Statistical Inference for Stochastic Processes (2023) 26:279–330
https://doi.org/10.1007/s11203-023-09288-w

Parameter estimation for ergodic linear SDEs from partial and discrete observations

Masahiro Kurisaki[1]

Received: 23 March 2022 / Accepted: 31 March 2023 / Published online: 28 April 2023

Abstract

We consider a problem of parameter estimation for the state space model described by linear stochastic differential equations. We assume that an unobservable Ornstein–Uhlenbeck process drives another observable process by the linear stochastic differential equation, and these two processes depend on some unknown parameters. We construct the quasi-maximum likelihood estimator of the unknown parameters and show asymptotic properties of the estimator.

Keywords Partially observed linear model · State space model · Hidden Ornstein Uhlenbeck model · Kalman–Bucy filter · Quasi-likelihood analysis

1 Introduction

On the probability space $(\Omega, \mathcal{F}, P)$ with a complete and right-continuous filtration $\{\mathcal{F}_t\}$, we consider a $(d_1 + d_2)$-dimensional Gaussian process (X_t, Y_t) satisfying the following stochastic differential equations:

$$dX_t = -a(\theta_2)X_t dt + b(\theta_2)dW_t^1, \tag{1.1}$$

$$dY_t = c(\theta_2)X_t dt + \sigma(\theta_1)dW_t^2, \tag{1.2}$$

where W^1 and W^2 are independent d_1 and d_2-dimensional $\{\mathcal{F}_t\}$-Wiener processes, (X_0, Y_0) is a gaussian random variable independent of W^1 and W^2, $\theta_1 \in \Theta_1 \subset \mathbb{R}^{m_1}$ and $\theta_2 \in \Theta_2 \subset \mathbb{R}^{m_2}$ are unknown parameters, and $a, b : \Theta_2 \to M_{d_1}(\mathbb{R}), c : \Theta_2 \to M_{d_2,d_1}(\mathbb{R})$ and $\sigma : \Theta_1 \to M_{d_2}(\mathbb{R})$ are known functions. Here $M_{m,n}(\mathbb{R})$ is the set of $m \times n$ matrices over $\mathbb{R}$ and $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$, Θ_1 and Θ_2 are known parameter spaces. The solution of (1.1) is an Ornstein–Uhlenbeck process and it has an ergodic property.

We assume that the process X is unobservable, and the purpose of this article is to construct estimators of θ_1 and θ_2 based on discrete observations of Y; we assume discrete observations $Y_{t_0}, Y_{t_1}, Y_{t_2}, \cdots, Y_{t_n}$ where $t_i = ih_n$ for some $h_n > 0$, instead of considering the continuous observation $\{Y_t\}_{0 \le t \le T}$. The discrete observation case is much more complicated and

✉ Masahiro Kurisaki
makurisaki@g.ecc.u-tokyo.ac.jp

1 Graduate School of Mathematical Sciences, The University of Tokyo, Tokyo, Japan

interesting because of the construction of the estimator for θ_2, which is the main object of this paper and described in detail in Sect. 3, and also because there is no need to estimate θ_1 (Kutoyants 2004, 2019a).

Note that we can not identify $b(\theta_2)$ and $c(\theta_2)$ simultaneously from observation of $\{Y_t\}$. In fact, the system

$$dX_t = -a(\theta_2)X_t dt + 2b(\theta_2)dW_t^1$$
$$dY_t = \frac{1}{2}c(\theta_2)X_t dt + \sigma(\theta_1)dW_t^2$$

generates the same $\{Y_t\}$ as (1.1) and (1.2). Therefore, we need to impose some restrictions on a, b, c, σ and the dimensions of the parameter spaces.

When θ_1 and θ_2 are known, one can estimate the unobservable state $\{X_t\}$ from observations of $\{Y_t\}$ by the following well-known Kalman–Bucy filter.

Theorem 1.1 (Theorem 10.2, Liptser and Shiriaev 2001a) *In* (1.1) *and* (1.2)*, let* $\sigma(\theta)\sigma(\theta)'$ *be positive definite, where the prime means the transpose. Then* $m_t = E[X_t|\{Y_t\}_{0\le s\le t}]$ *and* $\gamma_t = E[(X_t - m_t)(X_t - m_t)']$ *are the solutions of the equations*

$$dm_t = -a(\theta_2)m_t dt + \gamma_t c(\theta_2)'\{\sigma(\theta_1)\sigma(\theta_1)'\}^{-1}\{dY_t - c(\theta_2)m_t dt\}, \tag{1.3}$$

$$\frac{d\gamma_t}{dt} = -a(\theta_2)\gamma_t - \gamma_t a(\theta_2)' - \gamma_t c(\theta_2)'\{\sigma(\theta_1)\sigma(\theta_1)'\}^{-1}c(\theta_2)\gamma_t + b(\theta_2)b(\theta_2)'. \tag{1.4}$$

Equation (1.4) is the matrix Riccati equation, which has been examined in the theory of linear quadratic control (Sontag 2013). It is known that (1.4) has the unique positive-semidefinite solution (Liptser and Shiriaev 2001a). Moreover, under proper conditions, one can show that the corresponding algebraic Riccati equation

$$-a(\theta_2)\gamma - \gamma a(\theta_2)' - \gamma c(\theta_2)'\{\sigma(\theta_1)\sigma(\theta_1)'\}^{-1}c(\theta_2)\gamma + b(\theta_2)b(\theta_2)' = O \tag{1.5}$$

has the maximal and minimal solutions (Coppel 1974; Zhou et al. 1996), and the solution of (1.4) converges to the maximal solution of (1.5) at an exponential rate (Leipnik 1985). Further details on this topic will be discussed in Sect. 3.

There are already several studies on parameter estimation in the system (1.1) and (1.2) with the Kalman–Bucy filter. For example, Kutoyants (2004) discusses the ergodic case, Kutoyants (1994) and Kutoyants (2019b) small noise cases, and Kutoyants (2019a) the one-step estimator. However, all of them assume $d_1 = d_2 = 1$ and need continuous observation of Y. The continuous observation case is simpler, because we do not have to estimate θ_1. In fact, we have

$${Y_t}^2 - {Y_0}^2 = 2\int_0^t Y_s dY_s + \sigma(\theta_1)^2 t$$

by Itô's formula and (1.1), and therefore we can get the exact value of $\sigma(\theta_1)$.

On the other hand, parametric inference for discretely observed stochastic differential equations without an unobservable process has been studied for decades (for example Sørensen 2002; Shimizu and Yoshida 2006; Yoshida 1992). Especially, Yoshida (2011) developed Ibragimov–Khasminskii theory (Ibragimov and Has' Minskii 1981) into the quasi-likelihood analysis, and investigated the behavior of the quasi-likelihood estimator and the adaptive Bayes estimator in the ergodic diffusion process. Quasi-likelihood analysis is helpful to discretely observed cases, and many works have been derived from it: see Uchida and Yoshida (2012) for the non-ergodic case, Ogihara and Yoshida (2011) for the jump case, Masuda (2019) for the Lévy driven case, Gloter and Yoshida (2021) for the degenerate case, Kamatani and Uchida (2015) for the multi-step estimator, and Nakakita et al. (2021) for the case with observation noises.

This paper also makes use of quasi-likelihood analysis to investigate the behaviors of our estimators. In Sect. 2, we describe the more precise setup and present asymptotic properties of our estimators, which are main results of this paper. Then we go on to proofs of these results in Sects. 3 and 5. We first discuss the estimation of θ_2 in Sect. 3 because it is the main part of this article, whereas estimation of θ_1 is quite parallel to the usual case without an unobservable variable. We also examine the Riccati differential equation (1.4) and algebraic Riccati equation (1.5) in Sect. 3. In Sect. 4, we discuss the special case where $d_1 = d_2 = 1$. In the one-dimensional case, we can reduce our assumptions to simpler ones. In Sect. 5, we discuss estimation of θ_1. Finally, we show in Sect. 6 the result of computational simulation by YUIMA (Brouste et al. 2014), an package on R, and suggest a way to improve our estimators when the wrong initial value is given.

2 Notations, assumptions and main results

Let $\theta_1^* \in \mathbb{R}^{m_1}$ and $\theta_2^* \in \mathbb{R}^{m_2}$ be the true values of θ_1 and θ_2, respectively, and define the $(d_1 + d_2)$-dimensional Gaussian process (X_t, Y_t) by

$$dX_t = -a(\theta_2^*)X_t dt + b(\theta_2^*)dW_t^1, \tag{2.1}$$

$$dY_t = c(\theta_2^*)X_t dt + \sigma(\theta_1^*)dW_t^2, \tag{2.2}$$

where $W_1, W_2, X_0, Y_0, a, b, c$ and σ are the same as Sect. 1; $a, b : \Theta_2 \to M_{d_1}(\mathbb{R})$, $c : \Theta_2 \to M_{d_2,d_1}(\mathbb{R})$ and $\sigma : \Theta_1 \to M_{d_2}(\mathbb{R})$. In this article, we have access to the discrete observations Y_{ih_n} $(i = 0, 1, \cdots, n)$, where h_n is some positive constant, and we construct the estimators of θ_1 and θ_2 based on the observations.

We assume that $\Theta_1 \subset \mathbb{R}^{m_1}$ and $\Theta_2 \subset \mathbb{R}^{m_2}$ are open bounded subsets and that the Sobolev embedding inequality holds on $\Theta = \Theta_1 \times \Theta_2$; for any $p > m_1 + m_2$ and $f \in C^1(\Theta)$, there exists some constant C depending only on Θ such that

$$\sup_{\theta \in \Theta} |f(\theta)| \leq C \left(\|f\|_{L^p} + \|\partial_{\theta_i} f\|_{L^p} \right). \tag{2.3}$$

For example, if each Θ_i $(i = 1, 2)$ has a Lipchitz boundary, this inequality is valid (Leoni 2017).

Let $Z(\theta)$ $(\theta \in \Theta = \Theta_1 \times \Theta_2)$ be a class of random variables, where $Z(\theta)$ is continuously differentiable with respect to θ. Then by (2.3) and Fubini's theorem, we get for any $p > m_1 + m_2$

$$\begin{aligned} E\left[\sup_{\theta \in \Theta} |Z(\theta)|^p\right] &\leq C2^{p-1} \left(E\left[\int_{\Theta_i} |Z(\theta)|^p d\theta_i + \int_{\Theta} |\partial_\theta Z(\theta)|^p d\theta_i \right] \right) \\ &= C2^{p-1} \left(\int_{\Theta_i} E[|Z(\theta)|^p] d\theta + \int_{\Theta} E[|\partial_\theta Z(\theta)|^p] d\theta \right) \\ &\leq C_p \sup_{\theta \in \Theta} \left(E[|Z(\theta)|^p] + E[|\partial_\theta Z(\theta)|^p] \right), \end{aligned}$$

where C_p is some constant depending on p and Θ. This result will be frequently referred to in the following sections.

In what follows, we use the following notations:

- $\mathbb{R}_+ = [0, \infty)$, $\mathbb{N} = \{1, 2, \cdots\}$.

- $\Theta = \Theta_1 \times \Theta_2, \theta_1 = (\theta_1^1, \cdots, \theta_1^{m_1}), \theta_2 = (\theta_2^1, \cdots, \theta_2^{m_2}), \theta^* = (\theta_1^*, \theta_2^*)$.
- For any subset $\Xi \subset \mathbb{R}^m$, $\overline{\Xi}$ is the closure of Ξ.
- For every set of matrices A, B and C, A' is the transpose of A, $A^{\otimes 2} = AA'$, $A[B, C] = B'AC$ and $A[B^{\otimes 2}] = B'AB$.
- For every matrix $A, |A|$ is the Frobenius norm of A. Namely, if $A = (a_{ij})_{1\leq i\leq n, 1\leq j\leq m}$, $|A|$ is defined by

$$|A| = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2}.$$

- For every matrix A, $\lambda_{\min}(A)$ denotes the smallest real part of eigenvalues of matrix A.
- For every symmetric matrix A and $B \in M_d(\mathbb{R})$, $A > B$ (resp. $A \geq B$) means that $A - B$ is positive (resp. semi-positive) definite.
- For any open subset $\Xi \subset \mathbb{R}^m$ and $A : \Xi \to M_d(\mathbb{R})$ of class C^k, $\partial_\xi^k A(\xi)$ denotes the k-dimensional tensor on $M_d(\mathbb{R})$ whose $(j_1, j_2, \cdots, j_k)$ entry is $\frac{\partial}{\partial \xi_{j_1}} \cdots \frac{\partial}{\partial \xi_{j_k}} A(\theta_i)$, where $1 \leq j_1, \cdots, j_k \leq m$ and $\xi = (\xi_1, \cdots, \xi_m)$.
- For every k-dimensional tensor A with $(i_1, i_2, \cdots, i_k)$ entry $A_{i_1\cdots i_k} \in M_d(\mathbb{R})$ and every matrix $B \in M_d(\mathbb{R})$, AB denotes the tensor whose $(i_1, i_2, \cdots, i_k)$ entry is $A_{i_1\cdots i_k}B$. BA is also defined in the same way.
- For any partially differentiable function $f : \Theta_2 \to \mathbb{R}^{d_2}$ and $S \in M_{d_2}(\mathbb{R})$, $S[\partial_{\theta_2}^{\otimes 2}]f(\theta)$ is the matrix whose (i, j)-entry is $\frac{\partial}{\partial \theta_2^i} f(\theta_2) S_{ij} \frac{\partial}{\partial \theta_2^j} f(\theta_2)$.
- If both A and B are matrices with $M_d(\mathbb{R})$ entries, AB is the normal product of matrices.
- For every matrix A on $M_d(\mathbb{R})$ with (i, j) entry $A_{ij} \in M_d(\mathbb{R})$, $\mathrm{Tr}A$ is a matrix on $\mathbb{R}$ with (i, j) entry $\mathrm{Tr}A_{ij}$.
- For every stochastic process Z, $\Delta_i Z = Z_{t_i} - Z_{t_{i-1}}$.
- We write a^*, b^*, c^*, σ^* and Σ^* for $a(\theta_2^*), b(\theta_2^*), c(\theta_2^*), \sigma(\theta_1^*)$ and $\Sigma(\theta_1^*)$.
- We omit the subscript n in h_n and just write h when there is no ambiguity.
- We designate $\sigma(\theta_1)\sigma(\theta_1)'$ as $\Sigma(\theta_1)$.
- C denotes a generic positive constant. When C depends on some parameter p, we might use C_p instead of C.

Moreover, we need the following assumptions:

[A1] $nh_n \to \infty,\ nh_n{}^2 \to 0$ as $n \to \infty$. Moreover, we assume $h_n \leq 1$ for every $n \in \mathbb{N}$.
[A2] a, b, c and σ are of class C^4.

Then we can extend a, b, c and σ to continuous functions on $\overline{\Theta}_1$ and $\overline{\Theta}_2$.

[A3]

$$\inf_{\theta_2\in\overline{\Theta}_2} \lambda_{\min}(a(\theta_2)) > 0$$

$$\inf_{\theta_2\in\overline{\Theta}_2} \lambda_{\min}(b(\theta_2)^{\otimes 2}) > 0$$

$$\inf_{\theta_1\in\overline{\Theta}_1} \lambda_{\min}(\Sigma(\theta_1)) > 0.$$

[A4] For any $\theta_1 \in \overline{\Theta}_1$ and $\theta_2 \in \overline{\Theta}_2$, the pair of matrix $(a(\theta_2)', \Sigma(\theta_1)[c(\theta_2)^{\otimes 2}])$ is controllable; i.e. the matrix

$$\left(\Sigma(\theta_1)[c(\theta_2)^{\otimes 2}]\ a(\theta_2)'\Sigma(\theta_1)[c(\theta_2)^{\otimes 2}]\ \cdots\ a(\theta_2)'^{d_1}\Sigma(\theta_1)[c(\theta_2)^{\otimes 2}]\right)$$

has full row rank.
Moreover, the eigenvalues of the matrix

$$H(\theta_1,\theta_2)=\begin{pmatrix} a(\theta_2)' & \Sigma(\theta_1)^{-1}[c(\theta_2)^{\otimes 2}] \\ b(\theta_2)^{\otimes 2} & -a(\theta_2)\end{pmatrix} \tag{2.4}$$

are uniformly bounded away from the imaginary axis; i.e. there are some constant $C>0$ such that for any $\theta_1\in\overline{\Theta}_1$ and $\theta_2\in\overline{\Theta}_2$ and eigenvalue λ of $H(\theta_1,\theta_2)$, it holds

$$|\mathrm{Re}(\lambda)|>C.$$

By Assumption [A4] and the corollary of Theorem 6 in Coppel (1974), for every $\theta_1\in\overline{\Theta}_1$ and $\theta_2\in\overline{\Theta}_2$, Eq. (1.5) has the maximal solution $\gamma=\gamma_+(\theta_1,\theta_2)$ and minimal solution $\gamma=\gamma_-(\theta_1,\theta_2)$, where $\gamma_+(\theta_1,\theta_2)>\gamma_-(\theta_1,\theta_2)$. The meaning of the maximal and minimal solutions is that for any symmetric solution γ of (1.5), it holds $\gamma_-\le\gamma\le\gamma_+$.

Now we define $\mathbb{Y}_1$ and $\mathbb{Y}_2$ by

$$\mathbb{Y}_1(\theta_1)=-\frac{1}{2}\left\{\mathrm{Tr}\,\Sigma(\theta_1)^{-1}\Sigma(\theta_1^*)-d_1+\log\frac{\det\Sigma(\theta_1)}{\det\Sigma(\theta_1^*)}\right\} \tag{2.5}$$

and

$$\begin{aligned}\mathbb{Y}_2(\theta_2)=-\frac{1}{2}\mathrm{Tr}\int_0^\infty \Sigma^{*-1}\Bigg[\Bigg\{\int_0^s & c(\theta_2)\exp(-\alpha(\theta_2)u)\gamma_+(\theta_1^*,\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\\ &\times\exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}du\\ &+c(\theta_2)\exp(-\alpha(\theta_2)s)\gamma_+(\theta_1^*,\theta_2)c(\theta_2)'\\ &-c^*\exp(-a^*s)\gamma_+(\theta^*)c^{*\prime}\Bigg\}^{\otimes 2}\Bigg][(\sigma^{*\prime -1})^{\otimes 2}]ds,\end{aligned} \tag{2.6}$$

respectively, where

$$\alpha(\theta_2)=a(\theta_2)+\gamma_+(\theta_1^*,\theta_2)\Sigma(\theta_1^*)^{-1}[c(\theta_2)^{\otimes 2}], \tag{2.7}$$

and assume the following condition.

[A5] There is some positive constant $C>0$ satisfying

$$\mathbb{Y}_1(\theta_1)\le -C|\theta_1-\theta_1^*|^2 \tag{2.8}$$

and

$$\mathbb{Y}_2(\theta_2)\le -C|\theta_2-\theta_2^*|^2. \tag{2.9}$$

Remark By (2.7), it holds

$$\begin{aligned}&\int_0^s c^*\exp(-\alpha(\theta_2^*)u)\gamma_+(\theta^*)c^{*\prime}\Sigma^{*-1}c^*\exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}du\\ &=\int_0^s c^*\exp(-\alpha(\theta_2^*)u)\left\{\alpha(\theta_2^*)-a^*\right\}\exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}du\\ &=c^*\exp(-\alpha(\theta_2^*)s)-c(\theta_2)\exp(-a^*s)\gamma_+(\theta^*)c^{*\prime},\end{aligned}$$

and therefore $\mathbb{Y}_2(\theta_2)$ has the following expression:

$$\mathbb{Y}_2(\theta_2)=-\frac{1}{2}\mathrm{Tr}\int_0^\infty \Sigma^{*-1}\Bigg[\Bigg\{\int_0^s\{c(\theta_2)\exp(-\alpha(\theta_2)u)\gamma_+(\theta_1^*,\theta_2)c(\theta_2)'\Sigma^{*-1}c^*$$

$$
\begin{aligned}
&- c^* \exp(-\alpha(\theta_2^*)u)\gamma_+(\theta^*)c^{*\prime}\Sigma^{*-1}c^*\}du \\
&+ c(\theta_2)\exp(-\alpha(\theta_2)s)\gamma_+(\theta_1^*, \theta_2)c(\theta_2)' \\
&-c^* \exp(-\alpha(\theta_2^*)s)\gamma_+(\theta^*)c^{*\prime} \Big\}^{\otimes 2}\Big][(\sigma^{*\prime -1})^{\otimes 2}]ds.
\end{aligned}
$$

In particular, we have $\mathbb{Y}_2(\theta_2^*) = 0$.

Under these assumptions above, we set

$$
\mathbb{H}_n^1(\theta_1) = -\frac{1}{2}\sum_{j=1}^n \left\{ \frac{1}{h}\Sigma^{-1}(\theta_1)[(\Delta_j Y)^{\otimes 2}] + \log\det\Sigma(\theta_1) \right\} \tag{2.10}
$$

and

$$
\Gamma^1 = \frac{1}{2}\left[\mathrm{Tr}\{\Sigma^{*-1}\partial_{\theta_1}\Sigma(\theta_1^*)\}\right]^{\otimes 2},
$$

and we define our estimator of θ_1 as the maximizer of $\mathbb{H}_n^1(\theta_1)$. Note that $\mathrm{Tr}\{\Sigma^{*-1}\partial_{\theta_1}\Sigma(\theta_1^*)\}$ is a vector whose jth entry is $\mathrm{Tr}\left\{\Sigma^{*-1}\dfrac{\partial}{\partial\theta_1^j}\Sigma(\theta_1^*)\right\}$. Then the following theorem holds:

Theorem 2.1 *We assume [A1]–[A5], and for each $n \in \mathbb{N}$, let $\hat{\theta}_1^n$ be a random variable satisfying*

$$
\mathbb{H}_n^1(\hat{\theta}_1^n) = \max_{\theta_1 \in \overline{\Theta}_1} \mathbb{H}_n^1(\theta_1).
$$

Then for every $p > 0$ and any continuous function $f : \mathbb{R}^d \to \mathbb{R}$ such that

$$
\limsup_{|x|\to\infty} \frac{|f(x)|}{|x|^p} < \infty,
$$

it holds that

$$
E[f(\sqrt{n}(\hat{\theta}_1^n - \theta_1^*))] \to E[f(Z)] \ (n \to \infty),
$$

where $Z \sim N(0, (\Gamma^1)^{-1})$.

