\[
|D^{-1/2}\Delta V(\theta)|^2 = \sum_{i,l} |I_i^l|^{-1} |\Delta_i^l V(\theta)|^2 \leq C \sum_{i,l} |I_i^l|^{-1} |I_i^l|^2 = C \sum_{i,l} |I_i^l| \leq Cnh_n.
\]

Furthermore, Lemma 3.5, (3.4), (3.10), and the equation \(E_{\Pi}[\Delta X^c(\Delta X^c)^\top] = S_n(\sigma_0)\) yield
\[
E_{\Pi}[|\Psi_1|^2] = 4(\Delta \bar{V})^\top S_n^{-1}(\sigma) E_{\Pi}[\Delta X^c(\Delta X^c)^\top] S_n^{-1}(\sigma) \Delta \bar{V} = O_p(nh_n) = o_p(n).
\]

Then, we obtain (3.7) by (3.8)–(3.11) and Lemma 3.1.

Next, we show that
\[
(\Delta X^c)^\top S_n^{-1}(\sigma) \Delta X^c - \text{tr}(S_n^{-1}(\sigma)S_n(\sigma_0)) = \bar{R}_n(\sqrt{n}).
\]

Itô’s formula yields
\[
(\Delta X^c)^\top S_n^{-1}(\sigma) \Delta X^c - \text{tr}(S_n^{-1}(\sigma)S_n(\sigma_0))
= \sum_{i,j} [S_n^{-1}(\sigma)]_{ij} (\Delta_i X^c \Delta_j X^c - [S_n(\sigma_0)]_{ij})
= \sum_{i,j} [S_n^{-1}(\sigma)]_{ij} \left\{ \int_{I_i} \Delta_{i,t} X^c dX_{t}^{c,\psi(i)} + \int_{I_j} \Delta_{i,t} X^c dX_{t}^{c,\psi(j)} \right\}
= 2 \sum_{i,j} [S_n^{-1}(\sigma)]_{ij} \int_{I_i} \Delta_{i,t} X^c dX_{t}^{c,\psi(i)},
\]

where \(X_{t}^{c,l}\) is the \(l\)-th component of \(X_{t}^{c}\).
Since $\langle \Delta_i X^c, \Delta_j X^c \rangle_t = \int_{[0,t] \cap I_i \cap I_j} \Psi(i), \Psi(j) \, dt$, together with the Burkholder–Davis–Gundy inequality, we have

$$E \left[ \left( \sum_{i,j} [S_n^{-1}(\sigma)]_{ij} \int_{I_i} \Delta_j X_t \, dX^c_t, \Psi(i) \right)^q \right]$$

$$\leq C_q \sum_{l=1}^2 E \left[ \left( \sum_{i,j} [S_n^{-1}(\sigma)]_{ij} [S_n^{-1}(\sigma)]_{i,j} \int_{I_i} \Delta_j X_t \Delta_j X_t \, dX^c_t \, [\Sigma_t]_{\Psi(i), \Psi(i)} \right)^{q/2} \right]$$

$$+ C_q E \left[ \left( \sum_{i,i_1,i_2,j,j_2} [S_n^{-1}(\sigma)]_{i,j} [S_n^{-1}(\sigma)]_{i_2,j_2} \times \int_{I_i \cap I_j} \Delta_j X_t \Delta_j X_t \, dX^c_t \, [\Sigma_t]_{\Psi(i_1), \Psi(i_2)} \right)^{q/2} \right]$$

$$\leq C_q E \left[ \left( \sum_{i,i_1,i_2,j,j_2} [S_n^{-1}(\sigma)]_{i,j_1} [S_n^{-1}(\sigma)]_{i_2,j_2} \sup_t |[\Sigma_t]_{\Psi(i_1), \Psi(i_2)}| \sup_t |I_i \cap I_j| \sup_t |\Delta_j X_t| \right)^{q/2} \right]$$

$$\leq C_q E \left[ \left( \|D^{1/2} \mathrm{Abs}(S_n^{-1})(|I_i \cap I_j|)_{ij} \mathrm{Abs}(S_n^{-1}) \| D^{1/2} \| \sum_t \sup_t |\Delta_i X_t^c|^{2} \right)^{q/2} \right].$$

Together with Lemmas 3.3 and 3.5, the triangle inequality for $L^{q/2}$ that

$$|I_i \cap I_j| = \left[ D^{1/2} \left( \begin{array}{c} \mathcal{E}_M1 \\ G^T \\ \mathcal{E}_M2 \end{array} \right) D^{1/2} \right]_{ij},$$

and that

$$\|\mathrm{Abs}(S_n^{-1})\|^2 = \sup_{|x|=1} |\mathrm{Abs}(S_n^{-1})x|^2$$

$$= \sup_{|x|=1} \left( \sum_j [S_n^{-1}]_{ij} |x_j|^2 \right)^2$$

$$\leq C \sup_{|x|=1} \left( \sum_j [D^{-1/2} \hat{S}_n^{-1} D^{-1/2}]_{ij} |x_j|^2 \right)^2$$

$$\leq C \|D^{-1/2} \hat{S}_n^{-1} D^{-1/2}\|^2$$

by Lemma 3.5, we have

$$E \left[ \left( \sum_{i,j} [S_n^{-1}(\sigma)]_{ij} \int_{I_i} \Delta_j X_t \, dX^c_t, \Psi(i) \right)^q \right]$$
\[
\leq C_q (1 - \tilde{\rho}_n)^{-q} E_{\Pi} \left[ \left( \sum_i \sup_t |\Delta_{i,t} X^c|^2 \right)^{q/2} \right] \\
\leq C_q (1 - \tilde{\rho}_n)^{-q} \left( \sum_i E_{\Pi}[|\Delta_{i,t} X^c|^q]^{2/q} \right)^{q/2} \\
\leq C_q M^{q/2} (1 - \tilde{\rho}_n)^{-q}
\]

on \{\tilde{\rho}_n < 1\} for \(q \geq 1\). Then, thanks to (3.2), (3.4), (3.13) and Lemma 3.1, we obtain (3.12).

(3.12), (3.7), Sobolev’s inequality, and similar estimates for \(\partial^k_\sigma (H_n^1(\sigma) - H_n^1(\sigma_0))\) yield

\[
\partial^k_\sigma (H_n^1(\sigma) - H_n^1(\sigma_0)) \\
= -\frac{1}{2} \partial^k_\sigma \text{tr}(S_n(\sigma_0)(S_n^{-1}(\sigma) - S_n^{-1}(\sigma_0))) - \frac{1}{2} \partial^k_\sigma \log \det S_n(\sigma) \det S_n(\sigma_0) + \sqrt{n} \hat{\epsilon}_{n}(\sigma) \\
\]

for \(k \in \{0, 1, 2, 3\}\).

For \((s_k)_{k=0}^\infty \in \mathcal{S}\), let \(\hat{A}^1_{k,p} = e^{(k)}_{(k)} (GG^\top)^p\) and \(\hat{A}^2_{k,p} = e^{(k)}_{(k)} (G^\top G)^p\) for \(p \in \mathbb{Z}_+\) and \(1 \leq k \leq q_n\). The following lemma is used when we specify the limit of \(n^{-1}(H_n^1(\sigma) - H_n^1(\sigma_0))\) in the next proposition.

**Lemma 3.7** Assume (A2) and (A4). Then, for any \(p \geq 1\)

\[
n^{-1} \max_{1 \leq k \leq q_n} |\text{tr}(\hat{A}^1_{k,p}) - \text{tr}(\hat{A}^2_{k,p})| \overset{p}{\to} 0
\]

as \(n \to \infty\).

**Proof** By the definition of \(\hat{A}^l_{k,p}\), we obtain

\[
|\text{tr}(\hat{A}^1_{k,p}) - \text{tr}(\hat{A}^2_{k,p})| \\
= \left| \sum_{i: \sup I^1_i \in (s_k-1,s_k]} [(GG^\top)^p]_{ii} - \sum_{j: \sup I^2_j \in (s_k-1,s_k]} [(G^\top G)^p]_{jj} \right| \\
= \left| \sum_{i: \sup I^1_i \in (s_k-1,s_k]} \sum_{i':j} [(GG^\top)^{p-1}]_{i'i'j'} [G]_{i'j} [G^\top]_{jj} - \sum_{j: \sup I^2_j \in (s_k-1,s_k]} \sum_{i,i''} [G^\top]_{ji} [(GG^\top)^{p-1}]_{ii''} [G]_{i'j} \right|.
\]

Two summands in the right-hand side coincide when both \(\sup I^1_i\) and \(\sup I^2_j\) are included or not included in \((s_k-1, s_k]\). In other cases, we have min_{\mu=0,1} |\sup I^1_i -
$s_{k-u} \leq r_n$ if $[G^\top]_{ji} > 0$. Therefore, we obtain

$$|\text{tr}(\hat{A}_{k,p}^1) - \text{tr}(\hat{A}_{k,p}^2)| \leq \left( \sum_{i,j; \text{ sup } I_1 \notin (s_{k-1},s_k)} + \sum_{i,j; \text{ sup } I_2 \notin (s_{k-1},s_k)} \right) \times \sum_{i'} [(GG^\top)^{p-1}]_{ii'} [G]_{i'j} [G^\top]_{ji} \leq \sum_{i; \text{ min}_u = 0,1} [(GG^\top)^{p}]_{ii}.$$  

Thanks to (A2) and (A4), the right-hand side of the above inequality is equal to $O_p(h_n^{-1}) = o_p(n)$.

Let $\mathcal{Y}_1(\sigma) = \lim_{T \to \infty} (T^{-1} \int_0^T y_{1,t}(\sigma) dt)$, where

$$y_{1,t}(\sigma) = -\frac{1}{2} A(\rho_t) \sum_{l=1}^2 B_{l,t}^2 + A(\rho_t) \frac{B_{1,t} B_{2,t} \rho_{t,0}}{\rho_t} + \sum_{l=1}^2 a_0 \left( \frac{1}{2} - \frac{1}{2} B_{l,t}^2 + \log B_{l,t} \right) + \int_{\rho_{t,0}}^{\rho_t} A(\rho) \frac{d\rho}{\rho}.$$  

The limit $\mathcal{Y}_1(\sigma)$ exists under (A1), (A3), and (A4).

**Proposition 3.8** Assume (A1)–(A4). Then

$$\sup_{\sigma \in \Theta_1} |n^{-1} \partial_\sigma^k (H_n^1(\sigma) - H_n^1(\sigma_0)) - \partial_\sigma^k \mathcal{Y}_1(\sigma)| \to 0 \quad (3.14)$$

as $n \to \infty$ for $k \in \{0, 1, 2, 3\}$.

**Proof** Let $A_p^1 = (\tilde{G} \tilde{G}^\top)^p$, $A_p^2 = (\tilde{G}^\top \tilde{G})^p$, $\hat{\Sigma}_{i,j}^l = \hat{\Sigma}_{i,j}^l(\sigma_0)$, and $\hat{\Sigma}_{i,j}^{1,2} = \hat{\Sigma}_{i,j}^{1,2}(\sigma_0)$. Thanks to (A1), for any $\epsilon > 0$, there exists $\delta > 0$, such that $|t - s| < \delta$ implies

$$|\rho_t - \rho_s| \lor |\Sigma_t - \Sigma_s| \lor |\mu_t - \mu_s| < \epsilon \quad (3.15)$$

for any $\sigma$ and $\theta$. We fix such $\delta > 0$, and fix a partition $s_k = k\delta/2$. Then, (3.5) and (A4) yield

$$n^{-1} \text{tr}(S_n^{-1}(\sigma)(S_n(\sigma_0) - S_n(\sigma))) = \frac{1}{n} \text{tr}\left( S_n^{-1}(\sigma) \left( \text{diag}(\hat{\Sigma}_{i,0}^l - \hat{\Sigma}_{i,j}^l) + \hat{\Sigma}_{i,j,0}^{1,2} - \hat{\Sigma}_{i,j}^{1,2} \right) \right) \text{diag}(\hat{\Sigma}_{j,0}^{1,2} - \hat{\Sigma}_{j,j}^{1,2}) = \frac{1}{n} \sum_{p=0}^{\infty} \left\{ \sum_{l=1}^2 \text{tr} \left( \text{diag}\left( \frac{\hat{\Sigma}_{i,0}^l - \hat{\Sigma}_{i,j}^l}{\hat{\Sigma}_{i}^l} \right) A_p^1 \right) - 2 \text{tr} \left( A_p^1 \tilde{G} \left\{ \hat{\Sigma}_{i,j,0}^{1,2} - \hat{\Sigma}_{i,j}^{1,2} \right\} \right) \right\},$$

\(\Box\) Springer
\[
\frac{1}{n} \sum_{p=0}^{q_n+1} \sum_{k=1}^{2} \left\{ \sum_{l=1}^{2} \text{tr} \left( \text{diag} \left( \left( \frac{\hat{\Sigma}_{i,0}^{l}}{\hat{\Sigma}_{i}^{l}} - 1 \right) \right) \right) E_{(k)} A_{p}^{l} \right\} \\
-2\text{tr} \left( E_{(k)} A_{p}^{l} \tilde{G} \left\{ \frac{\hat{\Sigma}_{i,j,0}^{l,2} - \hat{\Sigma}_{i,j}^{l,2}}{\left( \hat{\Sigma}_{i}^{l} \right)^{1/2} \left( \hat{\Sigma}_{j}^{l} \right)^{1/2}} \right\} \right)
\] (3.16)

if \( \tilde{\rho}_n < 1 \).

