Quasi-Likelihood Analysis for Diffusion Processes and Diffusion Processes with Jumps

その他のタイトル	拡散過程及びジャンプ型拡散過程に対する疑似尤度解析
学位授与年月日	2013-07-15
URL	http://doi.org/10.15083/00007218

博士論文

論文題目：

Quasi-Likelihood Analysis for Diffusion Processes and Diffusion Processes with Jumps

(拡散過程及びジャンプ型拡散過程に対する疑似尤度解析)

氏名：　荻原　哲平

Preface

In this paper, we study asymptotic behaviors of quasi-maximum likelihood estimators and Bayes type estimators for parameterized diffusion processes and diffusion processes with jumps.

Let a d-dimensional stochastic process $X = \{X_t\}_{0\leq t<\infty}$ satisfy a stochastic differential equation :

$$dX_t = \mu(X_t,\theta)dt + b(X_t,\sigma)dW_t, \quad t\in[0,\infty), \tag{1}$$

where $\{W_t\}_{0\leq t\leq T}$ is a multi-dimensional standard Wiener process, μ is a vector-valued function and b is a matrix-valued function. Then we consider the problem of estimating true values θ_* and σ_* of parameter θ and σ, respectively. We may consider two sampling schemes of X : continuous time observations and discrete time observations. From a practical viewpoint, discrete time observations are realistic because it is difficult to observe a continuous path of X. The asymptotic theory of statistical estimation have been well developed for discretely observed diffusion processes.

Suppose that X is ergodic and discrete samples $\{X_{kh_n}\}_{k=0}^n$ are observed for some $h_n > 0$. Then it is well-known that the quasi-maximum likelihood estimators $\hat\theta_n$ and $\hat\sigma_n$ generated by the Euler-Maruyama type quasi-log-likelihood function have consistency :

$$(\hat\sigma_n,\hat\theta_n)\to^p(\sigma_*,\theta_*), \tag{2}$$

asymptotic normality :

$$(\sqrt{n}(\hat\sigma_n-\sigma_*),\sqrt{nh_n}(\hat\theta_n-\theta_*))\to^d\zeta, \tag{3}$$

and moment convergence :

$$E[\mathsf{f}(\sqrt{n}(\hat\sigma_n-\sigma_*),\sqrt{nh_n}(\hat\theta_n-\theta_*))]\to\mathbb{E}[\mathsf{f}(\zeta)] \tag{4}$$

as $n\to\infty, h_n\to 0, nh_n\to\infty$ for any continuous function f of at most polynomial growth, where ζ is a multi-dimensional normal random variable. Similar results as (2)-(4) hold when one replaces $\hat\theta_n$ and $\hat\sigma_n$ with Bayes type estimators $\tilde\theta_n$ and $\tilde\sigma_n$, respectively.

Though the above results are obtained under the assumption $T = nh_n\to\infty$, there are also asymptotic results when the end time T is fixed. We denote the observations of X by $\{X_{kT/n}\}_{k=0}^n$, then the quasi-maximum likelihood estimator $\hat\sigma_n$ based on the quasi-likelihood function has consistency :

$$\hat\sigma_n\to^p\sigma_*, \tag{5}$$

asymptotic mixed normality :

$$\sqrt{n}(\hat\sigma_n-\sigma_*)\to^{s\text{-}\mathcal{L}}\Gamma_0^{-1/2}\mathcal{N}_0, \tag{6}$$

moment convergence :

$$E[\mathsf{f}(\sqrt{n}(\hat\sigma_n-\sigma_*))]\to\mathbb{E}[\mathsf{f}(\Gamma_0^{-1/2}\mathcal{N}_0)] \tag{7}$$

for any continuous function f of at most polynomial growth, where $\to^{s\text{-}\mathcal{L}}$ denotes stable convergence, Γ_0 is a symmetric positive definite random matrix, $\mathcal{N}_0$ is a multi-dimensional standard normal random variable independent of Γ_0. Similar results also hold true for Bayes type estimators $\tilde\sigma_n$.

In this paper, we extend these results for regularly sampled diffusion processes to nonsynchronously observed diffusion processes and regularly sampled diffusion processes with jumps.

In Chapter 1, we consider a two-dimensional stochastic process $Y = \{(Y_t^1, Y_t^2)\}_{0\leq t\leq T}$ satisfying a stochastic integral equation :

$$Y_t = Y_0 + \int_0^t \mu_s ds + \int_0^t b(X_s,\sigma)dW_s, \quad t \in [0,T], \tag{8}$$

where $\{W_t\}_{0\leq t\leq T}$ is a two-dimensional standard Wiener process, $\{\mu_t\}_{0\leq t\leq T}$ and $X = \{X_t\}_{0\leq t\leq T}$ are two-dimensional and n_2-dimensional stochastic processes, respectively, and b is a 2×2 matrix-valued function. As an example, Y becomes a diffusion process if $\mu_t = \mu(t, Y_t)$ and $X_t = (t, Y_t)$.

We investigate asymptotic behaviors of a quasi-maximum likelihood estimator and a Bayes type estimator when the end time $T > 0$ is fixed, observation times $\{S^i\}_i, \{T^j\}_j$ and $\{\mathcal{T}_k^j\}_j$ of $\{Y_t^1\}, \{Y_t^2\}$ and $\{X_t^k\}$, respectively, are random and nonsynchronous and $\max_{i,j,k}(|S^i - S^{i-1}| \vee |T^j - T^{j-1}| \vee |\mathcal{T}_k^j - \mathcal{T}_k^{j-1}|) \to^p 0$.

The problem of nonsynchronous observations appears when one estimates covariation of two security log-prices by high-frequency financial data. Security log-prices are observed when transactions occur. Then observations are inevitably nonsynchronous because transactions for different securities occur at different time points. Linear interpolation or '*previous tick*' is a natural method to solve this problem. However, it is known that these simple methods of '*synchronization*' cause serious bias of the estimator. Recently, there are large numbers of studies about various covariance estimators solving this problem.

However, previous works about estimation problems of nonsynchronous observations are mainly focused on nonparametric methods. In this paper, we consider Y given by (8) and prove similar results to (5)-(7), that is, a quasi-maximum likelihood estimator and a Bayes type estimator have consistency, asymptotic mixed normality and convergence of moments.

Theory of the random field of likelihood ratio enables us to reduce the asymptotic behavior of estimators to more tractable asymptotic properties of the quasi-likelihood function H_n. To specify the asymptotic behavior of H_n, we assume that certain functions of observation times converge in probability. Asymptotic variance of estimation error described by these limit function. See $[A3]$ in Section 1.3 for details. Thus the effects of nonsynchronous observations appear in asymptotic variance of estimators.

In the two previous settings of synchronously observed diffusion processes, quasi-maximum likelihood estimators and Bayes type estimators are asymptotically efficient, that is, they attains the minimal asymptotic variance. Asymptotic efficiency of the quasi-maximum likelihood estimators and the Bayes type estimators is unknown for nonsynchronously observed diffusion processes. However, Example 1.1 shows that the quasi-maximum likelihood estimator has lower estimation error than that of a nonparametric estimator of quadratic covariation $\langle Y^1, Y^2\rangle_T$. Thus performance of our estimator is preferable as a parametric estimator.

In Chapter 2, we consider a d-dimensional stochastic process $X = \{X_t\}_{0\leq t<\infty}$ satisfying

$$dX_t = a(X_{t-},\theta)dt + b(X_{t-},\sigma)dW_t + \int_E c(X_{t-},z,\theta)p(dt,dz), \quad t \in [0,\infty), \tag{9}$$

where $\{W_t\}_{0\leq t\leq\infty}$ is a d-dimensional standard Wiener process, p is a Poisson random measure, a, b and c are Borel functions and $E = \mathbb{R}^d \setminus \{0\}$, $\sigma \in \Pi$ and $\theta \in \Theta$ are parameters and Π and Θ are bounded open sets in Euclidean spaces. We assume that X is ergodic and discrete samples $\{X_{kh_n}\}_{k=0}^n$ are observed for some $h_n > 0$. We study asymptotic behaviors of a quasi-maximum likelihood estimator and a Bayes type estimator as $n \to \infty, h_n \to 0, nh_n \to \infty$.

Shimizu and Yoshida [46] proposed a method using a threshold detecting whether jumps occur in an interval $((k-1)h_n, kh_n]$ by the value $|X_{kh_n} - X_{(k-1)h_n}|$. They constructed a quasi-log-likelihood function by using this threshold and proved consistency and asymptotic normality of the quasi-maximum likelihood estimators. In Chapter 2, we use an improved quasi-log-likelihood function $H_n(\sigma,\theta)$ and prove consistency, asymptotic normality and moment convergence of the quasi-maximum likelihood estimators and the Bayes type estimators. In particular, results of the asymptotic behavior of Bayes type estimators for diffusion processes with jumps are new to the best of my knowledge. Moment convergence of estimators plays important roles in developments of theory of information criteria and asymptotic expansion. Shimizu and Yoshida [46] assumed that the Lévy measure f satisfies $|f(z)| \leq C|z|^\gamma$ for some $C > 0$ and $\gamma > 3$ near the origin. We weaken the assumption and our assumption is satisfied by many distributions whose density function f is bounded near the origin.

To discuss the asymptotic properties of estimators, we prove polynomial type large deviation inequalities :

For any $L > 0$ there exists $C_L > 0$ such that

$$\begin{aligned} P\Big[\sup_{(u_1,\theta)\in V_n^1(r)\times\Theta} \exp\Big\{H_n\Big(\sigma^* + \frac{u_1}{\sqrt{n}}, \theta\Big) - H_n(\sigma^*, \theta)\Big\} \geq e^{-\frac{r}{2}}\Big] &\leq \frac{C_L}{r^L}, \\ P\Big[\sup_{u_2\in V_n^2(r)} \exp\Big\{H_n\Big(\hat{\sigma}_n, \theta^* + \frac{u_2}{\sqrt{nh_n}}) - H_n(\hat{\sigma}_n, \theta^*)\Big\} \geq e^{-\frac{r}{2}}\Big] &\leq \frac{C_L}{r^L} \end{aligned} \tag{10}$$

for any $r > 0$, where σ^* and θ^* are true values of σ and θ, respectively, $\hat{\sigma}_n$ is the quasi-maximum likelihood estimator for the parameter σ, and

$$V_n^1(r) = \{u_1; \sigma^* + n^{-1/2}u_1 \in \Pi, |u_1| \geq r\}, \quad V_n^2(r) = \{u_2; \theta^* + (nh_n)^{-1/2}u_2 \in \Theta, |u_2| \geq r\}.$$

The inequalities (10) yield

$$P[|\sqrt{n}(\hat{\sigma}_n - \sigma_*)| \geq r] \leq \frac{C_L}{r^L}, \quad P[|\sqrt{nh_n}(\hat{\theta}_n - \theta_*)| \geq r] \leq \frac{C_L}{r^L}$$

for quasi-maximum likelihood estimators $\hat{\sigma}_n$ and $\hat{\theta}_n$, and so give moment estimates of estimators. These estimates play an important role in the proof of consistency and asymptotic normality of Bayes type estimators and moment convergence of estimators.

Results in Chapters 1 and 2 are published in [35] and [34], respectively.

I am very grateful to Professor Nakahiro Yoshida who was my supervisor for his valuable instructions and advices. I learned many things about statistical and mathematical theories from him. I would also like to thank Professor Masayuki Uchida for giving useful comments and encouraging me.

Contents

1 QLA for nonsynchronously observed diffusion processes **9**
1.1 Introduction 9
1.2 Construction of a quasi-likelihood function 11
1.3 The limit of H_n and observation times 15
1.3.1 Convergence conditions of functions of the observation times 16
1.3.2 The limit of H_n 18
1.4 Separation of the limit of H_n 20
1.5 Asymptotic properties of the quasi-maximum-likelihood estimator and the Bayes type estimator 22
1.6 Sufficient conditions for the conditions about the observation times 23
1.7 Proofs 26
1.7.1 Proof of Proposition 1.2 26
1.7.2 Proof of Proposition 1.3 28
1.7.3 Proof of Lemmas 1.7 and 1.8 40
1.7.4 Proof of Proposition 1.5 and Theorem 1.2 42
1.7.5 Proof of Propositions 1.6 - 1.9 48

2 QLA for diffusion processes with jumps **53**
2.1 Introduction 53
2.2 Quasi-likelihood and asymptotic property of the estimators 54
2.3 Proof of the main results 61
2.3.1 Polynomial type large deviation inequality and the quasi-likelihood analysis 61
2.3.2 Proof of Theorems 2.1 and 2.2 64

Chapter 1

Quasi-Likelihood Analysis for Nonsynchronously Observed Diffusion Processes

1.1 Introduction

Given a probability space $(\Omega, \mathcal{F}, P)$ with a right-continuous filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$, we consider a stochastic regression model specified by the following equation :

$$Y_t = Y_0 + \int_0^t \mu_s ds + \int_0^t b(X_s, \sigma) dW_s, \quad t \in [0,T], \tag{1.1}$$

where $Y = \{Y_t\}_{0\leq t\leq T} = \{(Y_t^1, Y_t^2)\}_{0\leq t\leq T}$ is a two-dimensional $\mathbf{F}$-adapted process, $\{W_t\}_{0\leq t\leq T}$ is a two-dimensional standard $\mathbf{F}$-Wiener process, $b = (b^{ij})_{1\leq i,j\leq 2} : \mathbb{R}^{n_2} \times \Lambda \to \mathbb{R}^2 \otimes \mathbb{R}^2$ is a Borel function, $\mu = \{\mu_t\}$ and $X = \{X_t\}$ are $\mathbf{F}$-progressively measurable processes with values in $\mathbb{R}^2$ and $\mathbb{R}^{n_2}$, respectively, $\sigma \in \Lambda$, and Λ is a bounded open subset of $\mathbb{R}^{n_1}$. For example, if $\mu_t = \mu(t, Y_t)$ and $X_t = (t, Y_t)$, then $\{Y_t\}$ is a time-inhomogeneous diffusion process.

Our purpose is to estimate the true value σ_* of parameter $\sigma \in \Lambda$ by nonsynchronous observations $\{Y_{S^i}^1\}_i$, $\{Y_{T^j}^2\}_j$ and $\{X_{\mathcal{T}_k^j}^k\}_{j,k}$, where $\{S^i\}_i, \{T^j\}_j$ and $\{\mathcal{T}_k^j\}_{j,k}$ are observation times of Y^1, Y^2 and X, respectively. In our setting, μ is completely unobservable and unknown.

The problem of nonsynchronous observations appears in the analysis of high-frequency financial data. Recently, as availability of intraday security prices gets increase, the analysis of high-frequency data becomes more significant. In particular, the realized volatility has been studied actively as an estimator of security returns' volatility.

In the study of portfolio risk management of financial assets, the quadratic covariation of two security log-prices is also a significant risk measure. Therefore estimation of quadratic covariation with high-frequency data has also been studied by many authors. One problem of estimation is nonsynchronous trading. The observation times of two different security prices do not necessarily coincide with each other.

If $Y^1 = \{Y_t^1\}_{0\leq t\leq T}$ and $Y^2 = \{Y_t^2\}_{0\leq t\leq T}$ are synchronously observed at some stopping times $\{S^i\}$, then the realized covariance between Y^1 and Y^2 converges to $\langle Y^1, Y^2\rangle_T$ in probability as $\max_i |S^i - S^{i-1}| \to^p 0$. When observation times of Y^1 and Y^2 are nonsynchronous, to calculate the realized covariance, we need to synchronize the data by some method. However, the realized covariance has serious bias if we use a simple synchronizing method such as *previous-tick* interpolation or linear interpolation. Epps [13] first indicated this phenomenon by U.S. stock data analysis, and this phenomenon is called the *Epps Effect*.

To solve this problem, Malliavin and Mancino [29] proposed a Fourier analytic method, and Hayashi and Yoshida [17] proposed an estimator based on overlapping of observation intervals. In sequent papers [18, 19], Hayashi and Yoshida studied the asymptotic distribution of estimation error of their estimator and proved asymptotic mixed normality. There also exist some works about estimation of the quadratic covariation with nonsynchronous data contaminated by market microstructure noise. We refer the reader to Barndorff-Nielsen et al. [4] for a kernel based method, Christensen, Kinnebrock and Podolskij [9] for a pre-averaged Hayashi-Yoshida estimator, Aït-Sahalia, Fan and Xiu [3] for a method with the maximum likelihood estimator of a model with deterministic diffusion coefficients, and Bibinger [5, 6] for a multiscale estimator.

With respect to the problem of nonsynchronous observations, nonparametric approaches have been studied mainly. In this work, we use a quasi-likelihood function, that approximates the likelihood function in diffusion cases and construct a quasi-maximum likelihood estimator and a Bayes type estimator for a parametric stochastic regression model with nonsynchronous observations. The asymptotic behavior of estimators will be investigated when the end time T is fixed and $\max_{i,j,j',k}|S^i-S^{i-1}|\vee|T^j-T^{j-1}|\vee|\mathcal{T}_k^{j'}-\mathcal{T}_k^{j'-1}|\to 0$ in probability. Hence our method can be applied not only to estimating the quadratic covariation but also to identifying nonlinear structure of the process Y.

There exist many studies about asymptotic theory of parametric estimation for stochastic differential equations with high-frequency data. Among many studies in a long history, we refer the reader to Prakasa Rao [36,37], Yoshida [52–54], Kessler [24] under ergodicity, Shimizu and Yoshida [46], Ogihara and Yoshida [34] for jump diffusion processes, Masuda [33] for Ornstein-Uhlenbeck processes driven by heavy-tailed symmetric Lévy processes, Sørensen and Uchida [47], Uchida [48, 49] for perturbed diffusions, Dohnal [11], Genon-Catalot and Jacod [14, 15], Gobet [16], Uchida and Yoshida [50, 51] for the fixed interval case.

One of the most useful approaches to study asymptotic behaviors of quasi-maximum likelihood estimators and Bayes type estimators is the theory of random field of likelihood ratios initiated by Ibragimov and Has'minskii [20–22]. Their theory enabled to reduce the problem of asymptotic behaviors of estimators to more tractable properties of the random field of likelihood ratios. In [22], they applied their theory to independent observations and white Gaussian noise models. Kutoyants [25–28] developed Ibragimov-Has'minskii's theory for diffusion processes and point processes. Yoshida [54] investigated polynomial type large deviation inequalities to apply Ibragimov-Has'minskii's theory and discussed consistency and asymptotic normality of quasi-maximum likelihood estimators and Bayes type estimators for ergodic diffusion processes. This scheme was also applied to jump diffusion processes in Ogihara and Yoshida [34], Ornstein-Ohlenbeck processes driven by heavy-tailed symmetric Lévy processes in Masuda [33], and diffusion processes in the fixed interval in Uchida and Yoshida [50, 51].

In this work, we construct a quasi-log-likelihood function for the stochastic regression model (1.1) with nonsynchronous observations. Then we will show consistency, asymptotic mixed normality and the convergence of moments of the quasi-maximum likelihood estimator and the Bayes type estimator with the aid of polynomial type large deviation inequalities. The advantage of our approach is to obtain asymptotic mixed normality, exact representation of asymptotic variance and convergence of moments of the estimators. The convergence of moments of the estimators is important, e.g., when we investigate the asymptotic expansion and the theory of information criteria. Moreover, our method does not require any synchronization methods.

When the sampling scheme is synchronous and equi-spaced : $S^i=T^i=iT/n$, Gobet [16] showed local asymptotic mixed normality of the likelihood function of observations and obtained the asymptotic minimax bound for the variance of estimators. In the case of nonsynchronous observations, we expect that local asymptotic mixed normality of the likelihood function holds and our estimators attain the asymptotic minimax bound since our quasi-likelihood function seems to be asymptotically equivalent with the *true* likelihood function and our quasi-likelihood ratio has a limit distribution of LAMN type. However, these problems are not proved in this paper and left as future work.

This chapter is organized as follows. In Section 1.2, we construct a quasi-log-likelihood function H_n and discuss its non-degeneracy. Section 1.3 gives the asymptotic behavior of H_n. Section 1.3.1 deals with two equivalent conditions of the asymptotic behavior of observation times $\{S^i\}$ and $\{T^j\}$ to control the asymptotic behavior of H_n. In Section 1.3.2, we specify the limit of H_n and estimate the rate of convergence. Section 1.4 studies the degree of separation of the limit of H_n, which is necessary to prove asymptotic properties of the quasi-maximum likelihood estimator and the Bayes type estimator. We also introduce sufficient conditions for the condition of separation. In Section 1.5, our main results about asymptotic properties for estimators are stated. Section 1.6 introduces easily tractable sufficient conditions for assumptions about the observation times in the main theorems. Proofs are collected in Section 1.7.

1.2 Construction of a quasi-likelihood function

In this section, we define a quasi-log-likelihood function H_n to construct a quasi-maximum likelihood estimator and a Bayes type estimator.

First, we define some notations. For a real number a, $[a]$ denote the maximum integer which is not greater than a. For a matrix A, $A^\star$ denotes transpose of A and $\| A \|$ represents the norm of A as a linear map. We often regard a p-dimensional vector v as a $p \times 1$ matrix. $\mathcal{E}_p$ denotes unit matrix of size p. We set $\sup \emptyset = -\infty$, $\inf \emptyset = +\infty$ and $2\mathbb{N} = \{2k; k \in \mathbb{N}\}$. For $M \in \mathbb{N}$ and $K \subset \mathbb{R}^M$, $\bar{K}$ denotes the closure of K. For a set $K \subset \Omega$, K^c denotes the complementary set of K. For an interval $K \subset [0,T]$ and a stochastic process $\{\mathsf{Z}_t\}_{0 \le t \le T}$, we denote $L(K) = \inf K, R(K) = \sup K$, $\mathsf{Z}(K) = \mathsf{Z}_{R(K)} - \mathsf{Z}_{L(K)}$, $K_t = K \cap [0,t)$ and $|K| = R(K) - L(K)$. Let $b^i(x,\sigma) = (b^{i1}(x,\sigma), b^{i2}(x,\sigma))^\star$ $(i = 1,2)$. For a vector $\kappa = (\kappa_1, \cdots, \kappa_M)$, we denote $\partial_\kappa^k = (\frac{\partial^k}{\partial \kappa_{i_1} \cdots \partial \kappa_{i_k}})_{i_1, \cdots i_k = 1}^M$. We denote $|x|^2 = \sum_{i_1,\cdots,i_M} |x_{i_1,\cdots,i_M}|^2$ for $x = \{x_{i_1,\cdots,i_M}\}_{i_1,\cdots,i_M}$.

Let Λ satisfy Sobolev's inequality, that is, for any $p > n_1$, there exists $C > 0$ such that

$$\sup_{x \in \Lambda} |u(x)| \le C \sum_{k=0,1} \| \partial_x^k u(x) \|_p$$

for $u \in C^1(\Lambda)$. It is the case if Λ has Lipschitz boundary. See Adams [1], Adams and Fournier [2] for more details.

We recall the definition of stable convergence. Given an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ of $(\Omega, \mathcal{F}, P)$, let $\{\mathsf{Z}_n\}_{n \in \mathbb{N}}$ and Z be random variables on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ with values in a metric space E. Then we say that Z_n stably converges in law to Z, and write $\mathsf{Z}_n \to^{s\text{-}\mathcal{L}} \mathsf{Z}$, if $E[\mathsf{Y} f(\mathsf{Z}_n)] \to E[\mathsf{Y} f(\mathsf{Z})]$ as $n \to \infty$ for any bounded continuous function $f : E \to \mathbb{R}$ and any bounded variable Y on $(\Omega, \mathcal{F})$. See Jacod [23] for more details.

For $1 \le k \le n_2$, let observation times $\{S^i\}_i$, $\{T^j\}_j$ and $\{\mathcal{T}_k^j\}_{j,k}$ be strictly increasing with respect to i or j almost surely and satisfy $S^0 = T^0 = \mathcal{T}_k^0 = 0$, $S^i = \inf\{t \ge 0; N_t^1 \ge i\} \wedge T$, $T^j = \inf\{t \ge 0; N_t^2 \ge j\} \wedge T$, and $\mathcal{T}_k^j = \inf\{t \ge 0; N_t^{k+2} \ge j\} \wedge T$ for $i, j \ge 1$, where $\{N_t^{k'}\}_t$ are simple point processes, that is, $\{N_t^{k'}\}$ is a càdlàg $\mathbb{Z}_+$-valued stochastic process whose jumps are equal to 1 and $N_0^{k'} = 0$ $(1 \le k' \le n_2 + 2)$. These observations and point processes depend on a positive integer $n \in \mathbb{N}$. Let $\Pi = \Pi_n = ((S^i)_i, (T^j)_j, (\mathcal{T}_k^j)_{j,k})$, $l_n = N_{T-}^1 + 1, m_n = N_{T-}^2 + 1, m_n^k = N_{T-}^{k+2} + 1$ for $1 \le k \le n_2$, then $l_n, m_n, \{m_n^k\}_{k=1}^{n_2}$ are observation counts. We also assume $\{\Pi_n\}_{n \in \mathbb{N}}$ are independent of $\mathcal{F}_T$. Denote $I^i = [S^{i-1}, S^i)$ $(1 \le i \le l_n)$, $J^j = [T^{j-1}, T^j)$ $(1 \le j \le m_n)$,

$$r_n = \max_{i,j}(|I^i| \vee |J^j|) \vee \max_{1 \le k \le n_2} \max_{1 \le j \le m_n^k} |\mathcal{T}_k^j - \mathcal{T}_k^{j-1}|,$$

and $\mathcal{T}_k'(K) = \max\{\mathcal{T}_k^j; j \in \mathbb{Z}_+, \mathcal{T}_k^j \le L(K)\}$ for $1 \le k \le n_2$ and an interval $K \subset [0,T]$. Let $\{b_n\}_{n \in \mathbb{N}}$ be a sequence of positive numbers such that $b_n \ge 1$ $(n \in \mathbb{N})$ and $b_n \to \infty$ as $n \to \infty$. $\{b_n\}$ represents order of observation counts. Conditions for $\{b_n\}$ are given in [$A2$-q, δ], [$A3'$-q, η], [$A4$-q, δ] later.

For a function $g : \mathbb{R}^{n_2} \times \Lambda \to \mathbb{R}$, let $g_t = g(X_t, \sigma), g_{t,*} = g(X_t, \sigma_*)$, $g_{K,t} = g(\{X^k_{\mathcal{T}_k'(K) \wedge t}\}_k, \sigma)$ and $g_K = g_{K,T}$ for interval $K \subset [0,T]$. We use the symbol C for a generic positive constant which is independent of n and p, and is varying from line to line.

We assume the following conditions.

[$A1$]

1. The mapping $b : \mathbb{R}^{n_2} \times \Lambda \to \mathbb{R}^2 \otimes \mathbb{R}^2$ has the continuous derivative $\partial_x^j \partial_\sigma^i b$ and $\partial_\sigma^i b$ can be continuously extended to $\mathbb{R}^{n_2} \times \bar{\Lambda}$ for $0 \le j \le 3$ and $0 \le i \le 4$. Moreover,

$$\sup_{\sigma \in \Lambda} |\partial_x^j \partial_\sigma^i b(x,\sigma)| \le C(1 + |x|)^C$$

for $0 \le j \le 3$, $0 \le i \le 4$ and $x \in \mathbb{R}^{n_2}$.

2. There exists $\epsilon > 0$ such that $\det bb^\star(x,\sigma) \ge \epsilon$ for $(x,\sigma) \in \mathbb{R}^{n_2} \times \Lambda$.
3. $|b(x,\sigma) - b(y,\sigma)| \le C|x - y|$ for $x, y \in \mathbb{R}^{n_2}$ and $\sigma \in \Lambda$.
4. $Y_0 \in \cap_{q > 0} L^q(\Omega)$.

5. There exists $\gamma \in (0,1)$ such that

$$\sup_{0\leq t\leq T} E[|\mu_t|^q] < \infty \quad \text{and} \quad \sup_{0\leq s<t\leq T} \frac{E[|\mu_t - \mu_s|^q]}{|t-s|^{q\gamma}} < \infty$$

for any $q > 0$.

6. There exists $n_3 \in \mathbb{Z}_+$ such that X can be decomposed as

$$X_t = X_0 + \int_0^t \tilde{b}_s^1 ds + \int_0^t \tilde{b}_s^2 dW_s + \int_0^t \tilde{b}_s^3 d\hat{W}_s,$$

where

$$\tilde{b}_t^i = \tilde{b}_0^i + \int_0^t \hat{b}_s^{i1} ds + \int_0^t \hat{b}_s^{i2} dW_s + \int_0^t \hat{b}_s^{i3} d\hat{W}_s, \ (i = 2, 3)$$

$\{\tilde{b}_t^i\}_{0\leq t\leq T}$ $(1 \leq i \leq 3)$ and $\{\hat{b}_t^{ij}\}_{0\leq t\leq T}$ $(2 \leq i \leq 3, 1 \leq j \leq 3)$ are $\mathbf{F}$-progressively measurable processes, $\{\hat{W}_t\}_{0\leq t\leq T}$ is an n_3-dimensional standard $\mathbf{F}$-Wiener process independent of $\{W_t\}$ and $E[\sup_{0\leq t\leq T}(|\hat{b}_t^{ij}| \vee |\tilde{b}_t^i| \vee |X_0|)^p] < \infty$ for any i, j and $p > 0$. We ignore the terms $\tilde{b}_t^3$, $\int_0^t \tilde{b}_s^3 d\hat{W}_s$ and $\int_0^t \hat{b}_s^{i3} d\hat{W}_s$ when $n_3 = 0$.

Our setting contains the case where X or Y_0 depends on σ and main results hold in this case. However, if X or Y_0 depends on σ, our estimator $\hat{\sigma}_n$ may not be the quasi-maximum likelihood estimator since we need to consider the density of observations $\{X^k_{T^j_k}\}$ or Y_0. Nevertheless, we use the terms "quasi-maximum likelihood estimator" and "Bayes type estimator" in this case. If $X_t = (t, Y_t)$ and Y_0 does not depend on σ, we can see $\hat{\sigma}_n$ is the maximum likelihood type estimator.

Under $[A1]$ 2, there exists $\epsilon > 0$ such that

$$\det(b_t b_t^\star) = |b_t^1|^2 |b_t^2|^2 (1 - \rho_t^2) \geq \epsilon \tag{1.2}$$

for $\rho(x,\sigma) = b^1 \cdot b^2 |b^1|^{-1} |b^2|^{-1}(x,\sigma)$. Therefore

$$\bar{\rho} = \sup_{t\in[0,T],\sigma\in\Lambda} |\rho_t| < 1 \quad a.s. \tag{1.3}$$

by $[A1]$ 1.

Let us denote

$$S(\sigma) = \begin{pmatrix} \mathrm{diag}(\{|b_I^1|^2\}_I) & \left\{ b_I^1 \cdot b_J^2 \frac{|I\cap J|}{\sqrt{|I|}\sqrt{|J|}} \right\}_{IJ} \\ \left\{ b_I^1 \cdot b_J^2 \frac{|I\cap J|}{\sqrt{|I|}\sqrt{|J|}} \right\}_{JI} & \mathrm{diag}(\{|b_J^2|^2\}_J) \end{pmatrix}$$

and define a quasi-log-likelihood function $H_n = H_n(\sigma)$ of $((Y^1(I)/\sqrt{|I|})_I^\star, (Y^2(J)/\sqrt{|J|})_J^\star)$ by

$$H_n = -\frac{1}{2}\left(\left(\frac{Y^1(I)}{\sqrt{|I|}}\right)_I^\star, \left(\frac{Y^2(J)}{\sqrt{|J|}}\right)_J^\star \right) S^{-1} \left(\left(\frac{Y^1(I)}{\sqrt{|I|}}\right)_I^\star, \left(\frac{Y^2(J)}{\sqrt{|J|}}\right)_J^\star \right)^\star - \frac{1}{2}\log\det S$$

when $\det S > 0$. If $X_t \equiv t$ and $\mu \equiv 0$, S is the covariance matrix for the Euler-Maruyama type approximation $((\tilde{Y}^1(I))_I, (\tilde{Y}^2(J))_J)$ of $((Y^1(I))_I, (Y^2(J))_J)$ defined by $\tilde{Y}^1(I) = b^1(L(I)) \cdot W(I)$, $Y^2(J) = b^2(L(J)) \cdot W(J)$. Though H_n is the quasi-log-likelihood function for $\mu \equiv 0$, we can see that the effect of drift term μ in a quasi-likelihood function can be ignored asymptotically. So H_n is applicable for general cases.

Remark 1.1. *Though the quasi-likelihood function H_n is defined as functions on $(\Omega, \mathcal{F}, P)$, we often regard it as a function on the state space. We adopt the same thing to the quasi-maximum likelihood estimator and the Bayes type estimator.*

When the sampling scheme is synchronous, we have uniform non-degeneracy of S by the condition $[A1]$ 2. However, in the case of nonsynchronous observations, the problem becomes more complicated since the observation times of diffusion coefficients are not the same for Y^1 and Y^2. However, the following proposition ensures that H_n is well-defined under $[A1]$ 2.

Proposition 1.1. *Assume* $[A1]$ 2. *Then* $\det S(\sigma) > 0$ *almost surely for any* $\sigma \in \Lambda$.

Proof. Fix $\omega \in \Omega$. It is sufficient to show that S is positive definite. Let $((u_I)_I, (v_J)_J)$ be a real vector satisfying

$$((u_I)_I, (v_J)_J) S((u_I)_I, (v_J)_J)^\star = 0.$$

We assume that $((u_I)_I, (v_J)_J)$ has a non-zero element and show this leads to a contradiction.

Let $\{\tilde{W}_t\}$ be a two-dimensional standard Wiener process on some probability space, and $\{M_t\}_{0\le t\le T}$ be a stochastic process on the same probability space, satisfying

$$M_t = \sum_I \frac{u_I}{\sqrt{|I|}} b_I^1 \cdot \tilde{W}(I_t) + \sum_J \frac{v_J}{\sqrt{|J|}} b_J^2 \cdot \tilde{W}(J_t).$$

Then $\{M_t\}$ is a martingale satisfying

$$\langle M\rangle_t = \sum_I \frac{u_I^2}{|I|}|b_I^1|^2|I_t| + \sum_J \frac{v_J^2}{|J|}|b_J^2|^2|J_t| + 2\sum_{I,J} u_I v_J b_I^1 \cdot b_J^2 \frac{|(I\cap J)_t|}{\sqrt{|I||J|}}.$$

Since

$$\langle M\rangle_T = ((u_I)_I, (v_J)_J) S((u_I)_I, (v_J)_J)^\star = 0,$$

it follows that $\langle M\rangle_t = 0$ for $0 \le t \le T$.

We may assume some I satisfies $L(I) = \min\{L(I); u_I \ne 0\} \wedge \min\{L(J); v_J \ne 0\}$ without loss of generality. We fix this I below.

First, we consider the case that $L(I) < \min\{L(J); v_J \ne 0\}$. Then

$$\langle M\rangle_{L(I)+\delta} = |b_I^1|^2 u_I^2 \delta / |I| = 0$$

for sufficiently small $\delta > 0$. Therefore we have $|b_I^1| = 0$, which contradicts $[A1]$ 2.

In the case that $L(I) = L(J)$ for some J with $v_J \ne 0$, we obtain

$$\langle M\rangle_{L(I)+\delta} = |b_I^1|^2 \frac{u_I^2}{|I|}\delta + |b_J^2|^2 \frac{v_J^2}{|J|}\delta + 2\frac{u_I v_J}{\sqrt{|I||J|}} b_I^1 \cdot b_J^2 \delta = 0 \tag{1.4}$$

for sufficiently small $\delta > 0$. Since $L(I) = L(J)$, we obtain $b_J^2 = b_I^2$. Therefore

$$\left(\frac{u_I}{\sqrt{|I|}}, \frac{v_J}{\sqrt{|J|}}\right) b_I b_I^\star \left(\frac{u_I}{\sqrt{|I|}}, \frac{v_J}{\sqrt{|J|}}\right)^\star = 0$$

by (1.4). This contradicts the fact that $b_I b_I^\star$ is positive definite by $[A1]$ 2. □

Let

$$\bar{\rho}_t = \sup_{0\le s\le t} |\rho_s|, \quad \rho_{I,J,t} = \frac{b_{I,t}^1 \cdot b_{J,t}^2}{|b_{I,t}^1||b_{J,t}^2|}, \quad \tilde{\rho}_n(t) = \sup_{\sigma, I, J; I\cap J \ne \emptyset} |\rho_{I,J,t}| \vee \bar{\rho}_t,$$

To discuss asymptotic behavior of the quasi-likelihood, we need a more precise estimate for non-degeneracy of S. To this end, we will estimate $1 - \tilde{\rho}_n(t)$ from below. Assuming $[A1]$ and $r_n \to^p 0$ $(n \to \infty)$, we have $\sup_t |\bar{\rho}_t - \tilde{\rho}_n(t)| \to^p 0$ $(n \to \infty)$ by uniform continuity of b^1 and b^2 with respect to t and σ for fixed ω. Therefore $\lim_{n\to\infty} P[\sup_t \tilde{\rho}_n(t) \ge 1] = 0$ by (1.3).

We need a stronger estimate for $\tilde{\rho}_n$. For stochastic processes $\{s_n(t)\}_{0\le t\le T, n\in\mathbb{N}}$, we consider the following condition:

[S] There exists $M \in \mathbb{N}$, stochastic processes $\{\bar{s}_n(t,x)\}$ and a $\sigma(\Pi_n)$-measurable $\mathbb{R}^M$-valued random variable X such that $s_n(t) = \bar{s}_n(t;\mathsf{X})$, $\bar{s}_n(t,x)$ is continuous with respect to (t,x) a.s., $\bar{s}_n(0,x) \leq 1-|\rho_0|$, $t \mapsto \bar{s}_n(t,x)$ is non-increasing and $\{\bar{s}_n(t,x)\}_{0\leq t\leq T}$ is a $[0,1]$-valued $\mathbf{F}$-adapted process for $n \in \mathbb{N}$ and $x \in \mathbb{R}^M$.

Let

$$\tau_n = \tau(s_n) = \inf\{t \in [0,T]; \tilde{\rho}_n(t) \geq 1 - s_n(t)\} \wedge T.$$

We consider the following condition for $q > 0$ and $\xi > 0$.

[S-q,ξ] $\{s_n(t)\}_{t,n}$ satisfies [S], $P[\tau(s_n) < T] = O(b_n^{-\xi})$ and $\sup_n E[(s_n(T))^{-q}] < \infty$.

Define $\hat{S} = \hat{S}(\sigma; s_n)$ and $\hat{H}_n = \hat{H}_n(\sigma; s_n)$ similarly to S and H_n respectively, substituting b^1_{I,τ_n} for b^1_I and b^2_{J,τ_n} for b^2_J in the definition of S. Under [S-q,ξ], it is easy to see that $\sup_\sigma |H_n - \hat{H}_n| \to^p 0$ as $n \to \infty$. To investigate asymptotic properties of estimators, it is convenient to use $\hat{H}_n$.

If $\sup_t \tilde{\rho}_n(t) < c$ almost surely for some $0 < c < 1$, we can set $s_n \equiv 1-c$. However, in general, we need the following conditions to obtain $\{s_n\}$ for $q > 0$ and $\delta \in (0,1)$.

[$A2$] $r_n \to^p 0$ as $n \to \infty$.

[$A2$-q,δ] $E[r_n^q] = O(b_n^{-\delta q})$.

The following lemma gives examples of $\{s_n\}$ in general cases.

Lemma 1.1. *Let $P > 1, q' > 0$, $0 < \delta < 1$ and $s_n(t) = (1-\bar{\rho}_t)/P$. Assume* [$A1$], [$A2$-$q',\delta$]. *Then $\{s_n\}$ satisfies* [S-q,ξ] *for any $q > 0$ and $0 < \xi < \delta q'$.*

Proof. It is clear that $\{s_n\}$ satisfies [S]. Since $1/(1-|\rho_t|) \leq 2|b^1_t|^2|b^2_t|^2/\epsilon$ by (1.2), we obtain $\sup_n E[(s_n(T))^{-q}] < \infty$ for any $q > 0$ by [$A1$].

Moreover, let $\eta = (\delta - \xi/q')/9$ and $q_0 \geq \xi/\eta$, then by [$A1$] and the mean value theorem, we obtain

$$\sup_{0\leq t\leq T}\left|\tilde{\rho}_n(t) - \bar{\rho}_t\right| \leq C\sup_t(1+|X_t|)^C \times \sup_{|t-s|\leq r_n}|X_t - X_s|.$$

For $t \in (0,T)$, we have

$$\begin{aligned} 1-\tilde{\rho}_n(t) \leq s_n(t) \quad &\Rightarrow \quad \frac{1-\tilde{\rho}_n(t)}{1-\bar{\rho}_t} \leq \frac{1}{P} \Rightarrow 1 - \frac{1}{P} \leq \frac{\tilde{\rho}_n(t)-\bar{\rho}_t}{1-\bar{\rho}_t} \\ &\Rightarrow \quad 1-\bar{\rho}_t \leq b_n^{-\eta} \quad \text{or} \quad b_n^{-\eta}(1-1/P) \leq (\tilde{\rho}_n(t) - \bar{\rho}_t). \end{aligned}$$

From this relation, we see that

$$\begin{aligned} P[\tau_n < T] &= P[\text{There exists } t \in [0,T) \text{ s.t. } 1-\tilde{\rho}_n(t) \leq s_n(t)] \\ &\leq b_n^{-q_0\eta}E[1/(1-\bar{\rho}_T)^{q_0}] + P[b_n^{-\eta}(1-1/P) \leq b_n^{2\eta}r_n^{1/3}] \\ &\quad + P\left[C\sup_t(1+|X_t|)^C \vee \sup_{s\neq t}\frac{|X_t - X_s|}{|t-s|^{1/3}} \geq b_n^\eta\right]. \end{aligned}$$

Then by [$A1$], [$A2$-q',δ] and Kolmogorov criterion([39] Chapter I, Theorem (2.1)), we obtain

$$P[\tau_n < T] \quad \leq \quad E[(b_n^{3\eta}(1-1/P)^{-1}r_n^{1/3})^{3q'}] + O(b_n^{-\xi}) = O(b_n^{-\xi}).$$

□

From now on, we fix $\{s_n\}$ which satisfy [S-q,ξ] for some $q > 0$ and $\xi > 0$ unless otherwise indicated.

Next, we expand $\hat{H}_n$. We denote

$$\begin{aligned} D &= \mathrm{diag}(\{|b^1_{I,\tau_n}|\}_I, \{|b^2_{J,\tau_n}|\}_J), \quad L = \left\{\rho_{I,J,\tau_n}\frac{|I\cap J|}{\sqrt{|I||J|}}\right\}_{I,J}, \\ \tilde{L} &= \begin{pmatrix} 0 & L \\ L^\star & 0 \end{pmatrix}, \quad Z = \left(\left(\frac{Y^1(I)}{|b^1_{I,\tau_n}|\sqrt{|I|}}\right)^\star_I, \left(\frac{Y^2(J)}{|b^2_{J,\tau_n}|\sqrt{|J|}}\right)^\star_J\right)^\star. \end{aligned}$$

Since $\hat{S} = D(\mathcal{E}_{l_n+m_n} + \tilde{L})D$,

$$\hat{H}_n = -\frac{1}{2}Z^\star MZ - \log\det D + \frac{1}{2}\log\det M$$

for $M = (\mathcal{E}_{l_n+m_n} + \tilde{L})^{-1}$. Moreover, for $G = \{|I \cap J|/\sqrt{|I||J|}\}_{IJ}$, we obtain

$$\| \tilde{L} \|^2 = \| \{\rho_{I,J,\tau_n} G_{IJ}\}_{IJ} \|^2 \vee \| \{\rho_{I,J,\tau_n} G_{IJ}\}_{JI} \|^2 \leq (1 - s_n(T))^2 (\| G \|^2 \vee \| G^\star \|^2).$$

Lemma 1.2. *For any $n \in \mathbb{N}$ and $\omega \in \Omega$, all the eigenvalues of the symmetric matrices $GG^\star, G^\star G$ are in $[0,1]$. In particular, $\| G \| \vee \| G^\star \| \leq 1$.*

Proof. Fix $\omega \in \Omega$. We denote by $\{\lambda_i\}_{i=1}^{l_n}$ the eigenvalues of $GG^\star$. Obviously, $0 \leq \lambda_i$ $(1 \leq i \leq l_n)$. Let $\{\tilde{W}_t\}_t$ be a one-dimensional standard Wiener process on some probability space and

$$\Sigma_1 = \begin{pmatrix} \mathcal{E}_{l_n} & G \\ G^\star & \mathcal{E}_{m_n} \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} \mathcal{E}_{l_n} & -G \\ 0 & \mathcal{E}_{m_n} \end{pmatrix},$$

then Σ_1 is the covariance matrix of $((\tilde{W}(I)/\sqrt{|I|})_I, ((\tilde{W}(J)/\sqrt{|J|})_J)$ and

$$\Sigma_2 \Sigma_1 \Sigma_2^\star = \begin{pmatrix} \mathcal{E}_{l_n} - GG^\star & 0 \\ 0 & \mathcal{E}_{m_n} \end{pmatrix}.$$

Since Σ_1 is non-negative definite, $\mathcal{E}_{l_n} - GG^\star$ is also non-negative definite, and hence $1 - \lambda_i \geq 0$ $(1 \leq i \leq l_n)$. Therefore we conclude $0 \leq \lambda_i \leq 1$ $(1 \leq i \leq l_n)$.

In particular, we have $\| G^\star \|^2 = \| GG^\star \| = \max_i \lambda_i \leq 1$. The same conclusion can be drawn for $G^\star G$ and $\| G \|$. □

Since $\| \tilde{L} \| \leq 1 - s_n(T)$ by Lemma 1.2, $\sum_{p=0}^\infty (-1)^p \tilde{L}^p$ exists almost surely and this gives $M = \sum_{p=0}^\infty (-1)^p \tilde{L}^p$, under $[S\text{-}q, \xi]$. Moreover, we obtain

$$\max_{1 \leq k \leq l_n+m_n} |\eta_k| = \| \tilde{L} \| \leq 1 - s_n(T),$$

where $\{\eta_k\}_{k=1}^{l_n+m_n}$ are the eigenvalues of $\tilde{L}$. Hence

$$\log\det(\mathcal{E}_{l_n+m_n} + \tilde{L}) = \sum_{k=1}^{l_n+m_n} \log(1 + \eta_k) = \sum_{k=1}^{l_n+m_n} \sum_{p=1}^\infty \frac{(-1)^{p+1}\eta_k^p}{p} = \sum_{p=1}^\infty \frac{(-1)^{p+1}}{p} \mathrm{tr}(\tilde{L}^p)$$

almost surely. Therefore

$$\begin{aligned} \hat{H}_n &= -\frac{1}{2}Z^\star \left(\sum_{p=0}^\infty (-1)^p \tilde{L}^p \right) Z - \log\det D + \frac{1}{2}\sum_{p=1}^\infty \frac{(-1)^p}{p}\mathrm{tr}(\tilde{L}^p) \\ &= -\frac{1}{2}Z^\star \sum_{p=0}^\infty \begin{pmatrix} (LL^\star)^p & -(LL^\star)^p L \\ -(L^\star L)^p L^\star & (L^\star L)^p \end{pmatrix} Z - \log\det D + \frac{1}{2}\sum_{p=1}^\infty \frac{(-1)^p}{p}\mathrm{tr}(\tilde{L}^p) \qquad (1.5) \end{aligned}$$

almost surely.

1.3 The limit of H_n and observation times

In this section, we investigate the asymptotic behavior of H_n and $\hat{H}_n$ to apply Ibragimov-Has'minskii's theory. To obtain these estimates, we need some convergence conditions ($[A3]$, $[A3']$ and $[A3'\text{-}q, \eta]$ given in Section 1.3.1) for the observation times. Proposition 1.3 in Section 1.3.2 will give asymptotic properties of H_n and $\hat{H}_n$.

1.3.1 Convergence conditions of functions of the observation times

Since M is a functional of $\{\rho_{I^i,J^j,\tau_n}\}_{i,j}$, we can write

$$M = \begin{pmatrix} M^{11}(\{\rho_{I^i,J^j,\tau_n}\}) & M^{12}(\{\rho_{I^i,J^j,\tau_n}\}) \\ (M^{12}(\{\rho_{I^i,J^j,\tau_n}\}))^\star & M^{22}(\{\rho_{I^i,J^j,\tau_n}\}) \end{pmatrix}.$$

Let A_n be defined as

$$A_n(\mathcal{C}^1,\mathcal{C}^2,\mathcal{C}^3,\mathcal{C}^4) = \mathrm{tr}(M^{11}(\mathcal{C}^4)\mathcal{C}^1) + 2\mathrm{tr}((M^{12}(\mathcal{C}^4))^\star\mathcal{C}^3) + \mathrm{tr}(M^{22}(\mathcal{C}^4)\mathcal{C}^2),$$

where $\mathcal{C}^1,\mathcal{C}^2,\mathcal{C}^3,\mathcal{C}^4$ are complex matrices of size $l_n\times l_n, m_n\times m_n, l_n\times m_n$ and $l_n\times m_n$, respectively, and the absolute value of each element of $\mathcal{C}^4$ is less than 1. Then we see $Z^\star MZ$ can be rewritten as

$$Z^\star MZ = A_n(\{Z_iZ_{i'}\}_{i,i'=1}^{l_n}, \{Z_{j+l_n}Z_{j'+l_n}\}_{j,j'=1}^{m_n}, \{Z_iZ_{j+l_n}\}_{i,j}, \{\rho_{I^i,J^j,\tau_n}\}_{i,j}).$$

Let $\mathbf{1}$ denote an $l_n\times m_n$ matrix with all elements equal 1, $\{\nu_n^{p,i}\}_{n\in\mathbb{N},p\in\mathbb{Z}_+,i=1,2}$ be random measures on $[0,T)$ which satisfy

$$\nu_n^{p,1}([0,t)) = b_n^{-1}\sum_I((GG^\star)^p)_{II}1_{\{L(I)\in[0,t)\}},\ \nu_n^{p,2}([0,t)) = b_n^{-1}\sum_J((G^\star G)^p)_{JJ}1_{\{L(J)\in[0,t)\}},$$

and

$$\mathcal{E}^1(t) = \{\delta_{i,i'}1_{\{I^i\cap[0,t)\neq\emptyset\}}\}_{i,i'=1}^{l_n},\quad \mathcal{E}^2(t) = \{\delta_{j,j'}1_{\{J^j\cap[0,t)\neq\emptyset\}}\}_{j,j'=1}^{m_n},$$

where δ denotes the Kronecker delta function. Moreover, for $p\in\mathbb{Z}_+$ and $i=1,2$, let

$$\Psi^{p,i}(f,g) \quad = \quad \Psi^{p,i,n}(f,g) = \int_0^T f(s)\nu_n^{p,i}(ds) - \int_0^T f(s)g(s)ds,$$

for $\mathbb{R}$-valued functions f,g on $[0,T]$ such that f is càdlàg and g is Lebesgue integrable. Note that $b_n^{-1}A_n(\mathcal{E}^1(t),\mathbf{0},\mathbf{0},z\mathbf{1}) = \sum_{p=0}^\infty z^{2p}\nu_n^{p,1}([0,t))$, $b_n^{-1}A_n(\mathbf{0},\mathcal{E}^2(t),\mathbf{0},z\mathbf{1}) = \sum_{p=0}^\infty z^{2p}\nu_n^{p,2}([0,t))$, $\Psi^{0,1}(1_{[0,t)},g) = b_n^{-1}\sum_I 1_{\{L(I)\in[0,t)\}} - \int_0^t g(s)ds$ and $\Psi^{0,2}(1_{[0,t)},g) = b_n^{-1}\sum_J 1_{\{L(J)\in[0,t)\}} - \int_0^t g(s)ds$ for $z\in\mathbb{C}$, $|z|<1$ and $t\in(0,T]$.

To obtain convergence of H_n, we consider the following condition.

[*A3*] There exist $\sigma(\{\Pi_n\}_n)$-measurable left-continuous processes $a_0(t)$ and $c_0(t)$ such that $\int_0^T a_0(t)dt\vee\int_0^T c_0(t)dt < \infty$ almost surely and

$$\Psi^{0,1}(1_{[0,t)},a_0)\vee\Psi^{0,2}(1_{[0,t)},c_0) \to^p 0 \quad \text{as} \quad n\to\infty \tag{1.6}$$

for any $t\in(0,T]$. Moreover, at least one of the following conditions holds true.

1. There exist $\eta\in(0,1)$ and a $\sigma(\{\Pi_n\}_n)$-measurable process $a(z,t)$ such that a is continuous with respect to z and left-continuous with respect to t, $\int_0^T a(z,t)dt<\infty$ and $b_n^{-1}A_n(\mathcal{E}^1(t),\mathbf{0},\mathbf{0},z\mathbf{1}) \to^p \int_0^t a(z,s)ds$ as $n\to\infty$ for $z\in\mathbb{C}, |z|<\eta$ and $t\in(0,T]$.
2. There exist $\eta\in(0,1)$ and a $\sigma(\{\Pi_n\}_n)$-measurable process $c(z,t)$ such that c is continuous with respect to z and left-continuous with respect to t, $\int_0^T c(z,t)dt<\infty$ and $b_n^{-1}A_n(\mathbf{0},\mathcal{E}^2(t),\mathbf{0},z\mathbf{1}) \to^p \int_0^t c(z,s)ds$ as $n\to\infty$ for $z\in\mathbb{C}, |z|<\eta$ and $t\in(0,T]$.

In particular, $\{l_n/b_n\}_n$ and $\{m_n/b_n\}_n$ are tight under (1.6).

$A_n(\mathcal{E}^1(T),\mathbf{0},\mathbf{0},z\mathbf{1})$ and $A_n(\mathbf{0},\mathcal{E}^2(T),\mathbf{0},z\mathbf{1})$ appear in an asymptotically equivalent representation of H_n when $b(x,\sigma)$ does not depend on x and $\mu_t\equiv 0$. Therefore convergence conditions for observation times like [*A3*] 1 and 2 are natural conditions to specify the limit of H_n.

[*A3′*] There exist $\sigma(\{\Pi_n\}_n)$-measurable left-continuous processes $a_0(t)$ and $c_0(t)$ such that $\int_0^T a_0(t)dt\vee\int_0^T c_0(t)dt < \infty$ almost surely and (1.6) holds for any $t\in(0,T]$. Moreover, at least one of the following conditions holds true.

1. For any $p\in\mathbb{N}$, there exists a $\sigma(\{\Pi_n\}_n)$-measurable left-continuous process $a_p(t)$ such that $\int_0^T a_p(t)dt < \infty$ a.s. and for any $t\in(0,T]$, $\Psi^{p,1}(1_{[0,t)},a_p)\to^p 0$ as $n\to\infty$.

2. For any $p \in \mathbb{N}$, there exists a $\sigma(\{\Pi_n\}_n)$-measurable left-continuous process $c_p(t)$ such that $\int_0^T c_p(t)dt < \infty$ a.s. and for any $t \in (0,T]$, $\Psi^{p,2}(1_{[0,t)}, c_p) \to^p 0$ as $n \to \infty$.

As we will show later in Proposition 1.2, $[A3]$ and $[A3']$ are equivalent under $[A2]$.

Let $q > 2$ and $\eta \in (0,1)$. For $\alpha \in (0,1/2)$ and $f : [0,T] \to \mathbb{R}$, $\alpha-$Hölder continuous, we denote $\omega_\alpha(f) = \sup_{t \neq s} |f_t - f_s|/|t-s|^\alpha$.