In particular, it holds that

$$
\sqrt{n}(\hat{\theta}_1^n - \theta_1^*) \xrightarrow{d} N(0, (\Gamma^1)^{-1}) \ (n \to \infty).
$$

Next we construct the estimator of θ_2, which is the main object of this article. Recall that under Assumption [A4], (1.5) has the maximal solution $\gamma_+(\theta_1, \theta_2)$. Now we replace γ_t with $\gamma_+(\theta_1, \theta_2)$ in (1.3), and define $m_t(\theta_1, \theta_2; m_0)$ by

$$
\begin{cases}
dm_t = -a(\theta_2)m_t dt + \gamma_+(\theta_1, \theta_2)c(\theta_2)'\{\sigma(\theta_1)\sigma(\theta_1)'\}^{-1}\{dY_t - c(\theta_2)m_t dt\} \\
m_0(\theta_1, \theta_2; m_0) = m_0,
\end{cases} \tag{2.11}
$$

where $m_0 \in \mathbb{R}^{d_1}$ is an arbitrary initial estimated value of X_0.

Due to Itô's formula, the solution of (2.11) can be written as

$$
\begin{aligned}
m_t(\theta_1, \theta_2) &= \exp(-\alpha(\theta_1, \theta_2)t)\, m_0 \\
&\quad + \int_0^t \exp(-\alpha(\theta_1, \theta_2)(t-s))\, \gamma_+(\theta_1, \theta_2)c(\theta_2)'\Sigma(\theta_1)^{-1}dY_s,
\end{aligned} \tag{2.12}
$$

where

$$\alpha(\theta_1, \theta_2) = a(\theta_2) + \gamma_+(\theta_1, \theta_2)\Sigma(\theta_1)^{-1}[c(\theta_2)^{\otimes 2}]. \tag{2.13}$$

The eigenvalues of $\alpha(\theta_1, \theta_2)$ coincides with those of $H(\theta_1, \theta_2)$ in (2.4) with positive real part (see Zhou et al. 1996), so there exists some constant $C > 0$ such that for any $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$,

$$\inf_{\lambda \in \sigma(\alpha(\theta_1, \theta_2))} \mathrm{Re}\lambda > C,$$

where $\sigma(\alpha(\theta_1, \theta_2))$ is the set of all eigenvalues of $\alpha(\theta_1, \theta_2)$.

According to (2.12), we set for $i, n \in \mathbb{N}$,

$$\begin{aligned}\hat{m}_i^n(\theta_2; m_0) &= \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)t_i\right) m_0 \\ &\quad + \sum_{j=1}^{i} \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)(t_i - t_{j-1})\right) \gamma_+(\hat{\theta}_1^n, \theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_j Y, \end{aligned} \tag{2.14}$$

$$\begin{aligned}\mathbb{H}_n^2(\theta_2; m_0) &= \frac{1}{2}\sum_{i=1}^{n}\Big\{-h\Sigma(\hat{\theta}_1^n)^{-1}[(c(\theta_2)\hat{m}_{j-1}^n(\theta_2))^{\otimes 2}] \\ &\quad + \hat{m}_{j-1}^n(\theta_2)'c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_j Y + \Delta_j Y'\Sigma(\hat{\theta}_1^n)^{-1}c(\theta_2)\hat{m}_{j-1}^n(\theta_2)\Big\}, \end{aligned} \tag{2.15}$$

and

$$\begin{aligned}\Gamma^2 &= \mathrm{Tr}\int_0^\infty \Sigma^{*-1}[\partial_{\theta_2}^{\otimes 2}]\Big\{\int_0^s c(\theta_2)\exp(-\alpha(\theta_2)u)\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^* \\ &\quad \times \exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}du \\ &\quad + c(\theta_2)\exp(-\alpha(\theta_2)s)\gamma_+(\theta_2)c(\theta_2)'\Big\}\Big|_{\theta_2=\theta_2^*} ds, \end{aligned} \tag{2.16}$$

where $\hat{\theta}_1^n$ is the estimator of θ_1 defined in Theorem 2.1. Then the following theorem holds:

Theorem 2.2 *We assume [A1]–[A5], and let $m_0 \in \mathbb{R}^{d_1}$ be an arbitrary initial value and $\hat{\theta}_2^n = \hat{\theta}_2^n(m_0)$ be a random variable satisfying*

$$\mathbb{H}_n^2(\hat{\theta}_2^n) = \max_{\theta_2 \in \overline{\Theta}_2} \mathbb{H}_n^2(\theta_2)$$

for each $n \in \mathbb{N}$. Moreover, let Γ^2 be positive definite. Then for any $p > 0$ and continuous function $f : \mathbb{R}^d \to \mathbb{R}$ such that

$$\limsup_{|x|\to\infty} \frac{|f(x)|}{|x|^p} < \infty,$$

it holds that

$$E[f(\sqrt{t_n}(\hat{\theta}_2^n - \theta_2^*))] \to E[f(Z)] \ (n \to \infty),$$

where $Z \sim N(0, (\Gamma^2)^{-1})$.

In particular, it holds that

$$\sqrt{t_n}(\hat{\theta}_2^n - \theta_2^*) \xrightarrow{d} N(0, (\Gamma^2)^{-1}) \ (n \to \infty).$$

Remark (1) In order to calculate $\hat{m}_i^n$, one can use the autoregressive formula

$$\hat{m}_{i+1}^n(\theta_2; m_0) = \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)h\right)\hat{m}_i^n(\theta_2; m_0) + \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)h\right)\gamma_+(\hat{\theta}_1^n, \theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_{i+1}Y.$$

(2) One can obtain $\gamma(\theta_1, \theta_2)$ in the following way (see Zhou et al. 1996 for details). Let $v_1, v_2, \cdots, v_{d_1}$ be generalized eigenvectors of $H(\theta_1, \theta_2)$ in (2.4) with positive real part eigenvalues. Note that $H(\theta_1, \theta_2)$ has d_1 eigenvalues (with multiplicity) in the right half-plane and d_1 in the left half-plane. We define the matrices $X_1(\theta_1, \theta_2)$ and $X_2(\theta_1, \theta_2)$ by

$$\begin{pmatrix} v_1 & v_2 & \cdots & v_{d_1} \end{pmatrix} = \begin{pmatrix} X_1(\theta_1, \theta_2) \\ X_2(\theta_1, \theta_2) \end{pmatrix}.$$

Then $X_1(\theta_1, \theta_2)$ is invertible and it holds $\gamma_+(\theta_1, \theta_2) = X_2(\theta_1, \theta_2)X_1(\theta_1, \theta_2)^{-1}$.

(3) $\mathbb{H}^2(\theta_2)$ can be interpreted as a approximated log-likelihood function with θ_1 given. In fact, if $X_t = X_t(\theta)$ and $Y_t = Y_t(\theta)$ are generated by (1.1) and (1.2), and we set $m_0 = E[X_0|Y_0]$ and $\gamma_0 = E[(m_0 - X_0)^{\otimes 2}]$, then it follows $m_t(\theta) = E[X_t(\theta)|\{Y_s(\theta)\}_{0\le s\le t}]$ by Theorem 1.1. Thus by the innovation theorem (Kallianpur 1980), we can replace $X_t(\theta)$ with $m_t(\theta)$ in Eq. (1.2), and consider the equation

$$dY_t(\theta) = c(\theta_2)m_t(\theta)dt + \sigma(\theta_1)d\overline{W}_t$$

where $\overline{W}$ is a d_2-dimensional Wiener process. We can approximate this equation as

$$\Delta_i Y(\theta) \approx c(\theta_2)m_{t_{i-1}}(\theta)h + \sigma(\theta_1)\Delta_i\overline{W},$$

when $h \approx 0$. Then we obtain the approximated likelihood function

$$p(\theta) \approx \prod_{i=1}^n \frac{1}{(2\pi h)^{\frac{d}{2}}\{\det\Sigma(\theta_1)\}^{\frac{1}{2}}} \times \exp\left(-\frac{1}{2h}\Sigma(\theta_1)^{-1}\left[(\Delta_i Y - c(\theta_2)m_{t_{i-1}}(\theta)h)^{\otimes 2}\right]\right).$$

(4) The condition the W^1 and W^2 are independent is not essential; according to Section 12 of Liptser and Shiriaev (2001b), Kalman–Bucy filter can be extended to the equation of the form

$$dX_t = \{a_0(\theta_2, Y_t) + a_1(\theta_2, Y_t)X_t\}dt + b_1(\theta_2, Y_t)dW_t^1 + b_2(\theta_2, Y_t)dW_t^2$$
$$dY_t = \{c_0(\theta_2, Y_t) + c_1(\theta_2, Y_t)X_t\}dt + \sigma(\theta_1, Y_t)dW_t^2.$$

However, this case is more complicated, and thus is left for future research.

3 Proof of Theorem 2.2

In this section, we write $m_t(\theta_2)$, $\hat{m}_i^n(\theta_2)$, $\mathbb{H}_n^2(\theta_2)$, $\gamma_+(\theta_2)$ and $\alpha(\theta_2)$ instead of $m_t(\theta_1^*, \theta_2; m_0)$, $\hat{m}_i^n(\theta_2; m_0)$, $\mathbb{H}_n^2(\theta_2; m_0)$, $\gamma_+(\theta_1^*, \theta_2)$ and $\alpha(\theta_1^*, \theta_2)$, respectively, for simplicity.

Moreover, let $m_t^* = E[X_t|\{Y_t\}_{0\le s\le t}]$ and $\gamma_t^* = E[(X_t - m_t)(X_t - m_t)']$. Then by Theorem 1.1, they are the solutions of

$$dm_t^* = -a^* m_t^* dt + \gamma_t^* c^{*\prime} \Sigma^{*-1}\{dY_t - c^* m_t^* dt\} \tag{3.1}$$

$$\frac{d\gamma_t^*}{dt} = -a^*\gamma_t^* - \gamma_t^*(a^*)' - \Sigma^{*-1}[(c^*\gamma_t^*)^{\otimes 2}] + b^{*\otimes 2}. \tag{3.2}$$

We start with preliminary lemmas, which is frequently referred to in proving inequalities.

Lemma 3.1 *Let $\{W_t\}$ be a d-dimensional $\{\mathcal{F}_t\}$-Wiener process.*

(1) *Let $f : \mathbb{R}_+ \to \mathbb{R}^m$ be a measurable function. Then for any $p \ge 1$ and $0 \le s \le t$, it holds*

$$\left(\int_s^t |f(u)|du\right)^p \le (t-s)^{p-1}\int_s^t |f(u)|^p du$$

(2) *Let $\{A_t\}$ be a $M_{k,d}(\mathbb{R})$-valued progressively measurable process and $\{W_t\}$ be a d-dimensional Wiener process. Then for every $0 \le s \le t \le T$ and $p \ge 2$, it holds*

$$E\left[\sup_{s\le t\le T}\left|\int_s^t A_u dW_u\right|^p\right] \le C_{p,d,k} E\left[\left(\int_s^T |A_u|^2 du\right)^{\frac{p}{2}}\right] \le C_{p,d,k}(T-s)^{\frac{p}{2}-1}\int_s^t E[|A_u|^p]du.$$

Proof (1) By Hölder's inequality, we obtain

$$\int_s^t |f(u)|du \le \left(\int_s^t |f(u)|^p du\right)^{\frac{1}{p}}\left(\int_s^t du\right)^{1-\frac{1}{p}} = (t-s)^{1-\frac{1}{p}}\left(\int_s^t |f(u)|^p du\right)^{\frac{1}{p}},$$

and it shows the desired inequality.

(2) Let $A_t^{(ij)}$ be the (i,j) entry of A_t, and $W_t^{(j)}$ be the jth element of W_t. Then the Burkholder-Davis-Gundy inequality gives

$$\begin{aligned}
E\left[\sup_{s\le t\le T}\left|\int_s^t A_u dW_u\right|^p\right] &= E\left[\sup_{s\le t\le T}\left\{\sum_{i=1}^k\left(\sum_{j=1}^d\int_s^t A_u^{(ij)}dW_u^{(j)}\right)^2\right\}^{\frac{p}{2}}\right] \\
&\le C_{p,d,k}\sum_{i=1}^k\sum_{j=1}^d E\left[\sup_{s\le t\le T}\left|\int_s^t A_u^{(ij)}dW_u^{(j)}\right|^p\right] \\
&\le C_{p,d,k}\sum_{i=1}^k\sum_{j=1}^d E\left[\left|\int_s^T (A_u^{(ij)})^2 du\right|^{\frac{p}{2}}\right] \\
&\le C_{p,d,k} E\left[\left|\int_s^T \sum_{i=1}^k\sum_{j=1}^d (A_u^{(ij)})^2 du\right|^{\frac{p}{2}}\right]
\end{aligned}$$

$$= C_{p,d,k} E\left[\left|\int_s^T |A_u|^2 du\right|^{\frac{p}{2}}\right].$$

Hence we have proved the first inequality, and together with (1) we obtain the second one.

□

Lemma 3.2 *Let A be a $d \times d$ matrix having eigenvalues $\lambda_1, \cdots, \lambda_k$. Then for all $\epsilon > 0$, there exists some constant $C_{\epsilon,d}$ depending on ϵ and d such that*

$$|\exp(At)| \leq C_{\epsilon,d}(1 + |A|^{d-1})e^{(\lambda_{\max}+\epsilon)t} \ (t \geq 0),$$

where

$$\lambda_{\max} = \max_{i=1,\cdots,k} \mathrm{Re}\lambda_k.$$

Proof Let

$$A = U^*(D + N)U, \ D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_d)$$

be a Schur decomposition of A, where $\lambda_1, \lambda_2, \cdots, \lambda_d$ are the eigenvalues of A, U is an unitary matrix, and N is a strictly upper triangular matrix. Then we have

$$\begin{aligned}
|\exp(At)| &= |\exp((D + N)t)| = |\exp(Dt)\exp(Nt)| \\
&\leq |\exp(Dt)||\exp(Nt)| \\
&\leq C_d e^{\lambda_{\max} t} \sum_{k=1}^{d-1} \frac{|N|^k}{k!} t^k \\
&\leq C_d e^{\lambda_{\max} t} \sum_{k=1}^{d-1} \frac{|A|^k}{k!} t^k \\
&\leq C_d e^{(\lambda_{\max}+\epsilon)t} \sum_{k=1}^{d-1} \frac{|A|^k}{k!} t^k e^{-\epsilon t} \\
&\leq C_{\epsilon,d}(1 + |A|^{d-1})e^{(\lambda_{\max}+\epsilon)t},
\end{aligned}$$

noting that U is unitary, $N^d = O$, and $|A| = |D + N| \geq N$. □

Lemma 3.3 *For any $s, t \geq 0$ such that $0 \leq t - s \leq 1$ and $p \geq 1$, it holds*

$$\sup_{t \geq 0} E[|X_t|^p] \leq C_p, \tag{3.3}$$

$$E[|Y_s - Y_t|^p] \leq C_p|s - t|^{\frac{p}{2}} \tag{3.4}$$

and

$$E[|X_s - X_t|^p] \leq C_p|s - t|^{\frac{p}{2}}. \tag{3.5}$$

Proof By Itô's formula, the solution of (2.1) can be expressed as

$$X_t = \exp(-a^*t)X_0 + \int_0^t \exp(-a^*(t - s))b^* dW_s^1, \tag{3.6}$$

where exp is the matrix exponential. Hence by Lemmas 3.1 and 3.2, we have

$$
\begin{aligned}
E[|X_t|^p] &\leq |\exp(-a^*t)|^p E[|X_0|^p] \\
&\quad + C_p \left(\int_0^t |b^*|^2 |\exp(-a^*(t-s))|^2 ds \right)^{\frac{p}{2}} \\
&\leq C_p e^{-\eta p t} + C_p \left(\int_0^t e^{-2\eta s} ds \right)^{\frac{p}{2}} \leq C_p
\end{aligned}
$$

for some constant $\eta > 0$. Therefore for $s \leq t$ we obtain

$$
\begin{aligned}
E[|Y_t - Y_s|^p] &= E\left[\left| c^* \int_s^t X_u du + \sigma^*(W_t^2 - W_s^2) \right|^p \right] \\
&\leq C_p \left((t-s)^{p-1} \int_s^t E[|X_u|^p] du + (t-s)^{\frac{p}{2}} \right) \\
&\leq C_p \left((t-s)^p + (t-s)^{\frac{p}{2}} \right) \\
&\leq C_p (t-s)^{\frac{p}{2}}.
\end{aligned}
$$

We can show (3.5) in the same way. □

We next discuss important properties of $\gamma_+(\theta_1, \theta_2)$ and γ_t^*.

Proposition 3.4 *The maximal solution of* (1.5) $\gamma_+(\theta_1, \theta_2)$ *is of class* C^4.

Proof Let $\theta^0 = (\theta_1^0, \theta_2^0) \in \Theta_1 \times \Theta_2$, and we consider the mapping $f : M_{d_1}(\mathbb{R}) \to M_{d_1}(\mathbb{R})$ such that

$$
f : X \mapsto a(\theta_2^0)X + Xa(\theta_2^0)' + \Sigma(\theta_1^0)^{-1}[(c(\theta_2^0)X)^{\otimes 2}] - b(\theta_2^0)^{\otimes 2}.
$$

Since for every $T \in M_{d_1}(\mathbb{R})$, we have

$$
\begin{aligned}
f(X+T) - f(T) &= \left\{ a(\theta_2^0) + X'\Sigma(\theta_1^0)^{-1}[c(\theta_2^0)^{\otimes 2}] \right\} T \\
&\quad + T \left\{ a(\theta_2^0)' + \Sigma(\theta_1)^{-1}[c(\theta_2^0)^{\otimes 2}]X \right\} \\
&\quad + \Sigma(\theta_1^0)^{-1}[(c(\theta_2^0)T)^{\otimes 2}]
\end{aligned}
$$

and

$$
\lim_{|T| \to 0} \frac{|\Sigma(\theta_1^0)^{-1}[(c(\theta_2^0)T)^{\otimes 2}]|}{|T|} = 0,
$$

the differential of f at $X = \gamma_+(\theta^0)$ is given by

$$
(df)_{\gamma_+(\theta^0)} : T \mapsto \alpha(\theta_0)T + T\alpha(\theta_0),
$$

where α is defined by (2.13).

If $(df)_{\gamma_+(\theta^0)}$ is not injective, $\alpha(\theta_0)$ has eigenvalues μ and λ such that $\mu + \bar{\lambda} = 0$ (see lemma 2.7 in Zhou et al. 1996). However, noting that $\gamma_+(\theta_1, \theta_2)$ is the unique symmetric solution of $f(X) = O$ such that $-\alpha(\theta_0)$ is stable (Coppel 1974; Zhou et al. 1996), there are no such eigenvalues. Therefore $(df)_{\gamma_+(\theta^0)}$ is injective, and by the implicit function theorem, there exists a neighborhood $U \subset \Theta_1 \times \Theta_2$ containing θ^0 and a mapping $\phi : U \to M_{d_1}(\mathbb{R})$ of class C^4 such that

$$
\phi(\theta^0) = \gamma_+(\theta^0), \quad f(\phi(\theta)) = O \ (\theta \in U).
$$

Since $-a(\theta_2) - \phi(\theta)\Sigma(\theta_1)^{-1}[c(\theta_2)^{\otimes 2}]$ is stable at $\theta = (\theta_1, \theta_2) = \theta^0$, it is also stable on a neighborhood of θ^0. Thus by the uniqueness of γ_+, we obtain $\gamma_+(\theta) = \phi(\theta)$ on that neighborhood and therefore the desired result. □

By this proposition, Theorem 2.1 and the mean value theorem, we get the following corollary.

Corollary 3.5 *For any $p \geq 1$, it holds*

$$E\left[\sup_{\theta_2 \in \Theta_2} |\gamma_+(\hat{\theta}_1^n, \theta_2) - \gamma_+(\theta_1^*, \theta_2)|^p\right]^{\frac{1}{p}} \leq Cn^{-\frac{1}{2}}$$

and

$$E\left[\sup_{\theta_2 \in \Theta_2} |\alpha(\hat{\theta}_1^n, \theta_2) - \alpha(\theta_1^*, \theta_2)|^p\right]^{\frac{1}{p}} \leq Cn^{-\frac{1}{2}}.$$

Proposition 3.6 *For every $\theta_1 \in \overline{\Theta}_1$ and $\theta_2 \in \overline{\Theta}_2$,*

$$\gamma_+(\theta_1, \theta_2) > 0 \tag{3.7}$$

and

$$\gamma_-(\theta_1, \theta_2) < 0. \tag{3.8}$$

Proof Noting that for A and $\gamma \in M_{d_1}(\mathbb{R})$,

$$\frac{d}{dt}(\exp(At)\gamma \exp(A't)) = \exp(At)(A\gamma + \gamma A')\exp(A't),$$

and the Eq. (1.5) is equivalent to

$$\begin{aligned}&\left\{a(\theta_2) + \gamma\Sigma(\theta_1)^{-1}[c(\theta_2)^{\otimes 2}]\right\}\gamma + \gamma\left\{a(\theta_2) + \gamma\Sigma(\theta_1)^{-1}[c(\theta_2)^{\otimes 2}]\right\}' \\ &\quad = \gamma\Sigma(\theta_1)^{-1}[c(\theta_2)^{\otimes 2}]\gamma + b(\theta_2)^{\otimes 2},\end{aligned}$$

we obtain

$$\begin{aligned}&\gamma_+(\theta_1, \theta_2) \\ &\quad = \int_{-\infty}^0 \exp(\alpha(\theta_1, \theta_2)t)\{\alpha(\theta_1, \theta_2)\gamma + \gamma\alpha(\theta_1, \theta_2)'\}\exp(\alpha(\theta_1, \theta_2)'t)dt \\ &\quad = \int_{-\infty}^0 \exp(\alpha(\theta_1, \theta_2)t)\left\{\Sigma(\theta_1)^{-1}[c(\theta_2)^{\otimes 2}][\gamma_+(\theta_1, \theta_2)^{\otimes 2}] + b(\theta_2)^{\otimes 2}\right\} \\ &\qquad \times \exp(\alpha(\theta_1, \theta_2)t)dt > 0\end{aligned}$$

by assumption [A3], (2.13) and the stability of $-\alpha(\theta_1, \theta_2)$. In the same way, we can show $\gamma_-(\theta_1, \theta_2) < 0$. □

Combining this result with assumption [A3], (2.13) and Lemma 3.2, we obtain the following corollary.

Corollary 3.7 *There exists some constant $C_1 > 0$ and $C_2 > 0$ such that*

$$\sup_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2} |\exp(-\alpha(\theta_1, \theta_2))| \leq C_1 e^{-C_2 t}.$$

Now we go on to the convergence of γ_t^*. Concerning the convergence rate of Riccati equations, Leipnik (1985) presents the following result.