Let \( \hat{\rho}_k = \rho_{sk-1} \) and \( \hat{B}_{k,l} = \left( \left[ \Sigma_{sk-1}(\sigma) \right]_{ll} / \left[ \Sigma_{sk-1}(\sigma) \right]_{ll} \right)^{1/2} \). Then, (3.15) yields that for any \( p \in \mathbb{Z}_{+} \), we have

\[
\| E_{(k)} A_{p}^{l} \|_{ij} - \hat{\rho}_k^{2p} [ \hat{A}_{k,p}^{l} ]_{ij} \leq C p \tilde{\rho}_n^{2p-1} \epsilon
\] (3.17)
on \( \{ 2pr_n < \delta / 2 \} \). Here, the factor \( p \) in the right-hand side appears, because we consider the difference between \( 2p \) products of \( \rho_{i'} j' \) and \( \hat{\rho}_k^{2p} \). Moreover, Lemma 3.4 and (3.4) yield

\[
\limsup_{n \to \infty} \max_{1 \leq k \leq q_n+1} \sum_{p=0}^\infty \| E_{(k)} A_{p}^{l} \| \leq C \limsup_{n \to \infty} \sum_{p=0}^\infty \hat{\rho}_n^{2p} < \infty
\] (3.18)

almost surely.

Then, together with (A2) and Lemma 3.7, we obtain

\[
n^{-1} \text{tr} \left( S_n^{-1}(\sigma)(S_n(\sigma) - S_n(\sigma)) \right) = \frac{1}{n} \sum_{p=0}^{q_n+1} \sum_{k=1}^{2} \left\{ \hat{\rho}_k^{2p} \sum_{l=1}^{2} (B_{k,l}^{2} - 1) \text{tr}(\hat{A}_{k,p}^{l}) - 2\hat{\rho}_k^{2p+1} (\hat{B}_{k,1} \hat{B}_{k,2} \hat{\rho}_k - \hat{\rho}_k) \text{tr}(\hat{A}_{k,p+1}^{l}) \right\} + \epsilon_n,
\] (3.19)

where \( \hat{\rho}_{k,0} = \rho_{sk-1}(\sigma) \), and \( (\epsilon_n)_{n=1}^{\infty} \) denotes a general sequence of random variables such that \( \limsup_{n \to \infty} |\epsilon_n| \to 0 \) as \( \delta \to 0 \).

Moreover, by (3.3) and Lemma 3.4, we can apply Lemma A.3 in Ogihara (2018) to \( S_n \). Then, we have

\[
\log \text{det} S_n(\sigma) = \log \text{det} \tilde{D} + \log \text{det} \left( E_M + \left( \begin{array}{cc} 0 & \tilde{G} \\ \tilde{G}^\top & 0 \end{array} \right) \right) \\
= \sum_{l=1}^{2} \sum_{i=1}^{M_l} \log \tilde{\Sigma}_i^l + \sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} \text{tr} \left( \left( \begin{array}{cc} 0 & \tilde{G} \\ \tilde{G}^\top & 0 \end{array} \right)^p \right) \\
= \sum_{l=1}^{2} \sum_{i=1}^{M_l} \log \tilde{\Sigma}_i^l - \sum_{p=1}^{\infty} \frac{1}{p} \text{tr}(\tilde{G}\tilde{G}^\top)^p
\]
if $\bar{\rho}_n < 1$. Therefore, thanks to (3.2) and (3.17), we obtain

$$
n^{-1} \log \frac{\det S_n(\sigma)}{\det S_n(\sigma_0)} = n^{-1} \sum_{i=1}^{2} \sum_{l=1}^{M_i} \log \frac{\hat{\Sigma}_i}{\Sigma_{l,0}} - n^{-1} \sum_{p=1}^{\infty} \frac{1}{p} \text{tr}((\hat{G} \hat{G}^\top)^p - (\hat{G} \hat{G}^\top)^p(\sigma_0))
$$

$$
= -n^{-1} \sum_{k=1}^{q_n} \left\{ \sum_{l=1}^{2} M_{l,k} \log \hat{B}^2_{k,l} + \sum_{p=1}^{\infty} \frac{\hat{\rho}^2_{k,p} - \hat{\rho}^2_{k,0}}{p} \text{tr}(\hat{A}^1_{k,p}) \right\} + \varepsilon_n.
$$

(3.20)

Lemma 3.7, (3.6), (3.19), and (3.20) yield

$$
H_n^1(\sigma) - H_n^1(\sigma_0) = -\frac{1}{2} \sum_{p=0}^{\infty} \sum_{k=1}^{q_n} \left\{ \hat{\rho}^2_{k,p} \left( \sum_{l=1}^{2} \hat{B}^2_{k,l} - 1 \right) \text{tr}(\hat{A}^1_{k,p}) \right\}
$$

$$
-2\hat{\rho}^2_{k,p+1} \left( \hat{B}_{k,1} \hat{B}_{k,2} \hat{\rho}_{k,0} - \hat{\rho}_{k} \right) \text{tr}(\hat{A}^1_{k,p+1})
$$

$$
+ \frac{1}{2} \sum_{k=1}^{q_n} \left\{ \sum_{l=1}^{2} M_{l,k} \log \hat{B}^2_{k,l} + \sum_{p=1}^{\infty} \frac{\hat{\rho}^2_{k,p} - \hat{\rho}^2_{k,0}}{p} \text{tr}(\hat{A}^1_{k,p}) \right\} + n\varepsilon_n
$$

$$
= \sum_{k=1}^{q_n} \frac{1}{2} \sum_{p=0}^{\infty} \sum_{l=1}^{2} \hat{\rho}^2_{k,p} \text{tr}(\hat{A}^1_{k,0})
$$

$$
+ \frac{1}{2} \sum_{l=1}^{2} M_{l,k} \log \hat{B}^2_{k,l} + \sum_{p=1}^{\infty} \frac{\hat{\rho}^2_{k,p} - \hat{\rho}^2_{k,0}}{2p} \text{tr}(\hat{A}^1_{k,p}) \right\} + n\varepsilon_n
$$

$$
= \sum_{k=1}^{q_n} \sum_{p=1}^{\infty} \hat{\rho}^2_{k,p} \left\{ \frac{1}{2} \sum_{l=1}^{2} \hat{B}^2_{k,l} \text{tr}(\hat{A}^1_{k,p}) + \frac{\hat{\rho}_{k,0}}{\hat{\rho}_{k}} \hat{B}_{k,1} \hat{B}_{k,2} \text{tr}(\hat{A}^1_{k,p}) \right\}
$$

$$
+ \frac{1}{2} \sum_{l=1}^{2} M_{l,k} \left\{ - \hat{B}^2_{k,l} + 1 + \log \hat{B}^2_{k,l} \right\}
$$

$$
+ \sum_{p=1}^{\infty} \frac{\hat{\rho}^2_{k,p} - \hat{\rho}^2_{k,0}}{2p} \text{tr}(\hat{A}^1_{k,p}) \right\} + n\varepsilon_n.
$$

(3.21)

Here, we used that $\text{tr}(\hat{A}^1_{k,0}) = \text{tr}(\epsilon^1_{(k)_p}) = M_{l,k}$.

Moreover, (A4) and (3.15) yield

$$
\left| \sum_{k=1}^{q_n} f(s_{k-1}) \text{tr}(\hat{A}^1_{k,p}) - h_n^{-1} \int_0^{nh_n} a_p f(t) \text{d}t \right|
$$

$$
\leq \left| \sum_{k=1}^{q_n} f(s_{k-1}) (\text{tr}(\hat{A}^1_{k,p}) - h_n^{-1} a_p (s_k - s_{k-1})) \right|
$$
\[ + \left| h_n^{-1} a_p^1 \sum_{k=1}^{q_n} \int_{s_k-1}^{s_k} (f(t) - f(s_k-1)) \, dt \right| + O_p(h_n^{-1}) \]

\[ \leq o_p(h_n^{-1}) \cdot q_n + C_p \epsilon n + O_p(h_n^{-1}) = o_p(n) + n \epsilon_n \]  

(3.22)

for \( p \geq 1 \) and any choice of \( f(t) = \rho_t^{2p} B_{1,t}^2, \rho_t^{2p-1} \rho_{t,0} B_{1,t} B_{2,t} \) and \( (\rho_t^{2p} - \rho_{t,0}^{2p}) / (2p) \).

Here, we used that \( q_n = O(nh_n) \) by the definition of \( (s_k)_{k=0}^\infty \in \mathcal{S} \). Similarly, we obtain

\[ \sum_{k=1}^{q_n} M_{l,k}(1 - B_{k,i}^2 + \log B_{k,i}^2) = h_n^{-1} \int_0^{nh_n} a_0^1(1 - B_{t,i}^2 + \log B_{t,i}^2) \, dt + n \epsilon_n. \]

Together with (3.21), (A3) and the equation

\[ \sum_{p=1}^\infty a_p^1 \left( \rho_t^{2p} - \rho_{t,0}^{2p} \right) / 2p = \sum_{p=1}^\infty a_p^1 \int_{\rho_{t,0}}^\rho \rho^{2p-1} \, d\rho = \int_{\rho_{t,0}}^\rho A(\rho) / \rho \, d\rho, \]

we obtain

\[ H_n^1(\sigma) - H_n^1(\sigma_0) = n \mathcal{Y}_1(\sigma) + n \epsilon_n. \]

The above arguments show that the supremum with respect to \( \sigma \) of the residual term in the above equation is also equal to \( n \epsilon_n \), and consequently, we obtain (3.14) with \( k = 0 \). Similarly, we obtain (3.14) with \( k \in \{1, 2, 3\} \). \( \square \)

**Proposition 3.9** Assume (A1)–(A4). Then, there exists a positive constant \( \chi \), such that

\[ \mathcal{Y}_1 \leq \liminf_{T \to \infty} \int_0^T \left\{ - \frac{1}{2} (a_0^1 \wedge a_0^2)(B_{1,t} - B_{2,t})^2 - \chi \{ a_1^1(\rho_t - \rho_{t,0})^2 \right. \\
\left. + a_0^1 \wedge a_0^2(B_{1,t} B_{2,t} - 1)^2 \} \right\} \, dt. \]

**Proof** The proof is based on the ideas of proof of Lemma 5 in Ogihara and Yoshida (2014). Let

\[ G_k = ([G]_{ij} \mathbb{1}_{\sup I_1^{ij}, \sup I_2^{ij} \in (s_{k-1}, s_k)})(G)_{ij}, \]

and let \( \bar{A}_{k,p}^j = (G_k G_k^\top)^p \) and \( \bar{A}_{k,p}^2 = (G_k^\top G_k)^p \). Let \( \bar{A}_k = \sum_{p=1}^\infty \rho_k^{2p} \text{tr}(\bar{A}_{k,p}^j) \) and \( \bar{B}_k = \sum_{p=1}^\infty (2p)^{p-1}(\rho_k^{2p} - \rho_k^{2p}) \text{tr}(\bar{A}_{k,p}^j) \). Similarly to the proof of Lemma 3.7, the difference between \( \text{tr}(\bar{A}_{k,p}^j) \) and \( \text{tr}(\bar{A}_{k,p}^j) \) comes from terms with \( \sup I_1^{ij} \) close to \( s_{k-1} \) or \( s_k \), and hence, we obtain

\[ \max_{1 \leq k \leq q_n} \left| \text{tr}(\bar{A}_{k,p}^j) - \text{tr}(\bar{A}_{k,p}^j) \right| = o_p(n). \]
Therefore, (3.21) yields

\[
\mathcal{Y}_1 = \frac{1}{n} \sum_{k=1}^{q_n} \left\{ -\frac{1}{2} (\tilde{B}_{k,1}^2 + \tilde{B}_{k,2}^2) \tilde{A}_k + \frac{\hat{\rho}_k,0}{\hat{\rho}_k} \tilde{B}_{k,1} \tilde{B}_{k,2} \tilde{A}_k \\
+ \frac{1}{2} \sum_{l=1}^{2} M_{l,k} (1 - \tilde{B}_{k,l}^2 + \log \tilde{B}_{k,l}^2) + \tilde{B}_k \right\} + e_n
\]

\[
= \frac{1}{n} \sum_{k=1}^{q_n} \left\{ -\frac{1}{2} (\tilde{B}_{k,1}^2 - \tilde{B}_{k,2}^2)^2 \tilde{A}_k + \tilde{B}_{k,1} \tilde{B}_{k,2} \left( \tilde{A}_k \frac{\hat{\rho}_k,0}{\hat{\rho}_k} - \tilde{A}_k \right) \\
+ \frac{1}{2} \sum_{l=1}^{2} M_{l,k} (1 - \tilde{B}_{k,l}^2 + \log \tilde{B}_{k,l}^2) + \tilde{B}_k \right\} + e_n.
\]

Then, since

\[
\frac{1}{2} \sum_{l=1}^{2} M_{l,k} (1 - \tilde{B}_{k,l}^2 + \log \tilde{B}_{k,l}^2)
= M_{1,k} \left( 1 - \frac{1}{2} - \frac{\tilde{B}_{k,2}^2}{2} + \log (\tilde{B}_{k,1} \tilde{B}_{k,2}) \right) + \frac{M_{2,k} - M_{1,k}}{2} \left( 1 - \frac{\tilde{B}_{k,2}^2}{2} + \log (\tilde{B}_{k,2}^2) \right)
= -\frac{1}{2} M_{1,k} (\tilde{B}_{k,1} - \tilde{B}_{k,2})^2 - M_{1,k} \tilde{B}_{k,1} \tilde{B}_{k,2} + M_{1,k} \left( 1 + \log (\tilde{B}_{k,1} \tilde{B}_{k,2}) \right) + \frac{M_{2,k} - M_{1,k}}{2} \left( 1 - \frac{\tilde{B}_{k,2}^2}{2} + \log (\tilde{B}_{k,2}^2) \right),
\]

and a similar estimate holds by switching the roles of $M_{1,k}$ and $M_{2,k}$, we have

\[
\mathcal{Y}_1 = n^{-1} \sum_{k=1}^{q_n} \left\{ -\frac{1}{2} (M_{1,k} + \tilde{A}_k)(\tilde{B}_{k,1} - \tilde{B}_{k,2})^2 + M_{1,k} (1 + \log (\tilde{B}_{k,1} \tilde{B}_{k,2})) \\
+ \tilde{B}_k + \frac{M_{2,k} - M_{1,k}}{2} \left( 1 - \frac{\tilde{B}_{k,2}^2}{2} + \log (\tilde{B}_{k,2}^2) \right) + \tilde{B}_{k,1} \tilde{B}_{k,2} \left( \tilde{A}_k \frac{\hat{\rho}_k,0}{\hat{\rho}_k} - \tilde{A}_k - M_{1,k} \right) \right\} + e_n
\]

\[
= n^{-1} \sum_{k=1}^{q_n} \left\{ -\frac{1}{2} (M_{2,k} + \tilde{A}_k)(\tilde{B}_{k,1} - \tilde{B}_{k,2})^2 + M_{2,k} (1 + \log (\tilde{B}_{k,1} \tilde{B}_{k,2})) \\
+ \tilde{B}_k + \frac{M_{1,k} - M_{2,k}}{2} \left( 1 - \frac{\tilde{B}_{k,1}^2}{2} + \log (\tilde{B}_{k,1}^2) \right) + \tilde{B}_{k,1} \tilde{B}_{k,2} \left( \tilde{A}_k \frac{\hat{\rho}_k,0}{\hat{\rho}_k} - \tilde{A}_k - M_{2,k} \right) \right\} + e_n.
\]