$[A3'$-$q,\eta]$ There exist $n_0 \in \mathbb{N}$, $\alpha \in (0, 1/2 - 1/q)$ and $\sigma(\{\Pi_n\}_n)$-measurable left-continuous processes $\{a_0(t)\}$, $\{c_0(t)\}$, $\{a_p(t)\}_{p\in\mathbb{N}}$ such that $\int_0^T (c_0 \vee a_p)(t)dt \in L^q(\Omega)$ for $p \in \mathbb{Z}_+$, $E[(l_n + m_n)^q] < \infty$ for $n \in \mathbb{N}$ and

$$\begin{aligned}
\sup_{n\geq n_0} E[(b_n^\eta|\Psi^{0,1}(f,a_0)|)^q] \vee E[(b_n^\eta|\Psi^{0,2}(f,c_0)|)^q] &\leq C\left(\sup_t |f_t|^q + \omega_\alpha(f)^q\right), \\
\max_{i=1,2}\sup_{p\in\mathbb{N}}\sup_{n\geq n_0} E[(b_n^\eta|\Psi^{p,i}(f,a_p)|)^q]/(1+p)^C &\leq C\left(\sup_t |f_t|^q + \omega_\alpha(f)^q\right)
\end{aligned}$$

for any α-Hölder continuous function f on $[0,T]$.

For $q > 2$ and $\eta \in (0,1)$, it can be shown that $[A3'$-$q,\eta]$ implies $[A3']$.

The following lemma is easy to check.

Lemma 1.3. *Let $\{\alpha_p\}_{p\in\mathbb{N}} \subset \mathbb{C}$ with $\sum_{p=1}^\infty |\alpha_p| < \infty$ and $\{\xi_p^n\}_{n,p\in\mathbb{N}}$ and $\{F_n\}_{n\in\mathbb{N}}$ be random variables satisfying $\xi_p^n \to^p 0$ $(n\to\infty)$ for $p \in \mathbb{N}$, $\{F_n\}_{n\in\mathbb{N}}$ are tight, and $|\xi_p^n| \leq F_n$, $(n, p \in \mathbb{N})$. Then $\sum_{p=1}^\infty \alpha_p \xi_p^n \to^p 0$ as $n \to \infty$.*

The equivalence of $[A3]$ 1 and $[A3]$ 2 is established by our next lemma.

Lemma 1.4. *Assume $[A2]$ and that there exist stochastic processes $a_0(t)$ and $c_0(t)$ such that $\int_0^T a_0(t)dt \vee \int_0^T c_0(t)dt < \infty$ a.s. and (1.6) holds for $t \in (0,T]$. Then $\nu_n^{p,1}([0,t)) - \nu_n^{p,2}([0,t)) \to^p 0$ as $n \to \infty$ for $t \in [0,T], p \geq 1$ and*

$$b_n^{-1}A_n(\mathcal{E}^1(t), \mathbf{0}, \mathbf{0}, z\mathbf{1}) - b_n^{-1}A_n(\mathbf{0}, \mathcal{E}^2(t), \mathbf{0}, z\mathbf{1}) \to^p \int_0^t (a_0 - c_0)(s)ds$$

as $n \to \infty$ for $z \in \mathbb{C}, |z| < 1$ and $t \in [0,T]$.

In particular, $[A3]$ 1 $\iff$ $[A3]$ 2, $[A3']$ 1 $\iff$ $[A3']$ 2 and $a_p \equiv c_p$ $dt \times P$-a.e. (t,ω) for $p \geq 1$ under the assumptions above.

Proof. Since $\| G \| \vee \| G^\star \| \leq 1$ by Lemma 1.2, we have $|((GG^\star)^p)_{II'}| \leq 1$, $|G_{IJ}| \leq 1$ for any I, I', J and $p \in \mathbb{Z}_+$. Then since $G^\star_{JI} \neq 0$ implies $I \cap J \neq \emptyset$, we obtain

$$\begin{aligned}
|\nu_n^{p,1}([0,t)) - \nu_n^{p,2}([0,t))| &= b_n^{-1}\Bigg| \sum_{I;L(I)\in[0,t)} \sum_{I'} \sum_J ((GG^\star)^{p-1})_{II'} G_{I'J} G^\star_{JI} \\
&\quad - \sum_{J;L(J)\in[0,t)} \sum_I \sum_{I'} G^\star_{JI} ((GG^\star)^{p-1})_{II'} G_{I'J} \Bigg| \\
&\leq 2b_n^{-1} \sum_{t-r_n \leq L(I) \leq t+r_n} 1 \to^p 0
\end{aligned}$$

as $n \to \infty$ for $p \geq 1$ by $[A2]$ and (1.6).

Since $|\nu_n^{p,1}([0,t)) - \nu_n^{p,2}([0,t))| \leq b_n^{-1}(l_n + m_n)$, the desired conclusions are given by tightness of $\{b_n^{-1}(l_n + m_n)\}_n$ and Lemma 1.3. □

Proposition 1.2. *$[A3]$ and $[A3']$ are equivalent under $[A2]$. Moreover, under $[A2]$ and $[A3]$, $a(\rho,t) = \sum_{p=0}^\infty a_p(t)\rho^{2p}$, $c(\rho,t) = \sum_{p=0}^\infty c_p(t)\rho^{2p}$ and*

$$b_n^{-1}A_n(x^2\mathcal{E}^1(t), y^2\mathcal{E}^2(t), xy\rho_*\mathcal{E}^1(t)G, \rho\mathbf{1}) \to^p \int_0^t A(x,y,\rho,\rho_*,s)ds$$

as $n \to \infty$ for $x, y \in \mathbb{R}, \rho, \rho_* \in (-1,1), t \in (0,T]$, where

$$A(x,y,\rho,\rho_*,t) = x^2 a(\rho,t) + y^2 c(\rho,t) - 2xy(a(\rho,t) - a_0(t))\rho_*/\rho 1_{\{\rho\neq 0\}}.$$

The convergent sequence which appears in Proposition 1.2 is asymptotically equivalent representation of $b_n^{-1}Z^\star MZ$ if $\mu_t \equiv 0$ and $b(x,\sigma)$ does not depend on x. Therefore, the convergence result in Proposition 1.2 is the convergence result of $b_n^{-1}Z^\star MZ$ with ignoring the structure of diffusion coefficients (b_t^1, b_t^2).

1.3.2 The limit of H_n

We discuss the asymptotic behavior of H_n under $[A3], [A3'\text{-}q,\eta]$.

First, we assume one more condition. Let $\mathcal{I}$ be a set of intervals defined by

$$\mathcal{I} = \{I^i\}_{i=1}^{l_n} \cup \{J^j\}_{j=1}^{m_n} \cup \{[\mathcal{T}_k^{j-1}, \mathcal{T}_k^j); 1 \leq k \leq n_2, 1 \leq j \leq m_n^k\}.$$

Let $\theta_{0,k} = I^k$ for $1 \leq k \leq l_n$, $\theta_{0,k} = J^{k-l_n}$ for $l_n < k \leq l_n + m_n$, and

$$\theta_{p,k} = \cup\{K_{2p}; K_1, \cdots, K_{2p} \in \mathcal{I}, K_1 \cap \theta_{0,k} \neq \emptyset, K_j \cap K_{j-1} \neq \emptyset\ (1 \leq j \leq 2p)\}$$

for $p \in \mathbb{N}$ and $1 \leq k \leq l_n + m_n$. Moreover, let $\Phi_{p,i} = \sum_k |\theta_{p,k}|^i$, $\bar{\Phi}_{p_1,p_2} = \sum_{k_1,k_2} |\theta_{p_1,k_1} \cap \theta_{p_2,k_2}|$ for $i \in \{1,2\}$ and $p, p_1, p_2 \in \mathbb{Z}_+$. For $q \geq 2$ and $\delta \geq 1$, we consider the following conditions.

$[A4]$ There exists $\delta' \geq 1$ such that

$$\left\{(b_n^{-1} \vee r_n^2)\sum_{p=0}^\infty \frac{(\Phi_{2p+2,1})^2}{(p+1)^{2\delta'}}\right\} \vee \left\{b_n^{-1}\sum_{p_1,p_2=0}^\infty \frac{\bar{\Phi}_{2p_1+3,2p_2+3}}{(p_1+1)^{\delta'}(p_2+1)^{\delta'}}\right\} \to^p 0$$

as $n \to \infty$.

$[A4\text{-}q,\delta]$

1.
$$\lim_{n\to\infty} E\left[(b_n^{-\frac{q}{2}} \vee r_n^q)\sum_{p=0}^\infty \frac{(\Phi_{2p+2,1})^q}{(p+1)^{q\delta}}\right] = 0.$$

2.
$$\lim_{n\to\infty} E\left[\left(b_n^{-1}\sum_{p_1,p_2=0}^\infty \frac{\bar{\Phi}_{2p_1+3,2p_2+3}}{(p_1+1)^{\delta}(p_2+1)^{\delta}}\right)^{\frac{q}{2}}\right] = 0.$$

We can see that $[A4\text{-}q,\delta]$ implies $[A4]$ for any $q \geq 2$ and $\delta \geq 1$ by Jensen's inequality. Moreover, we can use the following condition instead of $[A4]$.

$[A4']$ There exist positive constants $\delta_1, \delta_2, \delta_3$ such that $(3\delta_1 + 2\delta_3) \vee (\delta_1 + \delta_2) < 1$ and the following two conditions hold:

1. $\lim_{n\to\infty} P[r_n \geq b_n^{-1+\delta_1}] = 0$.
2.
$$\lim_{n\to\infty} b_n^2 \sup_{j_1,j_2\in\mathbb{N}, |j_1-j_2|\geq b_n^{\delta_2}} P\left[l_n \geq j_1 \vee j_2 \text{ and } \frac{|S^{j_2} - S^{j_1}|}{|j_2 - j_1|} \leq b_n^{-1-\delta_3}\right] = 0,$$
$$\lim_{n\to\infty} b_n^2 \sup_{j_1,j_2\in\mathbb{N}, |j_1-j_2|\geq b_n^{\delta_2}} P\left[m_n \geq j_1 \vee j_2 \text{ and } \frac{|T^{j_2} - T^{j_1}|}{|j_2 - j_1|} \leq b_n^{-1-\delta_3}\right] = 0.$$

Lemma 1.5. *Assume $[A4']$ and that $\{(l_n + m_n)/b_n\}_{n\in\mathbb{N}}$ is tight. Then $[A4]$ holds.*

Proof. Let $\phi(k)$ denote minimal $k' > l_n$ which satisfy $I^k \cap J^{k'-l_n} \neq \emptyset$ for $1 \leq k \leq l_n$,

$$\mathcal{U}_{j_1,j_2}^n = \left\{l_n \geq j_1 \vee j_2 \text{ and } \frac{|S^{j_2} - S^{j_1}|}{|j_2 - j_1|} \leq b_n^{-1-\delta_3}\right\}^c \bigcap \left\{m_n \geq j_1 \vee j_2 \text{ and } \frac{|T^{j_2} - T^{j_1}|}{|j_2 - j_1|} \leq b_n^{-1-\delta_3}\right\}^c$$

for $j_1, j_2 \in \mathbb{N}$ and

$$\bar{\mathcal{U}}_n = \{r_n < b_n^{-1+\delta_1}\} \cap \cap_{j_1,j_2\in\mathbb{N},|j_2-j_1|\geq b_n^{\delta_2}} \mathcal{U}^n_{j_1,j_2}.$$

Then on $\bar{\mathcal{U}}_n$, for $k_1 \leq l_n$, $l_n < k_2 \leq l_n + m_n$ and $p_1, p_2 \in \mathbb{Z}_+$ which satisfy $|\phi(k_1) - k_2| \geq b_n^{\delta_2}$ and $\theta_{p_1,k_1} \cap \theta_{p_2,k_2} \neq \emptyset$, we have

$$|\phi(k_1) - k_2| b_n^{-1-\delta_3} < |T^{\phi(k_1)} - T^{k_2}| \leq (2p_1 + 2p_2 + 2) r_n < (2p_1 + 2p_2 + 2) b_n^{-1+\delta_1}.$$

Therefore $|\phi(k_1) - k_2| b_n^{-\delta_1-\delta_3} < 2(p_1+1)(p_2+1)$.

Then by using the relation $|\theta_{p_1,k_1} \cap \theta_{p_2,k_2}| \leq \{(4p_1+1) \wedge (4p_2+1)\} r_n$, we obtain

$$\begin{aligned}
&\sum_{p_1,p_2=0}^{\infty} \sum_{k_1 \leq l_n, k_2 > l_n} \frac{|\theta_{p_1,k_1} \cap \theta_{p_2,k_2}|}{(p_1+1)^5 (p_2+1)^5} \\
\leq\ & C \sum_{k_1 \leq l_n} \left\{ \sum_{k_2 > l_n, |\phi(k_1)-k_2| \geq b_n^{\delta_2}} \sum_{p_1,p_2} \frac{\{(4p_1+1) \wedge (4p_2+1)\} b_n^{-1+\delta_1}}{(p_1+1)^3 (p_2+1)^3 |\phi(k_1) - k_2|^2 b_n^{-2\delta_1 - 2\delta_3}} + b_n^{\delta_2} b_n^{-1+\delta_1} \right\} \\
\leq\ & C b_n^{-1+(3\delta_1+2\delta_3)\vee(\delta_1+\delta_2)} \sum_{k_1 \leq l_n} \left(\sum_{k_2 > l_n, |\phi(k_1)-k_2| \geq b_n^{\delta_2}} \frac{1}{|\phi(k_1) - k_2|^2} + 1 \right) \\
\leq\ & C l_n b_n^{-1+(3\delta_1+2\delta_3)\vee(\delta_1+\delta_2)}
\end{aligned}$$

on $\bar{\mathcal{U}}_n$. Similar arguments for other combinations of k_1 and k_2 yield

$$\sum_{p_1,p_2=0}^{\infty} \frac{\bar{\Phi}_{2p_1+3,2p_2+3}}{(p_1+1)^5(p_2+1)^5} \leq C(l_n + m_n) b_n^{-1+(3\delta_1+2\delta_3)\vee(\delta_1+\delta_2)} \quad \text{on} \quad \bar{\mathcal{U}}_n.$$

On the other hand, for any $\epsilon > 0$ there exists a positive constant K such that $P[(l_n + m_n)/b_n > K] < \epsilon$ for any $n \in \mathbb{N}$ by tightness of $\{(l_n + m_n)/b_n\}_n$. Then we obtain $\limsup_{n\to\infty} P[\bar{\mathcal{U}}_n^c] \leq \epsilon$ by [A4′] and the inequality

$$P[\bar{\mathcal{U}}_n^c] \leq P[r_n \geq b_n^{-1+\delta_1}] + P[(l_n + m_n)/b_n > K] + \sum_{1 \leq j_1, j_2 \leq [K b_n], |j_2 - j_1| \geq b_n^{\delta_2}} P[(\mathcal{U}^n_{j_1,j_2})^c].$$

Since $\epsilon > 0$ is arbitrary, we have

$$b_n^{-1} \sum_{p_1,p_2=0}^{\infty} \frac{\bar{\Phi}_{2p_1+3,2p_2+3}}{(p_1+1)^5(p_2+1)^5} \to^p 0.$$

It is easier to prove the convergence about $\Phi_{2p+2,1}$. □

Let $B^i(x,\sigma) = |b^i(x,\sigma_*)|/|b^i(x,\sigma)|$ $(i = 1,2)$,

$$C(\rho,t) = \sum_{p=1}^{\infty} \frac{a_p(t)}{p} \rho^{2p} = \sum_{p=1}^{\infty} \frac{c_p(t)}{p} \rho^{2p},$$

and

$$h_t^{\infty}(\sigma) = -\frac{1}{2} A(B_t^1, B_t^2, \rho_t, \rho_{t,*}, t) - a_0 \log|b_t^1| - c_0 \log|b_t^2| + \frac{1}{2} C(\rho_t, t) \tag{1.7}$$

for $t \in [0,T], \rho \in (-1,1)$.

Proposition 1.3. 1. *Let $0 \leq v \leq 3$. Assume [A1] − [A3]. Then*

$$\sup_{\sigma\in\Lambda} \left| b_n^{-1} \partial_\sigma^v H_n(\sigma) - \int_0^T \partial_\sigma^v h_t^\infty(\sigma) dt \right| \to^p 0$$

as $n \to \infty$.

2. *Let $0 \leq v \leq 3$, $q \in 2\mathbb{N}$, $q > 2 \vee n_1$, $\delta > 1$, $\xi > 0$, $\eta \in (0,1)$ and $\{s_n\}_{n\in\mathbb{N}}$ be stochastic processes. Assume [A1], [A2], [A3′-q,η], [A4-$(2q),\delta$], [S-$((2v + 2[\delta] + 12)q),\xi$] for $\{s_n\}$, and that $\sup_n E[b_n^{-2q}(l_n + m_n)^{2q}] < \infty$. Then*

$$\sup_n E\left[\left(\sup_{\sigma\in\Lambda} b_n^{\eta'} \left| b_n^{-1} \partial_\sigma^v \hat{H}_n(\sigma; s_n) - \int_0^T \partial_\sigma^v h_t^\infty(\sigma) dt \right| \right)^q\right] < \infty$$

for $\eta' \leq \eta \wedge (1/2) \wedge (\xi/(2q))$.

1.4 Separation of the limit of H_n

We deal with Condition $[H]$ about separation of the limit of H_n which is necessary to apply Ibragimov-Has'minskii's theory. When the sampling scheme is synchronous and equi-spaced : $S^i = T^i = [b_n]^{-1} iT$ $(0 \leq i \leq [b_n])$, Uchida and Yoshida [51] discussed tractable sufficient conditions for Condition $[H0]$ of separation. In this section, we will confirm that $[H0]$ implies $[H]$ under certain conditions.

Under $[A1] - [A3]$, we define $\mathcal{Y}_n(\sigma;\sigma_*) = b_n^{-1}(H_n(\sigma) - H_n(\sigma_*))$, and $\mathcal{Y}(\sigma;\sigma_*)$ denotes the probability limit of $\mathcal{Y}_n(\sigma;\sigma_*)$. By Proposition 1.3, we obtain $\mathcal{Y}(\sigma;\sigma_*) = \int_0^T (h_t^\infty(\sigma) - h_t^\infty(\sigma_*))dt$.

Moreover, the equation (1.7) can be rewritten as

$$\begin{aligned} h_t^\infty(\sigma) &= -\frac{1}{2}(B_t^1)^2(a_0 + \mathcal{A}(\rho_t)) - \frac{1}{2}(B_t^2)^2(c_0 + \mathcal{A}(\rho_t)) + B_t^1 B_t^2 \mathcal{A}(\rho_t)\frac{\rho_{t,*}}{\rho_t} \\ &\quad -a_0 \log|b_t^1| - c_0 \log|b_t^2| + \int_0^{\rho_t} \frac{\mathcal{A}(\rho)}{\rho} d\rho, \end{aligned} \tag{1.8}$$

where $\mathcal{A}(\rho) = \mathcal{A}(\rho,t) = a(\rho,t) - a_0(t) = c(\rho,t) - c_0(t)$ and we regard $\mathcal{A}(\rho)/\rho = 0$ when $\rho = 0$. Since $B_{t,*}^1 = B_{t,*}^2 = 1$,

$$h_t^\infty(\sigma_*) = -\frac{1}{2}a_0 - \frac{1}{2}c_0 - a_0 \log|b_{t,*}^1| - c_0 \log|b_{t,*}^2| + \int_0^{\rho_{t,*}} \frac{\mathcal{A}(\rho)}{\rho} d\rho.$$

Therefore for $y_t(\sigma) = h_t^\infty(\sigma) - h_t^\infty(\sigma_*)$, it follows that

$$\begin{aligned} y_t(\sigma) &= -\frac{1}{2}(B_t^1)^2(a_0 + \mathcal{A}) - \frac{1}{2}(B_t^2)^2(c_0 + \mathcal{A}) + B_t^1 B_t^2 \mathcal{A}\frac{\rho_{t,*}}{\rho_t} + \frac{a_0}{2} + \frac{c_0}{2} \\ &\quad +a_0 \log B_t^1 + c_0 \log B_t^2 + \int_{\rho_{t,*}}^{\rho_t} \frac{\mathcal{A}}{\rho} d\rho \\ &= -\frac{1}{2}(a_0 + \mathcal{A})(B_t^1 - B_t^2)^2 + a_0 + a_0 \log B_t^1 B_t^2 + \int_{\rho_{t,*}}^{\rho_t} \frac{\mathcal{A}}{\rho} d\rho \\ &\quad +\frac{c_0 - a_0}{2}(1 - (B_t^2)^2 + \log(B_t^2)^2) + B_t^1 B_t^2(\mathcal{A}\rho_{t,*}/\rho_t - \mathcal{A} - a_0) \\ &= -\frac{1}{2}(c_0 + \mathcal{A})(B_t^1 - B_t^2)^2 + c_0 + c_0 \log B_t^1 B_t^2 + \int_{\rho_{t,*}}^{\rho_t} \frac{\mathcal{A}}{\rho} d\rho \\ &\quad +\frac{a_0 - c_0}{2}(1 - (B_t^1)^2 + \log(B_t^1)^2) + B_t^1 B_t^2(\mathcal{A}\rho_{t,*}/\rho_t - \mathcal{A} - c_0). \end{aligned} \tag{1.9}$$

Let $F(x) = 1 - x + \log x$ $(x > 0)$.

Lemma 1.6. *Let $\epsilon_1 \in (0, e^{-1}]$, $\epsilon_2 \geq 1$. Then*

$$-\log(1/\epsilon_1)(x-1)^2 \leq F(x) \leq -(x-1)^2/4\epsilon_2^2,$$

for $x \in [\epsilon_1, 1 + \epsilon_2]$.

Proof. For $0 < x \leq 1 + \epsilon_2$, since $-1/\epsilon_2 < (x-1)/\epsilon_2 \leq 1$, it follows that

$$F(x) \leq -((x-1)^2 \wedge 1)/4 \leq -(x-1)^2/4\epsilon_2^2.$$

On the other hand, for $x \geq \epsilon_1$, let $f(x) = F(x) + \log(1/\epsilon_1)(x-1)^2$ then since $f'(x) = (x-1)(2\log(1/\epsilon_1) - 1/x)$, we have $f(x) \geq f(1) \wedge f(\epsilon_1) \geq 0$. □

Let

$$\begin{aligned} f_1(t,x,\rho,\rho_*) &= a_0 + a_0 \log x + \int_{\rho_*}^{\rho} \mathcal{A}(\rho')/\rho' d\rho' + x(\mathcal{A}\rho_*/\rho - \mathcal{A} - a_0), \\ f_2(t,x,\rho,\rho_*) &= c_0 + c_0 \log x + \int_{\rho_*}^{\rho} \mathcal{A}(\rho')/\rho' d\rho' + x(\mathcal{A}\rho_*/\rho - \mathcal{A} - c_0), \end{aligned}$$

$$R = \max_{0\le i\le 4}\max_{0\le j\le 3}\max_{1\le p\le 2}\sup_{\sigma\in\Lambda, t_k\in[0,T]\ (1\le k\le n_2)}\left(|\partial_x^j\partial_\sigma^i b^p| \vee \left|\partial_x^j\partial_\sigma^i\frac{1}{|b^p|}\right|(\{X_{t_k}^k\}_{k=1}^{n_2},\sigma)\right)$$
$$\vee \max_{1\le i\le 3}\sup_{t\in[0,T]}|\tilde{b}_t^i| \vee \max_{2\le i\le 3, 1\le j\le 3}\sup_{t\in[0,T]}|\hat{b}_t^{ij}|,$$

and $C_1 = (1-\bar{\rho}_T^2)^2/(12R^8)$. Then $R \ge 1$ and $E[R^q] < \infty$ for any $q > 0$ under $[A1]$.

Lemma 1.7. *Assume* $[A1]-[A3]$. *Then*

$$(f_1 \vee f_2)(t, B_t^1 B_t^2, \rho_t, \rho_{t,*}) \le -C_1\left\{a_1(t)(\rho_t-\rho_{t,*})^2 + (a_0\wedge c_0)(t)(B_t^1B_t^2-1)^2\right\}$$

for $dt\times P$*-a.e.* $(t,\omega)\in[0,T]\times\Omega$.

We write $\mathcal{Y}_0$ for $\mathcal{Y}$ defined by using the same processes X, Y and the synchronous, equi-spaced sampling $S^i = T^i = \mathcal{T}_k^i = [b_n]^{-1}iT$ $(0\le i\le [b_n], 1\le k\le n_2)$. Let

$$\chi(\sigma_*) = \inf_{\sigma\ne\sigma_*}\frac{-\mathcal{Y}(\sigma;\sigma_*)}{|\sigma-\sigma_*|^2}, \quad \chi_0(\sigma_*) = \inf_{\sigma\ne\sigma_*}\frac{-\mathcal{Y}_0(\sigma;\sigma_*)}{|\sigma-\sigma_*|^2}.$$

Moreover, we consider the following conditions.

$[H]$ For every $L > 0$, there exists $c_L > 0$ such that for all $r > 0$, $P[\chi \le r^{-1}] \le c_L/r^L$.

$[H0]$ For every $L > 0$, there exists $c_L > 0$ such that for all $r > 0$, $P[\chi_0 \le r^{-1}] \le c_L/r^L$.

$[H']$ $\chi > 0$ almost surely.

Obviously, $[H]$ implies $[H']$.

Lemma 1.8. *Assume* $[A1]$. *Then there exists a positive random variable* $\mathcal{R}'$ *which does not depend on* σ, σ_*, *such that* $E[(\mathcal{R}')^q] < \infty$ *for any* $q > 0$ *and*

$$\mathcal{Y}_0(\sigma;\sigma_*) \ge -\mathcal{R}'\int_0^T\left\{(B_t^1-B_t^2)^2 + (B_t^1B_t^2-1)^2 + (\rho_t-\rho_{t,*})^2\right\}dt$$

for any $\sigma, \sigma_*\in\Lambda$.

The following proposition ensures that to prove $[H]$, it is enough to prove $[H0]$ which is the condition of separation for synchronous, equi-spaced observations.

Proposition 1.4. *Assume* $[A1]-[A3]$. *Then there exists a positive random variable* $\mathcal{R}$ *which do not depend on* σ, σ_* *such that* $E[\mathcal{R}^{-q}] < \infty$ *for any* $q \ge 1$ *and*

$$\mathcal{Y}(\sigma;\sigma_*) \le -\mathcal{R}\int_0^T\left\{(a_0\wedge c_0)\{(B_t^1-B_t^2)^2 + (B_t^1B_t^2-1)^2\} + a_1(\rho_t-\rho_{t,*})^2\right\}dt$$

for $\sigma,\sigma_*\in\Lambda$. *In particular, if* $E[(\operatorname{ess\,inf}_{t\in[0,T]}a_1(t))^{-q}] < \infty$ *for any* $q > 0$, $[H0]$ *implies* $[H]$.

Proof. In the equation (1.9), by using the second representation if $a_0 \le c_0$ and using the third representation if $a_0 > c_0$, we obtain

$$y_t(\sigma) \le -\frac{1}{2}(a_0\wedge c_0 + \mathcal{A})(B_t^1-B_t^2)^2 + (f_1\vee f_2)(t, B_t^1B_t^2, \rho_t, \rho_{t,*}).$$

By Lemma 1.7, we have

$$y_t(\sigma) \le -\frac{1}{2}(a_0\wedge c_0)(B_t^1-B_t^2)^2 - C_1 a_1(\rho_t-\rho_{t,*})^2 - C_1(a_0\wedge c_0)(B_t^1B_t^2-1)^2$$

for $dt\times P$-a.e. (t,ω), where $C_1 = (1-\bar{\rho}_T^2)^2/(12R^8)$. Therefore by integrating with respect to t,

$$\mathcal{Y}(\sigma;\sigma_*) \le -\mathcal{R}\int_0^T\left((a_0\wedge c_0)\{(B_t^1-B_t^2)^2 + (B_t^1B_t^2-1)^2\} + a_1(\rho_t-\rho_{t,*})^2\right)dt,$$

where $\mathcal{R} = C_1$. In particular, let $E[(\mathrm{ess\,inf}_{t\in[0,T]}a_1(t))^{-q}] < \infty$ for any $q > 0$ and $[H0]$ holds. We can see $a_0 \wedge c_0 \geq a_1$ for $dt \times P$-a.e. (t,ω) since $\nu^{1,i}([0,t)) \leq \nu^{0,i}([0,t))$ for any $t \in (0,T]$ and $i = 1,2$. Therefore by Lemma 1.8 we have

$$\begin{aligned}\mathcal{Y}(\sigma;\sigma_*) &\leq -\mathcal{R}\mathrm{ess\,inf}_t a_1(t)\int_0^T \left((B_t^1 - B_t^2)^2 + (B_t^1B_t^2 - 1)^2 + (\rho_t - \rho_{t,*})^2\right) dt \\ &\leq \mathcal{R}(\mathcal{R}')^{-1}\mathrm{ess\,inf}_t a_1(t)\mathcal{Y}_0(\sigma;\sigma_*)\end{aligned}$$

almost surely, where $\mathcal{R}'$ is defined in Lemma 1.8. Hence $\chi \geq \mathcal{R}(\mathcal{R}')^{-1}\mathrm{ess\,inf}_t a_1(t)\chi_0$ a.s. and for any $L > 0$, there exists a constant $c_L > 0$ such that

$$\begin{aligned}P[\chi \leq r^{-1}] &\leq P[\chi_0 \leq r^{-1/2}] + P[\mathcal{R}(\mathcal{R}')^{-1}\mathrm{ess\,inf}_t a_1(t) \leq r^{-1/2}] \\ &\leq \frac{c_{2L,0}}{r^L} + \frac{1}{r^L}E\left[\left(\mathcal{R}'\mathcal{R}^{-1}(\mathrm{ess\,inf}_t a_1(t))^{-1}\right)^{2L}\right] \leq \frac{c_L}{r^L},\end{aligned}$$

where $c_{2L,0}$ denotes the coefficient of r^{-2L} in $[H0]$. This gives $[H]$. □

Remark 1.2. *In the case of nonsynchronous observations, under* $[A1]$ *and* $[A3]$*, we can prove an inequality*

$$\mathcal{Y}(\sigma;\sigma_*) \geq -\mathcal{R}'\int_0^T \left((a_0 \vee c_0)\{(B_t^1 - B_t^2)^2 + (B_t^1B_t^2 - 1)^2\} + a_1(\rho_t - \rho_{t,*})^2\right) dt,$$

which corresponds to Lemma 1.8.

Remark 1.3. *By Proposition 1.4, it follows that*

$$\mathcal{Y}_0(\sigma;\sigma_*) \leq -\mathcal{R}\int_0^T \{(B_t^1 - B_t^2)^2 + (B_t^1B_t^2 - 1)^2 + (\rho_t - \rho_{t,*})^2\}dt.$$

On the other hand, we can see that there exists a positive random variable $\tilde{\mathcal{R}}$ *such that* $E[\tilde{\mathcal{R}}^q] < \infty$ *for any* $q > 0$ *and*

$$|(bb^\star)_t - (bb^\star)_{t,*}|^2 \leq \tilde{\mathcal{R}}\{(B_t^1 - B_t^2)^2 + (B_t^1B_t^2 - 1)^2 + (\rho_t - \rho_{t,*})^2\}$$

for any $t \in [0,T], \sigma, \sigma_* \in \Lambda$ *by using the inequality* $(x-1)^2 + (y-1)^2 \leq (x-y)^2 + 2(xy-1)^2$ $(x, y \geq 0)$*. Therefore* $[H0]$ *holds if there exists a constant* $\epsilon > 0$ *such that*

$$|(bb^\star)(x,\sigma_1) - (bb^\star)(x,\sigma_2)| \geq \epsilon|\sigma_1 - \sigma_2|$$

for any $x \in \mathbb{R}^{n_2}, \sigma_1, \sigma_2 \in \Lambda$*. Weaker sufficient conditions for* $[H0]$ *can be found in Uchida and Yoshida [51].*

1.5 Asymptotic properties of the quasi-maximum-likelihood estimator and the Bayes type estimator

In this section, we investigate consistency and asymptotic mixed normality of the quasi-maximum-likelihood estimator and the Bayes type estimator as main results.

Let the quasi-maximum likelihood estimator $\hat{\sigma}_n$ of the parameter σ_* be $\sigma \in \bar{\Lambda}$ which maximizes H_n. If maximizing points are not unique, we select so that $\hat{\sigma}_n$ is measurable.

Theorem 1.1. *Assume* $[A1]-[A3], [H']$*. Then* $\hat{\sigma}_n \to^p \sigma_*$ *as* $n \to \infty$*.*

Proof. By Proposition 1.3, we have $\sup_\sigma |\mathcal{Y}_n(\sigma;\sigma_*) - \mathcal{Y}(\sigma;\sigma_*)| \to^p 0$ as $n \to \infty$. On the other hand, by $[H']$, for any $\epsilon, \delta > 0$, there exists $\eta > 0$ such that $P[\chi \leq \eta] \leq \epsilon$. Since $\mathcal{Y}_n(\hat{\sigma}_n;\sigma_*) \geq 0$, it follows that

$$P[|\hat{\sigma}_n - \sigma_*| \geq \delta] \leq P[\chi \leq \eta] + P[\mathcal{Y}(\hat{\sigma}_n;\sigma_*) \leq -\eta\delta^2] \leq \epsilon + P[\sup_\sigma |\mathcal{Y}_n(\sigma;\sigma_*) - \mathcal{Y}(\sigma;\sigma_*)| \geq \eta\delta^2].$$

Therefore there exists $n_0 \in \mathbb{N}$ such that $P[|\hat{\sigma}_n - \sigma_*| \geq \delta] \leq 2\epsilon$ if $n \geq n_0$. □

Let $\{s_n\}_{n\in\mathbb{N}}$ be stochastic processes which satisfy $[S]$,

$$\Gamma = -\int_0^T \partial_\sigma^2 h_t^\infty(\sigma_*)dt,$$

$U_n(\sigma_*) = \{u \in \mathbb{R}^{n_1}; \sigma_* + b_n^{-1/2}u \in \Lambda\}$, $V_n(r,\sigma_*) = \{|u| \geq r\} \cap U_n(\sigma_*)$, and $\mathcal{Z}_n(u,\sigma_*) = \exp(\hat{H}_n(\sigma_* + b_n^{-1/2}u; s_n) - \hat{H}_n(\sigma_*; s_n))$ for $u \in U_n(\sigma)$.

Proposition 1.5. *(polynomial type large deviation inequalities) Let $L > 0$ and $\delta \in (0, 1/2)$. Assume for any $q > 0$ there exists $q' \in 2\mathbb{N}, q' > q$ and $\delta' \geq 1$ such that $[A1]$, $[A2]$, $[A3'$-$q', \delta]$, $[A4$-$q', \delta']$, $[H]$ and $[S$-$q', 2q'\delta]$ hold for $\{s_n\}$. Then there exists $C_L > 0$ such that*

$$P\Big[\sup_{u\in V_n(r,\sigma_*)} \mathcal{Z}_n(u,\sigma_*) \geq e^{-r/2}\Big] \leq \frac{C_L}{r^L}$$

for any $n \in \mathbb{N}$ and $r > 0$.

Let $\mathcal{N}$ be an n_1-dimensional standard normal random variable which is defined on an extension of $(\Omega, \mathcal{F}, P)$ and independent of $\mathcal{F}$. We use the same notation E for expectations on the extension of $(\Omega, \mathcal{F}, P)$.

Theorem 1.2. *1. Assume $[A1] - [A4], [H']$. Then $b_n^{1/2}(\hat{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \Gamma^{-1/2}\mathcal{N}$ as $n \to \infty$.*

2. Let $\delta \in (0, 1/2)$. Assume for any $q > 0$, there exists $q' \in 2\mathbb{N}, q' > q$ and $\delta' \geq 1$ such that $[A1]$, $[A2$-$q', \delta]$, $[A3'$-$q', \delta]$, $[A4$-$q', \delta']$, $[H]$ hold. Then $E[Y'\mathsf{f}(b_n^{1/2}(\hat{\sigma}_n - \sigma_*))] \to E[Y'\mathsf{f}(\Gamma^{-1/2}\mathcal{N})]$ as $n \to \infty$ for any continuous function f of at most polynomial growth and any bounded random variable Y' on $(\Omega, \mathcal{F})$.

We also consider the Bayes type estimator $\tilde{\sigma}_n$ for a prior density $\pi : \Lambda \to \mathbb{R}_+$ defined as

$$\tilde{\sigma}_n = \left(\int_\Lambda \exp(H_n(\sigma))\pi(\sigma)d\sigma\right)^{-1} \int_\Lambda \sigma \exp(H_n(\sigma))\pi(\sigma)d\sigma. \tag{1.10}$$

Theorem 1.3. *Let $\delta \in (0, 1/2)$. Assume for any $q > 0$ there exists $q' \in 2\mathbb{N}, q' > q$ and $\delta' \geq 1$ such that $[A1]$, $[A2$-$q', \delta]$, $[A3'$-$q', \delta]$, $[A4$-$q', \delta']$, $[H]$ hold, and that the prior density π is continuous and $0 < \inf_\sigma \pi(\sigma) \leq \sup_\sigma \pi(\sigma) < \infty$. Then $b_n^{1/2}(\tilde{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \Gamma^{-1/2}\mathcal{N}$ as $n \to \infty$. Moreover, $E[Y'\mathsf{f}(b_n^{1/2}(\tilde{\sigma}_n - \sigma_*))] \to E[Y'\mathsf{f}(\Gamma^{-1/2}\mathcal{N})]$ as $n \to \infty$ for any continuous function f of at most polynomial growth and any bounded random variable Y' on $(\Omega, \mathcal{F})$.*

1.6 Sufficient conditions for the conditions about the observation times

In this section, we will introduce tractable sufficient conditions for $[A2$-$q, \delta]$, $[A3'$-$q, \eta]$, $[A4$-$q, \delta]$, and the estimate with respect to $\mathrm{ess\,inf}_t a_1$ in Proposition 1.4.

Let $q > 0$. We consider the following conditions for point processes $\{N_t^i\}_{0\leq t\leq T, 1\leq i\leq n_2+2}$ which generate observations.

$[B1$-$q]$ There exists $n_0 \in \mathbb{N}$ such that

$$\sup_{n\geq n_0} \max_{1\leq i\leq n_2+2} \sup_{0\leq t\leq T-b_n^{-1}} E\left[(N^i_{t+b_n^{-1}} - N^i_t)^q\right] < \infty.$$

$[B2$-$q]$ There exists $n_0 \in \mathbb{N}$ such that

$$\limsup_{u\to\infty} \sup_{n\geq n_0} \max_{1\leq i\leq n_2+2} \sup_{0\leq t\leq T-ub_n^{-1}} u^q P[N^i_{t+ub_n^{-1}} - N^i_t = 0] < \infty.$$

For example, let $X \equiv Y$, $\{b_n\}$ be a positive integer valued sequence, $\{\bar{N}^1_t\}, \{\bar{N}^2_t\}$ be two independent homogeneous Poisson processes with positive intensities λ_1, λ_2 respectively, and stochastic processes $\{N^1_t\}, \{N^2_t\}$ satisfy $N^i_t = \bar{N}^i_{b_n t}$, $(i = 1, 2)$. Then [$B1$-q] obviously holds for any $q > 0$. Moreover, [$B2$-q] holds for any $q > 0$ since

$$\lim_{u\to\infty} u^q \sup_{i=1,2}\sup_{n\in\mathbb{N}} \sup_{0\le t\le T-ub_n^{-1}} P[N^i_{t+ub_n^{-1}} - N^i_t = 0] = \lim_{u\to\infty} u^q e^{-(\lambda_1\wedge\lambda_2)u} = 0.$$

We will investigate sufficient conditions of [$A3'$-q,η]. First, we denote $t_k = Tk/[b_n]$ $(0 \le k \le [b_n])$, $\mathcal{G}^n_{j,k} = \sigma(N^i_t - N^i_s; t_j \le s < t \le t_k, i = 1, 2)$ $(0 \le j < k \le [b_n])$, $\alpha^n_0 = 1/4$ and

$$\alpha^n_k = 0 \vee \sup_{1\le j_1,j_2\le[b_n]-1, j_2-j_1\ge k}\ \sup_{A\in\mathcal{G}^n_{0,j_1}}\ \sup_{B\in\mathcal{G}^n_{j_2,[b_n]}} |P(A\cap B) - P(A)P(B)| \tag{1.11}$$

for $k \in \mathbb{N}$.

Let $\zeta^{p,i}_n$ be measures which satisfy $\zeta^{p,i}_n([s,t)) = E[\nu^{p,i}_n([s,t))]$. Moreover, for $n_0 \in \mathbb{N}$, $\epsilon' > 0$, a Lebesgue integrable function $g : [0,T] \mapsto \mathbb{R}$, and a continuous function $f : [0,T] \mapsto \mathbb{R}$, we define

$$\bar{\Psi}^{p,i}_{n_0,\epsilon'}(f;g) = \sup_{n\ge n_0} b_n^{\epsilon'} \left| \int_0^T f_t \zeta^{p,i}_n(dt) - \int_0^T f_t g_t dt \right|.$$

Proposition 1.6. *Let $q \in 2\mathbb{N}$, $q > 2$, $\epsilon \in (0,1)$, $\delta > 0$ and $\beta \in (0, 1/2 - 1/q)$. Assume that [$B1$-$(q(1+\delta))$], [$B2$-$(q\epsilon)$] hold, $E[(N^1_T + N^2_T)^q] < \infty$ for $n \in \mathbb{N}$ and there exists $n_0 \in \mathbb{N}$ such that*

$$\mathcal{S} = \sup_{n\ge n_0} \sum_{k=0}^\infty (k+1)^{q-2+(q-1)/\delta} \alpha^n_k < \infty. \tag{1.12}$$

Moreover, assume that there exist $\epsilon' > 0$, $C > 0$, and left-continuous deterministic functions $a_0(t), c_0(t), a_p(t)$ ($p \in \mathbb{N}$) such that $\int_0^T a_p(t)dt < \infty$ ($p \in \mathbb{Z}_+$), $\int_0^T c_0(t)dt < \infty$ and

$$\bar{\Psi}^{0,1}_{n_0,\epsilon'}(f;a_0) \vee \bar{\Psi}^{0,2}_{n_0,\epsilon'}(f;c_0) \vee \max_{i=1,2}\sup_{p\in\mathbb{N}} \frac{\bar{\Psi}^{p,i}_{n_0,\epsilon'}(f;a_p)}{(p+1)^C} \le C(\sup_t |f_t| + \omega_\beta(f)) \tag{1.13}$$

for any β-Hölder continuous function $f : [0,T] \to \mathbb{R}$. Then [$A3'$-q,η] holds for $\eta = \epsilon' \wedge \beta \wedge (\delta\epsilon/(2(1+\delta-\delta\epsilon)))$ with α in [$A3'$-q,η] equal to β.

For example, let $\{\bar{N}^i_t\}_{t\ge0}$ be a point process where the distribution of $(\bar{N}^i_{t+t_k} - \bar{N}^i_{t+t_{k-1}})_{k=1}^M$ does not depend on $t \ge 0$ for $i = 1, 2$, $M \in \mathbb{N}$ and $0 \le t_0 < t_1 < \cdots < t_M$. Moreover, let $\{N^i_t\}_t$ satisfy $N^i_t = \bar{N}^i_{[b_n]t}$ for $t \in [0,T]$, $i = 1, 2$ and $n \in \mathbb{N}$. Then the relation (1.13) holds if [$B1$-2] and [$B2$-ϵ] hold for some $\epsilon \in (0, 2]$. In this case, we obtain $a_p = T^{-1}\lim_{n\to\infty} E[\nu^{p,1}_n([0,T))]$, $c_0 = T^{-1}\lim_{n\to\infty} E[\nu^{0,2}_n([0,T))]$, and $\epsilon' = (\epsilon/4)\wedge\beta$. In particular, $\{a_p\}_{p\in\mathbb{Z}_+}, c_0$ are constants.

For general $\{N^i_t\}$, the following propositions are sufficient conditions for [$A4$-q,δ], [$A2$-q,δ] and the estimate of $\mathrm{ess\,inf}_t a_1(t)$ in Proposition 1.4.

Proposition 1.7. *Let $q \in 2\mathbb{N}$, $q > 2$, $p'_1, p'_2 > 1$, $1/p'_1 + 1/p'_2 = 1$. Assume [$B1$-$(p'_1 q)$] and [$B2$-$(p'_2(q+2))$].*

1. *Then there exists $n_0 \in \mathbb{N}$ such that $\sup_{n\ge n_0} E[(\Phi_{p,1})^q] \le C(p+1)^{q+1}$ for any $p \in \mathbb{Z}_+$. In particular, [$A4$-$q', (1+3/q')$] 1 holds for any $q' \in [2,q)$.*
2. *Then there exists $n_0 \in \mathbb{N}$ such that*

$$\sup_{n\ge n_0} E[(\bar{\Phi}_{p_1,p_2})^{\frac{q}{2}}] \le C(p_1+1)^{\frac{q}{2}+1}(p_2+1)^{\frac{q}{2}+1}$$

for $p_1, p_2 \in \mathbb{Z}_+$. In particular, [$A4$-$q$, 3] 2 holds.

Proposition 1.8. *Let $q \in \mathbb{N}$ and we assume [$B2$-$(q+1)$]. Then there exists $n_0 \in \mathbb{N}$ such that $\sup_{n\ge n_0} E\left[b_n^{q-1} r_n^q\right] < \infty$.*

Proposition 1.9. *Assume there exists $n_1 \in \mathbb{N}$ and $q > 0$ such that $[A3']$ and $[B2\text{-}q]$ hold, $a_1(t)$ does not depend on t, $\{N^i_{t+[b_n]^{-1}T} - N^i_t\}_{0\le t\le T-[b_n]^{-1}T, n\ge n_1, i=1,2}$ is tight and α-mixing coefficients $\{\alpha^n_k\}_k$ defined by (1.11) satisfy $\sup_{n\ge n_1}\sum_{k=1}^\infty k\alpha^n_k < \infty$. Then there exists a constant $\delta > 0$ such that $a_1 \ge \delta$ almost surely.*

Finally, we state a corollary of main theorems.

We assume $\{\bar{N}^i_t\}_{t\ge 0}$ is an exponential α-mixing point process where $E[(\bar{N}^1_1 + \bar{N}^2_1)^q] < \infty$ for $q > 0$ and the distribution of $(\bar{N}^i_{t+t_k} - \bar{N}^i_{t+t_{k-1}})_{k=1}^M$ does not depend on $t \ge 0$ for $i = 1, 2$, $M \in \mathbb{N}$ and $0 \le t_0 < t_1 < \cdots < t_M$ for $1 \le i \le 2$. Let $\{N^i_t\}$ satisfy $N^i_t = \bar{N}^i_{[b_n]t}$ for $t \in [0, T]$, $i = 1, 2$ and $n \in \mathbb{N}$, $\hat{\sigma}_n$ be the quasi-maximum-likelihood estimator defined by H_n, $\pi : \Lambda \to \mathbb{R}_+$ be a continuous function and $\tilde{\sigma}_n$ be defined by (1.10).

Corollary 1.1. *Assume $[A1], [H0], [B1\text{-}q], [B2\text{-}q]$ for any $q > 2$.*

1. *Then $\hat{\sigma}_n \to^p \sigma_*$, $b_n^{1/2}(\hat{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \Gamma^{-1/2}\mathcal{N}$ and $E[Y'\mathsf{f}(b_n^{1/2}(\hat{\sigma}_n - \sigma_*))] \to E[Y'\mathsf{f}(\Gamma^{-1/2}\mathcal{N})]$ as $n \to \infty$ for any continuous function f of at most polynomial growth and any bounded random variable Y' on $(\Omega, \mathcal{F})$.*
2. *Assume that $0 < \inf_\sigma \pi(\sigma) \le \sup_\sigma \pi(\sigma) < \infty$. Then $\tilde{\sigma}_n \to^p \sigma_*$, $b_n^{1/2}(\tilde{\sigma}_n - \sigma_*) \to^{s\text{-}\mathcal{L}} \Gamma^{-1/2}\mathcal{N}$ and $E[Y'\mathsf{f}(b_n^{1/2}(\tilde{\sigma}_n - \sigma_*))] \to E[Y'\mathsf{f}(\Gamma^{-1/2}\mathcal{N})]$ as $n \to \infty$ for any continuous function f of at most polynomial growth and any bounded random variable Y' on $(\Omega, \mathcal{F})$.*

Example 1.1. *We consider a simple model with deterministic diffusion coefficients:*

$$\begin{cases} dY^1_t &= \sigma_1 dW^1_t \\ dY^2_t &= \sigma_3 dW^1_t + \sigma_2 dW^2_t \\ (Y^1_0, Y^2_0) &= (0,0) \end{cases}$$

where $\epsilon > 0$, $R' > \epsilon$ and $\sigma = (\sigma_1, \sigma_2, \sigma_3) \in \Lambda = (\epsilon, R') \times (\epsilon, R') \times (-R', R')$. Let $\{\bar{N}^1_t\}, \{\bar{N}^2_t\}$ be two independent homogeneous Poisson processes with positive intensities λ_1, λ_2 respectively and point processes $\{N^1_t\}, \{N^2_t\}$ which generate observations satisfy $N^{i,n}_t = \bar{N}^i_{nt}$ $(i = 1, 2)$.

Then we can easily check $[A1], [B1\text{-}q], [B2\text{-}q]$ hold for any $q > 2$. Since $(x + y)^2 \ge x^2/2 - y^2$ for $x, y \in \mathbb{R}$, we have

$$\begin{aligned} |bb^\star(\sigma) - bb^\star(\bar{\sigma})|^2 &\ge (\sigma_1^2 - \bar{\sigma}_1^2)^2 + \frac{\epsilon^2}{(R')^2}(\sigma_1\sigma_3 - \bar{\sigma}_1\bar{\sigma}_3)^2 + \frac{\epsilon^4}{16(R')^4}(\sigma_2^2 + \sigma_3^2 - \bar{\sigma}_2^2 - \bar{\sigma}_3^2)^2 \\ &\ge 4\epsilon^2(\sigma_1 - \bar{\sigma}_1)^2 + \frac{\epsilon^2}{(R')^2}\left\{\frac{\bar{\sigma}_1^2(\sigma_3 - \bar{\sigma}_3)^2}{2} - \sigma_3^2(\sigma_1 - \bar{\sigma}_1)^2\right\} \\ &\quad + \frac{\epsilon^4}{16(R')^4}\left\{\frac{(\sigma_2^2 - \bar{\sigma}_2^2)^2}{2} - (\sigma_3^2 - \bar{\sigma}_3^2)^2\right\} \ge \frac{\epsilon^6}{8(R')^4}|\sigma - \bar{\sigma}|^2 \end{aligned}$$

for $\sigma, \bar{\sigma} \in \Lambda$. Then by Remark 1.3, we obtain $[H0]$.

H_n can be written as

$$\begin{aligned} H_n(\sigma) &= -\frac{1}{2}\left(\left(\frac{Y^1(I^i)}{\sqrt{|I^i|}}\right)^\star_i, \left(\frac{Y^2(J^j)}{\sqrt{|J^j|}}\right)^\star_j\right)\begin{pmatrix} \sigma_1^2\mathcal{E}_{l_n} & \sigma_1\sigma_3 G \\ \sigma_1\sigma_3 G^\star & (\sigma_2^2 + \sigma_3^2)\mathcal{E}_{m_n}\end{pmatrix}^{-1} \\ &\quad \times\left(\left(\frac{Y^1(I^i)}{\sqrt{|I^i|}}\right)^\star_i, \left(\frac{Y^2(J^j)}{\sqrt{|J^j|}}\right)^\star_j\right)^\star - \frac{1}{2}\log\det\begin{pmatrix} \sigma_1^2\mathcal{E}_{l_n} & \sigma_1\sigma_3 G \\ \sigma_1\sigma_3 G^\star & (\sigma_2^2 + \sigma_3^2)\mathcal{E}_{m_n}\end{pmatrix}. \end{aligned}$$

By calculating σ which maximizes H_n, we obtain the quasi-maximum likelihood estimator $\hat{\sigma}_n = (\hat{\sigma}_{1,n}, \hat{\sigma}_{2,n}, \hat{\sigma}_{3,n})$. By Corollary 1.1, we have $\hat{\sigma}_n \to^p \sigma_*$, $\sqrt{n}(\hat{\sigma}_n - \sigma_*) \to^d N(0, \Gamma^{-1})$ as $n \to \infty$, where $\sigma_* = (\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*})$ is the true value. In this case, $\rho = \rho_t(\sigma_*)$ can be written as $\rho = \sigma_{3,*}/\sqrt{\sigma_{2,*}^2 + \sigma_{3,*}^2}$, $\{a_p\}_{p=0}^\infty, \{c_p\}_{p=0}^\infty$ in $[A3']$ become constants and a, c in $[A3]$ can be written as $a(\rho') - a_0 = c(\rho') - c_0 = \mathcal{A}(\rho') = \sum_{p=1}^\infty a_p \times (\rho')^{2p}$ for $\rho' \in (-1, 1)$. If $\rho \ne 0$, Γ and Γ^{-1} can be calculated by using Proposition 1.10 later as

$$\Gamma = T\begin{pmatrix} \frac{a_0 + a}{\sigma_{1,*}^2} & 0 & -\frac{\mathcal{A}}{\sigma_{1,*}\sigma_{3,*}} \\ 0 & \frac{2c(1-\rho^2)^2 + \partial_\rho\mathcal{A}\rho(1-\rho^2)^2}{\sigma_{2,*}^2} & \frac{2c\rho^2(1-\rho^2) - \partial_\rho\mathcal{A}\rho(1-\rho^2)^2}{\sigma_{2,*}\sigma_{3,*}} \\ -\frac{\mathcal{A}}{\sigma_{1,*}\sigma_{3,*}} & \frac{2c\rho^2(1-\rho^2) - \partial_\rho\mathcal{A}\rho(1-\rho^2)^2}{\sigma_{2,*}\sigma_{3,*}} & \frac{-\mathcal{A} + 2c\rho^4 + \partial_\rho\mathcal{A}\rho(1-\rho^2)^2}{\sigma_{3,*}^2}\end{pmatrix},$$

Table 1.1: Sample means of estimators for 10,000 independent simulated sample paths. $T = 1, (\lambda_1, \lambda_2) = (1, 1)$. The left table represents the result for $(\sigma_1, \sigma_2, \sigma_3) = (1, 1, 0.5)$ and the right table represents the result for $(\sigma_1, \sigma_2, \sigma_3) = (0.5, 2, 1)$. Sample standard deviations are given in parentheses.

n		50	100	500
	true value			
$\hat{\sigma}_{1,n}$	1	0.994 (0.102)	0.998 (0.070)	0.999 (0.031)
$\hat{\sigma}_{2,n}$	1	0.968 (0.129)	0.983 (0.091)	0.996 (0.040)
$\hat{\sigma}_{3,n}$	0.5	0.499 (0.224)	0.502 (0.154)	0.5 (0.067)
$\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$	0.5	0.5 (0.238)	0.503 (0.165)	0.5 (0.071)
HY_n	0.5	0.501 (0.336)	0.504 (0.236)	0.5 (0.106)
$\sqrt{v/n}$		0.228	0.161	0.072
$\sqrt{v_0/n}$		0.339	0.239	0.107

n		50	100	500
	true value			
$\hat{\sigma}_{1,n}$	0.5	0.497 (0.050)	0.499 (0.035)	0.499 (0.015)
$\hat{\sigma}_{2,n}$	2	1.936 (0.259)	1.968 (0.181)	1.995 (0.079)
$\hat{\sigma}_{3,n}$	1	0.986 (0.449)	0.996 (0.307)	0.997 (0.135)
$\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$	0.5	0.495 (0.239)	0.499 (0.164)	0.498 (0.072)
HY_n	0.5	0.498 (0.335)	0.499 (0.237)	0.498 (0.108)
$\sqrt{v/n}$		0.228	0.161	0.072
$\sqrt{v_0/n}$		0.339	0.239	0.107

$$\Gamma^{-1} = \frac{1}{T\{4ac\mathcal{A} + 2\partial_\rho \mathcal{A}\rho(a_0 c + c_0 a)\}}\mathrm{diag}(\{\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*}\})\mathcal{P}\mathrm{diag}(\{\sigma_{1,*}, \sigma_{2,*}, \sigma_{3,*}\}),$$

where

$$\mathcal{P} = \begin{pmatrix} -2c\mathcal{A} + (c_0 + c)\partial_\rho \mathcal{A}\rho & \mathcal{A}\{-\frac{2c\rho^2}{1-\rho^2} + \partial_\rho \mathcal{A}\rho\} & \mathcal{A}(2c + \partial_\rho \mathcal{A}\rho) \\ \mathcal{A}\{-\frac{2c\rho^2}{1-\rho^2} + \partial_\rho \mathcal{A}\rho\} & \frac{-2a\mathcal{A} + (a_0+a)\{2c\rho^4 + \partial_\rho \mathcal{A}\rho(1-\rho^2)^2\}}{(1-\rho^2)^2} & (a_0 + a)\{-\frac{2c\rho^2}{1-\rho^2} + \partial_\rho \mathcal{A}\rho\} \\ \mathcal{A}(2c + \partial_\rho \mathcal{A}\rho) & (a_0 + a)\{-\frac{2c\rho^2}{1-\rho^2} + \partial_\rho \mathcal{A}\rho\} & (a_0 + a)(2c + \partial_\rho \mathcal{A}\rho) \end{pmatrix}$$

and $\rho' = \rho$ *is substituted for* $a, c, \mathcal{A}, \partial_\rho \mathcal{A}$.

