Convergence rate of EM algorithm for SDEs with low regular drifts *

Jianhai Bao[a)], Xing Huang[a)], Shaoqin Zhang[b)]

[a)]Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

jianhaibao@tju.edu.cn

[a)]Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

xinghuang@tju.edu.cn

[b)]School of Statistics and Mathematics, Central University of Finance and Economics, Beijing 100081, China

zhangsq@cufe.edu.cn

May 20, 2022

Abstract

In this paper, by employing Gaussian type estimate of heat kernel, we establish Krylov's estimate and Khasminskii's estimate for EM algorithm. As applications, by taking Zvonkin's transformation into account, we investigate convergence rate of EM algorithm for a class of multidimensional SDEs with low regular drifts, which need not to be piecewise Lipschitz.

AMS subject Classification: 60H10, 34K26, 39B72.
Keywords: Zvonkin's transform, Euler-Maruyama approximation, low regular drift, Krylov's estimate

1 Introduction and main results

Strong/weak convergence of numerical schemes for stochastic differential equations (SDEs for short) with regular coefficients have been investigated considerably; see monographs e.g. [12]. As we know, (forward) Euler-Maruyama (EM for abbreviation) is the simplest algorithm to discretize SDEs whose coefficients are of linear growth. Whereas, EM scheme is invalid once the coefficients of SDEs involved are of nonlinear growth; see e.g. [10] for some illustrative counterexamples. Whence the other variants of EM scheme were designed to treat SDEs with non-globally Lipschitz condition; see [7, 8] for backward EM scheme, [2, 11] as for tamed EM algorithm, and [18] concerning truncated EM method, to name a few. Nowadays, convergence analysis of numerical algorithms for SDEs with irregular coefficients also receives much attention; see e.g. [5] for SDEs with Hölder continuous diffusions via the Yamada-Watanabe approximation approach, [31] for SDEs whose drift terms are Hölder continuous with the aid of Meyer-Tanaka's formula and estimates on local times,

*Supported in part by NNSFC (11771326, 11801406, 11831014, 11901604,12071340).

and [1, 24] for SDEs whose drifts are Hölder(-Dini) continuous by regularities of the corresponding backward Kolmogrov equations. In the past few years, numerical approximations of SDEs with discontinuous drifts have also gained a lot of interest; see, for instance, [6, 16, 17, 20, 21]. Up to now, most of the existing literatures above on strong approximations of SDEs with discontinuous drift coefficients are implemented under the additional assumption that the drift term is piecewise Lipschitz continuous.

Since the pioneer work of Zvonkin [34], the wellposedness for SDEs with irregular coefficients has been developed greatly in different manners; see e.g. [3, 4, 14, 30, 32] for SDEs driven by Brownian motions or jump processes, and e.g. [9, 25] for McKean-Vlasov (or distribution-dependent or mean-field) SDEs. So far, there also exist a number of literatures upon numerical simulations of SDEs with low regularity. In particular, [22] is concerned with strong convergence rate of EM scheme for SDEs with irregular coefficients, where the one-sided Lipschitz condition is imposed on the drift term. Subsequently, the one-sided Lipschitz condition put in [22] was dropped in [23] whereas the 1-dimensional SDEs are barely concerned. At this point, our goal in this paper has been evident. More precisely, motivated by the previous literatures, in this paper we aim to investigate convergence rate of EM for multidimensional SDEs with low regularity, where the drift terms need not to be piecewise Lipschitz continuity imposed in e.g. [6, 16, 17, 20, 21].

Now we consider the following SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad t \geq 0, \quad X_0 = x \in \mathbb{R}^d, \tag{1.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^m$, and $(W_t)_{t\geq 0}$ is an m-dimensional Brownian motion on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$. For the drift b and the diffusion σ, we assume

(A1) $\|b\|_\infty := \sup_{x\in\mathbb{R}^d} |b(x)| < \infty$ and there exists a constant $p > \frac{d}{2}$ such that $|b|^2 \in L^p(\mathbb{R}^d)$, the usual L^p-space on $\mathbb{R}^d$;

(A2) there exist constants $\gamma \geq 2, \alpha_\gamma > 0$ and a continuous decreasing function $\phi_\gamma : (0,\infty) \to (0,\infty)$ with $\int_0^l \phi_\gamma(s)\mathrm{d}s < \infty$ for arbitrary $l > 0$ such that

$$\frac{1}{s^{\frac{d}{2}}} \int_{\mathbb{R}^d} |b(x+y) - b(x+z)|^\gamma \mathrm{e}^{-\frac{1}{s}|x|^2} \mathrm{d}x \leq \phi_\gamma(s)|y-z|^{\alpha_\gamma}, \qquad y, z \in \mathbb{R}^d, \quad s > 0;$$

(A3) There exist constants $\breve{\lambda}_0, \hat{\lambda}_0, L_0 > 0$ such that

$$\breve{\lambda}_0|\xi|^2 \leq \langle(\sigma\sigma^*)(x)\xi, \xi\rangle \leq \hat{\lambda}_0|\xi|^2, \quad x, \xi \in \mathbb{R}^d, \tag{1.2}$$

$$\|\sigma(x) - \sigma(y)\|_{\mathrm{HS}} \leq L_0|x-y|, \quad x, y \in \mathbb{R}^d, \tag{1.3}$$

where σ^* means the transpose of σ and $\|\cdot\|_{\mathrm{HS}}$ stands for the Hilbert-Schmidt norm.