For $l \in \{1, 2\}$, let

\[
F_{l,k} = M_{l,k} (1 + \log (\tilde{B}_{k,1} \tilde{B}_{k,2})) + \tilde{B}_k + \tilde{B}_{k,1} \tilde{B}_{k,2} \left( \tilde{A}_k \frac{\hat{\rho}_k,0}{\hat{\rho}_k} - \tilde{A}_k - M_{l,k} \right).
\]
then since \(1 - x + \log x \leq 0\) for \(x > 0\), we obtain

\[
\mathcal{Y}_1 \leq n^{-1} \sum_{k=1}^{q_n} \left[ \left\{ -\frac{1}{2} (M_{1,k} + \tilde{A}_k) (\tilde{B}_{k,1} - \tilde{B}_{k,2})^2 + F_{1,k} \right\} 1_{\{M_{2,k} \geq M_{1,k}\}} + \right. \\
\left. + \left\{ -\frac{1}{2} (M_{2,k} + \tilde{A}_k) (\tilde{B}_{k,1} - \tilde{B}_{k,2})^2 + F_{2,k} \right\} 1_{\{M_{2,k} < M_{1,k}\}} \right] + e_n,
\]

and therefore, we have

\[
\mathcal{Y}_1 \leq n^{-1} \sum_{k=1}^{q_n} \left[ -\frac{1}{2} (M_{1,k} \wedge M_{2,k} + \tilde{A}_k) (\tilde{B}_{k,1} - \tilde{B}_{k,2})^2 + F_{1,k} \lor F_{2,k} \right] + e_n. \quad (3.23)
\]

Let \((\lambda_i^k)_{i=1}^{M_{1,k}}\) be all the eigenvalues of \(G_k G_k^\top\). Similarly to Lemma 3.3, we have \(0 \leq \lambda_i^k \leq 1\). Then, we have

\[
F_{1,k} = \sum_{i=1}^{M_{1,k}} \left\{ 1 + \log(\tilde{B}_{k,1} \tilde{B}_{k,2}) + \tilde{B}_{k,1} \tilde{B}_{k,2} \sum_{p=0}^{\infty} (\lambda_i^k)^{p+1} \hat{\rho}_k^{2p+1} \hat{\rho}_{k,0} - (\lambda_i^k)^p \hat{\rho}_k^{2p} \right\}
\]

\[
+ \sum_{p=1}^{\infty} (\lambda_i^k)^p \left( \frac{\hat{\rho}_k^{2p} - \hat{\rho}_{k,0}^{2p}}{2p} \right).
\]

Moreover, by setting \(g_i^k = \sqrt{1 - \lambda_i^k \hat{\rho}_k^{2} \tilde{g}_{i,0}^k} = \sqrt{1 - \lambda_i^k \hat{\rho}_k^{2}} \), and \(F(x) = 1 - x + \log x\), we have

\[
F_{1,k} = \sum_{i=1}^{M_{1,k}} \left\{ 1 + \tilde{B}_{k,1} \tilde{B}_{k,2} (g_i^k)^{-2} (\lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} - 1) + \log(\tilde{B}_{k,1} \tilde{B}_{k,2} g_{i,0}^k (g_i^k)^{-1}) \right\}
\]

\[
= \sum_{i=1}^{M_{1,k}} \left\{ \tilde{B}_{k,1} \tilde{B}_{k,2} (g_i^k)^{-2} (\lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} - 1) + \tilde{B}_{k,1} \tilde{B}_{k,2} g_{i,0}^k (g_i^k)^{-1} + F(\tilde{B}_{k,1} \tilde{B}_{k,2} g_{i,0}^k (g_i^k)^{-1}) \right\}.
\]

Here, we also used the expansion formulas \((1-x)^{-1} = \sum_{p=0}^{\infty} x^p \) and \(-\log(1-x) = \sum_{p=1}^{\infty} x^p / p \) for \(|x| < 1\).

Let

\[
\mathcal{R} = \sup_{t, \sigma, 0 \leq \ell \leq 4} (|\partial^l \Sigma_t|^{1/2} \lor |\partial^l \Sigma_t^{-1}|^{1/2}).
\]

Since \(g_i^k \leq 1\), \(0 \leq \lambda_i^k \leq 1\), and \(|\hat{\rho}_k| < 1\), we have

\[
(g_i^k)^{-2} (\lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} - 1) + g_{i,0}^k (g_i^k)^{-1} = \frac{(\lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} - 1 + g_{i,0}^k g_i^k)(1 - \lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} + g_{i,0}^k g_i^k)}{(g_i^k)^2 (1 - \lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} + g_{i,0}^k g_i^k)}
\]

\(\mathcal{R} \) Springer
By a similar argument we have

\[ \frac{(\lambda_i^k \hat{\rho}_k - \hat{\rho}_{k,0})^2}{(g_i^k)^2(1 - \lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} + \hat{g}_{i,0}^k g_i^k)} \leq \frac{\lambda_i^k (\hat{\rho}_k - \hat{\rho}_{k,0})^2}{(g_i^k)^2(1 - \lambda_i^k \hat{\rho}_k \hat{\rho}_{k,0} + \hat{g}_{i,0}^k g_i^k)} \leq \frac{\lambda_i^k}{3} (\hat{\rho}_k - \hat{\rho}_{k,0})^2. \]

Together with Lemma 11 in Ogihara and Yoshida (2014) and

\[ \hat{B}_{k,1} \hat{B}_{k,2} g_{i,0}^k = \frac{\hat{B}_{k,1} \hat{B}_{k,2} \sqrt{1 - \lambda_i^k \hat{\rho}_{k,0}^2} - \sqrt{\lambda_i^k \hat{\rho}_{k,0}^2}}{\sqrt{1 - \lambda_i^k \hat{\rho}_{k,0}^2}} \leq \hat{B}_{k,1} \hat{B}_{k,2} \leq \frac{\mathcal{R}^4}{\sqrt{1 - \hat{\rho}_{n,0}^2}}, \]

we have

\[ \hat{B}_{k,1} \hat{B}_{k,2} g_{i,0}^k (g_i^k)^{-1} - 1 = \frac{-\hat{B}_{k,1} \hat{B}_{k,2} \lambda_i^k (\hat{\rho}_k - \hat{\rho}_{k,0})^2 - \frac{1}{4} \hat{\rho}_{n,0}^2 \hat{B}_{k,1} \hat{B}_{k,2} g_{i,0}^k (g_i^k)^{-1} - 1)^2}{\lambda_i^k (\hat{\rho}_k - \hat{\rho}_{k,0})^2}. \]

Moreover, the inequality \( a^2 \geq (a + b)^2/2 - b^2 \) with \( a = \hat{B}_{k,1} \hat{B}_{k,2} g_{i,0}^k - g_i^k \) and \( b = g_i^k - g_{i,0}^k \) yields

\[ (\hat{B}_{k,1} \hat{B}_{k,2} g_{i,0}^k (g_i^k)^{-1} - 1)^2 \geq \frac{(\hat{B}_{k,1} \hat{B}_{k,2} g_{i,0}^k - g_i^k)^2}{2} (\hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 - \frac{(\hat{B}_{k,1} \hat{B}_{k,2} g_{i,0}^k - g_i^k)^2}{2} = \frac{1}{2} \left( \lambda_i^k \hat{\rho}_{k,0}^2 (\hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 + \frac{\lambda_i^k}{16} \hat{\rho}_{n,0}^2 \hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 \right) \geq \frac{1}{2} \left( \lambda_i^k \hat{\rho}_{k,0}^2 (\hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 + \frac{\lambda_i^k}{16} \hat{\rho}_{n,0}^2 \hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 \right), \]

and hence, we have

\[ F_{1,k} \leq \sum_{i=1}^{M_{1,k}} \left\{ -\frac{\hat{B}_{k,1} \hat{B}_{k,2} \lambda_i^k (\hat{\rho}_k - \hat{\rho}_{k,0})^2}{3} \left( \hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 + \frac{\lambda_i^k}{16} \hat{\rho}_{n,0}^2 \hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 \right) \right\} \]

\[ = \left( \frac{\hat{B}_{k,1} \hat{B}_{k,2}}{3} - \frac{1}{16} \mathcal{R}^8 \right) \left( \hat{A}_{k,1}^1 (\hat{\rho}_k - \hat{\rho}_{k,0})^2 - \frac{(1 - \hat{\rho}_{n,0}^2)^2}{8} \left( \hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 + \frac{\lambda_i^k}{16} \hat{\rho}_{n,0}^2 \hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 \right). \]

By a similar argument for \( F_{2,k} \), there exists a positive constant \( \bar{\chi} \) which does not depend on \( k \) nor \( n \), such that

\[ F_{1,k} \lor F_{2,k} \leq -\bar{\chi} (1 - \hat{\rho}_{n,0}^2)^2 \left( \hat{A}_{k,1}^1 (\hat{\rho}_k - \hat{\rho}_{k,0})^2 + \frac{M_{1,k} \hat{B}_{k,1} \hat{B}_{k,2} - 1)^2}{8} \right). \]
Together with (3.23), we have
\[
\mathcal{Y}_1 \leq n^{-1} \sum_{k=1}^{\eta n} \left\{ -\frac{1}{2} (M_{1,k} \wedge M_{2,k})(\hat{B}_{k,1} - \hat{B}_{k,2})^2 \\
- \tilde{\chi}(1 - \hat{\rho}_n^2) \left\{ \text{tr}(\hat{A}_{k,1}^\dagger)(\hat{\rho}_k - \hat{\rho}_{k,0})^2 + M_{1,k} \wedge M_{2,k}(\hat{B}_{k,1} \hat{B}_{k,2} - 1)^2 \right\} \right\} + e_n.
\]

By letting \( n \to \infty \), (A4) and (3.3) yield the conclusion. \( \square \)

(A6) and Remark 4 in Ogihara and Yoshida (2014) yield that
\[
\limsup_{T \to \infty} \frac{1}{T} \int_0^T \left\{ |B_{1,t} - B_{2,t}|^2 + |B_{1,t} B_{2,t} - 1|^2 + |\rho_t - \rho_{t,0}|^2 \right\} dt > 0,
\]
when \( \sigma \neq \sigma_0 \).

Then, by Proposition 3.9, we have \( \mathcal{Y}_1(\sigma) < 0 \) (note that \( a_0^1 \wedge a_0^2 \geq a_1^1 \) by (2.3) and a similar argument). Therefore, for any \( \delta > 0 \), there exists \( \eta > 0 \), such that
\[
\inf_{|\sigma - \sigma_0| \geq \delta} (-\mathcal{Y}_1(\sigma)) \geq \eta.
\]

Then, since \( H_n^1(\hat{\sigma}_n) - H_n^1(\sigma_0) \geq 0 \) by the definition, for any \( \epsilon > 0 \), we have
\[
P(|\hat{\sigma}_n - \sigma_0| \geq \delta) \leq P\left( \sup_{|\sigma - \sigma_0| \geq \delta} \left( H_n^1(\sigma) - H_n^1(\sigma_0) \right) \geq 0 \right) \leq P\left( \sup_{|\sigma - \sigma_0| \geq \delta} \left( n^{-1}(H_n^1(\sigma) - H_n^1(\sigma_0)) - \mathcal{Y}_1(\sigma) \right) \geq \eta \right) \leq P\left( \sup_{\sigma} |n^{-1}(H_n^1(\sigma) - H_n^1(\sigma_0)) - \mathcal{Y}_1(\sigma)| \geq \eta \right) < \epsilon \quad (3.24)
\]
for sufficiently large \( n \) by Proposition 3.8, which implies \( \hat{\sigma}_n \xrightarrow{P} \sigma_0 \) as \( n \to \infty \).