We can see the estimator $\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$ *for the cross variation* $\langle Y^1, Y^2\rangle_T = \sigma_{1,*}\sigma_{3,*}T$ *also has consistency and*

$$\sqrt{n}(\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T - \langle Y^1, Y^2\rangle_T) \to^d N(0, v)$$

as $n \to \infty$ *by using the delta method, where*

$$v = T\sigma_{1,*}^2\sigma_{3,*}^2 \frac{2a(\rho)c(\rho) + \partial_\rho \mathcal{A}(\rho)\rho(a(\rho) + c(\rho))}{-2a(\rho)c(\rho)\mathcal{A}(\rho) + \partial_\rho \mathcal{A}(\rho)\rho(a_0 c(\rho) + c_0 a(\rho))}.$$

By using the result in Hayashi and Yoshida [18], we can calculate the asymptotic variance of estimation error of the Hayashi-Yoshida estimator HY_n*. In the settings in this example, we obtain*

$$\sqrt{n}(\mathrm{HY}_n - \langle Y^1, Y^2\rangle_T) \to^d N(0, v_0)$$

as $n \to \infty$*, where*

$$v_0 = T\sigma_{1,*}^2\sigma_{3,*}^2\left\{(1 + \rho^{-2})\left(\frac{2}{\lambda_1} + \frac{2}{\lambda_2}\right) - \frac{2}{\lambda_1 + \lambda_2}\right\}.$$

We also simulate $\hat{\sigma}_n, \hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T, \mathrm{HY}_n$ *for some values of parameters. Table 1.1 represents the results. We can see that each estimators work well and sample standard deviation of* $\hat{\sigma}_{1,n}\hat{\sigma}_{3,n}T$ *is about two-thirds of that of* HY_n*. The lowest two rows represent numerical calculation results of asymptotic standard deviation of estimators and we can find these values are close to sample standard deviation of estimators.*

1.7 Proofs

1.7.1 Proof of Proposition 1.2

Proof of Proposition 1.2.
[*A*3′] ⇒ [*A*3]:

Since $\nu_n^{p,1}([0,t)) \le b_n^{-1} l_n$, $\{b_n^{-1}l_n\}_{n\in\mathbb{N}}$ is tight, $\int_0^t a_p(s)ds \le \int_0^t a_0(s)ds$ and $\nu_n^{p,1}([0,t))$ converges to $\int_0^t a_p(s)ds$ in probability by $[A3']$ for $p \in \mathbb{Z}_+$, we have $\sum_{p=0}^\infty z^{2p}\int_0^t a_p(s)ds < \infty$ and

$$b_n^{-1}A_n(\mathcal{E}^1(t), \mathbf{0}, \mathbf{0}, z\mathbf{1}) = \sum_{p=0}^\infty z^{2p}\nu_n^{p,1}([0,t)) \to^p \sum_{p=0}^\infty z^{2p}\int_0^t a_p(s)ds$$

for any $t \in (0,T]$ and $z \in \mathbb{C}, |z| < 1$ by Lemma 1.3.

$[A3] \Rightarrow [A3']$:

Fix $t \in (0,T]$. Let $\{f_n\}$ be functions on $\{z \in \mathbb{C}; |z| < 1\}$ satisfying

$$f_n(z) = b_n^{-1}A_n(\mathcal{E}^1(t), \mathbf{0}, \mathbf{0}, z\mathbf{1}) = \sum_{p=0}^\infty z^{2p}\nu_n^{p,1}([0,t)).$$

Then since $\nu_n^{p,1}([0,t)) \le b_n^{-1}l_n$, the power series in the right-hand side converges absolutely on $\{|z| < 1\}$. Consequently, f_n is a holomorphic function. Then we have

$$\nu_n^{p,1}([0,t)) = \frac{1}{2\pi i}\int_{|z|=\eta/3}\frac{f_n(z)}{z^{2p+1}}dz. \tag{1.14}$$

Let $f(z) = \int_0^t a(z,s)ds$. Since

$$|f_n(z)| \le \frac{l_n}{b_n}\frac{1}{1-(2/3)^2} = \frac{9l_n}{5b_n} \tag{1.15}$$

on $\{|z| \le 2/3\}$, $\{\sup_{|z|\le 2/3}|f_n(z)|\}$ are tight. Moreover, since $f_n(z) \to^p f(z)$ $(n \to \infty)$ for $z \in \mathbb{C}, |z| < \eta$, $\sup_{|z|\le\eta/2}|f(z)| \le (9/5)\int_0^T a_0(t)dt < \infty$, almost surely. Therefore $\{\sup_{|z|\le\eta/2}|f_n(z) - f(z)|\}$ are also tight.

Let $\Gamma : |z| = 2/3$. For any $z_1, z_2 \in \{|z| < 1/2\}$, the Cauchy integral formula gives

$$\begin{aligned} 2\pi|f_n(z_1) - f_n(z_2)| &= \left|\int_\Gamma\left(\frac{f_n(\xi)}{\xi - z_1} - \frac{f_n(\xi)}{\xi - z_2}\right)d\xi\right| = \left|\int_\Gamma\frac{f_n(\xi)(z_1 - z_2)}{(\xi - z_1)(\xi - z_2)}d\xi\right| \\ &\le |z_1 - z_2|\cdot 2\pi\cdot\frac{2}{3}\cdot 6^2\cdot\sup_{|z|\le 2/3}|f_n(z)| \le Cb_n^{-1}l_n|z_1 - z_2|. \end{aligned}$$

By the convergence $f_n(z) \to^p f(z)$ $(n \to \infty)$ for $z \in \mathbb{C}, |z| < \eta$, we obtain

$$|f(z_1) - f(z_2)| \quad\le\quad C|z_1 - z_2|\int_0^T a_0(s)ds \quad \text{a.s.}$$

for $z_1, z_2 \in \{z; |z| \le \eta/2\}$. Then for any $\epsilon > 0$, tightness of $\{b_n^{-1}l_n\}$ gives

$$\lim_{\eta'\to 0}\sup_n P\left[\sup_{z_1,z_2\in\{|z|\le\eta/2\},|z_1-z_2|<\eta'}|(f_n - f)(z_1) - (f_n - f)(z_2)| > \epsilon\right] = 0.$$

Then by the tightness of $\{\sup_{|z|\le\eta/2}|f_n(z) - f(z)|\}_n$ and tightness criterion in C space in Billingsley [8] which can be extended to the one in $C(\{|z| \le \eta/2\})$, $\{f_n - f\}_n$ is tight in $C(\{|z| \le \eta/2\})$. Therefore, since $f_n(z) \to^p f(z)$ as $n \to \infty$, we see that $\{f_n - f\}$ converge in probability to 0 in $C(\{|z| \le \eta/2\})$. Therefore by (1.14), we have

$$\nu_n^{p,1}([0,t)) \to^p \frac{1}{2\pi i}\int_{|z|=\eta/3}\frac{f(z)}{z^{2p+1}}dz$$

as $n \to \infty$ for $p \ge 1$.

By the equation $f(z) = \int_0^t a(z,s)ds$ and Fubini's theorem, there exists $a_p(s)$ such that $\int_0^T a_p(s)ds < \infty$ and $\nu_n^{p,1}([0,t)) \to^p \int_0^t a_p(s)ds$ as $n \to \infty$. We thus get $[A3']$.

Moreover, under $[A2]$ and $[A3]$, the above proof gives the relations between a, c and $\{a_p\}, \{c_p\}$ in the statement. The rest of the proof is easy since

$$\begin{aligned} & b_n^{-1} A_n(x^2\mathcal{E}^1(t), y^2\mathcal{E}^2(t), xy\rho_*\mathcal{E}^1(t)G, \rho\mathbf{1}) \\ = & \sum_{p=0}^{\infty} \rho^{2p}\{x^2\nu_n^{p,1}([0,t)) + y^2\nu_n^{p,2}([0,t)) - 2xy\rho_*\rho\nu_n^{p+1,1}([0,t))\}. \end{aligned}$$

□

1.7.2 Proof of Proposition 1.3

To prove Proposition 1.3, we use some Lemmas.

Lemma 1.9. *Let $q \in \mathbb{N}$, $M \in \mathbb{N} \cup \{\infty\}$, $(\Omega', \mathcal{F}', P')$ be a probability space, $\{F_j\}_{j=1}^M$ be random variables, and $\mathcal{G}$ be a sub σ-field of $\mathcal{F}'$.*

1. *Then $E'\big[|\sum_{j=1}^M F_j|^q|\mathcal{G}\big] \leq \big(\sum_{j=1}^M E'[|F_j|^q|\mathcal{G}]^{\frac{1}{q}}\big)^q$, where E' denotes expectation with respect to P'.*
2. *We assume $q \in 2\mathbb{N}$ and $\{\sum_{j=1}^k F_j\}_{0\leq k\leq M}$ is martingale with respect to some filtration. Then there exists a constant $C_q > 0$ which depends only q, such that $E'\big[|\sum_{j=1}^M F_j|^q\big] \leq C_q\big(\sum_{j=1}^M E'[|F_j|^q]^{\frac{2}{q}}\big)^{\frac{q}{2}}$.*

Proof. We expand the summation and use Hölder's inequality.

$$\begin{aligned} E'\Big[\Big|\sum_{j=1}^M F_j\Big|^q\Big|\mathcal{G}\Big] &\leq \sum_{j_1,\dots,j_q=1}^M E'[|F_{j_1}\dots F_{j_q}||\mathcal{G}] \leq \sum_{j_1,\dots,j_q=1}^M E'[|F_{j_1}|^q|\mathcal{G}]^{\frac{1}{q}} \dots E'[|F_{j_q}|^q|\mathcal{G}]^{\frac{1}{q}} \\ &\leq \Big(\sum_{j=1}^M E'[|F_j|^q|\mathcal{G}]^{\frac{1}{q}}\Big)^q. \end{aligned}$$

For 2., we use the Burkholder-Davis-Gundy inequality and apply 1. for $\mathcal{G} = \{\emptyset, \Omega'\}$. □

Lemma 1.10. *Let $\{G_p\}_{p\in\mathbb{Z}_+}$ be a sequence of positive numbers, $a \in \mathbb{N}, b, r, s \in \mathbb{Z}_+$ and $\rho \in [0,1)$. Then there exists a constant $C > 0$ which depends only on a, b, r, s such that*

$$\sum_{p=0}^{\infty} \rho^{a(p-b)\vee 0}(p+1)^s G_p \leq C(1-\rho)^{-(s+\frac{r+1}{2})}\Big(\sum_{p=0}^{\infty}\frac{G_p^2}{(p+1)^r}\Big)^{\frac{1}{2}}.$$

Proof. By the Cauchy-Schwarz inequality, we have

$$\sum_{p=0}^{\infty} \rho^{a(p-b)\vee 0}(p+1)^s G_p \leq \Big(\sum_{p=0}^{\infty}\rho^{2a(p-b)\vee 0}(p+1)^{2s+r}\Big)^{\frac{1}{2}}\Big(\sum_{p=0}^{\infty}\frac{G_p^2}{(p+1)^r}\Big)^{\frac{1}{2}}.$$

Then the conclusion follows since

$$\sum_{p=0}^{\infty}\rho^{2a(p-b)\vee 0}(p+1)^{2s+r} \leq C + C\sum_{p=0}^{\infty}\rho^{2ap}(p+1)^{2s+r} \leq C + C\frac{(2s+r)!}{(1-\rho^{2a})^{2s+r+1}}.$$

□

Lemma 1.11. *Let $(\Omega', \mathcal{F}', P')$ be probability space and $\{\mathcal{G}_n\}_{n\in\mathbb{N}} \subset \mathcal{F}'$ be sub σ-fields.*

1. *Let $\{X_n'\}_{n\in\mathbb{N}} \subset L^1(\Omega')$. Assume $E'[|X_n'||\mathcal{G}_n] \to^p 0$ $(n\to\infty)$. Then $X_n' \to^p 0$ $(n\to\infty)$.*
2. *Let $d_1, d_2 \in \mathbb{N}$, $p > d_1$, $\Lambda' \subset \mathbb{R}^{d_1}$ be a bounded open set and $X_n' : \Omega' \to C^1(\Lambda'; \mathbb{R}^{d_2})$ be random variables $(n \in \mathbb{N})$. Assume that Λ' satisfies Sobolev's inequality, $\{\sup_{\sigma\in\Lambda'}|X_n'(\sigma)|^p \vee |\partial_\sigma X_n'(\sigma)|^p\}_{n\in\mathbb{N}} \subset L^1(\Omega')$ and $\sup_{\sigma\in\mathbb{Q}^{d_1}\cap\Lambda'} E'[|\partial_\sigma X_n'(\sigma)|^p \vee |X_n'(\sigma)|^p|\mathcal{G}_n] \to^p 0$ as $n\to\infty$. Then $\sup_{\sigma\in\Lambda'}|X_n'(\sigma)| \to^p 0$ as $n\to\infty$.*

Proof. 1. For any $\epsilon, \delta > 0$, there exists $N \in \mathbb{N}$ such that $P'[E'[|X'_n||\mathcal{G}_n] \geq \epsilon\delta/2] < \epsilon/2$ for $n \geq N$. Therefore, for $n \geq N$, we have

$$\begin{aligned} P'[|X'_n| \geq \delta] &\leq P'[E'[|X'_n||\mathcal{G}_n] \geq \epsilon\delta/2] + P'[|X'_n| \geq \delta, E'[|X'_n||\mathcal{G}_n] < \epsilon\delta/2] \\ &\leq \frac{\epsilon}{2} + \frac{1}{\delta}E'[|X'_n|, E'[|X'_n||\mathcal{G}_n] < \epsilon\delta/2] \\ &= \frac{\epsilon}{2} + \frac{1}{\delta}E'[E'[|X'_n||\mathcal{G}_n], E'[|X'_n||\mathcal{G}_n] < \epsilon\delta/2] \leq \epsilon. \end{aligned}$$

2. First, by Sobolev's inequality, we have

$$E'\left[\left(\sup_{\sigma\in\Lambda'}|X'_n|\right)^p\middle|\mathcal{G}_n\right] \leq CE'\left[\int_{\Lambda'}|\partial_\sigma X'_n|^p d\sigma\middle|\mathcal{G}_n\right] + CE'\left[\int_{\Lambda'}|X'_n|^p d\sigma\middle|\mathcal{G}_n\right].$$

Moreover, for $v = 0, 1$ and $A \in \mathcal{G}_n$, it follows that

$$\begin{aligned} &E'\left[\int_{\Lambda'}|\partial_\sigma^v X'_n|^p d\sigma, A\right] = \int_{\Lambda'}E'[E'[|\partial_\sigma^v X'_n|^p|\mathcal{G}_n], A]d\sigma \\ \leq &\int_{\Lambda'}E'\left[\sup_{\sigma\in\mathbb{Q}^{d_1}\cap\Lambda'}E'[|\partial_\sigma^v X'_n|^p|\mathcal{G}_n], A\right]d\sigma \leq E'\left[\sup_{\sigma\in\mathbb{Q}^{d_1}\cap\Lambda'}E'[|\partial_\sigma^v X'_n|^p|\mathcal{G}_n]\cdot|\Lambda'|, A\right], \end{aligned}$$

where $|\Lambda'|$ denotes volume of Λ'. Since $A \in \mathcal{G}_n$ is arbitrary, we have

$$E'\left[\int_{\Lambda'}|\partial_\sigma^v X'_n|^p d\sigma\middle|\mathcal{G}_n\right] \leq \sup_{\sigma\in\mathbb{Q}^{d_1}\cap\Lambda'}E'[|\partial_\sigma^v X'_n|^p|\mathcal{G}_n]\cdot|\Lambda'| \quad \text{a.s.}$$

Therefore we obtain

$$E'\left[\left(\sup_\sigma|X'_n|\right)^p\middle|\mathcal{G}_n\right] \leq C|\Lambda'|\sum_{v=0}^1\sup_{\sigma\in\mathbb{Q}^{d_1}\cap\Lambda'}E'[|\partial_\sigma^v X'_n|^p|\mathcal{G}_n] \to^p 0$$

as $n \to \infty$. Then the proof is completed by 1. □

Lemma 1.12. *Let $(\Omega', \mathcal{F}', P')$ be a probability space, $T' > 0$, $q \in 2\mathbb{N}$, $\{\mathcal{F}'_t\}_{0\leq t\leq T'}$ be a filtration, $M \in \mathbb{N}$, $\{K^i\}_{i=1}^M$ be a deterministic partition of $[0, T']$ where $L(K^i) < L(K^j)$ for $i < j$. Let $(\tilde{W}_t^l, \mathcal{F}'_t)_{0\leq t\leq T'}$ be standard Brownian motions ($l = 1, 2, 3$), and $F_{i,j,k}$ be $\mathcal{F}'_{L(K^i)\wedge L(K^j)\wedge L(K^k)}$-measurable random variables. Assume $\langle\tilde{W}^{p_1}, \tilde{W}^{p_2}\rangle$ are deterministic for $1 \leq p_1 < p_2 \leq 3$. Then for $\Delta\tilde{W}_i^l = \tilde{W}^l(K^i)$, $F_{i,j}^1 = F_{i,j,j}, F_{i,j}^2 = F_{j,i,j}$ and $F_{i,j}^3 = F_{j,j,i}$, there exists a constant $C_q > 0$ which depends only on q such that*

$$\begin{aligned} E'\left[\left|\sum_{i,j,k}\Delta\tilde{W}_i^1\Delta\tilde{W}_j^2\Delta\tilde{W}_k^3 F_{i,j,k}\right|^q\right] &\leq C_q\left(\sum_{i,j,k}|K^i||K^j||K^k|E'[|F_{i,j,k}|^q]^{\frac{2}{q}}\right)^{\frac{q}{2}} \\ &+C_q\left(\sum_i|K^i|\left(\sum_{j\neq i}|K^j|\sum_{l=1}^3 E'[|F_{i,j}^l|^q]^{\frac{1}{q}}\right)^2\right)^{\frac{q}{2}}. \end{aligned}$$

Proof. In this proof, general constants denoted by C depend only on q.

Let us denote

$$\mathcal{H}_{i,j,k} = \Delta\tilde{W}_i^1\Delta\tilde{W}_j^2\Delta\tilde{W}_k^3 F_{i,j,k}, \quad \mathcal{H}_{i,j}^2 = \mathcal{H}_{i,j,j} + \mathcal{H}_{j,i,j} + \mathcal{H}_{j,j,i},$$
$$\mathcal{H}_{i,j,k}^3 = \mathcal{H}_{i,j,k} + \mathcal{H}_{i,k,j} + \mathcal{H}_{j,i,k} + \mathcal{H}_{j,k,i} + \mathcal{H}_{k,i,j} + \mathcal{H}_{k,j,i},$$

then it follows that

$$\sum_{i,j,k}\mathcal{H}_{i,j,k} = \sum_i\mathcal{H}_{i,i,i} + \sum_i\sum_{j<i}(\mathcal{H}_{i,j}^2 + \mathcal{H}_{j,i}^2) + \sum_i\sum_{j<i}\sum_{k<j}\mathcal{H}_{i,j,k}^3.$$

Since $\langle \tilde{W}^{p_1}, \tilde{W}^{p_2}\rangle$ is deterministic for $1\leq p_1<p_2\leq 3$, Itô's formula yields

$$E'[\mathcal{H}_{i,i,i}|\mathcal{F}_{L(K^i)}]=E'[\mathcal{H}^2_{i,j}|\mathcal{F}_{L(K^i)}]=E'[\mathcal{H}^3_{i,j,k}|\mathcal{F}_{L(K^i)}]=0$$

for $k<j<i$. Therefore by Lemma 1.9 we have

$$\begin{aligned}E'\left[\left|\sum_{i,j,k}\mathcal{H}_{i,j,k}\right|^q\right] \quad \leq \quad & C\bigg\{\sum_i E'[|\mathcal{H}_{i,i,i}|^q]^{\frac{2}{q}}+\sum_i E'\left[\left|\sum_{j<i}(\mathcal{H}^2_{i,j}+\mathcal{H}^2_{j,i}-E'[\mathcal{H}^2_{j,i}|\mathcal{F}'_{L(K^i)}])\right|^q\right]^{\frac{2}{q}} \\ &+\sum_i E'\left[\left|\sum_{j<i}\sum_{k<j}\mathcal{H}^3_{i,j,k}\right|^q\right]^{\frac{2}{q}}\bigg\}^{\frac{q}{2}}+CE'\left[\left|\sum_i\sum_{j<i}E'[\mathcal{H}^2_{j,i}|\mathcal{F}'_{L(K^i)}]\right|^q\right]. \qquad (1.16)\end{aligned}$$

We will estimate each term of the right-hand side of (1.16). First,

$$\sum_i E'[|\mathcal{H}_{i,i,i}|^q]^{\frac{2}{q}} \quad = \quad \sum_i E'[|F_{i,i,i}|^q E'[(\Delta\tilde{W}^1_i\Delta\tilde{W}^2_i\Delta\tilde{W}^3_i)^q|\mathcal{F}'_{L(K^i)}]]^{\frac{2}{q}}\leq C\sum_i|K^i|^3E'[|F_{i,i,i}|^q]^{\frac{2}{q}}. \quad (1.17)$$

Let

$$\mathcal{H}^{2,1}_{i,j}=\Delta\tilde{W}^2_j\Delta\tilde{W}^3_jF_{i,j,j},\quad \mathcal{H}^{2,2}_{i,j}=\Delta\tilde{W}^1_j\Delta\tilde{W}^3_jF_{j,i,j},\quad \mathcal{H}^{2,3}_{i,j}=\Delta\tilde{W}^1_j\Delta\tilde{W}^2_jF_{j,j,i}.$$

Since

$$E'\left[\left(\sum_{j<i}\mathcal{H}^{2,l}_{i,j}\right)^q\right]^{\frac{2}{q}}\leq C\sum_{j<i}E'\left[(\mathcal{H}^{2,l}_{i,j}-E'[\mathcal{H}^{2,l}_{i,j}|\mathcal{F}'_{L(K^j)}])^q\right]^{\frac{2}{q}}+C\left(\sum_{j<i}E'\left[E'[\mathcal{H}^{2,l}_{i,j}|\mathcal{F}'_{L(K^j)}]^q\right]^{\frac{1}{q}}\right)^2$$

for each i,l by Lemma 1.9, we obtain

$$\begin{aligned}&\sum_i E'\left[\left(\sum_{j<i}(\mathcal{H}^2_{i,j}+\mathcal{H}^2_{j,i}-E'[\mathcal{H}^2_{j,i}|\mathcal{F}'_{L(K^i)}])\right)^q\right]^{\frac{2}{q}} \\ \leq\quad & C\sum_i|K^i|E'\left[\sum_{l=1}^3\left(\sum_{j<i}\mathcal{H}^{2,l}_{i,j}\right)^q\right]^{\frac{2}{q}}+C\sum_i|K^i|^2E'\left[\sum_{l=1}^3\left(\sum_{j<i}\Delta\tilde{W}^l_jF^l_{j,i}\right)^q\right]^{\frac{2}{q}} \\ \leq\quad & C\sum_i|K^i|\sum_{j<i}|K^j|^2\sum_{l=1}^3E'[(F^l_{i,j})^q]^{\frac{2}{q}}+C\sum_i|K^i|\sum_{l=1}^3\left(\sum_{j<i}|K^j|E'[(F^l_{i,j})^q]^{\frac{1}{q}}\right)^2 \\ &+C\sum_i|K^i|^2\sum_{j<i}|K^j|\sum_{l=1}^3E'[(F^l_{j,i})^q]^{\frac{2}{q}}. \qquad (1.18)\end{aligned}$$

Moreover, let

$$\begin{aligned}\mathcal{H}^{3,1}_{i,j,k} &= \Delta\tilde{W}^2_j\Delta\tilde{W}^3_kF_{i,j,k}+\Delta\tilde{W}^2_k\Delta\tilde{W}^3_jF_{i,k,j}, \\ \mathcal{H}^{3,2}_{i,j,k} &= \Delta\tilde{W}^1_j\Delta\tilde{W}^3_kF_{j,i,k}+\Delta\tilde{W}^1_k\Delta\tilde{W}^3_jF_{k,i,j}, \\ \mathcal{H}^{3,3}_{i,j,k} &= \Delta\tilde{W}^1_j\Delta\tilde{W}^2_kF_{j,k,i}+\Delta\tilde{W}^1_k\Delta\tilde{W}^2_jF_{k,j,i},\end{aligned}$$

then by Lemma 1.9 we have

$$\begin{aligned}\sum_i E'\left[\left(\sum_{j<i}\sum_{k<j}\mathcal{H}^3_{i,j,k}\right)^q\right]^{\frac{2}{q}} \quad\leq\quad & C\sum_i|K^i|\sum_{l=1}^3E'\left[\left(\sum_{j<i}\sum_{k<j}\mathcal{H}^{3,l}_{i,j,k}\right)^q\right]^{\frac{2}{q}}\leq C\sum_i|K^i|\sum_{l=1}^3\sum_{j<i}E'\left[\left(\sum_{k<j}\mathcal{H}^{3,l}_{i,j,k}\right)^q\right]^{\frac{2}{q}} \\ \leq\quad & C\sum_i|K^i|\sum_{j<i}|K^j|\sum_{k<j}|K^k|\Big(E'[(F_{i,j,k})^q]^{\frac{2}{q}}+E'[(F_{i,k,j})^q]^{\frac{2}{q}} \\ &+E'[(F_{j,i,k})^q]^{\frac{2}{q}}+E'[(F_{j,k,i})^q]^{\frac{2}{q}}+E'[(F_{k,i,j})^q]^{\frac{2}{q}}+E'[(F_{k,j,i})^q]^{\frac{2}{q}}\Big). \qquad (1.19)\end{aligned}$$

Furthermore, let $g_1(K^i)=\langle\tilde{W}^2,\tilde{W}^3\rangle(K^i)$, $g_2(K^i)=\langle\tilde{W}^1,\tilde{W}^3\rangle(K^i)$, $g_3(K^i)=\langle\tilde{W}^1,\tilde{W}^2\rangle(K^i)$, then we obtain

$$E'\bigg[\bigg|\sum_i\sum_{j<i}E'[\mathcal{H}^2_{j,i}|\mathcal{F}'_{L(K^i)}]\bigg|^q\bigg]=E'\bigg[\bigg|\sum_i\sum_{j<i}\sum_{l=1}^3 g_l(K^i)\Delta\tilde{W}^l_jF^l_{j,i}\bigg|^q\bigg]\leq 3^q\sum_{l=1}^3E'\bigg[\bigg|\sum_j\bigg(\sum_{i>j}g(K^i)F^l_{j,i}\bigg)\Delta\tilde{W}^l_j\bigg|^q\bigg].$$

Hence Lemma 1.9 yields

$$\begin{aligned}E'\bigg[\bigg|\sum_i\sum_{j<i}E'[\mathcal{H}^2_{j,i}|\mathcal{F}'_{L(K^i)}]\bigg|^q\bigg] &\leq C\sum_{l=1}^3\bigg(\sum_j|K^j|E'\bigg[\bigg(\sum_{i>j}g_l(K^i)F^l_{j,i}\bigg)^q\bigg]^{\frac{2}{q}}\bigg)^{\frac{q}{2}}\\ &\leq C\bigg(\sum_j|K^j|\bigg(\sum_{i>j}|K^i|\sum_{l=1}^3E'[(F^l_{j,i})^q]^{\frac{1}{q}}\bigg)^2\bigg)^{\frac{q}{2}}. \qquad (1.20)\end{aligned}$$

By (1.16)-(1.20), we obtain the conclusion. □

For $p\geq 0$, we denote

$$L_p = \{\rho_{L(\theta_{[p/2],i}\cup\theta_{[p/2],j+l_n})\wedge\tau_n}G_{I^i,J^j}\}_{i,j},\quad \tilde{L}_p=\begin{pmatrix}0 & L_p\\ L_p^\star & 0\end{pmatrix},\quad \tilde{M}=\sum_{p=0}^\infty(-1)^p\tilde{L}_p^p,$$

$$\tilde{M}_p=\begin{pmatrix}(GG^\star)^p & (GG^\star)^pG\\ (G^\star G)^pG^\star & (G^\star G)^p\end{pmatrix},\quad \hat{Z}_{k,t}=\begin{cases}\int_{I^k_t}b^1_{s,*}dW_s/(|b^1_{I^k,\tau_n}|\sqrt{|I^k|}) & (k\leq l_n)\\ \int_{J^{k-l_n}_t}b^2_{s,*}dW_s/\Big(|b^2_{J^{k-l_n},\tau_n}|\sqrt{|J^{k-l_n}|}\Big) & (k>l_n)\end{cases}$$

and $D_t=\mathrm{diag}(\{|b^1_{I,\tau_n}|\}_{I\cap[0,t)\neq\emptyset},\{|b^2_{J,\tau_n}|\}_{J\cap[0,t)\neq\emptyset})$. Though $\{\hat{Z}_{k,t}\}_t$ is not necessarily a local martingale, we denote by $\langle\hat{Z}\rangle_t$ the quadratic variation of $\hat{Z}$ regarding Π as deterministic functions, that is, $\langle\hat{Z}\rangle_t$ be an $(l_n+m_n)\times(l_n+m_n)$ symmetric matrix with

$$(\langle\hat{Z}\rangle_t)_{k,k'}=\begin{cases}\int_{I^k_t}|b^1_{s,*}|^2ds\delta_{k,k'}/(|b^1_{I^k,\tau_n}|^2|I^k|) & (k,k'\leq l_n)\\ \int_{J^{k-l_n}_t}|b^2_{s,*}|^2ds\delta_{k,k'}/(|b^2_{J^{k-l_n},\tau_n}|^2|J^{k-l_n}|) & (k,k'>l_n)\\ \int_{I^k_t\cap J^{k'-l_n}_t}b^1_{s,*}\cdot b^2_{s,*}ds/(|b^1_{I^k,\tau_n}||b^2_{J^{k'-l_n},\tau_n}|\sqrt{|I^k||J^{k'-l_n}|} & (k\leq l_n,k'>l_n)\end{cases}$$

Moreover we define

$$\begin{aligned}\tilde{H}^1_n(t) &= \tilde{H}^1_{n,s_n}(t;\sigma)=-\frac{1}{2}\hat{Z}^\star_{\cdot,t}\tilde{M}\hat{Z}_{\cdot,t}-\log\det D_t+\frac{1}{2}\sum_{p=1}^\infty\frac{(-1)^p}{p}\sum_k(\tilde{L}^p_p)_{k,k}1_{\{\theta_{0,k}\cap[0,t)\neq\emptyset\}},\\ \tilde{H}^2_n(t) &= \tilde{H}^2_{n,s_n}(t;\sigma)=-\frac{1}{2}\mathrm{tr}(\tilde{M}\langle\hat{Z}\rangle_t)-\log\det D_t+\frac{1}{2}\sum_{p=1}^\infty\frac{(-1)^p}{p}\sum_k(\tilde{L}^p_p)_{k,k}1_{\{\theta_{0,k}\cap[0,t)\neq\emptyset\}},\end{aligned}$$

$$\tilde{H}^3_n(t)=\tilde{H}^3_{n,s_n}(t;\sigma)=b_n\sum_{p=0}^\infty\sum_{i=1}^2\int_0^t\mathcal{D}^i_p(s\wedge\tau(s_n),s;\sigma)\nu^{p,i}_n(ds),$$

$\hat{Z}_k=\hat{Z}_{k,T}$, and $\tilde{H}^i_n=\tilde{H}^i_n(T)$ $(1\leq i\leq 3)$, where

$$\mathcal{D}^i_p(s,t;\sigma)=\begin{cases}-\frac{|b^i_{t,*}|^2}{2|b^i_s|^2}-\log|b^i_s| & (p=0)\\ -\frac{|b^i_{t,*}|^2}{2|b^i_s|^2}\rho_s^{2p}+\frac{1}{2}\left(\frac{|b^1_{t,*}||b^2_{t,*}|}{|b^1_s||b^2_s|}+\frac{|b^i_{t,*}|}{|b^i_s|}-\frac{|b^{3-i}_{t,*}|}{|b^{3-i}_s|}\right)\rho_s^{2p-1}\rho_{t,*}+\frac{\rho_s^{2p}}{4p} & (p\geq 1)\end{cases}$$

for $i=1,2$. Then we have $\partial_\sigma\tilde{H}^3_{n,s_n}(t;\sigma_*)\equiv 0$ on $\{\tau(s_n)=T\}$.

Let $q\in 2\mathbb{N}$, $\gamma\in(0,1)$ be defined in $[A1]$ 5., $\Theta^1_p=\sup_{0\leq t\leq T}E[|\mu_t|^p]$, $\Theta^2_p=\sup_{0\leq s<t\leq T}E[|\mu_t-\mu_s|^p]/|t-s|^{p\gamma}$, and $\{s_n(t)\}_{0\leq t\leq T,n\in\mathbb{N}}$ be stochastic processes which satisfies $[S]$. Moreover, we define $\varphi_q(\{x_p\})=(\sum_{p=0}^\infty x_p)^q\vee(\sum_{p=0}^\infty x_p^{2q/(2q-1)})^{q-1/2}$ for $\{x_p\}_{p=0}^\infty\subset\mathbb{R}_+$ and $q\in 2\mathbb{N}$.

Lemma 1.13. *Assume [A1]. Let $r \in \mathbb{N}, r \geq 2$. Then there exists a constant $C > 0$ which depends only on q, r, n_2 and n_3 such that*

$$\begin{aligned} & E[|\partial_\sigma^v \hat{H}_n(\sigma; s_n) - \partial_\sigma^v \tilde{H}^1_{n,s_n}(T;\sigma)|^q|\Pi] \\ \leq\ & C(T^{\frac{3}{2}q} \vee 1)(E[R^C] + \Theta^1_C)E[R^C s_n(T)^{-(2v+2r+7)q}|\Pi]^{\frac{1}{2}} \\ & \times\bigg\{(T\vee 1)^{\frac{q}{2}} + ((r_n^{q(\gamma\wedge\frac{1}{2})}(l_n+m_n)^{\frac{q}{2}})\vee 1)\{(\Theta^1_{8q})^{\frac{1}{8}} + (\Theta^2_{8q})^{\frac{1}{8}}\} \\ & +\varphi_q\bigg(\bigg\{\frac{\sqrt{(l_n+m_n)\Phi_{2p+2,2}}\vee\Phi_{2p+2,1}}{(p+1)^r}\bigg\}_p\bigg) + \bigg(\sum_{p_1,p_2=0}^\infty \frac{\bar{\Phi}_{2p_1+3,2p_2+3}}{(p_1+1)^r(p_2+1)^r}\bigg)^{\frac{q}{2}}\bigg\} \end{aligned} \tag{1.21}$$

and

$$\begin{aligned} & E[|\partial_\sigma^v \tilde{H}^2_{n,s_n}(T;\sigma) - \partial_\sigma^v \tilde{H}^3_{n,s_n}(T;\sigma)|^q|\Pi] \\ \leq\ & CE[R^C(1-\bar{\rho}_T)^{-C}]\bigg\{\bigg(\sum_{p=0}^\infty \frac{\Phi_{2p+2,1}}{(p+1)^r}\bigg)^q + \bigg(\sum_{p_1,p_2=0}^\infty \frac{\bar{\Phi}_{2p_1+2,2p_2+2}}{(p_1+1)^r(p_2+1)^r}\bigg)^{\frac{q}{2}}\bigg\} \end{aligned}$$

for $0 \leq v \leq 4$ and $\sigma \in \Lambda$.

Proof. In this proof, general constants denoted by C depend only on q, r, n_2, n_3.

We first prove (1.21). Let

$$\tilde{\mu}_k = \begin{cases} \int_{I^k} \mu^1_s ds/(|b^1_{I^k,\tau_n}|\sqrt{|I^k|}) & (k \leq l_n) \\ \int_{J^{k-l_n}} \mu^2_s ds/(|b^2_{J^{k-l_n},\tau_n}|\sqrt{|J^{k-l_n}|}) & (k > l_n) \end{cases}$$

for $1 \leq k \leq l_n + m_n$. Then $\hat{H}_n(\sigma; s_n) - \tilde{H}^1_{n,s_n}(T;\sigma)$ can be decomposed as

$$\hat{H}_n(\sigma; s_n) - \tilde{H}^1_{n,s_n}(T;\sigma) = -\frac{1}{2}Z^\star(M-\tilde{M})Z - \tilde{\mu}^\star\tilde{M}(\hat{Z}+\frac{\tilde{\mu}}{2}) + \frac{1}{2}\sum_{p=1}^\infty \frac{(-1)^p}{p}(\mathrm{tr}(\tilde{L}^p) - \mathrm{tr}(\tilde{L}^p_p)).$$

We will estimate each term of the right-hand side of this equation.

First, we assume that Π is deterministic. Let $\mathcal{W}_t = (W_t, \hat{W}_t)$, then we can write

$$((\tilde{L})^{2p} - (\tilde{L})^{2p+1})_{kk'} - ((\tilde{L}_{2p})^{2p} - (\tilde{L}_{2p+1})^{2p+1})_{kk'} = \int_{\theta_{2p+2,k}} \xi^{k,k'}_{p,t}\cdot d\mathcal{W}_t + \int_{\theta_{2p+2,k}} \eta^{k,k'}_{p,t}dt,$$

by Itô's formula, where

$$\begin{aligned} |\partial_\sigma^v \xi^{k,k'}_{p,t}| &\leq 2\cdot 2\cdot 4^{v+1}(2p+1)^{v+1}\cdot n_2 R^{2v+3}(1-s_n(T))^{(2p-v-1)\vee 0}(\tilde{M}_p)_{kk'}1_{\theta_{2p+2,k}}(t), \\ |\partial_\sigma^v \eta^{k,k'}_{p,t}| &\leq 4\cdot 2\cdot\frac{1}{2}4^{v+2}(2p+1)^{v+2}\cdot n_2^2 R^{2v+6}(1-s_n(T))^{(2p-v-2)\vee 0}(\tilde{M}_p)_{kk'}1_{\theta_{2p+2,k}}(t). \end{aligned}$$

Moreover, let $\tilde{\xi}^{k,k'}_{p,t}$ be the one constructed by $\xi^{k,k'}_{p,t}$ substituting $L(\theta_{2p+2,k})\wedge\tau(s_n)$ for all times of X, $\tilde{b}^2$ and $\tilde{b}^3$, then we can write $\tilde{\xi}^{k,k'}_{p,t} = \sum_K \tilde{\xi}^{k,k'}_{p,K}1_K(t)$ for some random variables $\{\tilde{\xi}^{k,k'}_{p,K}\}_K$, where $\{K\}$ denotes the set of intervals obtained by unifying partitions $\{S^i\}$,$\{T^j\}$ and $\{\mathcal{T}^j_k\}$. Furthermore, let

$$\xi^{kk'}_{p,t} - \tilde{\xi}^{kk'}_{p,t} = \int_0^t \hat{\xi}^{k,k'}_{p,s}\cdot d\mathcal{W}_s + \int_0^t \hat{\eta}^{k,k'}_{p,s}ds, \tag{1.22}$$

then we have

$$\begin{aligned} |\partial_\sigma^v \hat{\xi}^{k,k'}_{p,s}| &\leq 2\cdot 4\cdot 4^{v+1}(2p+1)^{v+1}\cdot 4(2p+2)\cdot n_2^2 R^{2v+6}(1-s_n(T))^{(2p-v-2)\vee 0}(\tilde{M}_p)_{k,k'}1_{\theta_{2p+2,k}}(s), \\ |\partial_\sigma^v \hat{\eta}^{k,k'}_{p,s}| &\leq 2\cdot 8\cdot 4^{v+1}(2p+1)^{v+1}\cdot\frac{1}{2}4^2(2p+2)^2\cdot n_2^3 R^{2v+9}(1-s_n(T))^{(2p-v-3)\vee 0}(\tilde{M}_p)_{k,k'}1_{\theta_{2p+2,k}}(s). \end{aligned}$$

Therefore we can write

$$
\begin{aligned}
-\frac{1}{2}Z^\star(M-\tilde{M})Z &= -\frac{1}{2}\sum_{p=0}^\infty\sum_{k,k'}\left(\int\xi_{p,t}^{k,k'}\cdot d\mathcal{W}_t+\int\eta_{p,t}^{k,k'}dt\right)Z_kZ_{k'} \\
&= -\frac{1}{2}\sum_{p=0}^\infty\sum_{k,k'}\int\tilde{\xi}_{p,t}^{k,k'}\cdot d\mathcal{W}_t\tilde{Z}_{k,p}\tilde{Z}_{k',p}-\frac{1}{2}\sum_{p=0}^\infty\sum_{k,k'}\int(\xi_{p,t}^{k,k'}-\tilde{\xi}_{p,t}^{k,k'})\cdot d\mathcal{W}_tZ_kZ_{k'} \\
&\quad -\frac{1}{2}\sum_{p=0}^\infty\sum_{k,k'}\int\tilde{\xi}_{p,t}^{k,k'}\cdot d\mathcal{W}_t(Z_kZ_{k'}-\tilde{Z}_{k,p}\tilde{Z}_{k',p})-\frac{1}{2}\sum_{p=0}^\infty\sum_{k,k'}\int\eta_{p,t}^{k,k'}dtZ_kZ_{k'} \\
&= \mathcal{X}+R_1+R_2+R_3,
\end{aligned}
$$

where

$$
(\tilde{Z}_{k,p})_{k=1}^{l_n+m_n}=\left(\left(\frac{b^1_{L(\theta_{2p+2,i}),*}\cdot W(I^i)}{|b^1_{L(\theta_{2p+2,i})\wedge\tau_n}|\sqrt{|I^i|}}\right)_{i=1}^{l_n},\left(\frac{b^2_{L(\theta_{2p+2,j+l_n}),*}\cdot W(J^j)}{|b^2_{L(\theta_{2p+2,j+l_n})\wedge\tau_n}|\sqrt{|J^j|}}\right)_{j=1}^{m_n}\right).
$$

Let $k^1(K)$ denotes $k\le l_n$ which satisfies $K\subset I^k$, $k^2(K)$ denotes $k>l_n$ which satisfies $K\subset J^{k-l_n}$ and

$$
\tilde{B}_{k,p}=\begin{cases} b^1_{L(\theta_{2p+2,k}),*}/(|b^1_{L(\theta_{2p+2,k})\wedge\tau_n}|\sqrt{|I^k|}) & (k\le l_n) \\ b^2_{L(\theta_{2p+2,k}),*}/(|b^2_{L(\theta_{2p+2,k})\wedge\tau_n}|\sqrt{|J^{k-l_n}|}) & (l_n<k\le l_n+m_n). \end{cases}
$$

Then $\mathcal{X}$ can be rewritten as

$$
\begin{aligned}
\mathcal{X} &= -\frac{1}{2}\sum_{p=0}^\infty\sum_{k,k'}\sum_{K''}\tilde{\xi}_{p,K''}^{k,k'}\cdot\mathcal{W}(K'')\tilde{Z}_{k,p}\tilde{Z}_{k',p} \\
&= -\frac{1}{2}\sum_{K,K',K''}\sum_{i,j=1}^2\sum_{p=0}^\infty\tilde{\xi}_{p,K''}^{k^i(K),k^j(K')}\cdot\mathcal{W}(K'')\tilde{B}_{k^i(K),p}\cdot W(K)\tilde{B}_{k^j(K),p}\cdot W(K').
\end{aligned}
$$

Let

$$
F^{i,j,v}_{K,K',K''}=\frac{1}{2}\sum_{v_1,v_2,v_3\ge 0,v_1+v_2+v_3=v}\frac{v!}{v_1!v_2!v_3!}\sum_{p=0}^\infty|\partial_\sigma^{v_1}\tilde{\xi}_{p,K''}^{k^i(K),k^j(K')}||\partial_\sigma^{v_2}\tilde{B}_{k^i(K),p}||\partial_\sigma^{v_3}\tilde{B}_{k^j(K'),p}|
$$

for $1\le i,j\le 2$, then $F^{i,j,v}_{K,K',K''}=F^{j,i,v}_{K',K,K''}$. Hence for general Π and any $q\in 2\mathbb{N}$, we have

$$
\begin{aligned}
& E[|\partial_\sigma^v\mathcal{X}|^q|\Pi] \\
\le\ & C\sum_{i,j=1}^2\left\{\left(\sum_{K,K',K''}|K||K'||K''|E[(F^{i,j,v}_{K,K',K''})^q|\Pi]^{\frac{2}{q}}\right)^{\frac{q}{2}}\right. \\
& \left.+\left(\sum_K|K|\left(\sum_{K'}|K'|(2E[(F^{i,j,v}_{K,K',K'})^q|\Pi]^{\frac{1}{q}}+E[(F^{i,j,v}_{K',K',K})^q|\Pi]^{\frac{1}{q}}\right)^2\right)^{\frac{q}{2}}\right\} \\
\le\ & C\left(\sum_{k,k',K''}|K''|E\left[\left(\sum_{p=0}^\infty R^C(p+1)^{v+1}(1-s_n(T))^{(2p-5)\vee 0}(\tilde{M}_p)_{k,k'}1_{\theta_{2p+2,k}}(K'')\right)^q\middle|\Pi\right]^{\frac{2}{q}}\right)^{\frac{q}{2}} \\
& +C\left\{\sum_k|\theta_{0,k}|\left(\sum_{k'}|\theta_{0,k'}|\left\{E\left[\left(\sum_{p=0}^\infty R^C(p+1)^{v+1}\frac{(1-s_n(T))^{(2p-5)\vee 0}}{\sqrt{|\theta_{0,k}|}\sqrt{|\theta_{0,k'}|}}(\tilde{M}_p)_{k,k'}\right)^q\middle|\Pi\right]^{\frac{1}{q}}\right.\right.\right. \\
& \left.\left.\left.+E\left[\left(\sum_{p=0}^\infty R^C(p+1)^{v+1}\frac{(1-s_n(T))^{(2p-5)\vee 0}}{|\theta_{0,k'}|}1_{\theta_{2p+2,k}\cap\theta_{0,k'}\ne\emptyset}\right)^q\middle|\Pi\right]^{\frac{1}{q}}\right\}\right)^2\right\}^{\frac{q}{2}}
\end{aligned}
$$

by Lemma 1.12.

Moreover, by Lemma 1.10, we have

$$\begin{aligned}E[|\partial_\sigma^v\mathcal{X}|^q|\Pi] &\leq CE[R^C s_n(T)^{-(v+\frac{2r+3}{2})q}|\Pi]\bigg\{\bigg(\sum_{k,k',K''}|K''|\sum_{p=0}^\infty\frac{(\tilde{M}_p)_{k,k'}1_{\theta_{2p+2,k}(K'')}}{(p+1)^{2r}}\bigg)^{\frac{q}{2}} \\ &+\bigg(\sum_k\bigg(\sum_{k'}\sqrt{|\theta_{0,k'}|}\bigg(\sum_{p=0}^\infty\bigg(\frac{(\tilde{M}_p)_{k,k'}}{(p+1)^r}\bigg)^2\bigg)^{\frac{1}{2}}\bigg)^2\bigg)^{\frac{q}{2}} \\ &+\bigg(\sum_k|\theta_{0,k}|\bigg(\sum_{k'}\bigg(\sum_{p=0}^\infty\frac{1_{\theta_{2p+2,k}\cap\theta_{0,k'}\neq\emptyset}}{(p+1)^{2r}}\bigg)^{\frac{1}{2}}\bigg)^2\bigg)^{\frac{q}{2}}\bigg\}.\end{aligned}$$

Since $(\sum_{p\in\mathbb{N}}\alpha_p)^{1/q'}\leq\sum_{p\in\mathbb{N}}\alpha_p^{1/q'}$ for $\alpha_p\geq 0$ $(p\in\mathbb{N})$ and $q'>1$, we have

$$\begin{aligned}E[|\partial_\sigma^v\mathcal{X}|^q|\Pi] &\leq CE[R^C s_n(T)^{-(v+\frac{2r+3}{2})q}|\Pi]\bigg\{\bigg(\sum_{k,k'}\sum_{p=0}^\infty\frac{(\tilde{M}_p)_{k,k'}|\theta_{2p+2,k}|}{(p+1)^{2r}}\bigg)^{\frac{q}{2}} \\ &+\bigg(\sum_k\sum_{k_1',k_2'}\sqrt{|\theta_{0,k_1'}|}\sqrt{|\theta_{0,k_2'}|}\sum_{p_1,p_2=0}^\infty\frac{(\tilde{M}_{p_1})_{k,k_1'}(\tilde{M}_{p_2})_{k,k_2'}}{(p_1+1)^r(p_2+1)^r}\bigg)^{\frac{q}{2}} \\ &+\bigg(\sum_k|\theta_{0,k}|\sum_{k_1',k_2'}\sum_{p_1,p_2=0}^\infty\frac{1_{\theta_{2p_1+2,k_1'}\cap\theta_{0,k}\neq\emptyset}1_{\theta_{2p_2+2,k_2'}\cap\theta_{0,k}\neq\emptyset}}{(p_1+1)^r(p_2+1)^r}\bigg)^{\frac{q}{2}}\bigg\}\end{aligned}$$

Therefore, $E[|\partial_\sigma^v\mathcal{X}|^q|\Pi]$ is less than the right-hand side of (1.21) since $\parallel\tilde{M}_p\parallel\leq 2$,

$$\sum_k|\theta_{0,k}|1_{\theta_{2p_1+2,k_1'}\cap\theta_{0,k}\neq\emptyset}1_{\theta_{2p_2+2,k_2'}\cap\theta_{0,k}\neq\emptyset}\leq|\theta_{2p_1+3,k_1'}\cap\theta_{2p_2+3,k_2'}|,$$

$$\sum_{k,k'}(\tilde{M}_p)_{k,k'}|\theta_{2p+2,k}|\leq\parallel\tilde{M}_p\parallel(l_n+m_n)^{\frac{1}{2}}\Phi_{2p+2,2}^{\frac{1}{2}}\leq 2(l_n+m_n)^{\frac{1}{2}}\Phi_{2p+2,2}^{\frac{1}{2}},$$

and

$$\sum_k\sum_{k_1',k_2'}\sqrt{|\theta_{0,k_1'}|}\sqrt{|\theta_{0,k_2'}|}(\tilde{M}_{p_1})_{k,k_1'}(\tilde{M}_{p_2})_{k,k_2'}\leq\parallel\tilde{M}_{p_1}^\star\tilde{M}_{p_2}\parallel\sum_k|\theta_{0,k}|\leq 8T.$$

We will estimate $E[|R_2|^q|\Pi]$ in the next step. Since

$$E[(\partial_\sigma^v(Z_kZ_{k'}-\tilde{Z}_{k,p}\tilde{Z}_{k',p}))^{2q}|\Pi]^{\frac{1}{2q}}\leq C_q(T\vee 1)(E[R^C]+\Theta_C^1)(|\theta_{2p+2,k}|^{\frac{1}{2}}+|\theta_{2p+2,k'}|^{\frac{1}{2}}),$$

Lemma 1.9 and the Cauchy-Schwarz inequality yield

$$\begin{aligned}E[|\partial_\sigma^v R_2|^q|\Pi] &\leq \frac{1}{2^q}\bigg(\sum_{k,k'}\sum_{p=0}^\infty\sum_{v_1,v_2\geq 0,v_1+v_2=v}\frac{v!}{v_1!v_2!}E\bigg[\bigg(\int_{\theta_{2p+2,k}}\partial_\sigma^{v_1}\tilde{\xi}_{p,t}^{k,k'}\cdot d\mathcal{W}_t\bigg)^{2q}\bigg|\Pi\bigg]^{\frac{1}{2q}} \\ &\times E\bigg[(\partial_\sigma^{v_2}(Z_kZ_{k'}-\tilde{Z}_{k,p}\tilde{Z}_{k',p}))^{2q}\bigg|\Pi\bigg]^{\frac{1}{2q}}\bigg)^q \\ &\leq C\bigg(\sum_{k,k'}\sum_{p=0}^\infty(p+1)^{v+1}|\theta_{2p+2,k}|^{\frac{1}{2}}E[R^C(1-s_n(T))^{2q(2p-5)\vee 0}|\Pi]^{\frac{1}{2q}}(\tilde{M}_p)_{k,k'} \\ &\times(E[R^C]+\Theta_C^1)(T\vee 1)(|\theta_{2p+2,k}|^{\frac{1}{2}}+|\theta_{2p+2,k'}|^{\frac{1}{2}})\bigg)^q.\end{aligned}$$

Hence by the Hölder inequality, we have

$$\begin{aligned} E[|\partial_\sigma^v R_2|^q|\Pi] &\leq C(T^q\vee 1)(E[R^C]+\Theta_C^1)E[R^C s_n(T)^{-(2v+2r+3)q}|\Pi]^{\frac{1}{2}} \\ &\quad \times\bigg(\sum_{p=0}^\infty\bigg(\sum_{k,k'}\frac{(\tilde{M}_p)_{k,k'}(|\theta_{2p+2,k}|+|\theta_{2p+2,k}|^{1/2}|\theta_{2p+2,k'}|^{1/2})}{(p+1)^r}\bigg)^{\frac{2q}{2q-1}}\bigg)^{q-\frac{1}{2}} \\ &\leq C(T^q\vee 1)(E[R^C]+\Theta_C^1)E[R^C s_n(T)^{-(2v+2r+3)q}|\Pi]^{\frac{1}{2}} \\ &\quad \times\bigg\{\sum_{p=0}^\infty\bigg(\frac{((l_n+m_n)\Phi_{2p+2,2})^{1/2}\vee\Phi_{2p+2,1}}{(p+1)^r}\bigg)^{\frac{2q}{2q-1}}\bigg\}^{q-\frac{1}{2}}. \end{aligned}$$

Then $E[|R_2|^q|\Pi]$ is less than the right-hand side of (1.21).