Theorem 3.8 (Section 5, Leipnik 1985) *Let $A, B, C \in M_d(\mathbb{R})$ and consider the equation*

$$\frac{d}{dP}(t) = -A - P(t)B - B'P(t) - P(t)CP(t).$$

Moreover, assume C is symmetric, $C \leq 0$, (B, C) is controllable and the matrix

$$H = \begin{pmatrix} B & C \\ -A & -B' \end{pmatrix}$$

has no pure imaginary eigenvalues.
Then if $P_0 - P^+$ is non-singular, then it holds for any $\epsilon > 0$ that

$$|P(t) - P^-| \leq Ce^{2(r+\epsilon)t} \ (t \to \infty)$$

and if $P_0 - P^-$ is non-singular, then it holds for any $\epsilon > 0$ that

$$|P(t) - P^+| \leq Ce^{-2(r-\epsilon)t} \ (t \to -\infty),$$

where P^+ and P^- are the maximal and minimal solutions of the algebraic Riccati equation

$$A + PB + B'P + PCP = O$$

respectively, $r < 0$ is the maximum real part of the eigenvalues of $B + CP^+$.

Proposition 3.9 *For any $\epsilon > 0$, there exists some constant $C > 0$ such that*

$$|\gamma_t^* - \gamma_+(\theta^*)| \leq Ce^{-2\{\lambda_{\min}(\alpha(\theta_2^*))-\epsilon\}t}.$$

In particular, $|\gamma_t^*|$ is bounded.

Proof According to (3.2) and Theorem 3.8, it is enough show that $\gamma_0^* - \gamma_-(\theta^*)$ is non-singular, where $\gamma_-(\theta_1, \theta_2)$ is the minimal solution of (1.5). If we assume $\gamma_0^* - \gamma_-(\theta^*)$ is singular, there exists $x \in \mathbb{R}^{d_1}\backslash\{0\}$ such that $\{\gamma_0^* - \gamma_-(\theta^*)\}x = 0$, and we get $x\gamma_0^* x = x\gamma_-(\theta^*)x$. However, since $\gamma_0^* \geq 0$ and we have $\gamma_-(\theta^*) < 0$ by Proposition 3.6, that is a contradiction. □

Next we consider the innovation process

$$\overline{W}_t = (\sigma^*)^{-1}\left(Y_t - \int_0^t c^* m_s^* ds\right).$$

Note that the right-hand side is well-defined since $\{m_t^*\}$ has a progressively measurable modification, and that $\overline{W}_t$ is also a Wiener process (Kallianpur 1980). Since Y_t is the solution of

$$dY_t = c^* m_t^* dt + \sigma^* d\overline{W}_t, \tag{3.9}$$

we obtain together with (3.1)

$$dm_t^* = -a^* m_t^* dt + \gamma_t^* c^{*\prime} \sigma^{*\prime -1} d\overline{W}_t.$$

Therefore Itô's formula gives

$$m_t^* = \exp(-a^* t)m_0^* + \int_0^t \exp(-a^*(t-s))\gamma_s^* c^{*\prime} \sigma^{*\prime -1} d\overline{W}_s. \tag{3.10}$$

Moreover, using Proposition 3.9, we can show for any $p \geq 1$,

$$\sup_{t\geq 0} E[|m_t^*|^p] \leq C_p \tag{3.11}$$

and

$$\sup_{0\leq t-s\leq 1} E[|m_t^* - m_s^*|^p] \leq C_p (t-s)^{\frac{p}{2}} \tag{3.12}$$

in the same way as Lemma 3.3.

Lemma 3.10 *For $j = 0, 1, 2, \cdots$ and $\theta \in \Theta$, let $Z_j(\theta)$ be a $M_{k,l}(\mathbb{R})$-valued and $\mathcal{F}_{t_j}$-measurable random variable, and $U(\theta)$ be an $M_{l,d}(\mathbb{R})$-valued random variable. Moreover, we assume $Z_j(\theta)$ is continuously differentiable with respect to θ. Then for any $n \in \mathbb{N}$ and $p > m_1 + m_2$, it holds*

$$\begin{aligned} &E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^n Z_{j-1}(\theta)U(\theta)\Delta_j W\right|^p\right] \\ &\leq C_{d,k,l} E\left[\sup_{\theta\in\Theta}|U(\theta)|^{2p}\right]^{\frac{1}{2}} \\ &\quad \times \sup_{\theta\in\Theta}\left\{E\left[\left\{\sum_{j=1}^n |Z_{j-1}(\theta)|^2 h\right\}^p\right] + E\left[\left\{\sum_{j=1}^n \left|\partial_\theta Z_{j-1}(\theta)\right|^2 h\right\}^p\right]\right\}^{\frac{1}{2}}. \end{aligned}$$

Proof Let $Z_j^{(ij)}$, $U^{(ij)}$ and $(Z_jU)^{(ij)}$ be the (i, j) entries of Z_j, U and Z_jU, respectively, and $W^{(j)}$ be the jth element of $W^{(j)}$. Then we have

$$\begin{aligned} &E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^n Z_{j-1}(\theta)U(\theta)\Delta_j W\right|^p\right] \\ &= E\left[\sup_{\theta\in\Theta}\left\{\sum_{p=1}^k\left(\sum_{j=1}^n\sum_{q=1}^d\sum_{r=1}^l Z_{j-1}^{(pr)}(\theta)U^{(rq)}(\theta)\Delta_j W^{(q)}\right)^2\right\}^{\frac{p}{2}}\right] \\ &\leq C_{d,k,l} E\left[\sup_{\theta\in\Theta}|U(\theta)|^{2p}\right]^{\frac{1}{2}} \sum_{p=1}^k\sum_{r=1}^l E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^n Z_{j-1}^{(pr)}(\theta)\Delta_j W^{(q)}\right|^{2p}\right]^{\frac{1}{2}}. \end{aligned} \tag{3.13}$$

Moreover, the Sobolev inequality and the Burkholder-Davis-Gundy inequality gives

$$\begin{aligned} &E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^n Z_{j-1}^{(pr)}(\theta)\Delta_j W^{(q)}\right|^{2p}\right] \\ &\leq C_p \sup_{\theta\in\Theta}\left\{E\left[\left|\sum_{j=1}^n Z_{j-1}^{(pr)}(\theta)\Delta_j W^{(q)}\right|^{2p}\right] + E\left[\left|\sum_{j=1}^n \frac{\partial}{\partial\theta} Z_{j-1}^{(pr)}(\theta)\Delta_j W^{(q)}\right|^{2p}\right]\right\} \end{aligned}$$

$$\leq C_p \sup_{\theta\in\Theta}\left\{E\left[\left|\sum_{j=1}^n Z_{j-1}^{(pr)}(\theta)^2 h\right|^p\right] + E\left[\left|\sum_{j=1}^n \left\{\frac{\partial}{\partial\theta} Z_{j-1}^{(pr)}(\theta)\right\}^2 h\right|^p\right]\right\}$$
$$\leq C_p \sup_{\theta\in\Theta}\left\{E\left[\left(\sum_{j=1}^n |Z_{j-1}(\theta)|^2 h\right)^p\right] + E\left[\left(\sum_{j=1}^n \left|\partial_\theta Z_{j-1}(\theta)\right|^2 h\right)^p\right]\right\}. \tag{3.14}$$

By (3.13) and (3.14), we obtain the desired result. □

Lemma 3.11 *For every $\theta \in \Theta$, let $\{Z_t(\theta)\}$ be a $\mathbb{M}_{d,d_1}(\mathbb{R})$-valued progressively measurable process. Moreover, we assume $Z_t(\theta)$ is differentiable with respect to θ, and for any $T > 0$, $p > 0$ and $\theta, \theta' \in \Theta$*

$$\sup_{0\leq t\leq T} E\left[|Z_t(\theta) - Z_t(\theta')|^p\right] \leq C_{T,p}|\theta-\theta'|^p,$$
$$\sup_{0\leq t\leq T} E\left[|\partial_\theta Z_t(\theta) - \partial_\theta Z_t(\theta')|^p\right] \leq C_{T,p}|\theta-\theta'|^p.$$

Then $\{\xi_\cdot(\theta)\}_{\theta\in\Theta}$ with $\xi_t(\theta) = \int_0^t Z_t(\theta)d\overline{W}_s$ has a modification $\{\tilde{\xi}_\cdot(\theta)\}_{\theta\in\Theta}$ which is continuously differentiable with respect to θ. Moreover, it holds almost surely for any $t \geq 0$ and $\theta \in \Theta$

$$\partial_\theta \tilde{\xi}_t(\theta) = \int_0^t \partial_\theta Z_t(\theta)d\overline{W}_s.$$

Proof For any matrix valued function ϕ on $\mathbb{R}^{m_1+m_2}$ and $\epsilon > 0$, let

$$\Delta^j\phi(\theta;\epsilon) = \frac{1}{\epsilon}\{\xi_t(\theta+\epsilon e_j) - \xi_t(\theta)\},$$

where $e_1, \cdots, e_{m_1+m_2}$ is the standard basis of $\mathbb{R}^{m_1+m_2}$. Then for $\theta, \theta' \in \Theta, \epsilon, \epsilon' > 0$ and $p \geq 1$, we have

$$\sup_{0\leq t\leq T} E\left[\left|\Delta^j Z_t(\theta;\epsilon) - \Delta^j Z_t(\theta;\epsilon')\right|^p\right]$$
$$= \sup_{0\leq t\leq T} E\left[\left|\int_0^1 \frac{\partial}{\partial\theta^j} Z_t(\theta+u\epsilon e_j)du - \int_0^1 \frac{\partial}{\partial\theta^j} Z_s(\theta+u\epsilon' e_j)du\right|^p\right]$$
$$\leq \int_0^1 \sup_{0\leq t\leq T} E\left[\left|\frac{\partial}{\partial\theta^j} Z_t(\theta+u\epsilon e_j) - \frac{\partial}{\partial\theta^j} Z_t(\theta'+u\epsilon' e_j)\right|\right]du$$
$$\leq C_{p,T}(|\theta-\theta'| + |\epsilon-\epsilon'|),$$

where $\theta = (\theta^1, \cdots, \theta^{m_1+m_2})$.

Hence by Lemma 3.1, it follows for any $\theta, \theta' \in \Theta, \epsilon, \epsilon' > 0$ and $N \in \mathbb{N}$

$$E\left[\sup_{0\leq t\leq N}\left|\Delta^j\xi_t(\theta;\epsilon) - \Delta^j\xi_t(\theta';\epsilon')\right|^p\right]$$
$$= E\left[\sup_{0\leq t\leq N}\left|\int_0^t \{\Delta^j Z_t(\theta;\epsilon) - \Delta^j Z_t(\theta';\epsilon')\}d\overline{W}_s\right|^p\right]$$

$$\leq C_p N^{\frac{p}{2}-1} \int_0^N E\left[\left|\Delta^j Z_t(\theta;\epsilon) - \Delta^j Z_t(\theta';\epsilon')\right|^p\right] ds$$
$$\leq C_{p,N}(|\theta-\theta'| + |\epsilon-\epsilon'|).$$

Now for this $C_{p,N}$, we take a sequence $\alpha_N > 0$ ($N \in \mathbb{N}$) so that

$$S_p = \sum_{n=1}^{\infty} \alpha_N C_{p,N} < \infty, \quad \sum_{n=1}^{\infty} \alpha_N < \infty,$$

and define the norm on $C(\mathbb{R}_+; M_{d,d_1}(\mathbb{R}))$ by

$$\|A\| = \sum_{N=1}^{\infty} \alpha_N \left(\sup_{0\leq t\leq N} |A(s)| \wedge 1 \right).$$

Then the topology induced by this norm is equivalent to the topology of uniform convergence, and we have

$$E\left[\left\|\Delta^j \xi.(\theta;\epsilon) - \Delta^j \xi.(\theta';\epsilon')\right\|^p\right] \leq C_p(|\theta-\theta'| + |\epsilon-\epsilon'|). \tag{3.15}$$

Therefore, by the the Kolmogorov continuity theorem, $\{\Delta^j \xi.(\theta;\epsilon)\}_{\theta\in\Theta,0<|\epsilon|\leq 1}$ has a uniformly continuous modification $\{\zeta.(\theta;\epsilon)\}_{\theta\in\Theta,0<|\epsilon|\leq 1}$. Because of the uniform continuity, $\zeta.(\theta;\epsilon)$ can be extended to a continuous process on $\theta \in \Theta$, $|\epsilon| \leq 1$.

On the other hand, we can show in the same way that $\{\xi.(\theta;\epsilon)\}_{\theta\in\Theta}$ has a continuous modification $\{\tilde{\xi}.(\theta;\epsilon)\}_{\theta\in\Theta}$. Then $\Delta^j \tilde{\xi}.(\theta;\epsilon)$ and $\zeta.(\theta;\epsilon)$ are both continuous modifications of $\Delta^j \xi.(\theta;\epsilon)$, and thus they are indistinguishable. Therefore almost surely for any $t \geq 0$ and $\theta \in \Theta$,

$$\frac{\partial \tilde{\xi}_t}{\partial \theta_j}(\theta) = \lim_{\epsilon\to 0} \frac{\xi_t(\theta+\epsilon e_j) - \xi(\theta)}{\epsilon} = \lim_{\epsilon\to 0} \Delta^j \xi_t(\theta;\epsilon)$$

exists. The continuity of $\dfrac{\partial \xi_t}{\partial \theta_j}(\theta)$ follows from the continuity of $\zeta.(\theta,\epsilon)$.

Moreover, by the assumption and Lemma 3.1 (2), we have for $p \geq 2$,

$$E\left[\left|\int_0^t \left\{\frac{1}{\epsilon}(Z_s(\theta+\epsilon e_j) - Z_s(\theta)) - \frac{\partial}{\partial \theta^j} Z_s(\theta)\right\} d\overline{W}_s\right|^p\right]$$
$$= E\left[\left|\int_0^t \left\{\frac{\partial Z_s}{\partial \theta_j}(\theta+\eta_s \epsilon e_j) - \frac{\partial}{\partial \theta^j} Z_s(\theta)\right\} d\overline{W}_s\right|^p\right]$$
$$\leq C_p t^{\frac{p}{2}-1} \int_0^t E\left[\left|\frac{\partial Z_s}{\partial \theta_j}(\theta+\eta_s \epsilon e_j) - \frac{\partial Z_s}{\partial \theta^j}(\theta)\right|^p\right] ds$$
$$\leq C_{p,t}\epsilon \to 0 \ (\epsilon \to 0),$$

where $0 \leq \eta_s \leq 1$. This means

$$\Delta^j \xi_t(\theta;\epsilon) \to \int_0^s \frac{\partial}{\partial \theta^j} Z_s(\theta) ds \ \ (\epsilon \to 0)$$

in L^p, and hence there exists a subsequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ such that $\epsilon_n \to 0$ and

$$\Delta^j \xi_t(\theta;\epsilon_n) \xrightarrow{\text{a.s.}} \int_0^s \frac{\partial}{\partial \theta^j} Z_s(\theta) ds \ \ (n \to \infty).$$

Therefore we obtain almost surely

$$\frac{\partial}{\partial\theta^j}\tilde{\xi}_t(\theta) = \Delta^j\tilde{\xi}_t(\theta; 0) = \int_0^s \frac{\partial}{\partial\theta^j}Z_s(\theta)ds.$$

□

Lemma 3.12 (1) *For $j \in \mathbb{N}$, let $f_j : [t_{j-1}, t_j] \times \Theta \to M_{k,d_2}(\mathbb{R})$ be of class C^1. Then for any $p > m_1 + m_2$, it holds*

$$\begin{aligned} &E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)dY_s\right|^p\right]^{\frac{1}{p}} \\ &\le C_p\sup_{\theta\in\Theta}\left\{\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|f_j(s,\theta)|ds + \sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\partial_\theta f_j(s,\theta)|ds\right. \\ &\quad\left. + \left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|f_{j-1}(s,\theta)|^2ds\right)^{\frac{1}{2}} + \left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\partial_\theta f_{j-1}(s,\theta)|^2ds\right)^{\frac{1}{2}}\right\}, \end{aligned} \tag{3.16}$$

where C_p is a constant which depends only on p.

(2) *For $j = 0, 1, 2, \cdots$ and $\theta \in \Theta$, let $Z_j(\theta)$ be a $M_{k,l}(\mathbb{R})$-valued and $\mathcal{F}_{t_j}$-measurable random variable, and $U(\theta)$ be an $M_{l,d}(\mathbb{R})$-valued random variable. Moreover, we assume $Z_j(\theta)$ is continuously differentiable with respect to θ. Then for any $p > m_1 + m_2$, it holds*

$$\begin{aligned} &E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^{i}Z_{j-1}(\theta)U(\theta)\Delta_jY\right|^p\right]^{\frac{1}{p}} \\ &\le C_pE\left[\sup_{\theta\in\Theta}|U(\theta)|^{4p}\right]^{\frac{1}{4p}} \\ &\quad\times\sup_{\theta\in\Theta}\left\{\sum_{j=1}^{i}E\left[|Z_{j-1}(\theta)|^{2p}\right]^{\frac{1}{2p}}h + \sum_{j=1}^{i}E\left[|\partial_\theta Z_{j-1}(\theta)|^{2p}\right]^{\frac{1}{2p}}h\right\} \\ &\quad + C_pE\left[\sup_{\theta\in\Theta}|U(\theta)|^{2p}\right]^{\frac{1}{2p}} \\ &\quad\times\sup_{\theta\in\Theta}\left\{\sum_{j=1}^{n}E[|Z_{j-1}(\theta)|^{2p}]^{\frac{1}{p}}h + \sum_{j=1}^{n}E[\left|\partial_\theta Z_{j-1}(\theta)\right|^{2p}]^{\frac{1}{p}}h\right\}^{\frac{1}{2}}. \end{aligned} \tag{3.17}$$

Proof (1) By Lemma 3.11, we can assume for every j

$$\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)dY_s = \int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)c^*m_s^*ds + \int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)\sigma^*\overline{W}_s$$

is continuously differentiable, and

$$\partial_\theta\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)dY_s = \int_{t_{j-1}}^{t_j}\partial_\theta f_{j-1}(s,\theta)dY_s.$$

Therefore by the Sobolev inequality and (3.9),

$$
\begin{aligned}
&E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)dY_s\right|^p\right]^{\frac{1}{p}}\\
&\le C_p\sup_{\theta\in\Theta}\left(E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)dY_s\right|^p\right]^{\frac{1}{p}}\right.\\
&\quad\left.+E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} \partial_\theta f_{j-1}(s,\theta)dY_s\right|^p\right]^{\frac{1}{p}}\right)\\
&\le C_p\sup_{\theta\in\Theta}\left(E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)c^*m_s^*ds\right|^p\right]^{\frac{1}{p}}\right.\\
&\quad+E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)\sigma^*d\overline{W}_s\right|^p\right]^{\frac{1}{p}}\\
&\quad+E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} \partial_\theta f_{j-1}(s,\theta)c^*m_s^*ds\right|^p\right]^{\frac{1}{p}}\\
&\quad\left.+E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} \partial_\theta f_{j-1}(s,\theta)\sigma^*d\overline{W}_s\right|^p\right]^{\frac{1}{p}}\right). \qquad (3.18)
\end{aligned}
$$

In order to bound the first term of (3.18), we set

$$
f(s,\theta)=\sum_{j=1}^{i} f_j(s,\theta)1_{(t_{j-1},t_j]}(s).
$$

Then we have

$$
\begin{aligned}
&E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)m_s^*ds\right|^p\right]=E\left[\left|\int_0^{t_i} f(s,\theta)c^*m_s^*ds\right|^p\right]\\
&\le E\left[\left(\int_0^{t_i}|f(s,\theta)c^*m_s^*|ds\right)^p\right]\\
&\le |c^*|^pE\left[\left|\left(\int_0^{t_i}|f(s,\theta)|^{\frac{1}{p}}|m_s^*|^pds\right)^{\frac{1}{p}}\left(\int_0^{t_i}|f(s,\theta)|ds\right)^{1-\frac{1}{p}}\right|^p\right]\\
&\le |c^*|^p\left(\int_0^{t_i}|f(s,\theta)|ds\right)^{p-1}\int_0^{t_i}|f(s,\theta)|E[|m_s^*|^p]ds\\
&\le C_p\left(\int_0^{t_i}|f(s,\theta)|ds\right)^p=C_p\left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|f_j(s,\theta)|ds\right)^p.
\end{aligned}
$$

In the same way, it holds for the third term

$$E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}\partial_\theta Z_{j-1}(s,\theta)c^* m_s^* ds\right|^p\right]^{\frac{1}{p}} \le C_p\left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\partial_\theta f_j(s,\theta)|ds\right)^p.$$

Next by Lemma 3.1 (2), we obtain for the second term

$$E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} f_{j-1}(s,\theta)\sigma^* d\overline{W}_s\right|^p\right] = E\left[\left|\int_0^{t_i} f(s,\theta)\sigma^* d\overline{W}_s\right|^p\right]^{\frac{1}{p}}$$

$$\le C_p\left(\int_0^{t_i}\sum_{j=1}^{i}|f(s,\theta)|^2 ds\right)^{\frac{1}{2}} = C_p\left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|f_{j-1}(s,\theta)|^2 ds\right)^{\frac{1}{2}}$$

and in the same way it holds for the fourth term

$$E\left[\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}\partial_\theta f_{j-1}(s,\theta)d\overline{W}_s\right|^p\right] \le C_p\left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\partial_\theta f_{j-1}(s,\theta)|^2 ds\right)^{\frac{1}{2}}.$$

We complete the proof by the above inequalities.
(2) By the Sobolev inequality and (3.9),

$$\begin{aligned}
&E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^{i} Z_{j-1}(\theta)U(\theta)\Delta_j Y\right|^p\right]^{\frac{1}{p}} \\
&\le C_p\left(E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^{i} Z_{j-1}(\theta)\int_{t_{j-1}}^{t_j} U(\theta)c^* m_s^* ds\right|^p\right]^{\frac{1}{p}}\right. \\
&\quad\left. +E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^{i} Z_{j-1}(\theta)U(\theta)\sigma^*(W_{t_j}-W_{t_{j-1}})\right|^p\right]^{\frac{1}{p}}\right). \qquad (3.19)
\end{aligned}$$