### 3.3 Asymptotic normality of \( \hat{\sigma}_n \)

Let \( S_{n,0} = S_n(\sigma_0) \) and \( \Sigma_{t,0} = \Sigma_t(\sigma_0) \). (3.7) and the equation \( \partial_\sigma S_{n,0}^{-1} = -S_{n,0}^{-1} \partial_\sigma S_{n,0} S_{n,0}^{-1} \) imply
\[
\partial_\sigma H_n^1(\sigma_0) = -\frac{1}{2} (\Delta X^c)^\top \partial_\sigma S_{n,0}^{-1} \Delta X^c - \frac{1}{2} \text{tr}(\partial_\sigma S_{n,0} S_{n,0}^{-1}) + o_p(\sqrt{n})
= -\frac{1}{2} \text{tr}(\partial_\sigma S_{n,0}^{-1} (\Delta X^c)^\top (\Delta X^c) - S_{n,0}) + o_p(\sqrt{n}).
\] (3.25)

Let \( (L_n)_{n \in \mathbb{N}} \) be a sequence of positive integers such that \( L_n \to \infty \) and \( L_n n^\eta (nh_n)^{-1} \to 0 \) as \( n \to \infty \) for some \( \eta > 0 \). Let \( \hat{s}_k = k T_n / L_n \) for \( 0 \leq k \leq L_n \), let
Let $J^k = (\tilde{s}_{k-1}, \tilde{s}_k)$, and let $S_{n,0}^{(k)}$ be an $M \times M$ matrix satisfying

$$[S_{n,0}^{(k)}]_{ij} = \int_{I_i \cap I_j \cap J^k} [\Sigma_t,0] \psi(i),\psi(j) \, dt.$$  

For a two-dimensional stochastic process $(U_t)_{t \geq 0} = ((U_t^1, U_t^2))_{t \geq 0}$, let $\Delta_{i,t}^{(k)} U = U_{i,t}^l (S_{n,l}^0 \vee \tilde{s}_{k-1}) \wedge \tilde{s}_k \wedge t - U_{i,t}^l (S_{n,l}^0 \vee \tilde{s}_{k-1}) \wedge \tilde{s}_k \wedge t$, and let $\Delta_{i,t}^{(k)} U = \Delta_{\psi(i),t}^{(k)} U$ for $1 \leq i \leq M$. Let

$$\Delta_{i,t}^{(k)} U = \Delta_{i,t}^{(k)} U_0,$$

and let $\Delta^{(k)} U = (\Delta^{(k)} U_{i,t})_{1 \leq i \leq M}$. Let

$$X_k = -\frac{1}{2\sqrt{n}} \left\{ (\Delta^{(k)} X^c)^\top \partial_{\sigma} S_{n,0}^{-1} \Delta^{(k)} X^c - \text{tr}(\partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} \Delta^{(k)} X^c) \right\} - \frac{1}{\sqrt{n}} \sum_{k' < k} (\Delta^{(k)} X^c)^\top \partial_{\sigma} S_{n,0}^{-1} \Delta^{(k)} X^c.$$

Then, since $\Delta X^c = \sum_{k=1}^{L_n} \Delta^{(k)} X^c$ and $S_{n,0} = \sum_{k=1}^{L_n} S_{n,0}^{(k)}$, (3.25) yields

$$n^{-1/2} \partial_{\sigma} H_n^1(\sigma_0) = \sum_{k=1}^{L_n} X_k + o_p(1). \quad (3.26)$$

Moreover, Itô’s formula yields

$$\sqrt{n} X_k = -\frac{1}{2} \sum_{i,j} \left\{ \partial_{\sigma} S_{n,0}^{-1} \int_{I_i \cap J^k} 2 \int_{I_i \cap J^k} (\Delta^{(k)} X^c)^{\psi(i)} + 2 \sum_{k' < k} \int_{I_i \cap J^k} (\Delta^{(k)} X^c)^{\psi(i)} \right\}$$

$$= -\sum_{i,j} \left\{ \partial_{\sigma} S_{n,0}^{-1} \int_{I_i \cap J^k} (\Delta^{(k)} X^c)^{\psi(i)} \right\}. \quad (3.27)$$

Let $G_t = \mathcal{F}_t \vee \sigma(\{\Pi_n\})$ for $t \geq 0$. We will show

$$n^{-1/2} \partial_{\sigma} H_n^1(\sigma_0) \xrightarrow{d} N(0, \Gamma_1), \quad (3.28)$$

using Corollary 3.1 and the remark after that in Hall and Heyde (1980). For this purpose, it is sufficient to show

$$\sum_{k=1}^{L_n} E_k[X_k^2] \xrightarrow{p} \Gamma_1, \quad (3.29)$$

and

$$\sum_{k=1}^{L_n} E_k[X_k^4] \xrightarrow{p} 0, \quad (3.30)$$

by (3.26), where $E_k$ denotes the conditional expectation with respect to $G_{\tilde{s}_{k-1}}$.

We first show four auxiliary lemmas. Let $M_k = \#\{i; 1 \leq i \leq M, \sup I_i \in J^k\}$. 

 Springer
Lemma 3.10 Assume (A1). Then, there exists a positive constant $C$, such that $\|D^{-1/2}S_{n,0}^{(k)}D^{-1/2}\| \leq C$ and $\text{tr}(D^{-1/2}S_{n,0}^{(k)}D^{-1/2}) \leq C(\tilde{M}_k + 1)$ for any $1 \leq k \leq L_n$.

Proof Since

$$\|S_{n,0}^{(k)}ij\| \leq C \left[ D^{1/2} \left( \frac{\mathcal{E}_{M_1}}{G^T} G \right) D^{1/2} \right]_{ij},$$

Lemma 3.3 yields

$$\|D^{-1/2}S_{n,0}^{(k)}D^{-1/2}\| \leq C \left( \frac{\mathcal{E}_{M_1}}{G^T} G \right) \leq C.$$

Moreover, we have

$$\text{tr}(D^{-1/2}S_{n,0}^{(k)}D^{-1/2}) = C \sum_{i=1}^{M} \int_{I_i \cap J_k} |\sigma_{ij}^0 S_{n,0}^{(k)}ij| dt \leq C \sum_{i=1}^{M} 1_{\{I_i \cap J_k \neq \emptyset\}} \leq C(\tilde{M}_k + 1).$$

Lemma 3.11 Assume (A4) and that $nh_n L_n^{-1} \rightarrow \infty$ as $n \rightarrow \infty$. Then, $\{L_n n^{-1} \max_{1 \leq k \leq L_n} \tilde{M}_k\}_{n=1}^{\infty}$ is $P$-tight.

Proof Let $\mathcal{M}_n = [nh_n L_n^{-1}]$. We define a partition of $[0, \infty)$ by

$$s_j = \frac{nh_n j}{2L_n \mathcal{M}_n} \quad (j \geq 0).$$

Then, $(s_j)_{j=0}^{\infty} \in \mathcal{S}$ when $nh_n L_n^{-1} \geq 1$, and $(s_j)_{j=0}^{2L_n \mathcal{M}_n}$ is a subpartition of $(\tilde{s}_k)_{k=0}^{L_n}$.

For $M_{l,j}$ which corresponds to this partition ($\tilde{M}_k$ remains to be defined using $\tilde{s}_k$), we have

$$\tilde{M}_k = \sum_{l=1}^{2M_n k} \sum_{j=2M_n (k-1)+1}^{2M_n k} M_{l,j},$$

since $\tilde{s}_k = nh_n k L_n^{-1} = s_{2M_n k}$. Therefore, (A4) yields

$$\max_{1 \leq k \leq L_n} \tilde{M}_k \leq 4M_n \max_{l,j} M_{l,j} \leq C \mathcal{M}_n [h_n^{-1}(a_0^1 \lor a_0^2) + o_p(h_n^{-1})] = O_p(n L_n^{-1}).$$

Lemma 3.12 Assume (A1). Then

$$\|D^{-1/2}S_{n,0}^{(k)}D^{-1/2}S_{n,0}^{(k')}D^{-1/2}\| \leq C \frac{(Q_n + 1)\tilde{\rho}_n Q_n}{(1 - \tilde{\rho}_n)^2}$$

on $\{\tilde{\rho}_n < 1\}$ for $|k - k'| > 1$, where $Q_n = [r_n^{-1}(T_n/L_n - 2r_n)]$.

\(\Box\) Springer
Proof Using the expansion formula (3.5), we have

\[ S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} = -S_{n,0}^{(k)} S_{n,0}^{-1} \partial_{\sigma} S_{n,0} S_{n,0}^{(k)} \]

\[ = -S_{n,0}^{(k)} \tilde{D}^{-1/2} \sum_{p=0}^{\infty} (-1)^p \begin{pmatrix} 0 & \tilde{G} \\ \tilde{G}^T & 0 \end{pmatrix}^p \tilde{D}^{-1/2} \partial_{\sigma} S_{n,0} \tilde{D}^{-1/2} \]

\[ \times \sum_{q=0}^{\infty} (-1)^q \begin{pmatrix} 0 & \tilde{G} \\ \tilde{G}^T & 0 \end{pmatrix}^q \tilde{D}^{-1/2} S_{n,0} \]

\[ = \sum_{p,q=0}^{\infty} (-1)^{p+q} S_{n,0}^{(k)} \mathcal{E}_{p,q} \mathcal{E}_{p,q}^{(k)} S_{n,0} \quad (3.31) \]

if \( \bar{\rho}_n < 1 \), where

\[ \mathcal{E}_{p,q}^{(k)} = \tilde{D}^{-1/2} \begin{pmatrix} 0 & \tilde{G} \\ \tilde{G}^T & 0 \end{pmatrix}^p \tilde{D}^{-1/2} \partial_{\sigma} S_{n,0} \tilde{D}^{-1/2} \begin{pmatrix} 0 & \tilde{G} \\ \tilde{G}^T & 0 \end{pmatrix}^q \tilde{D}^{-1/2}. \]

We consider a necessary condition for

\[ [S_{n,0}^{(k)} \mathcal{E}_{p,q}^{(k)} S_{n,0}^{(k)}]_i' j' = \sum_{i,j} [S_{n,0}^{(k)}]_i' j' [\mathcal{E}_{p,q}^{(k)}]_i j' [S_{n,0}^{(k)}]_j j' \quad (3.32) \]

to be zero for any \( i' \) and \( j' \). We first observe that the element \( [\mathcal{E}_{p,q}^{(k)}]_i j \) is equal to zero if \( [S_{n,0}^{(k)}]_i j = 0 \), where

\[ \tilde{S} = \begin{pmatrix} \mathcal{E}_{M_1} & \tilde{G} \\ \tilde{G}^T & \mathcal{E}_{M_2} \end{pmatrix}. \]

Moreover, \( [S_{n,0}^{(k)}]_i' j' \neq 0 \) only if \( I_i \cap J^k \neq \emptyset \), and \( [S_{n,0}^{(k)}]_j j' \neq 0 \) only if \( I_j \cap J^k \neq \emptyset \). Since \( \inf_{x \in I_i, y \in I_j} |x - y| > T_n/L_n - 2r_n \) if \( I_i \cap J^k \neq \emptyset \) and \( I_j \cap J^k \neq \emptyset \), we have \( [\tilde{S}^r]_j = 0 \) for \( r \leq \mathcal{Q}_n \) when \( [S_{n,0}^{(k)}]_i' j' \neq 0 \) and \( [S_{n,0}^{(k)}]_j j' \neq 0 \). Therefore, \( [S_{n,0}^{(k)} \mathcal{E}_{p,q}^{(k)}]_i' j' = 0 \) for any \( i' \) and \( j' \) if \( p + q + 1 \leq \mathcal{Q}_n \).

Then, (3.31) and Lemmas 3.4, 3.5 and 3.10 yield

\[ \| \tilde{D}^{-1/2} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \tilde{D}^{-1/2} \|
\]
\[
\leq C \frac{(Q_n + 1) \hat{\rho}_n^Q_n}{(1 - \hat{\rho}_n)^2}
\]
on \{\hat{\rho}_n < 1\}. \qed

**Lemma 3.13** Let \(m \in \mathbb{N}\). Let \(V\) be an \(m \times m\) symmetric, positive definite matrix and \(A\) be a \(m \times m\) matrix. Let \(X\) be a random variable following \(N(0, V)\). Then

\[
E[(X^T AX)^2] = \text{tr}(AV)^2 + 2\text{tr}((AV)^2),
\]

\[
E[(X^T AX)^3] = \text{tr}(AV)^3 + 6\text{tr}(AV)\text{tr}((AV)^2) + 8\text{tr}((AV)^3),
\]

\[
E[(X^T AX)^4] = \text{tr}(AV)^4 + 12\text{tr}(AV)^2\text{tr}((AV)^2) + 12\text{tr}((AV)^2)^2
\]

\[+32\text{tr}(AV)\text{tr}((AV)^2) + 48\text{tr}((AV)^4).
\]

**Proof** We only show the result for \(E[(X^T AX)^4]\). Let \(U\) be an orthogonal matrix and \(\Lambda\) be a diagonal matrix satisfying \(UVU^T = \Lambda\). Then, we have \(UX \sim N(0, \Lambda)\), and

\[
E \left[\prod_{i=1}^{8} (UX)_{l_i}^2\right] = \sum_{(l_{2q-1}, l_{2q})^4} 4 \prod_{q=1}^{4} [\Lambda]_{l_{2q-1}, l_{2q}},
\]

where the summation of \((l_{2q-1}, l_{2q})^4\) is taken over all disjoint pairs of \(\{j_1, \ldots, j_8\}\). Then, by setting \(B = UAU^T\), we have

\[
E[(X^T AX)^4] = \sum_{j_1, \ldots, j_8} \sum_{(l_{2q-1}, l_{2q})_q} 4 \prod_{p=1}^{4} [B]_{j_{2p-1}, j_{2p}} \prod_{q=1}^{4} [\Lambda]_{l_{2q-1}, l_{2q}},
\]

Let \(n C_k = \frac{n!}{k!(n-k)!}\). Out of \(j_1, \ldots, j_8\), we connect \(j_{2p-1}\) to \(j_{2p}\) and \(l_{2q-1}\) to \(l_{2q}\) \((1 \leq p, q \leq 4)\). Then, the pattern of the connected components gives five different cases.

1. Four connected components (four components of size 2): only one case of the pairs \((l_{2q-1}, l_{2q})^4\) appears, which corresponds to \(\text{tr}((B\Lambda)^4)\).

2. Three connected components (a component of size 4 and two components of size 2):
   
   The choice of elements for a component of size 4 gives \(4C_2\) ways, and the choice of the pair \((l_{2q-1}, l_{2q})\) for this component gives two ways, and hence, \(4C_2 \times 2 = 12\) ways in total. This case corresponds to \(\text{tr}(B\Lambda)^2\text{tr}((B\Lambda)^2)\).