Furthermore, by Lemma 1.9 and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} E[|\partial_\sigma^v R_1|^q|\Pi] &\leq C\bigg(\sum_{k,k'}E\bigg[\bigg|\partial_\sigma^v\bigg(\int\sum_{p=0}^\infty(\xi_{p,t}^{k,k'}-\tilde{\xi}_{p,t}^{k,k'})\cdot d\mathcal{W}_t Z_k Z_{k'}\bigg)\bigg|^q\bigg|\Pi\bigg]^{\frac{1}{q}}\bigg)^q \\ &\leq C(E[R^C]+T^q(\Theta_{8q}^1)^{1/4})\bigg(\sum_{k,k'}\sum_{0\leq v_1\leq v}E\bigg[\bigg|\int\bigg(\sum_{p=0}^\infty\partial_\sigma^{v_1}(\xi_{p,t}^{k,k'}-\tilde{\xi}_{p,t}^{k,k'})\bigg)^2dt\bigg|^q\bigg|\Pi\bigg]^{\frac{1}{2q}}\bigg)^q \end{aligned} \tag{1.23}$$

Since Π is independent of $\{(X_t,(\tilde{b}_t^i)_i)\}_t$, we can choose conditional expectation for which $t\mapsto E[(\sum_{p=0}^\infty\partial_\sigma^{v_1}(\xi_{p,t}^{k,k'}-\tilde{\xi}_{p,t}^{k,k'}))^{2q}|\Pi]$ is Lebesgue integrable almost surely for $0\leq v_1\leq v$. Therefore, by (1.23) and similar argument to the proof of Lemma 1.9, we have

$$E[|\partial_\sigma^v R_1|^q|\Pi]\leq C(E[R^C]+T^q(\Theta_{8q}^1)^{1/4})\bigg(\sum_{k,k'}\sum_{0\leq v_1\leq v}\bigg(\int E\bigg[\bigg(\sum_{p=0}^\infty\partial_\sigma^{v_1}(\xi_{p,t}^{k,k'}-\tilde{\xi}_{p,t}^{k,k'})\bigg)^{2q}|\Pi\bigg]^{\frac{1}{q}}dt\bigg)^{\frac{1}{2}}\bigg)^q. \tag{1.24}$$

By Lemma 1.10, (1.24), (1.22) and the estimates after that, we have

$$\begin{aligned} &E[|\partial_\sigma^v R_1|^q|\Pi] \\ \leq\ & C(E[R^C]+T^q(\Theta_{8q}^1)^{\frac{1}{4}})\bigg(\sum_{k,k'}\bigg(\int E\bigg[\bigg(\sum_{p=0}^\infty(p+1)^{v+3}R^{2v+9} \\ &\times(1-s_n(T))^{(2p-7)\vee 0}(\tilde{M}_p)_{k,k'}|\theta_{2p+2,k}|^{\frac{1}{2}}(T^{\frac{1}{2}}\vee 1)1_{\theta_{2p+2,k'}}(t)\bigg)^{2q}\bigg|\Pi\bigg]^{\frac{1}{q}}dt\bigg)^{\frac{1}{2}}\bigg)^q \\ \leq\ & C(E[R^C]+T^q(\Theta_{8q}^1)^{\frac{1}{4}})(T^{\frac{q}{2}}\vee 1)E[R^C s_n(T)^{-(2v+2r+7)q}|\Pi]^{\frac{1}{2}}\bigg(\sum_{k,k'}\bigg(\sum_{p=0}^\infty\frac{((\tilde{M}_p)_{k,k'})^2|\theta_{2p+2,k}||\theta_{2p+2,k'}|}{(p+1)^{2r}}\bigg)^{\frac{1}{2}}\bigg)^q \\ \leq\ & C(E[R^C]+T^q(\Theta_{8q}^1)^{\frac{1}{4}})(T^{\frac{q}{2}}\vee 1)E[R^C s_n(T)^{-(2v+2r+7)q}|\Pi]^{\frac{1}{2}}\bigg(\sum_{p=0}^\infty\frac{\Phi_{2p+2,1}}{(p+1)^r}\bigg)^q. \end{aligned}$$

Then $E[|\partial_\sigma^v R_1|^q|\Pi]$ is less than the right-hand side of (1.21). Similarly, we can see

$$E[|\partial_\sigma^v R_3|^q]\leq CE[R^C s_n(T)^{-(2v+2r+5)q}|\Pi]^{\frac{1}{2}}(E[R^C]+T^q(\Theta_{8q}^1)^{\frac{1}{4}})\bigg(\sum_{p=0}^\infty\frac{\sqrt{(l_n+m_n)\Phi_{2p+2,2}}}{(p+1)^r}\bigg)^q.$$

Hence $E[|\partial_\sigma^v(Z^\star(M-\tilde{M})Z/2)|^q|\Pi]$ is less than the right-hand side of (1.21).

Similaly, we can see

$$E\bigg[\bigg|\partial_\sigma^v\bigg(\frac{1}{2}\sum_{p=1}^\infty\frac{(-1)^p}{p}(\mathrm{tr}(\tilde{L}^p)-\mathrm{tr}(\tilde{L}_p^p))\bigg)\bigg|^q\bigg|\Pi\bigg]$$

is less than the right-hand side of (1.21).

Moreover,

$$E[|\partial_\sigma^v(\tilde{\mu}^\star\tilde{M}\tilde{\mu}/2)|^q|\Pi] \leq CE\Bigg[\bigg(\sum_{v_1=0}^v|\partial_\sigma^{v_1}\tilde{\mu}|\bigg)^{2q}\bigg(\sum_{v_2=0}^v \parallel \partial_\sigma^{v_2}\tilde{M} \parallel\bigg)^q\bigg|\Pi\Bigg] \leq CE[(R^C s_n(T)^{-(v+1)q})^2|\Pi]^{\frac{1}{2}}(2T)^q(\Theta_{4q}^1)^{\frac{1}{2}}$$

since $\sum_{v_2=0}^v \parallel \partial_\sigma^{v_2}\tilde{M} \parallel \leq CR^{2v}\sum_{p=0}^\infty((2p)^v \vee 1)\bar{\rho}_{\tau_n}^{(2p-v)\vee 0} \leq CR^{2v}s_n(T)^{-(v+1)}$.

We will estimate $E[|\partial_\sigma^v(\tilde{\mu}^\star\tilde{M}\hat{Z})|^q|\Pi]$ at last. Let $\mathcal{L}_{p,k,k'} = ((\tilde{L}_{2p})^{2p} - (\tilde{L}_{2p+1})^{2p+1})_{k,k'}$. $\tilde{\mu}^\star\tilde{M}\hat{Z}$ can be decomposed as

$$\tilde{\mu}^\star\tilde{M}\hat{Z} = \sum_{k,k'}\sum_{p=0}^\infty \mathcal{L}_{p,k,k'}\tilde{\mu}_{p,k}\hat{Z}_{k'} + \sum_{k,k'}\sum_{p=0}^\infty \mathcal{L}_{p,k,k'}(\tilde{\mu}_k - \tilde{\mu}_{p,k})\hat{Z}_{k'} = \Xi_1 + \Xi_2,$$

where

$$\tilde{\mu}_{p,k} = \begin{cases} \mu^1_{L(\theta_{p+1,k})}\sqrt{|I^k|}/|b^1_{L(\theta_{p+1,k})\wedge\tau_n}| & (k \leq l_n) \\ \mu^2_{L(\theta_{p+1,k})}\sqrt{|J^{k-l_n}|}/|b^2_{L(\theta_{p+1,k})\wedge\tau_n}| & (k > l_n) \end{cases}$$

Then by the Burkholder-Davis-Gundy inequality, we obtain

$$\begin{aligned} E[|\partial_\sigma^v\Xi_1|^q|\Pi] &\leq C\sum_{v_1+v_2=v}E\Bigg[\bigg(\sum_{k'}\bigg(\sum_k\sum_{p=0}^\infty\partial_\sigma^{v_1}(\mathcal{L}_{p,k,k'}\tilde{\mu}_{p,k})\bigg)^2(\partial_\sigma^{v_2}\hat{Z}_{k'})^2\bigg)^{\frac{q}{2}}\bigg|\Pi\Bigg] \\ &\leq C\sum_{v_1+v_2=v}\Bigg(\sum_{k'}E\Bigg[\bigg(\sum_k\sum_{p=0}^\infty\partial_\sigma^{v_1}(\mathcal{L}_{p,k,k'}\tilde{\mu}_{p,k})\bigg)^q(\partial_\sigma^{v_2}\hat{Z}_{k'})^q\bigg|\Pi\Bigg]^{\frac{2}{q}}\Bigg)^{\frac{q}{2}} \\ &\leq CE[R^{4q}]^{\frac{1}{2}}\sum_{v_1=0}^v\Bigg(\sum_{k'}\bigg(\sum_k\sum_{p=0}^\infty E[(\partial_\sigma^{v_1}(\mathcal{L}_{p,k,k'}\tilde{\mu}_{p,k}))^{2q}|\Pi]^{\frac{1}{2q}}\bigg)^2\Bigg)^{\frac{q}{2}}. \end{aligned}$$

Since

$$E[(\partial_\sigma^{v_1}(\mathcal{L}_{p,k,k'}\tilde{\mu}_{p,k}))^{2q}|\Pi]^{\frac{1}{2q}} \leq C(2p+1)^vE[R^C\bar{\rho}_{\tau_n}^{4q(2p-5)}|\Pi]^{\frac{1}{4q}}(\tilde{M}_p)_{k,k'}(\Theta_{4q}^1)^{\frac{1}{4q}}\sqrt{|\theta_{0,k}|},$$

we obtain

$$\begin{aligned} E[|\partial_\sigma^v\Xi_1|^q|\Pi] &\leq CE[R^{4q}]^{\frac{1}{2}}(\Theta_{4q}^1)^{\frac{1}{4}}\Bigg(\sum_{k'}\sum_{k_1,k_2}\sum_{p_1,p_2=0}^\infty E[R^C\bar{\rho}_{\tau_n}^{4q(2p_1-5)}|\Pi]^{\frac{1}{4q}}(p_1+1)^v(p_2+1)^v \\ &\quad\times E[R^C\bar{\rho}_{\tau_n}^{4q(2p_2-5)}|\Pi]^{\frac{1}{4q}}(\tilde{M}_{p_1})_{k_1,k'}(\tilde{M}_{p_2})_{k_2,k'}\sqrt{|\theta_{0,k_1}|}\sqrt{|\theta_{0,k_2}|}\Bigg)^{\frac{q}{2}} \\ &\leq CT^{q/2}E[R^{4q}]^{\frac{1}{2}}(\Theta_{4q}^1)^{\frac{1}{4}}\bigg(\sum_{p=0}^\infty E[R^C\bar{\rho}_{\tau_n}^{4q(2p-5)}|\Pi]^{\frac{1}{4q}}(p+1)^v\bigg)^q. \end{aligned}$$

Then $E[|\partial_\sigma^v\Xi_1|^q|\Pi]$ is less than the right-hand side of (1.21) since

$$\begin{aligned} \sum_{p=0}^\infty E[R^C\bar{\rho}_{\tau_n}^{4q(2p-5)}|\Pi]^{\frac{1}{4q}}(p+1)^v &\leq \bigg(\sum_{p=0}^\infty E[R^C\bar{\rho}_{\tau_n}^{4q(2p-5)}|\Pi](p+1)^{q(4v+\frac{9}{2})}\bigg)^{\frac{1}{4q}}\bigg(\sum_{p=0}^\infty(p+1)^{-\frac{9}{8}}\bigg)^{1-\frac{1}{4q}} \\ &\leq CE[R^Cs_n(T)^{-q(4v+5)}|\Pi]^{1/4q}. \end{aligned}$$

On the other hand, Lemma 1.9 yields

$$\begin{aligned}
&E[|\partial_\sigma^v \Xi_2|^q|\Pi] \\
\leq\ & C\bigg(\sum_{p=0}^\infty \sum_{k,k'} E\bigg[\bigg\{R^C\bar\rho_{\tau_n}^{(2p-v)\vee 0}(2p+1)^v(\tilde M_p)_{k,k'}\bigg(\sum_{v_1=0}^v|\partial_\sigma^{v_1}(\tilde\mu_k-\tilde\mu_{p,k})|\bigg)\bigg(\sum_{v_2=0}^v|\partial_\sigma^{v_2}\hat Z_k|\bigg)\bigg\}^q\bigg|\Pi\bigg]^{\frac{1}{q}}\bigg)^q \\
\leq\ & CE[R^C]\bigg\{\sum_{p=0}\sum_{k,k'}(\tilde M_p)_{k,k'}E[R^C\bar\rho_{\tau_n}^{2q(2p-v)\vee 0}(2p+1)^{2qv}|\Pi]^{\frac{1}{2q}} \\
&\times\bigg(E[R^C]^{\frac{1}{8q}}\sqrt{|\theta_{0,k}|}(\Theta_{8q}^2)^{\frac{1}{8q}}|\theta_{p+1,k}|^\gamma+E[R^C]^{\frac{1}{8q}}\sqrt{|\theta_{0,k}|}(\Theta_{8q}^1)^{\frac{1}{8q}}(T^{\frac12}\vee 1)\sqrt{|\theta_{p+1,k}|}\bigg)\bigg\}^q \\
\leq\ & CE[R^C](T^{\frac{q}{2}}\vee 1)((\Theta_{8q}^1)^{\frac18}+(\Theta_{8q}^2)^{\frac18})\bigg(\sum_{p=0}^\infty E[R^C\bar\rho_{\tau_n}^{2q(2p-4)\vee 0}|\Pi](p+1)^{(2v+6)q}\bigg)^{\frac12} \\
&\times\bigg\{\sum_{p=0}^\infty\frac{1}{(p+1)^2}\bigg(\sum_{k,k'}(\tilde M_p)_{k,k'}\sqrt{|\theta_{0,k}|}\frac{((4p+1)r_n)^\gamma\vee((4p+1)r_n)^{1/2}}{p+1}\bigg)^{\frac{2q}{2q-1}}\bigg\}^{\frac{2q-1}{2}} \\
\leq\ & CE[R^C]r_n^{q(\gamma\wedge\frac12)}(T^{\frac32 q}\vee 1)(l_n+m_n)^{\frac{q}{2}}((\Theta_{8q}^1)^{\frac18}+(\Theta_{8q}^2)^{\frac18})E[R^Cs_n(T)^{-(2v+7)q}|\Pi]^{\frac12},
\end{aligned}$$

where we use the fact $r_n^\gamma\vee r_n^{\frac12}=T^\gamma(r_nT^{-1})^\gamma\vee T^{\frac12}(r_nT^{-1})^{\frac12}\leq(\sqrt{T}\vee 1)r_n^{\gamma\wedge\frac12}$. This complete the proof of (1.21).

We next estimate $E[|\partial_\sigma^v\tilde H_{n,s_n}^2(T;\sigma)-\partial_\sigma^v\tilde H_{n,s_n}^3(T;\sigma)|^q|\Pi]$. Let $\mathcal{J}(k)=1$ $(1\leq k\leq l_n)$, $\mathcal{J}(k)=2$ $(l_n<k\leq l_n+m_n)$ and $\check B_k^i=|b^i_{L(\theta_{0,k}),*}|/|b^i_{L(\theta_{0,k})\wedge\tau_n}|$ for $1\leq k\leq l_n+m_n, i=1,2$. For $p\in\mathbb{Z}_+$ and $1\leq k,k'\leq l_n+m_n$, we define $\{\check\xi_{p,t}^{k,k'}\},\{\check\eta_{p,t}^{k,k'}\}$ as if follows.

1. the case $k=k'$:

$$\begin{aligned}
&-\frac12((\tilde L_{2p})^{2p})_{k,k}(\langle\hat Z\rangle_T)_{k,k}+\frac12(\check B_k^{\mathcal{J}(k)})^2\rho_{L(\theta_{0,k})}^{2p}(\tilde M_p)_{k,k} \\
&-(\log|b^{\mathcal{J}(k)}_{\theta_{0,k},\tau_n}|-\log|b^{\mathcal{J}(k)}_{L(\theta_{0,k})\wedge\tau_n}|)1_{\{p=0\}}+\frac{1}{4p}(((\tilde L_{2p})^{2p})_{k,k}-\rho^{2p}_{L(\theta_{0,k})\wedge\tau_n}(\tilde M_p)_{k,k})1_{\{p\geq 1\}} \\
&=\int\check\xi_{p,t}^{k,k}\cdot d\mathcal{W}_t+\int\check\eta_{p,t}^{k,k}dt.
\end{aligned}$$

2. the case ($k\leq l_n$ and $k'>l_n$) or ($k>l_n$ and $k'\leq l_n$):

$$\begin{aligned}
&-\frac12\{(\check B_k^1\check B_k^2+\check B_k^{\mathcal{J}(k)})\rho^{2p+1}_{L(\theta_{0,k})\wedge\tau_n}\rho_{L(\theta_{0,k}),*}-\check B_{k'}^{\mathcal{J}(k)}\rho^{2p+1}_{L(\theta_{0,k'})\wedge\tau_n}\rho_{L(\theta_{0,k'}),*}\} \\
&\times(\tilde M_p)_{k,k'}(\tilde M_0)_{k,k'}+\frac12((\tilde L_{2p+1})^{2p+1})_{k,k'}(\langle Z\rangle_T)_{k,k'}=\int\check\xi_{p,t}^{k,k'}\cdot d\mathcal{W}_t+\int\check\eta_{p,t}^{k,k'}dt.
\end{aligned}$$

3. other case : We set $\check\xi_{p,t}^{k,k'}\equiv 0$ and $\check\eta_{p,t}^{k,k'}\equiv 0$.

Then by Itô's formula, we obtain

$$\begin{aligned}
|\sum_{k,k'}\partial_\sigma^v\check\xi_{p,t}^{k,k'}| &\leq CR^C(p+1)^{v+1}\bar\rho_T^{(2p-v-1)\vee 0}\sum_k\{(\tilde M_p)_{k,k}+(\tilde M_{p+1})_{k,k}\}1_{\theta_{2p+2,k}}(t), \\
|\sum_{k,k'}\partial_\sigma^v\check\eta_{p,t}^{k,k'}| &\leq CR^C(p+1)^{v+2}\bar\rho_T^{(2p-v-2)\vee 0}\sum_k\{(\tilde M_p)_{k,k}+(\tilde M_{p+1})_{k,k}\}1_{\theta_{2p+2,k}}(t).
\end{aligned}$$

Moreover, we have

$$\tilde H_{n,s_n}^2(T;\sigma)-\tilde H_{n,s_n}^3(T;\sigma)=\int\sum_{k,k'}\sum_{p=0}^\infty\check\xi_{p,t}^{k,k'}\cdot d\mathcal{W}_t+\int\sum_{k,k'}\sum_{p=0}^\infty\check\eta_{p,t}^{k,k'}dt.$$

Therefore we obtain the conclusion by Lemma 1.10. $\square$

Lemma 1.14. *1. Assume that $[A1],[A2]$ hold and $\{b_n^{-1}(l_n+m_n)\}_n$ is tight. Then*

$$\sup_\sigma b_n^{-1}|\partial_\sigma^v H_n(\sigma)-\partial_\sigma^v \tilde{H}^3_{n,s_n}(T;\sigma)| \to^p 0$$

as $n\to\infty$ for $0\leq v\leq 3$, where $s_n=r_n^{1/42}\wedge((1-|\rho_0|)/2)$.

2. Let $0\leq v\leq 3$, $q\in 2\mathbb{N}, q>n_1$, $\delta\geq 1$, and $\{s_n\}_{n\in\mathbb{N}}$ be stochastic processes which satisfy $[S]$. Assume that $[A1],[A2],[A4\text{-}(2q),\delta]$ hold,

$$\limsup_{n\to\infty} E[s_n(T)^{-(2v+2[\delta]+12)q}]<\infty \quad \text{and} \quad \limsup_{n\to\infty} E[b_n^{-2q}(l_n+m_n)^{2q}]<\infty. \tag{1.25}$$

Then there exists $n_0\in\mathbb{N}$ such that

$$\sup_{n\geq n_0} E\left[\left(\sup_\sigma b_n^{-1/2}|\partial_\sigma^v\hat{H}_n(\sigma;s_n)-\partial_\sigma^v\tilde{H}^3_{n,s_n}(T;\sigma)|\right)^q\right]<\infty.$$

Proof. We first prove 2. Since $\{r_n\}_n$ is bounded and $r_n\to^p 0$ as $n\to\infty$, we have $\lim_{n\to\infty}E[r_n^{q'}]=0$ for any $q'>0$. Then by $[A1],[A4\text{-}(2q),\delta],(1.25)$, Lemma 1.13 with $r=[\delta]+2$, Cauchy-Schwarz inequalities, Jensen's inequality and the estimate $\Phi_{2p+2,2}\leq r_n(8p+9)\Phi_{2p+2,1}$, we have $\lim_{n\to\infty}\sup_\sigma E[b_n^{-q/2}|\partial_\sigma^v(\hat{H}_n-\tilde{H}_n^1)|^q]=0$ for $0\leq v\leq 4$. Hence $\lim_{n\to\infty}E[b_n^{-q/2}\sup_\sigma|\partial_\sigma^v(\hat{H}_n-\tilde{H}_n^1)|^q]=0$ for $0\leq v\leq 3$ by Sobolev's inequality. Similarly, we have $\lim_{n\to\infty}E[b_n^{-q/2}\sup_\sigma|\partial_\sigma^v(\tilde{H}_n^2-\tilde{H}_n^3)|^q]=0$ for $0\leq v\leq 3$.

We estimate $\tilde{H}_n^1-\tilde{H}_n^2$ in the next step. Let $0\leq v\leq 4$ and Π be deterministic. By Itô's formula and symmetry of $\tilde{M}$, we have

$$\tilde{H}_n^1(t)-\tilde{H}_n^2(t)=-\frac{1}{2}\sum_{k,k'}\tilde{M}_{k,k'}\{\hat{Z}_{k,t}\hat{Z}_{k',t}-(\langle\hat{Z}\rangle_t)_{k,k'}\}=-\sum_{k,k'}\tilde{M}_{k,k'}\int_0^t\hat{Z}_{k,s}d\hat{Z}_{k',s}.$$

Therefore, $\{\partial_\sigma^v(\tilde{H}_n^1(t)-\tilde{H}_n^2(t))\}_{0\leq t\leq T}$ is the martingale. By the Burkholder-Davis-Gundy inequality, we obtain

$$E[|\partial_\sigma^v(\tilde{H}_n^1-\tilde{H}_n^2)|^q]\leq CE[\langle\partial_\sigma^v(\tilde{H}_n^1-\tilde{H}_n^2)\rangle_T^{q/2}] \quad (0\leq v\leq 4).$$

Moreover,

$$\langle\partial_\sigma^v(\tilde{H}_n^1-\tilde{H}_n^2)\rangle_T\leq CR^4\parallel\tilde{M}_0\parallel\left(\sum_{0\leq j_1+j_2\leq v}|(\partial_\sigma^{j_1}\hat{Z})^*|\times\parallel\{|\partial_\sigma^{j_2}\tilde{M}_{k,k'}|\}_{k,k'}\parallel\right)^2,$$

where $|(\partial_\sigma^{j_1}\hat{Z})^*|^2=\sum_k\sup_t|\partial_\sigma^{j_1}\hat{Z}_{k,t}|^2$. Since $E[|(\partial_\sigma^j\hat{Z})^*|^{2q}]\leq CE[R^{4q}](l_n+m_n)^q$, $\parallel\tilde{M}_0\parallel\leq 2$ and $\parallel\{|\partial_\sigma^j\tilde{M}_{k,k'}|\}_{k,k'}\parallel\leq CR^{2j}(1-\bar{\rho}_T)^{-j-1}$ for $0\leq j\leq 4$, we have

$$b_n^{-\frac{q}{2}}E[|\partial_\sigma^v(\tilde{H}_n^1-\tilde{H}_n^2)|^q]\leq CE[R^C]E[(b_n^{-1}(l_n+m_n))^{q/2}]E[R^C(1-\bar{\rho}_T)^{-2(v+1)q}]^{1/2}$$

for general Π. Then by Sobolev's inequality, there exists $n_0\in\mathbb{N}$ such that

$$\sup_{n\geq n_0}E\left[\left(\sup_\sigma b_n^{-\frac{1}{2}}|\partial_\sigma^v(\tilde{H}_n^1-\tilde{H}_n^2)|\right)^q\right]<\infty$$

for $0\leq v\leq 3$. This completes the proof of 2.

Finally, we prove 1. Since $|\Phi_{2p+2,i}|\leq C(p+1)^i r_n^i(l_n+m_n)$ and $|\Phi_{2p_1+3,2p_2+3}|\leq C(p_1+1)(p_2+1)(l_n+m_n)^2r_n$ for $p_1,p_2\in\mathbb{Z}_+$ and $i=1,2$, by Lemma 1.13 with $r=3$, we have

$$\sup_{\sigma\in\mathbb{Q}^{n_1}\cap\Lambda}E[|\partial_\sigma^v(\hat{H}_n(\sigma;s_n)-\tilde{H}^1_{n,s_n}(T;\sigma))|^q|\Pi]\leq C(1+r_n^{-\frac{q}{4}})(1+r_n^{q(\gamma\wedge\frac{1}{2})}(l_n+m_n)^{\frac{q}{2}}+r_n^{\frac{q}{2}}(l_n+m_n)^q)$$

for $q\in 2\mathbb{N}$, $q>n_1$ and $0\leq v\leq 4$. Therefore, by Lemma 1.11 2., the assumptions and the inequality $T=\sum_I|I|\leq r_nl_n$, we obtain $\{b_n^{-1}r_n^{-1}\}_n$ is tight and

$$\sup_\sigma b_n^{-1}|\partial_\sigma^v(\hat{H}_n(\sigma;s_n)-\tilde{H}^1_{n,s_n}(T;\sigma))|\to^p 0$$

as $n \to \infty$ for $0 \le v \le 3$. Similarly, we obtain $\sup_\sigma b_n^{-1}|\partial_\sigma^v(\tilde{H}^2_{n,s_n}(T;\sigma) - \tilde{H}^3_{n,s_n}(T;\sigma)| \to^p 0$ as $n \to \infty$ for $0 \le v \le 3$.

Moreover, similarly to the proof of 2., we have

$$E[|b_n^{-1}\partial_\sigma^v(\tilde{H}^1_n - \tilde{H}^2_n)|^q|\Pi] \le CE[R^C]b_n^{-q/2}(b_n^{-1}(l_n+m_n))^{q/2}E[R^C(1-\bar{\rho}_T)^{-C}]$$

for $q > n_1$ and $0 \le v \le 4$. Hence by Lemma 1.11 2., we have $b_n^{-1}\sup_\sigma|\partial_\sigma^v(\tilde{H}^1_n - \tilde{H}^2_n)| \to^p 0$ as $n \to \infty$ for $0 \le v \le 3$.

Moreover, since $P[\tau(s_n) < T] \to 0$ as $n \to \infty$, $b_n^{-1}\sup_\sigma|\partial_\sigma^v(H_n(\sigma) - \hat{H}_n(\sigma;s_n))| \to^p 0$ as $n \to \infty$, which completes the proof. □

Lemma 1.15. *Assume [A3′]. Then*

$$\sup_{\sigma\in\Lambda}|\Psi^{p,1}(f(\cdot,\sigma),a_p)| \to^p 0 \quad \text{and} \quad \sup_{\sigma\in\Lambda}|\Psi^{p,2}(f(\cdot,\sigma),c_p)| \to^p 0$$

as $n \to \infty$ for $p \in \mathbb{Z}_+$ and $f(t,\sigma)$: random variable defined on $[0,T]\times\bar{\Lambda}$ such that f is continuous with respect to (t,σ).

Proof. Let $\{f^k\}_k$ be step functions such that $\sup_{t,\sigma}|f(t,\sigma) - f^k(t,\sigma)| \to^p 0$ as $k \to \infty$. By [A3′], we obtain $\sup_\sigma|\int_0^T f^k(t,\sigma)\nu_n^{p,1}(dt) - \int_0^T f^k(t,\sigma)a_p(t)dt| \to^p 0$ as $n \to \infty$ for any $k \in \mathbb{N}$.

Since $\{\nu_n^{0,1}([0,T))\}_n$ is tight, for any $\epsilon,\delta > 0$, there exists $K \in \mathbb{N}$ such that

$$P\left[\sup_\sigma\left|\int_0^T (f-f^k)\nu_n^{p,1}(dt)\right| \vee \sup_\sigma\left|\int_0^T (f-f^k)a_p(t)dt\right| > \delta\right] < \epsilon \ (k \ge K, n \in \mathbb{N}).$$

Then there exists $N \in \mathbb{N}$ such that $P[\sup_\sigma|\Psi^{p,1}(f(\cdot,\sigma),a_p)| > 3\delta] < 2\epsilon$ for $n \ge N$. Similarly, we have $\sup_\sigma|\Psi^{p,2}(f(\cdot,\sigma),c_p)| \to^p 0$ as $n \to \infty$. □

Proof of Proposition 1.3.

We first prove 1. By Lemma 1.14, it is sufficient to show $\sup_\sigma|b_n^{-1}\partial_\sigma^v\tilde{H}^3_n(T;\sigma) - \int_0^T\partial_\sigma^v h_t^\infty(\sigma)dt| \to^p 0$ as $n \to \infty$ for $0 \le v \le 3$, where $s_n = r_n^{1/42} \wedge ((1-|\rho_0|)/2)$.

Since $P[\tau(s_n) < T] \to 0$ as $n \to \infty$,

$$\sum_{p=0}^\infty\sum_{i=1}^2\sup_\sigma\left|\int_0^T(\partial_\sigma^v\mathcal{D}_p^i(t\wedge\tau(s_n),t;\sigma) - \partial_\sigma^v\mathcal{D}_p^i(t,t;\sigma))\nu_n^{p,i}(dt)\right| \to^p 0$$

as $n \to \infty$ for $0 \le v \le 3$.

Moreover, by Lemma 1.15, $\sup_\sigma|\Psi^{p,1}(\partial_\sigma^v\mathcal{D}_p^1(\cdot,\cdot;\sigma),a_p)| \to^p 0$ as $n \to \infty$ for $p \in \mathbb{Z}_+$ and $0 \le v \le 3$.

Then by Lemma 1.3, the tightness of $\{\nu_n^{0,1}([0,T))\}_n$, and the estimates $\nu_n^{p,1}([0,T)) \le \nu_n^{0,1}([0,T))$ and $|\partial_\sigma^v\mathcal{D}_p^1(t,t,\sigma)| \le CR^C\bar{\rho}_T^{(2p-v)\vee 0}(p+1)^v$, we have

$$\sum_{p=0}^\infty\sup_\sigma|\Psi^{p,1}(\partial_\sigma^v\mathcal{D}_p^1(\cdot,\cdot;\sigma),a_p)| \to^p 0$$

as $n \to \infty$ for $0 \le v \le 3$. Similarly, we obtain

$$\sum_{p=0}^\infty\sup_\sigma|\Psi^{p,2}(\partial_\sigma^v\mathcal{D}_p^2(\cdot,\cdot;\sigma),c_p)| \to^p 0$$

as $n \to \infty$ for $0 \le v \le 3$. Since $h_t^\infty(\sigma) = \sum_{p=0}^\infty(\mathcal{D}_p^1(t,t;\sigma)a_p(t) + \mathcal{D}_p^2(t,t;\sigma)c_p(t))$, we obtain 1.

We next prove 2. First, $[S\text{-}((2v+2[\delta]+12)q),\xi]$ and the estimate $\nu_n^{p,i}([0,T)) \le \nu^{0,i}([0,T)) \le b_n^{-1}(l_n+m_n)$ $(p \in \mathbb{Z}_+)$ yield

$$\sup_n E\left[\sup_\sigma\left|b_n^{\frac{\xi}{2q}}\sum_{p=0}^\infty\sum_{i=1}^2\int_0^T\left\{\partial_\sigma^v\mathcal{D}_p^i(t\wedge\tau(s_n),t;\sigma) - \partial_\sigma^v\mathcal{D}_p^i(t,t;\sigma)\right\}\nu_n^{p,i}(dt)\right|^q\right] < \infty$$

for $0 \leq v \leq 3$.

Then by Lemma 1.14, it is sufficient to show that there exists $n_0 \in \mathbb{N}$ such that

$$\sup_{n \geq n_0} E\left[\sup_{\sigma}\left|b_n^{\eta}\sum_{p=0}^{\infty}\left\{\Psi^{p,1}(\partial_\sigma^v \mathcal{D}_p^1(\cdot,\cdot;\sigma),a_p) + \Psi^{p,2}(\partial_\sigma^v \mathcal{D}_p^2(\cdot,\cdot;\sigma),c_p)\right\}\right|^q\right] < \infty.$$

By $[A3'\text{-}q,\eta]$ and independence of $\{\Pi_n\}_n$ and X, we have

$$\begin{aligned} & \sup_{n \geq n_0} E\left[\left|b_n^{\eta}\sum_{p=0}^{\infty}\Psi^{p,1}(\partial_\sigma^v \mathcal{D}_p^1(\cdot,\cdot;\sigma),a_p)\right|^q\right] \\ \leq \ & C\sum_{p=0}^{\infty}\frac{1}{(p+1)^2}\sup_{n \geq n_0} E[(b_n^{\eta}(p+1)^2|\Psi^{p,1}(\partial_\sigma^v \mathcal{D}_p^1(\cdot,\cdot;\sigma),a_p)|)^q] \\ \leq \ & C\sum_{p=0}^{\infty}(p+1)^C E\left[\sup_t |\partial_\sigma^v \mathcal{D}_p^1(t,t;\sigma)|^q + \omega_\alpha(\partial_\sigma^v \mathcal{D}_p^1(\cdot,\cdot;\sigma))^q\right] \end{aligned} \tag{1.26}$$

for $0 \leq v \leq 4$, α in $[A3'\text{-}q,\eta]$ and n_0 which is renewed if necessary.

By Itô's formula, we obtain

$$E[|\partial_\sigma^v \mathcal{D}_p^1(t,t;\sigma) - \partial_\sigma^v \mathcal{D}_p^1(s,s;\sigma)|^q] \leq CE[((p+1)^{v+2}\bar{\rho}_T^{(2p-v-2)\vee 0}R^C)^q]|t-s|^{q/2}$$

for $s < t$.

Hence by Kolmogorov's criterion([39] Theorem (2.1)) and its proof, we have

$$E[\omega_\alpha(\partial_\sigma^v \mathcal{D}_p^1(\cdot,\cdot;\sigma))^q] \leq CE[((p+1)^{v+2}\bar{\rho}_T^{(2p-v-2)\vee 0}R^C)^q]. \tag{1.27}$$

(1.26),(1.27) yield $\sup_\sigma \sup_{n \geq n_0} E[|b_n^{\eta}\sum_{p=0}^{\infty}\Psi^{p,1}(\partial_\sigma^v \mathcal{D}_p^1(\cdot,\cdot;\sigma),a_p)|^q] < \infty$. Then by Sobolev's inequality, we have $\sup_{n \geq n_0} E[\sup_\sigma |b_n^{\eta}\sum_{p=0}^{\infty}\Psi^{p,1}(\partial_\sigma^v \mathcal{D}_p^1(\cdot,\cdot;\sigma),a_p)|^q] < \infty$ for $0 \leq v \leq 3$. Similarly, there exists $n_1 \in \mathbb{N}$ such that
$\sup_{n \geq n_1} E[\sup_\sigma |b_n^{\eta}\sum_{p=0}^{\infty}\Psi^{p,2}(\partial_\sigma^v \mathcal{D}_p^2(\cdot,\cdot;\sigma),c_p)|^q] < \infty$ for $0 \leq v \leq 3$. □

1.7.3 Proof of Lemmas 1.7 and 1.8

Proof of Lemma 1.7.

Let $G_{[s,t)} = \{G_{I,J}\}_{L(I),L(J) \in [s,t)}$ for $0 \leq s < t \leq T$, $\{\lambda_i'\}_{i=1}^{l'}$ be the eigenvalues of $G_{[s,t)}G_{[s,t)}^{\star}$ and $f_1^{(s)}(t) = f_1(t, B_s^1 B_s^2, \rho_s, \rho_{s,*})$. Since $|b_n^{-1}\mathrm{tr}((G_{[s,t)}G_{[s,t)}^{\star})^p) - \nu_n^{p,1}([s,t))| \to^p 0$ as $n \to \infty$ by a similar argument to the proof of Lemma 1.4, we have

$$\int_s^t a_p(u)du = \text{P-}\lim_{n\to\infty} b_n^{-1}\sum_{i=1}^{l'}(\lambda_i')^p,$$

where P- lim denotes the limit in probability. Moreover, similarly to the proof of Lemma 1.2, we have $\sup_i |\lambda_i'| \leq 1$.

Let $g_i = g_i(\rho_s) = \sqrt{1-\lambda_i'\rho_s^2}$, $g_{i,*} = g_i(\rho_{s,*})$. Then since

$$\begin{aligned} \mathcal{A}(\rho)\frac{\rho_*}{\rho} - \mathcal{A}(\rho) - a_0 &= \sum_{p=1}^{\infty}a_p\rho^{2p}\left(\frac{\rho_*}{\rho}-1\right) - a_0 = \sum_{p=0}^{\infty}a_{p+1}\rho^{2p+1}\rho_* - \sum_{p=0}^{\infty}a_p\rho^{2p}, \\ \int_{\rho_*}^{\rho}\frac{\mathcal{A}(\rho')}{\rho'}d\rho' &= \frac{1}{2}\sum_{p=1}^{\infty}\frac{a_p}{p}(\rho^{2p}-\rho_*^{2p}) \end{aligned}$$

for $\rho, \rho_* \in (-1,1)$, we have

$$
\begin{aligned}
&\int_s^t f_1^{(s)}(u)du \\
= \;& \text{P-}\lim_{n\to\infty} b_n^{-1} \sum_{i=1}^{l'} \left\{ 1 + \log(B_s^1 B_s^2) + B_s^1 B_s^2 \sum_{p=0}^{\infty} \left((\lambda_i')^{p+1} \rho_s^{2p+1} \rho_{s,*} - (\lambda_i')^p \rho_s^{2p} \right) + \frac{1}{2} \sum_{p=1}^{\infty} \frac{(\lambda_i')^p}{p} (\rho_s^{2p} - \rho_{s,*}^{2p}) \right\} \\
= \;& \text{P-}\lim_{n\to\infty} b_n^{-1} \sum_{i=1}^{l'} \left\{ 1 + B_s^1 B_s^2 g_i^{-2} (\lambda_i' \rho_s \rho_{s,*} - 1) + \log(B_s^1 B_s^2 g_{i,*} g_i^{-1}) \right\} \\
= \;& \text{P-}\lim_{n\to\infty} b_n^{-1} \sum_{i=1}^{l'} \left\{ B_s^1 B_s^2 g_i^{-2} (\lambda_i' \rho_s \rho_{s,*} - 1) + B_s^1 B_s^2 g_{i,*} g_i^{-1} + F(B_s^1 B_s^2 g_{i,*} g_i^{-1}) \right\} \qquad (1.28)
\end{aligned}
$$

by Lemma 1.3. Since

$$
\begin{aligned}
g_i^{-2}(\lambda_i' \rho_s \rho_{s,*} - 1) + g_{i,*} g_i^{-1} &= -\frac{(\lambda_i' \rho_s \rho_{s,*} - 1)^2 - g_{i,*}^2 g_i^2}{g_i^2 (1 - \lambda_i' \rho_s \rho_{s,*} + g_{i,*} g_i)} = -\frac{\lambda_i' (\rho_s - \rho_{s,*})^2}{g_i^2 (1 - \lambda_i' \rho_s \rho_{s,*} + g_{i,*} g_i)} \\
&\le -\lambda_i' (\rho_s - \rho_{s,*})^2 / 3 \qquad (1.29)
\end{aligned}
$$

and $B_s^1 B_s^2 g_{i,*} g_i^{-1} - 1 \le R^4 / \sqrt{1 - \bar{\rho}_T^2}$, it follows that

$$
\int_s^t f_1^{(s)}(u)du \le \text{P-}\lim_{n\to\infty} b_n^{-1} \sum_{i=1}^{l'} \left\{ -\frac{B_s^1 B_s^2 \lambda_i'}{3} (\rho_s - \rho_{s,*})^2 - \frac{1 - \bar{\rho}_T^2}{4R^8} (B_s^1 B_s^2 g_{i,*} g_i^{-1} - 1)^2 \right\}
$$

from (1.28),(1.29) and Lemma 1.6.

Moreover, since

$$
\begin{aligned}
(B_s^1 B_s^2 g_{i,*} g_i^{-1} - 1)^2 &\ge (B_s^1 B_s^2 g_{i,*} - g_i)^2 \ge g_{i,*}^2 (B_s^1 B_s^2 - 1)^2 / 2 - (g_i - g_{i,*})^2 \\
&= g_{i,*}^2 (B_s^1 B_s^2 - 1)^2 / 2 - (\lambda_i')^2 (\rho_s - \rho_{s,*})^2 (\rho_s + \rho_{s,*})^2 / (g_i + g_{i,*})^2 \\
&\ge (1 - \bar{\rho}_T^2)(B_s^1 B_s^2 - 1)^2 / 2 - \lambda_i' (\rho_s - \rho_{s,*})^2 / (1 - \bar{\rho}_T^2),
\end{aligned}
$$

we obtain

$$
\begin{aligned}
&\int_s^t f_1^{(s)}(u)du \\
\le \;& -\frac{B_s^1 B_s^2 (\rho_s - \rho_{s,*})^2}{3} \int_s^t a_1(u)du - \frac{1 - \bar{\rho}_T^2}{4R^8} \left\{ \frac{(1 - \bar{\rho}_T^2)(B_s^1 B_s^2 - 1)^2}{2} \int_s^t a_0(u)du - \frac{(\rho_s - \rho_{s,*})^2}{1 - \bar{\rho}_T^2} \int_s^t a_1(u)du \right\} \\
= \;& -\left(\frac{B_s^1 B_s^2}{3} - \frac{1}{4R^8} \right) (\rho_s - \rho_{s,*})^2 \int_s^t a_1(u)du - \frac{(1 - \bar{\rho}_T^2)^2}{8R^8} (B_s^1 B_s^2 - 1)^2 \int_s^t a_0(u)du \\
\le \;& -C_1 \int_s^t \left\{ a_1(u)(\rho_s - \rho_{s,*})^2 + a_0(u)(B_s^1 B_s^2 - 1)^2 \right\} du.
\end{aligned}
$$

Since $s < t$ is arbitrary, we obtain

$$
\int_s^t f_1(u, B_u^1 B_u^2, \rho_u, \rho_{u,*})du \le -C_1 \int_s^t \left\{ a_1(u)(\rho_u - \rho_{u,*})^2 + a_0(u)(B_u^1 B_u^2 - 1)^2 \right\} du.
$$

Then we have

$$
f_1(t, B_t^1 B_t^2, \rho_t, \rho_{t,*}) \le -C_1 \left\{ a_1(t)(\rho_t - \rho_{t,*})^2 + a_0(t)(B_t^1 B_t^2 - 1)^2 \right\} \quad dt \times P\text{- a.e. } (t, \omega).
$$

Similar argument using the eigenvalues of $G_{[s,t)}^\star G_{[s,t)}$ instead of that of $G_{[s,t)} G_{[s,t)}^\star$ yields

$$
f_2(t, B_t^1 B_t^2, \rho_t, \rho_{t,*}) \le -C_1 \left\{ a_1(t)(\rho_t - \rho_{t,*})^2 + c_0(t)(B_t^1 B_t^2 - 1)^2 \right\} \quad dt \times P\text{- a.e. } (t, \omega). \quad \square
$$

Proof of Lemma 1.8.

For the case that observation intervals $\{I\},\{J\}$ are synchronous and equi-spaced : $|I|=|J|=T/[b_n]$, we obtain $a_0\equiv c_0\equiv a_1\equiv 1$, $\mathcal{A}(\rho)=\rho^2/(1-\rho^2)$. Let us denote y_t by $y_{t,0}$ for the synchronous, equi-spaced sampling case, then by (1.9) we have

$$\begin{aligned} y_{t,0} &= -\frac{(B_t^1-B_t^2)^2}{2(1-\rho_t^2)}+1+\log B_t^1B_t^2+\frac{1}{2}\log\frac{1-\rho_{t,*}^2}{1-\rho_t^2}+B_t^1B_t^2\frac{\rho_t\rho_{t,*}-1}{1-\rho_t^2} \\ &= -\frac{(B_t^1-B_t^2)^2}{2(1-\rho_t^2)}+F\left(B_t^1B_t^2\sqrt{\frac{1-\rho_{t,*}^2}{1-\rho_t^2}}\right)+B_t^1B_t^2\left(\frac{\rho_t\rho_{t,*}-1}{1-\rho_t^2}+\sqrt{\frac{1-\rho_{t,*}^2}{1-\rho_t^2}}\right). \end{aligned}$$

Since $B_t^1B_t^2\sqrt{1-\rho_{t,*}^2}/\sqrt{1-\rho_t^2}\geq R^{-4}\sqrt{1-\bar{\rho}_T^2}$, by Lemma 1.6 and similar argument to (1.29), it follows that

$$y_{t,0}\geq -\frac{(B_t^1-B_t^2)^2}{2(1-\bar{\rho}_T^2)}-\left(\log\frac{R^4}{\sqrt{1-\bar{\rho}_T^2}}\vee 1\right)\left(B_t^1B_t^2\sqrt{\frac{1-\rho_{t,*}^2}{1-\rho_t^2}}-1\right)^2-R^4\frac{(\rho_t-\rho_{t,*})^2}{(1-\bar{\rho}_T^2)^2}.$$

Since

$$\begin{aligned} (B_t^1B_t^2\sqrt{1-\rho_{t,*}^2}/\sqrt{1-\rho_t^2}-1)^2 &\leq \frac{2(B_t^1B_t^2-1)^2+2(\sqrt{1-\rho_{t,*}^2}-\sqrt{1-\rho_t^2})^2}{1-\bar{\rho}_T^2} \\ &\leq \frac{2(B_t^1B_t^2-1)^2}{1-\bar{\rho}_T^2}+\frac{(\rho_{t,*}-\rho_t)^2(\rho_{t,*}+\rho_t)^2}{2(1-\bar{\rho}_T^2)^2}, \end{aligned}$$

there exists a positive random variable $\mathcal{R}'$ which does not depend on σ,σ_*,t such that $E[(\mathcal{R}')^q]<\infty$ for any $q>0$ and

$$y_{t,0}\geq -\mathcal{R}'\left\{(B_t^1-B_t^2)^2+(B_t^1B_t^2-1)^2+(\rho_t-\rho_{t,*})^2\right\}.$$

By integrating with respect to t, we have the desired conclusion. □

1.7.4 Proof of Proposition 1.5 and Theorem 1.2

Proof of Proposition 1.5.

We use Theorem 2 in Yoshida [54].

Let $\beta_1=\delta$, $\beta_2=1/2-\delta$, $0<\rho_2'<\delta, 0<\alpha<1\wedge(\rho_2'/2)$, $\beta=\alpha/(1-\alpha)$ and $0<\rho_1'<1\wedge\beta\wedge(2\beta_1/(1-\alpha))$. Let

$$\hat{\mathcal{Y}}_n(\sigma;\sigma_*)=b_n^{-1}(\hat{H}_n(\sigma)-\hat{H}_n(\sigma_*)),\quad \hat{\Gamma}_n(\sigma)=-b_n^{-1}\partial_\sigma^2\hat{H}_n(\sigma),$$

then it is sufficient to prove the following five conditions for any $L>0$.

1. There exists $c_L>0$ such that for any $r>0$, we have $P[\chi\leq r^{-(\rho_2'-2\alpha)}]\leq c_L/r^L$ and $P[\{r^{-\rho_1'}|u|^2\leq u^\star\Gamma u/4 \text{ for any } u\in\mathbb{R}^{n_1}\}^c]\leq c_L/r^L$.
2. For $M_1=L(1-\rho_1')^{-1}$, $\sup_n E[(b_n^{-1/2}|\partial_\sigma\hat{H}_n(\sigma_*)|)^{M_1}]<\infty$.
3. For $M_2=L(1-2\beta_2-\rho_2')^{-1}$,
$$\sup_n E\left[\left(\sup_\sigma b_n^{\frac{1}{2}-\beta_2}|\hat{\mathcal{Y}}_n(\sigma;\sigma_*)-\mathcal{Y}(\sigma;\sigma_*)|\right)^{M_2}\right]<\infty.$$
4. For $M_3=L(\beta-\rho_1')^{-1}$, $\sup_n E[(b_n^{-1}\sup_\sigma|\partial_\sigma^3\hat{H}_n(\sigma)|)^{M_3}]<\infty$.
5. For $M_4=L(2\beta_1/(1-\alpha)-\rho_1')^{-1}$, $\sup_n E[(b_n^{\beta_1}|\hat{\Gamma}_n(\sigma_*)-\Gamma|)^{M_4}]<\infty$.

By using Taylor's formula for $h_t^\infty(\sigma) - h_t^\infty(\sigma_*)$, we obtain $\chi \le \inf_{u\in\mathbb{R}^{n_1}\setminus\{0\}} u^\star\Gamma u/(2|u|^2)$. Then $[H]$ yields 1. Moreover, 3. and 5. obviously hold by Proposition 1.3 2. By Proposition 1.3 and the estimate $E[(\sup_\sigma|\int_0^T \partial_\sigma^3 h_t^\infty(\sigma)dt|)^{M_3}] < \infty$, 4. also holds. Finally, Lemma 1.14, $[S\text{-}q', 2q'\delta]$ for some sufficiently large q' and the estimate $\partial_\sigma \tilde{H}_n^3(T;\sigma_*) \equiv 0$ on $\{\tau(s_n) = T\}$ show 2. □

Proposition 1.10. *Assume $[A1]-[A4]$. Then $(\mathcal{V}_n(u_1),\cdots,\mathcal{V}_n(u_k)) \to^{s\text{-}\mathcal{L}} (\mathcal{V}(u_1),\cdots,\mathcal{V}(u_k))$ as $n\to\infty$ for $k\in\mathbb{N}$, $u_1,\cdots,u_k\in\mathbb{R}^{n_1}$, where $\mathcal{V}_n(u) = b_n^{-1/2}\partial_\sigma H_n(\sigma_*)u + b_n^{-1}u^\star\partial_\sigma^2 H_n(\sigma_*)u/2$, $\mathcal{V}(u) = u^\star\Gamma^{1/2}\mathcal{N} - u^\star\Gamma u/2$ and $\mathcal{N}$ is defined before the statement of Theorem 1.2. Moreover,*

$$\begin{aligned}\partial_\sigma^2 h_t^\infty(\sigma_*) &= \mathcal{A}(\rho_{t,*})\left(\frac{\partial_\sigma\rho_{t,*}}{\rho_{t,*}} - \partial_\sigma B_{t,*}^1 - \partial_\sigma B_{t,*}^2\right)^2 - \partial_\rho\mathcal{A}(\rho_{t,*})\frac{(\partial_\sigma\rho_{t,*})^2}{\rho_{t,*}} \\ &\quad -2(a_0(t)+\mathcal{A}(\rho_{t,*}))(\partial_\sigma B_{t,*}^1)^2 - 2(c_0(t)+\mathcal{A}(\rho_{t,*}))(\partial_\sigma B_{t,*}^2)^2.\end{aligned}$$

Proof. By (1.8) we have

$$\begin{aligned}\partial_\sigma h_t^\infty &= -\partial_\sigma B_t^1 B_t^1(a_0+\mathcal{A}(\rho_t)) - \frac{1}{2}(B_t^1)^2\partial_\sigma(\mathcal{A}(\rho_t)) - \partial_\sigma B_t^2 B_t^2(c_0+\mathcal{A}(\rho_t)) - \frac{1}{2}(B_t^2)^2\partial_\sigma(\mathcal{A}(\rho_t)) \\ &\quad +(\partial_\sigma B_t^1 B_t^2 + B_t^1\partial_\sigma B_t^2)\mathcal{A}\frac{\rho_{t,*}}{\rho_t} + B_t^1 B_t^2\rho_{t,*}\partial_\sigma\left(\frac{\mathcal{A}(\rho_t)}{\rho_t}\right) + a_0\frac{\partial_\sigma B_t^1}{B_t^1} + c_0\frac{\partial_\sigma B_t^2}{B_t^2} + \frac{\mathcal{A}(\rho_t)}{\rho_t}\partial_\sigma\rho_t \\ &= \partial_\sigma B_t^1\mathcal{A}(\rho_t)\left(B_t^2\frac{\rho_{t,*}}{\rho_t} - B_t^1\right) + \partial_\sigma B_t^2\mathcal{A}(\rho_t)\left(B_t^1\frac{\rho_{t,*}}{\rho_t} - B_t^2\right) + a_0\partial_\sigma B_t^1\left(\frac{1}{B_t^1} - B_t^1\right) \\ &\quad + c_0\partial_\sigma B_t^2\left(\frac{1}{B_t^2} - B_t^2\right) + \partial_\sigma(\mathcal{A}(\rho_t))\left(B_t^1 B_t^2\frac{\rho_{t,*}}{\rho_t} - \frac{(B_t^1)^2}{2} - \frac{(B_t^2)^2}{2}\right) + \mathcal{A}\frac{\partial_\sigma\rho_t}{\rho_t}\left(1 - B_t^1 B_t^2\frac{\rho_{t,*}}{\rho_t}\right).\end{aligned}$$

Since $B_{t,*}^1 = B_{t,*}^2 = 1$ and each term of the right-hand side of the previous equation has a factor which equals 0 if we substitute $\sigma = \sigma_*$, it follows that

$$\begin{aligned}\partial_\sigma^2 h_t^\infty(\sigma_*) &= (\partial_\sigma B_{t,*}^1\partial_\sigma B_{t,*}^2 + \partial_\sigma B_{t,*}^2\partial_\sigma B_{t,*}^1)\mathcal{A}_* - ((\partial_\sigma B_{t,*}^1)^2 + (\partial_\sigma B_{t,*}^2)^2)\mathcal{A}_* - (\partial_\sigma B_{t,*}^1 + \partial_\sigma B_{t,*}^2)\mathcal{A}_*\frac{\partial_\sigma\rho_{t,*}}{\rho_{t,*}} \\ &\quad -2a_0(\partial_\sigma B_{t,*}^1)^2 - 2c_0(\partial_\sigma B_{t,*}^2)^2 - \partial_\rho\mathcal{A}(\rho_{t,*})\frac{(\partial_\sigma\rho_{t,*})^2}{\rho_{t,*}} + \mathcal{A}_*\frac{\partial_\sigma\rho_{t,*}}{\rho_{t,*}}\left(\frac{\partial_\sigma\rho_{t,*}}{\rho_{t,*}} - \partial_\sigma B_{t,*}^1 - \partial_\sigma B_{t,*}^2\right) \\ &= \mathcal{A}_*\left(\frac{\partial_\sigma\rho_{t,*}}{\rho_{t,*}} - \partial_\sigma B_{t,*}^1 - \partial_\sigma B_{t,*}^2\right)^2 - \partial_\rho\mathcal{A}(\rho_{t,*})\frac{(\partial_\sigma\rho_{t,*})^2}{\rho_{t,*}} - 2(a_0+\mathcal{A}_*)(\partial_\sigma B_{t,*}^1)^2 - 2(c_0+\mathcal{A}_*)(\partial_\sigma B_{t,*}^2)^2,\end{aligned}$$

where $\mathcal{A}_* = \mathcal{A}(\rho_{t,*})$.

On the other hand, for $u\in\mathbb{R}^{n_1}$, let $s_n(t) = (1-\bar{\rho}_t)/2$, $\Upsilon_1 = b_n^{-1/2}(\partial_\sigma H(\sigma_*) - \partial_\sigma\hat{H}_n(\sigma_*;s_n))u$, $\Upsilon_2 = b_n^{-1/2}(\partial_\sigma\hat{H}_n(\sigma_*;s_n) + \sum_{i=1}^3(-1)^i\partial_\sigma\tilde{H}_{n,s_n}^i(T;\sigma_*))u$, $\Upsilon_3 = b_n^{-1}u^\star\partial_\sigma^2 H_n(\sigma_*)u/2 + u^\star\Gamma u/2$, $\Upsilon_4 = b_n^{-1/2}\partial_\sigma\tilde{H}_{n,s_n}^3(T;\sigma_*)$ and $\tilde{\mathcal{X}}_t = \tilde{\mathcal{X}}_{t,n}(u) = b_n^{-1/2}(\partial_\sigma\tilde{H}_{n,s_n}^1(t;\sigma_*) - \partial_\sigma\tilde{H}_{n,s_n}^2(t;\sigma_*))u$. Then

$$\mathcal{V}_n(u) = \tilde{\mathcal{X}}_{T,n}(u) - \frac{1}{2}u^\star\Gamma u + \sum_{j=1}^4\Upsilon_j.$$

As $n\to\infty$, since $P[\tau(s_n) < T]\to 0$, we have $\Upsilon_1 \to^p 0$. By $[A1]-[A4]$ and Lemmas 1.11 and 1.13 with $q = 2$, we have $\Upsilon_2 \to^p 0$. Furthermore, we obtain $\Upsilon_3 \to^p 0$ by Proposition 1.3. Moreover, $\Upsilon_4 \to^p 0$ since $P[\tau(s_n) < T]\to 0$ and $\partial_\sigma\tilde{H}_{n,s_n}^3(T;\sigma_*) \equiv 0$ on $\{\tau(s_n) = T\}$.