For the first term of the right-hand side, it follows from Lemma 3.1 (1), (3.11) and the Sobolev inequality

$$\begin{aligned}
&E\left[\sup_{\theta\in\Theta}\left|\sum_{j=1}^{i} Z_{j-1}(\theta)\int_{t_{j-1}}^{t_j} U(\theta)c^* m_s^* ds\right|^p\right]^{\frac{1}{p}} \\
&\le E\left[\sum_{j=1}^{i}\left|\sup_{\theta\in\Theta}\int_{t_{j-1}}^{t_j} Z_{j-1}(\theta)U(\theta)c^* m_s^* ds\right|^p\right]^{\frac{1}{p}} \\
&\le \left(\sum_{j=1}^{i} h^{p-1}\int_{t_{j-1}}^{t_j} E\left[\sup_{\theta\in\Theta}|Z_{j-1}(\theta)U(\theta)c^* m_s^*|^p\right] ds\right)^{\frac{1}{p}}
\end{aligned}$$

$$
\begin{aligned}
&\leq |c|^* \left(\sum_{j=1}^{i} h^{p-1} \int_{t_{j-1}}^{t_j} E\left[\sup_{\theta\in\Theta} |Z_{j-1}(\theta)|^{2p} \right]^{\frac{1}{2}} \right. \\
&\quad \left. \times E\left[\sup_{\theta\in\Theta} |U(\theta)|^{4p} \right]^{\frac{1}{4}} E[|m_s^*|^{4p}]^{\frac{1}{4}} ds \right)^{\frac{1}{p}} \\
&\leq C_p E\left[\sup_{\theta\in\Theta} |U(\theta)|^{4p} \right]^{\frac{1}{4}} \left(\sum_{j=1}^{i} h^p E\left[\sup_{\theta\in\Theta} |Z_{j-1}(\theta)|^{2p} \right]^{\frac{1}{2}} \right)^{\frac{1}{p}} \\
&\leq C_p E\left[\sup_{\theta\in\Theta} |U(\theta)|^{4p} \right]^{\frac{1}{4}} \\
&\quad \times \left(\sum_{j=1}^{i} h^p \sup_{\theta\in\Theta} \left\{ E\left[|Z_{j-1}(\theta)|^{2p} \right] + E\left[|\partial_\theta Z_{j-1}(\theta)|^{2p} \right] \right\}^{\frac{1}{2}} \right)^{\frac{1}{p}} \\
&\leq C_p E\left[\sup_{\theta\in\Theta} |U(\theta)|^{4p} \right]^{\frac{1}{4}} \\
&\quad \times \sup_{\theta\in\Theta} \sum_{j=1}^{i} \left\{ E\left[|Z_{j-1}(\theta)|^{2p} \right]^{\frac{1}{2p}} + E\left[|\partial_\theta Z_{j-1}(\theta)|^{2p} \right]^{\frac{1}{2p}} \right\} h.
\end{aligned}
$$

As for the second term, we have by Lemma 3.10

$$
\begin{aligned}
&E\left[\sup_{\theta\in\Theta} \left| \sum_{j=1}^{i} Z_{j-1}(\theta) U(\theta) \sigma^* (W_{t_j} - W_{t_{j-1}}) \right|^p \right]^{\frac{1}{p}} \\
&\leq C_p E\left[\sup_{\theta\in\Theta} |U(\theta)|^{2p} \right]^{\frac{1}{2p}} \\
&\quad \times \sup_{\theta\in\Theta} \left\{ E\left[\left(\sum_{j=1}^{n} |Z_{j-1}(\theta)|^2 h \right)^p \right] + E\left[\left(\sum_{j=1}^{n} \left| \partial_\theta Z_{j-1}(\theta) \right|^2 h \right)^p \right] \right\}^{\frac{1}{2p}} \\
&\leq C_p E\left[\sup_{\theta\in\Theta} |U(\theta)|^{2p} \right]^{\frac{1}{2p}} \\
&\quad \times \sup_{\theta\in\Theta} \left\{ E\left[\left(\sum_{j=1}^{n} |Z_{j-1}(\theta)|^2 h \right)^p \right]^{\frac{1}{p}} + E\left[\left(\sum_{j=1}^{n} \left| \partial_\theta Z_{j-1}(\theta) \right|^2 h \right)^p \right]^{\frac{1}{p}} \right\}^{\frac{1}{2}} \\
&\leq C_p E\left[\sup_{\theta\in\Theta} |U(\theta)|^{2p} \right]^{\frac{1}{2p}} \\
&\quad \times \sup_{\theta\in\Theta} \left\{ \sum_{j=1}^{n} E[|Z_{j-1}(\theta)|^{2p}]^{\frac{1}{p}} h + \sum_{j=1}^{n} E[\left|\partial_\theta Z_{j-1}(\theta)\right|^{2p}]^{\frac{1}{p}} h \right\}^{\frac{1}{2}}.
\end{aligned}
$$

Thus we completed the proof. □

Proposition 3.13 *For any $p > m_1 + m_2$, it holds*

$$\sup_{i\in\mathbb{N}} E\left[\sup_{\theta_2\in\Theta_2} |\hat{m}_i^n(\theta_2)|^p\right] < \infty$$

$$\sup_{i\in\mathbb{N}} E\left[\sup_{\theta_2\in\Theta_2} \left|\partial_{\theta_2}\hat{m}_i^n(\theta_2)\right|^p\right] < \infty$$

$$\sup_{i\in\mathbb{N}} E\left[\sup_{\theta_2\in\Theta_2} \left|\partial_{\theta_2}^2\hat{m}_i^n(\theta_2)\right|^p\right] < \infty$$

and

$$\sup_{i\in\mathbb{N}} E\left[\sup_{\theta_2\in\Theta_2} \left|\partial_{\theta_2}^3\hat{m}_i^n(\theta_2)\right|^p\right] < \infty.$$

Proof We only prove the first one; the rest can be shown in the same way. By (2.14) and the stability of $-\alpha(\theta_1, \theta_2)$, it is enough to show

$$\sup_{i\in\mathbb{N}} E\left[\left|\sup_{\theta=(\theta_1,\theta_2)\in\Theta} \sum_{j=1}^{i} \exp\left(-\alpha(\theta)(t_i - t_{j-1})\right) \gamma_+(\theta)c(\theta_2)'\Sigma(\theta_1)^{-1}\Delta_j Y\right|\right] < \infty.$$

To accomplish this, it is enough to show

$$\sum_{j=1}^{i} \left|\exp\left(-\alpha(\theta)(t_i - t_{j-1})\right) \gamma_+(\theta)c(\theta_2)'\Sigma(\theta_1)^{-1}\right| h < C \tag{3.20}$$

$$\sum_{j=1}^{i} \left|\exp\left(-\alpha(\theta)(t_i - t_{j-1})\right) \gamma_+(\theta)c(\theta_2)'\Sigma(\theta_1)^{-1}\right|^2 h < C \tag{3.21}$$

$$\sum_{j=1}^{i} \left|\partial_\theta \left\{\exp\left(-\alpha(\theta)(t_i - t_{j-1})\right) \gamma_+(\theta)c(\theta_2)'\Sigma(\theta_1)^{-1}\right\}\right| h < C \tag{3.22}$$

$$\sum_{j=1}^{i} \left|\partial_\theta \left\{\exp\left(-\alpha(\theta)(t_i - t_{j-1})\right) \gamma_+(\theta)c(\theta_2)'\Sigma(\theta_1)^{-1}\right\}\right|^2 h < C \tag{3.23}$$

according to (3.17).

These can be shown by using Corollary 3.7 and noting that it holds by Haber (2018)

$$\begin{aligned} &\left|\partial_\theta \exp\left(-\alpha(\theta)(t_i - t_{j-1})\right)\right| \\ &= \left|-\int_0^1 \exp(-s\alpha(\theta)(t_i - t_{j-1}))\partial_\theta\alpha(\theta)(t_i - t_{j-1}) \exp(-(1-s)\alpha(\theta)(t_i - t_{j-1}))ds\right| \\ &\le C(t_i - t_{j-1})e^{-C(t_i - t_{j-1})}. \end{aligned}$$

□

Proposition 3.14 *For any $n, i \in \mathbb{N}$ and $p > m_1 + m_2$,*

$$E\left[\sup_{\theta_2\in\Theta_2} |m_{t_i}(\theta_2) - \hat{m}_i^n(\theta_2)|^p\right]^{\frac{1}{p}} \le C_p(n^{-\frac{1}{2}} + h).$$

Proof By (2.12) and (2.14), we have

$$\begin{aligned}
&E\left[\sup_{\theta_2\in\Theta_2} |m_{t_i}(\theta_2) - \hat{m}_i^n(\theta_2)|^p\right]^{\frac{1}{p}} \\
&\le E\left[\sup_{\theta_2\in\Theta_2} \left|\left\{\exp(-\alpha(\theta_1^*, \theta_2)t) - \exp(-\alpha(\hat{\theta}_1^n, \theta_2)t)\right\} m_0\right|^p\right]^{\frac{1}{p}} \\
&\quad + E\left[\sup_{\theta_2\in\Theta_2} \left|\int_0^{t_i} \exp\left(-\alpha(\theta_1^*, \theta_2)(t_i - s)\right) \gamma_+(\theta_1^*, \theta_2)c(\theta_2)'\Sigma^{*-1} dY_s\right.\right. \\
&\quad \left.\left. - \sum_{j=1}^{i} \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)(t_i - t_{j-1})\right) \gamma_+(\hat{\theta}_1^n, \theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_j Y\right|^p\right]^{\frac{1}{p}} \\
&\le E\left[\sup_{\theta_2\in\Theta_2} \left|\left\{\exp(-\alpha(\theta_1^*, \theta_2)t) - \exp(-\alpha(\hat{\theta}_1^n, \theta_2)t)\right\} m_0\right|^p\right]^{\frac{1}{p}} \\
&\quad + E\left[\sup_{\theta_2\in\Theta_2} \left|\int_0^{t_i} \exp\left(-\alpha(\theta_1^*, \theta_2)(t_i - s)\right) \gamma_+(\theta_1^*, \theta_2)c(\theta_2)'\Sigma^{*-1} dY_s\right.\right. \\
&\quad \left.\left. - \sum_{j=1}^{i} \exp\left(-\alpha(\theta_1^*, \theta_2)(t_i - t_{j-1})\right) \gamma_+(\theta_1^*, \theta_2)c(\theta_2)'\Sigma^{*-1}\Delta_j Y\right|^p\right]^{\frac{1}{p}} \\
&\quad + E\left[\sup_{\theta_2\in\Theta_2} \left|\sum_{j=1}^{i} \exp\left(-\alpha(\theta_1^*, \theta_2)(t_i - t_{j-1})\right)\right.\right. \\
&\quad \left.\left.\left\{\gamma_+(\theta_1^*, \theta_2)c(\theta_2)'\Sigma^{*-1} - \gamma_+(\hat{\theta}_1^n, \theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\right\} \Delta_j Y\right|^p\right]^{\frac{1}{p}} \\
&\quad + E\left[\sup_{\theta_2\in\Theta_2} \left|\sum_{j=1}^{i} \left\{\exp\left(-\alpha(\theta_1^*, \theta_2)(t_i - t_{j-1})\right)\right.\right.\right. \\
&\quad \left.\left.\left. - \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)(t_i - t_{j-1})\right)\right\} \gamma_+(\hat{\theta}_1^n, \theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_j Y\right|^p\right]^{\frac{1}{p}}. \qquad (3.24)
\end{aligned}$$

The first term of the right-hand side can be bounded by the mean value theorem and Theorem 2.1:

$$\begin{aligned}
&E\left[\sup_{\theta_2\in\Theta_2} \left|\left\{\exp(-\alpha(\theta_1^*, \theta_2)t) - \exp(-\alpha(\hat{\theta}_1^n, \theta_2)t)\right\} m_0\right|^p\right]^{\frac{1}{p}} \\
&\le CE\left[|\hat{\theta}_1^n - \theta_1^*|^p\right]^{\frac{1}{p}} \le Cn^{-\frac{1}{2}}.
\end{aligned} \qquad (3.25)$$

Next we evaluate the second term using (3.16). Noting that by the mean value theorem and Lemma 3.2, we have

$$\begin{aligned}
&\left|\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s)) - \exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))\right| \\
&\quad = \left|\alpha(\theta_1^*,\theta_2)\exp(-\alpha(\theta_1^*,\theta_2)(t_i-u))(s-t_{j-1})\right| \\
&\quad \le Ce^{-C(t_i-u)}(s-t_{j-1}) \\
&\quad \le Ce^{-C(t_i-s)}h
\end{aligned}$$

and

$$\begin{aligned}
&\left|(t_i-s)\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s)) - (t_i-t_{j-1})\exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))\right| \\
&\quad \le |(t_{j-1}-s)\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s)| \\
&\qquad + (t_i-t_{j-1})\left|\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s)) - \exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))\right| \\
&\quad \le Ce^{-C(t_i-s)}h,
\end{aligned}$$

where $t_{j-1} \le u \le s \le t_j$, it follows from (3.16)

$$\begin{aligned}
&E\left[\sup_{\theta_2\in\Theta_2}\left|\int_0^{t_i}\exp\left(-\alpha(\theta_1^*,\theta_2)(t_i-s)\right)\gamma_+(\theta_1^*,\theta_2)c(\theta_2)'\Sigma^{*-1}dY_s\right.\right. \\
&\qquad \left.\left. -\sum_{j=1}^{i}\exp\left(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1})\right)\gamma_+(\theta_1^*,\theta_2)c(\theta_2)'\Sigma^{*-1}\Delta_j Y\right|^p\right]^{\frac{1}{p}} \\
&\quad = E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}\left\{\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s))\right.\right.\right. \\
&\qquad \left.\left.\left. -\exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))\right\}\gamma_+(\theta_1^*,\theta_2)c(\theta_2)'\Sigma^{*-1}dY_s\right|^p\right]^{\frac{1}{p}} \\
&\quad \le C_p\sup_{\theta_2\in\Theta_2}\left\{\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s))\right. \\
&\qquad -\exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))|ds \\
&\qquad +\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\partial_{\theta_2}\alpha(\theta_1^*,\theta_2)(t_i-s)\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s)) \\
&\qquad -\partial_{\theta_2}\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1})\exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))|ds \\
&\qquad +\left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s)) - \exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))|^2 ds\right)^{\frac{1}{2}} \\
&\qquad +\left(\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}|\partial_{\theta_2}\alpha(\theta_1^*,\theta_2)(t_i-s)\exp(-\alpha(\theta_1^*,\theta_2)(t_i-s))\right. \\
&\qquad \left.\left. -\partial_{\theta_2}\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1})\exp(-\alpha(\theta_1^*,\theta_2)(t_i-t_{j-1}))|^2 ds\right)^{\frac{1}{2}}\right\}
\end{aligned}$$

$$
\begin{aligned}
&\le C_p \sum_{j=1}^{i} \int_{t_{j-1}}^{t_j} e^{-C(t_i-s)} dsh + C_p \left(\sum_{j=1}^{i} e^{-C(t_i-s)} h^2 \right)^{\frac{1}{2}} \\
&\le C_p \int_0^{t_i} e^{-C(t_i-s)} dsh + C_p \left(\int_0^{t_i} e^{-C(t_i-s)} dsh^2 \right)^{\frac{1}{2}} \\
&\le C_p h. \qquad (3.26)
\end{aligned}
$$

As for the third term, in the same way as Proposition 3.13, we have

$$
E\left[\sup_{\theta_2\in\Theta_2} \left| \sum_{j=1}^{i} \exp\left(-\alpha(\theta_1^*, \theta_2)(t_i - t_{j-1})\right) \right.\right.
$$

$$
\left.\left. \left\{ \gamma_+(\theta_1^*, \theta_2)c(\theta_2)'\Sigma^{*-1} - \gamma_+(\hat{\theta}_1^n, \theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1} \right\} \Delta_j Y \right|^p \right]^{\frac{1}{p}} \le C_p n^{-\frac{1}{2}},
$$

since it holds

$$
E\left[\sup_{\theta_2\in\Theta_2} \left| \gamma_+(\theta_1^*, \theta_2)c(\theta_2)'\Sigma^{*-1} - \gamma_+(\hat{\theta}_1^n, \theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1} \right|^p \right]^{\frac{1}{p}} \le C_p n^{-\frac{1}{2}} \qquad (3.27)
$$

by the mean value theorem and Theorem 2.1.

Finally, we consider the forth term of (3.24). Noting that it follows from Lemma 3.2 and the stability of $-\alpha(\theta_1, \theta_2)$,

$$
\begin{aligned}
&\left| \exp\left(-\left[\alpha(\theta_1, \theta_2) + \left\{\alpha(\theta_1^*, \theta_2) - \alpha(\theta_1, \theta_2)\right\} u\right](t_i - t_{j-1})\right) \right| \\
&\quad = \left| \exp\left(-\alpha(\theta_1, \theta_2)(1-u)(t_i - t_{j-1})\right) \right| \left| \exp\left(\alpha(\theta_1^*, \theta_2)u(t_i - t_{j-1})\right) \right| \\
&\quad \le C e^{-C(1-u)(t_i - t_{j-1})} e^{-Cu(t_i - t_{j-1})} = C e^{-C(t_i - t_{j-1})},
\end{aligned}
$$

we have

$$
\sum_{j=1}^{i} \left| (t_i - t_{j-1}) \right.
$$

$$
\left. \int_0^1 \exp\left(-\left[\alpha(\theta_1, \theta_2) + \left\{\alpha(\theta_1^*, \theta_2) - \alpha(\theta_1, \theta_2)\right\} u\right](t_i - t_{j-1})\right) du \right| h \le C.
$$

In the same way, we obtain the boundedness of

$$
\sum_{j=1}^{i} \left| (t_i - t_{j-1}) \right.
$$

$$
\left. \times \partial_{(\theta_1, \theta_2)} \int_0^1 \exp\left(-\left[\alpha(\theta_1, \theta_2) + \left\{\alpha(\theta_1^*, \theta_2) - \alpha(\theta_1, \theta_2)\right\} u\right](t_i - t_{j-1})\right) du \right| h,
$$

$$
\sum_{j=1}^{i} \left| (t_i - t_{j-1}) \right.
$$

$$
\left. \times \int_0^1 \exp\left(-\left[\alpha(\theta_1, \theta_2) + \left\{\alpha(\theta_1^*, \theta_2) - \alpha(\theta_1, \theta_2)\right\} u\right](t_i - t_{j-1})\right) du \right|^2 h
$$

and

$$\sum_{j=1}^{i} \Big| (t_i - t_{j-1}) \times \partial_{(\theta_1,\theta_2)} \int_0^1 \exp\left(-\left[\alpha(\theta_1,\theta_2) + \left\{\alpha(\theta_1^*,\theta_2) - \alpha(\theta_1,\theta_2)\right\} u\right](t_i - t_{j-1})\right) du \Big|^2 h.$$

Thus by (3.17) we obtain

$$\sum_{j=1}^{i} E\left[\sup_{\theta_2\in\Theta_2} \Big| (t_i - t_{j-1}) \int_0^1 \exp\left(-\left[\alpha(\hat{\theta}_1^n,\theta_2) + \left\{\alpha(\theta_1^*,\theta_2) - \alpha(\hat{\theta}_1^n,\theta_2)\right\} u\right](t_i - t_{j-1})\right) du \; \gamma_+(\hat{\theta}_1^n,\theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_j Y \Big|^p\right] \le C_p.$$

Therefore it follows

$$E\left[\sup_{\theta_2\in\Theta_2} \Big| \sum_{j=1}^{i} \left\{\exp\left(-\alpha(\theta_1^*,\theta_2)(t_i - t_{j-1})\right) - \exp\left(-\alpha(\hat{\theta}_1^n,\theta_2)(t_i - t_{j-1})\right)\right\} \gamma_+(\hat{\theta}_1^n,\theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_j Y \Big|^p\right]^{\frac{1}{p}} \le C_p n^{-\frac{1}{2}}. \tag{3.28}$$

Now we completed the proof by (3.24)–(3.28). □

Next, we replace m_0^* and γ_s^* with m_0 and $\gamma_+(\theta^*)$ in (3.10), and introduce

$$\tilde{m}_t^* = \exp(-a^* t)m_0 + \int_0^t \exp(-a^*(t-s))\gamma_+(\theta^*)c^{*\prime}{\sigma^{*\prime}}^{-1} d\overline{W}_s. \tag{3.29}$$

Furthermore, we consider for every $n, i \in \mathbb{N}$,

$$\tilde{Y}_t = Y_0 + \int_0^t c^* \tilde{m}_s^* ds + \sigma^* \overline{W}_t, \tag{3.30}$$

$$\tilde{m}_i^n(\theta_2) = \exp\left(-\alpha(\hat{\theta}_1^n,\theta_2)t_i\right) m_0 + \sum_{j=1}^{i} \exp\left(-\alpha(\hat{\theta}_1^n,\theta_2)(t_i - t_{j-1})\right) \gamma_+(\hat{\theta}_1^n,\theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\Delta_j \tilde{Y}. \tag{3.31}$$

and

$$\tilde{\Delta}_i Y = c^* \tilde{m}_{i-1}(\theta_2^*)h + \sigma^* \Delta_i \overline{W}. \tag{3.32}$$

Then in the same way as Proposition 3.13, it holds for any $p > m_1 + m_2$

$$\sup_{i\in\mathbb{N}} E\left[\sup_{\theta_2\in\Theta_2} |\partial_{\theta_2}^k \tilde{m}_i^n(\theta_2)|^p\right] < \infty \; (k = 0, 1, 2, 3). \tag{3.33}$$

Proposition 3.15 *For any $p > 0$ and $t \geq 0$, it holds*

$$E\left[|m_t^* - \tilde{m}_t^*|^p\right]^{\frac{1}{p}} \leq C_p e^{-Ct}.$$

Proof Use (3.10), (3.29), Lemmas 3.1 and 3.2, Proposition 3.9 and the stability of α. □

Proposition 3.16 *Let $A : \Theta \to M_{d_1,d_2}(\mathbb{R})$ be a continuous mapping. Then for any $i, n \in \mathbb{N}$, $p > 0$ and $k = 0, 1, 2, \cdots$, it holds*

$$E\left[\sup_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}\left|\sum_{j=1}^{i}(t_i - t_{j-1})^k \exp\left(-\alpha(\theta_1,\theta_2)(t_i - t_{j-1})\right)\right.\right.$$
$$\left.\left.\times A(\theta_1,\theta_2)(\Delta_j\tilde{Y} - \Delta_j Y)\right|^p\right]^{\frac{1}{p}} \leq C_{p,k,A}e^{-Ct_i}.$$

Proof It can be shown that by Lemma 3.1 and Proposition 3.15, since

$$\Delta_j\tilde{Y} - \Delta_j Y = c^*\int_{t_{j-1}}^{t_j}\{\tilde{m}_s^* - m_s^*\}ds.$$

□

By (2.14) and (3.31), we obtain the following corollaries.