3. Two connected components (two components of size 4): The choice of elements for each component gives \(4C_2\) ways, excluding duplicates, and the choice of the pair \((l_{2q-1}, l_{2q})\) for each component gives two ways, and hence, \(4C_2 \times 2 \times 2 = 12\) ways in total. This case corresponds to \(\text{tr}((B\Lambda)^2)^2\).

4. Two connected components (a component of size 6 and a component of size 2): The choice of elements for a component of size 6 gives \(4C_1\) ways, and the choice of the pair \((l_{2q-1}, l_{2q})\) for this component gives \(4 \times 2 = 8\) ways, and hence \(4C_1 \times 8 = 32\) ways in total. This case corresponds to \(\text{tr}(B\Lambda)\text{tr}((B\Lambda)^2)\).
5. One connected component (a component of size 8): The choice of the pair 
\((l_2q-1, l_2q)\) gives \(6 \times 4 \times 2 = 48\) ways. This case corresponds to \(\text{tr}((B\Lambda)^4)\).

Then, we obtain the conclusion.

\[\square\]

**Proposition 3.14** Assume (A1)–(A4) and (A6). Then

\[
n^{-1/2} \partial_\sigma H_n^1(\sigma_0) \xrightarrow{d} N(0, \Gamma_1)
\]

as \(n \to \infty\).

**Proof** It is sufficient to show (3.29) and (3.30). Let \(\mathfrak{A}_k = (\Delta^{(k)} X^c)^\top \partial_\sigma S_{n,0}^{-1} \Delta^{(k)} X^c\) and \(\mathfrak{B}_k = \partial_\sigma S_{n,0}^{-1} S_{n,0}^{(k)}\). By the definition of \(\mathfrak{A}_k\), we have

\[
\sum_{k=1}^{L_n} E_k[A_k^4]
\]

\[
\leq \frac{C}{n^2} \sum_{k=1}^{L_n} \left\{ E_k\left[\left((\Delta^{(k)} X^c)^\top \partial_\sigma S_{n,0}^{-1} \Delta^{(k)} X^c - \text{tr}(\partial_\sigma S_{n,0}^{-1} S_{n,0}^{(k)})\right)^4\right]\right.
\]

\[
+ E_k\left[\left(\sum_{k' < k} (\Delta^{(k')} X^c)^\top \partial_\sigma S_{n,0}^{-1} \Delta^{(k')} X^c\right)^4\right]\}
\]

\[
= \frac{C}{n^2} \sum_{k=1}^{L_n} \left\{ E_k[A_k^4] - 4 E_k[A_k^3] \text{tr}(\mathfrak{B}_k) + 6 E_k[A_k^2] \text{tr}(\mathfrak{B}_k)^2 - 4 \text{tr}(\mathfrak{B}_k)^4 + \text{tr}(\mathfrak{B}_k)^4\right\}
\]

\[
+ \frac{C}{n^2} \sum_{k=1}^{L_n} \left\{ \left(\sum_{k' < k} \Delta^{(k')} X^c\right)^\top \partial_\sigma S_{n,0}^{-1} S_{n,0}^{(k)} \partial_\sigma S_{n,0}^{-1} \left(\sum_{k' < k} \Delta^{(k')} X^c\right)\right\}^2. \quad (3.33)
\]

Thanks to Lemmas 3.13, 3.10, 3.11 and 3.5, (3.4), and Lemma A.1 in Ogihara (2018), the first term in the right-hand side is calculated as

\[
\frac{C}{n^2} \sum_{k=1}^{L_n} \left\{ \text{tr}(\mathfrak{B}_k)^4 + 12 \text{tr}(\mathfrak{B}_k)^2 \text{tr}(\mathfrak{B}_k^2) + 12 \text{tr}(\mathfrak{B}_k^2)^2 + 32 \text{tr}(\mathfrak{B}_k) \text{tr}(\mathfrak{B}_k^2) + 48 \text{tr}(\mathfrak{B}_k^4) - 4 \text{tr}(\mathfrak{B}_k)^3 + 6 \text{tr}(\mathfrak{B}_k) \text{tr}(\mathfrak{B}_k^2) + 8 \text{tr}(\mathfrak{B}_k^3)\right\}
\]

\[
+ 2 \text{tr}(\mathfrak{B}_k^2) \right\} - 3 \text{tr}(\mathfrak{B}_k)^4\}
\]

\[
= \frac{C}{n^2} \sum_{k=1}^{L_n} \left\{ 48 \text{tr}(\mathfrak{B}_k^4) + 12 \text{tr}(\mathfrak{B}_k^2)^2\right\}
\]

\[
\leq \frac{C}{n^2} (\max_k \tilde{M}_k + 1)^2 L_n(1 - \tilde{\rho}_n)^{-8} 1_{\tilde{\rho}_n < 1} + o_p(1) \xrightarrow{P} 0. \quad (3.34)
\]
Moreover, Lemma 3.13 yields

\[
E \Pi \left[ \frac{C}{n^2} \sum_{k=1}^{L} \left\{ \left( \sum_{k' < k} \Delta^{(k')} X^{c} \right)^{\top} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} \left( \sum_{k' < k} \Delta^{(k')} X^{c} \right) \right\}^2 \right]
\]

\[
\leq \frac{C}{n^2} \sum_{k=1}^{L} \sum_{k_1', k_2' < k} \left\{ |\text{tr}(\partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k_1')})| + |\text{tr}(\partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k_2')})| \right\}.
\]

(3.35)

If \( k_1' < k - 1 \), Lemmas 3.5 and 3.12, Lemma A.1 in Ogihara (2018) and the equation \( \partial_{\sigma} S_{n,0}^{-1} = -S_{n,0}^{-1} \sigma_{n,0} S_{n,0}^{-1} \) yield

\[
|\text{tr}(\partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k_1')})| = |\text{tr}(\tilde{D}^{1/2} S_{n,0}^{-1} \tilde{D}^{1/2} \tilde{D}^{-1/2} S_{n,0}^{-1} \tilde{D}^{1/2} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k_1')})| \leq C M Q_{n} \tilde{\rho}_{n}^{Q_{n}} (1 - \tilde{\rho}_{n})^{-4}
\]

on \{ \tilde{\rho}_{n} < 1 \}. Here, we used that \( |\text{tr}(\tilde{D}^{1/2} S_{n,0}^{-1} \tilde{D}^{1/2})| \leq M \cdot \| \tilde{D}^{1/2} S_{n,0}^{-1} \tilde{D}^{1/2} \| \leq C M (1 - \tilde{\rho}_{n})^{-1}. \) Similarly, we obtain

\[
|\text{tr}(\partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k_2')})| \leq C M Q_{n} \tilde{\rho}_{n}^{Q_{n}} (1 - \tilde{\rho}_{n})^{-8}.
\]

Since \( \tilde{\rho}_{n}^{Q_{n}} \) converges to zero very fast if \( \tilde{\rho}_{n} < 1 \) and \( r_{n} \leq 1 \), together with (A2) and (3.3), the summation for of the terms with \( k_1' < k - 1 \) or \( k_2' < k - 1 \) in the right-hand side of (3.35) is equal to \( o_{p}(1) \).

Then, together with Lemmas 3.5, 3.10, and 3.11, and Lemma A.1 in Ogihara (2018), we obtain

\[
E \Pi \left[ \frac{C}{n^2} \sum_{k=1}^{L} \left\{ \left( \sum_{k' < k} \Delta^{(k')} X^{c} \right)^{\top} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} \left( \sum_{k' < k} \Delta^{(k')} X^{c} \right) \right\}^2 \right]
\]

\[
\leq \frac{C}{n^2} \sum_{k=1}^{L} \left\{ |\text{tr}(\partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k_1)})| + |\text{tr}(\partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k)} \partial_{\sigma} S_{n,0}^{-1} S_{n,0}^{(k_2)})| \right\} + o_{p}(1)
\]

\[
= O_{p} \left( \frac{L_{n}}{n^2} \max_{k} \tilde{M}_{k} + 1 \right)^{2} + o_{p}(1) \xrightarrow{p} 0
\]

(3.36)

as \( n \to \infty \). Then, (3.33), (3.34), and (3.36) yield (3.30).
Next, we show (3.29). Let $I_{i,j}^k = I_i \cap I_j \cap J^k$. Then, (3.27) yields

$$\sum_{k=1}^{L_n} E_k[X_k^2] = \frac{1}{n} \sum_{k=1}^{L_n} \sum_{i_1, j_1} \sum_{i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \times \int_{I_{i,j}^k} [\Sigma_{t,0}]_{\psi(i_1), \psi(i_2)} E_k[\Delta_{j_1,t} X^c \Delta_{j_2,t} X^c] dt$$

$$= \frac{1}{n} \sum_{k=1}^{L_n} \sum_{i_1, j_1} \sum_{i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \times \int_{I_{i,j}^k} [\Sigma_{t,0}]_{\psi(i_1), \psi(i_2)} \int_{I_j \cap [0,t]} \int_{I_j \cap [0,t]} [\Sigma_{s,0}]_{\psi(j_1), \psi(j_2)} ds dt.$$  (3.37)

We can decompose

$$\int_{I_{i,j}^k} [\Sigma_{t,0}]_{\psi(i_1), \psi(i_2)} \int_{I_j \cap [0,t]} \int_{I_j \cap [0,t]} [\Sigma_{s,0}]_{\psi(j_1), \psi(j_2)} ds dt$$

$$= \int_0^{T_n} F_{i_1,j_2}^k(t) \int_0^t F_{i_1,j_2}^k(s) ds dt + \sum_{k' < k} \mathcal{F}_{i_1,j_2} \mathcal{F}_{i_1,j_2}^{k'},$$

where $F_{i,j}^k(t) = [\Sigma_{t,0}]_{\psi(i), \psi(j)} I_{i,j}^k(t)$, and $\mathcal{F}_{i,j} = \int_0^{T_n} F_{i,j}^k(t) dt$. Moreover, switching the roles of $i_1, i_2$ and $j_1, j_2$, we obtain

$$\sum_{i_1, j_1} \sum_{i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \int_0^{T_n} F_{i_1,j_2}^k(t) \int_0^t F_{i_1,j_2}^k(s) ds dt$$

$$= \sum_{i_1, j_1} \sum_{i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \times \frac{1}{2} \left\{ \int_0^{T_n} F_{i_1,j_2}^k(t) \int_0^t F_{i_1,j_2}^k(s) ds dt + \int_0^{T_n} F_{i_1,j_2}^k(t) \int_0^t F_{j_1,j_2}^k(s) ds dt \right\}$$

$$= \frac{1}{2} \sum_{i_1, j_1} \sum_{i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \int_0^{T_n} F_{i_1,j_2}^k(t) \int_0^t F_{j_1,j_2}^k(s) ds dt$$

$$+ \int_0^{T_n} F_{i_1,j_2}^k(s) \int_0^{T_n} F_{j_1,j_2}^k(t) dt ds \right\}$$

$$= \frac{1}{2} \sum_{i_1, j_1} \sum_{i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \mathcal{F}_{i_1,j_2} \mathcal{F}_{i_1,j_2}^{k} \mathcal{F}_{j_1,j_2}^{k}.$$
Therefore, we have

\[
\sum_{k=1}^{L_n} E_k[\chi_k^2] = \frac{1}{2n} \sum_{k=1}^{L_n} \sum_{i_1, j_1, i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \left( \mathcal{F}_{i_1, i_2} \mathcal{F}_{j_1, j_2} + 2 \sum_{k' < k} \mathcal{F}_{i_1, i_2} \mathcal{F}_{j_1, j_2} \right)
\]

\[
= \frac{1}{2n} \sum_{k, k' = 1}^{L_n} \sum_{i_1, j_1, i_2, j_2} [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \mathcal{F}_{i_1, i_2} \mathcal{F}_{j_1, j_2}
\]

\[
= \frac{1}{2n} \sum_{i_1, j_1, i_2, j_2} \left( [\partial_\sigma S_{n,0}^{-1}]_{i_1, j_1} [\partial_\sigma S_{n,0}^{-1}]_{i_2, j_2} \int_{I_1 \cap I_2} [\Sigma_{t,0}]_{\psi(i_1), \psi(i_2)} dt 
\right.
\]

\[
\times \int_{I_1 \cap I_2} [\Sigma_{t,0}]_{\psi(j_1), \psi(j_2)} ds
\]

\[
= \frac{1}{2n} \text{tr}((\partial_\sigma S_{n,0}^{-1} S_{n,0})^2).
\]

(3.38)

\[\partial_\sigma S_{n,0}^{-1} S_{n,0}\) corresponds to \(\hat{D}(t)\) in the proof (p. 2993) of Proposition 10 of Ogihara and Yoshida (2014). Then, by a similar step to the proof of Proposition 10 in Ogihara and Yoshida (2014), we have (3.29).

\[\square\]

**Proposition 3.15** Assume (A1)–(A4) and (A6). Then, \(\Gamma_1\) is positive definite and

\[\sqrt{n}(\hat{\sigma}_n - \sigma_0) \xrightarrow{d} N(0, \Gamma_1^{-1})\]

as \(n \to \infty\).