Below, we make some comments on the assumptions **(A2)** and **(A3)**.

Remark 1.1. *If ϕ_γ is bounded, then we can replace $\phi_\gamma(s)$ in* **(A2)** *by $\sup_{s\in[0,T]} \phi_\gamma(s)$ which is automatically decreasing. Let*

$$\omega_{n,\delta}(\phi_\gamma) = \sup_{x,y\in[n\delta,(n+1)\delta]} |\phi_\gamma(x) - \phi_\gamma(y)|.$$

Instead of that ϕ_γ is decreasing, we can assume that ϕ_γ satisfies

$$\sup_{0<\delta\le 1}\left(\delta\sum_{k=1}^{\lfloor T/\delta\rfloor}\omega_{k,\delta}(\phi_\gamma)\right)<+\infty. \tag{1.4}$$

Then for any $\kappa_0>0$

$$\begin{aligned}\sum_{k=1}^{\lfloor T/\delta\rfloor}\phi_\gamma(\kappa_0 k\delta)\delta &\le \sum_{k=1}^{\lfloor T/\delta\rfloor}\int_{k\delta}^{(k+1)\delta}\phi_\gamma(\kappa_0 t)\mathrm{d}t+\delta\sum_{k=1}^{\lfloor T/\delta\rfloor}\omega_{n,\delta}(\phi_\gamma(\kappa_0\cdot))\\ &\le \kappa_0^{-1}\left(\int_0^{\kappa_0 T}\phi_\gamma(t)\mathrm{d}t+\sup_{0<\delta\le 1}\left(\delta\sum_{k=1}^{\lfloor T/\delta\rfloor}\omega_{n,\delta}(\phi_\gamma)\right)\right)\\ &<\infty.\end{aligned}$$

It is not easy to check (1.4) for ϕ_γ with $\lim_{\delta\to 0^+}\phi_\gamma(\delta)=+\infty$. However, if ϕ_γ is decreasing, then

$$\sup_{0<\delta\le 1}\left(\delta\sum_{k=1}^{\lfloor T/\delta\rfloor}\omega_{n,\delta}(\phi_\gamma)\right)=\sup_{0<\delta\le 1}\left(\delta\sum_{k=1}^{\lfloor T/\delta\rfloor}(\phi_\gamma(k\delta)-\phi_\gamma((k+1)\delta))\right)=\delta\phi_\gamma(\delta).$$

Hence, in this case, (1.4) holds if and only if there is $C>0$ such that $\phi_\gamma(x)\le\frac{C}{x}$.

Remark 1.2. *For $x\in\mathbb{R}^d$, let $\|\sigma(x)\|_{\rm op}=\sup_{|y|\le 1}|\sigma(x)y|$, the operator norm of $\sigma(x)$. By the Cauchy-Schwarz inequality, it follows from (1.2) that*

$$\|\sigma(x)\|_{\rm op}^2=\sum_{i=1}^d\sup_{|y|\le 1}\langle y,\sigma(x)^*e_i\rangle^2\le\|\sigma^*(x)\|_{\rm HS}^2=\sum_{i=1}^d\langle(\sigma\sigma)(x)^*e_i,e_i\rangle\le d\hat\lambda_0,\qquad x\in\mathbb{R}^d,$$

where $e_i, i=1,\cdots,d$, is the orthogonal basis of $\mathbb{R}^d$. Then, we arrive at

$$\|\sigma(x)\|_{\rm op}\le\|\sigma(x)\|_{\rm HS}=\|\sigma^*(x)\|_{\rm HS}\le\sqrt{d\hat\lambda_0},\qquad x\in\mathbb{R}^d. \tag{1.5}$$