Corollary 3.17 *For any $i, n \in \mathbb{N}$, $p > 0$ and $k = 0, 1, 2, 3, 4$, it holds*

$$E\left[\sup_{\theta_2\in\Theta_2}\left|\partial_{\theta_2}^k\{\tilde{m}_i^n(\theta) - \hat{m}_i^n(\theta)\}\right|^p\right]^{\frac{1}{p}} \leq C_p e^{-Ct_i}.$$

Corollary 3.18 *For any $i, n \in \mathbb{N}$ and $p > m_1 + m_2$, it holds*

$$E\left[|\Delta_i Y - \tilde{\Delta}_i Y|^p\right]^{\frac{1}{p}} \leq C_p(h^{\frac{3}{2}} + n^{-\frac{1}{2}}h + e^{-Ct_i}h).$$

Proof This result directly follows from By (3.9), (3.32), (3.12), Lemma 3.1, Propositions 3.14 and 3.15 and Corollary 3.17. □

Proposition 3.19 *Let $A : \Theta \to M_{d_1,d_2}(\mathbb{R})$ be a continuous mapping. Then for any $n \in \mathbb{N}$, $p > m_1 + m_2$ and $k = 0, 1, 2, 3$*

$$E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^{n}\partial_{\theta_2}^k\{\hat{m}_{i-1}^n(\theta_2) - \tilde{m}_{i-1}^n(\theta_2)\}A(\theta_2)\Delta_i Y\right|^p\right] < C_p.$$

Proof By (2.14) and (3.31), we have

$$\hat{m}_{i-1}^n(\theta_2) - \tilde{m}_{i-1}^n(\theta_2)$$
$$= \sum_{j=1}^{i}\exp\left(-\alpha(\hat{\theta}_1^n,\theta_2)(t_i - t_{j-1})\right)\gamma_+(\hat{\theta}_1^n,\theta_2)c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}(\Delta_j Y - \Delta_j\tilde{Y}).$$

Hence for every $k = 0, 1, 2, 3$, $\partial_{\theta_2}^k \{\hat{m}_{i-1}^n(\theta_2) - \tilde{m}_{i-1}^n(\theta_2)\}$ is a sum of the form

$$\sum_{j=1}^{i} \partial_{\theta_2}^l \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)(t_i - t_{j-1})\right) A_1(\theta)(\Delta_j Y - \Delta_j \tilde{Y}) \qquad (l = 0, 1, 2, 3),$$

where A_1 is a $M_{d_1,d_2}(\mathbb{R})$-valued k-dimensional tensor of class C^1. Thus if we set

$$\Phi(\theta) = \sum_{j=1}^{i} \partial_{\theta_2}^l \exp\left(-\alpha(\theta_1, \theta_2)(t_i - t_{j-1})\right) A_1(\theta)(\Delta_j Y - \Delta_j \tilde{Y}),$$

it is enough to show

$$E\left[\sup_{\theta\in\Theta}\left|\sum_{i=1}^{n} \Phi(\theta)\Delta_i Y\right|^p\right]^{\frac{1}{p}} < C_p. \tag{3.34}$$

By Haber (2018), we have

$$E\left[\sup_{\theta\in\Theta} |\Phi(\theta)|^p\right]^{\frac{1}{p}} \leq C_p e^{-Ct_i}$$

and

$$E\left[\sup_{\theta\in\Theta} |\partial_\theta \Phi(\theta)|^p\right]^{\frac{1}{p}} \leq C_p e^{-Ct_i}.$$

Thus it holds by (3.17) and Proposition 3.16

$$E\left[\sup_{\theta\in\Theta}\left|\sum_{i=1}^{n} \Phi(\theta)\Delta_i Y\right|^p\right]^{\frac{1}{p}} \leq C_p \sum_{i=1}^{n} e^{-Ct_i} \leq C_p.$$

Hence we obtain (3.34). □

Proposition 3.20 *Let Z be a $M_{d_2}(\mathbb{R})$-valued random variable. Then for any $n \in \mathbb{N}, k = 0, 1, 2, 3$ and $p > m_1 + m_2$ it holds*

$$E\left[\left|\sup_{\theta_2\in\Theta_2} \sum_{i=1}^{n} \partial_{\theta_2}^k \{\hat{m}_{i-1}^n(\theta_2)' c(\theta_2)'\} Z \Delta_j Y\right|^p\right]^{\frac{1}{p}} \leq C_p \left(E\left[|A|^{4p}\right]^{\frac{1}{4p}} nh + E\left[|A|^{2p}\right]^{\frac{1}{2p}} (nh)^{\frac{1}{2}}\right).$$

Proof By (2.14), $\partial_{\theta_2}^k \{\hat{m}_i^n(\theta_2)' c(\theta_2)'\}$ is a sum of the form

$$A_i(\hat{\theta}_1^n, \theta_2) \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2) t_i\right) + \sum_{l=0}^{k} \sum_{j=1}^{i} \partial_{\theta_2}^l \exp\left(-\alpha(\hat{\theta}_1^n, \theta_2)(t_i - t_{j-1})\right) B_i(\hat{\theta}_1^n, \theta_2)\Delta_j Y,$$

where A_i and B_i are k-dimensional tensor valued continuously differentiable mappings on Θ. Thus if we set

$$\Psi_i(\theta) = \Psi_i(\theta_1, \theta_2) = A_i(\theta) \exp(-\alpha(\theta) t_i) + \sum_{l=0}^{k} \sum_{j=1}^{i} \partial_{\theta_2}^l \exp\left(-\alpha(\theta_1, \theta_2)(t_i - t_{j-1})\right) B_i(\theta) \Delta_j Y,$$

it is enough to show

$$E\left[\left|\sup_{\theta\in\Theta} \sum_{i=1}^{n} \Psi_{i-1}(\theta) Z \Delta_j Y\right|^p\right]^{\frac{1}{p}} \leq C_p \left(E\left[|A|^{4p}\right]^{\frac{1}{4p}} nh + E\left[|A|^{2p}\right]^{\frac{1}{2p}} (nh)^{\frac{1}{2}}\right). \quad (3.35)$$

In the same way as Proposition 3.13, we first obtain

$$E\left[|\Psi_i(\theta)|^p\right] \leq C_p$$

and

$$E\left[|\partial_\theta \Psi_i(\theta)|^p\right] \leq C_p.$$

Therefore noting that $\Psi_i(\theta)$ is $\mathcal{F}_{t_{i-1}}$-measurable, we obtain (3.35) by (3.17). □

Next, we define $\tilde{\mathbb{H}}_n^2$, $\tilde{\Delta}_n^2$, $\tilde{\Gamma}_n^2$ and $\tilde{\mathbb{Y}}_n^2$ by

$$\tilde{\mathbb{H}}_n^2(\theta_2) = \frac{1}{2} \sum_{i=1}^{n} \Big\{-h \Sigma^{*-1}[(c(\theta_2)\tilde{m}_{i-1}^n(\theta_2))^{\otimes 2}] + \tilde{m}_{i-1}^n(\theta_2)' c(\theta_2)' \Sigma^{*-1} \tilde{\Delta}_j Y + \tilde{\Delta}_j Y' \Sigma^{*-1} c(\theta_2) \tilde{m}_{i-1}^n(\theta_2)\Big\} \quad (3.36)$$

$$\tilde{\mathbb{Y}}_n^2(\theta_2) = \frac{1}{t_n}\{\tilde{\mathbb{H}}_n^2(\theta_2) - \tilde{\mathbb{H}}_n^2(\theta_2^*)\} \quad (3.37)$$

$$\tilde{\Delta}_n^2 = \frac{1}{\sqrt{t_n}} \partial_\theta \tilde{\mathbb{H}}_n^2(\theta_2^*), \quad (3.38)$$

and

$$\tilde{\Gamma}_n^2 = -\frac{1}{t_n} \partial_\theta^2 \tilde{\mathbb{H}}_n^2(\theta_2^*), \quad (3.39)$$

respectively.

Proposition 3.21 *For any $n \in \mathbb{N}$, $p > m_1 + m_2$ and $k = 0, 1, 2, 3$, it holds*

$$E\left[\sup_{\theta_2\in\Theta_2} \left|\partial_{\theta_2}^k \{\mathbb{H}_n(\theta_2) - \tilde{\mathbb{H}}_n(\theta_2)\}\right|^p\right]^{\frac{1}{p}} \leq C_p (nh^{\frac{3}{2}} + n^{\frac{1}{2}} h + 1).$$

Proof We only consider the case of $k = 0$. The rest is the same. By (2.15) and (3.36),

$$E\left[\sup_{\theta_2\in\Theta_2} \left|\mathbb{H}_n(\theta_2) - \tilde{\mathbb{H}}_n(\theta_2)\right|^p\right]^{\frac{1}{p}} \leq E\left[\sup_{\theta_2\in\Theta_2} \left|\frac{1}{2} h \sum_{i=1}^{n} \{\Sigma(\hat{\theta}_1^n)^{-1} - \Sigma^{*-1}\}[(c(\theta_2)\hat{m}_{j-1}^n(\theta_2))^{\otimes 2}]\right|^p\right]^{\frac{1}{p}}$$

$$+ E\left[\sup_{\theta_2\in\Theta_2}\left|\frac{1}{2}\sum_{i=1}^{n}\hat{m}_{j-1}^{n}(\theta_2)'c(\theta_2)'\{\Sigma(\hat{\theta}_1^n)^{-1}-\Sigma^{*-1}\}\Delta_j Y\right|^p\right]^{\frac{1}{p}}$$

$$+ E\left[\sup_{\theta_2\in\Theta_2}\left|\frac{1}{2}\sum_{i=1}^{n}\Delta_j Y'\{\Sigma(\hat{\theta}_1^n)^{-1}-\Sigma^{*-1}\}c(\theta_2)\hat{m}_{j-1}^{n}(\theta_2)\right|^p\right]^{\frac{1}{p}}$$

$$+ E\left[\sup_{\theta_2\in\Theta_2}\left|\frac{1}{2}h\sum_{i=1}^{n}\left\{\Sigma^{*-1}[(c(\theta_2)\hat{m}_{j-1}^{n}(\theta_2))^{\otimes 2}]\right.\right.\right.$$
$$\left.\left.\left.-\Sigma^{*-1}[(c(\theta_2)\tilde{m}_{j-1}^{n}(\theta_2))^{\otimes 2}]\right\}\right|^p\right]^{\frac{1}{p}}$$

$$+ E\left[\sup_{\theta_2\in\Theta_2}\left|\frac{1}{2}\sum_{i=1}^{n}\{\hat{m}_{j-1}^{n}(\theta_2)'c(\theta_2)'\Sigma^{*-1}\Delta_j Y\right.\right.$$
$$\left.\left.-c(\theta_2)\tilde{m}_{j-1}^{n}(\theta_2)\Sigma^{*-1}\tilde{\Delta}_j Y\}\right|^p\right]^{\frac{1}{p}}$$

$$+ E\left[\sup_{\theta_2\in\Theta_2}\left|\frac{1}{2}\sum_{i=1}^{n}\{\Delta_j Y'\Sigma^{*-1}c(\theta_2)\hat{m}_{j-1}^{n}(\theta_2)\right.\right.$$
$$\left.\left.-\tilde{\Delta}_j Y'\Sigma^{*-1}c(\theta_2)\tilde{m}_{j-1}^{n}(\theta_2)\}\right|^p\right]^{\frac{1}{p}}. \tag{3.40}$$

For the first three terms of the right-hand side, we have by Theorem 2.1 and Proposition 3.13

$$E\left[\sup_{\theta_2\in\Theta_2}\left|\frac{1}{2}h\sum_{i=1}^{n}\{\Sigma(\hat{\theta}_1^n)^{-1}-\Sigma^{*-1}\}[(c(\theta_2)\hat{m}_{j-1}^{n}(\theta_2))^{\otimes 2}]\right|^p\right]^{\frac{1}{p}}$$
$$\leq C_p n^{\frac{1}{2}}h,$$

and by Proposition 3.20

$$E\left[\sup_{\theta_2\in\Theta_2}\left|\frac{1}{2}\sum_{i=1}^{n}\hat{m}_{j-1}^{n}(\theta_2)'c(\theta_2)'\{\Sigma(\hat{\theta}_1^n)^{-1}-\Sigma^{*-1}\}\Delta_j Y\right|^p\right]^{\frac{1}{p}}$$
$$\leq C_p n^{-\frac{1}{2}}\{nh+(nh)^{\frac{1}{2}}\}\leq C_p(n^{-\frac{1}{2}}h+h^{\frac{1}{2}}).$$

In the same way, the third term can be bounded by $C_p(n^{-\frac{1}{2}}h+h^{\frac{1}{2}})$.

Furthermore, making use of Proposition 3.13, (3.33) and Corollary 3.17, we can bound the fourth term by $C_p\sum_{i=1}^{n}he^{-Ct_i}\leq C_p h$, noting that

$$\Sigma(\hat{\theta}_1^n)^{-1}[(c(\theta_2)\hat{m}_{j-1}^{n}(\theta_2))^{\otimes 2}]-\Sigma(\hat{\theta}_1^n)^{-1}[(c(\theta_2)\tilde{m}_{j-1}^{n}(\theta_2))^{\otimes 2}]$$
$$=\{\hat{m}_{j-1}^{n}(\theta_2)+\tilde{m}_{j-1}^{n}(\theta_2)\}'c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}c(\theta_2)\{\hat{m}_{j-1}^{n}(\theta_2)-\tilde{m}_{j-1}^{n}(\theta_2)\}$$
$$+\{\hat{m}_{j-1}^{n}(\theta_2)-\tilde{m}_{j-1}^{n}(\theta_2)\}'c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\tilde{m}_{j-1}^{n}(\theta_2)$$
$$+\tilde{m}_{j-1}^{n}(\theta_2)'c(\theta_2)'\Sigma(\hat{\theta}_1^n)^{-1}\{\hat{m}_{j-1}^{n}(\theta_2)-\tilde{m}_{j-1}^{n}(\theta_2)\}.$$

Finally, the last two terms can be bounded by $C_p + C_p \sum_{i=1}^{n}(h^{\frac{3}{2}} + n^{-\frac{1}{2}}h + e^{-Ct_i}h) \leq C_p(1 + nh^{\frac{3}{2}} + n^{\frac{1}{2}}h + h)$ due to the Corollary 3.18, Proposition 3.19 and the identity

$$\begin{aligned}&\hat{m}^n_{j-1}(\theta_2)'c(\theta_2)'\Sigma(\hat{\theta}^n_1)^{-1}\Delta_j Y - \tilde{m}^n_{j-1}(\theta_2)'c(\theta_2)'\Sigma(\hat{\theta}^n_1)^{-1}\tilde{\Delta}_j Y\\ &\quad = \{\hat{m}^n_{j-1}(\theta_2) - \tilde{m}^n_{j-1}(\theta_2)\}'c(\theta_2)'\Sigma(\hat{\theta}^n_1)^{-1}\Delta_j Y\\ &\qquad + \tilde{m}^n_{j-1}(\theta_2)'c(\theta_2)'\Sigma(\hat{\theta}^n_1)^{-1}\{\Delta_j Y - \tilde{\Delta}_j Y\}.\end{aligned}$$

Putting it all together, we obtain

$$\begin{aligned}E\left[\sup_{\theta_2\in\Theta_2}\left|\mathbb{H}_n(\theta_2) - \tilde{\mathbb{H}}_n(\theta_2)\right|^p\right]^{\frac{1}{p}} &\leq C_p(1 + nh^{\frac{3}{2}} + n^{\frac{1}{2}}h + h^{\frac{1}{2}} + h)\\ &\leq C_p(1 + nh^{\frac{3}{2}} + n^{\frac{1}{2}}h).\end{aligned}$$

□

Proposition 3.22 *For any $p \geq 2$, it holds*

$$\sup_{n\in\mathbb{N}} E\left[|\tilde{\Delta}_n|^p\right] < \infty.$$

Proof If we set $\tilde{M}^n_j(\theta_2) = c(\theta_2)\tilde{m}^n_j(\theta)$, we have

$$\begin{aligned}\tilde{\Delta}^2_n &= \frac{1}{2\sqrt{t_n}}\sum_{i=1}^{n}\left\{\partial_{\theta_2}\tilde{M}^n_{i-1}(\theta_2)'\sigma^{*\prime -1}\Delta_i\overline{W} + \Delta_i\overline{W}'\sigma^{*-1}\partial_{\theta_2}\tilde{M}^n_{i-1}(\theta_2)\right\}\\ &= \frac{1}{\sqrt{t_n}}\sum_{i=1}^{n}\left\{\partial_{\theta_2}\tilde{M}^n_i(\theta_2)'\sigma^{*\prime -1}\Delta_i\overline{W}\right\}.\end{aligned} \tag{3.41}$$

by (3.32), (3.38) and (3.51). Thus by Lemma 3.1 and (3.33),

$$\begin{aligned}E\left[|\tilde{\Delta}_n|^p\right]^{\frac{2}{p}} &\leq \frac{1}{t_n^{\frac{p}{2}}}E\left[\left(\sum_{i=1}^{n}|\partial_{\theta_2}\tilde{M}^n_i(\theta_2)'\sigma^{*\prime -1}|^2 h\right)^{\frac{p}{2}}\right]^{\frac{2}{p}}\\ &\leq \frac{1}{t_n^{\frac{p}{2}}}C_p\sum_{i=1}^{n}E\left[|\partial_{\theta_2}\tilde{M}^n_i(\theta_2)'\sigma^{*\prime -1}|^2\right]^{\frac{2}{p}}h\\ &\leq \frac{1}{t_n^{\frac{p}{2}}} \times C_p nh = C_p.\end{aligned}$$

□

Next, we define the process $\{\mu_t\}$ by replacing Y with $\tilde{Y}$ (therefore m^*_t with $m_t(\theta^*)$ and γ^*_t with $\gamma_+(\theta^*)$) in (2.12);

$$\begin{aligned}\mu_t(\theta_2) &= \exp\left(-\alpha(\theta_2)t\right)m_0\\ &\quad + \int_0^t \exp\left(-\alpha(\theta_2)(t-s)\right)\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}d\tilde{Y}_s.\end{aligned} \tag{3.42}$$

Then as m_t is the solution of (2.11), so μ_t is the solution of

$$\begin{cases} d\mu_t(\theta_2) = -\alpha(\theta_2)\mu_t dt + \gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}d\tilde{Y}_t \\ \mu_0(\theta_2) = m_0. \end{cases} \tag{3.43}$$

Moreover, it holds $\mu_t(\theta_2^*) = \tilde{m}_t^*$ since by (3.31) $\tilde{m}_t^*$ is the solution of

$$d\tilde{m}_t^* = -a^*\tilde{m}_t^* + \gamma_+(\theta^*)c^{*\prime}\sigma^{*\prime -1}d\overline{W}_t,$$

which is equivalent to

$$d\tilde{m}_t^* = -\alpha(\theta_2^*)\tilde{m}_t^* dt + \gamma_+(\theta^*)c^{*\prime}\Sigma^{*-1}d\tilde{Y}_t.$$

Moreover, just as Proposition 3.14, the following proposition holds:

Proposition 3.23 *For any $n, i \in \mathbb{N}$ and $p > m_1 + m_2$, we have*

$$E\left[\sup_{\theta_2\in\Theta_2} |\mu_{t_i}(\theta_2) - \tilde{m}_i^n(\theta_2)|^p\right]^{\frac{1}{p}} \leq C_p(n^{-\frac{1}{2}} + h).$$

Together with (3.33), we obtain the following corollary.

Corollary 3.24 *For any $i \in \mathbb{N}$ and $p > m_1 + m_2$, we have*

$$E\left[\sup_{\theta_2\in\Theta_2} |\mu_{t_i}(\theta_2)|^p\right]^{\frac{1}{p}} \leq C_p.$$

Proposition 3.25

$$E[\Sigma^{*-1}[\{c(\theta_2)\mu_t(\theta_2) - c(\theta_2^*)\mu_t(\theta_2^*)\}^{\otimes 2}]], = -2\mathbb{Y}(\theta_2) + O(e^{-t_i})$$

where $O(e^{-t})$ is some continuous function $r : \Theta \to \mathbb{R}$ such that

$$|r(\theta)| \leq Ce^{-Ct}.$$

Proof By (3.42) and (3.31), we have

$$\begin{aligned}
\mu_t(\theta) &= \exp(-\alpha(\theta_2)t)m_0 \\
&\quad + \int_0^t \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\tilde{m}_s ds \\
&\quad + \int_0^t \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\sigma^{*\prime -1}d\overline{W}_s \\
&= \exp(-\alpha(\theta_2)t)m_0 \\
&\quad + \int_0^t \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^* \\
&\quad \times \left\{\exp(-a^*s)m_0 + \int_0^s \exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}\sigma^{*\prime -1}d\overline{W}_u\right\} ds \\
&\quad + \int_0^t \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\sigma^{*\prime -1}d\overline{W}_s \\
&= \exp(-\alpha(\theta_2)t)m_0 \\
&\quad + \int_0^t \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\exp(-a^*s)m_0 ds
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^t\int_0^s \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\\
&\times\exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}\sigma^{*\prime-1}d\overline{W}_u ds\\
&+\int_0^t \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\sigma^{*\prime-1}d\overline{W}_s\\
=\;&\exp(-\alpha(\theta_2)t)m_0\\
&+\int_0^t \exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\exp(-a^*s)m_0 ds\\
&+\int_0^t\left\{\int_s^t \exp(-\alpha(\theta_2)(t-u))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\right.\\
&\exp(-a^*(u-s))\gamma_+(\theta^*)c^{*\prime}\sigma^{*\prime-1}du\\
&\left.+\exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\sigma^{*\prime-1}\right\}d\overline{W}_s.
\end{aligned}
\tag{3.44}
$$

Therefore

$$
\begin{aligned}
&E[\Sigma^{*-1}[\{c(\theta_2)\mu_t(\theta_2)-c(\theta_2^*)\mu_t(\theta_2^*)\}^{\otimes 2}]]\\
&=E[\Sigma^{*-1}[\{c(\theta_2)\mu_t(\theta_2)-c(\theta_2^*)\tilde{m}_t^*\}^{\otimes 2}]]\\
&=E\Bigg[\Sigma^{*-1}\Bigg[\Bigg\{\int_0^t\Bigg\{\int_s^t c(\theta_2)\exp(-\alpha(\theta_2)(t-u))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\\
&\quad\times\exp(-a^*(u-s))\gamma_+(\theta^*)c^{*\prime}du\\
&\quad+c(\theta_2)\exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\\
&\quad-c^*\exp(-a^*(t-s))\gamma_+(\theta^*)c^{*\prime}\Bigg\}\sigma^{*\prime-1}d\overline{W}_s\Bigg\}^{\otimes 2}\Bigg]\Bigg]+O(e^{-Ct})\\
&=\mathrm{Tr}\int_0^t\Sigma^{*-1}\Bigg[\Bigg\{\int_s^t c(\theta_2)\exp(-\alpha(\theta_2)(t-u))\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\\
&\quad\times\exp(-a^*(u-s))\gamma_+(\theta^*)c^{*\prime}du\\
&\quad+c(\theta_2)\exp(-\alpha(\theta_2)(t-s))\gamma_+(\theta_2)c(\theta_2)'\\
&\quad-c^*\exp(-a^*(t-s))\gamma_+(\theta^*)c^{*\prime}\Bigg\}^{\otimes 2}\Bigg][(\sigma^{*\prime-1})^{\otimes 2}]ds+O(e^{-Ct})\\
&=\mathrm{Tr}\int_0^t\Sigma^{*-1}\Bigg[\Bigg\{\int_0^s c(\theta_2)\exp(-\alpha(\theta_2)u)\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\\
&\quad\times\exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}du\\
&\quad+c(\theta_2)\exp(-\alpha(\theta_2)s)\gamma_+(\theta_2)c(\theta_2)'\\
&\quad-c^*\exp(-a^*s)\gamma_+(\theta^*)c^{*\prime}\Bigg\}^{\otimes 2}\Bigg][(\sigma^{*\prime-1})^{\otimes 2}]ds+O(e^{-Ct}).
\end{aligned}
$$