**Proof** Proposition 3.9, (A6), and Remark 4 in Ogihara and Yoshida (2014) yield

\[\mathcal{V}_1(\sigma) \leq -c|\sigma - \sigma_0|^2\]

(3.39)

for some positive constant \(c\). Moreover, \(\mathcal{V}_1(\sigma_0) = 0\) by \(B_{1,t,0} = 1\), and \(\partial_\sigma \mathcal{V}_1(\sigma_0) = 0\) by

\[\partial_\sigma y_{1,t}(\sigma_0) = -\partial_\rho \mathcal{A}(\rho_{t,0}) \partial_\sigma \rho_{t,0} - \frac{1}{2} \mathcal{A}(\rho_{t,0}) \sum_{l=1}^{2} \partial_\sigma B_{l,t,0}
\]

\[+ \partial_\rho \mathcal{A}(\rho_{t,0}) \partial_\sigma \rho_{t,0} - \mathcal{A}(\rho_{t,0}) \frac{\partial_\sigma \rho_{t,0}}{\rho_{t,0}}
\]

\[+ (\partial_\sigma B_{1,t,0} + \partial_\sigma B_{2,t,0}) \mathcal{A}(\rho_{t,0}) + \sum_{l=1}^{2} a_0^l (-\partial_\sigma B_{l,t,0} + \partial_\sigma B_{l,t,0})
\]

\[+ \frac{\mathcal{A}(\rho_{t,0})}{\rho_{t,0}} \partial_\sigma \rho_{t,0}
\]

\[= 0.
\]
Then, Taylor’s formula yields
\[ Y_1(\sigma) = (\sigma - \sigma_0)^\top \partial_\sigma^2 Y_1(\sigma_0)(\sigma - \sigma_0) + o(|\sigma - \sigma_0|^2). \]

Therefore, considering \( \sigma \) sufficiently close to \( \sigma_0 \), \( \Gamma_1 = -\partial_\sigma^2 Y_1(\sigma_0) \) should be positive definite by (3.39).

By Taylor’s formula and the equation \( \partial_\sigma H_1^1(\hat{\sigma}_n) = 0 \), we have
\[-\partial_\sigma H_1^1(\sigma_0) = \partial_\sigma H_1^1(\hat{\sigma}_n) - \partial_\sigma H_1^1(\sigma_0) = \int_0^1 \partial_\sigma^2 H_1^1(\sigma_t) dt (\hat{\sigma}_n - \sigma_0) = \partial_\sigma^2 H_1^1(\sigma_0)(\hat{\sigma}_n - \sigma_0) + (\hat{\sigma}_n - \sigma_0)^\top \int_0^1 (1 - t) \partial_\sigma^3 H_1^1(\sigma_t) dt (\hat{\sigma}_n - \sigma_0), \]

where \( \sigma_t = t\hat{\sigma}_n + (1 - t)\sigma_0 \).

Therefore, we obtain
\[ \sqrt{n}(\hat{\sigma}_n - \sigma_0) = \left\{ -\frac{1}{n} \partial_\sigma^2 H_1^1(\sigma_0) - \frac{1}{n} \int_0^1 (1 - t) \partial_\sigma^3 H_1^1(\sigma_t) dt (\hat{\sigma}_n - \sigma_0) \right\}^{-1} \cdot \frac{1}{\sqrt{n}} \partial_\sigma H_1^1(\sigma_0). \]

(3.40)

Since Proposition 3.8 yields
\[ -\frac{1}{n} \partial_\sigma^2 H_1^1(\sigma_0) \overset{p}{\to} -\partial_\sigma^2 Y_1(\sigma_0) = \Gamma_1, \]

and
\[ \left\{ \sup_{\sigma} \left| \frac{1}{n} \partial_\sigma^3 H_1^1(\sigma) \right| \right\}_{n \in \mathbb{N}} \]

is \( P \)-tight, together with Proposition 3.14, we conclude
\[ \sqrt{n}(\hat{\sigma}_n - \sigma_0) \overset{d}{\to} N(0, \Gamma_1^{-1}). \]

(3.41)

\[ \Box \]

### 3.4 Consistency of \( \hat{\theta}_p \)

Let
\[ Y_2(\theta) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \sum_{p=0}^\infty \left\{ -\frac{1}{2} \sum_{l=1}^2 f_p^{1+l} \rho_{1,0}^{2p} \phi_{l,t}^2 + f_p^{12} \rho_{1,0}^{2p+1} \phi_{1,t} \phi_{2,t} \right\} \, dt, \]

which exists under (A1), (A3), and (A5).
Proposition 3.16 Assume (A1)–(A6). Then

$$\sup_{\theta \in \Theta_2} \left| (nh_n)^{-1} \delta^k_\theta (H^2_n(\theta) - H^2_n(\theta_0)) - \delta^k_\theta \mathcal{Y}_2(\theta) \right| \xrightarrow{P} 0 \quad (3.42)$$

as $n \to \infty$ for $k \in \{0, 1, 2, 3\}$.

Proof We first show that

$$\hat{X}(\theta)^\top S_n^{-1}(\hat{\sigma}_n)\hat{X}(\theta)$$

$$= \Delta X^\top S_n^{-1}(\hat{\sigma}_n)\Delta X - 2\Delta V(\theta)^\top S_{n,0}^{-1}\Delta X^c - \Delta V(\theta)^\top S_{n,0}^{-1}(2\Delta V(\theta_0)$$

$$- \Delta V(\theta)) + \sqrt{nh_n} \hat{e}_n(\theta), \quad (3.43)$$

where $(\hat{e}_n(\theta))_{n=1}^\infty$ denotes a general sequence of random variables, such that

$$\sup_{\theta} |\hat{e}_n(\theta)| \xrightarrow{P} 0 \text{ as } n \to \infty.$$

Lemma 3.5 and (3.10) yield

$$E_\Pi \left[ (\Delta V(\theta)^\top \delta^k_\sigma S_{n,0}^{-1}\Delta X^c)^2 \right]$$

$$= \sum_{i_1,j_1,i_2,j_2} \sum_{\{\delta^k_\sigma S_{n,0}\}} \sum_{\{\delta^k_\sigma S_{n,0}\}} \Delta_{i_1} V(\theta) \Delta_{i_2} V(\theta) \left\{ E_\Pi [\Delta_{j_1} X^c \Delta_{j_2} X^c] \right\}$$

$$\leq C |D|^{1/2} \|\Delta V(\theta)\|^2 \|D^{1/2} \delta^k_\sigma S_{n,0}^{-1} D^{1/2}\|^2 \|D^{-1/2} S_{n,0} D^{-1/2}\|$$

$$\leq Cnh_n (1 - \hat{\rho}_n)^{-2k-2} \quad (3.44)$$

on $\{\hat{\rho}_n < 1\}$.

Since (3.2), Lemma 3.2 and Taylor’s formula yield

$$E_\Pi [\|D^{-1/2} \Delta X\|^2] = \sum_{i=1}^M \frac{E_\Pi [\|\Delta_i X\|^2]}{|I_i|} \leq C(M + nh_n) = O_p(n).$$

$$\hat{X}(\theta)^\top S_n^{-1}(\hat{\sigma}_n)\hat{X}(\theta) - \Delta X^\top S_n^{-1}(\hat{\sigma}_n)\Delta X = -\Delta V(\theta)^\top S_n^{-1}(\hat{\sigma}_n)(2\Delta X - \Delta V(\theta)),$$

and

$$S_n^{-1}(\hat{\sigma}_n) = S_{n,0}^{-1}(\hat{\sigma}_n - \sigma_0) \partial_{\sigma} S_{n,0}^{-1}(\hat{\sigma}_n - \sigma_0)^\top \int_0^1 (1-u) \partial_{\sigma} S_{n,0}^{-1}(u \hat{\sigma}_n + (1-u)\sigma_0) du (\hat{\sigma}_n - \sigma_0) \quad (3.46)$$

(3.4), (3.10), (3.41), (3.45), and Lemma 3.5 simply
\[ \begin{align*}
\sup_{\theta} | & \bar{X}(\theta)^T S_n^{-1}(\hat{\sigma}_n) \bar{X}(\theta) - \Delta X^T S_n^{-1}(\hat{\sigma}_n) \Delta X \\
& + \Delta V(\theta)^T \left\{ S_{n,0}^{-1} + (\hat{\sigma}_n - \sigma_0) \partial_\sigma S_{n,0}^{-1} \right\} (2\Delta X - \Delta V(\theta)) | \\
= & \sup_{\theta} \left| \Delta V(\theta)^T \int_0^1 (1 - u) \sum_{i,j} \partial_{\sigma_i} \partial_{\sigma_j} S_n^{-1}(u\hat{\sigma}_n) \right| \\
& + (1 - u)\sigma_0)[\hat{\sigma}_n - \sigma_0], [\hat{\sigma}_n - \sigma_0] J du (2\Delta X - \Delta V(\theta)) \\
\leq & \sup_{\theta} |D^{-1/2} \Delta V(\theta)| \cdot \sup_{\theta} |D^{-1/2} (2\Delta X - \Delta V(\theta))| \cdot |\hat{\sigma}_n - \sigma_0|^2 \\
& \times \sum_{ij} \left\| \int_0^1 (1 - u) D^{1/2} \partial_{\sigma_i} \partial_{\sigma_j} S_n^{-1}(u\hat{\sigma}_n + (1 - u)\sigma_0) D^{1/2} du \right\| \\
= & O_p(\sqrt{n h_n} \cdot \sqrt{n} \cdot (n^{-1/2})^2 \cdot 1) = o_p(\sqrt{n h_n}). \quad (3.47)
\end{align*} \]

Thanks to (3.4), (3.10), (3.41), and Lemma 3.5, we have

\[ \begin{align*}
\sup_{\theta} | & \Delta V(\theta)^T \left\{ (\hat{\sigma}_n - \sigma_0) \partial_\sigma S_{n,0}^{-1} \right\} (2\Delta X - \Delta V(\theta)) | \\
= & \sup_{\theta} |\Delta V(\theta)^T \left\{ (\hat{\sigma}_n - \sigma_0) \partial_\sigma S_{n,0}^{-1} \right\} (2\Delta X + 2\Delta V(\theta) - \Delta V(\theta)) | \\
\leq & \sup_{\theta} |2\Delta V(\theta)^T \left\{ (\hat{\sigma}_n - \sigma_0) \partial_\sigma S_{n,0}^{-1} \right\} \Delta X^c | \\
& + C \sup_{\theta} |D^{-1/2} \Delta V(\theta)|^2 \left\| D^{1/2} \partial_\sigma S_n^{-1} D^{1/2} \right\| |\hat{\sigma}_n - \sigma_0| \\
\leq & |\hat{\sigma}_n - \sigma_0| \sup_{\theta} |2\Delta V(\theta)^T \partial_\sigma S_n^{-1} \Delta X^c | + O_p(n h_n) \cdot O_p(n^{-1/2}). \quad (3.48)
\end{align*} \]

For \( k \in (0, 1) \) and \( q \geq 1 \), the Burkholder–Davis–Gundy inequality, Lemma 3.5 and a similar estimate to (3.10) yield

\[ \sup_{\theta} E_{\Pi} \left[ |\partial_{\theta}^k \Delta V(\theta)^T \partial_\sigma S_n^{-1} \Delta X^c|^q \right]^{1/q} \]

\[ \leq C_q \sup_{\theta} \sum_{l=1}^2 E_{\Pi} \left[ \left| \sum_i [\partial_\sigma S_n^{-1} \partial_{\theta}^k \Delta V(\theta)]_{i+(l-1)M_1} \Delta_i^l X^c \right|^q \right]^{1/q} \]

\[ \leq C_q \sup_{\theta} \sum_{l=1}^2 \left( \sum_i \left[ \partial_\sigma S_n^{-1} \partial_{\theta}^k \Delta V(\theta) \right]_{i+(l-1)M_1} \left| I_i \right|^l \right)^{1/2} \]

\[ = C_q \sup_{\theta} \left( \partial_{\theta}^k \Delta V(\theta)^T \partial_\sigma S_n^{-1} D \partial_\sigma S_n^{-1} \partial_{\theta}^k \Delta V(\theta) \right)^{1/2} \]

\[ \leq C_q \sqrt{n h_n} (1 - \tilde{\rho}_n)^{-2}. \quad (3.49) \]
Together with (3.4), (3.41), (3.48), and Sobolev’s inequality, we have
\[
\sup_{\theta} |\Delta V(\theta)^{\top} \left\{ (\hat{\sigma}_n - \sigma_0)\partial_\theta S_{n,0}^{-1} \right\} (2\Delta X - \Delta V(\theta))| = o_p(\sqrt{n h_n}). \tag{3.50}
\]

Then, (3.47) and (3.50) yield (3.43).