Then it is sufficient to show

$$\sum_{i=1}^k v_i(\tilde{\mathcal{X}}_{T,n}(u_i) - \frac{1}{2}u_i^\star\Gamma u_i) \to^{s\text{-}\mathcal{L}} \sum_{i=1}^k v_i\mathcal{V}(u_i)$$

as $n\to\infty$ for any $v_1,\cdots,v_k\in\mathbb{R}$ and $u_1,\cdots,u_k\in\mathbb{R}^{n_1}$.

Let $\mathcal{F}_t^\dagger = \cap_{t'>t}\{\mathcal{F}_{t'}\bigvee\sigma(\{\Pi_n\}_n)\}$ for $t\in[0,T)$ and $\mathcal{F}_T^\dagger=\mathcal{F}_T\bigvee\sigma(\{\Pi_n\}_n)$. Then $\{W_t,\mathcal{F}_t^\dagger\}_{0\leq t\leq T}$ is also a Wiener process and $\{\tilde{\mathcal{X}}_t(u),\mathcal{F}_t^\dagger\}_t$ is a martingale for $u\in\mathbb{R}^{n_1}$. By Theorem 2-1 of Jacod [23], it is sufficient to show that

$$\langle\tilde{\mathcal{X}}_{\cdot,n}(u)\rangle_t\to^p u^\star\Gamma_t u,\quad \langle\tilde{\mathcal{X}}_{\cdot,n}(u),W\rangle_t\to^p 0,\quad \langle\tilde{\mathcal{X}}_{\cdot,n}(u),N'\rangle_t\to^p 0$$

as $n\to\infty$ for any $t\in[0,T]$, $u\in\mathbb{R}^{n_1}$ and $N'\in\mathcal{M}_b(W^\perp)$, where $\Gamma_t=-\int_0^t\partial_\sigma^2 h_s^\infty(\sigma_*)ds$ and $\mathcal{M}_b(W^\perp)$ is the class of all bounded $\mathcal{F}_t^\dagger$-martingales which are orthogonal to W.

By Itô's formula and symmetry of $\tilde{M}$, we obtain

$$\tilde{\mathcal{X}}_t=-b_n^{-\frac{1}{2}}\sum_{k_1,k_2}\partial_\sigma\left\{\tilde{M}_{k_1,k_2}\int_0^t\hat{Z}_{k_1,s}d\hat{Z}_{k_2,s}\right\}\Big|_{\sigma=\sigma_*}u.$$

Hence it is obvious that $\langle\tilde{\mathcal{X}},N'\rangle_t=0$ for all $N'\in\mathcal{M}_b(W^\perp)$.

Moreover,

$$\begin{aligned}\langle\tilde{\mathcal{X}},W^i\rangle_t &= -b_n^{-\frac{1}{2}}\sum_{k_1,k_2}\sum_{v_1+v_2+v_3=1}\partial_\sigma^{v_1}\tilde{M}_{k_1,k_2}\int_0^t\partial_\sigma^{v_2}\hat{Z}_{k_1,s}d\langle\partial_\sigma^{v_3}\hat{Z}_{k_2},W^i\rangle_s u\\ &= -b_n^{-\frac{1}{2}}\sum_k\int_0^t\frac{\int_{(\theta_{0,k})_s}b_{v,*}^{\mathcal{J}(k)}dW_v}{\sqrt{|\theta_{0,k}|}}\mathcal{B}_{k,s}^i ds+o_p(1)\end{aligned}$$

for $i=1,2$, where $\mathcal{J}(k)=1$ $(1\leq k\leq l_n)$, $\mathcal{J}(k)=2$ $(l_n<k\leq l_n+m_n)$ and

$$\mathcal{B}_{k,s}^i=\sum_{k_2}\sum_{v_1+v_2+v_3=1}\partial_\sigma^{v_1}\tilde{M}_{k,k_2}\partial_\sigma^{v_2}\left(|b^{\mathcal{J}(k)}_{\theta_{0,k},\tau(s_n)}|^{-1}\right)b^{\mathcal{J}(k_2),i}_{L(\theta_{0,k}),*}\partial_\sigma^{v_3}\left(|b^{\mathcal{J}(k_2)}_{L(\theta_{0,k})\wedge\tau(s_n)}|^{-1}\right)\frac{1_{\theta_{0,k_2}}(s)}{\sqrt{|\theta_{0,k_2}|}}u.$$

On the other hand, we have

$$\begin{aligned}&E\left[\left|-b_n^{-\frac{1}{2}}\sum_k\int_0^t\frac{\int_{(\theta_{0,k})_s}b_{v,*}^{\mathcal{J}(k)}dW_v}{\sqrt{|\theta_{0,k}|}}\mathcal{B}_{k,s}^i ds\right|^2\right]\\ =\ & b_n^{-1}E\left[\sum_{k,k'}\int_0^t\int_0^t\frac{\int_{(\theta_{0,k})_{s_1}\cap(\theta_{0,k'})_{s_2}}b_{v,*}^{\mathcal{J}(k)}\cdot b_{v,*}^{\mathcal{J}(k')}dv}{\sqrt{|\theta_{0,k}|}\sqrt{|\theta_{0,k'}|}}\mathcal{B}_{k,s_1}^i\mathcal{B}_{k',s_2}^i ds_1ds_2\right]\\ \leq\ & b_n^{-1}E\left[R^C\sum_{k,k'}(\tilde{M}_0)_{k,k'}\int_0^t\int_0^t|\mathcal{B}_{k,s_1}^i||\mathcal{B}_{k',s_2}^i|ds_1ds_2\right]\leq b_n^{-1}E\left[R^C\sum_k\left(\int_0^t|\mathcal{B}_{k,s}^i|ds\right)^2\right]\to 0\end{aligned}$$

as $n\to\infty$ since $|\partial_\sigma^v(\tilde{M})_{k,k'}|\leq CR^{2v}(1-\bar{\rho}_T)^{-(v+5/2)}M'_{k,k'}$, where $M'_{k,k'}=\sum_{p=0}^\infty(\tilde{M}_p)_{k,k'}/(p+1)^2$. Hence we have $\langle\tilde{\mathcal{X}},W\rangle_t\to^p 0$ as $n\to\infty$ for any $t\in[0,T]$.

Then it is sufficient to show $\langle\tilde{\mathcal{X}}(u)\rangle_t\to^p u^\star\Gamma_t u$ as $n\to\infty$ for any $t\in[0,T]$ and $u\in\mathbb{R}^{n_1}$.

$$\begin{aligned}&\langle\tilde{\mathcal{X}}\rangle_t\\ =\ & b_n^{-1}u^\star\sum_{k_1,k_2,k_3,k_4}\int_0^t\partial_\sigma\left(\tilde{M}_{k_1,k_2}\hat{Z}_{k_1,s}\frac{b_{s,*}^{\mathcal{J}(k_2)}}{|b^{\mathcal{J}(k_2)}_{\theta_{0,k_2},\tau(s_n)}|}\right)\partial_\sigma\left(\tilde{M}_{k_3,k_4}\hat{Z}_{k_3,s}\frac{b_{s,*}^{\mathcal{J}(k_4)}}{|b^{\mathcal{J}(k_4)}_{\theta_{0,k_4},\tau(s_n)}|}\right)\frac{1_{\theta_{0,k_2}\cap\theta_{0,k_4}}(s)}{\sqrt{|\theta_{0,k_2}|}\sqrt{|\theta_{0,k_4}|}}ds\Big|_{\sigma=\sigma_*}u\\ =\ & b_n^{-1}\sum_{k_1,k_2,k_3,k_4}\int_0^t\tilde{\mathcal{B}}_{k_1,k_2}\cdot\tilde{\mathcal{B}}_{k_3,k_4}\frac{Z'_{k_1,s}Z'_{k_3,s}}{\sqrt{|\theta_{0,k_1}|}\sqrt{|\theta_{0,k_3}|}}\frac{1_{\theta_{0,k_2}\cap\theta_{0,k_4}}(s)}{\sqrt{|\theta_{0,k_2}|}\sqrt{|\theta_{0,k_4}|}}ds+o_p(1),\end{aligned}\tag{1.30}$$

where $\tilde{\mathcal{B}}_{k_1,k_2}=\partial_\sigma\big(\tilde{M}_{k_1,k_2}|b^{\mathcal{J}(k_2)}_{L(\theta_{0,k_1})\wedge\tau(s_n)}|^{-1}|b^{\mathcal{J}(k_1)}_{\theta_{0,k_1},\tau(s_n)}|^{-1}\big)|_{\sigma=\sigma_*}b^{\mathcal{J}(k_2)}_{L(\theta_{0,k_1}),*}u$ and $Z'_{k,s}=\int_{(\theta_{0,k})_s}b_{v,*}^{\mathcal{J}(k)}dW_v$.

Itô's formula yields

$$Z'_{k_1,s}Z'_{k_3,s}=\int_0^s Z'_{k_1,v}dZ'_{k_3,v}+\int_0^s Z'_{k_3,v}dZ'_{k_1,v}+\langle Z'_{k_1},Z'_{k_3}\rangle_s.\tag{1.31}$$

Moreover, let

$$F_k(v,s)=\sum_{k_1,k_2,k_4}\tilde{\mathcal{B}}_{k_1,k_2}\cdot\tilde{\mathcal{B}}_{k,k_4}\frac{Z'_{k_1,v}}{\sqrt{|\theta_{0,k_1}|}\sqrt{|\theta_{0,k}|}}\frac{1_{\theta_{0,k_2}\cap\theta_{0,k_4}}(s)}{\sqrt{|\theta_{0,k_2}|}\sqrt{|\theta_{0,k_4}|}},$$

then we have

$$\sup_v|F_k(v,s)|\leq CR^{10}(1-\bar{\rho}_T)^{-7}\sum_{k_1,k_2}\sum_{k_4}\frac{M'_{k_1,k_2}M'_{k,k_4}1_{\theta_{0,k_2}\cap\theta_{0,k_4}}(s)}{\sqrt{|\theta_{0,k_1}|}\sqrt{|\theta_{0,k_2}|}\sqrt{|\theta_{0,k}|}\sqrt{|\theta_{0,k_4}|}}\sup_v|Z'_{k_1,v}||u|^2$$

and therefore

$$\int_0^t E[\sup_v|F_k(v,s)|^4|\Pi_n]^{\frac{1}{4}}ds\leq C|u|^2|\theta_{0,k}|^{-1/2}\sum_{k_1}(M'\tilde{M}_0M')_{k_1,k}.$$

Then we obtain

$$\begin{aligned}
&E\Bigg[\bigg(\sum_k\int_0^t\int_0^s F_k(v,s)dZ'_{k,v}ds\bigg)^2\bigg|\Pi_n\Bigg]\\
=\ &\sum_{k,k'}\int_0^t\int_0^t E\bigg[\int_0^{s_1\wedge s_2}F_k(v,s_1)F_{k'}(v,s_2)d\langle Z'_k,Z'_{k'}\rangle_v\bigg|\Pi_n\bigg]ds_1ds_2\\
\leq\ &E[R^4]^{\frac{1}{2}}\sum_{k,k'}|\theta_{0,k}\cap\theta_{0,k'}|\int_0^t E[\sup_v|F_k(v,s_1)|^4|\Pi_n]^{\frac{1}{4}}ds_1\int_0^t E[\sup_v|F_{k'}(v,s_2)|^4|\Pi_n]^{\frac{1}{4}}ds_2\\
\leq\ &C|u|^4\sum_{k,k'}(\tilde{M}_0)_{k,k'}\sum_{k_1,k_1'}(M'\tilde{M}_0M')_{k_1,k}(M'\tilde{M}_0M')_{k_1',k'}\leq C|u|^4(l_n+m_n)=o_p(b_n^2).
\end{aligned}$$

Hence we obtain

$$b_n^{-1}\sum_k\int_0^t\int_0^s F_k(v,s)dZ'_{k,v}ds\to^p 0 \tag{1.32}$$

as $n\to\infty$ by Lemma 1.11.

By (1.30)-(1.32), we have

$$\langle\tilde{\mathcal{X}}\rangle_t=b_n^{-1}\sum_{k_1,k_2,k_3,k_4}\int_0^t\tilde{\mathcal{B}}_{k_1,k_2}\cdot\tilde{\mathcal{B}}_{k_3,k_4}\frac{\langle Z'_{k_1},Z'_{k_3}\rangle_s}{\sqrt{|\theta_{0,k_1}|}\sqrt{|\theta_{0,k_3}|}}\frac{1_{\theta_{0,k_2}\cap\theta_{0,k_4}}(s)}{\sqrt{|\theta_{0,k_2}|}\sqrt{|\theta_{0,k_4}|}}ds+o_p(1).$$

Let

$$\hat{L}_p(\rho_1,\rho_2)=\begin{pmatrix}\rho_1(GG^\star)^p & -\rho_2(GG^\star)^pG\\ -\rho_2(G^\star G)^pG^\star & \rho_1(G^\star G)^p\end{pmatrix},\hat{B}(x,y)=\begin{pmatrix}x\mathcal{E}_{l_n} & 0\\ 0 & y\mathcal{E}_{m_n}\end{pmatrix},$$

$$\mathcal{D}'(t)=\sum_{p=0}^\infty\partial_\sigma\Big\{\hat{B}(B_t^1,B_t^2)\hat{L}_p(\rho_t^{2p},\rho_t^{2p+1})\hat{B}(B_t^1,B_t^2)\Big\}\Big|_{\sigma=\sigma_*}$$

and

$$\hat{\mathcal{D}}(t)\ =\ \begin{pmatrix}\hat{\mathcal{D}}_{11}(t) & \hat{\mathcal{D}}_{12}(t)\\ \hat{\mathcal{D}}_{21}(t) & \hat{\mathcal{D}}_{22}(t)\end{pmatrix}=\mathcal{D}'(t)\hat{L}_0(1,-\rho_{t,*}),$$

where $\mathcal{E}_l$ denotes the unit matrix of size l. Then by $[A2]$ and the estimate $P[\tau(s_n)<T]\to 0$, we obtain

$$\begin{aligned}
\langle\tilde{\mathcal{X}}\rangle_t\ =\ &b_n^{-1}u^\star\sum_{k_1,k_2,k_3,k_4}(\mathcal{D}'(L(\theta_{0,k_1})))_{k_1,k_2}(\mathcal{D}'(L(\theta_{0,k_1})))_{k_3,k_4}\frac{b^{\mathcal{J}(k_1)}_{L(\theta_{0,k_1}),*}\cdot b^{\mathcal{J}(k_3)}_{L(\theta_{0,k_1}),*}}{|b^{\mathcal{J}(k_1)}_{L(\theta_{0,k_1}),*}||b^{\mathcal{J}(k_3)}_{L(\theta_{0,k_1}),*}|}\\
&\times\frac{b^{\mathcal{J}(k_2)}_{L(\theta_{0,k_1}),*}\cdot b^{\mathcal{J}(k_4)}_{L(\theta_{0,k_1}),*}}{|b^{\mathcal{J}(k_2)}_{L(\theta_{0,k_1}),*}||b^{\mathcal{J}(k_4)}_{L(\theta_{0,k_1}),*}|}\frac{\int_0^t\int_0^s 1_{\theta_{0,k_1}\cap\theta_{0,k_3}}(v)1_{\theta_{0,k_2}\cap\theta_{0,k_4}}(s)dvds}{\sqrt{|\theta_{0,k_1}|}\sqrt{|\theta_{0,k_2}|}\sqrt{|\theta_{0,k_3}|}\sqrt{|\theta_{0,k_4}|}}u+o_p(1).
\end{aligned}$$

Since for intervals K_1, K_2, we have

$$\int_0^t\int_0^s 1_{K_1}(v)1_{K_2}(s)dvds + \int_0^t\int_0^s 1_{K_2}(v)1_{K_1}(s)dvds = |(K_1)_t||(K_2)_t|,$$

then by symmetry of $\mathcal{D}'$, we have

$$\begin{aligned}\langle\tilde{\mathcal{X}}\rangle_t &= \frac{1}{2}b_n^{-1}u^\star\sum_{k_1,k_2,k_3,k_4}(\mathcal{D}'(L(\theta_{0,k_1})))_{k_1,k_2}(\mathcal{D}'(L(\theta_{0,k_1})))_{k_3,k_4}\frac{b^{\mathcal{J}(k_1)}_{L(\theta_{0,k_1}),*}\cdot b^{\mathcal{J}(k_3)}_{L(\theta_{0,k_1}),*}}{|b^{\mathcal{J}(k_1)}_{L(\theta_{0,k_1}),*}||b^{\mathcal{J}(k_3)}_{L(\theta_{0,k_1}),*}|}\\ &\quad\times\frac{b^{\mathcal{J}(k_2)}_{L(\theta_{0,k_1}),*}\cdot b^{\mathcal{J}(k_4)}_{L(\theta_{0,k_1}),*}}{|b^{\mathcal{J}(k_2)}_{L(\theta_{0,k_1}),*}||b^{\mathcal{J}(k_4)}_{L(\theta_{0,k_1}),*}|}\frac{|(\theta_{0,k_1}\cap\theta_{0,k_3})_t||(\theta_{0,k_2}\cap\theta_{0,k_4})_t|}{\sqrt{|\theta_{0,k_1}|}\sqrt{|\theta_{0,k_2}|}\sqrt{|\theta_{0,k_3}|}\sqrt{|\theta_{0,k_4}|}}u+o_p(1)\\ &= \frac{1}{2}b_n^{-1}u^\star\sum_{k;L(\theta_{0,k})\in[0,t)}((\hat{\mathcal{D}}(L(\theta_{0,k})))^2)_{k,k}u+o_p(1).\end{aligned}$$

On the other hand, for $p\in\mathbb{Z}_+$, $x,y\in\mathbb{R}$, $\rho_1,\rho_2,\rho_*\in[-1,1]$, we can write

$$\hat{B}(x,y)\hat{L}_p(\rho_1,\rho_2)\hat{B}(x,y)\hat{L}_0(1,-\rho_*) = \begin{pmatrix} x^2\rho_1(GG^\star)^p - xy\rho_2\rho_*(GG^\star)^{p+1} & (x^2\rho_1\rho_* - xy\rho_2)(GG^\star)^pG \\ (y^2\rho_1\rho_* - xy\rho_2)(G^\star G)^pG^\star & y^2\rho_1(GG^\star)^p - xy\rho_2\rho_*(G^\star G)^{p+1}\end{pmatrix}.$$

Then for $\mathcal{Q}_t^1 = (\partial_\sigma B^1_{t,*} - \partial_\sigma B^2_{t,*}) + \partial_\sigma\rho_{t,*}/\rho_{t,*}$, $\mathcal{Q}_t^2 = (\partial_\sigma B^2_{t,*} - \partial_\sigma B^1_{t,*}) + \partial_\sigma\rho_{t,*}/\rho_{t,*}$, we have

$$\begin{aligned}&\hat{\mathcal{D}}_{11}(t)\\ &= \sum_{p=0}^\infty\Big\{(2\partial_\sigma B^1_{t,*}\rho_{t,*}^{2p} + 2p\partial_\sigma\rho_{t,*}\rho_{t,*}^{2p-1})(GG^\star)^p - ((\partial_\sigma B^1_{t,*}+\partial_\sigma B^2_{t,*})\rho_{t,*}^{2p+2} + (2p+1)\partial_\sigma\rho_{t,*}\rho_{t,*}^{2p+1})(GG^\star)^{p+1}\Big\}\\ &= 2\partial_\sigma B^1_{t,*}\mathcal{E}_{l_n} + \mathcal{Q}_t^1\sum_{p=1}^\infty\rho_{t,*}^{2p}(GG^\star)^p.\end{aligned}$$

Similarly, we have

$$\hat{\mathcal{D}}_{22}(t) = 2\partial_\sigma B^2_{t,*}\mathcal{E}_{m_n} + \mathcal{Q}_t^2\sum_{p=1}^\infty\rho_{t,*}^{2p}(G^\star G)^p,$$

and

$$\hat{\mathcal{D}}_{12}(t) = -\mathcal{Q}_t^2\sum_{p=0}^\infty\rho_{t,*}^{2p+1}(GG^\star)^pG,\quad \hat{\mathcal{D}}_{21}(t) = -\mathcal{Q}_t^1\sum_{p=0}^\infty\rho_{t,*}^{2p+1}(G^\star G)^pG^\star.$$

Then by the estimate $a_p\equiv c_p$ $(p\geq 1)$, $[A3]$ and Lemma 1.15, it follows that

$$\begin{aligned}\langle\tilde{\mathcal{X}}\rangle_t &= \frac{1}{2}b_n^{-1}u^\star\Big\{\sum_{i;L(I^i)\in[0,t)}(\hat{\mathcal{D}}_{11}^2+\hat{\mathcal{D}}_{12}\hat{\mathcal{D}}_{21})_{ii}(L(I^i)) + \sum_{j;L(J^j)\in[0,t)}(\hat{\mathcal{D}}_{22}^2+\hat{\mathcal{D}}_{21}\hat{\mathcal{D}}_{12})_{jj}(L(J^j))\Big\}u+o_p(1)\\ &= u^\star\int_0^t\Big\{2(\partial_\sigma B^1_{s,*})^2a_0(s) + 2(\partial_\sigma B^2_{s,*})^2c_0(s) + (\partial_\sigma B^1_{s,*}\mathcal{Q}_s^1 + \mathcal{Q}_s^1\partial_\sigma B^1_{s,*})\mathcal{A}(\rho_{s,*})\\ &\quad+(\partial_\sigma B^2_{s,*}\mathcal{Q}_s^2 + \mathcal{Q}_s^2\partial_\sigma B^2_{s,*})\mathcal{A}(\rho_{s,*}) + \frac{\mathcal{Q}_s^1\mathcal{Q}_s^2+\mathcal{Q}_s^2\mathcal{Q}_s^1}{2}\sum_{p_1=0}^\infty\sum_{p_2=0}^\infty\rho_{s,*}^{2p_1+2p_2+2}a_{p_1+p_2+1}(s)\\ &\quad+\frac{(\mathcal{Q}_s^1)^2+(\mathcal{Q}_s^2)^2}{2}\sum_{p_1=1}^\infty\sum_{p_2=1}^\infty\rho_{s,*}^{2p_1+2p_2}a_{p_1+p_2}(s)\Big\}dsu+o_p(1).\end{aligned}$$

Since

$$\sum_{p_1=0}^\infty\sum_{p_2=0}^\infty\rho_{s,*}^{2p_1+2p_2+2}a_{p_1+p_2+1}(s) = \frac{\partial_\rho\mathcal{A}(\rho_{s,*})\rho_{s,*}}{2},\quad \sum_{p_1=1}^\infty\sum_{p_2=1}^\infty\rho_{s,*}^{2p_1+2p_2}a_{p_1+p_2}(s) = \frac{\partial_\rho\mathcal{A}(\rho_{s,*})\rho_{s,*}}{2} - \mathcal{A}(\rho_{s,*}),$$

we have

$$\begin{aligned}\langle\tilde{\mathcal{X}}\rangle_t &= u^\star\int_0^t\Big\{2(a_0(s)+\mathcal{A}(\rho_{s,*}))(\partial_\sigma B^1_{s,*})^2+2(c_0(s)+\mathcal{A}(\rho_{s,*}))(\partial_\sigma B^2_{s,*})^2+\frac{(\mathcal{Q}^1_s+\mathcal{Q}^2_s)^2}{4}\partial_\rho\mathcal{A}(\rho_{s,*})\rho_{s,*}\\ &\quad -\frac{\mathcal{A}(\rho_{s,*})}{2}\{(\mathcal{Q}^1_s-2\partial_\sigma B^1_{s,*})^2+(\mathcal{Q}^2_s-2\partial_\sigma B^2_{s,*})^2\}\Big\}dsu+o_p(1)\\ &= u^\star\Gamma_t u+o_p(1).\end{aligned}$$

□

Proof of Theorem 1.2.

1. Since Λ is open, there exists $\epsilon>0$ such that $O(\epsilon,\sigma_*)=\{\sigma;|\sigma-\sigma_*|<\epsilon\}\subset\Lambda$. For $\hat{\sigma}_n\in O(\epsilon,\sigma_*)$,we have

$$-\partial_\sigma H_n(\sigma_*)=\int_0^1\partial_\sigma^2 H_n(\sigma_*+u(\hat{\sigma}_n-\sigma_*))(\hat{\sigma}_n-\sigma_*)du$$

since $\partial_\sigma H_n(\hat{\sigma}_n)=0$. Therefore, for $\tilde{\Gamma}_n=-b_n^{-1}\int_0^1\partial_\sigma^2 H_n(\sigma_*+u(\hat{\sigma}_n-\sigma_*))du$, we obtain $b_n^{1/2}(\hat{\sigma}_n-\sigma_*)=\tilde{\Gamma}_n^{-1}b_n^{-1/2}\partial_\sigma H_n(\sigma_*)$ on $\{\det\tilde{\Gamma}_n\neq 0$ and $\hat{\sigma}_n\in O(\epsilon,\sigma_*)\}$. Then since Proposition 1.3 and Theorem 1.1 yield $P[\det\tilde{\Gamma}_n=0]\to 0$, $P[\hat{\sigma}_n\in O(\epsilon,\sigma_*)^c]\to 0$ and $\tilde{\Gamma}_n^{-1}1_{\{\det\tilde{\Gamma}_n\neq 0\}}\to^p\Gamma^{-1}$, we have $b_n^{1/2}(\hat{\sigma}_n-\sigma_*)\to^{s\text{-}\mathcal{L}}\Gamma^{-1/2}\mathcal{N}$ by Proposition 1.10.

2. Let $s_n(t)=(1-\bar{\rho}_t)/2$ for $n\in\mathbb{N}$ and $t\in[0,T]$ and $\{\sigma'_n\}_{n\in\mathbb{N}}$ be random variables where σ'_n maximizes $\hat{H}_n(\cdot;s_n)$ and $\sigma'_n\equiv\hat{\sigma}_n$ on $\{\tau(s_n)=T\}$. We first show the statement of Theorem 1.2 replacing $\hat{\sigma}_n$ with σ'_n.

To this end, we extend $\mathcal{Z}_n(\cdot;\sigma_*)$ to a continuous function which is defined on $\mathbb{R}^{n_1}$, tend to zero as $|u|\to\infty$, and has the same supremum as $\mathcal{Z}_n(\cdot;\sigma_*)$. We denote the extension of $\mathcal{Z}_n(\cdot;\sigma_*)$ by the same symbol.

Let $\mathcal{Z}(u,\sigma_*)=\exp(u^\star\Gamma^{1/2}\mathcal{N}-u^\star\Gamma u/2)$ and $B(R')=\{u;|u|\leq R'\}$ for $R'>0$. Then it is sufficient to show that $\limsup_{n\to\infty}E[|b_n^{1/2}(\sigma'_n-\sigma_*)|^p]<\infty$ for any $p>2$ and $\mathcal{Z}_n(\cdot,\sigma_*)\to^{s\text{-}\mathcal{L}}\mathcal{Z}(\cdot,\sigma_*)$ in $C(B(R'))$ as $n\to\infty$ for any $R'>0$, by virtue of Theorem 5 and Remark 5 in Yoshida [54].

By Lemmas 1.14 and 1.1 and Proposition 1.3, for any $R'>0$, there exists $n_0\in\mathbb{N}$ such that

$$\sup_{n\geq n_0}E\left[\sup_{u\in C(B(R'))}|\partial_u\log\mathcal{Z}_n(u;\sigma_*)|\right]<\infty.$$

Then by Propositions 1.3 and 1.10 and tightness criterion in C space in Billingsley [8] which can be extended to the one in $C(B(R'))$, it follows that $\log\mathcal{Z}_n(\cdot,\sigma_*)\to^{s\text{-}\mathcal{L}}\log\mathcal{Z}(\cdot,\sigma_*)$ in $C(B(R'))$ as $n\to\infty$.

On the other hand, for any $p>2$, let $L>p$, then by Proposition 1.5 and Lemma 1.1, we have

$$P[|b_n^{\frac{1}{2}}(\sigma'_n-\sigma_*)|\geq r]\leq P\left[\sup_{u\in V_n(r,\sigma_*)}\mathcal{Z}_n(u,\sigma_*)\geq 1\right]\leq\frac{C_L}{r^L}\quad(r>0).$$

Therefore we obtain $\sup_n E[|b_n^{1/2}(\sigma'_n-\sigma_*)|^p]<\infty$. This complete the proof of the statement of Theorem 1.2 for σ'_n.

We will prove the statement for $\hat{\sigma}_n$. By $[A1],[A2\text{-}q,\delta]$ for any $q>2\vee n_1$, and Lemma 1.1, we have $P[\tau(s_n)<T]=O(b_n^{-\xi})$ for any $\xi>0$. Then it follows that $b_n^{1/2}(\hat{\sigma}_n-\sigma_*)\to^{s\text{-}\mathcal{L}}\Gamma^{-1/2}\mathcal{N}$ as $n\to\infty$ by the result for σ'_n and the inequality

$$P[\sigma'_n\neq\hat{\sigma}_n]\leq P[\tau(s_n)<T]=O(b_n^{-\xi})$$

for any $\xi>0$.

Moreover, for any continuous function f of at most polynomial growth, we have

$$|E[\mathsf{f}(b_n^{1/2}(\hat{\sigma}_n-\sigma_*))]-E[\mathsf{f}(b_n^{1/2}(\sigma'_n-\sigma_*))]|\leq C(1+b_n^{1/2}R'')^C P[\sigma'_n\neq\hat{\sigma}_n]\to 0$$

as $n\to\infty$, where R'' denotes the diameter of the parameter space Λ. □

Proof of Theorem 1.3. Similarly to the argument in the proof of Theorem 1.2, we have $P[H_n \equiv \hat{H}_n(\cdot; s_n)] = 1 - O(b_n^{-\xi})$ for any $\xi > 0$, where $s_n(t) = (1-\bar{\rho}_t)/2$. Then by virtue of Theorem 10 in Yoshida [54], it is sufficient to show that there exists $n_0' \in \mathbb{N}$ such that

$$\sup_{n \geq n_0'} E\bigg[\bigg(\int_{U_n(\sigma_*)} \mathcal{Z}_n(u)\pi(\sigma_* + b_n^{-1/2}u)du\bigg)^{-1}\bigg] < \infty. \tag{1.33}$$

By Proposition 1.3 and Lemmas 1.14 and 1.1, for any $\delta > 0$, there exists $p \in 2\mathbb{N}, p > n_1 \vee 2$, $n_0' \in \mathbb{N}$ and $C_0 > 0$ such that

$$\sup_{n \geq n_0'} E[|\hat{H}_n(\sigma_* + b_n^{-1/2}u) - \hat{H}_n(\sigma_*)|^p] \leq C_0|u|^p$$

for any $u \in U'(\delta)$ where $U'(\delta) = \{u \in \mathbb{R}^{n_1}; |u_i| \leq \delta \ (i = 1, \ldots, n_1)\}$. Then we have (1.33) by Lemma 2 in Yoshida [54]. □

1.7.5 Proof of Propositions 1.6 - 1.9

First, we look back Rosenthal-type inequalities in Doukhan and Louhichi [12] (Theorem 3 and Lemma 7).

Theorem 1.4. *(Rosenthal-type inequalities) Let $q \geq 2$ and $q \in \mathbb{N}$. Let $\{X_n'\}_{n \in \mathbb{N}}$ be a centered process, $\alpha_0 = 1/4$ and*

$$\alpha_k = \sup_{i,j \in \mathbb{N}, j-i \geq k} \sup_{A \in \sigma(X_l'; l \leq i)} \sup_{B \in \sigma(X_m'; m \geq j)} |P(A \cap B) - P(A)P(B)|$$

for $k \in \mathbb{N}$. Suppose $\alpha_k \to 0$ $(k \to \infty)$. Then

$$\bigg|E\bigg[\bigg(\sum_{j=1}^n X_j'\bigg)^q\bigg]\bigg| \leq \frac{2^q(2q-2)!}{(q-1)!}\bigg\{\bigg(\sum_{i=1}^n \int_0^1 (\alpha^{-1}(u) \wedge n)^{q-1} Q_{X_i'}^q(u)du\bigg) \vee \bigg(\sum_{i=1}^n \int_0^1 (\alpha^{-1}(u) \wedge n) Q_{X_i'}^2(u)du\bigg)^{\frac{q}{2}}\bigg\},$$

where $\alpha^{-1}(u) = \sum_{k=0}^\infty 1_{\{\alpha_k > u\}}$ and $Q_{X'}(s) = \inf\{t > 0, P[|X'| > t] \leq s\}$.

Proof of Proposition 1.6.

In this proof, general constants denoted by C do not depend on n, p, f.

By Lemma 1.2, we obtain

$$((GG^\star)^p)_{II} \vee ((G^\star G)^p)_{JJ} \leq \| (GG^\star)^p \| \vee \| (G^\star G)^p \| \leq 1$$

for $p \in \mathbb{Z}_+$. Hence $b_n\nu_n^{p,i}([t_{k-1}, t_k)) \leq N_{t_k}^i - N_{t_{k-1}}^i + 1$ for $1 \leq k \leq [b_n]$, $p \in \mathbb{Z}_+$ and $i = 1, 2$. Therefore we obtain

$$\sup_{1 \leq k \leq [b_n]} E[\max_{p \in \mathbb{Z}_+, i=1,2} |b_n\nu_n^{p,i}([t_{k-1}, t_k))|^{q(1+\delta)}] \leq \sup_k E[\max_{i=1,2}(N_{t_k}^i - N_{t_{k-1}}^i + 1)^{q(1+\delta)}] \leq C \tag{1.34}$$

for sufficiently large n by $[B1\text{-}(q(1+\delta))]$.

For $h > 0$ and $k \in \mathbb{N}$, let

$$\begin{aligned} A_{h,t}^{p,+} &= \cap_{i=1,2} \cap_{l \in [1, p \wedge h^{-1}(T-t)] \cap \mathbb{N}} \{\omega; N_{t+lh}^i - N_{t+(l-1)h}^i > 0\}, \\ A_{h,t}^{p,-} &= \cap_{i=1,2} \cap_{l \in [1, p \wedge h^{-1}t] \cap \mathbb{N}} \{\omega; N_{t-(l-1)h}^i - N_{t-lh}^i > 0\}, \\ A_{k,h}^p &:= A_{[b_n]^{-1}hT, t_k}^{2p+1,+} \cap A_{[b_n]^{-1}hT, t_{k-1}}^{2p+1,-}, \end{aligned}$$

where $\cap_\emptyset = \Omega$.

Fix $p \in \mathbb{Z}_+$, $i = 1, 2$ and a $\beta-$Hölder continous function f on $[0, T]$. Then we have

$$\nu_n^{p,i}([t_{k-1}, t_k))1_{A_{k,h}^p} \in \mathcal{G}_{(k-2-[(2p+1)h]) \vee 0, (k+[(2p+1)h]+1) \wedge [b_n]}^n.$$

Let $\alpha^{-1}(u) = \sum_{k=0}^\infty 1_{\{\alpha_k^n > u\}}$, $f_k^n = f_{t_{k-1}^n}$, $\delta' = (1+\delta)/(2(1+\delta-\epsilon\delta))$ and

$$X_k' = b_n f_k^n \{\nu_n^{p,i}([t_{k-1}, t_k))1_{A_{k,b_n^{\delta'}}^p} - E[\nu_n^{p,i}([t_{k-1}, t_k))1_{A_{k,b_n^{\delta'}}^p}]\},$$

then by Rosenthal-type inequalities, we obtain

$$\begin{aligned} E\left[\left|b_n^{-1}\sum_{k=1}^{[b_n]}X'_k\right|^q\right] &\leq b_n^{-q}\frac{2^{q/2}(2q-2)!}{(q-1)!}\Big\{\Big(\sum_{k=1}^{[b_n]}\int_0^{\frac{1}{4}}(\alpha^{-1}(u)+2[(2p+1)b_n^{\delta'}]+3)^{q-1}Q^q_{X'_k}(u)du\Big) \\ &\quad \vee\Big(\sum_{k=1}^{[b_n]}\int_0^{\frac{1}{4}}(\alpha^{-1}(u)+2[(2p+1)b_n^{\delta'}]+3)Q^2_{X'_k}(u)du\Big)^{\frac{q}{2}}\Big\} \\ &\leq Cb_n^{-q}[b_n]^{q/2}\sup_k\int_0^{\frac{1}{4}}(\alpha^{-1}(u)+2[(2p+1)b_n^{\delta'}]+3)^{q-1}Q^q_{X'_k}(u)du \\ &\leq C(p+1)^{q-1}b_n^{q\delta'-\frac{q}{2}}\Big(\int_0^1(\alpha^{-1}(u))^{\frac{(1+\delta)(q-1)}{\delta}}du\Big)^{\frac{\delta}{1+\delta}}\Big(\sup_k\int_0^1 Q^{q(1+\delta)}_{X'_k}(u)du\Big)^{\frac{1}{1+\delta}}. \end{aligned}$$

For sufficiently large n, since (1.12) and (1.34) hold, $\int_0^1 Q^{q(1+\delta')}_{X'_k}(u)du = E[|X'_k|^{q(1+\delta)}]$, $(x+1)^{q'}-x^{q'}\leq q'(x+1)^{q'-1}$ $(x\geq 0, q'\geq 1)$ and $\alpha^{-1}(u)=k'$ if $\alpha^n_{k'}\leq u<\alpha^n_{k'-1}$, we have

$$\int_0^1(\alpha^{-1}(u))^{q'}du = \sum_{k=1}^\infty k^{q'}(\alpha^n_{k-1}-\alpha^n_k)\leq q'\sum_{k=0}^\infty(k+1)^{q'-1}\alpha^n_k$$

for $q'\geq 1$ and

$$E\left[\left|b_n^{-1}\sum_{k=1}^{[b_n]}X'_k\right|^q\right]\leq C(p+1)^{q-1}b_n^{q\delta'-\frac{q}{2}}\sup_t|f_t|^q.$$

On the other hand,

$$E\Big[\Big|\sum_{k=1}^{[b_n]}f^n_k\nu^{p,i}_n([t_{k-1},t_k))1_{(A^p_{k,b_n^{\delta'}})^c}\Big|^q\Big]\leq[b_n]^{-1}\sum_{k=1}^{[b_n]}\sup_k E[|N^i_{t_k}-N^i_{t_{k-1}}+1|^{q(1+\delta)}]^{\frac{1}{1+\delta}}\sup_t|f_t|^qP[(A^p_{k,b_n^{\delta'}})^c]^{\frac{\delta}{1+\delta}}.$$

Moreover, by $[B2\text{-}(q\epsilon)]$, we obtain

$$P[(A^p_{k,b_n^{\delta'}})^c]\leq 4(2p+1)\sup_{i=1,2}\sup_t P[N^i_{t+[b_n]^{-1}b_n^{\delta'}T}-N^i_t=0]\leq C(p+1)b_n^{-q\epsilon\delta'}.$$

Hence we have

$$E\Big[\Big|\sum_{k=1}^{[b_n]}f^n_k\nu^{p,i}_n([t_{k-1},t_k))(1_{A^p_{k,b_n^{\delta'}}}-1)\Big|^q\Big]\leq C(p+1)b_n^{-\frac{q\epsilon\delta\delta'}{1+\delta}}\sup_t|f_t|^q.$$

Therefore we obtain

$$E\Big[\Big|\sum_{k=1}^{[b_n]}f^n_k(\nu^{p,i}_n([t_{k-1},t_k))-\zeta^{p,i}_n([t_{k-1},t_k)))\Big|^q\Big]\leq C(p+1)^{q-1}b_n^{-q\eta}\sup_t|f_t|^q. \tag{1.35}$$

Furthermore, Hölder continuity of f and (1.34) yield

$$E\Big[\Big|\sum_{k=1}^{[b_n]}\int_{t^n_{k-1}}^{t^n_k}(f_t-f^n_k)d\nu^{p,i}_n\Big|^q\Big]\leq[b_n]^{q-1}\omega_\beta(f)^q\sum_{k=1}^{[b_n]}(T[b_n]^{-1})^{q\beta}E[\nu^{p,i}_n([t_{k-1},t_k))^q]\leq Cb_n^{-q\beta}\omega_\beta(f)^q. \tag{1.36}$$

By (1.35) and (1.36), we have

$$E\left[\left|\int_0^T f_td\nu^{p,i}_n-\int_0^T f_td\zeta^{p,i}_n\right|^q\right]\leq C(p+1)^{q-1}b_n^{-q\eta}\big\{\sup_t|f_t|^q+\omega_\beta(f)^q\big\}. \tag{1.37}$$

Since $p \in \mathbb{Z}_+$ and $i = 1, 2$ are arbitrary, we obtain $[A3'\text{-}q, \eta]$ by (1.13) and (1.37). □

Proof of Proposition 1.7.

1. For $h > 0$ and $1 \leq i \leq [b_n]$, let

$$\begin{aligned}
\hat{A}^{p,+}_{h,t} &= \cap_{r=1}^{n_2+2} \cap_{l \in [1, p \wedge h^{-1}(T-t)] \cap \mathbb{N}} \{\omega; N^r_{t+lh} - N^r_{t+(l-1)h} > 0\}, \\
\hat{A}^{p,-}_{h,t} &= \cap_{r=1}^{n_2+2} \cap_{l \in [1, p \wedge h^{-1}t] \cap \mathbb{N}} \{\omega; N^r_{t-(l-1)h} - N^r_{t-lh} > 0\}, \\
\hat{A}^p_{i,h} &:= \hat{A}^{2p+1,+}_{[b_n]^{-1}hT, t_i} \cap \hat{A}^{2p+1,-}_{[b_n]^{-1}hT, t_{i-1}}.
\end{aligned}$$

Then $\omega \in \hat{A}^p_{i,j}$ and $t_{i-1} < R(\theta_{0,k}) \leq t_i$ imply $|\theta_{p,k}| \leq j(4p+2)[b_n]^{-1}T$ for $\omega \in \Omega$, $1 \leq k \leq l_n + m_n$, $n \in \mathbb{N}$, $0 \leq i \leq [b_n]$ and $j \in \mathbb{N}$. Moreover, $\hat{A}^p_{i,j} = \Omega$ if j is sufficiently large for each i and p. Therfore, for $\Delta N_i = N^1_{t_i} - N^1_{t_{i-1}} + N^2_{t_i} - N^2_{t_{i-1}}$ and $\dot{A}^p_{i,j} = \hat{A}^p_{i,j} \setminus \cup_{j'=0}^{j-1} \hat{A}^p_{i,j'}$, we obtain

$$\begin{aligned}
E[(\Phi_{p,1})^q] &= E\left[\left(\sum_{i=1}^{[b_n]} \sum_{k; R(\theta_{0,k}) \in (t_{i-1}, t_i]} |\theta_{p,k}| \sum_{j=1}^{\infty} 1_{\dot{A}^p_{i,j}}\right)^q\right] \leq E\left[\left(\sum_{i=1}^{[b_n]} \sum_{j=1}^{\infty} j \cdot (4p+2)[b_n]^{-1} T \Delta N_i 1_{\dot{A}^p_{i,j}}\right)^q\right] \\
&\leq [b_n]^{q-1} \sum_{i=1}^{[b_n]} \sum_{j=1}^{\infty} j^q \cdot (4p+2)^q T^q [b_n]^{-q} E[(\Delta N_i)^q 1_{\dot{A}^p_{i,j}}]
\end{aligned}$$

for $p \in \mathbb{Z}_+$, since $\{\dot{A}^p_{i,j}\}_{j \in \mathbb{N}}$ are disjoint. Then by $[B1\text{-}(p'_1 q)], [B2\text{-}(p'_2(q+2)]$, the Hölder inequality and a similar estimate for $P[(A^p_{k,h})^c]$ in the proof of Proposition 1.6, we have

$$E[(\Phi_{p,1})^q] \leq C[b_n]^{-1} \sum_{i=1}^{[b_n]} \sum_{j=1}^{\infty} j^q (4p+2)^q P[(\hat{A}^p_{i,j-1})^c]^{1/p'_2} \leq C(p+1)^q \sum_{j=1}^{\infty} j^q \{C(p+1) j^{-p'_2(q+2)}\}^{\frac{1}{p'_2}} \leq C(p+1)^{q+1}$$

for sufficiently large n.

In particular, by the Hölder inequality and Jensen's inequality, we have

$$E\left[r_n^{q'} \sum_{p=0}^{\infty} \frac{(\Phi_{2p+2})^{q'}}{(p+1)^{q'+3}}\right] \leq E[r_n^{\frac{qq'}{q-q'}}]^{\frac{q-q'}{q}} E\left[\sum_{p=0}^{\infty} \frac{1}{(p+1)^2} \left(\frac{(\Phi_{2p+2})^q}{(p+1)^{q+\frac{q}{q'}}}\right)\right].$$

Therefore $[A4\text{-}q', (1+3/q')]$ holds since $r_n \to^p 0$ by the next Proposition 1.8.

2. The proof is similar to that of 1. For sufficiently large n, we have

$$\begin{aligned}
E[(\bar{\Phi}_{p_1,p_2})^{q/2}] &\leq [b_n]^{\frac{q}{2}-1} E\left[\sum_{i=1}^{[b_n]} \sum_{j=1}^{\infty} \left(\sum_{R(\theta_{0,k_1}) \in (t_{i-1}, t_i]} \sum_{k_2} |\theta_{p_1,k_1}| \wedge |\theta_{p_2,k_2}| 1_{\theta_{p_1+p_2,k_1} \cap \theta_{0,k_2} \neq \emptyset}\right)^{\frac{q}{2}} 1_{\dot{A}^{p_1+2p_2+1}_{i,j}}\right] \\
&\leq [b_n]^{\frac{q}{2}-1} E\left[\sum_{i=1}^{[b_n]} \sum_{j=1}^{\infty} \{(4p_1+2)j \wedge (4p_2+2)j\}^{\frac{q}{2}} [b_n]^{-\frac{q}{2}} T^{\frac{q}{2}} (\Delta N_i)^{\frac{q}{2}}\right. \\
&\quad \left. \times \left(\sum_{v=1}^{2} (N^v_{(t_i+(2p_1+2p_2+1)j[b_n]^{-1}T) \wedge T} - N^v_{(t_{i-1}-(2p_1+2p_2+2)j[b_n]^{-1}T) \vee 0})\right)^{\frac{q}{2}} 1_{\dot{A}^{p_1+2p_2+1}_{i,j}}\right] \\
&\leq C[b_n]^{-1} \sum_{i=1}^{[b_n]} \sum_{j=1}^{\infty} \{(4p_1+2)j \wedge (4p_2+2)j\}^{\frac{q}{2}} \{(4p_1+4p_2+3)j+1\}^{\frac{q}{2}} P[(\hat{A}^{p_1+2p_2+1}_{i,j-1})^c]^{\frac{1}{p'_2}}.
\end{aligned}$$

Since $(a \wedge b)(a+b) \leq 2ab$ $(a, b \geq 1)$, we obtain

$$E[(\bar{\Phi}_{p_1,p_2})^{q/2}] \leq C \sum_{j=1}^{\infty} (p_1+1)^{\frac{q}{2}} (p_2+1)^{\frac{q}{2}} j^q \{C(p_1+p_2+1) j^{-p'_2(q+2)}\}^{\frac{1}{p'_2}} \leq C(p_1+1)^{q/2+1} (p_2+1)^{q/2+1},$$

which completes the proof. □

Proof of Proposition 1.8.

Let $A'_j = \hat{A}^{[b_n j^{-1}T],+}_{jb_n^{-1},0}$ for $j \in \mathbb{N}$. Then since $r_n \leq 2jb_n^{-1}$ on A'_j, for sufficiently large n, we have

$$E[r_n^q] = E\left[r_n^q \sum_{j=1}^\infty 1_{A'_j \setminus \cup_{j'=0}^{j-1} A'_{j'}}\right] \leq \sum_{j=1}^\infty (2jb_n^{-1})^q P[(A'_{j-1})^c] \leq Cb_n^{-q} \sum_{j=1}^\infty j^q \cdot [b_n j^{-1}T] \cdot j^{-q-1} \leq Cb_n^{1-q},$$

where $A'_0 = \emptyset$. □

Proof of Proposition 1.9.

By $[B2\text{-}q]$, there exists $N \in \mathbb{N}$ such that

$$\sup_{n \geq n_0} \max_{i=1,2} \sup_{0 \leq t \leq T - N[b_n]^{-1}T} P[N^i_{t+N[b_n]^{-1}T} - N^i_t = 0] \leq \frac{1}{12}. \tag{1.38}$$

For $M = [b_n/3N]$, $h = [b_n]^{-1}T$ and $s_k = 3kNh$, we have

$$a_1 = \frac{1}{T}\int_0^T a_1 dt = \frac{1}{T} \operatorname*{P\text{-}lim}_{n\to\infty} b_n^{-1} \sum_{I,J} \frac{|I \cap J|^2}{|I||J|} = \frac{1}{T} \operatorname*{P\text{-}lim}_{n\to\infty} b_n^{-1} \sum_{k=1}^M \sum_{I,J;L(I)\in[s_{k-1},s_k)} \frac{|I \cap J|^2}{|I||J|}.$$

Let

$$\bar{A}^0_k = \emptyset, \quad \bar{A}^j_k = A^{2,+}_{jh,s_k} \cap A^{1,-}_{jh,s_{k-1}}, \quad \ddot{A}^j_k = \bar{A}^j_k \setminus \cup_{j'=0}^{j-1} \bar{A}^{j'}_k,$$

and

$$E_k = \cap_{l=1}^3 \{N^1_{s_{k-1}+lNh} - N^1_{s_{k-1}+(l-1)Nh} > 0\}$$

for $1 \leq k \leq M$ and $j \in \mathbb{N}$. Then for sufficiently large j, $\bar{A}^j_k = \Omega$. Moreover, for sufficiently large n, we have $\inf_{1\leq k\leq M} P[E_k] \geq 3/4$ by (1.38) and

$$\begin{aligned}
\sum_{I,J;L(I)\in[s_{k-1},s_k)} \frac{|I \cap J|^2}{|I||J|} &\geq \sum_{j=1}^\infty \sum_{I,J;L(I)\in[s_{k-1},s_k)} \frac{|I \cap J|^2 1_{\ddot{A}^j_k}}{((3N+j)h)((3N+3j)h)} \\
&\geq \sum_{j=1}^\infty \frac{(N+j)^{-2}}{9h^2} 1_{\ddot{A}^j_k} \sum_{I,J;L(I)\in[s_{k-1},s_k)} |I \cap J|^2 1_{E_k}.
\end{aligned}$$

For $r \in \mathbb{N}$ and $u > 0$, we have $x_1^2 + \ldots + x_r^2 \geq u^2/r$ when $x_i \geq 0$ $(1 \leq i \leq r)$, $x_1 + \ldots + x_r \geq u$. Hence

$$\sum_{I,J;L(I)\in[s_{k-1},s_k)} \frac{|I \cap J|^2}{|I||J|} \geq \sum_{j=1}^\infty \frac{(N+j)^{-2}}{9h^2} \frac{(Nh)^2 1_{\ddot{A}^j_k} 1_{E_k}}{\Delta N^1_k + \Delta N^2_k + 1},$$

where $\Delta N^i_k = N^i_{s_k} - N^i_{s_{k-1}}$ $(1 \leq k \leq M, i = 1,2)$. Then we obtain

$$\frac{b_n^{-1}}{T} \sum_{I,J} \frac{|I \cap J|^2}{|I||J|} \geq b_n^{-1} \sum_{j=1}^\infty \sum_{k=1}^M X'_{j,k} \quad \text{a.s., where} \quad X'_{j,k} = \frac{N^2}{9Tj(N+j)^2} \frac{1_{\ddot{A}^j_k} 1_{E_k}}{\Delta N^1_k + \Delta N^2_k + 1}. \tag{1.39}$$

On the other hand, Theorem 1.4 and a similar argument to the proof of Proposition 1.6 yield

$$E\left[\left|\sum_{k=1}^M (X'_{j,k} - E[X'_{j,k}])\right|^2\right] \leq \frac{Cb_n}{j(N+j)^4}$$

for $j \in \mathbb{N}$ and sufficiently large n. Therefore

$$b_n^{-1} \sum_{j=1}^\infty \sum_{k=1}^M (X'_{j,k} - E[X'_{j,k}]) \to^p 0 \tag{1.40}$$

as $n \to \infty$.

(1.39) and (1.40) yield

$$a_1 \geq \limsup_{n\to\infty} b_n^{-1} \sum_{j=1}^{\infty} \sum_{k=1}^{M} E[X'_{j,k}]. \tag{1.41}$$

Furthermore, since $\{\Delta N_k^i\}_{1\leq k\leq M, i=1,2, n\geq n_1}$ are tight by the assumption, there exists $R' > 0$ such that $\sup_{n\geq n_1,k,i} P\left[\Delta N_k^i > R'\right] < 1/8$. Consequently,

$$\sup_{n\geq n_1,k} P\left[(\Delta N_k^1 + \Delta N_k^2 + 1)^{-1} < (2R'+1)^{-1}\right] < 1/4. \tag{1.42}$$

On the other hand, by $[B2\text{-}q]$, we obtain

$$P\left[\cup_{j=J+1}^{\infty} \ddot{A}_k^j\right] \quad \leq \quad P[(\bar{A}_k^J)^c] \leq 6 \sup_{n\geq n_0,t,i} P[N_{t+Jh}^i - N_t^i = 0] \leq CJ^{-q}$$

for $J \in \mathbb{N}$ and $n \geq n_0$. Thus, there exists J which does not depend on n, k such that

$$P\left[\cup_{j=1}^{J} \ddot{A}_k^j\right] = 1 - P\left[\cup_{j=J+1}^{\infty} \ddot{A}_k^j\right] \geq \frac{3}{4}. \tag{1.43}$$

Therefore by (1.41),(1.42),(1.43) and the estimate $\inf_{1\leq k\leq M} P[E_k] \geq 3/4$, we obtain

$$a_1 \geq \frac{N^2}{9TJ(N+J)^2} \limsup_{n\to\infty} b_n^{-1} M \frac{1}{2R'+1} \cdot \frac{1}{4} = \frac{N^2(2R'+1)^{-1}}{36TJ(N+J)^2} \frac{1}{3N}.$$

□

Chapter 2

Quasi-Likelihood Analysis for diffusion processes with jumps

2.1 Introduction

Given a probability space $(\Omega, \mathcal{F}, P_{\alpha^*})$ with filtration $\mathbf{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$, let $X = \{X_t\}_{t\in\mathbb{R}_+}$ be a d-dimensional *càdlàg* $\mathbf{F}$-adapted process satisfying the stochastic differential equation

$$\begin{cases} dX_t &= a(X_{t-}, \theta^*)dt + b(X_{t-}, \sigma^*)dW_t + \int_E c(X_{t-}, z, \theta^*)p(dt, dz), \\ X_0 &= x_0, \end{cases} \tag{2.1}$$

where x_0 is a random variable, $\{W_t\}_{t\in\mathbb{R}_+}$ is an r-dimensional standard $\mathbf{F}$-Brownian motion, and $p(dt, dz)$ is a Poisson random measure on $\mathbb{R}_+ \times E$, $E = \mathbb{R}^d \backslash \{0\}$, with compensator $q^{\theta^*}(dt, dz) = P_{\alpha^*}[p(dt, dz)]$. We assume $\mathcal{F}_t$, $\sigma(W_u - W_t; u \geq t)$ and $\sigma(p(A); A \subset (t, \infty) \times E$ is a Borel set) are independent for any $t \geq 0$. We denote $\alpha^* = (\sigma^*, \theta^*)$ for the two statistical parameters $\sigma^* \in \Pi \subset \mathbb{R}^{d_1}$ and $\theta^* \in \Theta \subset \mathbb{R}^{d_2}$ which are unknown to the observer. On the other hand, the coefficients a and c are assumed to be known $\mathbb{R}^d$-valued Borel functions defined on $\mathbb{R}^d \times \Theta$ and $\mathbb{R}^d \times E \times \Theta$ respectively, and b is a known $\mathbb{R}^d \otimes \mathbb{R}^r$-valued Borel function defined on $\mathbb{R}^d \times \Pi$.

We want to estimate $\alpha^* = (\sigma^*, \theta^*)$ from the discrete observations $\{X_{t_i^n}\}_{0\leq i\leq n}$, where $t_i^n = ih$, $h = h_n$. In this chapter, we will present a quasi-likelihood analysis for jump diffusion processes. First we propose a quasi-likelihood function and then show the asymptotic normality of the quasi-maximum likelihood estimator and a Bayes type estimator based on it.