Now we have

$$
\begin{aligned}
&\int_0^s|c(\theta_2)\exp(-\alpha(\theta_2)u)\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}|du\\
&\quad\le\int_0^s C_pe^{-Cu}e^{-C(s-u)}du\le C_pse^{-Cs}\le C_pe^{-Cs}
\end{aligned}
$$

and thus by (2.6)

$$
\begin{aligned}
&|E[\Sigma^{*-1}[\{c(\theta_2)\mu_t(\theta_2)-c(\theta_2^*)\mu_t(\theta_2^*)\}^{\otimes 2}]]+2\mathbb{Y}^2(\theta_2)| \\
&\quad=\left|\int_t^\infty \Sigma^{*-1}\left[\left\{\int_0^s c(\theta_2)\exp(-\alpha(\theta_2)u)\gamma_+(\theta_2)c(\theta_2)'\Sigma^{*-1}c^*\right.\right.\right. \\
&\qquad\times\exp(-a^*(s-u))\gamma_+(\theta^*)c^{*\prime}du \\
&\qquad+c(\theta_2)\exp(-\alpha(\theta_2)s)\gamma_+(\theta_2)c(\theta_2)' \\
&\qquad\left.\left.\left.-c^*\exp(-a^*s)\gamma_+(\theta^*)c^{*\prime}\right\}^{\otimes 2}\right][(\sigma^{*\prime -1})^{\otimes 2}]ds\right| \\
&\quad\le Ce^{-Ct}.
\end{aligned}
$$

□

Proposition 3.26 *For any $n\in\mathbb{N}$ and $p>m_1+m_2$, it holds*

$$
E\left[\sup_{\theta_2\in\Theta_2}|\tilde{\mathbb{Y}}_n^2(\theta_2)-\mathbb{Y}^2(\theta_2)|^p\right]^{\frac{1}{p}}\le C_p\left(h+n^{-\frac{1}{2}}+t_n{}^{-\frac{1}{2}}\right).
$$

Proof By (3.30) and (3.37)

$$
\begin{aligned}
\tilde{\mathbb{Y}}_n^2(\theta_2)&=\frac{1}{2t_n}\sum_{i=1}^n\Big\{-h\Sigma^{*-1}[(c(\theta_2)\tilde{m}_{i-1}^n(\theta_2))^{\otimes 2}]+h\Sigma^{*-1}[(c^*\tilde{m}_{i-1}^n(\theta_2^*))^{\otimes 2}] \\
&\quad+\{\tilde{m}_{i-1}^n(\theta_2)'c(\theta_2)'-\tilde{m}_{i-1}^n(\theta_2^*)'c^{*\prime}\}\Sigma^{*-1}(c^*\tilde{m}_{i-1}(\theta_2^*)h+\sigma^*\Delta_j\overline{W}) \\
&\quad+(\tilde{m}_{i-1}(\theta_2^*)c^{*\prime}h+\Delta_j\overline{W}'\sigma^{*\prime})\Sigma^{*-1}\{c(\theta_2)\tilde{m}_{i-1}^n(\theta_2)-c^*\tilde{m}_{i-1}^n(\theta_2^*)\}\Big\} \\
&=\frac{1}{2t_n}\sum_{i=1}^n\Big\{-h\Sigma^{*-1}[(c(\theta_2)\tilde{m}_{i-1}^n(\theta_2)-c^*\tilde{m}_{i-1}^n(\theta_2^*))^{\otimes 2}] \\
&\quad+\{\tilde{m}_{i-1}^n(\theta_2)'c(\theta_2)'-\tilde{m}_{i-1}^n(\theta_2^*)'c^{*\prime}\}\Sigma^{*-1}\sigma^*\Delta_j\overline{W}) \\
&\quad+\Delta_j\overline{W}'\sigma^{*\prime}\Sigma^{*-1}\{c(\theta_2)\tilde{m}_{i-1}^n(\theta_2)-c^*\tilde{m}_{i-1}^n(\theta_2^*)\}\Big\}.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
&E\left[\sup_{\theta_2\in\Theta_2}|\tilde{\mathbb{Y}}_n^2(\theta_2)-\mathbb{Y}^2(\theta_2)|^p\right]^{\frac{1}{p}} \\
&\le\frac{h}{2t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\Sigma^{*-1}[(c(\theta_2)\tilde{m}_{i-1}^n(\theta_2)-c^*\tilde{m}_{i-1}^n(\theta_2^*))^{\otimes 2}]\right.\right. \\
&\quad\left.\left.-\Sigma^{*-1}[(c(\theta_2)\mu_{t_{i-1}}(\theta_2)-c^*\mu_{t_{i-1}}(\theta_2^*))^{\otimes 2}]\right|^p\right]^{\frac{1}{p}} \\
&\quad+\frac{h}{2t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\Big\{\Sigma^{*-1}[(c(\theta_2)\mu_{t_{i-1}}(\theta_2)-c^*\mu_{t_{i-1}}(\theta_2^*))^{\otimes 2}]\right.\right. \\
&\quad\left.\left.+2\mathbb{Y}^2(\theta)\Big\}\right|^p\right]^{\frac{1}{p}}
\end{aligned}
$$

$$+\frac{1}{2t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\{\tilde{m}_{i-1}^n(\theta_2)'c(\theta_2)'-\tilde{m}_{i-1}^n(\theta_2^*)'c^{*\prime}\}\Sigma^{*-1}\sigma^*\Delta_j\overline{W}\right|^p\right]^{\frac{1}{p}}$$
$$+\frac{1}{2t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\Delta_j\overline{W}'\sigma^{*\prime}\Sigma^{*-1}\{c(\theta_2)\tilde{m}_{i-1}^n(\theta_2)-c^*\tilde{m}_{i-1}^n(\theta_2^*)\right|^p\right]^{\frac{1}{p}}. \tag{3.45}$$

For the first term of this, making use of Proposition 3.23, Corollary 3.24 and (3.33), we obtain

$$\begin{aligned}
&\frac{h}{2t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\Sigma^{*-1}[(c(\theta_2)\tilde{m}_{i-1}^n(\theta_2)-c^*\tilde{m}_{i-1}^n(\theta_2^*))^{\otimes 2}]\right.\right.\\
&\left.\left.\quad-\Sigma^{*-1}[(c(\theta_2)\mu_{t_{i-1}}(\theta_2)-c^*\mu_{t_{i-1}}(\theta_2^*))^{\otimes 2}]\right|^p\right]^{\frac{1}{p}} \qquad (3.46)\\
&\quad\le C_p\frac{h}{2t_n}\times(n^{-\frac{1}{2}}+h)\times n\le C_p(n^{-\frac{1}{2}}+h),
\end{aligned}$$

just as we evaluated the fourth term of (3.40).

Now we consider the second term. Due to the proof of Proposition 3.25, $c(\theta_2)\mu_{t_i}^n(\theta_2)-c^*\mu_{t_i}^n(\theta_2^*)$ has the form

$$c(\theta_2)\mu_{t_i}^n(\theta_2)-c^*\mu_t^n(\theta_2^*)=p_i(\theta_2)+\int_0^{t_i}q_i(s;\theta_2)d\overline{W}_s$$

where

$$\begin{aligned}
p_i(\theta_2)&=\exp(-\alpha(\theta_2)t_i)m_0-\exp(-\alpha(\theta_2^*)t_i)m_0\\
&\quad+\int_0^{t_i}\{\exp(-\alpha(\theta_2)(t_i-s))\gamma_+(\theta_2)c(\theta_2)'\\
&\quad-\exp(-\alpha(\theta_2^*)(t_i-s))\gamma_+(\theta_2^*)c^{*\prime}\}\Sigma^{*-1}c^*\exp(-a^*s)m_0ds,\\
q_i(s;\theta_2)&=\int_s^{t_i}c(\theta_2)\exp(-\alpha(\theta_2)(t_i-u))\gamma_+(\theta_2)c(\theta_2)'\\
&\quad\times\Sigma^{*-1}c^*\exp(-a^*(u-s))\gamma_+(\theta^*)c^{*\prime}du\\
&\quad+c(\theta_2)\exp(-\alpha(\theta_2)(t_i-s))\gamma_+(\theta_2)c(\theta_2)'\\
&\quad-c^*\exp(-a^*(t_i-s))\gamma_+(\theta^*)c^{*\prime}.
\end{aligned}$$

Then if we set $\nu_t^i(\theta_2)=p_i(\theta_2)+\int_0^t q_i(s;\theta_2)d\overline{W}_s$, Itô's formula gives

$$\begin{aligned}
&\Sigma^{*-1}\left[\left\{c(\theta_2)\mu_{t_i}^n(\theta_2)-c^*\mu_t^n(\theta_2^*)\right\}^{\otimes 2}\right]\\
&=\Sigma^{*-1}[(\nu_{t_i}^i(\theta_2))^{\otimes 2}]=\int_0^{t_i}\Sigma^{*-1}[(\nu_{t_i}^i(\theta_2))^{\otimes 2}]\\
&=\Sigma^{*-1}[p_i(\theta_2)^{\otimes 2}]+2\int_0^{t_i}\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s\\
&\quad+\mathrm{Tr}\int_0^{t_i}\Sigma^{*-1}[q_i(s;\theta_2)^{\otimes 2}]ds
\end{aligned}$$

$$= E\left[\Sigma^{*-1}\left[\left\{c(\theta_2)\mu^n_{t_i}(\theta_2) - c^*\mu^n_t(\theta_2^*)\right\}^{\otimes 2}\right]\right]$$
$$+ 2\int_0^{t_i} \nu^i_s(\theta_2)(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s$$
$$= -2\mathbb{Y}^2(\theta_2) + 2\int_0^{t_i} \nu^i_s(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s + O(e^{-Ct_i}).$$

Therefore

$$\frac{h}{2t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\left\{\Sigma^{*-1}[(c(\theta_2)\mu^n_{t_{i-1}}(\theta_2) - c^*\mu^n_{t_{i-1}}(\theta_2^*))^{\otimes 2}] + 2\mathbb{Y}^2(\theta)\right\}\right|^p\right]^{\frac{1}{p}}$$
$$\le \frac{h}{t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\int_0^{t_i} \nu^i_s(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s\right|^p\right]^{\frac{1}{p}} + \frac{1}{2t_n}\sum_{i=1}^n Ce^{-Ct_i}h \tag{3.47}$$
$$\le \frac{h}{t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^n\int_0^{t_i} \nu^i_s(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s\right|^p\right]^{\frac{1}{p}} + \frac{C}{t_n}.$$

Now by Lemma 3.11 and the continuos differentiability of p_i and q_i, we can assume $\nu^i_t(\theta_2)$ is continuously differentiable with respect to θ_2 and almost surely

$$\partial_{\theta_2}\nu^i_t(\theta_2) = \partial_{\theta_2}p_i(\theta_2) + \int_0^t \partial_{\theta_2}q_i(s;\theta_2)ds.$$

Thus by Lemma 3.1 (2) we obtain for any $T > 0$, $p \ge 2$ and $\theta_2, \theta_2' \in \Theta_2$

$$\sup_{0\le t\le T} E\left[|\nu^i_t(\theta_2) - \nu^i_t(\theta_2')|^p\right] \le C_p|\theta_2 - \theta_2'|^p$$

and

$$\sup_{0\le t\le T} E\left[|\partial_{\theta_2}\nu^i_t(\theta_2) - \partial_{\theta_2}\nu^i_t(\theta_2')|^p\right] \le C_p|\theta_2 - \theta_2'|^p.$$

Then again by Lemma 3.11, $\int_0^{t_i} \nu^i_s(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s$ is continuously differentiable and we have almost surely

$$\partial_{\theta_2}\int_0^{t_i} \nu^i_s(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s = \int_0^{t_i} \partial_{\theta_2}\{\nu^i_s(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)\}d\overline{W}_s.$$

Therefore the Sobolev inequality gives for any $p > m_1 + m_2$

$$\begin{aligned} &E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^{n}\int_0^{t_i} \nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s\right|^p\right]^{\frac{1}{p}} \\ &= E\left[\sup_{\theta_2\in\Theta_2}\left|\int_0^{t_n}\sum_{i=1}^{n}\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)1_{[0,t_i]}(s)d\overline{W}_s\right|^p\right]^{\frac{1}{p}} \\ &\le C_p\sup_{\theta_2\in\Theta_2}E\left[\left|\int_0^{t_n}\sum_{i=1}^{n}\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)1_{[0,t_i]}(s)d\overline{W}_s\right|^p\right]^{\frac{1}{p}} \\ &\quad + C_p\sup_{\theta_2\in\Theta_2}E\left[\left|\int_0^{t_n}\sum_{i=1}^{n}\partial_{\theta_2}\{\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)\}1_{[0,t_i]}(s)d\overline{W}_s\right|^p\right]^{\frac{1}{p}}. \end{aligned} \tag{3.48}$$

Now we have $|p_t(\theta_2)| \le Ce^{-Ct_i}$, $|q_i(s;\theta_2)| \le Ce^{-C(t_i-s)}$ and hence

$$E\left[|\nu_s^i(\theta_2)|^p\right] \le C_p.$$

Thus we obtain

$$E\left[\left|\sum_{i=1}^{n}\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)1_{[0,t_i]}(s)\right|^p\right]^{\frac{1}{p}} \le \frac{C_p}{h},$$

and therefore by Lemma 3.1

$$E\left[\left|\int_0^{t_n}\sum_{i=1}^{n}\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)1_{[0,t_i]}(s)d\overline{W}_s\right|^p\right] \le \frac{C_p}{h}t_n^{\frac{p}{2}}.$$

In the same way, we obtain

$$E\left[\left|\int_0^{t_n}\sum_{i=1}^{n}\partial_{\theta_2}\{\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)\}1_{[0,t_i]}(s)d\overline{W}_s\right|^p\right]^{\frac{1}{p}} \le \frac{C_p}{h}t_n^{\frac{p}{2}}.$$

Hence by (3.48), it follows

$$E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^{n}\int_0^{t_i}\nu_s^i(\theta_2)'\Sigma^{*-1}q_i(s;\theta_2)d\overline{W}_s\right|^p\right]^{\frac{1}{p}} \le \frac{C_p}{h}t_n^{\frac{p}{2}},$$

and therefore by (3.47)

$$\begin{aligned} &\frac{h}{2t_n}E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^{n}\Sigma^{*-1}[(c(\theta_2)\mu_{t_{i-1}}^n(\theta_2)-c^*\mu_{t_{i-1}}^n(\theta_2^*))^{\otimes 2}]-\mathbb{Y}^2(\theta)\right|^p\right]^{\frac{1}{p}} \\ &\le C_p\frac{h}{t_n}\frac{t_n^{\frac{p}{2}}}{h}+\frac{C}{t_n}\le C_p\frac{1}{t_n^{\frac{1}{2}}}. \end{aligned} \tag{3.49}$$

Finally, as for the third and fourth terms in (3.45), by the Sobolev inequality, Lemma 3.1 and (3.33) it holds

$$E\left[\sup_{\theta_2\in\Theta_2}\left|\sum_{i=1}^{n}\{\tilde{m}_{i-1}^n(\theta_2)'c(\theta_2)' - \tilde{m}_{i-1}^n(\theta_2^*)'c^{*\prime}\}\Sigma^{*-1}\sigma^*\Delta_j\overline{W}\right|^p\right] \leq C_p t_n^{\frac{p}{2}}. \tag{3.50}$$

We obtain the desired result by putting (3.45), (3.46), (3.49) and (3.50) together. □

Now we set

$$\tilde{M}_j^n(\theta_2) = c(\theta_2)\tilde{m}_j^n(\theta). \tag{3.51}$$

Then by (3.32) and (3.39), we have

$$\tilde{\Gamma}_n^2 = \frac{1}{t_n}\sum_{i=1}^{n}\Big\{\Sigma^{*-1}[\partial_{\theta_2}^{\otimes 2}]\tilde{M}_i^n(\theta^*)h - \partial_{\theta_2}^2\tilde{M}_{i-1}^n(\theta^*)'\sigma^{*\prime -1}\Delta_j\overline{W} - \Delta_j\overline{W}'\sigma^{*-1}\partial_{\theta_2}^2\tilde{M}_{i-1}^n(\theta^*)\Big\}.$$

Moreover, by (2.16) and (3.44), we obtain the following results in the same way as Propositions 3.25 and 3.26:

$$E\left[\Sigma^{*-1}[\partial_{\theta_2}^{\otimes 2}]\tilde{M}_i^n(\theta^*)\right] = \Gamma^2 + O(e^{-Ct_i}) \tag{3.52}$$

$$E\left[\left|\frac{1}{t_n}\sum_{i=1}^{n}\Sigma^{*-1}[\partial_{\theta_2}^{\otimes 2}]\tilde{M}_i^n(\theta^*)h - \Gamma^2\right|^p\right] \leq C_p\left(h^p + n^{-\frac{1}{2}p} + \frac{1}{t_n^{\frac{p}{2}}}\right) \tag{3.53}$$

$$E\left[|\tilde{\Gamma}_n - \Gamma^2|^p\right] \leq C_p\left(h^p + n^{-\frac{1}{2}p} + \frac{1}{t_n^{\frac{p}{2}}}\right). \tag{3.54}$$

Proposition 3.27 *It holds*

$$\tilde{\Delta}_n^2 \xrightarrow{d} N(0, \Gamma^2).$$

Proof Since $\tilde{\Delta}_n^2$ is given by the formula (3.41), we set

$$\xi_i^n = \frac{1}{\sqrt{t_n}}\partial_{\theta_2}\tilde{M}_{i-1}^n(\theta_2^*)'\sigma^{*\prime -1}\Delta_i\overline{W}.$$

Then $(\xi_i^n)^{\otimes 2}$ is the matrix whose (i, j) entry is

$$\frac{1}{t_n}\frac{\partial}{\partial\theta_2^i}\tilde{M}_{i-1}^n(\theta_2^*)'\sigma^{*\prime -1}\Delta_i\overline{W}\frac{\partial}{\partial\theta_2^j}\tilde{M}_{i-1}^n(\theta_2^*)'\sigma^{*\prime -1}\Delta_i\overline{W} = \frac{1}{t_n}\frac{\partial}{\partial\theta_2^i}\tilde{M}_{i-1}^n(\theta_2^*)'\sigma^{*\prime -1}\Delta_i\overline{W}\Delta_i\overline{W}'\sigma^{*-1}\frac{\partial}{\partial\theta_2^j}\tilde{M}_{i-1}^n(\theta_2^*).$$

Hence it follows from (3.53)

$$\sum_{i=1}^{n}E\left[(\xi_i^n)^{\otimes 2}|\mathcal{F}_{t_{i-1}}\right] = \sum_{i=1}^{n}E\left[\Sigma^{*-1}[\partial_{\theta_2}^{\otimes 2}]\tilde{M}_{i-1}^n(\theta_2)|\mathcal{F}_{t_{i-1}}\right] \xrightarrow{P} \Gamma^2 \ (n\to\infty).$$

Moreover, we have for $\epsilon > 0$

$$\begin{aligned}
&\sum_{i=1}^{n} E[|\xi_i^n|^2 1_{\{|\xi_i^n|>\epsilon\}}|\mathcal{F}_{t_{i-1}}] \\
&\leq \sum_{i=1}^{n} E[|\xi_i^n|^4|\mathcal{F}_{t_{i-1}}]^{\frac{1}{2}} P(|\xi_i^n| > \epsilon|\mathcal{F}_{t_{i-1}})^{\frac{1}{2}} \\
&\leq \sum_{i=1}^{n} E[|\xi_i^n|^4|\mathcal{F}_{t_{i-1}}]^{\frac{1}{2}} \times \frac{1}{\epsilon^2} E[|\xi_i^n|^4|\mathcal{F}_{t_{i-1}}]^{\frac{1}{2}} \\
&= \sum_{i=1}^{n} \frac{|\sigma^{*-1}|^4}{\epsilon^2 {t_n}^2} |\partial_\theta \tilde{M}_{i-1}^n(\theta^*)|^4 E[(\Delta_i \overline{W})^4] \\
&\leq \frac{|\sigma^{*-1}|^4}{\epsilon^2 {t_n}^2} \sum_{i=1}^{n} |\partial_\theta \tilde{M}_{i-1}^n(\theta^*)|^4 h^2,
\end{aligned}$$

and hence

$$\begin{aligned}
E\left[\sum_{i=1}^{n} E[|\xi_i^n|^2 1_{\{|\xi_i^n|>\epsilon\}}|\mathcal{F}_{t_{i-1}}]\right] &\leq \sum_{i=1}^{n} \frac{|\sigma^{*-1}|^4}{\epsilon^2 {t_n}^2} E[|\partial_\theta \tilde{M}_{i-1}^n(\theta^*)|^4] h^2 \\
&\leq C_\epsilon \sum_{i=1}^{n} \frac{1}{{t_n}^2} h^2 = \frac{C_\epsilon}{n} \to 0 \ (n \to \infty).
\end{aligned}$$

Therefore we obtain the desired result by the martingale central limit theorem. □

Proposition 3.28 *For any $p > m_1 + m_2$, it holds*

$$\sup_{n\in\mathbb{N}} E\left[\sup_{\theta_2\in\Theta_2} \left|\frac{1}{t_n} \partial_{\theta_2}^3 \mathbb{H}_n^2(\theta_2)\right|^p\right]^{\frac{1}{p}} < \infty.$$

Proof Use (2.15) to divide the left-hand side into three parts, and then apply (3.13) and Proposition 3.20. □