Applying (3.43) to \(\theta\) and \(\theta = \theta_0\), we have
\[
H^2_n(\theta) - H^2_n(\theta_0)
= \Delta(V(\theta) - V(\theta_0))^{\top} S_{n,0}^{-1} \Delta X + \frac{1}{2} \Delta V(\theta)^{\top} S_{n,0}^{-1} \Delta V(\theta)
- \frac{1}{2} \Delta V(\theta_0)^{\top} S_{n,0}^{-1} \Delta V(\theta_0) + \sqrt{n h_n} \epsilon_n(\theta)
= \Delta(V(\theta) - V(\theta_0))^{\top} S_{n,0}^{-1} \Delta X - \frac{1}{2} \Delta(V(\theta) - V(\theta_0))^{\top} S_{n,0}^{-1} \Delta(V(\theta) - V(\theta_0)) + \sqrt{n h_n} \epsilon_n(\theta), \tag{3.51}
\]

and hence, by similar estimates to (3.49), we have
\[
\sup_{\theta} \left| H^2_n(\theta) - H^2_n(\theta_0) + \frac{1}{2} \Delta(V(\theta) - V(\theta_0))^{\top} S_{n,0}^{-1} \Delta(V(\theta) - V(\theta_0)) \right| = O_p(\sqrt{n h_n}). \tag{3.52}
\]

Then, (3.5), (3.17), and a similar argument to (3.16) yield
\[
\Delta(V(\theta) - V(\theta_0))^{\top} S_{n,0}^{-1} \Delta(V(\theta) - V(\theta_0))
= \Delta(V(\theta) - V(\theta_0))^{\top} \mathcal{D}^{-1/2}(\sigma_0) \sum_{p=0}^{\infty} \left( (\hat{G}^T \hat{G})^p \hat{G} - (\hat{G}^T \hat{G})^p \right) \mathcal{D}^{-1/2}(\sigma_0) \Delta(V(\theta) - V(\theta_0))
= \sum_{p=0}^{\infty} \sum_{k=1}^{q_n} \rho_{k,0}^2 \left\{ \sum_{l=1}^{2} (\hat{\phi}_{1,1} - \hat{\phi}_{1,0})^2 \hat{A}_{k,1}^{\top} \hat{A}_{k,1} - 2 \hat{\phi}_{1,0} \hat{\phi}_{1,1} \hat{A}_{k,1}^{\top} \hat{A}_{k,1}^l G_k \hat{\phi}_{1,0} G_k \hat{\phi}_{1,0} + n h_n \epsilon_n \right\},
\]

where \(\hat{A}_{k,1} = \hat{A}_{k,1}(k)\). Together with (A3), (A5), (3.52), and a similar argument to (3.22), we obtain
\[
\sup_{\theta} \left| (n h_n)^{-1} (H^2_n(\theta) - H^2_n(\theta_0)) - \gamma_2(\theta) \right| \overset{P}{\to} 0 \tag{3.53}
\]
as \(n \to \infty\). Similar estimates for \((n h_n)^{-1} \partial_\theta^k (H^2_n(\theta) - H^2_n(\theta_0)) (k \in \{0, 1, 2, 3, 4\})\) yield the conclusion. \(\square\)

**Proposition 3.17** Assume (A1)–(A6). Then, \(\hat{\theta}_n \overset{P}{\to} \theta_0\) as \(n \to \infty\).

**Proof** By Lemma 3.5, we have
\[
\mathcal{D}^{1/2} S_{n,0}^{-1/2} \mathcal{D}^{1/2} \geq \left\| \mathcal{D}^{-1/2} S_{n,0} \mathcal{D}^{-1/2} \right\|^{-1} \mathcal{E}_M \geq C \mathcal{E}_M. \tag{3.54}
\]
Therefore, together with (3.15) and (3.16), we obtain

\[
- \frac{1}{2} \Delta(V(\theta) - V(\theta_0))^\top S_{n,0}^{-1} \Delta(V(\theta) - V(\theta_0)) \\
\leq - C \Delta(V(\theta) - V(\theta_0))^\top D^{-1} \Delta(V(\theta) - V(\theta_0)) \\
= C \sum_i |I_i|^{-1} \left( \int_{I_i} (\mu_{\psi(i)}(\theta) - \mu_{\psi(i)}(\theta_0)) \, dt \right)^2 \\
= - C \sum_{k=1}^{q_n} \sum_i (\mu_{\psi(i)}(\theta) - \mu_{\psi(i)}(\theta_0))^2 |I_i^k \cap J^k| + nh_ne_n \\
= - C \int_0^{T_n} |\mu_t(\theta) - \mu_t(\theta_0)|^2 \, dt + nh_ne_n. 
\]

(3.55)

Hence, we have

\[
\mathcal{Y}_2(\theta) \leq - C \limsup_{T \to \infty} \left( \frac{1}{T} \int_0^T |\mu_t(\theta) - \mu_t(\theta_0)|^2 \, dt \right). 
\]

(3.56)

Assumption (A6) yields that for any \( \theta \in \Theta \)

\[
\mathcal{Y}_2(\theta) \leq 0, \quad \text{and} \quad \mathcal{Y}_2(\theta) = 0 \quad \text{if and only if} \quad \theta = \theta_0; 
\]

(3.57)

(3.42), (3.57) together with a similar estimate to (3.24), we have the conclusion. \qed

3.5 Asymptotic normality of \( \hat{\theta}_n \)

**Proof of Theorem 2.3** By the definition of \( H_n^2(\theta) \), we obtain

\[
\partial_\theta H_n^2(\theta_0) = \partial_\theta \Delta V(\theta_0)^\top S_{n,0}^{-1}(\hat{\sigma}_n) \hat{X}(\theta_0) = \partial_\theta \Delta V(\theta_0)^\top S_{n,0}^{-1}(\hat{\sigma}_n) \Delta X^c. 
\]

By a similar argument to the derivation of (3.43), we can replace \( S_{n,0}^{-1}(\hat{\sigma}_n) \) in the right-hand side of the above equation by \( S_{n,0}^{-1} \) with approximation error equal to \( o_p(\sqrt{nh_n}) \). Then, we have

\[
\partial_\theta H_n^2(\theta_0) = \partial_\theta \Delta V(\theta_0)^\top S_{n,0}^{-1} \Delta X^c + o_p(\sqrt{nh_n}). 
\]

Let

\[
\hat{X}_k = \frac{1}{\sqrt{nh_n}} \partial_\theta \Delta V(\theta_0) S_{n,0}^{-1} \Delta^{(k)} X^c
\]

Springer
for $1 \leq k \leq L_n$. Then, we have

$$(nh_n)^{-1/2} \partial_\theta H_n^2(\theta_0) = \sum_{k=1}^{L_n} \hat{X}_k + o_p(1). \quad (3.58)$$

Lemma 3.5 and a similar argument to (3.10) yield

$$\sum_{k=1}^{L_n} E_k[\hat{X}_k^2] = \frac{3}{2} \frac{n^2 h_n^2}{\sqrt{p}} \sum_{k=1}^{L_n} \left\{ \partial_\theta \Delta V(\theta_0)^\top S_{n,0}^{-1} S_{n,0}^{-1} \partial_\theta \Delta V(\theta_0) \right\}^2$$

$$\leq \frac{C}{n^2 h_n^2} |\mathcal{D}|^{-1/2} |\partial_\theta V(\theta_0)|^2 \|\mathcal{D}^{1/2} S_{n,0}^{-1} \mathcal{D}^{1/2}\| \sum_{k=1}^{L_n} \|\mathcal{D}^{-1/2} S_{n,0}^{(k)} \mathcal{D}^{-1/2}\|$$

$$\leq \frac{CL_n}{nh_n} (1 - \tilde{p}_n)^2 \xrightarrow{p} 0.$$

Moreover, (3.5), (A5), and a similar argument to the proof of Proposition 3.8 yield

$$\sum_{k=1}^{L_n} E_k[\hat{X}_k^2] = \frac{1}{nh_n} \sum_{k=1}^{L_n} \sum_{i_1, j_1, i_2, j_2} S_{n,0}^{-1} S_{n,0}^{-1} S_{n,0}^{-1} \Delta_{i_1} \partial_\theta V(\theta_0) \Delta_{i_2} \partial_\theta V(\theta_0) [S_{n,0}^{(k)}]_{i_1, j_1, i_2, j_2}$$

$$= \frac{1}{nh_n} \Delta \partial_\theta V(\theta_0)^\top S_{n,0}^{-1} S_{n,0}^{-1} \Delta \partial_\theta V(\theta_0)$$

$$= \frac{1}{nh_n} \sum_{p=0}^{\infty} \sum_{k=1}^{q_n} \rho^2_{k,0} \left\{ \sum_{l=1}^{2} \partial_\theta \phi_{l, s_{k-1}}(\theta_0) \mathcal{J}_l^\top \mathcal{A}_{k,p} \mathcal{J}_l \right\} + e_n$$

$$\xrightarrow{p} \Gamma_2.$$

Therefore, (3.58) and the martingale central limit theorem (Corollary 3.1 and the remark after that in Hall & Heyde, 1980) yield

$$(nh_n)^{-1/2} \partial_\theta H_n^2(\theta_0) = \sum_{k=1}^{L_n} \hat{X}_k + o_p(1) \xrightarrow{d} N(0, \Gamma_2). \quad (3.59)$$

By (3.56) and (A6), there exists a positive constant $c$, such that $\mathcal{Y}_2(\theta) \leq -c|\theta - \theta_0|^2$.

Then, $\Gamma_2 = -\partial_\theta^2 \mathcal{Y}_2(\theta_0)$ is positive definite, since $\mathcal{Y}_2(\theta_0) = 0$ and $\partial_\theta \mathcal{Y}_2(\theta_0) = 0$.

Therefore, a similar estimate to Sect. 3.3, $P$-tightness of $\{(nh_n)^{-1} \sup_\theta |\partial_\theta^2 H_n^2(\theta)|\}$, and the equation $- (nh_n)^{-1} \partial_\theta^2 H_n^2(\theta_0) \xrightarrow{p} \Gamma_2$ yield

$$\sqrt{T_n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, \Gamma_2^{-1}).$$
(3.40) and a similar equation for \( \sqrt{nh_n(\hat{\theta}_n - \theta_0)} \) yield

\[
\left( \sqrt{n}(\hat{\sigma}_n - \sigma_0), \sqrt{nh_n}(\hat{\theta}_n - \theta_0) \right) = \left( n^{-1/2} \Gamma_1^{-1} \partial_{\sigma} H_n^1(\sigma_0), T_n^{-1/2} \Gamma_2^{-1} \partial_{\theta} H_n^2(\theta_0) \right) + o_p(1)
\]

\[
= \sum_{k=1}^{L_n} (\Gamma_1^{-1} \hat{\lambda}_k, \Gamma_2^{-1} \hat{\lambda}_k) + o_p(1).
\]

(3.60)

Then, since \( \sum_{k=1}^{L_n} E_k[\lambda_k \hat{\lambda}_k] = 0 \), we obtain

\[
(\sqrt{n}(\hat{\sigma}_n - \sigma_0), \sqrt{nh_n}(\hat{\theta}_n - \theta_0)) \xrightarrow{d} N(0, \Gamma_n^{-1}).
\]

3.6 Proofs of the results in Sects. 2.3 and 2.4

Proof of Theorem 2.5 Let \( \sigma_{tu} = \sigma_0 + t\epsilon_n u \) for \( u \in \mathbb{R}^d \) and \( t \in [0, 1] \), and let

\[
H_n(\sigma, \theta) = -\frac{1}{2} \bar{X}(\theta)^\top S_n^{-1}(\sigma) \bar{X}(\theta) - \frac{1}{2} \log \det S_n(\sigma).
\]

Then, we have

\[
H_n(\sigma_u, \theta_u) = u^\top \epsilon_n \int_0^1 \partial_{\alpha} H_n(\sigma_{tu}, \theta_{tu}) dt
\]

\[
= u^\top \epsilon_n \partial_{\alpha} H_n(\sigma_0, \theta_0) + \frac{1}{2} u^\top \epsilon_n \partial^2_{\alpha} H_n(\sigma_0, \theta_0) \epsilon_n u
\]

\[
+ \sum_{i, j, k} \int_0^1 \frac{(1 - s)^2}{2} \partial_{\alpha_i} \partial_{\alpha_j} \partial_{\alpha_k} H_n(\sigma_{su}, \theta_{su}) ds [\epsilon_n u_i][\epsilon_n u_j][\epsilon_n u_k].
\]

By similar arguments to Propositions 3.8 and 3.14, and Sects. 3.4 and 3.5, we obtain

\[
\sum_{i, j, k} \int_0^1 \frac{(1 - s)^2}{2} \partial_{\alpha_i} \partial_{\alpha_j} \partial_{\alpha_k} H_n(\sigma_{su}, \theta_{su}) ds [\epsilon_n u_i][\epsilon_n u_j][\epsilon_n u_k] \xrightarrow{p} 0,
\]

\[
\Delta_n := \epsilon_n \partial_{\alpha} H_n(\sigma_0, \theta_0) \xrightarrow{d} N(0, \mathcal{E}_d),
\]

\[-\epsilon_n \partial^2_{\alpha} H_n(\sigma_0, \theta_0) \epsilon_n \xrightarrow{p} \mathcal{E}_d.
\]

Therefore, we have the desired conclusion. \( \square \)

Remark 3.18 We can show that \((\hat{\sigma}_n, \hat{\theta}_n)\) is a regular estimator by the proof of Theorem 2.5, (3.60), and Theorem 2 in Jeganathan (1982).

Outline of the proof of Proposition 2.7

The proof is similar to the proof of Proposition 6 in Ogihara and Yoshida (2014). \( P \)-tightness of \( \{h_n M_{l, q_n + 1}\}_{n=1}^{\infty} \) immediately follows from (B1-1). Fix \( 1 \leq j \leq q_n \).
Then, using the mixing property (2.4) for $\mathcal{N}_t^{m,l}$, we obtain the following result; there exists $\eta > 0$, such that for any $q \geq 4$, there exists $C_q > 0$ that does not depend on $j$, such that

$$E\left[|h_n \text{tr}(\mathcal{E}_{(j)}^1 (GG^\top)^p) - E[h_n \text{tr}(\mathcal{E}_{(j)}^1 (GG^\top)^p)]|^q\right] \leq C_q (p + 1)^{q-1} h_n^{q \eta}.$$  

(The above inequality corresponds to (31) in Ogihara and Yoshida (2014). This is obtained by defining $b_n = h_n^{-1}$, $t_k = s_j - 1 + k[h_n^{-1}]^{-1}(s_j - s_{j-1})$ for $0 \leq k \leq [h_n^{-1}]$, and $X_k' = \text{tr}(\mathcal{E}_{(j,k)}^1 (GG^\top)^p)\mathbb{1}_{A^p_{k,b_n'}} - E[\text{tr}(\mathcal{E}_{(j,k)}^1 (GG^\top)^p)\mathbb{1}_{A^p_{k,b_n'}}]$ in the proof of Proposition 6 in Ogihara and Yoshida (2014), where $\mathcal{E}_{(j,k)}^i$ is an $M_i \times M_i$ matrix satisfying $[\mathcal{E}_{(j,k)}^i]_{ii} = 1$ if $i = i'$ and $sup_{t_i} \in (t_{k-1}, t_k)$, and otherwise, $[\mathcal{E}_{(j,k)}^i]_{ii} = 0$.)