Recently, jump diffusion models are becoming powerful tools to model various stochastic phenomena in many areas such as econometrics, physics, biology, and so on. Numbers of studies worked with jump diffusion models; for example among vast literature, we refer the reader to Prakasa Rao [38] and Cont and Tankov [10] and also references therein.

An earlier work of estimation of discretely observed jump diffusions is in Shimizu and Yoshida [43, 45, 46]. They proved consistency and asymptotic normality of the quasi-maximum likelihood estimator, under $h \to 0$, $nh \to \infty$ and $nh^2 \to 0$. Differently from diffusion models, one of the difficulties caused by the existence of jumps is that the observer cannot distinguish the increments of the data with jumps from those without jumps though classification of increments is necessary to assign each increment to the diffusion/jump likelihood function, when a likelihood analysis is executed. To solve the problem, they proposed a discrimination filter, which enabled to discriminate asymptotically between increments with jumps and increments without jumps. After that, Shimizu studied M-estimation for infinite activity jump processes in [42], and nonparametric estimation of density of Lévy measure in [43]. It should be noted that Mancini [30] independently presented consistent estimation of the characteristics of jumps for Poisson-diffusion model.

We construct a quasi-likelihood analysis for stochastic differential equations with jumps. For this attempt, we will take a general approach by the convergence of the statistical random field associated with the quasi-likelihood function. In the philosophy of the Ibragimov-Has'minskii-Kutoyants program (Ibragimov and Has'minskii [20–22] and Kutoyants [25–28]), we expect that the asymptotic properties of statistics based on the quasi-likelihood function can be derived in a unified way once the convergence of the statistical random field is established with a large deviation estimate for it. Though the Ibragimov-Has'minskii-Kutoyants program features it, the large deviation inequality becomes a technical obstacle, as explained in Yoshida [54]. However recently we obtain a methodology to produce a polynomial type large deviation inequality systematically. It is a general way independent of any particular nature of stochastic processes; inevitably we can apply it to the quasi-likelihood analysis for jump-diffusion processes, as we will do it in this article. We refer the reader to Yoshida [54] for details and a construction of the quasi-likelihood analysis for diffusion processes.

The Bayesian analysis reflects the advantage of the quasi-likelihood analysis so constructed. That is, we propose a Bayesian type estimator for the jump-diffusion and derive its asymptotic behavior. It becomes possible by our methodology. We also obtain certain tail probability estimates of the quasi-maximum likelihood and Bayesian estimators, which yield the convergence of moments of those estimators. The convergence of moments plays an essential role in key steps of theoretical statistics, for example, the asymptotic expansion of statistics, the theory of information criteria and the theory of prediction. For instance, the correction term of AIC, which is defined as the bias of the estimated Kullback-Leibler divergence of the predictive distribution, is validated as the expectation of the square of the scaled maximum likelihood estimator though mathematical backing in this rigorous sense has been neglected in most of literature even in i.i.d. settings.

In Section 2.2, we describe a quasi-likelihood function for the jump-diffusion model, and state the assumptions. Shimizu and Yoshida [46] assumed that the Lévy density f_θ satisfies $f_\theta(z)1_{\{|z|<r\}} \leq K|z|^\gamma$ for some $\gamma > 3$, $r > 0$ and $K > 0$. The condition on γ in Shimizu and Yoshida [46] is weakened in this paper to admit models in which $f_\theta(z) = O(1)$ near the origin, such as a normal distribution. We present the main theorems on the asymptotic normality and the convergence of moments of any order for the quasi-maximum likelihood estimator and a Bayesian type estimator with respect to the quasi-likelihood. Section 2.3.1 gives an exposition of the polynomial type large deviation theory in Yoshida [54]. Then the main theorems are proved in Section 2.3.2.

2.2 Quasi-likelihood and asymptotic property of the estimators

In this section, we will present the main results on the asymptotic properties of the quasi-maximum likelihood estimator and the Bayesian type estimator. In order to define the quasi-likelihood function (2.3)-(2.4) below, we introduce certain functions, for which we need precise descriptions in a sequence of assumptions below. The quasi-likelihood function looks involved due to the truncation function φ_n. However it is unavoidable in general because the sampled data are only available and the substitution of them for the continuously observed data breaks the function of the compensator; consequently we would meet many problems of divergence without truncations. On the other hand, too strong truncation would cause lack of efficiency. The balance is important and it is far from straightforward. This is the reason why we set a rather long sequence of assumptions before introducing the quasi-likelihood function. However, in some "good" cases, we can omit φ_n. See Condition $[H10]$.

Now we detail the setting and notation. We assume that q^{θ^*} has a representation $q^{\theta^*}(dt, dz) = f_{\theta^*}(z)dzdt$ with a density $f_\theta(z)$ disintegrated as $f_\theta(z) = \lambda(\theta)F_\theta(z)$ by a nonnegative function $\lambda(\theta)$ and a probability density F_θ for $\theta \in \Theta$. For a vector $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_l)$, we denote $\partial_\kappa = (\frac{\partial}{\partial\kappa_1}, \frac{\partial}{\partial\kappa_2}, \ldots, \frac{\partial}{\partial\kappa_l})$, $\partial_\kappa^2 = (\frac{\partial^2}{\partial\kappa_i\partial\kappa_j})_{1\leq i,j\leq l}$ and $\partial_\kappa^3 = (\frac{\partial^3}{\partial\kappa_i\partial\kappa_j\partial\kappa_k})_{1\leq i,j,k\leq l}$. It will be assumed that the full parameter space $\Xi = \Pi \times \Theta$ is a bounded open subset of $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, and that Ξ, Π, and Θ admit Sobolev's inequality. An open set $U \subset \mathbb{R}^m$ is said to admit Sobolev's inequality if for any $p > m$, there exists a positive constant C depending U and p, such that

$$\sup_{x\in U} |u(x)| \quad \leq \quad C \sum_{k=0,1} \|\partial_x^k u(x)\|_p$$

for all $u \in C^1(U)$. It is the case if U has a Lipschitz boundary.

We will use the following notation: $\beta(x, \sigma) = b(x, \sigma)b^T(x, \sigma)$ for $\sigma \in \Pi$. For $S \subset \mathbb{R}^m$, $\bar{S}$ denotes closure of S. For $\kappa = (\kappa_{ij})_{1\leq i,j\leq l}$ and $\iota = (\iota_{ijk})_{1\leq i,j,k\leq l}$, we define $|\kappa| = \sqrt{\sum_{i,j=1}^l \kappa_{ij}^2}$ and $|\iota| = \sqrt{\sum_{i,j,k=1}^l \iota_{ijk}^2}$, respectively. For a function g defined on $\mathbb{R}^d \times \Xi$, we write $g_{i-1}(\alpha) = g(X_{t_{i-1}^n}, \alpha)$, $\Delta X_i^n = X_{t_i^n} - X_{t_{i-1}^n}$, $\Delta X_t = X_t - X_{t-}$, $\bar{X}_{i,n}(\theta) = \Delta X_i^n - ha_{i-1}(\theta)$ and $\bar{X}_{i,n} = \bar{X}_{i,n}(\theta^*)$. We abuse the notation and write $\mathcal{F}_{i-1} = \mathcal{F}_{t_{i-1}^n}$.

Let $k \in \mathbb{N}$ and u_n be a sequence of positive numbers. Denote by $R = R_n : \Xi \times \mathbb{R} \times \mathbb{R}^d \times \Omega \to \mathbb{R}^k$ a sequence of random functions for which there exists a constant C such that $|R(\alpha, u_n, x)| \le u_n C(1+|x|)^C$ for all $\alpha \in \Xi, x \in \mathbb{R}^d, n \in \mathbb{N}, \omega \in \Omega$. Moreover, in the case that $k = 1$, let $\tilde{R}(\alpha, u_n, x) = 1 - R(\alpha, u_n, x)$. We use the symbol C for a generic positive constant varying from line to line. The symbols $R(\alpha, u_n, x)$ and $\tilde{R}(\alpha, u_n, x)$ are also used to express generic variables that satisfy the inequality as above. Let $\hat{C}(\mathbb{R}^m)$ be the space of continuous functions on $\mathbb{R}^m$ that tend to zero as $|x| \to \infty$. Equip $\hat{C}(\mathbb{R}^m)$ with the supremum norm. For nonnegative sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$, $a_n \preceq b_n$ means that there exists $C > 0$ such that $a_n \le Cb_n$ for all $n \in \mathbb{N}$. We denote E for expectation with respect to P_{α^*}. Let $n^{-3/5} \preceq h \preceq n^{-4/7}$ for $h = h_n > 0$, $0 < b < 1/8$ and let $\{\epsilon_n\}$ be a sequence of positive number such that $\epsilon_n \to 0$,

$$\frac{\sqrt{n}h}{\epsilon_n^2} \vee \frac{h^b}{\epsilon_n} \preceq 1, \quad \text{and} \quad 1 \preceq n^3 h^4 \epsilon_n^{16}. \tag{2.2}$$

For example, $b = 1/10$ and $\epsilon_n := h^{1/16}$.

We consider the following conditions to obtain the main results.

[H1] For some constant L and function $\zeta(z)$ of at most polynomial growth in z,

$$|a(x,\theta^*) - a(y,\theta^*)| + |b(x,\sigma^*) - b(y,\sigma^*)| \le L|x-y|,$$

$$|c(x,z,\theta^*) - c(y,z,\theta^*)| \le \zeta(z)|x-y|, \quad |c(x,z,\theta^*)| \le \zeta(z)(1+|x|).$$

[H2] The process $\{X_t\}$ has the exponential α-mixing property, i.e., there exists $c > 0$ such that,

$$\sup_{t\in\mathbb{R}_+} \sup_{A\in\sigma[X_r:r\le t], B\in\sigma[X_r:r\ge t+h]} |P_{\alpha^*}[A\cap B] - P_{\alpha^*}[A]P_{\alpha^*}[B]| \le \frac{1}{c}e^{-ch} \quad (h>0).$$

Moreover, we will assume the stationarity of X for simplicity.

The ergodicity of X follows from [H2]. Denote by $\pi(dx)$ the invariant probability measure, i.e.,

$$\frac{1}{T}\int_0^T \mathsf{f}(X_t)dt \to^P \int \mathsf{f}(x)\pi(dx)$$

as $T \to \infty$ for any π-integrable function f. For the exponential mixing property [H2] of a jump diffusion process and the following condition [H3], we refer the reader to Masuda [32] and [31].

[H3] For every $p \ge 1$, $\sup_{t\ge 0} E[|X_t|^p] < \infty$.

[H4] For each $\sigma \in \Pi$, the derivatives $\partial_x^k b(x,\sigma)$ $(k = 0,1,2)$ exist on $\mathbb{R}^d$ and they are continuous in x. Moreover, for fixed x, the derivatives $\partial_\theta^l a(x,\theta)$ and $\partial_\sigma^l b(x,\sigma)$ $(l = 0,1,2,3,4)$ exist and continuous on Θ and Π respectively, and a and b can be continuously extended to $\bar{\Theta}$ and $\bar{\Pi}$ respectively, for any $x \in \mathbb{R}^d$. Furthermore, a, b, and their derivatives are of at most polynomial growth in x uniformly in α:

$$|\partial_x^k b(x,\sigma)|, |\partial_\theta^l a(x,\theta)|, |\partial_\sigma^l b(x,\sigma)| \le C(1+|x|)^C \ (x \in \mathbb{R}^d, \alpha \in \Xi),$$

for $l = 0,1,2,3,4$, and $k = 1,2$.

[H5] There exist constants $r > 0$ and $K > 0$ such that $f_{\theta^*}(z)1_{\{|z|\le r\}} \le K|z|^{1-d}$, and that

$$\sup_{\theta\in\Theta} \int |z|^p f_\theta(z)dz < \infty$$

for all $p \ge 1$.

[H6] For each (θ, x), the mapping $z \mapsto y = c(x, z, \theta)$ is an injection from E into E and has an inverse $z = c^{-1}(x, y, \theta)$ from the image of c onto E, which is differentiable with respect to y. Furthermore, the set $B(x) := \mathrm{Im}(c(x, \cdot, \theta)) = \{y \in E;$ there exists $z \in E,$ such that $y = c(x, z, \theta)\}$ is open and independent of $\theta \in \Theta$, and the set $\{(x, y) \in \mathbb{R}^d \times E; x \in \mathbb{R}^d, y \in B(x)\}$ is a Borel set. Moreover, for the absolute value $J(x, y, \theta)$ of the Jacobian of $c^{-1}(x, y, \theta)$ and

$$\Psi_\theta(y, x) = f_\theta(c^{-1}(x, y, \theta))J(x, y, \theta) \quad (y \in B(x), x \in \mathbb{R}^d, \theta \in \Theta),$$

the set $A(x) = \{y \in B(x); \Psi_\theta(y, x) \neq 0\}$ does not depend on θ.

[H7] There exists positive constants c_0 and r_1 such that $|y| \geq c_0|c^{-1}(x, y, \theta^*)|$ for any $x \in \mathbb{R}^d$ and $y \in B(x) \cap \{y; |y| \leq r_1\}$.

[H8] $\inf_{x \in \mathbb{R}^d, \sigma \in \Pi} \det \beta(x, \sigma) > 0$.

We assume that Ψ_θ of [H6] admits an extension to $E \times \mathbb{R}^d$ in such a way that Ψ_θ satisfies [H9]-[H12] below.

[H9] The function $\Psi_\theta(y, x)$ has derivatives $\partial_\theta^k \Psi_\theta(y, x), \partial_y \partial_\theta^k \Psi_\theta(y, x) (k = 0, 1, 2, 3, 4)$ in $(x, y, \theta) \in \mathbb{R}^d \times E \times \Theta$ which is continuous in y, and for $x \in \mathbb{R}^d, y \in E$, $\Psi_\theta(y, x)$ can be continuously extended to $\bar{\Theta}$. Moreover, for $x \in \mathbb{R}^d$ and $k = 0, 1, 2, 3, 4$,

$$\begin{aligned}
\int_{B(x)} \sup_{\theta \in \Theta} |\partial_\theta^k \Psi_\theta(y, x)| dy &\leq C(1 + |x|)^C, \\
\int_{A(x)} \sup_{\theta \in \Theta} |\partial_\theta^k \log \Psi_\theta(y, x)| \times \Psi_{\theta^*}(y, x) dy &\leq C(1 + |x|)^C, \\
\sup_{\theta \in \Theta} \int_{A(x)} |\partial_\theta^k \log \Psi_\theta(y, x)|^2 \times \Psi_{\theta^*}(y, x) dy &\leq C(1 + |x|)^C, \\
\sup_{\theta \in \Theta} \int_{A(x)} |\partial_\theta \log \Psi_\theta(y, x)|^l \times \Psi_{\theta^*}(y, x) dy &\leq C(1 + |x|)^C \quad (l = 3, 4).
\end{aligned}$$

Let $\partial_\theta^k \log \Psi_\theta(y, x)\Big|_{\Psi_\theta(y,x)=0}, \partial_y \partial_\theta^k \log \Psi_\theta(y, x)\Big|_{\Psi_\theta(y,x)=0} = -\infty$ (k=0,1,2,3,4).
As anticipated at the beginning of this section, we need a sequence of truncation functions.

[H10] At least one of the following two conditions holds true.

1. There exists a sequence of real valued Borel functions $\{\varphi_n(x, y)\}_{n \in \mathbb{N}}$ on $\mathbb{R}^d \times E$, possessing the following properties: $0 \leq \varphi_n \leq 1$, and there exists $M > 0$ such that $\varphi_n(x, y) = 0$ whenever $(x, y) \in D_n$, where

$$\begin{aligned}
D_n &= \cup_{k=0}^4 \left\{ (x, y) \in \mathbb{R}^d \times E; \sup_{\theta \in \Theta} |\partial_\theta^k \log \Psi_\theta(y, x)| \geq \frac{M(1 + |x|)^M}{\epsilon_n^{k \vee 1}} \right\} \\
&\quad \cup \cup_{k=0}^4 \left\{ (x, y) \in \mathbb{R}^d \times E; \sup_{\theta \in \Theta} |\partial_y \partial_\theta^k \log \Psi_\theta(y, x)| \geq \frac{M(1 + |x|)^M (1 + |y|)^M}{\epsilon_n^{k+1}} \right\}.
\end{aligned}$$

   Moreover, φ_n is differentiable with respect to y, $\partial_y \varphi_n$ is continuous in y,

$$\partial_y \varphi_n = 0 \quad \text{on } D_n, \quad \text{and} \quad \sup_{x \in \mathbb{R}^d, y \in E} |\partial_y \varphi_n| = O(\epsilon_n^{-1}).$$

2. It holds that $|\partial_y^l \partial_\theta^k \log \Psi_\theta(x, y)| \leq C(1+|y|)^C (1+|x|)^C$ $(x \in \mathbb{R}^d, y \in E, \theta \in \Theta, l = 0, 1,\ k = 0, 1, 2, 3, 4)$. In this case, we set $\varphi_n \equiv 1$ and $\epsilon_n = h^{\frac{1}{16} \wedge b}$.

[H11] There exists $a > 0$ such that

$$\sup_{\theta \in \Theta} \int_{A(x)} |\partial_\theta^k \log \Psi_\theta(y, x)| \times \Psi_{\theta^*}(y, x)(1 - \varphi_n(x, y)) dy \leq Ch^a (1 + |x|)^C (k = 0, 2).$$

$$\sup_{\theta\in\Theta}\int_{B(x)}|\Psi_\theta(y,x)-\Psi_{\theta^*}(y,x)|(1-\varphi_n(x,y))dy \le Ch^a(1+|x|)^C.$$

$$\int_{B(x)}|\partial_\theta^2\Psi_{\theta^*}(y,x)|(1-\varphi_n(x,y))dy \le Ch^a(1+|x|)^C.$$

The condition [H9]-[H11] are rather complicated in the case that [H10] 2 does not hold. The conditions [H9]-[H11] are satisfied if there exist constants $a_1, a_2, a_3, a_4 > 0$ such that the following [G1] and [G2] are satisfied.

[G1] The function $\Psi_\theta(y,x)$ has derivatives $\partial_\theta^k\Psi_\theta(y,x), \partial_y\partial_\theta^k\Psi_\theta(y,x)$ $(k=0,1,2,3,4)$ in $(x,y,\theta)\in\mathbb{R}^d\times E\times\Theta$ which is continuous in y, and for $x\in\mathbb{R}^d, y\in E$, $\Psi_\theta(y,x)$ can be continuously extended to $\bar{\Theta}$. Moreover,

$$\begin{aligned}
\sup_{\theta\in\Theta}|\partial_\theta^k\log\Psi_\theta(y,x)| &\le C\Big\{|y|^{a_1}\vee|\log|y||^{a_2}\Big\}(1+|x|)^C \ (y\in A(x)),\\
\sup_{\theta\in\Theta}|\partial_y\partial_\theta^k\log\Psi_\theta(y,x)| &\le C\Big\{|y|^{a_1}\vee\frac{1}{|y|}\Big\}(1+|x|)^C \ (y\in A(x)),\\
\sup_{\theta\in\Theta}|\partial_\theta^k\Psi_\theta(y,x)| &\le Ce^{-a_3|y|}(1+|\log|y||)^{a_4}(1+|x|)^C \ (y\in B(x)),
\end{aligned}$$

for $x\in\mathbb{R}^d$, $k=0,1,2,3,4$.

[G2] There exist a sequence of real valued Borel functions $\{\varphi_n(x,y)\}_{n\in\mathbb{N}}$ on $\mathbb{R}^d\times E$ and positive constants $\{c_i\}_{i=1}^5$ such that $c_1<c_2$, $c_3<c_4$, $\epsilon_n\preceq h^{c_5}$, $0\le\varphi_n\le1$, and for $x\in\mathbb{R}^d$ and $n\in\mathbb{N}$,

$$\begin{aligned}
\varphi_n(x,y) &= 0 \quad\text{if}\quad |y|\ge\frac{c_4}{\epsilon_n^{1/a_1}} \quad\text{or}\quad |y|\le c_1\epsilon_n,\\
\varphi_n(x,y) &= 1 \quad\text{if}\quad c_2\epsilon_n\le|y|\le\frac{c_3}{\epsilon_n^{1/a_1}},
\end{aligned}$$

for large n. Moreover, φ_n is differentiable with respect to y, $\partial_y\varphi_n$ is continuous in y and

$$\sup_{x\in\mathbb{R}^d,y\in E}|\partial_y\varphi_n| = O(\epsilon_n^{-1}).$$

Define Y^1 and Y^2 as follows:

$$Y^1(\sigma;\sigma^*) = \frac{1}{2}\int tr\left(I_d-\beta^{-1}(x,\sigma)\beta(x,\sigma^*)\right)\pi(dx)-\frac{1}{2}\int\log\frac{\det\beta(x,\sigma)}{\det\beta(x,\sigma^*)}\pi(dx),$$

and

$$\begin{aligned}
Y^2(\theta;\alpha^*) &= -\frac{1}{2}\int(a(x,\theta)-a(x,\theta^*))^T\beta^{-1}(x,\sigma^*)(a(x,\theta)-a(x,\theta^*))\pi(dx)\\
&\quad+\int\int_{A(x)}(\log\Psi_\theta(y,x)-\log\Psi_{\theta^*}(y,x))\Psi_{\theta^*}(y,x)dy\pi(dx)\\
&\quad-\int\int_{B(x)}(\Psi_\theta(y,x)-\Psi_{\theta^*}(y,x))dy\pi(dx)\\
&= -\frac{1}{2}\int(a(x,\theta)-a(x,\theta^*))^T\beta^{-1}(x,\sigma^*)(a(x,\theta)-a(x,\theta^*))\pi(dx)\\
&\quad+\int\int_{A(x)}(\log\Psi_\theta(y,x)-\log\Psi_{\theta^*}(y,x))\Psi_{\theta^*}(y,x)dy\pi(dx)-(\lambda(\theta)-\lambda(\theta^*)).
\end{aligned}$$

[H12] There exist positive constants $\chi(\alpha^*)$ and $\chi'(\alpha^*)$ such that

$$Y^1(\sigma;\sigma^*)\le-\chi(\alpha^*)|\sigma-\sigma^*|^2 \quad\text{for all}\quad \sigma\in\Pi,$$

and

$$Y^2(\theta;\alpha^*)\le-\chi'(\alpha^*)|\theta-\theta^*|^2 \quad\text{for all}\quad \theta\in\Theta.$$

Let $D > 0$ be a constant[1] and let ρ satisfy

$$\frac{3}{8} + b \leq \rho < \frac{1}{2},$$

where b is the constant in the definition of ϵ_n. Our quasi-likelihood is given by

$$L_n(\alpha) \quad = \quad \exp(H_n(\alpha)), \tag{2.3}$$

where

$$\begin{aligned} H_n(\alpha) \quad = \quad & -\frac{1}{2h}\sum_{i=1}^{n} \bar{X}_{i,n}^T(\theta)\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n| \leq Dh^\rho\}} \\ & -\frac{1}{2}\sum_{i=1}^{n} \log\det\beta_{i-1}(\sigma)1_{\{|\Delta X_i^n| \leq Dh^\rho\}} \\ & +\sum_{i=1}^{n}\{\log\Psi_\theta(\Delta X_i^n, X_{t_{i-1}^n})\}\varphi_n(X_{t_{i-1}^n}, \Delta X_i^n)1_{\{|\Delta X_i^n| > Dh^\rho\}} \\ & -h\sum_{i=1}^{n}\int_{B(X_{t_{i-1}^n})} \Psi_\theta(y, X_{t_{i-1}^n})\varphi_n(X_{t_{i-1}^n}, y)dy. \end{aligned} \tag{2.4}$$

The above quasi-likelihood function is slightly different from that of in Shimizu and Yoshida [46] by technical reasons.

The intuitive meaning of H_n is the following. If no jumps occur in the interval $(t_{i-1}^n, t_i^n]$ then we have

$$\Delta X_i^n = \int_{t_{i-1}^n}^{t_i^n} a(X_{t-}, \theta^*)dt + \int_{t_{i-1}^n}^{t_i^n} b(X_{t-}, \sigma^*)dW_t \sim a_{i-1}(\theta^*)h_n + b_{i-1}(\sigma^*)(W_{t_i^n} - W_{t_{i-1}^n}), \tag{2.5}$$

and the conditional log-likelihood function of the variable of the right-hand side corresponds to the first two terms of H_n without a constant term.

If the process $\{X_t\}$ jumps only once in the interval $(t_{i-1}^n, t_i^n]$ at the time τ_i^n and $c(x, z, \theta) \equiv z$, then we obtain

$$\Delta X_i^n \sim \Delta X_{\tau_i^n}, \tag{2.6}$$

for sufficiently large n since the diffusion part of ΔX_i^n is negligible compared with the jump part. The distribution of the variable of the right-hand side of (2.6) corresponds F_{θ^*} and we have

$$\log F_{\theta^*}(\Delta X_i^n) = \log\Psi_{\theta^*}(\Delta X_i^n) - \log\lambda(\theta^*). \tag{2.7}$$

The first term of (2.7) corresponds approximately to the third term of H_n.

Therefore, H_n is the quasi-log-likelihood function that if $|\Delta X_i^n| \leq Dh^\rho$ then H_n judges no jumps occur in the interval $(t_{i-1}^n, t_i^n]$, and if $|\Delta X_i^n| > Dh^\rho$ then H_n judges a jump occurs in the interval $(t_{i-1}^n, t_i^n]$. We refer the reader to Shimizu and Yoshida [46] for more detail of the rationale of the derivation of H_n.

We define $\Gamma^1(\sigma^*)$ and $\Gamma^2(\alpha^*)$ as follows:

$$\begin{aligned} \Gamma^1(\sigma^*) \quad &= \quad \frac{1}{2}\int tr(\partial_\sigma^2\beta^{-1}(x,\sigma^*)\beta(x,\sigma^*))\pi(dx) + \frac{1}{2}\int \partial_\sigma^2 \log\det\beta(x,\sigma^*)\pi(dx) \\ &= \quad \frac{1}{2}\int tr(\beta^{-1}\partial_\sigma\beta\beta^{-1}\partial_\sigma\beta(x,\sigma^*))\pi(dx) \end{aligned}$$

[1]We may set $D = 1$ in what follows. It is theoretically the same, however the choice of D will have a practical meaning.

$$\begin{aligned}\Gamma^2(\alpha^*) &= \int \partial_\theta a^T(x,\theta^*)\beta^{-1}(x,\sigma^*)\partial_\theta a(x,\theta^*)\pi(dx) + \int\int_{B(x)} \partial_\theta^2 \Psi_{\theta^*}(x,y)dy\pi(dx) \\ &\quad - \int\int_{A(x)} \partial_\theta^2 \log\Psi_{\theta^*}(y,x)\Psi_{\theta^*}(y,x)dy\pi(dx). \\ &= \int \partial_\theta a^T(x,\theta^*)\beta^{-1}(x,\sigma^*)\partial_\theta a(x,\theta^*)\pi(dx) \\ &\quad + \int\int_{A(x)} \frac{(\partial_\theta\Psi_{\theta^*}(y,x))^2}{\Psi_{\theta^*}(y,x)} dy\pi(dx).\end{aligned}$$

Let $(\hat\sigma_n,\hat\theta_n)$ be a quasi-maximum likelihood estimator for H_n, i.e., $(\hat\sigma_n,\hat\theta_n)$ is a random variable and $H_n(\hat\sigma_n,\hat\theta_n) = \max_{\sigma\in\bar\Pi,\theta\in\bar\Theta} H_n(\sigma,\theta)$. Let $\hat u_n = (\sqrt{n}(\hat\sigma_n-\sigma^*), \sqrt{nh}(\hat\theta_n-\theta^*))$, and $\check u$ be a random vector which follows a normal distribution $N(0, \mathrm{diag}(\Gamma^1(\alpha^*)^{-1}, \Gamma^2(\alpha^*)^{-1}))$.

Theorem 2.1. *Suppose that $[H1]-[H12]$ are fulfilled. Then $\hat u_n \to^d \check u$ as $n\to\infty$. Moreover $E[\mathsf{f}(\hat u_n)] \to \mathbb{E}[\mathsf{f}(\check u)]$ as $n\to\infty$ for any continuous function f of at most polynomial growth.*

We also discuss consistency and asymptotic normality of the **adaptive Bayes type estimator**. Let $\pi_{1,n},\pi_{2,n}$ be prior densities of σ and θ respectively and parameter spaces Π and Θ are convex sets. We assume $0 < \inf_{\sigma\in\Pi,\theta\in\Theta,n\in\mathbb{N}}(\pi_{1,n}\wedge\pi_{2,n}) \le \sup_{\sigma\in\Pi,\theta\in\Theta,n\in\mathbb{N}}(\pi_{1,n}\vee\pi_{2,n}) < \infty$, and $\{\pi_{k,n}\}_{n\in\mathbb{N}}$ is equicontinuous ($k=1,2$). Then the adaptive Bayes type estimators $(\tilde\sigma_n,\tilde\theta_n)$ for (σ,θ) with respect to the quadratic loss function are defined inductively by

$$\begin{aligned}\tilde\sigma_n &= \int_\Pi \sigma\exp\left(H_n(\sigma,\theta^\star)\right)\pi_{1,n}(\sigma)d\sigma \Big/ \left\{\int_\Pi \exp\left(H_n(\sigma,\theta^\star)\right)\pi_{1,n}(\sigma)d\sigma\right\} \\ \tilde\theta_n &= \int_\Theta \theta\exp\left(H_n(\tilde\sigma_n,\theta)\right)\pi_{2,n}(\theta)d\theta \Big/ \left\{\int_\Theta \exp\left(H_n(\tilde\sigma_n,\theta)\right)\pi_{2,n}(\theta)d\theta\right\}\end{aligned}$$

where $\theta^\star$ is an arbitrary dummy value of θ. Let

$$\tilde u_n = (\sqrt{n}(\tilde\sigma_n-\sigma^*), \sqrt{nh}(\tilde\theta_n-\theta^*)).$$

Theorem 2.2. *Suppose that $[H1]-[H12]$ are satisfied. Then $\tilde u_n \to^d \check u$ as $n\to\infty$. Moreover, $E[\mathsf{f}(\tilde u_n)] \to \mathbb{E}[\mathsf{f}(\check u)]$ as $n\to\infty$ for any continuous function f of at most polynomial growth.*

Remark 2.1. *As there is flexibility to choose parameters D, ρ, π_k, $\bar\theta^\star$ and the function φ_n which satisfies the assumptions of Theorems 2.1 and 2.2, the choice affects estimation results for finite n in practice. In particular, the choice of the threshold Dh^ρ is important because we might detect no jumps in any interval if we set a too high threshold, and we might judge that jumps occur in all intervals if we set a too low threshold. The problem of the choice of the threshold seems difficult and there seems no deciding theory yet. However Shimizu [44] argued about a method of selecting a threshold with a certain criterion.*

Remark 2.2. *Though the quasi-likelihood function H_n and the estimators $(\hat\sigma_n,\hat\theta_n)$ and $(\tilde\sigma_n,\tilde\theta_n)$ are defined as functions on $(\Omega,\mathcal{F},P_{\alpha^*})$, we often regard them as functions on the state space.*

We will show some examples of models which satisfy $[H1]-[H12]$.

Example [Lévy OU processes] Let $d=1$, $\epsilon_n = h^{1/16}$, $0 < R_i^- < R_i^+$ $(1\le i\le 5)$. Suppose that $\{X_t\}$ satisfies

$$dX_t = -aX_{t-}dt + \sigma dW_t + \int_E zp(dt,dz),$$

where $(a,\sigma)\in(R_1^-,R_1^+)\times(R_2^-,R_2^+)$ and X_0 follows the invariant probability measure π.

(i) Let

$$f_\theta(z) = \lambda \frac{\alpha^\beta}{\Gamma(\beta)} z^{\beta-1} e^{-\alpha z} 1_{\{z>0\}},$$

where $\theta = (a, \lambda, \alpha, \beta) \in (R_1^-, R_1^+) \times (R_3^-, R_3^+) \times (R_4^-, R_4^+) \times (1, R_5^+)$. Then

$$\log \Psi_\theta(y) = \log \lambda + \beta \log \alpha - \log \Gamma(\beta) + (\beta - 1) \log y - \alpha y \quad (y > 0).$$

Moreover, we take a function $\rho(x)$ such that $\rho \in C^1(\mathbb{R})$, $0 \le \rho \le 1$, $\rho \equiv 1$ on $\{|x| \le 1\}$ and $\rho \equiv 0$ on $\{|x| \ge 2\}$, and

$$\varphi_n(x, y) = \begin{cases} 1 - \rho(\frac{y}{\epsilon_n}) - \rho(\frac{1}{y\epsilon_n}) & (y > 0) \\ 0 & (y < 0) \end{cases}$$

for large n. Then this model satisfies $[H1] - [H12]$. The symbols Γ^1 and Γ^2 become

$$\Gamma^1(\alpha^*) = \frac{2}{(\sigma^*)^2}$$

and

$$\Gamma^2(\alpha^*) = \begin{pmatrix} \mu_2/(\sigma^*)^2 & 0 & 0 & 0 \\ 0 & 1/\lambda^* & 0 & 0 \\ 0 & 0 & \lambda^*\beta^*/(\alpha^*)^2 & -\lambda^*/\alpha^* \\ 0 & 0 & -\lambda^*/\alpha^* & \lambda^*\{\Gamma(\beta^*)\Gamma''(\beta^*) - (\Gamma'(\beta^*))^2\}/(\Gamma(\beta^*))^2 \end{pmatrix},$$

respectively, where $\mu_2 = \int x^2 \pi(dx)$. For the Levy OU process, π can be calculated explicitly. See Sato [40] and [41] and Wolfe [55].

(ii) Let

$$f_\theta(z) = \lambda \frac{1}{\sqrt{2\pi\theta_2}} \exp\Big\{ - \frac{(z-\theta_1)^2}{2\theta_2} \Big\},$$

where $R_4, R_4' > 0$, $\theta = (a, \lambda, \theta_1, \theta_2) \in (R_1^-, R_1^+) \times (R_3^-, R_3^+) \times (-R_4', R_4) \times (R_5^-, R_5^+)$. Then

$$\log \Psi_\theta(y, x) = \log \lambda - \frac{(y-\theta_1)^2}{2\theta_2} - \frac{1}{2} \log(2\pi\theta_2).$$

In this case, $[H10]$ 2. holds. So we can set $\varphi_n \equiv 1$. Then this model satisfies $[H1] - [H12]$. Let $n_1 = \sum_{i=1}^n 1_{\{|\Delta X_i^n| > Dh^\rho\}}$. Then the quasi-maximum likelihood estimator $(\hat\sigma_n, \hat a_n, \hat\lambda_n, \hat\theta_{1,n}, \hat\theta_{2,n})$ can be calculated as

$$\begin{aligned} \hat\sigma_n^2 &= \frac{1}{h(n-n_1)} \sum_{i=1}^n (\Delta X_i^n + \hat a_n h X_{t_{i-1}^n})^2 1_{\{|\Delta X_i^n| \le Dh^\rho\}}, \\ \hat a_n &= -\left(\sum_{i=1}^n X_{t_{i-1}^n} \Delta X_i^n 1_{\{|\Delta X_i^n| \le Dh^\rho\}} \right) \Big/ \left(h \sum_{i=1}^n X_{t_{i-1}^n}^2 1_{\{|\Delta X_i^n| \le Dh^\rho\}} \right), \\ \hat\lambda_n &= \frac{n_1}{nh}, \quad \hat\theta_{1,n} = \frac{1}{n_1} \sum_{i=1}^n \Delta X_i^n 1_{\{|\Delta X_i^n| > Dh^\rho\}}, \quad \hat\theta_{2,n} = \frac{1}{n_1} \sum_{i=1}^n (\Delta X_i^n - \hat\theta_{1,n})^2 1_{\{|\Delta X_i^n| > Dh^\rho\}}, \end{aligned}$$

if $0 < n_1 < n$ and the parameter space Ξ is sufficiently large to contain this point. The symbols Γ^1 and Γ^2 become

$$\Gamma^1(\alpha^*) = \frac{2}{(\sigma^*)^2}, \quad \Gamma^2(\alpha^*) = diag(\frac{\mu_2}{(\sigma^*)^2}, \frac{1}{\lambda^*}, \frac{\lambda^*}{\theta_2^*}, \frac{\lambda^*}{2(\theta_2^*)^2}),$$

where $\mu_2 = \int x^2 \pi(dx)$. So by Theorem 2.1, the asymptotic distribution of $(\sqrt{n}(\hat\sigma_n - \sigma^*), \sqrt{nh}(\hat\theta_n - \theta^*))$ for the quasi-maximum likelihood estimator $(\hat\sigma_n, \hat\theta_n)$ becomes $N(0, diag((\sigma^*)^2/2, (\sigma^*)^2/\mu_2, \lambda^*, \theta_2^*/\lambda^*, 2(\theta_2^*)^2/\lambda^*))$.

2.3 Proof of the main results

In this section, we will prove Theorems 2.1 and 2.2. For this purpose, let us begin with a few basic results on the polynomial type large deviation inequality and its applications to the quasi-likelihood analysis. This scheme is applicable to various stochastic structures, in particular it works well for nonlinear stochastic processes.

2.3.1 Polynomial type large deviation inequality and the quasi-likelihood analysis

To show consistency and asymptotic normality of the quasi-maximum likelihood estimator and the Bayes type estimator, we will use the method in Yoshida [54].

Let $\Theta \subset \mathbb{R}^m$ be bounded open set admitting Sobolev's inequality, while $\mathcal{T}$ is an arbitrary set. We apply the quasi-likelihood analysis based on a random field $H_n(\theta,\tau) : \Omega \times \bar{\Theta} \times \mathcal{T} \to \mathbb{R}$, that is C^3 on Θ and continuous on $\bar{\Theta}$ for every $\omega \in \Omega$ and $\tau \in \mathcal{T}$. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers such that $a_n \to 0$ as $n \to \infty$. Set $b_n = a_n^{-2}$, $\xi^* = (\theta^*, \tau^*) \in \Theta \times \mathcal{T}$, and $U_n = \{u \in \mathbb{R}^m; \theta^* + a_n u \in \Theta\}$. We consider the ratio of the quasi-likelihood functions defined by

$$Z_n(u,\tau;\theta^*) = \exp\{H_n(\theta^* + a_n u, \tau) - H_n(\theta^*,\tau)\} \quad (u \in U_n, \tau \in \mathcal{T}).$$

Corresponding to the log likelihood function and the observed information in likelihood analysis, we also set

$$Y_n(\theta,\tau;\theta^*) = \frac{1}{b_n}(H_n(\theta,\tau) - H_n(\theta^*,\tau)),$$

and $\Gamma_n(\theta,\tau) = -b_n^{-1}\partial_\theta^2 H_n(\theta,\tau)$ $(\theta \in \Theta, \tau \in \mathcal{T})$.

The key of our arguments is the so-called polynomial type large deviation inequality. In order to derive it, we will assume several conditions, however, they are rather mild compared with the assumptions to ensure the usual exponential type large deviation inequality. In the conditions stated below, $\Gamma(\tau;\xi^*)$ and $Y(\theta,\tau;\theta^*)$ are given deterministic functions and the latter satisfies $Y(\theta^*,\tau;\theta^*) = 0$. Suppose that $L > 0$, $\alpha > 0$, $\beta = \alpha/(1-\alpha)$, $0 < \beta_1 < 1/2$, $\beta_2 \geq 0$, $0 < \rho_1 < 1$ and $\rho_2 > 0$. The following conditions [P1]-[P5] are the conditions [A1″],[A4′],[A6],[B1],[B2] in Yoshida [54], respectively.

[P1] For $M_3 = L(\beta - \rho_1)^{-1}$,

$$\sup_{n\in\mathbb{N}} E\left[\left(b_n^{-1} \sup_{(\theta,\tau)\in\Theta\times\mathcal{T}} |\partial_\theta^3 H_n(\theta,\tau)|\right)^{M_3}\right] < \infty.$$

Moreover, for $M_4 = L(\frac{2\beta_1}{1-\alpha} - \rho_1)^{-1}$,

$$\sup_{n\in\mathbb{N}} E\left[\sup_{\tau\in\mathcal{T}} (b_n^{\beta_1}|\Gamma_n(\theta^*,\tau) - \Gamma(\tau;\xi^*)|)^{M_4}\right] < \infty.$$

[P2] $\rho_1 < \beta \wedge \frac{2\beta_1}{1-\alpha}$, $\alpha < \rho_2/2$, and $1 - 2\beta_2 - \rho_2 > 0$.

[P3] For $M_1 = L(1-\rho_1)^{-1}$,

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\tau\in\mathcal{T}} |a_n \partial_\theta H_n(\theta^*,\tau)|\right)^{M_1}\right] < \infty.$$

For $M_2 = L(1 - 2\beta_2 - \rho_2)^{-1}$,

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta,\ \tau\in\mathcal{T}} b_n^{\frac{1}{2}-\beta_2}|Y_n(\theta,\tau;\theta^*) - Y(\theta,\tau;\theta^*)|\right)^{M_2}\right] < \infty.$$

[P4] The matrix $\Gamma(\tau;\xi^*)$ is positive definite uniformly in $\tau \in \mathcal{T}$.

[P5] There exists a deterministic and positive number χ such that

$$Y(\theta,\tau;\theta^*) = Y(\theta,\tau;\theta^*) - Y(\theta^*,\tau;\theta^*) \leq -\chi|\theta-\theta^*|^2$$

for all $\theta \in \Theta$ and all $\tau \in \mathcal{T}$.

The following theorems are Theorems 3, 5 and 10 of Yoshida [54]. Here we give a simplified version of them.

Theorem 2.3. *Suppose that [P1]-[P5] are satisfied. Then there exists a constant $C_L > 0$ such that*

$$P_{\xi^*}\left[\sup_{(u,\tau)\in V_n(r)\times\mathcal{T}} Z_n(u,\tau;\theta^*) \geq e^{-r^{2-(\rho_1\vee\rho_2)/2}}\right] \leq \frac{C_L}{r^L} \tag{2.8}$$

for all $n \in \mathbb{N}$ and $r > 0$ where $\sup\emptyset = -\infty$, and

$$V_n(r) = \{u \in \mathbb{R}^m; \theta^* + a_n u \in \Theta, |u| \geq r\}.$$

Let $\Theta_k \subset \mathbb{R}^{m_k}$ be bounded open set $(k = 1, 2, \dots, K)$, $\Theta = \Theta_1 \times \Theta_2 \times \dots \times \Theta_K$, $m = \sum_{k=1}^K m_k$. Let $\{a_n^k\}_{n\in\mathbb{N}}$ be positive sequence such that $a_n^k \to 0$ $(n\to\infty)$, $a_n = diag(a_n^1 I_{m_1}, \dots, a_n^K I_{m_K})$. Set

$$Z_n(u;\theta^*) \quad = \quad \exp\{H_n(\theta^* + a_n u) - H_n(\theta^*)\} \quad (u \in U_n),$$

where I_l denotes a unit matrix of rank l. We extend $Z_n(\cdot;\theta^*)$ to a function in $\hat{C}(\mathbb{R}^m)$ so that its norm is not greater than that of $Z_n(\cdot;\theta^*)$, and denote it by the same symbol. Let $\hat{\theta}_n$ be a random variable and $H_n(\hat{\theta}_n) = \max_{\theta\in\bar{\Theta}} H_n(\theta)$, and let $\hat{u}_n = ((a_n)^{-1}(\hat{\theta}_n - \theta^*))$.

Write $B(R) = \{u \in \mathbb{R}^d; |u| \leq R\}$ for $R > 0$.

Theorem 2.4. *Assume the following conditions.*

(1) *There exists a random function $Z(\cdot;\theta^*)$ in $\hat{C}(\mathbb{R}^m)$ such that for every $R > 0$, $Z_n(\cdot;\theta^*) \to^d Z(\cdot;\theta^*)$ in $C(B(R))$ as $n \to \infty$.*

(2) *There exists a measurable mapping $\hat{u}$ that is a unique maximum point of $Z(\cdot;\theta^*)$ a.s.*

Moreover, we assume that for any $L > 0$, $\limsup_{n\to\infty} ||\hat{u}_n||_L < \infty$.

Then $\hat{u}_n \to^d \hat{u}$ as $n \to \infty$, and $E[\mathsf{f}(\hat{u}_n)] \to \mathbb{E}[\mathsf{f}(\hat{u})]$ as $n \to \infty$ for any continuous function f of at most polynomial growth.

The adaptive Bayes type estimator for parameters θ_k $(k = 1, \dots, K)$ is generally defined as follows. Let $\pi_{k,n}$ be a prior density of the parameter θ_k for each $k = 1, \dots, K$. Let $\underline{\theta}_k = (\theta_1, \theta_2, \dots, \theta_k)$ and $\bar{\theta}_k = (\theta_k, \theta_{k+1}, \dots, \theta_K)$. We assume $0 < \inf_{\theta_k\in\Theta_k, n\in\mathbb{N}} \pi_{k,n} \leq \sup_{\theta_k\in\Theta_k, n\in\mathbb{N}} \pi_{k,n} < \infty$, and $\{\pi_{k,n}\}_{n\in\mathbb{N}}$ is equicontinuous $(k = 1, \dots, K)$. Then the adaptive Bayes type estimators $(\tilde{\theta}_{k,n})_{k=1,\dots,K}$ for parameters $(\theta_k)_{k=1,\dots,K}$ with respect to the quadratic loss function are defined inductively by

$$\begin{aligned}\tilde{\theta}_{k,n} \quad &= \quad \left\{\int_{\Theta_k} \exp\left(H_n(\underline{\tilde{\theta}}_{k-1,n}, \theta_k, \bar{\theta}^\star_{k+1})\right)\pi_{k,n}(\theta_k)d\theta_k\right\}^{-1} \\ &\quad\times \int_{\Theta_k} \theta_k \exp\left(H_n(\underline{\tilde{\theta}}_{k-1,n}, \theta_k, \bar{\theta}^\star_{k+1})\right)\pi_{k,n}(\theta_k)d\theta_k,\end{aligned}$$

where $\bar{\theta}^\star_{k+1}$ is a known dummy value of $\bar{\theta}_{k+1}$. By convention, we neglect $\underline{\theta}_0$ and $\bar{\theta}_{K+1}$.

Let $\Theta, \{\Theta_k\}_{k=1}^K, H_n(\theta)$ and $\{a_n^k\}_{n\in\mathbb{N}}$ be the same setting as above and we assume Θ_k is convex for $k = 1, 2, \dots, K$. We denote $V_n^k(r, \theta_k^*) = \{u_k \in \mathbb{R}^{m_k}; \theta_k^* + a_n^k u_k \in \Theta_k, |u_k| \geq r\}$, $u = (u_1, \dots, u_K)$, and

$$Z_n^k(u_k; \underline{\theta}_{k-1}, \theta_k^*, \bar{\theta}_{k+1}) = \exp\{H_n(\underline{\theta}_{k-1}, \theta_k^* + a_n^k u_k, \bar{\theta}_{k+1}) - H_n(\underline{\theta}_{k-1}, \theta_k^*, \bar{\theta}_{k+1})\}.$$

Let $(\tilde{\theta}_{k,n})_{k=1,\dots,K}$ be adaptive Bayes type estimators for parameters $(\theta_k)_{k=1,\dots,K}$ with prior densities $\{\pi_{k,n}\}_{n\in\mathbb{N},1\le k\le K}$. We denote

$$
\begin{aligned}
\tilde{u}_n^k &= (a_n^k)^{-1}(\tilde{\theta}_{k,n}-\theta_k^*)\\
&= \left(\int_{\mathbb{U}_n^k(\theta_k^*)} Z_n^k(u_k;\underline{\tilde{\theta}}_{k-1,n},\theta_k^*,\bar{\theta}_{k+1}^\star)\pi_{k,n}(\theta_k^*+a_n^k u_k)du_k\right)^{-1}\\
&\quad\times\int_{\mathbb{U}_n^k(\theta_k^*)} u_k Z_n^k(u_k;\underline{\tilde{\theta}}_{k-1,n},\theta_k^*,\bar{\theta}_{k+1}^\star)\pi_{k,n}(\theta_k^*+a_n^k u_k)du_k,\\
\tilde{u}_k &= \left(\int_{\mathbb{R}^{m_k}} Z^k(u_k;\theta^*)du_k\right)^{-1}\int_{\mathbb{R}^{m_k}} u_k Z^k(u_k;\theta^*)du_k,
\end{aligned}
$$

$\tilde{u}_n=(\tilde{u}_n^1,\dots,\tilde{u}_n^K)$ and $\tilde{u}=(\tilde{u}_1,\dots,\tilde{u}_K)$, where $\mathbb{U}_n^k(\theta_k^*)=\{u_k\in\mathbb{R}^{m_k};\theta_k^*+a_n^k u_k\in\Theta_k\}$ and Z^k is a random field $(k=1,\dots,K)$.

Theorem 2.5. *Assume the following conditions.*

(1) *For every $k=1,\dots,K$, $Z^k(\cdot;\theta^*)\in\hat{C}(\mathbb{R}^{m_k})$ and for every $R>0$, $(Z_n^k(u_k;\underline{\tilde{\theta}}_{k-1,n},\theta_k^*,\bar{\theta}_{k+1}^\star))_{k=1,\dots,K}\to^d (Z^k(u_k;\theta^*))_{k=1,\dots,K}$ in $C(\{u;|u|\le R\})$ as $n\to\infty$.*

(2) *For any $L>1$, there exists $C_L>0$ such that*

$$
P_{\xi^*}\left[\sup_{u_k\in V_n^k(r,\theta_k^*)} Z_n^k(u_k;\underline{\tilde{\theta}}_{k-1,n},\theta_k^*,\bar{\theta}_{k+1}^\star)\ge e^{-\frac{r}{2}}\right]\le\frac{C_L}{r^L}
$$

for all $n\in\mathbb{N}$, $r>0$, and $k=1,2,\dots,K$.

(3) *For some $N\in\mathbb{N}$,*

$$
\sup_{n\ge N} E\left[\left(\int_{\mathbb{U}_n^k(\theta_k^*)} Z_n^k(u_k;\underline{\tilde{\theta}}_{k-1,n},\theta_k^*,\bar{\theta}_{k+1}^\star)\pi_{k,n}(\theta_k^*+a_n u_k)du\right)^{-1}\right]<\infty.
$$

Then

$$
\tilde{u}_n\to^d\tilde{u}\quad\text{as}\quad n\to\infty,
$$

and

$$
E[\mathsf{f}(\tilde{u}_n)]\to\mathbb{E}[\mathsf{f}(\tilde{u})],
$$

as $n\to\infty$ for any continuous function f of at most polynomial growth.

Theorem 2.3 implies that the polynomial type large deviation inequality (2.8) is obtained by some moment conditions on the contrast function H_n and its derivatives and regularity conditions $[P4]$ and $[P5]$. The polynomial type large deviation inequality will be necessary in application of Theorem 2.4 to check that $\limsup\parallel\hat{u}_n\parallel_L<\infty$ for any $L>0$ and in application of Theorem 2.5 to check Condition (2) above. These conditions are immediate consequence of (2.8) and control a probability that $|\hat{u}_n|$ becomes large. This control plays an essential role in the proof of convergence of moments of any order for $\hat{u}_n$ and the convergence of the Bayes type estimator. The rest of this chapter is mainly devoted to verifying the moment conditions $[P1]$ and $[P3]$ for parameters σ and θ in order to obtain the polynomial type large deviation inequality.

2.3.2 Proof of Theorems 2.1 and 2.2

As the first step, we will apply Theorem 2.3 to σ as "θ" and θ as "τ" there. In this case, "Γ_n" and "Y_n" in Theorem 2.3 are as follows:

$$\begin{aligned}\Gamma_n^1(\sigma,\theta) &= -\frac{1}{n}\partial_\sigma^2 H_n(\sigma,\theta)\\ &= \frac{1}{2nh}\sum_{i=1}^n \bar{X}_{i,n}^T(\theta)\partial_\sigma^2\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad+\frac{1}{2n}\sum_{i=1}^n \partial_\sigma^2\log\det\beta_{i-1}(\sigma)1_{\{|\Delta X_i^n|\le Dh^\rho\}},\end{aligned}$$

and

$$\begin{aligned}Y_n(\sigma,\theta;\sigma^*) &= \frac{1}{n}\{H_n(\sigma,\theta)-H_n(\sigma^*,\theta)\}\\ &= -\frac{1}{2nh}\sum_{i=1}^n \bar{X}_{i,n}^T(\theta)\{\beta_{i-1}^{-1}(\sigma)-\beta_{i-1}^{-1}(\sigma^*)\}\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad-\frac{1}{2n}\sum_{i=1}^n\log\frac{\det\beta_{i-1}(\sigma)}{\det\beta_{i-1}(\sigma^*)}1_{\{|\Delta X_i^n|\le Dh^\rho\}}.\end{aligned}$$

At the second stage of the proof, we will consider the random field $\theta\mapsto H_n(\hat{\sigma}_n,\theta)$ for the estimator of θ. When applying Theorem 2.3, as "Γ_n" and "Y_n", we take the ones given by

$$\begin{aligned}\Gamma_n^2(\theta) &= -\frac{1}{nh}\partial_\theta^2 H_n(\hat{\sigma}_n,\theta)\\ &= -\frac{1}{nh}\sum_{i=1}^n\{\partial_\theta^2 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\hat{\sigma}_n)\bar{X}_{i,n}(\theta)-h\partial_\theta a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\hat{\sigma}_n)\partial_\theta a_{i-1}(\theta)\}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad-\frac{1}{nh}\sum_{i=1}^n\{\partial_\theta^2\log\Psi_\theta(\Delta X_i^n,X_{t_{i-1}^n})\}\varphi_n(X_{t_{i-1}^n},\Delta X_i^n)1_{\{|\Delta X_i^n|>Dh^\rho\}}\\ &\quad+\frac{1}{n}\sum_{i=1}^n\int_{B(X_{t_{i-1}^n})}\partial_\theta^2\Psi_\theta(y,X_{t_{i-1}^n})\varphi_n(X_{t_{i-1}^n},y)dy,\end{aligned}$$

and

$$\begin{aligned}Y_n(\theta;\alpha^*) &= \frac{1}{nh}\left(H_n(\hat{\sigma}_n,\theta)-H_n(\hat{\sigma}_n,\theta^*)\right)\\ &= -\frac{1}{2nh^2}\sum_{i=1}^n(\bar{X}_{i,n}^T(\theta)\beta_{i-1}^{-1}(\hat{\sigma}_n)\bar{X}_{i,n}(\theta)-\bar{X}_{i,n}^T\beta_{i-1}^{-1}(\hat{\sigma}_n)\bar{X}_{i,n})1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad+\frac{1}{nh}\sum_{i=1}^n\{\log\Psi_\theta(\Delta X_i^n,X_{t_{i-1}^n})-\log\Psi_{\theta^*}(\Delta X_i^n,X_{t_{i-1}^n})\}\varphi_n(X_{t_{i-1}^n},\Delta X_i^n)1_{\{|\Delta X_i^n|>Dh^\rho\}}\\ &\quad-\frac{1}{n}\sum_{i=1}^n\int_{B(X_{t_{i-1}^n})}(\Psi_\theta(y,X_{t_{i-1}^n})-\Psi_{\theta^*}(y,X_{t_{i-1}^n}))\varphi_n(X_{t_{i-1}^n},y)dy.\end{aligned}$$

To prove main theorems, we need several lemmas. Let $J_i^n=p((t_{i-1}^n,t_i^n]\times E)$, $Z_t=\int_{[0,t]\times E}z\,p(dt,dz)$, and $\epsilon:=\frac{1}{6}$, then $n^{2\epsilon}\preceq nh$ for large n. We set $\tau_i^n=\inf\{t;|\Delta X_t|>0,t_{i-1}^n<t\le t_i^n\}$ and $\nu_i^n=\sup\{t;|\Delta X_t|>0,t_{i-1}^n<t\le t_i^n\}$. If the infimum or supremum on the right-hand side does not exist, then we define the random times to equal t_i^n. Let $C_{i,0}^n=\{J_i^n=0,|\Delta X_i^n|\le Dh^\rho\}$, $C_{i,1}^n=\{J_i^n=1,|\Delta X_i^n|\le Dh^\rho\}$, $C_{i,2}^n=\{J_i^n\ge 2,|\Delta X_i^n|\le Dh^\rho\}$, $D_{i,0}^n=\{J_i^n=0,|\Delta X_i^n|>Dh^\rho\}$, $D_{i,1}^n=\{J_i^n=1,|\Delta X_i^n|>Dh^\rho\}$, $D_{i,2}^n=\{J_i^n\ge 2,|\Delta X_i^n|>Dh^\rho\}$.