Proof of Theorem 2.2 We set Δ_n^2, Γ_n^2 and $\mathbb{Y}_n^2$ by

$$\mathbb{Y}_n^2(\theta_2) = \frac{1}{t_n}\{\mathbb{H}_n^2(\theta_2) - \mathbb{H}_n^2(\theta_2^*)\} \tag{3.55}$$

$$\Delta_n^2 = \frac{1}{\sqrt{t_n}} \partial_\theta \mathbb{H}_n^2(\theta_2^*) \tag{3.56}$$

$$\Gamma_n^2 = -\frac{1}{t_n} \partial_\theta^2 \mathbb{H}_n^2(\theta_2^*). \tag{3.57}$$

Then by Proposition 3.21 for any $n \in \mathbb{N}$ and $p > m_1 + m_2$, it holds

$$E\left[\left|\Delta_n^2 - \tilde{\Delta}_n^2\right|^p\right]^{\frac{1}{p}} \leq C_p \left(n^{\frac{1}{2}} + h^{\frac{1}{2}} + (nh)^{-1}\right) \tag{3.58}$$

$$E\left[\left|\Gamma_n^2 - \tilde{\Gamma}_n^2\right|^p\right]^{\frac{1}{p}} \leq C_p \left(h^{\frac{1}{2}} + n^{-\frac{1}{2}} + (nh)^{-1}\right) \tag{3.59}$$

and

$$E\left[\sup_{\theta_2\in\Theta_2}\left|\mathbb{Y}_n^2(\theta_2)-\tilde{\mathbb{Y}}_n^2(\theta_2)\right|^p\right]^{\frac{1}{p}}\leq C_p\left(h^{\frac{1}{2}}+n^{-\frac{1}{2}}+(nh)^{-1}\right).$$

Together with Proposition 3.22, (3.54) and Proposition 3.26, we have for any $p > m_1+m_2$ (therefore for any $p > 0$)

$$\sup_{n\in\mathbb{N}} E\left[|\Delta_n^2|^p\right]^{\frac{1}{p}} < \infty, \tag{3.60}$$

$$\sup_{n\in\mathbb{N}} E\left[\left|t_n^{\frac{1}{2}}(\Gamma_n^2-\Gamma^2)\right|^p\right]^{\frac{1}{p}} < \infty \tag{3.61}$$

and

$$\sup_{n\in\mathbb{N}} E\left[\sup_{\theta_2\in\Theta_2}\left|t_n^{\frac{1}{2}}(\mathbb{Y}_n^2(\theta)-\mathbb{Y}^2(\theta_2))\right|^p\right]^{\frac{1}{p}} < \infty. \tag{3.62}$$

Moreover, by Proposition 3.27 and (3.58) we obtain

$$\Delta_n \xrightarrow{d} N(0,\Gamma^2). \tag{3.63}$$

Then we have proved the theorem by the assumption [A5], Proposition 3.28, (3.60)–(3.63) and Theorem 5 in Yoshida (2011). □

4 One-dimensional case

In this section, we consider the special case where $d_1 = d_2 = 1$. In this case, $a(\theta_2)$, $b(\theta_2)$, $c(\theta_2)$ and $\sigma(\theta_1)$ are scalar valued, so we set $m_1 = 1$ and $\sigma(\theta_1) = \theta_1$. Moreover, we assume $\Theta_1 \subset (\epsilon,\infty)$ for some $\epsilon > 0$. In the one-dimensional case, (1.5) is reduced to

$$\frac{c(\theta_2)^2}{{\theta_1}^2}\gamma^2+2a(\theta_2)\gamma+b(\theta_2)^2=0,$$

and the larger solution of this is

$$\gamma_+(\theta_1,\theta_2)=\frac{{\theta_1}^2a(\theta_2)}{c(\theta_2)^2}\left(\sqrt{1+\frac{b(\theta_2)^2c(\theta_2)^2}{{\theta_1}^2a(\theta_2)^2}}-1\right).$$

Thus we have

$$\alpha(\theta_1,\theta_2)=\sqrt{a(\theta_2)^2+\frac{b(\theta_2)^2c(\theta_2)^2}{{\theta_1}^2}} \tag{4.1}$$

by (2.13). Furthermore, the eigenvalues of $H(\theta_1,\theta_2)$ in Assumption [A4] is $\pm\alpha(\theta_1,\theta_2)$ and hence one can remove Assumption [A4].

As for the estimation of θ_1, one can obtain the explicit expression of $\hat{\theta}_1^n$. In fact, we have

$$\mathbb{H}_n^1(\theta_1)=-\frac{1}{2}\sum_{j=1}^n\left\{\frac{1}{h{\theta_1}^2}(\Delta_j Y)^2+2\log\theta_1\right\}$$

and hence

$$\frac{d}{d\theta_1}\mathbb{H}_n^1(\theta_1) = \frac{1}{h{\theta_1}^3}\sum_{j=1}^{n}(\Delta_j Y)^2 - \frac{n}{\theta_1}.$$

Thus we obtain the formula

$$\hat{\theta}_1^n = \left(\frac{1}{t_n}\sum_{j=1}^{n}(\Delta_j Y)^2\right)^{\frac{1}{2}}.$$

Moreover, $\mathbb{Y}_1(\theta_1)$ and Γ^1 can be written as

$$\mathbb{Y}_1(\theta_1) = -\frac{1}{2}\left(\frac{\theta_1^{*2}}{{\theta_1}^2} - 1 - 2\log\frac{\theta_1^*}{\theta_1}\right)$$

and

$$\Gamma^1 = \frac{1}{2}\left(\frac{2\theta_1^*}{\theta_1^{*2}}\right)^2 = \frac{2}{\theta_1^{*2}}.$$

Therefore noting that $x^2 - 1 - 2\log x \geq (x-1)^2$ $(x \geq 0)$ we have

$$\mathbb{Y}_1(\theta_1) \leq -\frac{1}{2}\left(\frac{\theta_1^*}{\theta_1} - 1\right)^2 \leq -\frac{(\theta_1 - \theta_1^*)^2}{2\epsilon^2}$$

and hence (2.8) holds.

As for the estimation of θ_2, since we have

$$\gamma(\theta_1,\theta_2) = \frac{{\theta_1}^2}{c(\theta_2)^2}\left\{\alpha(\theta_1,\theta_2) - a(\theta_2)\right\} \tag{4.2}$$

by (2.13), we obtain for $\alpha(\theta_2) \neq a^*$

$$\begin{aligned}
\mathbb{Y}_2(\theta_2) &= -\frac{1}{2}\int_0^\infty \left\{-\frac{\{a(\theta_2)-a^*\}(\alpha(\theta_2^*)-a^*)}{\alpha(\theta_2)-a^*}e^{-a^*s}\right.\\
&\quad \left. + \frac{\{\alpha(\theta_2)-\alpha(\theta_2^*)\}\{\alpha(\theta_2)-a(\theta_2)\}}{\alpha(\theta_2)-a^*}e^{-\alpha(\theta_2)s}\right\}^2 ds\\
&= -\frac{1}{4a^*\alpha(\theta_2)\{a^*+\alpha(\theta_2)\}}\\
&\quad \times\left[\{a^*\alpha(\theta_2) - a(\theta_2)\alpha(\theta_2^*)\}^2 + a^*\alpha(\theta_2)\left\{\alpha(\theta_2) - a(\theta_2) - \alpha(\theta_2^*) + a^*\right\}^2\right]\\
&= -\frac{a^*a(\theta)^2}{4\alpha(\theta_2)\{a^*+\alpha(\theta_2)\}}\left\{\frac{\alpha(\theta_2)}{a(\theta_2)} - \frac{\alpha(\theta_2^*)}{a(\theta_2^*)}\right\}^2\\
&\quad - \frac{1}{4a^*\alpha(\theta_2)\{a^*+\alpha(\theta_2)\}}\left\{\alpha(\theta_2) - a(\theta_2) - \alpha(\theta_2^*) + a^*\right\}^2,
\end{aligned} \tag{4.3}$$

making use of (2.6) and the identity

$$\begin{aligned}
\int_0^\infty \left(pe^{-\alpha t} + qe^{-\beta t}\right)^2 dt &= \frac{p^2}{2\alpha} + \frac{2pq}{\alpha+\beta} + \frac{q^2}{2\beta}\\
&= \frac{1}{2\alpha\beta}\left\{(\alpha q + \beta p)^2 + \alpha\beta(p-q)^2\right\},
\end{aligned}$$

where $\alpha, \beta > 0$ and $p, q \in \mathbb{R}$. Even if $\alpha(\theta_2) = a^*$, we obtain the same formula by letting $a^* \to \alpha(\theta_2)$ in (4.3).

Now we obtain a sufficient condition for (2.9) by the following proposition.

Proposition 4.1 *Assume [A3],* $\inf_{\theta_2 \in \Theta_2} |c(\theta_2)| > C$ *and*

$$|a(\theta_2) - a(\theta_2^*)| + |\alpha(\theta_2) - \alpha(\theta_2^*)| \geq C|\theta_2 - \theta_2^*|. \tag{4.4}$$

Then it holds

$$Y(\theta_2) \leq -C|\theta_2 - \theta_2^*|^2. \tag{4.5}$$

Proof Let us assume there is no constant C satisfying (4.5). Then there exists some sequence $\theta_2^{(n)} \in \Theta_2$ $(n \in \mathbb{N})$ such that

$$\left| \frac{\alpha(\theta_2^{(n)})}{a(\theta_2^{(n)})} - \frac{\alpha(\theta_2^*)}{a(\theta_2^*)} \right| \leq \frac{1}{n} |\theta_2^{(n)} - \theta_2^*|$$

and

$$\left| \alpha(\theta_2^{(n)}) - a(\theta_2^{(n)}) - \alpha(\theta_2^*) + a^* \right| \leq \frac{1}{n} |\theta_2^{(n)} - \theta_2^*|.$$

Thus if we set

$$A(\theta_2) = \alpha(\theta_2) - a(\theta_2)$$

and

$$B(\theta_2) = \frac{\alpha(\theta_2)}{a(\theta_2)},$$

it follows that

$$\begin{aligned} |a(\theta_2^{(n)}) - a(\theta_2^*)| &= \left| \frac{A(\theta_2^{(n)})}{B(\theta_2^{(n)}) - 1} - \frac{A(\theta_2^*)}{B(\theta_2^*) - 1} \right| \\ &\leq \frac{|A(\theta_2^{(n)}) - A(\theta_2^*)|}{B(\theta_2^{(n)}) - 1} + \frac{|A(\theta_2^*)||B(\theta_2^{(n)}) - B(\theta_2^*)|}{\{B(\theta_2^{(n)}) - 1\}\{B(\theta_2^*) - 1\}} \\ &\leq \frac{C}{n} |\theta_2^{(n)} - \theta_2^*|, \end{aligned}$$

noting that it holds $B(\theta_2) - 1 > C$ by the assumptions and (4.1). In the same way, we have

$$|\alpha(\theta_2^{(n)}) - \alpha(\theta_2^*)| \leq \frac{C}{n} |\theta_2^{(n)} - \theta_2^*|,$$

but these contradict (4.4). □

We similarly obtain the explicit expression of Γ^2 by (4.2):

$$\begin{aligned} \Gamma^2 = \frac{1}{\theta_1^{*2}} \int_0^\infty [\partial_{\theta_2}^{\otimes 2}] \Bigg\{ \int_0^s & c(\theta_2) \exp(-\alpha(\theta_2)u) \gamma_+(\theta_2) c(\theta_2)' \Sigma^{*-1} c^* \\ & \exp(-a^*(s-u)) \gamma_+(\theta^*) c^{*\prime} du \\ & + c(\theta_2) \exp(-\alpha(\theta_2)s) \gamma_+(\theta_2) c(\theta_2)' \Bigg\} \Bigg|_{\theta_2 = \theta_2^*} ds \end{aligned}$$

$$= \int_0^\infty \{\partial_\theta \alpha(\theta^*) e^{-\alpha^* s} - \partial_\theta a(\theta^*) e^{-a^* s}\}^{\otimes 2} ds$$
$$= \frac{\{\partial_{\theta_2}\alpha(\theta^*)\}^{\otimes 2}}{2\alpha^*} + \frac{\{\partial_{\theta_2}a(\theta^*)\}^{\otimes 2}}{2a^*}$$
$$- \frac{\partial_{\theta_2}\alpha(\theta^*) \otimes \partial_{\theta_2}a(\theta^*) + \partial_{\theta_2}a(\theta^*) \otimes \partial_{\theta_2}\alpha(\theta^*)}{\alpha^* + a^*}$$
$$= \frac{1}{2\alpha^*}\left(\partial_{\theta_2}\alpha(\theta^*) - \frac{2\alpha^*}{\alpha^* + a^*}\partial_{\theta_2}a(\theta^*)\right)^{\otimes 2} + \frac{(\alpha^*)^2 + (a^*)^2}{2(\alpha^* + a^*)a^*}\{\partial_{\theta_2}a(\theta^*)\}^{\otimes 2}.$$

Hence Γ^2 is positive definite if and only if $\{\partial_{\theta_2}a(\theta^*)\}^{\otimes 2}$ or $\{\partial_{\theta_2}\alpha(\theta^*)\}^{\otimes 2}$ is positive definite. This does not happen if $m_2 \geq 3$; in fact, one can take $x \in \mathbb{R}^{m_2}$ so that $x'\partial_{\theta_2}a(\theta^*) = x'\partial_{\theta_2}\alpha(\theta^*)$ if $m_2 \geq 3$. Thus we need to assume $m_2 \leq 2$ in the one-dimensional case.

Putting it all together, we obtain the following result.

Theorem 4.2 *Let $m_1 = 1, m_2 \leq 2, \sigma(\theta_1) = \theta_2$ and $\Theta_1 \subset (\epsilon, \infty)$ for some $\epsilon > 0$. Moreover, we assume [A1], [A2] and the following conditions:*

*[**B1**]*

$$\inf_{\theta_2 \in \Theta_2} a(\theta_2) > 0$$
$$\inf_{\theta_2 \in \Theta_2} |b(\theta_2)| > 0$$
$$\inf_{\theta_2 \in \Theta_2} |c(\theta_2)| > 0$$

*[**B2**] For any $\theta_1 \in \Theta_1$ and $\theta_2, \theta_2^* \in \Theta_2$,*

$$|a(\theta_2, \theta_1) - a(\theta_2^*, \theta_1)| + |\alpha(\theta_2, \theta_1) - \alpha(\theta_2^*, \theta_1)| \geq C_{\theta_1}|\theta_2 - \theta_2^*|.$$

*[**B3**] For any $\theta \in \Theta$, $\{\partial_{\theta_2}a(\theta)\}^{\otimes 2}$ or $\{\partial_{\theta_2}\alpha(\theta)\}^{\otimes 2}$ is positive definite.*

(1) If we set

$$\hat{\theta}_1^n = \left(\frac{1}{t_n}\sum_{j=1}^n (\Delta_j Y)^2\right)^{\frac{1}{2}},$$

then for every $p > 0$ and any continuous function $f : \mathbb{R}^d \to \mathbb{R}$ such that

$$\limsup_{|x|\to\infty} \frac{|f(x)|}{|x|^p} < \infty,$$

it holds that

$$E[f(\sqrt{n}(\hat{\theta}_1^n - \theta_1^*))] \to E[f(Z)] \ (n \to \infty),$$

where $Z \sim N\left(0, \frac{\theta_1^{*2}}{2}\right)$.

In particular, it holds that

$$\sqrt{n}(\hat{\theta}_1^n - \theta_1^*) \xrightarrow{d} N\left(0, \frac{\theta_1^{*2}}{2}\right) \ (n \to \infty).$$

(2) Let us define $\gamma_+(\theta_1,\theta_2)$ and $\alpha(\theta_1,\theta_2)$ by

$$\gamma_+(\theta_1,\theta_2)=\frac{{\theta_1}^2 a(\theta_2)}{c(\theta_2)^2}\left(\sqrt{1+\frac{b(\theta_2)^2c(\theta_2)^2}{{\theta_1}^2a(\theta_2)^2}}-1\right).$$

and

$$\alpha(\theta_1,\theta_2)=\sqrt{a(\theta_2)^2+\frac{b(\theta_2)^2c(\theta_2)^2}{{\theta_1}^2}},$$

respectively, and set

$$\hat{m}_i^n(\theta_2;m_0)=e^{-\alpha(\hat{\theta}_1^n,\theta_2)t_i}m_0 + \frac{1}{(\hat{\theta}_1^n)^2}\sum_{j=1}^{i}e^{-\alpha(\hat{\theta}_1^n,\theta_2)(t_i-t_{j-1})}\gamma_+(\hat{\theta}_1^n,\theta_2)c(\theta_2)\Delta_jY,$$

$$\mathbb{H}_n^2(\theta_2;m_0)=\frac{1}{2}\sum_{i=1}^{n}\left\{-h(c(\theta_2)\hat{m}_{j-1}^n(\theta_2))^2+2\hat{m}_{j-1}^n(\theta_2)c(\theta_2)\Delta_jY\right\},$$

and

$$\Gamma^2=\frac{\{\partial_{\theta_2}\alpha(\theta^*)\}^{\otimes 2}}{2\alpha^*}+\frac{\{\partial_{\theta_2}a(\theta^*)\}^{\otimes 2}}{2a(\theta^*)} - \frac{\partial_{\theta_2}\alpha(\theta^*)\otimes\partial_{\theta_2}a(\theta^*)+\partial_{\theta_2}a(\theta^*)\otimes\partial_{\theta_2}\alpha(\theta^*)}{\alpha(\theta^*)+a(\theta^*)},$$

where $m_0\in\mathbb{R}^{d_1}$ is an arbitrary initial value.

Then, if $\hat{\theta}_2^n=\hat{\theta}_2^n(m_0)$ is a random variable satisfying

$$\mathbb{H}_n^2(\hat{\theta}_2^n)=\max_{\theta_2\in\overline{\Theta}_2}\mathbb{H}_n^2(\theta_2)$$

for each $n\in\mathbb{N}$, then for any $p>0$ and continuous function $f:\mathbb{R}^d\to\mathbb{R}$ such that

$$\limsup_{|x|\to\infty}\frac{|f(x)|}{|x|^p}<\infty,$$

it holds that

$$E[f(\sqrt{t_n}(\hat{\theta}_2^n-\theta_2^*))]\to E[f(Z)]\ (n\to\infty),$$

where $Z\sim N(0,(\Gamma^2)^{-1})$.

In particular, it holds that

$$\sqrt{t_n}(\hat{\theta}_2^n-\theta_2^*)\xrightarrow{d}N(0,(\Gamma^2)^{-1})\ (n\to\infty).$$

5 Proof of Theorem 2.1

In this section, we prove Theorem 2.1, which can be proved in the same way as the diffusion case in Yoshida (2011).

Lemma 5.1 *For every $p \geq 2$ and $A \in M_{d_1}(\mathbb{R})$, it holds*

$$E\left[\left|\sum_{j=1}^{n} A[(\Delta_j Y)^{\otimes 2}] - \sum_{j=1}^{n} A[(\sigma^* \Delta_j W^2)^{\otimes 2}]\right|^p\right]^{\frac{1}{p}} \leq C_p |A| (nh^2 + n^{\frac{1}{2}} h^{\frac{3}{2}}).$$

Proof First we get

$$\begin{aligned}
&E\left[\left|\sum_{j=1}^{n} A[(\Delta_j Y)^{\otimes 2}] - \sum_{j=1}^{n} A[(\sigma^* \Delta_j W^2)^{\otimes 2}]\right|^p\right]^{\frac{1}{p}} \\
&= E\left[\left|\sum_{j=1}^{n} A\left[\left(c^* \int_{t_{j-1}}^{t_j} X_t ds + \sigma^* \Delta_j W^2\right)^{\otimes 2}\right] - \sum_{j=1}^{n} A[(\sigma^* \Delta_j W^2)^{\otimes 2}]\right|^p\right]^{\frac{1}{p}} \\
&\leq E\left[\left|\sum_{j=1}^{n} A\left[\left(c^* \int_{t_{j-1}}^{t_j} X_t ds\right)^{\otimes 2}\right]\right|^p\right]^{\frac{1}{p}} \\
&\quad + E\left[\left|\sum_{j=1}^{n} A\left[c^* \int_{t_{j-1}}^{t_j} X_t ds, \sigma^* \Delta_j W^2\right]\right|^p\right]^{\frac{1}{p}} \\
&\quad + E\left[\left|\sum_{j=1}^{n} A\left[\sigma^* \Delta_j W^2, c^* \int_{t_{j-1}}^{t_j} X_t ds\right]\right|^p\right]^{\frac{1}{p}}.
\end{aligned}$$

For the first term of the rightest-hand side, we obtain by Lemmas 3.1 and 3.3

$$\begin{aligned}
&E\left[\left|\sum_{j=1}^{n} A\left[\left(c^* \int_{t_{j-1}}^{t_j} X_t ds\right)^{\otimes 2}\right]\right|^p\right]^{\frac{1}{p}} \\
&\leq |A||c^*| \sum_{j=1}^{n} E\left[\left|\int_{t_{j-1}}^{t_j} X_t ds\right|^{2p}\right]^{\frac{1}{p}} \\
&\leq |A||c^*| \sum_{j=1}^{n} h^{2p-1} \left(\int_{t_{j-1}}^{t_j} E[|X_t|^{2p}]\right)^{\frac{1}{p}} ds \\
&\leq C_p |A| nh^2.
\end{aligned}$$

For the second and third terms, it holds

$$\begin{aligned}
&E\left[\left|\sum_{j=1}^{n} A\left[c^* \int_{t_{j-1}}^{t_j} X_t ds, \sigma^* \Delta_j W^2\right]\right|^p\right]^{\frac{1}{p}} \\
&\leq E\left[\left|\sum_{j=1}^{n} A\left[c^* \int_{t_{j-1}}^{t_j} (X_t - X_{t_{j-1}}) ds, \sigma^* \Delta_j W^2\right]\right|^p\right]^{\frac{1}{p}}
\end{aligned}$$

$$
\begin{aligned}
&+ E\left[\left|\sum_{j=1}^{n} A\left[c^* X_{t_{j-1}} h, \sigma^* \Delta_j W^2\right]\right|^p\right]^{\frac{1}{p}} \\
&\leq |A||c^*||\sigma^*| \sum_{j=1}^{n} E\left[\left|\int_{t_{j-1}}^{t_j} (X_t - X_{t_{j-1}}) ds\right|^p |\Delta_j W^2|^p\right]^{\frac{1}{p}} \\
&+ hE\left[\left|\sum_{j=1}^{n} X'_{t_{j-1}} c^{*\prime} A \sigma^* \Delta_j W^2\right|^p\right]^{\frac{1}{p}} \\
&\leq |A||c^*||\sigma^*| \sum_{j=1}^{n} E\left[\left|\int_{t_{j-1}}^{t_j} (X_t - X_{t_{j-1}}) ds\right|^{2p}\right]^{\frac{1}{2p}} E\left[|\Delta_j W^2|^{2p}\right]^{\frac{1}{2p}} \\
&+ hE\left[\left|\sum_{j=1}^{n} X'_{t_{j-1}} c^{*\prime} A \sigma^* \Delta_j W^2\right|^p\right]^{\frac{1}{p}} \\
&\leq C_p |A| \sum_{j=1}^{n} \left(h^{2p-1} \int_{t_{j-1}}^{t_j} E[|X_u - X_{t_{j-1}}|^{2p}] du\right)^{\frac{1}{2p}} h^{\frac{1}{2}} \\
&+ C_p |A| h \left((nh)^{\frac{p}{2}-1} \sum_{j=1}^{n} E[|X_{t_{j-1}}|^p] h\right)^{\frac{1}{p}} \\
&\leq C_p |A| (nh^2 + n^{\frac{1}{2}} h^{\frac{3}{2}}).
\end{aligned}
$$