Under (**A1**) and (**A3**), (1.1) has a unique strong solution $(X_t)_{t\ge 0}$; see, for instance, [9, Lemma 3.1]. (**A2**) is imposed to reveal the convergence rate of EM scheme corresponding to (1.1), which is defined as below: for any $\delta\in(0,1)$,

$$\mathrm{d}X_t^{(\delta)}=b(X_{t_\delta}^{(\delta)})\mathrm{d}t+\sigma(X_{t_\delta}^{(\delta)})\mathrm{d}W_t,\quad t\ge 0,\quad X_0^{(\delta)}=X_0 \tag{1.6}$$

with $t_\delta:=\lfloor t/\delta\rfloor\delta$, where $\lfloor t/\delta\rfloor$ denotes the integer part of t/δ. We emphasize that $(X_{k\delta}^{(\delta)})_{k\ge 0}$ is a homogeneous Markov process; see e.g. [19, Theorem 6.14]. For $t\ge s$ and $x\in\mathbb{R}^d$, denote $p^{(\delta)}(s,t,x,\cdot)$ by the transition density of $X_t^{(\delta)}$ with the starting point $X_s^{(\delta)}=x$. Set

$$\mathscr{K}_1:=\Big\{(p,q)\in(1,\infty)\times(1,\infty):\frac{d}{p}+\frac{2}{q}<2\Big\},\qquad \gamma_0:=\frac{1}{1-1/q-d/2p},\quad (p,q)\in\mathscr{K}_1,$$
$$\mathscr{K}_2:=\Big\{(p,q)\in(1,\infty)\times(1,\infty):\frac{d}{p}+\frac{1}{q}<1\Big\}.$$

Our first main result in this paper is stated as follows.

Theorem 1.3. *Assume* (**A1**)-(**A3**). *Then, for* $\beta \in (0, \gamma)$, $(p, q) \in \mathscr{K}$ *and* $T > 0$, *there exist constants* $C_1, C_2 > 0$ *independent of* δ *such that*

$$\mathbb{E}\Big(\sup_{0\le t\le T}|X_t - X_t^{(\delta)}|^\beta\Big) \le C_1 \exp\Big(C_2\Big(1 + \||b|^2\|_{L^p}^{\gamma_0}\Big)\Big)\Big(\delta^{\frac{\beta}{2}} + \delta^{\frac{\alpha\gamma\beta}{2\gamma}}\Big). \tag{1.7}$$

Compared with [22], in Theorem 1.3 we get rid of the one-sided Lipschitz condition put on the drift coefficients. On the other hand, [23] is extended to the multidimensional setup. We point out that an $\mathcal{A}$ approximation is given in advance in [22, 23] to approximate the drift term. So, with contrast to the assumption set in [22, 23], the assumption (**A2**) is much more explicit. On the other hand, by a close inspection of the argument of Lemma 2.2 below, the assumption (**A2**) can indeed be replaced by the other alternatives. For instance, (**A2**) may be taken the place of (**A2'**) below.

(A2') there exist constants $\gamma \ge 2, \beta_\gamma, \theta_\gamma > 0$ such that for some constant $C > 0$,

$$\frac{1}{(rs)^{d/2}} \sup_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d\times\mathbb{R}^d} |b(x) - b(y)|^\gamma e^{-\frac{|x-z|^2}{s}} e^{-\frac{|y-x|^2}{r}} dy dx \le C r^{\theta_\gamma} s^{\beta_\gamma - 1}, \quad s, r > 0.$$

The drift b satisfying **(A2')** is said to be the Gaussian-Besov class with the index $(\beta_\gamma, \theta_\gamma)$, written as $GB^\gamma_{\beta_\gamma,\theta_\gamma}(\mathbb{R}^d)$. The index θ_γ is used to characterize the order of continuity and β_γ is used to characterize the type of continuity. Remark that functions with the same order of continuity may enjoy different type continuity; see, for instance, $f(x) = |x|^{\frac{1}{2}}$ with $(1, 1/2)$, and $f(x) = \mathbf{1}_{[c,d]}(x), c, d \in \mathbb{R}$, with $(1/2, 1/2)$. We refer to Example 4.2 below for the drift $b \in GB^2_{\beta_2,\theta_2}(\mathbb{R}^d)$. For $\theta \in (0, 1)$ and $p \ge 1$, let $W^{\theta,p}(\mathbb{R}^d)$ be the fractional order Sobolev space on $\mathbb{R}^d$. Nevertheless, $W^{\theta,p}(\mathbb{R}^d) \subsetneq GB^2_{1-\frac{d}{p},\theta}(\mathbb{R}^d)$, $\theta > 0, p \in [2, \infty)\cap(d, \infty)$; see Example 4.3 for more details. Furthermore, [29, Example 2.3] shows that the drift b constructed therein satisfies (**A2'**) but need not to be piecewise Lipschitz continuous (see e.g. [16, 17]).

In Theorem 1.3, the integrable condition (i.e., $|b|^2 \in L^p(\mathbb{R}^d)$) seems to be a little bit restrictive, which rules out some typical examples, e.g., $b(x) = \mathbf{1}_{[0,\infty)}(x)$. In the sequel, by implementing a truncation argument, the integrable condition can indeed be dropped. In such a setup (i.e., without integrable condition), we can still derive the convergence rate of the EM algorithm, which is presented as below.