Lemma 2.1. (Shimizu and Yoshida [46]) *Assume* $[H1], [H3]$ *and* $[H5]$. *Then for* $0 \le \rho < 1/2$, $D > 0$ *and any* $p \ge 1$,

$$\begin{aligned} P_{\alpha^*}\left[\sup_{t_{i-1}^n \le t < \tau_i^n} |X_t - X_{t_{i-1}^n}| > \frac{Dh^\rho}{2} |\mathcal{F}_{i-1}\right] &\le R(\alpha, h^p, X_{t_{i-1}^n}), \\ P_{\alpha^*}\left[\sup_{\nu_i^n \le t < t_i^n} |X_{t_i^n} - X_t| > \frac{Dh^\rho}{2} |\mathcal{F}_{i-1}\right] &\le R(\alpha, h^p, X_{t_{i-1}^n}), \end{aligned}$$

where $\sup \emptyset = -\infty$. *Each function* R *does not depend on* i.

Remark 2.3. $[H5]$ *is a little weaker condition than the corresponding condition in Shimizu and Yoshida [46]. However, by reading the proof of the corresponding lemma in Shimizu and Yoshida [46] carefully, we can verify that [H5] is a sufficient condition to prove Lemma 2.1. A similar argument holds for [H4] and [H5] of Lemma 2.3.*

Lemma 2.2. *Assume* $[H1], [H3], [H5], [H6]$ *and* $[H7]$. *Let* $\frac{3}{8} + b \le \rho < \frac{1}{2}$, *where* b *is the constant appearing in* (2.2). *Then for any* $p \ge 1$, *as* $n \to \infty$

$$\begin{aligned} P_{\alpha^*}[C_{i,0}^n|\mathcal{F}_{i-1}] &= \tilde{R}(\alpha, h, X_{t_{i-1}^n}) \\ P_{\alpha^*}[D_{i,0}^n|\mathcal{F}_{i-1}] &= R(\alpha, h^p, X_{t_{i-1}^n}) \\ P_{\alpha^*}[C_{i,1}^n|\mathcal{F}_{i-1}] &= R(\alpha, h^{11/8+b}, X_{t_{i-1}^n}) \\ P_{\alpha^*}[D_{i,1}^n|\mathcal{F}_{i-1}] &= \lambda(\alpha^*) h \tilde{R}(\alpha, h^{3/8+b}, X_{t_{i-1}^n}) \\ P_{\alpha^*}[C_{i,2}^n|\mathcal{F}_{i-1}] &\le \lambda(\alpha^*)^2 h^2 \\ P_{\alpha^*}[D_{i,2}^n|\mathcal{F}_{i-1}] &\le \lambda(\alpha^*)^2 h^2. \end{aligned}$$

Proof. The proof is almost the same as the proof of Lemma 2.2. in Shimizu and Yoshida [46]. First, it is obvious that $P_{\alpha^*}[C_{i,2}^n|\mathcal{F}_{i-1}] \le \lambda(\alpha^*)^2 h^2$, and $P_{\alpha^*}[D_{i,2}^n|\mathcal{F}_{i-1}] \le \lambda(\alpha^*)^2 h^2$. On $C_{i,1}^n$,

$$\begin{aligned} P_{\alpha^*}[C_{i,1}^n|\mathcal{F}_{i-1}] \le\ & P_{\alpha^*}\Big[|(X_{t_i^n} - X_{\tau_i^n}) + (X_{\tau_i^n -} - X_{t_{i-1}^n}) + \Delta X_{\tau_i^n}| \le Dh^\rho, |\Delta Z_{\tau_i^n}| > \frac{2Dh^\rho}{c_0}, J_i^n = 1|\mathcal{F}_{i-1}\Big] \\ & + P_{\alpha^*}\Big[|\Delta Z_{\tau_i^n}| \le \frac{2Dh^\rho}{c_0}, J_i^n = 1|\mathcal{F}_{i-1}\Big], \end{aligned}$$

where $\Delta Z_{\tau_i^n}$ has density F_{θ^*} under $\mathcal{F}_{i-1}$ and c_0 is the constant in condition [H7]. If $|(X_{t_i^n} - X_{\tau_i^n}) + (X_{\tau_i^n -} - X_{t_{i-1}^n}) + \Delta X_{\tau_i^n}| \le Dh^\rho$ and $|\Delta X_{\tau_i^n}|$ is small enough, then by [H7], we have

$$|X_{t_i^n} - X_{\tau_i^n}| + |X_{\tau_i^n -} - X_{t_{i-1}^n}| \ge c_0 |\Delta Z_{\tau_i^n}| - Dh^\rho.$$

Therefore, by Lemma 2.1, we have for large n,

$$\begin{aligned} P_{\alpha^*}[C_{i,1}^n|\mathcal{F}_{i-1}] &\le P_{\alpha^*}\Big[\sup_{t \in [t_{i-1}^n, \tau_i^n)} |X_t - X_{t_{i-1}^n}| + \sup_{t \in [\nu_i^n, t_i^n]} |X_{t_i^n} - X_t| > Dh^\rho |\mathcal{F}_{i-1}\Big] \\ &\quad + \lambda(\alpha^*) h e^{-\lambda(\alpha^*) h} \int_{|z| \le 2Dh^\rho / c_0} \frac{K|z|^{1-d}}{\lambda(\alpha^*)} dz \\ &\le R(\alpha, h^p, X_{t_{i-1}^n}) + Ch^{\rho+1} \\ &= R(\alpha, h^{11/8+b}, X_{t_{i-1}^n}) \end{aligned}$$

if we take $p \ge \rho + 1$.

For $D_{i,0}^n$, by applying Lemma 2.1 again, we have

$$\begin{aligned} P_{\alpha^*}[D_{i,0}^n|\mathcal{F}_{i-1}] &= P_{\alpha^*}[|X_{\tau_i^n} - X_{t_{i-1}^n}| > Dh^\rho, \tau_i^n = t_i^n|\mathcal{F}_{i-1}] \\ &= R(\alpha, h^p, X_{t_{i-1}^n}). \end{aligned}$$

Finally,

$$
\begin{aligned}
P_{\alpha^*}[C^n_{i,0}|\mathcal{F}_{i-1}] &= P_{\alpha^*}[J^n_i = 0|\mathcal{F}_{i-1}] - P_{\alpha^*}[D^n_{i,0}|\mathcal{F}_{i-1}] \\
&= e^{-\lambda(\alpha^*)h} - R(\alpha, h^p, X_{t^n_{i-1}}) \\
&= \tilde{R}(\alpha, h, X_{t^n_{i-1}}),
\end{aligned}
$$

and

$$
\begin{aligned}
P_{\alpha^*}[D^n_{i,1}|\mathcal{F}_{i-1}] &= P_{\alpha^*}[J^n_i = 1|\mathcal{F}_{i-1}] - P_{\alpha^*}[C^n_{i,1}|\mathcal{F}_{i-1}] \\
&= \lambda(\alpha^*)he^{-\lambda(\alpha^*)h}\tilde{R}(\alpha, h^{3/8+b}, X_{t^n_{i-1}}).
\end{aligned}
$$

□

Lemma 2.3. (Shimizu and Yoshida [46]) *Assume* $[H1], [H3] - [H7]$. *Then for* $k_j = 1, 2, \ldots, d$ $(j = 1, 2, 3, 4)$,

$$
\begin{aligned}
E[\bar{X}^{(k_1)}_{i,n} 1_{C^n_{i,0}}|\mathcal{F}_{i-1}] &= R(\alpha, h\sqrt{h}, X_{t^n_{i-1}}), \\
E[\bar{X}^{(k_1)}_{i,n}\bar{X}^{(k_2)}_{i,n} 1_{C^n_{i,0}}|\mathcal{F}_{i-1}] &= h\beta^{(k_1,k_2)}_{i-1}(\sigma^*) + R(\alpha, h^2, X_{t^n_{i-1}}), \\
E[\bar{X}^{(k_1)}_{i,n}\bar{X}^{(k_2)}_{i,n}\bar{X}^{(k_3)}_{i,n} 1_{C^n_{i,0}}|\mathcal{F}_{i-1}] &= R(\alpha, h^2, X_{t^n_{i-1}}), \\
E[\bar{X}^{(k_1)}_{i,n}\bar{X}^{(k_2)}_{i,n}\bar{X}^{(k_3)}_{i,n}\bar{X}^{(k_4)}_{i,n} 1_{C^n_{i,0}}|\mathcal{F}_{i-1}] &= h^2(\beta^{(k_1,k_2)}_{i-1}\beta^{(k_3,k_4)}_{i-1} + \beta^{(k_1,k_3)}_{i-1}\beta^{(k_2,k_4)}_{i-1} + \beta^{(k_1,k_4)}_{i-1}\beta^{(k_2,k_3)}_{i-1}) \\
&\quad + R(\alpha, h^3, X_{t^n_{i-1}}),
\end{aligned}
$$

where $\bar{X}^{(k)}_{i,n}$ *and* $\beta^{(k,l)}_{i-1}$ *denote the elements of the vector* $\bar{X}_{i,n}$ *and the matrix* β_{i-1},*respectively* $(1 \leq k, l \leq d)$.

Before proceeding to the next step, since we have defined several parameters and their relationships, we list up those relations again for convenience of reference:

$$
\begin{aligned}
& n^{-3/5} \preceq h \preceq n^{-4/7},\ 0 < b < 1/8,\ \epsilon_n \to 0,\ \frac{\sqrt{n}h}{\epsilon_n^2} \vee \frac{h^b}{\epsilon_n} \preceq 1, \\
& 1 \preceq n^3h^4\epsilon_n^{16},\ \frac{3}{8} + b \leq \rho < \frac{1}{2},\ \epsilon = \frac{1}{6},\ \text{and } n^{2\epsilon} \preceq nh,
\end{aligned}
$$

as $n \to \infty$.

Proposition 2.1. (Yoshida [54]) *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space,* $\{F_j\}_{j\in\mathbb{N}}$ *be a stationary process with mean* 0 *and suppose that for some* $0 < h < 1$ *and* $C > 0$,

$$
\sup_{j\in\mathbb{N}} \sup_{A\in\sigma[F_l;l\leq j], B\in\sigma[F_l;l\geq j+k]} |P[A\cap B] - P[A]P[B]| \leq C\exp(-hk)
$$

for all $k \in \mathbb{N}$ *and that for every* $p \geq 2$, $\sup_{j\in\mathbb{N}} ||F_j||_p \leq C_p$ *for some constant* C_p *depending on* p *but independent of* h. *Then for some constant* $C' = C'(C, p, C_{p+1}) < \infty$ *independent of* h *and the sequence* $\{F_j\}$,

$$
E\left[\sup_{j=1,\ldots,n}\left|\sum_{i=1}^j F_i\right|^p\right] \leq C'\left[(nh^{-1})^{\frac{p}{2}} + nh^{1-p}\right]
$$

for all $n \in \mathbb{N}$.

The following proposition is stronger than the ergodic property for the sum of the function of the jump-diffusion process with the exponential mixing property.

Proposition 2.2. *Suppose that* $[H2]$ *and* $[H3]$ *are satisfied, and Borel functions* $F_n : \mathbb{R}^d \times \mathbb{R}^d \times \Xi$ *satisfy* $|F_n(x, y, \alpha)| \leq C(1 + |x|)^C$ $(n \in \mathbb{N}, x, y \in \mathbb{R}^d, \alpha \in \Xi)$ *for some constant* $C > 0$. *Then for every* $p \geq 2$,

$$
\sup_{\alpha\in\Xi}\sup_{n\in\mathbb{N}} E\left[\left|n^\epsilon\frac{1}{n}\sum_{i=1}^n \left\{F_n(X_{t^n_{i-1}}, \Delta X^n_i, \alpha) - E[F_n(X_{t^n_{i-1}}, \Delta X^n_i, \alpha)]\right\}\right|^p\right] \quad < \quad \infty.
$$

Proof. By Proposition 2.1, we have

$$\begin{aligned}
&\sup_{\alpha\in\Xi}\sup_{n\in\mathbb{N}} E\left[\left|n^{\epsilon}\frac{1}{n}\sum_{i=1}^{n}\left\{F_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)-E[F_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)]\right\}\right|^p\right]\\
\le\;& C\sup_{n\in\mathbb{N}} n^{p(-1+\epsilon)}\left\{(nh^{-1})^{\frac{p}{2}}+nh^{1-p}\right\}\\
=\;& C\sup_{n\in\mathbb{N}}\left(\frac{n^{2\epsilon}}{nh}\right)^{\frac{p}{2}}\left\{1+(nh)^{1-\frac{p}{2}}\right\}<\infty.
\end{aligned}$$

□

The following proposition takes an essential role in estimating the third term of H_n, later in Lemmas 2.6 and 2.7. This proposition is the key element of the proof of Theorems 2.1 and 2.2.

Proposition 2.3. *Let $k\in\mathbb{N}$ and $p\ge 2^{k-1}$. Suppose $\{\mathcal{F}_i\}_{0\le i\le n}$ is a filtration on a probability space $(\Omega,\mathcal{F},P)$, and $\{F_i\}_{1\le i\le n}$ is a random sequence adapted to $\{\mathcal{F}_i\}$ such that $E[|F_i|^p]<\infty$ $(1\le i\le n)$. Then*

$$\begin{aligned}
E\left[|\sum_{i=1}^{n}F_i|^p\right] &\le C_{p,k}E\left[|\sum_{i=1}^{n}\psi_i^{k+1}(F_i)|^{\frac{p}{2^k}}\right]+C_{p,k}\sum_{l=1}^{k}E\left[|\sum_{i=1}^{n}E[\psi_i^l(F_i)|\mathcal{F}_{i-1}]|^{\frac{p}{2^{l-1}}}\right],\\
E\left[|\sum_{i=1}^{n}\{F_i-E[F_i|\mathcal{F}_{i-1}]\}|^p\right] &\le C_{p,k}E\left[|\sum_{i=1}^{n}\psi_i^{k+1}(F_i)|^{\frac{p}{2^k}}\right]+C_{p,k}\sum_{l=2}^{k}E\left[|\sum_{i=1}^{n}E[\psi_i^l(F_i)|\mathcal{F}_{i-1}]|^{\frac{p}{2^{l-1}}}\right],
\end{aligned}$$

where $C_{p,k}$ is a constant depending only on p and k, and

$$\begin{aligned}
\psi_i^1(F) &= F\\
\psi_i^{l+1}(F) &= \{\psi_i^l(F)-E[\psi_i^l(F)|\mathcal{F}_{i-1}]\}^2\ (l\in\mathbb{N}).
\end{aligned}$$

For $k=2,3$, we have

$$\begin{aligned}
\psi_i^2(F) &= F^2-2FE[F|\mathcal{F}_{i-1}]+(E[F|\mathcal{F}_{i-1}])^2\\
E[\psi_i^2(F)|\mathcal{F}_{i-1}] &= E[F^2|\mathcal{F}_{i-1}]-(E[F|\mathcal{F}_{i-1}])^2\\
\psi_i^3(F) &= \left\{F^2-2FE[F|\mathcal{F}_{i-1}]+2(E[F|\mathcal{F}_{i-1}])^2-E[F^2|\mathcal{F}_{i-1}]\right\}^2\\
&= F^4-4F^3E[F|\mathcal{F}_{i-1}]+8F^2(E[F|\mathcal{F}_{i-1}])^2-8F(E[F|\mathcal{F}_{i-1}])^3+4(E[F|\mathcal{F}_{i-1}])^4\\
&\quad -2F^2E[F^2|\mathcal{F}_{i-1}]+4FE[F|\mathcal{F}_{i-1}]E[F^2|\mathcal{F}_{i-1}]-4(E[F|\mathcal{F}_{i-1}])^2E[F^2|\mathcal{F}_{i-1}]\\
&\quad +(E[F^2|\mathcal{F}_{i-1}])^2.
\end{aligned}$$

These equations are used repeatedly later to apply Proposition 2.3 for $k=2$.

Proof of Proposition 2.3. We will prove the second inequality by induction on k. The first one is a trivial consequence of the second inequality. For $k=1$, since $\{\sum_{i=1}^m\{F_i-E[F_i|\mathcal{F}_{i-1}]\}\}_{0\le m\le n}$ is a martingale, by the Burkholder-Davis-Gundy inequality,

$$E\left[|\sum_{i=1}^{n}\{F_i-E[F_i|\mathcal{F}_{i-1}]\}|^p\right]\le C_pE\left[|\sum_{i=1}^{n}\{F_i-E[F_i|\mathcal{F}_{i-1}]\}^2|^{\frac{p}{2}}\right].$$

Suppose the second inequality holds for k. Suppose that $p\ge 2^k$. Then since $\{\sum_{i=1}^m\{\psi_i^{k+1}(F_i)-E[\psi_i^{k+1}(F_i)|\mathcal{F}_{i-1}]\}\}_{0\le m\le n}$ is martingale, by the Burkholder-Davis-Gundy inequality and induction hypothesis, we obtain

$$
\begin{aligned}
E\left[\left|\sum_{i=1}^{n}\{F_i - E[F_i|\mathcal{F}_{i-1}]\}\right|^p\right] &\leq C_{p,k}E\left[\left|\sum_{i=1}^{n}\psi_i^{k+1}(F_i)\right|^{\frac{p}{2^k}}\right] + C_{p,k}\sum_{l=2}^{k}E\left[\left|\sum_{i=1}^{n}E[\psi_i^l(F_i)|\mathcal{F}_{i-1}]\right|^{\frac{p}{2^{l-1}}}\right] \\
&\leq C_{p,k}2^{\frac{p}{2^k}}E\left[\left|\sum_{i=1}^{n}\{\psi_i^{k+1}(F_i) - E[\psi_i^{k+1}(F_i)|\mathcal{F}_{i-1}]\}\right|^{\frac{p}{2^k}}\right] \\
&\quad + C_{p,k}2^{\frac{p}{2^k}}E\left[\left|\sum_{i=1}^{n}E[\psi_i^{k+1}(F_i)|\mathcal{F}_{i-1}]\right|^{\frac{p}{2^k}}\right] + C_{p,k}\sum_{l=2}^{k}E\left[\left|\sum_{i=1}^{n}E[\psi_i^l(F_i)|\mathcal{F}_{i-1}]\right|^{\frac{p}{2^{l-1}}}\right] \\
&\leq C_{p,k+1}E\left[\left|\sum_{i=1}^{n}\{\psi_i^{k+1}(F_i) - E[\psi_i^{k+1}(F_i)|\mathcal{F}_{i-1}]\}^2\right|^{\frac{p}{2^{k+1}}}\right] \\
&\quad + C_{p,k+1}\sum_{l=2}^{k+1}E\left[\left|\sum_{i=1}^{n}E[\psi_i^l(F_i)|\mathcal{F}_{i-1}]\right|^{\frac{p}{2^{l-1}}}\right] \\
&= C_{p,k+1}E\left[\left|\sum_{i=1}^{n}\psi_i^{k+2}(F_i)\right|^{\frac{p}{2^{k+1}}}\right] + C_{p,k+1}\sum_{l=2}^{k+1}E\left[\left|\sum_{i=1}^{n}E[\psi_i^l(F_i)|\mathcal{F}_{i-1}]\right|^{\frac{p}{2^{l-1}}}\right]. \quad \square
\end{aligned}
$$

Proposition 2.4. *Suppose that [H1],[H3],[H5]-[H7] are satisfied. Let $\{u_n\}$ and $\{v_n\}$ be the sequences of positive numbers and $g_n(x, y, \alpha)$ ($n \in \mathbb{N}, \alpha \in \Xi$) be Borel functions. Assume g_n is differentiable with respect to $y \in E$, $\partial_y g_n$ is continous in y, and*

$$
|\partial_y g_n(x, y, \alpha)| \quad \leq \quad C v_n(1 + |y|)^C(1 + |x|)^C,
$$

for $n \in \mathbb{N}$, $\alpha \in \Xi$, $x \in \mathbb{R}^d, y \in E$. Moreover, assume at least one of the following two conditions holds true.

1. *$|g_n(x, y, \alpha)| \leq C u_n(1 + |x|)^C$ ($n \in \mathbb{N}$, $\alpha \in \Xi$, $x \in \mathbb{R}^d, y \in E$).*
2. *$|g_n(x, y, \alpha)| \leq C(1 + |y|)^C(1 + |x|)^C$ ($n \in \mathbb{N}$, $\alpha \in \Xi$, $x \in \mathbb{R}^d, y \in E$) and there exists $p > 0$ such that $u_n^{-p} \preceq h$.*

Then

$$
\begin{aligned}
\frac{1}{h}E[g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)1_{\{|\Delta X_i^n| > Dh^\rho\}}|\mathcal{F}_{i-1}] &= \int_{B(X_{t_{i-1}^n})} g_n(X_{t_{i-1}^n}, y, \alpha)\Psi_{\theta^*}(y, X_{t_{i-1}^n})dy \\
&\quad + R(\alpha, h^{3/8+b}u_n \vee \sqrt{h}v_n, X_{t_{i-1}^n}) \ (n \in \mathbb{N}).
\end{aligned}
$$

Proof. The proof is similar to that of Proposition 3.6 of Shimizu and Yoshida [46]. First, by Proposition 3.1 of Shimizu and Yoshida [46], it holds for $k \in \mathbb{N}, k \geq 2$ that

$$
E[|\Delta X_i^n|^k|\mathcal{F}_{i-1}] = R(\alpha, h, X_{t_{i-1}^n}).
$$

So if Condition 2 holds, then

$$
\begin{aligned}
E[|g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)|1_{\{|g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)| > u_n\}}|\mathcal{F}_{i-1}] &\leq u_n^{1-2p}E[|g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)|^{2p}|\mathcal{F}_{i-1}] \\
&\leq R(\alpha, h^2 u_n, X_{t_{i-1}^n})E[(1 + |\Delta X_i^n|)^{2Cp}|\mathcal{F}_{i-1}] \\
&= R(\alpha, h^2 u_n, X_{t_{i-1}^n}). \qquad (2.9)
\end{aligned}
$$

Therefore by Lemma 2.2, we have

$$
\begin{aligned}
& E\left[|g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)|1_{D_{i,0}^n \cup D_{i,2}^n}|\mathcal{F}_{i-1}\right] \\
= \ & E\left[|g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)|(1_{(D_{i,0}^n \cup D_{i,2}^n) \cap \{|g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)| > u_n\}} + 1_{(D_{i,0}^n \cup D_{i,2}^n) \cap \{|g_n(X_{t_{i-1}^n}, \Delta X_i^n, \alpha)| \leq u_n\}})|\mathcal{F}_{i-1}\right] \\
= \ & R(\alpha, h^2 u_n, X_{t_{i-1}^n}). \qquad (2.10)
\end{aligned}
$$

Consequently we obtain

$$\begin{aligned}
&\frac{1}{h}E\left[g_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)1_{\{|\Delta X_i^n|>Dh^\rho\}}|\mathcal{F}_{i-1}\right] \\
= \ &\frac{1}{h}E\left[g_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)1_{D_{i,1}^n}|\mathcal{F}_{i-1}\right]+R(\alpha,hu_n,X_{t_{i-1}^n}). \qquad (2.11)
\end{aligned}$$

In the case that Condition 1 holds, Lemma 2.2 leads (2.11).

In both cases, we have for $q\in\mathbb{N}$ and $q\geq 2$,

$$E[|X_{t_i^n}-X_{\tau_i^n}|^q\frac{1_{\{J_i^n=1\}}}{P_{\alpha^*}[J_i^n=1]}|\mathcal{F}_{i-1}]=R(\alpha,h,X_{t_{i-1}^n}),$$

and

$$E[|X_{\tau_i^n-}-X_{t_{i-1}^n}|^q\frac{1_{\{J_i^n=1\}}}{P_{\alpha^*}[J_i^n=1]}|\mathcal{F}_{i-1}]=R(\alpha,h,X_{t_{i-1}^n}).$$

Let $\xi_i^n(t)=t\Delta X_i^n+(1-t)\Delta X_{\tau_i^n}$ and $G_i^n=\{|\Delta X_i^n|\leq|X_{t_i^n}-X_{\nu_i^n}|+|X_{\tau_i^n-}-X_{t_{i-1}^n}|\}$, then it holds that $\xi_i^n(t)\neq 0$ $(0\leq t\leq 1)$ on $D_{i,1}^n\cap(G_i^n)^c$. So by the Cauchy-Schwartz inequality, we have

$$\begin{aligned}
&\frac{1}{h}E\left[\left|g_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)-g_n(X_{t_{i-1}^n},\Delta X_{\tau_i^n},\alpha)\right|1_{D_{i,1}^n\cap(G_i^n)^c}\middle|\mathcal{F}_{i-1}\right] \\
\leq \ &\lambda(\alpha^*)\Big(E\Big[|\int_0^1\partial_y g_n(X_{t_{i-1}^n},\xi_i^n(t),\alpha)dt|^2\frac{1_{D_{i,1}^n\cap(G_i^n)^c}}{P[J_i^n=1]}|\mathcal{F}_{i-1}\Big]\Big)^{\frac{1}{2}} \\
&\times\Big(E\Big[(|X_{t_i^n}-X_{\tau_i^n}|+|X_{\tau_i^n-}-X_{t_{i-1}^n}|)^2\frac{1_{D_{i,1}^n\cap(G_i^n)^c}}{P[J_i^n=1]}|\mathcal{F}_{i-1}\Big]\Big)^{\frac{1}{2}} \\
= \ &R(\alpha,\sqrt{h}v_n,X_{t_{i-1}^n}). \qquad (2.12)
\end{aligned}$$

Moreover, by Lemma 2.1, we have for any $p>0$,

$$P_{\alpha^*}[D_{i,1}^n\cap G_i^n|\mathcal{F}_{i-1}]\leq P_{\alpha^*}[|X_{t_i^n}-X_{\nu_i^n}|+|X_{\tau_i^n-}-X_{t_{i-1}^n}|>Dh^\rho|\mathcal{F}_{i-1}]\leq R(\alpha,h^p,X_{t_{i-1}^n}). \qquad (2.13)$$

Therefore by similar argument to the derivation of (2.11) with an equation

$$E[(|X_{\tau_i^n-}-X_{t_{i-1}^n}|+|X_{t_i^n}-X_{\nu_i^n}|)^p|\mathcal{F}_{i-1}]=R(\alpha,h,X_{t_{i-1}^n})=R(\alpha,1,X_{t_{i-1}^n}),$$

where $p\geq 2$, we obtain

$$\frac{1}{h}E\left[\left|g_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)-g_n(X_{t_{i-1}^n},\Delta X_{\tau_i^n},\alpha)\right|1_{D_{i,1}^n\cap G_i^n}\middle|\mathcal{F}_{i-1}\right]=R(\alpha,hu_n,X_{t_{i-1}^n}). \qquad (2.14)$$

Then (2.11), (2.12), and (2.14) yield

$$\begin{aligned}
&\frac{1}{h}E\left[g_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)1_{\{|\Delta X_i^n|>Dh^\rho\}}|\mathcal{F}_{i-1}\right] \\
= \ &\frac{1}{h}E\Big[g_n(X_{t_{i-1}^n},\Delta X_{\tau_i^n},\alpha)1_{D_{i,1}^n}|\mathcal{F}_{i-1}\Big]+R(\alpha,hu_n\vee\sqrt{h}v_n,X_{t_{i-1}^n}). \qquad (2.15)
\end{aligned}$$

The equation (2.15), Lemma 2.2 and a similar argument to the derivation of (2.14) yield

$$\begin{aligned}
&\frac{1}{h}E\left[g_n(X_{t_{i-1}^n},\Delta X_i^n,\alpha)1_{\{|\Delta X_i^n|>Dh^\rho\}}|\mathcal{F}_{i-1}\right] \\
= \ &\frac{1}{h}E\Big[g_n(X_{t_{i-1}^n},\Delta X_{\tau_i^n},\alpha)1_{\{J_i^n=1\}}|\mathcal{F}_{i-1}\Big]+R(\alpha,h^{3/8+b}u_n\vee\sqrt{h}v_n,X_{t_{i-1}^n}). \qquad (2.16)
\end{aligned}$$

Let $\tilde{\xi}_i^n(t) = tc(X_{\tau_i^n-}, \Delta Z_{\tau_i^n}, \theta^*) + (1-t)c_{i-1}(\Delta Z_{\tau_i^n}, \theta^*)$. If $d \geq 2$, we have

$$\begin{aligned} & \frac{1}{h} E\left[\left| g_n(X_{t_{i-1}^n}, c_{i-1}(\Delta Z_{\tau_i^n}, \theta^*), \alpha) - g_n(X_{t_{i-1}^n}, c(X_{\tau_i^n-}, \Delta Z_{\tau_i^n}, \theta^*), \alpha) \right| 1_{\{J_i^n=1\}} \middle| \mathcal{F}_{i-1} \right] \\ = \; & R(\alpha, \sqrt{h} v_n, X_{t_{i-1}^n}), \end{aligned} \tag{2.17}$$

by a similar argument to the derivation of (2.12) and modifying $\tilde{\xi}_i^n(t)$ so that $\tilde{\xi}_i^n$ bypass the origin and the length of modification is $O(h\sqrt{h})$, if necessary. If $d = 1$ and $J_i^n = 1$, then since the function $x \mapsto c(x, \Delta Z_{\tau_i^n}, \theta^*)$ is continuous and $c(x, \Delta Z_{\tau_i^n}, \theta^*) \neq 0$ $(x \in \mathbb{R}^d)$, we have $\tilde{\xi}_i^n(t) \neq 0$ $(0 \leq t \leq 1)$. So similarly, (2.17) holds.

Then by (2.17), we can rewrite the right-hand side of (2.16) by changing residual terms as

$$\begin{aligned} & \frac{1}{h} E\Big[g_n(X_{t_{i-1}^n}, c_{i-1}(\Delta Z_{\tau_i^n}, \theta^*), \alpha) 1_{\{J_i^n=1\}} | \mathcal{F}_{i-1} \Big] + R(\alpha, h^{3/8+b} u_n \vee \sqrt{h} v_n, X_{t_{i-1}^n}) \\ = \; & \frac{1}{h} E\Big[\int_{t_{i-1}^n}^{t_i^n} \int g_n(X_{t_{i-1}^n}, c_{i-1}(z, \theta^*), \alpha) p(ds, dz) | \mathcal{F}_{i-1} \Big] + R(\alpha, h^{3/8+b} u_n \vee \sqrt{h} v_n, X_{t_{i-1}^n}) \\ = \; & \frac{1}{h} E\Big[\int_{t_{i-1}^n}^{t_i^n} \int g_n(X_{t_{i-1}^n}, c_{i-1}(z, \theta^*), \alpha) q^{\theta^*}(ds, dz) | \mathcal{F}_{i-1} \Big] + R(\alpha, h^{3/8+b} u_n \vee \sqrt{h} v_n, X_{t_{i-1}^n}) \\ = \; & E\Big[\int g_n(X_{t_{i-1}^n}, c_{i-1}(z, \theta^*), \alpha) f_{\theta^*}(z) dz | \mathcal{F}_{i-1} \Big] + R(\alpha, h^{3/8+b} u_n \vee \sqrt{h} v_n, X_{t_{i-1}^n}) \\ = \; & \int_{B(X_{t_{i-1}^n})} g_n(X_{t_{i-1}^n}, y, \alpha) \Psi_{\theta^*}(y, X_{t_{i-1}^n}) dy + R(\alpha, h^{3/8+b} u_n \vee \sqrt{h} v_n, X_{t_{i-1}^n}). \end{aligned}$$

□

Remark 2.4. Let

$$\mathbb{D}_n^{(k)}(x, y, \theta) \;\; = \;\; \partial_\theta^k \log \Psi_\theta(y, x) \varphi_n(x, y) \;\; (k = 0, 1, 2, 3, 4).$$

Then since $h^b/\epsilon_n \preceq 1$, by Proposition 2.4, $[H1], [H3], [H5] - [H7], [H9]$ and $[H10]$, we have

$$\begin{aligned} \frac{1}{h} E\left[\mathbb{D}_n^{(k)}(X_{t_{i-1}^n}, \Delta X_i^n, \theta) 1_{\{|\Delta X_i^n| > D h^\rho\}} | \mathcal{F}_{i-1} \right] &= \int_{B(X_{t_{i-1}^n})} \mathbb{D}_n^{(k)}(X_{t_{i-1}^n}, y, \theta) \Psi_{\theta^*}(y, X_{t_{i-1}^n}) dy \\ &\quad + R(\alpha, \frac{h^{3/8}}{\epsilon_n^{k \vee 1 - 1}} \vee \frac{\sqrt{h}}{\epsilon_n^{k \vee 1 + 1}}, X_{t_{i-1}^n}) \quad (k = 0, 1, 2, 3, 4). \\ \frac{1}{h} E\left[\left| \mathbb{D}_n^{(k)}(X_{t_{i-1}^n}, \Delta X_i^n, \theta) \right|^2 1_{\{|\Delta X_i^n| > D h^\rho\}} | \mathcal{F}_{i-1} \right] &= \int_{B(X_{t_{i-1}^n})} \left| \mathbb{D}_n^{(k)}(X_{t_{i-1}^n}, y, \theta) \right|^2 \Psi_{\theta^*}(y, X_{t_{i-1}^n}) dy \\ &\quad + R(\alpha, \frac{h^{3/8}}{\epsilon_n^{2(k \vee 1) - 1}} \vee \frac{\sqrt{h}}{\epsilon_n^{2(k \vee 1) + 1}}, X_{t_{i-1}^n}) \quad (k = 0, 1, 2, 3, 4). \\ E\left[\left| \mathbb{D}_n^{(1)}(X_{t_{i-1}^n}, \Delta X_i^n, \theta) \right|^l 1_{\{|\Delta X_i^n| > D h^\rho\}} | \mathcal{F}_{i-1} \right] &= R(\alpha, h, X_{t_{i-1}^n}) \quad (l = 3, 4). \end{aligned}$$

For the last equation, we use

$$\begin{aligned} \frac{1}{h} E\left[\left| \mathbb{D}_n^{(1)}(X_{t_{i-1}^n}, \Delta X_i^n, \theta) \right|^l 1_{\{|\Delta X_i^n| > D h^\rho\}} | \mathcal{F}_{i-1} \right] &= \int_{B(X_{t_{i-1}^n})} \left| \mathbb{D}_n^{(1)}(X_{t_{i-1}^n}, y, \theta) \right|^l \Psi_{\theta^*}(y, X_{t_{i-1}^n}) dy \\ &\quad + R(\alpha, \frac{h^{\frac{3}{8}+b}}{\epsilon_n^l} \vee \frac{\sqrt{h}}{\epsilon_n^{l+1}}, X_{t_{i-1}^n}) \\ &= R(\alpha, 1, X_{t_{i-1}^n}) + R(\alpha, \left(\frac{\sqrt{h}}{\epsilon_n^4} \right)^{\frac{3}{4}} \vee \frac{\sqrt{h}}{\epsilon_n^5}, X_{t_{i-1}^n}) \end{aligned}$$

and

$$\frac{\sqrt{h}}{\epsilon_n^5} \leq \frac{\sqrt{h}}{\epsilon_n^6} \preceq \frac{1}{\sqrt{n^3h^5}} \preceq 1. \tag{2.18}$$

We will verify the conditions of Theorem 2.3 with $(\theta, \tau) \mapsto (\sigma, \theta)$ where (θ, τ) are parameters in Section 2.3. First, we verify $[P1]$.

Lemma 2.4. *Assume* $[H1]-[H8]$*. Then for any* $p > d_1 + d_2$*,*

(1) $\sup_{n\in\mathbb{N}} E\left[\left(\frac{1}{n}\sup_{\alpha\in\Xi}|\partial_\sigma^3 H_n(\sigma,\theta)|\right)^p\right] < \infty,$

(2) $\sup_{n\in\mathbb{N}} E\left[\left(n^\epsilon \sup_{\theta\in\Theta}|\Gamma_n^1(\sigma^*,\theta) - \Gamma^1(\sigma^*)|\right)^p\right] < \infty.$

Proof. (1) From Lemmas 2.2 and 2.3, it follows that

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\Big[\Big(\frac{1}{n}\sup_{\alpha\in\Xi}|\partial_\sigma^3 H_n(\sigma,\theta)|\Big)^p\Big] \\
= \;& \sup_{n\in\mathbb{N}} E\Big[\Big(\frac{1}{n}\sup_{\alpha\in\Xi}\Big|\frac{1}{2h}\sum_{i=1}^n \bar{X}_{i,n}^T(\theta)\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\leq Dh^\rho\}} \\
&+\frac{1}{2}\sum_{i=1}^n \partial_\sigma^3 \log\det\beta_{i-1}(\sigma)1_{\{|\Delta X_i^n|\leq Dh^\rho\}}\Big|\Big)^p\Big] \\
\leq \;& C\sup_{n\in\mathbb{N}} E\Big[\Big(\frac{1}{n}\sup_{\sigma\in\Pi}\Big|\frac{1}{2h}\sum_{i=1}^n \bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}\Big|\Big)^p\Big] + C \\
\leq \;& C\sup_{n\in\mathbb{N}} E\Big[\Big(\frac{1}{n}\sup_{\sigma\in\Pi}\Big|\frac{1}{2h}\sum_{i=1}^n \big\{\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}} \\
&\qquad - E[\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|\mathcal{F}_{i-1}]\big\}\Big|\Big)^p\Big] + C.
\end{aligned}$$

Moreover, by the Burkholder-Davis-Gundy inequality,

$$\begin{aligned}
&\sup_{\sigma\in\Pi}\sup_{n\in\mathbb{N}} E\Big[\Big(\frac{1}{n}\Big|\frac{1}{2h}\sum_{i=1}^n \big\{\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}} \\
&\qquad - E[\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|\mathcal{F}_{i-1}]\big\}\Big|\Big)^p\Big] \\
\leq \;& C\sup_{n\in\mathbb{N}}\left(\frac{nh^{4\rho}}{n^2h^2}\right)^{\frac{p}{2}} < \infty,
\end{aligned} \tag{2.19}$$

because $\rho > 1/4$. Similarly,

$$\begin{aligned}
&\sup_{\sigma\in\Pi}\sup_{n\in\mathbb{N}} E\Big[\Big(\frac{1}{n}\Big|\frac{1}{2h}\sum_{i=1}^n \frac{\partial}{\partial\sigma}\big\{\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}} \\
&-E[\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|\mathcal{F}_{i-1}]\big\}\Big|\Big)^p\Big] < \infty.
\end{aligned} \tag{2.20}$$

Then by (2.19), (2.20) and Sobolev's inequality, we have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\Big[\sup_{\alpha\in\Xi}\Big(\frac{1}{n}\Big|\frac{1}{2h}\sum_{i=1}^n \big\{\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}} \\
&\qquad - E[\bar{X}_{i,n}^T\partial_\sigma^3\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|\mathcal{F}_{i-1}]\big\}\Big|\Big)^p\Big] < \infty.
\end{aligned}$$

This completes the proof of (1). □

(2) By the definition of Γ_n^1 and Γ^1, we have

$$\Gamma_n^1(\sigma^*,\theta)-\Gamma^1(\sigma^*)=\Lambda_1+\Lambda_2,$$

where

$$\begin{aligned}\Lambda_1 &= \frac{1}{2nh}\sum_{i=1}^n \bar{X}_{i,n}^T(\theta)\partial_\sigma^2\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad -\frac{1}{2}\int tr(\partial_\sigma^2\beta^{-1}(x,\sigma^*)\beta(x,\sigma^*))\pi(dx),\end{aligned}$$

and

$$\begin{aligned}\Lambda_2 &= \frac{1}{2n}\sum_{i=1}^n \partial_\sigma^2\log\det\beta_{i-1}(\sigma^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad -\frac{1}{2}\int \partial_\sigma^2\log\det\beta(x,\sigma^*)\pi(dx).\end{aligned}$$

We will estimate Λ_1 and Λ_2. In order to estimate Λ_1, we first notice that

$$\begin{aligned}&\sup_{n\in\mathbb{N}}E\left[\sup_{\theta\in\Theta}\left|n^{\frac{\rho}{2}}\frac{1}{n}\sum_{i=1}^n\frac{1}{2h}\left(\bar{X}_{i,n}^T(\theta)\partial_\sigma^2\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta)-\bar{X}_{i,n}^T\partial_\sigma^2\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}\right)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right|^p\right]\\ \le\;& C\sup_{n\in\mathbb{N}}\left(n^{\frac{\rho}{2}}h^\rho\right)^p<\infty. \qquad (2.21)\end{aligned}$$

Furthermore, the Burkholder-Davis-Gundy inequality implies

$$\begin{aligned}&\sup_{n\in\mathbb{N}}E\Big[\Big|n^\epsilon\frac{1}{n}\sum_{i=1}^n\big(\frac{1}{2h}\bar{X}_{i,n}^T\partial_\sigma^2\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad -E[\frac{1}{2h}\bar{X}_{i,n}^T\partial_\sigma^2\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}]\big)\Big|^p\Big]\\ \le\;& C\sup_{n\in\mathbb{N}}\left(\frac{n^{2\epsilon}h^{4\rho}}{nh^2}\right)^{\frac{p}{2}}=C\sup_{n\in\mathbb{N}}\left(\frac{n^{2\epsilon}}{nh}\cdot h^{4\rho-1}\right)^{\frac{p}{2}}<\infty. \qquad (2.22)\end{aligned}$$

Moreover, by Lemmas 2.2, 2.3 and Proposition 2.2, we have

$$\begin{aligned}&\sup_{n\in\mathbb{N}}E\Big[\Big(n^\epsilon\Big|\frac{1}{n}\sum_{i=1}^n E[\frac{1}{2h}\bar{X}_{i,n}^T\partial_\sigma^2\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}]\\ &\quad -\frac{1}{2}\int tr(\partial_\sigma^2\beta^{-1}(x,\sigma^*)\beta(x,\sigma^*))\pi(dx)\Big|\Big)^p\Big]<\infty. \qquad (2.23)\end{aligned}$$

From (2.21), (2.22) and (2.23), it follows that

$$\sup_{n\in\mathbb{N}}E\left[\left(n^\epsilon\sup_{\theta\in\Theta}|\Lambda_1|\right)^p\right]<\infty. \qquad (2.24)$$

Next we will estimate Λ_2. By using Proposition 2.2, we have

$$\sup_{n\in\mathbb{N}}E\left[\left|n^\epsilon\frac{1}{2n}\sum_{i=1}^n\left(\partial_\sigma^2\log\det\beta_{i-1}(\sigma^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}-E[\partial_\sigma^2\log\det\beta_{i-1}(\sigma^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}]\right)\right|^p\right]<\infty. \qquad (2.25)$$

Moreover, by Lemma 2.2, we have

$$\sup_{n\in\mathbb{N}} E\left[\left|\sqrt{n}\left(\frac{1}{2n}\sum_{i=1}^{n} E[\partial_\sigma^2 \log\det\beta_{i-1}(\sigma^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}] - \frac{1}{2}\int \partial_\sigma^2 \log\det\beta(x,\sigma^*)\pi(dx)\right)\right|^p\right] < \infty. \quad (2.26)$$

From (2.25) and (2.26), it follows that

$$\sup_{n\in\mathbb{N}} E\left[\left|n^\epsilon \Lambda_2\right|^p\right] < \infty. \quad (2.27)$$

Inequalities (2.24) and (2.27) complete the proof. □

By this lemma, we can verify [P1] of Theorem 2.3 for any $L > 1$ if we take $\beta_1 > 0$ small enough. Conditions [P4] and [P5] are easily verified by [$H12$]. Next, we will verify [P3].

Lemma 2.5. *Assume* [$H1$] $-$ [$H8$]. *Then for any* $p > d_1 + d_2$,

(1) $\sup_{n\in\mathbb{N}} E\left[\sup_{\theta\in\Theta}|\frac{1}{\sqrt{n}}\partial_\sigma H_n(\sigma^*,\theta)|^p\right] < \infty$,

(2) $\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\alpha\in\Xi} n^\epsilon |Y_n(\sigma,\theta:\sigma^*) - Y^1(\sigma:\sigma^*)|\right)^p\right] < \infty$.

Proof. (1) By the definition of H_n, we have

$$\begin{aligned} -\frac{2}{\sqrt{n}}\partial_\sigma H_n(\sigma^*,\theta) \quad &= \quad \frac{1}{\sqrt{n}h}\sum_{i=1}^{n} \bar{X}_{i,n}^T(\theta)\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}} \\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial_\sigma \log\det\beta_{i-1}(\sigma^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}. \end{aligned} \quad (2.28)$$

We will estimate the right-hand side of (2.28). First, we notice that

$$\begin{aligned} &\sup_{n\in\mathbb{N}} E\left[\sup_{\theta\in\Theta}\left|\frac{1}{\sqrt{n}h}\sum_{i=1}^{n}\left(\bar{X}_{i,n}^T(\theta)\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta) - \bar{X}_{i,n}^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}\right)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right|^p\right] \\ = \quad &\sup_{n\in\mathbb{N}} E\left[\sup_{\theta\in\Theta}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)(\bar{X}_{i,n} + \bar{X}_{i,n}(\theta))1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right|^p\right] \\ \le \quad &C\sup_{n\in\mathbb{N}} E\left[\sup_{\theta\in\Theta}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{2(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right.\right.\right. \\ &\qquad\left.\left.\left. - E[2(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}]\right\}\right|^p\right] \\ &+C\sup_{n\in\mathbb{N}} E\left[\sup_{\theta\in\Theta}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\left[2(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right]\right|^p\right] + C. \end{aligned} \quad (2.29)$$

The second term of the right-hand side is finite because of Lemmas 2.2 and 2.3.

Moreover, the summation in the first term becomes a martingale. Then, by the Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned} &\sup_{n\in\mathbb{N}}\sup_{\theta\in\Theta} E\left[\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right.\right.\right. \\ &\qquad\left.\left.\left. - E[(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}]\right\}\right|^p\right] < \infty. \end{aligned} \quad (2.30)$$

Similarly,

$$\begin{aligned} &\sup_{n\in\mathbb{N}}\sup_{\theta\in\Theta} E\left[\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\left\{(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right.\right.\right. \\ &\qquad\left.\left.\left. - E[(a_{i-1}(\theta^*) - a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}]\right\}\right|^p\right] < \infty. \end{aligned} \quad (2.31)$$

Then by (2.30), (2.31), and Sobolev's inequality, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big\{(a_{i-1}(\theta^*)-a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}} - E[(a_{i-1}(\theta^*)-a_{i-1}(\theta))^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}]\Big\}\Big|^p\Big] < \infty. \quad (2.32)$$

Therefore, (2.29) and (2.32) yield

$$\sup_{n\in\mathbb{N}} E\left[\sup_{\theta\in\Theta}\left|\frac{1}{\sqrt{n}h}\sum_{i=1}^{n}\Big\{\bar{X}_{i,n}^T(\theta)\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta)-\bar{X}_{i,n}^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}\Big\}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right|^p\right] < \infty. \quad (2.33)$$

Let $K_{i,n}=\big\{\frac{1}{h}\bar{X}_{i,n}^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}+\partial_\sigma\log\det\beta_{i-1}(\sigma^*)\big\}1_{\{|\Delta X_i^n|\le Dh^\rho\}}$. Then, to complete the proof, it is sufficient to estimate the summation of $K_{i,n}$. For this purpose, we will use Proposition 2.3. First, by Lemmas 2.2 and 2.3, we have

$$\begin{aligned} E\Big[\frac{1}{h}\bar{X}_{i,n}^T\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\Big] &= tr(\partial_\sigma\beta_{i-1}^{-1}(\sigma^*)\beta_{i-1}(\sigma^*))+R(\alpha,h,X_{t_{i-1}^n}) \\ &= -\partial_\sigma\log\det\beta_{i-1}(\sigma^*)+R(\alpha,h,X_{t_{i-1}^n}). \end{aligned}$$

Therefore, it follows that

$$\sup_{n\in\mathbb{N}} E\big[\big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\left[K_{i,n}|\mathcal{F}_{i-1}\right]\big|^p\big] < \infty. \quad (2.34)$$

Moreover, by Lemmas 2.2 and 2.3, we have

$$\sup_{n\in\mathbb{N}} E\big[\big|\frac{1}{n}\sum_{i=1}^{n}E\left[\psi_i^2(K_{i,n})|\mathcal{F}_{i-1}\right]\big|^{\frac{p}{2}}\big] < \infty. \quad (2.35)$$

So by Proposition 2.3 for $k=2$, we have

$$\begin{aligned} \sup_{n\in\mathbb{N}} E\big[\big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}K_{i,n}\big|^p\big] &\le C\sup_{n\in\mathbb{N}} E\big[\big|\frac{1}{n^2}\sum_{i=1}^{n}\psi_i^3(K_{i,n})\big|^{\frac{p}{4}}\big]+C \\ &\le C\sup_{n\in\mathbb{N}}\left(\frac{h^{8\rho}}{nh^4}\right)^{\frac{p}{4}}+C\le C\sup_{n\in\mathbb{N}}\left(\frac{1}{nh}\cdot h^{8\rho-3}\right)^{\frac{p}{4}}+C<\infty. \end{aligned} \quad (2.36)$$

By (2.28), (2.33) and (2.36), we have

$$\sup_{n\in\mathbb{N}} E\left[\sup_{\theta\in\Theta}\big|\frac{1}{\sqrt{n}}\partial_\sigma H_n(\sigma^*,\theta)\big|^p\right] < \infty.$$

□

(2) By the definition of Y_n and Y_1, we have

$$Y_n(\sigma,\theta;\sigma^*)-Y^1(\sigma;\sigma^*)=\Lambda_3+\Lambda_4, \quad (2.37)$$

where

$$\begin{aligned} \Lambda_3 &= -\frac{1}{2nh}\sum_{i=1}^{n}\bar{X}_{i,n}^T(\theta)\{\beta_{i-1}^{-1}(\sigma)-\beta_{i-1}^{-1}(\sigma^*)\}\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}} \\ &\quad -\frac{1}{2}\int tr\left(I_d-\beta^{-1}(x,\sigma)\beta(x,\sigma^*)\right)\pi(dx), \end{aligned}$$

and

$$\Lambda_4=-\frac{1}{2n}\sum_{i=1}^{n}\log\frac{\det\beta_{i-1}(\sigma)}{\det\beta_{i-1}(\sigma^*)}1_{\{|\Delta X_i^n|\le Dh^\rho\}}+\frac{1}{2}\int\log\frac{\det\beta(x,\sigma)}{\det\beta(x,\sigma^*)}\pi(dx).$$

We first estimate Λ_3. By the Burkholder-Davis-Gundy inequality and Sobolev's inequality, we have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\bigg[\sup_{\alpha\in\Xi}\bigg|n^\epsilon \frac{1}{2nh}\sum_{i=1}^n \Big\{\bar{X}_{i,n}^T(\theta)\{\beta_{i-1}^{-1}(\sigma)-\beta_{i-1}^{-1}(\sigma^*)\}\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}} \\
&\qquad - E\left[\bar{X}_{i,n}^T(\theta)\{\beta_{i-1}^{-1}(\sigma)-\beta_{i-1}^{-1}(\sigma^*)\}\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right]\Big\}\bigg|^p\bigg] \\
\le\ & C\sup_{n\in\mathbb{N}}\left(\frac{n^{2\epsilon}h^{4\rho}}{nh^2}\right)^{\frac{p}{2}} < \infty. \qquad (2.38)
\end{aligned}$$

Moreover,

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\bigg[\sup_{\alpha\in\Xi}\bigg|n^\epsilon \frac{1}{2nh}\sum_{i=1}^n \Big\{E\left[\bar{X}_{i,n}^T(\theta)\{\beta_{i-1}^{-1}(\sigma)-\beta_{i-1}^{-1}(\sigma^*)\}\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right] \\
&\qquad - E\left[\bar{X}_{i,n}^T\{\beta_{i-1}^{-1}(\sigma)-\beta_{i-1}^{-1}(\sigma^*)\}\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right]\Big\}\bigg|^p\bigg] \\
\le\ & \sup_{n\in\mathbb{N}} C\,(n^\epsilon h)^p < \infty. \qquad (2.39)
\end{aligned}$$

Furthermore, by Lemmas 2.2 and 2.3, Proposition 2.2 and Sobolev's inequality, we have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\bigg[\sup_{\alpha\in\Xi}\bigg|n^\epsilon\Big\{-\frac{1}{2nh}\sum_{i=1}^n E\left[\bar{X}_{i,n}^T\{\beta_{i-1}^{-1}(\sigma)-\beta_{i-1}^{-1}(\sigma^*)\}\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right] \\
&\qquad -\frac{1}{2}\int tr\left(I_d-\beta^{-1}(x,\sigma)\beta(x,\sigma^*)\right)\pi(dx)\Big\}\bigg|^p\bigg] < \infty. \qquad (2.40)
\end{aligned}$$

From (2.38), (2.39) and (2.40), it follows that

$$\sup_{n\in\mathbb{N}} E\left[\sup_{\alpha\in\Xi}|n^\epsilon\Lambda_3|^p\right] < \infty. \qquad (2.41)$$

Next, we will estimate Λ_4. By Proposition 2.2 and Sobolev's inequality, we have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\bigg[\sup_{\sigma\in\Pi}\bigg|n^\epsilon\frac{1}{2n}\sum_{i=1}^n\Big\{\log\frac{\det\beta_{i-1}(\sigma)}{\det\beta_{i-1}(\sigma^*)}1_{\{|\Delta X_i^n|\le Dh^\rho\}} \\
&\qquad - E\Big[\log\frac{\det\beta_{i-1}(\sigma)}{\det\beta_{i-1}(\sigma^*)}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big]\Big\}\bigg|^p\bigg] < \infty. \qquad (2.42)
\end{aligned}$$

Moreover, by Lemma 2.2, we have

$$\frac{1}{2n}\sum_{i=1}^n E\left[\log\frac{\det\beta_{i-1}(\sigma)}{\det\beta_{i-1}(\sigma^*)}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right] = \frac{1}{2}\int\log\frac{\det\beta(x,\sigma)}{\det\beta(x,\sigma^*)}\pi(dx)+O(h). \qquad (2.43)$$

From (2.42) and (2.43), it follows that

$$\sup_{n\in\mathbb{N}} E\left[\sup_{\alpha\in\Xi}|n^\epsilon\Lambda_4|^p\right] < \infty. \qquad (2.44)$$

Inequalities (2.41) and (2.44) complete the proof. □

By Lemmas 2.4 and 2.5, we can verify conditions of Theorem 2.3 for any $L > 0$, by setting parameters so that

$$\begin{aligned}
\frac{1}{2} &> \beta_2 \geq \frac{1}{2} - \epsilon, \\
0 &< \rho_2 < 1 - 2\beta_2, \\
0 &< \alpha < \frac{\rho_2}{2}, \\
\beta &= \frac{\alpha}{1-\alpha}, \\
0 &< \beta_1 \leq \epsilon, \\
0 &< \rho_1 < 1 \wedge \beta \wedge \frac{2\beta_1}{1-\alpha}.
\end{aligned}$$

So by Theorem 2.3, we have

$$P_{\alpha^*}\left[\sup_{(u_1,\theta)\in V_n^1(r)\times\Theta} Z_n^1(u_1,\theta;\sigma^*) \geq e^{-\frac{r}{2}}\right] \leq \frac{C_L}{r^L}, \tag{2.45}$$

for all $n \in \mathbb{N}$ and $r > 0$, where $C_L > 0$ is a constant,

$$V_n^1(r) := \{u_1 \in \mathbb{R}^{d_1}; \sigma^* + \frac{u_1}{\sqrt{n}} \in \Pi, |u_1| \geq r\},$$

and

$$Z_n^1(u_1,\theta;\sigma^*) := \exp\left(H_n(\sigma^* + \frac{u_1}{\sqrt{n}},\theta) - H_n(\sigma^*,\theta)\right).$$

Then for any $L > 0$,

$$\begin{aligned}
P_{\alpha^*}\left[|\sqrt{n}(\hat{\sigma}_n - \sigma^*)| \geq r\right] &\leq P_{\alpha^*}\left[\sup_{(u_1,\theta)\in V_n^1(r)\times\Theta} Z_n^1(u_1,\theta;\sigma^*) \geq 1\right] \\
&\leq \frac{C_L}{r^L}.
\end{aligned}$$

Therefore for any $p > 0$, let $L > p$, then it follows that

$$\begin{aligned}
\sup_{n\in\mathbb{N}} E\left[|\sqrt{n}(\hat{\sigma}_n - \sigma^*)|^p\right] &\leq \sup_{n\in\mathbb{N}} p\int_0^\infty r^{p-1} P_{\alpha^*}\left[|\sqrt{n}(\hat{\sigma}_n - \sigma^*)| \geq r\right] dr \\
&\leq p\int_0^\infty r^{p-1}(1 \wedge \frac{C_L}{r^L}) dr < \infty. \qquad (2.46)
\end{aligned}$$

Next, we will verify the conditions of Theorem 2.3 with θ in place of "θ" and $\mathcal{T}$ being a point.