Therefore, by setting sufficiently large $q$, so that $nh_n^{1+q\eta} \to 0$, we have

$$E\left[\max_{1 \leq j \leq q_n} |h_n \text{tr}(\mathcal{E}_{(j)}^1 (GG^\top)^p) - E[h_n \text{tr}(\mathcal{E}_{(j)}^1 (GG^\top)^p)]|^q\right] \leq E\left[\sum_{j=1}^{q_n} |h_n \text{tr}(\mathcal{E}_{(j)}^1 (GG^\top)^p) - E[h_n \text{tr}(\mathcal{E}_{(j)}^1 (GG^\top)^p)]|^q\right] = O(q_n \cdot h_n^{q \eta} \to 0).$$

Here, we used that for any partition $(s_k)_{k=0}^\infty \in \mathcal{G}$, we have $q_n \leq nh_n/\epsilon + 1$ with $\epsilon = \inf_{k \geq 1} |s_k - s_{k-1}| > 0$, which implies $q_n = O(nh_n)$. Together with the assumptions, we obtain the conclusion.

**Outline of the proof of Proposition 2.8**

Similarly to the previous proposition, using the idea of Proposition 6 in Ogihara and Yoshida (2014) and the mixing property (2.4) for $\mathcal{N}_t^{m,l}$, we have that there exists $\eta > 0$, such that for any $q \geq 4$, there exists $C_q > 0$, such that

$$E\left[|\mathcal{J}_1^\top \mathcal{E}_{(j)}^1 (GG^\top)^p \mathcal{J}_1 - E[\mathcal{J}_1^\top \mathcal{E}_{(j)}^1 (GG^\top)^p \mathcal{J}_1]|^q\right] \leq C_q (p + 1)^{q-1} h_n^{q \eta}$$

for $1 \leq j \leq q_n$. (We define $b_n$ and $t_k$ the same as the previous proposition, and define

$$X_k' = [h_n]^{-1}\mathcal{J}_1^\top \mathcal{E}_{(j,k)}^1 (GG^\top)^p \mathcal{J}_1 \mathbb{1}_{A^p_{k,b_n'}} - E[[h_n]^{-1}\mathcal{J}_1^\top \mathcal{E}_{(j,k)}^1 (GG^\top)^p \mathcal{J}_1 \mathbb{1}_{A^p_{k,b_n'}}].$$

Together with the assumptions and similar estimates for $\mathcal{J}_1 \mathcal{E}_{(j)}^1 (GG^\top)^p \mathcal{J}_2$ and $\mathcal{J}_2 \mathcal{E}_{(j)}^2 (GG^\top)^p \mathcal{J}_2$, we obtain the conclusion.

**Outline of the proof of Proposition 2.9**

We can show the results by a similar approach to the proof of Proposition 9 in Ogihara and Yoshida (2014). Roughly speaking, under (B2-$q$), the probability $P(\mathcal{N}_t^{m,l} - \hat{\mathcal{G}} \to$
\( \mathcal{N}_{t^m,l} = 0 \) is small enough to estimate the denominator of

\[
\sum_{i,j} \frac{|I_i^1 \cap I_j^2|^2}{|I_i^1||I_j^2|}
\]

for sufficiently large \( N \). Then, we obtain estimates for the numerator using an inequality \( x_1^2 + \cdots + x_n^2 \geq R^2/n \) when \( x_1 + \cdots + x_n = R \).

**Proof of Lemma 2.10** We only show

\[
\max_{1 \leq k \leq q_n} |h_n E[\text{tr}(E_{(k)} (GG^T)^p)] - a_p^1 (s_k - s_{k-1})| \to 0.
\]

The other results are similarly obtained.

(2.4) is satisfied, because \( \alpha_k^p \leq c_1 e^{-c_2 k} \) for some positive constants \( c_1 \) and \( c_2 \).

Let \( \bar{t}_i^l \) be \( i \)-th jump time of \( \bar{\mathcal{N}}^d \). Then, we have \( \bar{S}^n,l_i = h_n \bar{t}_i^l \). Let \( \bar{G} \) be a matrix with infinity size defined by

\[
[\bar{G}]_{ij} = \frac{|(\bar{t}_i^{l-1}, \bar{t}_i^l] \cap (\bar{t}_j^{l-1}, \bar{t}_j^l]|}{\sqrt{\bar{t}_i^l - \bar{t}_i^{l-1}} \sqrt{\bar{t}_j^l - \bar{t}_j^{l-1}}}
\]

for \( i, j \geq 1 \).

For \( k \in \mathbb{N} \), let

\[
\mathfrak{S}^{p, n}_{k} = \sum_{i ; \bar{t}_i^{-1} \in [k-1,k)} [(\bar{G}G^T)^p]_{ii}, \quad \mathfrak{S}^{p, n}_{k} = \sum_{i ; \bar{S}^{n,l}_i \in [(k-1)h_n, kh_n)} [(G^T)^p]_{ii}.
\]

The following idea is based on Section 7.5 of Ogihara and Yoshida (2014). Roughly speaking, if there are sufficient observations around the interval \([k-1,k)\), we can apply mixing property of \( \mathcal{N}^m,l_{t} \) to \( \mathfrak{S}^{p, n}_{k} \). On the following sets \( A_{k, r}^p \) and \( \tilde{A}_{k, r}^p \), we have sufficient observations of \( \mathcal{N}^m,l_{t} \) and \( \mathcal{N}^d \). Let \( U_{t+r} = U_{t+r} - U_{t+r(j-1)} \) for a stochastic process \( (U_t)_{t \geq 0} \), and let

\[
A_{k, r}^p = \bigcap_{l=1,2} \left\{ \bigcap_{1 \leq j \leq 2p+1} \left\{ \tilde{\mathcal{A}}_{j, r}^h_{l,k} \mathcal{N}^n,l \geq 0 \right\} \bigcap_{-2p \leq j \leq 0} \left\{ \tilde{\mathcal{A}}_{j, r}^h_{l,k-1} \mathcal{N}^n,l \geq 0 \right\} \right\},
\]

\[
\tilde{A}_{k, r}^p = \bigcap_{l=1,2} \left\{ \bigcap_{1 \leq j \leq 2p+1} \left\{ \tilde{\mathcal{A}}_{j, r}^h_{l,k} \mathcal{N}^d \geq 0 \right\} \bigcap_{-2p \leq j \leq 0} \left\{ \tilde{\mathcal{A}}_{j, r}^h_{l,k-1} \mathcal{N}^d \geq 0 \right\} \right\}.
\]

(3.61)

Then, we obtain

\[
E[\mathfrak{S}^{p, n}_{k} \tilde{A}_{k, r}^p] = E[\mathfrak{S}^{p, n}_{k} \tilde{A}_{k, r}^p] \quad \text{if } k \wedge k' \geq rp + 1,
\]
\[ E[\mathfrak{G}^{n,p}_{k} 1_{A_{k,r}^p}] = E[\mathfrak{G}^{n,p}_{k'} 1_{A_{k,r}^p}] \quad \text{if } rp + 1 \leq k, k' \leq n - rp. \]

We also have \( P((\hat{A}_{k,r}^p)^c) \leq C(p + 1)r^{-q} \) by (B2-\(q\)). For any \( \epsilon > 0 \), there exists \( r > 0 \), such that

\[
P((\hat{A}_{k,r}^p)^c) < \epsilon/2. \tag{3.62}
\]

Therefore, \( \{E[\mathfrak{G}^{n,p}_k]\}_k \) is a Cauchy sequence, and hence, the limit \( a^1_p = \lim_{k \to \infty} E[\mathfrak{G}^{p}_k] \) exists for \( p \in \mathbb{N} \). Moreover, we see existence of

\[
a^l_0 = \lim_{k \to \infty} E[\hat{N}^l_k - \bar{N}^l_{k-1}] = E[\hat{N}^l_1 - \bar{N}^l_0]
\]

for \( l \in \{1, 2\} \).

Furthermore, for any \( \epsilon > 0 \), there exists \( r > 0 \), such that

\[
P((\hat{A}_{k,r}^p)^c) < \epsilon \quad \text{and} \quad |E[\mathfrak{G}^{p}_k] - a^1_p| < \epsilon \tag{3.63}
\]

for \( k \geq [rp] \). We also have

\[
E[\mathfrak{G}^{p}_k 1_{A_{k,r}^p}] = [\mathfrak{G}^{n,p}_k 1_{A_{k,r}^p}] \tag{3.64}
\]

for \( rp + 1 \leq k \leq n - rp \), since

\[
\sup I^l_i \in (s_{j-1}, s_j) \iff \bar{I}^l_i \in (h^{-1}s_{j-1}, h^{-1}s_j].
\]

Let \( r_j = [h^{-1}s_j] \). Then, since \( |\mathfrak{G}^{n,p}_k| \leq \sum_{i:s_{i-1,j}\in((k-1)h_n,kh_n]} 1 \leq E[\hat{N}^1_1], (3.63), (3.64) \), and the Cauchy–Schwarz inequality yield

\[
|h_n(s_j - s_{j-1})^{-1} E[\text{tr}(\mathcal{E}(j)(GG^\top)^p)] - a^1_p| \\
\leq h_n(s_j - s_{j-1})^{-1} E\left[ \sum_{k=r_j-1+1}^{r_j} \mathfrak{G}^{n,p}_k \right] - a^1_p | + 2h_n(s_j - s_{j-1})^{-1} E[\hat{N}^1_1]
\]

\[
\leq \left| \frac{1}{r_j - r_{j-1}} \right| E\left[ \sum_{k=r_j-1+1}^{r_j} \mathfrak{G}^{n,p}_k \right] - a^1_p | + Ch_n(s_j - s_{j-1})^{-1}
\]

\[
\leq \frac{1}{r_j - r_{j-1}} \sum_{k=r_j-1+1}^{r_j} \left| E[\mathfrak{G}^{n,p}_k 1_{A_{k,h}^p}] + E[\mathfrak{G}^{n,p}_k 1_{(A_{k,h}^p)^c}] - a^1_p \right| + Ch_n(s_j - s_{j-1})^{-1}
\]

\[
\leq \frac{1}{r_j - r_{j-1}} \sum_{k=r_j-1+1}^{r_j} \left( |E[\mathfrak{G}^{p}_k] - a^1_p| + 2E[\hat{N}^1_1]^2]^{1/2} \sqrt{\epsilon} \right) + Ch_n(s_j - s_{j-1})^{-1}
\]

\[
\leq \epsilon + 2E[\hat{N}^1_1]^2]^{1/2} \sqrt{\epsilon} + Ch_n(s_j - s_{j-1})^{-1}
\]

for \( 1 < j < q_n \). To get the corresponding inequality for \( j = 1, q_n \), we replace the summation range of \( k \) in the above inequality with the range from \( r_{j-1} + [rp] + 2 \)
to $r_j$ when $j = 1$, and with the range from $r_{j-1}$ to $r_j - \lfloor rp \rfloor - 1$ when $j = q_n$. Boundedness of $\{E[h_n M_{t,q_n+1}]\}_{n \in \mathbb{N}}$ is shown using the same techniques. Then, we have the conclusion.}

\section*{Acknowledgements}
On behalf of all authors, the corresponding author states that there is no conflict of interest. The author is grateful to the journal editor and two anonymous referees for careful reading and constructive comments, which led to improvements of this paper.

\section*{Funding}
Open access funding provided by The University of Tokyo. This work was supported by Japan Society for the Promotion of Science (Grant no. 21H00997).

\section*{Data Availability}
Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

\section*{Open Access}
This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

\section*{References}
Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. \textit{Journal of Econometrics}, 162(2), 149–169.

Bibinger, M., Hautsch, N., Malec, P., & Reiss, M. (2014). Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency. \textit{Annals of Statistics}, 42(4), 80–114.

Christensen, K., Kinnebrock, S., & Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. \textit{Journal of Econometrics}, 159(1), 116–133.

Florens-Zmirou, D. (1989). Approximate discrete-time schemes for statistics of diffusion processes. \textit{Statistics}, 20(4), 547–557.

Hall, P., & Heyde C. C. (1980). \textit{Martingale limit theory and its application}. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers]. Probability and Mathematical Statistics.

Hayashi, T., & Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. \textit{Bernoulli}, 11(2), 359–379.

Hayashi, T., & Yoshida, N. (2008). Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. \textit{Annals of the Institute of Statistical Mathematics}, 60(2), 367–406.

Hayashi, T., & Yoshida, N. (2011). Nonsynchronous covariation process and limit theorems. \textit{Stochastic Processes and Their Applications}, 121(10), 2416–2454.

Ibragimov, I. A., & Has’minskiĭ, R. Z. (1981). \textit{Statistical estimation. Applications of Mathematics} (Vol. 16). Springer. Asymptotic theory, Translated from the Russian by Samuel Kotz.

Jeganathan, P. (1982). On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. \textit{Sankhyā Series A}, 44(2), 173–212.

Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. \textit{Scandinavian Journal of Statistics}, 24(2), 211–229.

Malliavin, P., & Mancino, M. E. (2002). Fourier series method for measurement of multivariate volatilities. \textit{Finance and Stochastics}, 6(1), 49–61.

Malliavin, P., & Mancino, M. E. (2009). A Fourier transform method for nonparametric estimation of multivariate volatility. \textit{Annals of Statistics}, 37(4), 1983–2010.

Ogihara, T. (2015). Local asymptotic mixed normality property for nonsynchronously observed diffusion processes. \textit{Bernoulli}, 21(4), 2024–2072.
Ogihara, T. (2018). Parametric inference for nonsynchronously observed diffusion processes in the presence of market microstructure noise. *Bernoulli, 24*(4B), 3318–3383.

Ogihara, T., & Uehara, Y. (2022). Local asymptotic normality for ergodic jump-diffusion processes via transition density approximation. *Bernoulli (to appear).*

Ogihara, T., & Yoshida, N. (2014). Quasi-likelihood analysis for nonsynchronously observed diffusion processes. *Stochastic Processes and Their Applications, 124*(9), 2954–3008.

Uchida, M., & Yoshida, N. (2012). Adaptive estimation of an ergodic diffusion process based on sampled data. *Stochastic Processes and Their Applications, 122*(8), 2885–2924.

Yoshida, N. (1992). Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis, 41*(2), 220–242.

Yoshida, N. (2011). Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. *Annals of the Institute of Statistical Mathematics, 63*(3), 431–479.