Therefore we get the desired result. □

Lemma 5.2 *Let* $A_k \in M_{d_1}(\mathbb{R})$ $(k = 1, 2, \cdots, d)$, $A = (A_1, \cdots, A_d)$ *and*

$$
M_n(A) = \sum_{j=1}^{n} \left\{\frac{1}{h} A[(\Delta_j W^2)^{\otimes 2}] - \mathrm{Tr} A\right\}.
$$

Then it holds that

$$
E[|M_n(A)|^p] \leq C_p |A| \sqrt{n} \tag{5.1}
$$

and

$$
\frac{1}{\sqrt{n}} M_n \xrightarrow{d} N(0, 2(\mathrm{Tr} A)^{\otimes 2}) \ (n \to \infty). \tag{5.2}
$$

Proof On account of

$$
E\left[\frac{1}{h} A[(\Delta_j W^2)^{\otimes 2}] - \mathrm{Tr} A \middle| \mathcal{F}_{t_{j-1}}\right] = \frac{1}{h} E\left[A[(\Delta_j W^2)^{\otimes 2}]\right] - \mathrm{Tr} A = 0,
$$

$\{A[(\Delta_j W^2)^{\otimes 2}]/h - \mathrm{Tr}A\}_j$ is a martingale difference sequence with respect to $\{\mathcal{F}_{t_j}\}$. Hence the Burkholder inequality gives

$$\begin{aligned}
E[|M_n|^p]^{\frac{p}{2}} &\leq C_p E\left[\left|\sum_{j=1}^n \frac{1}{h}A[(\Delta_j W^2)^{\otimes 2}] - \mathrm{Tr}A\right|^{2p}\right]^{\frac{1}{2p}} \\
&\leq C_p \sum_{j=1}^n E\left[\left|\frac{1}{h}A[(\Delta_j W^2)^{\otimes 2}] - \mathrm{Tr}A\right|^{2p}\right]^{\frac{1}{2p}} \\
&\leq C_p n\left\{|A|^{2p} E\left[\left|\frac{1}{h}W_h^2\right|^{4p} + |A|^2 p\right]\right\} \\
&\leq C_p |A| n,
\end{aligned}$$

and we obtain (5.1).
Moreover, due to the fact that $\{A[(\Delta_j W^2)^{\otimes 2}]/h - \mathrm{Tr}A\}_j$ is independent and identically distributed, we have

$$\begin{aligned}
&E\left[\left(\frac{1}{h}A[(\Delta_j W^2)^{\otimes 2}] - \mathrm{Tr}A\right)^{\otimes 2}\right] \\
&= \frac{1}{h^2}E\left[\{A[(\Delta_j W^2)^{\otimes 2}]\}^{\otimes 2}\right] - \frac{1}{h}E[A[(\Delta_j W^2)^{\otimes 2}]]'\mathrm{Tr}A \\
&\quad - (\mathrm{Tr}A)'\frac{1}{h}E[A[(\Delta_j W^2)^{\otimes 2}]] + (\mathrm{Tr}A)^{\otimes 2} \\
&= 3(\mathrm{Tr}A)^{\otimes 2} - 2(\mathrm{Tr}A)^{\otimes 2} + (\mathrm{Tr}A)^{\otimes 2} = 2(\mathrm{Tr}A)^{\otimes 2}.
\end{aligned}$$

Thus we obtain (5.2). □

By Lemmas 5.1 and 5.2, we get the following lemma.

Lemma 5.3 *Let* $A_k \in M_{d_1}(\mathbb{R})$ $(k = 1, 2, \cdots, d)$, $A = (A_1, \cdots, A_d)$ *and*

$$L_n(A) = \sum_{j=1}^n \left\{\frac{1}{h}A[(\Delta_j Y)^{\otimes 2}] - \mathrm{Tr}A\Sigma\right\}.$$

Then it holds that

$$E[|L_n(A)|^p] \leq C_p|A|(nh + n^{\frac{1}{2}}h^{\frac{1}{2}} + n^{\frac{1}{2}}) \tag{5.3}$$

and

$$\frac{1}{\sqrt{n}}L_n \xrightarrow{d} N(0, 2(\mathrm{Tr}A\Sigma)^{\otimes 2})\ (n \to \infty). \tag{5.4}$$

Lemma 5.4 *For every* $p > 0$, *it holds*

$$\sup_{n\in\mathbb{N}} E\left[\left|\frac{1}{n}\sup_{\theta_1\in\Theta_1} \partial_{\theta_1}^3 \mathbb{H}_n^1(\theta_1)\right|^p\right] < \infty$$

Proof It is enough to prove the inequality for sufficiently large p. By Lemmas 5.2 and 5.1 and Assumptions [A2] and [A4] we get

$$\begin{aligned}
&E\left[\left|\frac{1}{n}\partial_{\theta_1}^3\mathbb{H}_n^1(\theta_1)\right|^p\right]^{\frac{1}{p}} \\
&= E\left[\left|\frac{1}{2n}\sum_{j=1}^n\left\{\frac{1}{h}\partial_{\theta_1}^3\Sigma^{-1}(\theta_1)[(\Delta_j Y)^{\otimes 2}] + \partial_{\theta_1}^3\log\det\Sigma(\theta_1)\right\}\right|^p\right]^{\frac{1}{p}} \\
&\leq E\left[\left|\frac{1}{2nh}\sum_{j=1}^n\left\{\partial_{\theta_1}^3\Sigma^{-1}(\theta_1)[(\Delta_j Y)^{\otimes 2} - \mathrm{Tr}\partial_{\theta_1}^3\Sigma^{-1}(\theta_1)\right\}\right|^p\right]^{\frac{1}{p}} \\
&\quad + \frac{1}{2}\left\{|\mathrm{Tr}\partial_{\theta_1}^3\Sigma^{-1}(\theta_1)| + |\partial_{\theta_1}^3\log\det\Sigma(\theta_1)|\right\} \\
&\leq C_p|\partial_{\theta_1}^3\Sigma^{-1}(\theta_1)|(h + n^{-\frac{1}{2}}h^{\frac{1}{2}} + n^{-\frac{1}{2}}) \\
&\quad + \frac{1}{2}\left\{|\mathrm{Tr}\partial_{\theta_1}^3\Sigma^{-1}(\theta_1)| + |\partial_{\theta_1}^3\log\det\Sigma(\theta_1)|\right\} \\
&\leq C_p,
\end{aligned}$$

and similarly

$$E\left[\left|\frac{1}{n}\partial_{\theta_1}^4\mathbb{H}_n^1(\theta_1)\right|^p\right]^{\frac{1}{p}} \leq C_p.$$

Thus we get the desired result for $p > d_1$ by the Sobolev inequality. □

Proof of Theorem 2.1 Let

$$\Delta_n^1 = \frac{1}{\sqrt{n}}\partial_{\theta_1}\mathbb{H}_n^1(\theta_1^*),$$
$$\Gamma_n^1 = -\frac{1}{n}\partial_{\theta_1}^2\mathbb{H}_n^1(\theta_1^*)$$

and

$$\mathbb{Y}_n^1(\theta_1) = \frac{1}{n}\{\mathbb{H}_n^1(\theta_1) - \mathbb{H}_n^1(\theta_1^*)\}.$$

Then

$$\begin{aligned}
\Delta_n^1 &= -\frac{1}{2\sqrt{n}}\sum_{j=1}^n\left\{\frac{1}{h}\partial_{\theta_1}\Sigma^{-1}(\theta_1^*)[(\Delta_j Y)^{\otimes 2}] + \frac{\partial_{\theta_1}\det\Sigma(\theta_1^*)}{\det\Sigma^*}\right\} \\
&= \frac{1}{2\sqrt{n}}\sum_{j=1}^n\left\{\frac{1}{h}\Sigma^{*-1}\partial_{\theta_1}\Sigma(\theta_1^*)\Sigma^{*-1}[(\Delta_j Y)^{\otimes 2}] - \mathrm{Tr}\Sigma^{*-1}\partial_{\theta_1}\Sigma(\theta_1^*)\right\},
\end{aligned}$$

and hence by Lemma 5.3, we obtain

$$E[|\Delta_n^1|^p] \leq C_p \tag{5.5}$$

and

$$\Delta_n^1 \xrightarrow{d} N(0, (\Gamma^1)^{-1})\ (n \to \infty). \tag{5.6}$$

By the same lemma, it follows that

$$\begin{aligned} &E\left[\left|n^{\frac{1}{2}}(\Gamma_n^1-\Gamma^1)\right|^p\right]^{\frac{1}{p}} \\ &= E\left[\left|n^{\frac{1}{2}}\left[\frac{1}{n}\sum_{j=1}^n\left\{\frac{1}{h}\partial_{\theta_1}^2\Sigma^{-1}(\theta_1^*)[(\Delta_j Y)^{\otimes 2}]+\partial_{\theta_1}^2\log\det\Sigma(\theta_1^*)\right\}-\Gamma^1\right]\right|^p\right]^{\frac{1}{p}} \\ &\le C_p(n^{\frac{1}{2}}h^2+h^{\frac{3}{2}}+1)=O(1), \end{aligned} \tag{5.7}$$

noting that

$$\begin{aligned} &\mathrm{Tr}\partial_{\theta_1}^2\Sigma^{-1}(\theta_1^*)\Sigma^* \\ &= \mathrm{Tr}\left\{2\{\Sigma^{*-1}\partial_{\theta_1}\Sigma(\theta_1^*)\}^{\otimes 2}-\Sigma^{-1}\frac{\partial}{\partial\theta_1^i}\frac{\partial}{\partial\theta_1^j}\Sigma(\theta_1^*)\right\} \end{aligned}$$

is equal to $-\partial_{\theta_1}^2\log\det\Sigma(\theta_1^*)+\Gamma^1$.
Moreover, we can show

$$\begin{aligned} &\sup_{\theta_1\in\Theta_1} E\left[(n^{\frac{1}{2}}|\mathbb{Y}_n^1(\theta_1)-\mathbb{Y}^1(\theta_1)|)^p\right]^{\frac{1}{p}} \\ &= \sup_{\theta_1\in\Theta_1} E\left[\left|-\frac{1}{2\sqrt{n}}\sum_{j=1}^n\left\{\frac{1}{h}\{\Sigma^{-1}(\theta_1)-\Sigma^{-1}(\theta_1^*)\}[(\Delta_j Y)^{\otimes 2}]\right.\right.\right. \\ &\qquad \left.\left.\left. -\mathrm{Tr}\{\Sigma(\theta_1)^{-1}-I\}\right\}\right|^p\right]^{\frac{1}{p}} \\ &\le C_p(n^{\frac{1}{2}}h+h+1)=O(1) \end{aligned}$$

and in the same way

$$\sup_{\theta_1\in\Theta_1} E\left[(n^{\frac{1}{2}}|\partial_{\theta_1}\{\mathbb{Y}_n^1(\theta_1)-\mathbb{Y}^1(\theta_1)\}|)^p\right]^{\frac{1}{p}}=O(1).$$

Thus by the Sobolev inequality, it holds for $p>d_1$

$$E\left[\sup_{\theta_1\in\Theta_1}(n^{\frac{1}{2}}|\mathbb{Y}_n^1(\theta_1)-\mathbb{Y}^1(\theta_1)|)^p\right]^{\frac{1}{p}}=O(1). \tag{5.8}$$

Then we have proved the theorem by the assumption [A5], Lemma 5.4, (5.5), (5.6), (5.7), (5.8) and Theorem 5 in Yoshida (2011). □

6 Simulations

In this section, we will verify the results of the previous sections by computational simulations. We set $d_1=d_2=1$ and consider the equations

$$\begin{cases} dX_t=-aX_tdt+bdW_t^1 \\ dY_t=X_tdt+\sigma dW_t^2 \end{cases}$$

Table 1 The summary of the simulation results with $n = 10^6, h = 0.0001$

	m_0	Remove	σ	a	b
True value (standard error)	$X_0 = 0$		0.02	1.5	0.3
			(1.414×10^{-5})	(0.2115)	(0.01324)
Estimation (i)	0	0		1.495	0.3011
				(0.2123)	(0.01338)
Estimation (ii)	1	0	0.02007	1.715	0.3249
			(1.640×10^{-5})	(0.2452)	(0.01338)
Estimation (iii)	1	100		1.706	0.3239
				(0.2436)	(0.01333)
Estimation (iv)	1	1000		1.535	0.3059
				(0.2177)	(0.01304)

with $X_0 = Y_0 = 0$, where we want to estimate $\theta_1 = \sigma$ and $\theta_2 = (a, b)$ from observations of Y_t.

We generated sample data Y_{t_i} $(i = 0, 1, \cdots, n)$ with $n = 10^6$, $h = 0.0001$ and true value of $(a^*, b^*, \sigma) = (1.5, 0.3, 0.02)$, and performed 10000 Monte Carlo replications. Recall that for the estimation of θ_2, we first calculate $\hat{m}$ by (2.14), and we have to choose its initial value m_0. Although we proved that Theorem 2.2 holds for an arbitrary choice of m_0, this value is a substitute for $E[X_0|Y_0]$, and thus in practice the choice of m_0 is very important as will be shown in the following. Also, the choice of the number of terms to drop, which is explained below, is relevant.

Taking these facts into account, we calculated the estimator of θ_2 in the following ways for each simulated data.

Estimation (i) $m_0 = 0$.

Estimation (ii) $m_0 = 1$.

Estimation (iii) $m_0 = 1$ and removed first 100 terms of $\hat{m}_i^n$; i.e. we replaced $\mathbb{H}_n^2(\theta_2; m_0)$ with

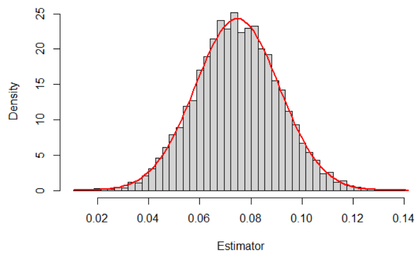
$$\frac{1}{2}\sum_{i=101}^{n}\left\{-h(c(\theta_2)\hat{m}_{j-1}^n(\theta_2))^2 + 2\hat{m}_{j-1}^n(\theta_2)c(\theta_2)\Delta_j Y\right\}.$$

Estimation (iv) $m_0 = 1$ and removed first 1000 terms of $\hat{m}_i^n$; i.e. we replaced $\mathbb{H}_n^2(\theta_2; m_0)$ with

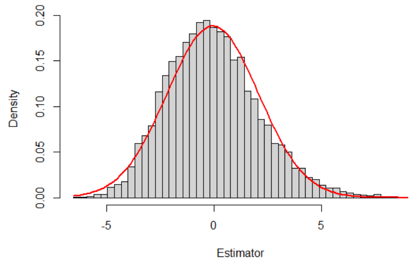
$$\frac{1}{2}\sum_{i=1001}^{n}\left\{-h(c(\theta_2)\hat{m}_{j-1}^n(\theta_2))^2 + 2\hat{m}_{j-1}^n(\theta_2)c(\theta_2)\Delta_j Y\right\}.$$

Table 1 shows the means and standard deviations of each estimators, and one can observe asymptotic normalities of them in Fig. 1. Note that the difference of four estimations are not relevant to the estimation of θ_1.

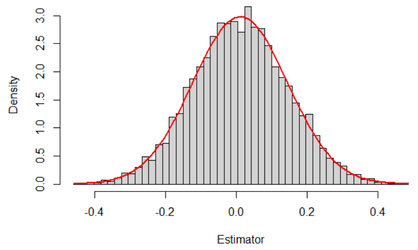
We can see from the results of first estimation and the corresponding histograms that the estimators behaved in accordance with the theory. At the same time, the following results shows that the wrong value of m_0 can significantly impact the accuracy of our estimator, but it can be improved by leaving out first several terms of $\hat{m}_i^n$. One can figure out the reason why this modification of removing first terms works by looking at Fig. 2; it shows that $\hat{m}_i^n(\theta^*)$



(a) Estimated values of $\sqrt{n}(\sigma - \sigma^*)$.



(b) Estimated values of $\sqrt{t_n}(a - a^*)$.



(c) Estimated values of $\sqrt{t_n}(b - b^*)$.

Fig. 1 Histograms of normalized estimators in Estimation (i). The red lines are the density of the normal distributions with means and standard deviations of data

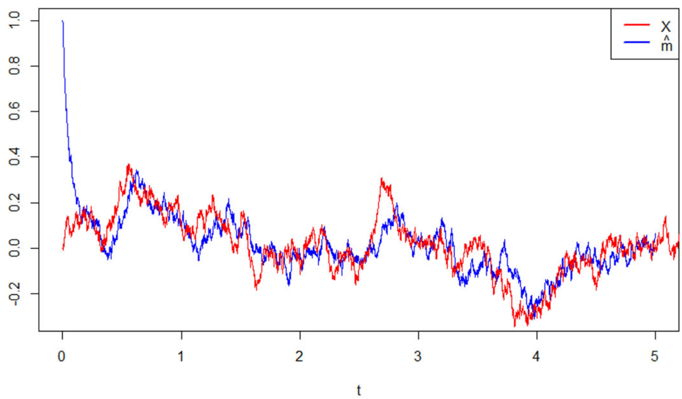


Fig. 2 A path of X_t and $\hat{m}_i^n(\theta^*)$ with $m_0 = 1$

with $m_0 = 1$ well approximate X_{t_i} except at the beginning. However, according to the result of Estimations (iii) and (iv), removing first 100 terms is not enough to improve estimators, whereas significant improvement is made by removing 1,000 terms. Thus, it is important choose m_0 which is closer to $E[X_0|Y_0]$ when some information of X_0 is available. Note that if X_0 and Y_0 are independent, then the best choice of m_0 is $E[X_0|Y_0] = E[X_0]$.

On the other hand, when you have no information of X_0, it will be interesting to consider the way to decide how many terms of $\hat{m}_i^n(\theta)$ should be removed. One possible way is to increase the removing number and calculate estimators until they converge. However, this method is computationally intensive, and more theoretical way will be needed.

The data and script that supports the findings of this study are available in the supplementary material of this article.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11203-023-09288-w.

Acknowledgements The author is grateful to N. Yoshida for important advice and useful discussions. I also thank Y. Koike for his help with accomplishing the computational simulations.

Funding Open access funding provided by The University of Tokyo.

Declarations

Conflict of interest The corresponding author states that there is no conflict of interest.

References

Brouste A, Fukasawa M, Hino H, Iacus SM, Kamatani K, Koike Y, Masuda H, Nomura R, Ogihara T, Shimuzu Y, Uchida M, Yoshida N (2014) The Yuima project: a computational framework for simulation and inference of stochastic differential equations. J Stat Softw 57(4):1–51. https://doi.org/10.18637/jss.v057.i04

Coppel W (1974) Matrix quadratic equations. Bull Aust Math Soc 10(3):377–401

Gloter A, Yoshida N (2021) Adaptive estimation for degenerate diffusion processes. Electron J Stat 15(1):1424–1472

Haber HE (2018) Notes on the matrix exponential and logarithm

Ibragimov IA, Has' Minskii RZ (1981) Statistical estimation: asymptotic theory. Springer, Berlin

Kallianpur G (1980) Stochastic filtering theory. Springer, New York

Kamatani K, Uchida M (2015) Hybrid multi-step estimators for stochastic differential equations based on sampled data. Stat Inference Stoch Process 18(2):177–204

Kutoyants YA (1994) Identification of dynamical systems with small noise. Springer, Berlin

Kutoyants YA (2004) Statistical inference for ergodic diffusion processes. Springer, Berlin

Kutoyants YA (2019) On parameter estimation of hidden ergodic Ornstein–Uhlenbeck process. Electron J Stat 13(2):4508–4526. https://doi.org/10.1214/19-EJS1631

Kutoyants YA (2019) On parameter estimation of the hidden Ornstein–Uhlenbeck process. J Multivar Anal 169:248–263. https://doi.org/10.1016/j.jmva.2018.09.008

Leipnik R (1985) A canonical form and solution for the matrix Riccati differential equation. ANZIAM J 26(3):355–361

Leoni G (2017) A first course in Sobolev spaces. American Mathematical Society, Philadelphia

Liptser RS, Shiriaev AN (2001) Statistics of random processes—I. General theory. Springer, Berlin

Liptser RS, Shiriaev AN (2001) Statistics of random processes—II. Applications. Springer, Berlin

Masuda H (2019) Non-Gaussian quasi-likelihood estimation of sde driven by locally stable Lévy process. Stoch Process Appl 129(3):1013–1059

Nakakita SH, Kaino Y, Uchida M (2021) Quasi-likelihood analysis and bayes-type estimators of an ergodic diffusion plus noise. Ann Inst Stat Math 73(1):177–225

Ogihara T, Yoshida N (2011) Quasi-likelihood analysis for the stochastic differential equation with jumps. Stat Inference Stoch Process 14(3):189–229

Shimizu Y, Yoshida N (2006) Estimation of parameters for diffusion processes with jumps from discrete observations. Stat Inference Stoch Process 9(3):227–277

Sontag ED (2013) Mathematical control theory: deterministic finite dimensional systems, vol 6. Springer, Berlin

Sørensen H (2002) Estimation of diffusion parameters for discretely observed diffusion processes. Bernoulli 66:491–508

Uchida M, Yoshida N (2012) Adaptive estimation of an ergodic diffusion process based on sampled data. Stoch Process Appl 122(8):2885–2924

Yoshida N (1992) Estimation for diffusion processes from discrete observation. J Multivar Anal 41(2):220–242. https://doi.org/10.1016/0047-259X(92)90068-Q

Yoshida N (2011) Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. Ann Inst Stat Math 63(3):431–479

Zhou K, Doyle J, Glover K (1996) Robust and optimal control. Feher/Prentice Hall Digital and Prentice Hall, Englewood Cliffs