Lemma 2.6. *Assume* $[H1] - [H12]$. *Let* $p_1 > d_2 \vee 2$, $p_2 \geq 2$ *and* $0 < \beta_1 \leq a \wedge \epsilon$, *where* a *is the constant appearing in [H11]. Then*

(1) $\sup_{n\in\mathbb{N}} E\left[\left(\frac{1}{nh}\sup_{\theta\in\Theta}|\partial_\theta^3 H_n(\hat{\sigma}_n,\theta)|\right)^{p_1}\right] < \infty$,

(2) $\sup_{n\in\mathbb{N}} E\left[\left((nh)^{\beta_1}|\Gamma_n^2(\theta^*) - \Gamma^2(\alpha^*)|\right)^{p_2}\right] < \infty$.

Proof. (1) By differentiating H_n with respect to θ three times, we have

$$\begin{aligned}
& \frac{1}{nh}\partial_\theta^3 H_n(\hat{\sigma}_n,\theta) \\
= \quad & \frac{1}{nh}\sum_{i=1}^n \partial_\theta^3 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\hat{\sigma}_n)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}} \\
& -\frac{3}{n}\sum_{i=1}^n \partial_\theta^2 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\hat{\sigma}_n)\partial_\theta a_{i-1}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}} \\
+ \quad & \frac{1}{nh}\sum_{i=1}^n \mathbb{D}_n^{(3)}(X_{t_{i-1}},\Delta X_i^n,\theta)1_{\{|\Delta X_i^n|>Dh^\rho\}} - \frac{1}{n}\sum_{i=1}^n\int_{B(X_{t_{i-1}^n})}\partial_\theta^3\Psi_\theta(y,X_{t_{i-1}})\varphi_n(X_{t_{i-1}},y)dy. \quad (2.47)
\end{aligned}$$

We will estimate each term of the right-hand side of (2.47). First, it is easily shown that

$$\sup_{n\in\mathbb{N}} E\left[\left(\frac{3}{n}\sup_{\theta\in\Theta}\Big|\sum_{i=1}^n \partial_\theta^2 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\hat{\sigma}_n)\partial_\theta a_{i-1}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big|\right)^{p_1}\right] < \infty,$$

and

$$\sup_{n\in\mathbb{N}} E\left[\left(\frac{1}{n}\sup_{\theta\in\Theta}\Big|\sum_{i=1}^n\int_{B(X_{t_{i-1}^n})}\partial_\theta^3\Psi_\theta(y,X_{t_{i-1}})\varphi_n(X_{t_{i-1}},y)dy\Big|\right)^{p_1}\right] < \infty,$$

so the second term and the forth term of the right-hand side of (2.47) are estimated.

To estimate the first term of (2.47), let $\delta > 0$ and $\{\sigma\in\mathbb{R}^{d_1}; |\sigma-\sigma^*|\le\delta\}\subset\Pi$. Then by (2.46), mean-value theorem and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
& \sup_{n\in\mathbb{N}} E\Big[\Big(\sup_{\theta\in\Theta}\Big|\frac{1}{nh}\sum_{i=1}^n \partial_\theta^3 a_{i-1}^T(\theta)(\beta_{i-1}^{-1}(\hat{\sigma}_n)-\beta_{i-1}^{-1}(\sigma^*))\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}1_{\{|\hat{\sigma}_n-\sigma^*|\le\delta\}}\Big|\Big)^{p_1}\Big] \\
\le \quad & C\sup_{n\in\mathbb{N}}\left(\frac{h^\rho}{\sqrt{nh}}\right)^{p_1} < \infty. \quad (2.48)
\end{aligned}$$

By (2.46), it is easy to show that

$$\sup_{n\in\mathbb{N}} E\Big[\Big(\sup_{\theta\in\Theta}\Big|\frac{1}{nh}\sum_{i=1}^n \partial_\theta^3 a_{i-1}^T(\theta)(\beta_{i-1}^{-1}(\hat{\sigma}_n)-\beta_{i-1}^{-1}(\sigma^*))\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}1_{\{|\hat{\sigma}_n-\sigma^*|>\delta\}}\Big|\Big)^{p_1}\Big] < \infty. \quad (2.49)$$

Moreover, the Burkholder-Davis-Gundy inequality and Sobolev's inequality yield

$$\begin{aligned}
& \sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|\frac{1}{nh}\sum_{i=1}^n\Big\{\partial_\theta^3 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}} \\
& \qquad - E\left[\partial_\theta^3 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right]\Big\}\Big|^{p_1}\Big] \\
\le \quad & C\sup_{n\in\mathbb{N}}\left(\frac{h^{2\rho}}{nh^2}\right)^{\frac{p_1}{2}} \le C\sup_{n\in\mathbb{N}}\frac{1}{n^{p_1\epsilon}h^{p_1\epsilon}} < \infty. \quad (2.50)
\end{aligned}$$

Furthermore, by Lemmas 2.2 and 2.3, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|\frac{1}{nh}\sum_{i=1}^n E\left[\partial_\theta^3 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right]\Big|^{p_1}\Big] < \infty. \quad (2.51)$$

Inequalities (2.48), (2.49), (2.50) and (2.51) give

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|\frac{1}{nh}\sum_{i=1}^n \partial_\theta^3 a_{i-1}^T(\theta)\beta_{i-1}^{-1}(\hat{\sigma}_n)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big|^{p_1}\Big] < \infty. \quad (2.52)$$

This completes the estimate of the first term of (2.47).

Finally, we will estimate the third term of (2.47). By Remark 2.4 and (2.18), we have

$$E[\mathbb{D}_n^{(3)}(X_{t_{i-1}}, \Delta X_i^n, \theta)1_{\{|\Delta X_i^n|>Dh^\rho\}}|\mathcal{F}_{i-1}] = R(\alpha, h, X_{t_{i-1}^n}), \tag{2.53}$$

and

$$E[|\mathbb{D}_n^{(3)}(X_{t_{i-1}}, \Delta X_i^n, \theta)|^2 1_{\{|\Delta X_i^n|>Dh^\rho\}}|\mathcal{F}_{i-1}] = R(\alpha, \frac{h^{11/8}}{\epsilon_n^5} \vee \frac{h\sqrt{h}}{\epsilon_n^7} \vee h, X_{t_{i-1}^n}). \tag{2.54}$$

So by using Proposition 2.3 for $k = 2$, we have

$$\begin{aligned}
&\sup_{\theta\in\Theta}\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{nh}\sum_{i=1}^n \mathbb{D}_n^{(3)}(X_{t_{i-1}}, \Delta X_i^n, \theta)1_{\{|\Delta X_i^n|>Dh^\rho\}}\Big|^{p_1}\Big] \\
\leq\ & \sup_{\theta\in\Theta}\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{n^4h^4}\sum_{i=1}^n \psi_i^3\big(\mathbb{D}_n^{(3)}(X_{t_{i-1}}, \Delta X_i^n, \theta)1_{\{|\Delta X_i^n|>Dh^\rho\}}\big)\Big|^{\frac{p_1}{4}}\Big] \\
&+C\sup_{n\in\mathbb{N}}\left(\frac{nh}{nh}\right)^{p_1} + C\sup_{n\in\mathbb{N}}\left(\frac{n}{n^2h^2}\times\left(\frac{h^{11/8}}{\epsilon_n^5}\vee\frac{h\sqrt{h}}{\epsilon_n^7}\vee h\right)\right)^{\frac{p_1}{2}} \\
\leq\ & C\sup_{n\in\mathbb{N}}\left(\frac{1}{n^3h^4\epsilon_n^{12}}\right)^{\frac{p_1}{4}} + C + C\sup_{n\in\mathbb{N}}\left(\frac{1}{nh\epsilon_n}\right)^{\frac{p_1}{2}} < \infty.
\end{aligned} \tag{2.55}$$

Similarly,

$$\begin{aligned}
&\sup_{\theta\in\Theta}\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{nh}\sum_{i=1}^n \partial_\theta\mathbb{D}_n^{(3)}(X_{t_{i-1}}, \Delta X_i^n, \theta)1_{\{|\Delta X_i^n|>Dh^\rho\}}\Big|^{p_1}\Big] \\
\leq\ & C\sup_{n\in\mathbb{N}}\left(\frac{1}{n^3h^4\epsilon_n^{16}}\right)^{\frac{p_1}{4}} + C\sup_{n\in\mathbb{N}}\left(\frac{1}{nh\epsilon_n^3}\right)^{\frac{p_1}{2}} + C < \infty.
\end{aligned} \tag{2.56}$$

By (2.55), (2.56) and Sobolev's inequality, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|\frac{1}{nh}\sum_{i=1}^n \mathbb{D}_n^{(3)}(X_{t_{i-1}}, \Delta X_i^n, \theta)1_{\{|\Delta X_i^n|>Dh^\rho\}}\Big|^{p_1}\Big] < \infty. \tag{2.57}$$

This completes the estimate of the third term. □

(2) By the definition, we obtain

$$\Gamma_n^2(\theta^*) - \Gamma^2(\alpha^*) = \Lambda_5 + \Lambda_6 + \Lambda_7 + \Lambda_8,$$

where

$$\Lambda_5 = -\frac{1}{nh}\sum_{i=1}^n \partial_\theta^2 a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\hat{\sigma}_n)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}},$$

$$\begin{aligned}
\Lambda_6 \ =\ & \frac{1}{n}\sum_{i=1}^n \partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\hat{\sigma}_n)\partial_\theta a_{i-1}(\theta^*)1_{\{|\Delta X_i^n|\leq Dh^\rho\}} \\
&-\int \partial_\theta a^T(x,\theta^*)\beta^{-1}(x,\sigma^*)\partial_\theta a(x,\theta^*)\pi(dx),
\end{aligned}$$

$$\Lambda_7 = -\frac{1}{nh}\sum_{i=1}^n \mathbb{D}_n^{(2)}(X_{t_{i-1}^n}, \Delta X_i^n, \theta^*)1_{\{|\Delta X_i^n|>Dh^\rho\}} + \int\int_{A(x)} \partial_\theta^2 \log\Psi_{\theta^*}(y,x)\Psi_{\theta^*}(y,x)dy\pi(dx),$$

and

$$\Lambda_8 = \frac{1}{n}\sum_{i=1}^n \int_{B(X_{t_{i-1}^n})} \partial_\theta^2\Psi_{\theta^*}(y, X_{t_{i-1}^n})\varphi_n(X_{t_{i-1}^n}, y)dy - \int\int_{B(x)} \partial_\theta^2\Psi_{\theta^*}(x,y)dy\pi(dx).$$

To estimate Λ_5, first by (2.46), we have

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^{\beta_1}\frac{1}{nh}\big|\sum_{i=1}^{n}\Big(\partial_\theta^2 a_{i-1}^T(\theta^*)(\beta_{i-1}^{-1}(\hat\sigma_n)-\beta_{i-1}^{-1}(\sigma^*))\bar X_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\big|\Big)^{p_2}\Big]\\ \le\;& C\sup_{n\in\mathbb{N}}\left(\frac{h^\rho}{\sqrt{nh}}(nh)^{\beta_1}\right)^{p_2}\le C\sup_{n\in\mathbb{N}}\left(\frac{1}{\sqrt{nh}}\frac{h^{\rho-\frac13}}{h^\epsilon}(nh)^{\beta_1}\right)^{p_2}\le C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^{\beta_1}}{n^\epsilon h^\epsilon}\right)^{p_2}<\infty. \end{aligned}\tag{2.58}$$

Moreover, since by Lemmas 2.2 and 2.3, we have

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^{\beta_1}\frac{1}{nh}\big|\sum_{i=1}^{n}E[\partial_\theta^2 a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma^*)\bar X_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}]\Big|\Big)^{p_2}\Big]\\ \le\;& C\sup_{n\in\mathbb{N}}\Big((nh)^{\beta_1}\sqrt{h}\Big)^{p_2}<\infty,\end{aligned}$$

and

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^{2\beta_1}\frac{1}{n^2h^2}\big|\sum_{i=1}^{n}E[\psi_i^2(\partial_\theta^2 a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma^*)\bar X_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}})|\mathcal{F}_{i-1}]\Big|\Big)^{\frac{p_2}{2}}\Big]\\ \le\;& C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^{2\beta_1}}{nh}\right)^{\frac{p_2}{2}}<\infty,\end{aligned}$$

therefore, it follows from Proposition 2.3 for $k=2$ that

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^{\beta_1}\frac{1}{nh}\big|\sum_{i=1}^{n}\partial_\theta^2 a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma^*)\bar X_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big|\Big)^{p_2}\Big]\\ \le\;& C\sup_{n\in\mathbb{N}} E\Big[\Big(\frac{(nh)^{4\beta_1}}{(nh)^4}\big|\sum_{i=1}^{n}\psi_i^3\left(\partial_\theta^2 a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma^*)\bar X_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right)\Big|\Big)^{\frac{p_2}{4}}\Big]+C\\ \le\;& C\sup_{n\in\mathbb{N}}\Big(\frac{(nh)^{4\beta_1}}{n^3h^4}\Big)^{\frac{p_2}{4}}+C\le C\sup_{n\in\mathbb{N}}\left(\frac{1}{n^2h^3}\right)^{\frac{p_2}{4}}+C<\infty.\end{aligned}\tag{2.59}$$

Inequalities (2.58) and (2.59) give

$$\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^{\beta_1}\big|\Lambda_5\big|\Big)^{p_2}\Big]<\infty.$$

Next, we will estimate Λ_6. First by (2.46), we have

$$\sup_{n\in\mathbb{N}} E\Big[\Big(\sqrt{nh}\Big|\frac{1}{n}\sum_{i=1}^{n}\partial_\theta a_{i-1}^T(\theta^*)\big\{\beta_{i-1}^{-1}(\hat\sigma_n)-\beta_{i-1}^{-1}(\sigma^*)\big\}\partial_\theta a_{i-1}(\theta^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big|\Big)^{p_2}\Big]<\infty.\tag{2.60}$$

Moreover, Proposition 2.2 yields

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^\epsilon\big|\frac{1}{n}\sum_{i=1}^{n}\big(\partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma^*)\partial_\theta a_{i-1}(\theta^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &-E\left[\partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma^*)\partial_\theta a_{i-1}(\theta^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right]\big)\big|\Big)^{p_2}\Big]<\infty.\end{aligned}\tag{2.61}$$

Furthermore, by Lemma 2.2, we have

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^\epsilon\big|\frac{1}{n}\sum_{i=1}^{n}E\left[\partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma^*)\partial_\theta a_{i-1}(\theta^*)1_{\{|\Delta X_i^n|\le Dh^\rho\}}\right]\\ &-\int\partial_\theta a^T(x,\theta^*)\beta^{-1}(x,\sigma^*)\partial_\theta a(x,\theta^*)\pi(dx)\big|\Big)^{p_2}\Big]<\infty.\end{aligned}\tag{2.62}$$

Inequalities (2.60), (2.61) and (2.62) give

$$\sup_{n\in\mathbb{N}} E\Big[\Big((nh)^{\epsilon}|\Lambda_6|\Big)^{p_2}\Big] < \infty.$$

On the other hand, Proposition 2.2 and [$H11$] lead

$$\sup_{n\in\mathbb{N}} E\left[\left((nh)^{\beta_1}|\Lambda_8|\right)^{p_2}\right] < \infty.$$

Finally, we will estimate Λ_7. By using Proposition 2.3 for $k=2$ and Remark 2.4, we have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\Bigg[\Bigg((nh)^{\beta_1}\Big|\frac{1}{nh}\sum_{i=1}^{n}\Big\{\mathbb{D}_n^{(2)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)1_{\{|\Delta X_i^n|>Dh^\rho\}}\\
&\qquad\qquad - E\big[\mathbb{D}_n^{(2)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)1_{\{|\Delta X_i^n|>Dh^\rho\}}|\mathcal{F}_{i-1}\big]\Big\}\Big|\Bigg)^{p_2}\Bigg]\\
\leq\ & C\sup_{n\in\mathbb{N}}\Big(\frac{(nh)^{4\beta_1}}{n^3h^4\epsilon_n^8}\Big)^{\frac{p_2}{4}} + C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^{2\beta_1}}{nh}\right)^{\frac{p_2}{2}} \leq C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^{4\beta_1}}{n^{\frac{3}{2}}h^2}\right)^{\frac{p_2}{4}} + C\\
\leq\ & C\sup_{n\in\mathbb{N}}\left(n^{\frac{12}{7}\beta_1-\frac{3}{10}}\right)^{\frac{p_2}{4}} + C < \infty. \qquad (2.63)
\end{aligned}$$

Moreover, by [H11], Proposition 2.2 and Remark 2.4, we have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\Bigg[\Bigg((nh)^{\beta_1}\Big|\frac{1}{nh}\sum_{i=1}^{n}E\big[\mathbb{D}_n^{(2)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)1_{\{|\Delta X_i^n|>Dh^\rho\}}|\mathcal{F}_{i-1}\big]\\
&\qquad\qquad -\int\int_{A(x)}\partial_\theta^2\log\Psi_{\theta^*}(y,x)\Psi_{\theta^*}(y,x)dy\pi(dx)\Big|\Bigg)^{p_2}\Bigg] < \infty. \qquad (2.64)
\end{aligned}$$

Inequalities (2.63) and (2.64) give

$$\sup_{n\in\mathbb{N}} E\big[\big((nh)^{\beta_1}|\Lambda_7|\big)^{p_2}\big] < \infty.$$

□

By Lemma 2.6, we can verify [P1] of Theorem 2.3 for the random field $\theta\mapsto H_n(\hat\sigma_n,\theta)$, if β_1 is small enough. By [$H12$], we can verify [P4] and [P5]. We will verify [P3] by the following lemma.

Lemma 2.7. *Assume [$H1$] − [$H12$]. Then for $p_3 > d_1\vee 4$, $p_4 > d_2\vee 2$, and $\frac{1}{2}>\beta_2\geq(\frac{1}{2}-\epsilon)\vee(\frac{1}{2}-a)$,*

(1) $\sup_{n\in\mathbb{N}} E\left[\big|\frac{1}{\sqrt{nh}}\partial_\theta H_n(\hat\sigma_n,\theta^*)\big|^{p_3}\right] < \infty$,

(2) $\sup_{n\in\mathbb{N}} E\left[\Big(\sup_{\theta\in\Theta}(nh)^{\frac{1}{2}-\beta_2}|Y_n(\theta;\alpha^*)-Y^2(\theta;\alpha^*)|\Big)^{p_4}\right] < \infty$.

Proof. (1) By the definition of H_n,

$$\begin{aligned}
\frac{1}{\sqrt{nh}}\partial_\theta H_n(\hat\sigma_n,\theta^*) &= \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\hat\sigma_n)\bar X_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}\\
&\quad+\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\mathbb{D}_n^{(1)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)1_{\{|\Delta X_i^n|>Dh^\rho\}}\\
&\quad-\frac{\sqrt{h}}{\sqrt{n}}\sum_{i=1}^{n}\int_{B(X_{t_{i-1}^n})}\partial_\theta\Psi_{\theta^*}(y,X_{t_{i-1}^n})\varphi_n(X_{t_{i-1}^n},y)dy. \qquad (2.65)
\end{aligned}$$

First, we will estimate the first term of the right-hand side of (2.65). Since

$$E[\partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma)\bar X_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|\mathcal{F}_{i-1}] = R(\alpha,h\sqrt{h},X_{t_{i-1}^n}),$$

and

$$E[|\partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|^2|\mathcal{F}_{i-1}] = R(\alpha, h, X_{t_{i-1}^n}),$$

by using Proposition 2.3 for $k=2$, we have

$$\sup_{\sigma\in\Pi}\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{\sqrt{nh}}\sum_{i=1}^n \partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}\Big|^{p_3}\Big] \leq C\sup_{n\in\mathbb{N}}\left(\frac{h^{4\rho}}{nh^2}\right)^{\frac{p_3}{4}} + C < \infty. \tag{2.66}$$

Similarly, we obtain

$$\sup_{\sigma\in\Pi}\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{\sqrt{nh}}\sum_{i=1}^n \partial_\theta a_{i-1}^T(\theta^*)\partial_\sigma\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}\Big|^{p_3}\Big] < \infty. \tag{2.67}$$

Therefore by Sobolev's inequality, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\sigma\in\Pi}\Big|\frac{1}{\sqrt{nh}}\sum_{i=1}^n \partial_\theta a_{i-1}^T(\theta^*)\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\leq Dh^\rho\}}\Big|^{p_3}\Big] < \infty. \tag{2.68}$$

Inequality (2.68) complete the estimate of the first term of the right-hand side of (2.65).

Next, we will estimate the second term and the third term of the right-hand side of (2.65). Let

$$M_i^n = \mathbb{D}_n^{(1)}(X_{t_{i-1}^n}, \Delta X_i^n, \theta^*)1_{\{|\Delta X_i^n|>Dh^\rho\}} - h\int_{B(X_{t_{i-1}^n})}\partial_\theta\Psi_{\theta^*}(y, X_{t_{i-1}^n})\varphi_n(X_{t_{i-1}^n}, y)dy.$$

Then by Remark 2.4, we have

$$E[M_i^n|\mathcal{F}_{i-1}] = R(\alpha, h^{11/8}\vee\frac{h\sqrt{h}}{\epsilon_n^2}, X_{t_{i-1}^n}),$$

$$E[(M_i^n)^l|\mathcal{F}_{i-1}] = R(\alpha, h, X_{t_{i-1}^n}) \quad (l=2,3,4).$$

So it follows that

$$\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{\sqrt{nh}}\sum_{i=1}^n E[M_i^n|\mathcal{F}_{i-1}]\Big|^{p_3}\Big] = C\sup_{n\in\mathbb{N}}\left(\frac{nh^{11/8}}{\sqrt{nh}}\vee\frac{nh\sqrt{h}}{\sqrt{nh}\epsilon_n^2}\right)^{p_3} < \infty,$$

$$\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{nh}\sum_{i=1}^n E[\psi_i^2(M_i^n)|\mathcal{F}_{i-1}]\Big|^{\frac{p_3}{2}}\Big] < \infty,$$

and

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{n^2h^2}\sum_{i=1}^n E[\psi_i^3(M_i^n)|\mathcal{F}_{i-1}]\Big|^{\frac{p_3}{4}}\Big] \\
=\ &\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{n^2h^2}\sum_{i=1}^n E\Big[\big\{(M_i^n)^2 - 2M_i^nE[M_i^n|\mathcal{F}_{i-1}] + 2E[M_i^n|\mathcal{F}_{i-1}]^2 - E[(M_i^n)^2|\mathcal{F}_{i-1}]\big\}^2|\mathcal{F}_{i-1}\Big]\Big|^{\frac{p_3}{4}}\Big] \\
=\ &C\sup_{n\in\mathbb{N}}\left(\frac{1}{nh}\right)^{\frac{p_3}{4}} < \infty.
\end{aligned}$$

Therefore by using Proposition 2.3 for $k=3$, we have

$$\sup_{n\in\mathbb{N}} E\Big[\Big|\frac{1}{\sqrt{nh}}\sum_{i=1}^n M_i^n\Big|^{p_3}\Big] \leq C\sup_{n\in\mathbb{N}}\left(\frac{1}{n^3h^4\epsilon_n^8}\right)^{\frac{p_3}{8}} + C < \infty.$$

This completes the proof. □

(2) By the definition of Y_n and Y^2,

$$Y_n(\theta;\alpha^*) - Y^2(\theta;\alpha^*) = \Lambda_9 + \Lambda_{10} + \Lambda_{11} + \Lambda_{12}$$

where

$$\Lambda_9 = \frac{1}{nh}\sum_{i=1}^{n}(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\hat{\sigma}_n)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}},$$

$$\begin{aligned}\Lambda_{10} &= -\frac{1}{2n}\sum_{i=1}^{n}(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\hat{\sigma}_n)(a_{i-1}(\theta)-a_{i-1}(\theta^*))1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\quad+\frac{1}{2}\int(a(x,\theta)-a(x,\theta^*))\beta^{-1}(x,\sigma^*)(a(x,\theta)-a(x,\theta^*))\pi(dx),\end{aligned}$$

$$\begin{aligned}\Lambda_{11} &= \frac{1}{nh}\sum_{i=1}^{n}\{\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta)-\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)\}1_{\{|\Delta X_i^n|>Dh^\rho\}}\\ &\quad-\int\int_{A(x)}(\log\Psi_\theta(y,x)-\log\Psi_{\theta^*}(y,x))\Psi_{\theta^*}(y,x)dy\pi(dx),\end{aligned}$$

and

$$\begin{aligned}\Lambda_{12} &= -\frac{1}{n}\sum_{i=1}^{n}\int_{B(X_{t_{i-1}^n})}(\Psi_\theta(y,X_{t_{i-1}^n})-\Psi_{\theta^*}(y,X_{t_{i-1}^n}))\varphi_n(X_{t_{i-1}^n},y)dy\\ &\quad+\int\int_{B(x)}(\Psi_\theta(y,x)-\Psi_{\theta^*}(y,x))dy\pi(dx).\end{aligned}$$

To estimate Λ_9, we first notice that by (2.46),

$$\begin{aligned}&\sup_{n\in\mathbb{N}}E\Big[\sup_{\theta\in\Theta}\Big|(nh)^\epsilon\frac{1}{nh}\sum_{i=1}^{n}(a_{i-1}(\theta)-a_{i-1}(\theta^*))(\beta_{i-1}^{-1}(\hat{\sigma}_n)-\beta_{i-1}^{-1}(\sigma^*))\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big|^{p_4}\Big]\\ \le\;& C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^\epsilon h^\rho}{\sqrt{nh}}\right)^{p_4}=C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^\epsilon}{\sqrt{nh}}h^{\rho-\frac{1}{2}}\right)^{p_4}\le C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^\epsilon}{n^\epsilon h^\epsilon}h^{\epsilon+\rho-\frac{1}{2}}\right)^{p_4}<\infty. \qquad (2.69)\end{aligned}$$

Moreover, by the Burkholder-Davis-Gundy inequality, Lemmas 2.2 and 2.3 and Sobolev's inequality, we have

$$\begin{aligned}&\sup_{n\in\mathbb{N}}E\Big[\sup_{\theta\in\Theta}\Big|(nh)^\epsilon\frac{1}{nh}\sum_{i=1}^{n}(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big|^{p_4}\Big]\\ \le\;& C\sup_{n\in\mathbb{N}}E\Big[\sup_{\theta\in\Theta}\Big|(nh)^\epsilon\frac{1}{nh}\sum_{i=1}^{n}\Big\{(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\qquad-E\left[(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right]\Big\}\Big|^{p_4}\Big]\\ &+C\sup_{n\in\mathbb{N}}E\Big[\sup_{\theta\in\Theta}\Big|(nh)^\epsilon\frac{1}{nh}\sum_{i=1}^{n}E\left[(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)\bar{X}_{i,n}1_{\{|\Delta X_i^n|\le Dh^\rho\}}|\mathcal{F}_{i-1}\right]\Big|^{p_4}\Big]\\ \le\;& C\sup_{n\in\mathbb{N}}\left((nh)^{2\epsilon}\frac{h^{2\rho}}{nh^2}\right)^{\frac{p_4}{2}}+C\sup_{n\in\mathbb{N}}(nh^2)^{\epsilon p_4}\le C\sup_{n\in\mathbb{N}}\left((nh)^{\frac{1}{3}}\frac{h^{\frac{3}{4}}}{nh^2}\right)^{\frac{p_4}{2}}+C\sup_{n\in\mathbb{N}}(nh^2)^{\epsilon p_4}<\infty. \qquad (2.70)\end{aligned}$$

Inequalities (2.69) and (2.70) give

$$\sup_{n\in\mathbb{N}}E\Big[\Big|\sup_{\theta\in\Theta}(nh)^\epsilon\Lambda_9\Big|^{p_4}\Big]<\infty.$$

Next, to estimate Λ_{10}, we have by (2.46),

$$\begin{aligned}&\sup_{n\in\mathbb{N}}E\Big[\sup_{\theta\in\Theta}\Big|\sqrt{nh}\frac{1}{2n}\sum_{i=1}^{n}\Big((a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\hat{\sigma}_n)(a_{i-1}(\theta)-a_{i-1}(\theta^*))1_{\{|\Delta X_i^n|\le Dh^\rho\}}\\ &\qquad-(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)(a_{i-1}(\theta)-a_{i-1}(\theta^*))1_{\{|\Delta X_i^n|\le Dh^\rho\}}\Big)\Big|^{p_4}\Big]<\infty. \qquad (2.71)\end{aligned}$$

Moreover, by using Proposition 2.2 and Sobolev's inequality, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|(nh)^{\epsilon}\frac{1}{2n}\sum_{i=1}^{n}\Big((a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)(a_{i-1}(\theta)-a_{i-1}(\theta^*))1_{\{|\Delta X_i^n|\le Dh^{\rho}\}}$$
$$-E\left[(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)(a_{i-1}(\theta)-a_{i-1}(\theta^*))1_{\{|\Delta X_i^n|\le Dh^{\rho}\}}\right]\Big)\Big|^{p_4}\Big]<\infty. \tag{2.72}$$

Furthermore, Lemma 2.2 yields

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|\sqrt{nh}\Big(\frac{1}{2n}\sum_{i=1}^{n}E\left[(a_{i-1}(\theta)-a_{i-1}(\theta^*))\beta_{i-1}^{-1}(\sigma^*)(a_{i-1}(\theta)-a_{i-1}(\theta^*))1_{\{|\Delta X_i^n|\le Dh^{\rho}\}}\right]$$
$$-\frac{1}{2}\int(a(x,\theta)-a(x,\theta^*))\beta^{-1}(x,\sigma^*)(a(x,\theta)-a(x,\theta^*))\pi(dx)\Big)\Big|^{p_4}\Big]<\infty. \tag{2.73}$$

Inequalities (2.71), (2.72) and (2.73) give

$$\sup_{n\in\mathbb{N}} E\Big[\Big|\sup_{\theta\in\Theta}(nh)^{\epsilon}\Lambda_{10}\Big|^{p_4}\Big]<\infty.$$

To estimate Λ_{11}, by using Proposition 2.3 for $k=2$, Remark 2.4, and Sobolev's inequality, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|\frac{(nh)^{\epsilon}}{nh}\sum_{i=1}^{n}\Big(\big\{\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta)-\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)\big\}1_{\{|\Delta X_i^n|>Dh^{\rho}\}}$$
$$-E\left[\big\{\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta)-\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)\big\}1_{\{|\Delta X_i^n|>Dh^{\rho}\}}|\mathcal{F}_{i-1}\right]\Big)\Big|^{p_4}\Big]$$
$$\le C\sup_{n\in\mathbb{N}}\left(\frac{(nh)^{4\epsilon}}{n^3h^4\epsilon_n^4}\right)^{\frac{p_4}{4}}+C=C\sup_{n\in\mathbb{N}}\left(\left(\frac{1}{n^3h^4\epsilon_n^{16}}\right)^{\frac14}\cdot\frac{(nh)^{\frac23}}{n^{\frac94}h^3}\right)^{\frac{p_4}{4}}+C<\infty. \tag{2.74}$$

Moreover, by Proposition 2.2, Remark 2.4, $[H11]$, and Sobolev's inequality, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|(nh)^{a\wedge\epsilon}\Big(\frac{1}{nh}\sum_{i=1}^{n}E\left[\{\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta)-\mathbb{D}_n^{(0)}(X_{t_{i-1}^n},\Delta X_i^n,\theta^*)\}1_{\{|\Delta X_i^n|>Dh^{\rho}\}}|\mathcal{F}_{i-1}\right]$$
$$-\int\int_{A(x)}(\log\Psi_{\theta}(y,x)-\log\Psi_{\theta^*}(y,x))\Psi_{\theta^*}(y,x)dy\pi(dx)\Big)\Big|^{p_4}\Big]<\infty. \tag{2.75}$$

Inequalities (2.74) and (2.75) lead

$$\sup_{n\in\mathbb{N}} E\Big[\Big|\sup_{\theta\in\Theta}(nh)^{a\wedge\epsilon}\Lambda_{11}\Big|^{p_4}\Big]<\infty.$$

Finally, to estimate Λ_{12}, by Proposition 2.2 and Sobolev's inequality, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|(nh)^{\epsilon}\frac{1}{n}\sum_{i=1}^{n}\Big(\int_{B(X_{t_{i-1}^n})}(\Psi_{\theta}(y,X_{t_{i-1}^n})-\Psi_{\theta^*}(y,X_{t_{i-1}^n}))\varphi_n(X_{t_{i-1}^n},y)dy$$
$$-\int\int_{B(x)}(\Psi_{\theta}(y,x)-\Psi_{\theta^*}(y,x))\varphi_n(x,y)dy\pi(dx)\Big)\Big|^{p_4}\Big]<\infty. \tag{2.76}$$

By $[H11]$, we have

$$\sup_{n\in\mathbb{N}} E\Big[\sup_{\theta\in\Theta}\Big|(nh)^{a}\int\int_{B(x)}(\Psi_{\theta}(y,x)-\Psi_{\theta^*}(y,x))(1-\varphi_n(x,y))dy\pi(dx)\Big)\Big|^{p_4}\Big]$$
$$\le C\sup_{n\in\mathbb{N}}((nh)^a h^a)^{p_4}<\infty. \tag{2.77}$$

Inequalities (2.76) and (2.77) give

$$\sup_{n\in\mathbb{N}} E\Big[\Big|\sup_{\theta\in\Theta}(nh)^{a\wedge\epsilon}\Lambda_{12}\Big|^{p_4}\Big]<\infty.$$

This completes the proof. □

Using Lemmas 2.6 and 2.7, we can verify the condition of Theorem 2.3 for any $L > 0$, with the parameters satisfying

$$\begin{aligned}
(\frac{1}{2}-\epsilon)\vee(\frac{1}{2}-a) &\leq \beta_2 < \frac{1}{2}, \\
0 &< \rho_2 < 1-2\beta_2, \\
0 &< \alpha < \frac{\rho_2}{2}, \\
\beta &= \frac{\alpha}{1-\alpha}, \\
0 &< \beta_1 \leq a \wedge \epsilon, \\
0 &< \rho_1 < 1 \wedge \beta \wedge \frac{2\beta_1}{1-\alpha}.
\end{aligned}$$

So by Theorem 2.3, there exists $C_L > 0$ such that

$$P_{\alpha^*}\left[\sup_{u_2 \in V_n^2(r)} Z_n^2(u_2;\hat{\sigma}_n,\theta^*) \geq e^{-\frac{r}{2}}\right] \leq \frac{C_L}{r^L}, \tag{2.78}$$

for all $n \in \mathbb{N}$ and $r > 0$, where

$$V_n^2(r) := \{u_2 \in \mathbb{R}^{d_2}; \theta^* + \frac{u_2}{\sqrt{nh}} \in \Theta, |u_2| \geq r\},$$

and

$$Z_n^2(u_2,\sigma;\theta^*) := \exp\left(H_n(\sigma,\theta^* + \frac{u_2}{\sqrt{nh}}) - H_n(\sigma,\theta^*)\right).$$

Proof of Theorem 2.1. We will prove Theorem 2.1 by using Theorem 2.4. To use Theorem 2.4, we first prove weak convergence of $Z_n(u_1,u_2;\alpha^*)$. By the Burkholder-Davis-Gundy inequality and Sobolev's inequality, we have for any $p \geq 1$,

$$\begin{aligned}
& E\left[\sup_{\alpha\in\Xi}\Big|\frac{1}{\sqrt{n}}\partial_\sigma\partial_\theta H_n(\sigma,\theta)\Big|^p\right] \\
= \ & E\left[\sup_{\alpha\in\Xi}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n \partial_\theta a_{i-1}(\theta)\partial_\sigma\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\leq Dh^\rho\}}\Big|^p\right] \\
\leq \ & CE\Big[\sup_{\alpha\in\Xi}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n\{\partial_\theta a_{i-1}(\theta)\partial_\sigma\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\leq Dh^\rho\}} \\
& \quad - E[\partial_\theta a_{i-1}(\theta)\partial_\sigma\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|\mathcal{F}_{i-1}]\}\Big|^p\Big] \\
& +CE\left[\sup_{\alpha\in\Xi}\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^n E[\partial_\theta a_{i-1}(\theta)\partial_\sigma\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)1_{\{|\Delta X_i^n|\leq Dh^\rho\}}|\mathcal{F}_{i-1}]\Big|^p\right] \\
\to \ & 0 \quad (n\to\infty).
\end{aligned}$$

Therefore, by using Lemmas 2.4 and 2.5, we have

$$\begin{aligned}
\log Z_n^1(u_1;\theta,\sigma^*) &= H_n(\sigma^* + \frac{u_1}{\sqrt{n}},\theta) - H_n(\sigma^*,\theta) \\
&= \frac{1}{\sqrt{n}}\partial_\sigma H_n(\sigma^*,\theta)[u_1] + \frac{1}{2n}\partial_\sigma^2 H_n(\sigma^*,\theta)[(u_1)^{\otimes 2}] \\
&\quad + \int_0^1 \frac{(1-t)^2}{2}\partial_\sigma^3 H_n(\sigma^* + \frac{u_1}{\sqrt{n}}t,\theta)\left[\left(\frac{u_1}{\sqrt{n}}\right)^{\otimes 3}\right]dt \\
&= \frac{1}{\sqrt{n}}\partial_\sigma H_n(\alpha^*)[u_1] - \frac{1}{2}\Gamma^1(\alpha^*)[(u_1)^{\otimes 2}] + o_p(1) \\
&\to^d \triangle_1[u_1] - \frac{1}{2}\Gamma^1(\alpha^*)[(u_1)^{\otimes 2}],
\end{aligned}$$

where $\triangle_1 \sim N_{d_1}(0, \Gamma^1(\alpha^*))$ by martingale central limit theorem. Similarly,

$$\begin{aligned}
\log Z_n^2(u_2;\sigma^*,\theta^*) &= H_n(\sigma^*,\theta^* + \frac{u_2}{\sqrt{nh}}) - H_n(\sigma^*,\theta^*) \\
&= \frac{1}{\sqrt{nh}}\partial_\theta H_n(\sigma^*,\theta^*)[u_2] + \frac{1}{2nh}\partial_\theta^2 H_n(\sigma^*,\theta^*) \\
&\quad + \int_0^1 \frac{(1-t)^2}{2}\partial_\theta^3 H_n(\sigma^*,\theta^* + \frac{u_2}{\sqrt{nh}}t)\left[\left(\frac{u_2}{\sqrt{nh}}\right)^{\otimes 3}\right] dt \\
&= \frac{1}{\sqrt{nh}}\partial_\theta H_n(\alpha^*)[u_2] - \frac{1}{2}\Gamma^2(\alpha^*)[(u_2)^{\otimes 2}] + o_p(1) \\
&\to^d \triangle_2\,[u_2] - \frac{1}{2}\Gamma^2(\alpha^*)[(u_2)^{\otimes 2}],
\end{aligned}$$

where $\triangle_2 \sim N_{d_2}(0, \Gamma^2(\alpha^*))$. Moreover, similar arguement yield

$$\begin{aligned}
&\log Z_n(u_1,u_2;\alpha^*) \\
= \;& \log Z_n^1(u_1;\theta^* + \frac{u_2}{\sqrt{nh}},\sigma^*) + \log Z_n^2(u_2;\sigma^*,\theta^*) \\
\to^d \;& \triangle_1\,[u_1] + \triangle_2[u_2] - \frac{1}{2}\Gamma^1(\alpha^*)[(u_1)^{\otimes 2}] - \frac{1}{2}\Gamma^2(\alpha^*)[(u_2)^{\otimes 2}] \quad (n\to\infty),
\end{aligned}$$

where $\triangle_1$ and $\triangle_2$ are independent.

On the other hand, let $B(R) = \{(u_1,u_2) \in \mathbb{R}^{d_1}\times\mathbb{R}^{d_2}; |u_1|^2 + |u_2|^2 \le R^2\}$, then the tightness of the family $\{\log Z_n(u_1,u_2;\alpha^*)|_{C(B(R))}; n\in\mathbb{N}\}$ follows if we show

$$\sup_{n\in\mathbb{N}} E\left[\sup_{u\in C(B(R))} |\partial_u \log Z_n(u_1,u_2;\alpha^*)|\right] < \infty \tag{2.79}$$

because of the tightness criterion in C space in Billingsley [8]. However, since

$$\begin{aligned}
&\partial_u \log Z_n(u_1,u_2;\alpha^*) \\
= \;& \frac{\partial}{\partial u}\left(H_n(\sigma^* + \frac{u_1}{\sqrt{n}},\theta^* + \frac{u_2}{\sqrt{nh}}) - H_n(\sigma^*,\theta^*)\right) \\
= \;& \left(\frac{1}{\sqrt{n}}\partial_\sigma H_n(\sigma^* + \frac{u_1}{\sqrt{n}},\theta^* + \frac{u_2}{\sqrt{nh}}), \frac{1}{\sqrt{nh}}\partial_\theta H_n(\sigma^* + \frac{u_1}{\sqrt{n}},\theta^* + \frac{u_2}{\sqrt{nh}})\right),
\end{aligned}$$

by using Lemmas 2.4 and 2.5, and similar to the proof of Lemmas 2.6 and 2.7, we can prove (2.79).

Then for

$$Z(u_1,u_2;\alpha^*) := \exp\left(\triangle_1[u_1] + \triangle_2[u_2] - \frac{1}{2}\Gamma^1(\alpha^*)[u_1^{\otimes 2}] - \frac{1}{2}\Gamma^2(\alpha^*)[u_2^{\otimes 2}]\right),$$

it follows that

$$Z_n(u_1,u_2;\alpha^*) \to^d Z(u_1,u_2;\alpha^*) \quad in \quad C(B(R)) \tag{2.80}$$

as $n\to\infty$. So the weak convergence of $Z_n(u_1,u_2;\alpha^*)$ is proved.

Next, we prove the moment condition of $\hat{u}_n$. By a similar argument to the derivation of (2.46), using (2.78), we have for any $p > 0$,

$$\sup_{n\in\mathbb{N}} E\left[(\sqrt{nh}(\hat{\theta}_n - \theta^*))^p\right] < \infty.$$

Therefore, for $\hat{u}_n := (\sqrt{n}(\hat{\sigma}_n - \sigma^*), \sqrt{nh}(\hat{\theta}_n - \theta^*))$,

$$\sup_{n\in\mathbb{N}} E\left[|\hat{u}_n|^p\right] < \infty. \tag{2.81}$$

By (2.80), (2.81) and Theorem 2.4, we have

$$\hat{u}_n \to^d \hat{u} := (\Gamma^1(\alpha^*)^{-1}\Delta_1, \Gamma^2(\alpha^*)^{-1}\Delta_2) \sim N(0, diag(\Gamma^1(\alpha^*)^{-1}, \Gamma^2(\alpha^*)^{-1})),$$

and

$$E[\mathsf{f}(\hat{u}_n)] \to \mathbb{E}[\mathsf{f}(\hat{u})],$$

as $n \to \infty$, for any continuous function f with at most polynomial growth. $\square$

Proof of Theorem 2.2. We will prove Theorem 2.2 by using Theorem 2.5. First, we prove weak convergence of $(Z_n^1(\cdot;\theta^\star,\sigma^*), Z_n^2(\cdot,\tilde{\sigma}_n,\theta^*))$. Let

$$Z^1(u_1;\alpha^*) = \exp\left(\triangle_1[u_1] - \frac{1}{2}\Gamma^1(\alpha^*)[u_1^{\otimes 2}]\right),$$

and

$$Z^2(u_2;\alpha^*) = \exp\left(\triangle_2[u_2] - \frac{1}{2}\Gamma^2(\alpha^*)[u_2^{\otimes 2}]\right),$$

then by the proof of Theorem 2.1, $\log Z_n^1(u_1;\theta^\star,\sigma^*) \to \log Z^1(u_1;\alpha^*)$ as $n \to \infty$. Next, By Lemmas 2.4, 2.5 and Lemma 2 in Yoshida [54], there exists $\delta > 0$ such that

$$\sup_{n \geq N} E\left[\left(\int_{|u_1| \leq \delta} Z_n^1(u_1;\theta^\star,\sigma^*)\pi_{1,n}(\sigma^* + \frac{u_1}{\sqrt{n}})du_1\right)^{-1}\right] < \infty. \tag{2.82}$$

So by using Theorem 2.5 for fixed $\theta^\star$, we have

$$\sup_{n \in \mathbb{N}} E[|\sqrt{n}(\tilde{\sigma}_n - \sigma^*)|^p] < \infty, \tag{2.83}$$

for any $p > 0$. Therefore, in a similar way to the proof of Lemmas 2.6 and 2.7 and discussion after them, we have for any $L > 0$,

$$P_{\alpha^*}\left[\sup_{u_2 \in V_n^2(r)} Z_n^2(u_2;\tilde{\sigma}_n,\theta^*) \geq e^{-\frac{r}{2}}\right] \leq \frac{C_L}{r^L}, \tag{2.84}$$

for any $n \in \mathbb{N}$ and $r > 0$.

Moreover, similarly to the proof of Theorem 2.1, it follows that

$$(Z_n^1(\cdot;\theta^\star,\sigma^*), Z_n^2(\cdot,\tilde{\sigma}_n,\theta^*)) \to^d (Z^1(\cdot;\alpha^*), Z^2(\cdot;\alpha^*)) \quad \text{in } C(B(R)) \quad (n \to \infty), \tag{2.85}$$

and there exists $\delta' > 0$ such that

$$\sup_{n \geq N} E\left[\left(\int_{|u_2| \leq \delta'} Z_n^2(u_2;\tilde{\sigma}_n,\theta^*)\pi_{2,n}(\theta^* + \frac{u_2}{\sqrt{nh}})du_2\right)^{-1}\right] < \infty. \tag{2.86}$$

By (2.45), (2.82), (2.84), (2.85), (2.86) and Theorem 2.5, we have

$$\begin{aligned} & \tilde{u}_n := (\sqrt{n}(\tilde{\sigma}_n - \sigma^*), \sqrt{nh}(\tilde{\theta}_n - \theta^*)) \\ \to^d \quad & \tilde{u} := (\Gamma^1(\alpha^*)^{-1}\Delta_1, \Gamma^2(\alpha^*)^{-1}\Delta_2) \sim N(0, diag(\Gamma^1(\alpha^*)^{-1}, \Gamma^2(\alpha^*)^{-1})), \end{aligned}$$

and

$$E[\mathsf{f}(\tilde{u}_n)] \to^d \mathbb{E}[\mathsf{f}(\tilde{u})],$$

as $n \to \infty$. for any continuous function f with at most polynomial growth. $\square$

Bibliography

[1] Adams, R. A.: Sobolev spaces. Pure and Applied Mathematics (Vol. 65). London, New York: Academic Press (1975)

[2] Adams, R. A. and Fournier, J. J. F. : Sobolev spaces. Second edition. Pure and Applied Mathematics (Amsterdam), 140. Elsevier / Academic Press, Amsterdam (2003)

[3] Aït-Sahalia, Y., Fan, J., and Xiu, D.: High-Frequency Covariance Estimates With Noisy and Asynchronous Financial Data. *Journal of the American Statistical Association*, **105**, 1504-1517. (2010)

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

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

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

[7] Bichteler, K. and Jacod, J.: Calcul de Malliavin pour les diffusions avec sauts: existence d'une densite dans le cas unidimensionnel, *Lecture Notes in Math*, Springer-Verlag, Berlin, 132-157 (1986)

[8] Billingsley, P.: Convergence of Probability Measures. Willey, New York (1968)

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

[10] Cont, R. and Tankov, P.: *Financial modeling with Jump Processes*, Financial Mathematics Series, Chapman & Hall/CRC, 2004

[11] Dohnal, G.: On estimating the diffusion coefficient, *Journal of Applied Probability*, **24**, 105-114. (1987)

[12] Doukhan, P. and Louhichi, S.: A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications* **84**, 313-342. (1999)

[13] Epps, T. W.: Comovements in Stock Prices in the Very Short Run. *Journal of the American Statistical Association*, **74**, 291-298. (1979)

[14] Genon-Catalot, V. and Jacod, J.: On the estimation of the diffusion coefficient for multidimensional diffusion processes, *Annales de I'IHP Probabilités et statistiques*, **29**, 119-151. (1993)

[15] Genon-Catalot, V. and Jacod, J.: Estimation of the diffusion coefficient for diffusion processes: random sampling, *Scandinavian Journal of Statistics*, **21**, 193-221. (1994)

[16] Gobet, E.: Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach, *Bernoulli*, **7**, 899-912. (2001)

[17] Hayashi, T. and Yoshida, N.: On covariance estimation of nonsynchronously observed diffusion processes. *Bernoulli*, **11**, 359-379. (2005)

[18] Hayashi, T. and Yoshida, N.: Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Annals of the Institute of Statistical Mathematics*, **60**, 367-406. (2008)

[19] Hayashi, T. and Yoshida, N.: Nonsynchronous covariation process and limit theorems. *Stochastic Processes and their Applications*, **121**, 2416-2454. (2011)

[20] Ibragimov, I.A., Has'minskii, R.Z.: The asymptotic behavior of certain statistical estimates in the smooth case. I. Investigation of the likelihood ratio. (Russian) *Teorija Verojatnostei i ee Primenenija*, **17**, 469-486. (1972)

[21] Ibragimov, I.A., Has'minskii, R.Z.: Asymptotic behavior of certain statistical estimates. II. Limit theorems for a posteriori density and for Bayesian estimates. (Russian) *Teorija Verojatnostei i ee Primenenija*, **18**, 78-93. (1973)

[22] Ibragimov, I.A., Has'minskii, R.Z.: *Statistical estimation: asymptotic theory.* Springer, New York, 1981

[23] Jacod, J.: On continuous conditional Gaussian martingales and stable convergence in law. Seminaire de Probabilites, XXXI, 232-246, Lecture Notes in Math., 1655, Springer, Berlin (1997)

[24] Kessler, M.: Estimation of an ergodic diffusion from discrete observations, *Scandinavian Journal of Statistics*, **24**, 211-229. (1997)

[25] Kutoyants, Yu. A.: *Parameter estimation for stochastic processes*, Translated and edited by B.L.S.Prakasa Rao, Herdermann, Berlin, 1984

[26] Kutoyants, Yu. A.: *Identification of dynamical systems with small noise*, Kluwer, Dordrecht Boston London, 1994

[27] Kutoyants, Yu. A.: *Statistical inference for spatial Poisson processes*, *Lecture Notes in Statistics*, 134. Springer, Berlin Heiderberg New York London Paris Tokyo Hong Kong, 1998

[28] Kutoyants, Yu. A.: *Statistical inference for ergodic diffusion processes*, Springer Series in Statistics. Springer-Verlag London, Ltd., London. 2004.

[29] Malliavin, P. and Mancino, M. E.: Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, **6**, 49-61. (2002)

[30] Mancini, C.: Estimation of the characteristics of the jumps of a general Poisson-diffusion model. Scand. Actuarial J., 1, 42-52 (2004)

[31] Masuda, H.: Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* 117, no. 1, 35–56 (2007)

[32] Masuda, H.: On stability of diffusions with compound-Poisson jumps. Bull. Inform. Cybernet. 40, 61–74 (2008)

[33] Masuda, H.: Approximate self-weighted LAD estimation of discretely observed ergodic Ornstein-Uhlenbeck processes, *Electronic Journal of Statistics*, **4**, 525-565. (2010)

[34] Ogihara, T. and Yoshida, N.: Quasi-likelihood analysis for the stochastic differential equation with jumps, *Statistical Inference for Stochastic Processes*, **14**, 189-229. (2011)

[35] Ogihara, T. and Yoshida, N.: Quasi-likelihood analysis for nonsynchronously observed diffusion processes, *Stochastic Processes and their Applications*, **124**, 2954-3008. (2014)

[36] Prakasa Rao, B.L.S.: Asymptotic theory for nonlinear learst squares estimator for diffusion processes, *Mathematische Operationsforschung Statistik. Series Statistics*, **14**, 195-209. (1983)

[37] Prakasa Rao, B.L.S.: Statistical inference from sampled data for stochastic processes, *Contemporary Mathematics - American Mathematical Society*, **80**, 249-284. (1988)

[38] Prakasa Rao, B. L. S.: *Semimartingales and Their Statistical Inference*, Monographs on Statistics and Applied Probability, 83, Chapman & Hall/CRC, 1999

[39] Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Springer, Berlin (1999)

[40] Sato, K.: *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, 1999

[41] Sato, K., Yamazato, M.: Operator-self-decomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type, *Stochastic Process. Appl.*, 17,73-100 (1984)

[42] Shimizu, Y.: M-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. *Stat. Inference Stoch. Process.* 9, no. 2, 179–225 (2006)

[43] Shimizu, Y.: Density estimation of Levy measures for discretely observed diffusion processes with jumps. *J. Japan Statist. Soc.* 36, 37-62 (2007)

[44] Shimizu, Y.: A practical inference for discretely observed jump-diffusions from finite samples. *J. Japan Statist. Soc.* 38, 391-413 (2008)

[45] Shimizu, Y., Yoshida, N.: Janpugata kakusankatei no risankansokukarano suiteinitsuite. (Estimation of diffusion processes with jumps based on discretely sampled observations.) The 2003 Japanese Joint Statistical Meeting, Meijo University, September 2-5, 2003.

[46] Shimizu, Y. and Yoshida, N.: Estimation of parameters for diffusion processes with jumps from discrete observations, *Statistical Inference for Stochastic Processes*, **9**, 227-277. (2006)

[47] Sørensen, M. and Uchida, M.: Small diffusion asymptotics for discretely sampled stochastic differential equations, *Bernoulli*, **9**, 1051-1069. (2003)

[48] Uchida, M.: Estimation for dynamical systems with small noise from discrete observations, *Journal of the Japan Statistical Society*, **33**, 157-167. (2003)

[49] Uchida, M.: Estimation for discretely observed small diffusions based on approximate martingale estimating functions, *Scandinavian Journal of Statistics*, **31**, 553-566. (2004)

[50] Uchida, M. and Yoshida, N.: Nondegeneracy of statistical random field and quasi likelihood analysis for diffusion. Research Memorandum 1149, Institute of Statistical Mathematics (2011)

[51] Uchida, M. and Yoshida, N.: Nondegeneracy of Random Field and Estimation of Diffusion. arXiv:1212.5715. (2012)

[52] Yoshida, N.: Estimation for diffusion processes from discrete observation, *Journal of Multivariate Analysis*, **41**, 220-242. (1992)

[53] Yoshida, N.: Polynomial type large deviation inequalities and convergence of statistical random fields, *ISM Research Memorandum 1021*, Institute of Statistical Mathematics. (2006)

[54] Yoshida, N.: Polynomial Type Large Deviation Inequalities and Quasi-Likelihood Analysis for Stochastic Differential Equations. *Ann. Inst. Statist. Math* **63**, 431-479. (2011)

[55] Wolfe, S. J.: On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$, *Stochastic Process. Appl.*, 12, 301-312 (1982)