Theorem 1.4. *Assume* $(\mathbf{A1}) - (\mathbf{A3})$ *without* $|b|^2 \in L^p(\mathbb{R}^d)$. *Then, for* $\beta \in (0, 2)$, $(p, q) \in \mathscr{K}_2$ *and* $T > 0$, *there exist constants* $C_1, C_2 > 0$ *independent of* δ *such that*

$$\mathbb{E}\Big(\sup_{0\le t\le T}|X_t - X_t^{(\delta)}|^\beta\Big) \le C_1\Big\{e^{C_2(-\frac{\beta}{2}(1\wedge\frac{\alpha\gamma}{\gamma})\log\delta)^{\frac{d\gamma_0}{2p}}} + 1\Big\}\Big(\delta^{\frac{\beta}{2}} + \delta^{\frac{\alpha\gamma\beta}{2\gamma}}\Big). \tag{1.8}$$

We remark that the right hand side of (1.8) approaches zero since

$$\lim_{\delta\to 0} e^{C_2(-\frac{\beta}{2}(1\wedge\frac{\alpha\gamma}{\gamma})\log\delta)^{\frac{d\gamma_0}{2p}}} \delta^{\frac{\beta}{2}(1\wedge\frac{\alpha\gamma}{\gamma})} = 0$$

due to the fact that $\lim_{x\to\infty} \frac{e^{C_2 x^{\frac{d\gamma_0}{2p}}}}{e^x} = 0$ whenever $(p, q) \in \mathscr{K}_2$.

The remainder of this paper is organized as follows. In Section 2, by employing Zvonkin's transform and establishing Krylov's estimate and Khaminskii's estimate for EM algorithm, which is based on Gaussian type estimate of heat kernel, we complete the proof of Theorem 1.3; In Section 3, we aim to finish the proof of Theorem 1.4 by adopting a truncation argument; In Section 4 we provide some illustrative examples to demonstrate our theory established; In the Appendix part, we show explicit upper bounds of the parameters associated with Gaussian type heat kernel estimates concerned with the exact solution and the EM scheme.

2 Proof of Theorem 1.3

Before finishing the proof of Theorem 1.3, we prepare several auxiliary lemmas. Set

$$
\begin{aligned}
\Lambda_1 :&= 2\Big\{\frac{\|b\|_\infty}{\sqrt{\check{\lambda}_0}} + 2\sqrt{d}L_0(\hat{\lambda}_0/\check{\lambda}_0)^2 + d^{\frac{d}{2}+1}d!(\hat{\lambda}_0/\check{\lambda}_0)^d L_0\Big\}\mathrm{e}^{\frac{\|b\|_\infty^2 T}{\hat{\lambda}_0}} \\
&\quad \vee \Big\{2\sqrt{\hat{\lambda}_0}\|b\|_\infty + \big(\|b\|_\infty^2 + 2\hat{\lambda}_0 L_0\sqrt{d}\big)\big(\sqrt{d}+2\big) + 2^{m+11}\check{\lambda}_0^{-1}\big(L_0 + 2\|b\|_\infty\big) \\
&\quad \times \Big(\|b\|_\infty^3 + (d\hat{\lambda}_0)^{\frac{3}{2}} + \check{\lambda}_0^{\frac{1}{2}}\big(\|b\|_\infty^2 + d\hat{\lambda}_0\big)\Big)\Big\}\frac{2^{\frac{d+1}{2}}}{\check{\lambda}_0}\mathrm{e}^{\frac{(\|b\|_\infty + \|b\|_\infty^2)T}{\hat{\lambda}_0}},
\end{aligned}
\tag{2.1}
$$

and

$$
\Lambda_2 := \mathrm{e}^{\frac{\|b\|_\infty T}{2\hat{\lambda}_0}}\sum_{i=0}^\infty \frac{\Big(\Lambda_1\sqrt{\pi T}\big((1+24d)\hat{\lambda}_0/\check{\lambda}_0\big)^d\Big)^i}{\Gamma(1+\frac{i}{2})},
\tag{2.2}
$$

where $\Gamma(\cdot)$ denotes the Gamma function. Due to Stirling's formula: $\Gamma(z+1) \sim \sqrt{2\pi z}(z/\mathrm{e})^z$, we have $\Lambda_2 < \infty$.

The lemma below provides an explicit upper bound of the transition kernel for $(X_t^{(\delta)})_{t\ge 0}$.

Lemma 2.1. *Under* (**A1**) *and* (**A3**),

$$
p^{(\delta)}(j\delta, t, x, y) \le \frac{\Lambda_3 \mathrm{e}^{-\frac{|y-x|^2}{\kappa_0(t-j\delta)}}}{(2\pi\check{\lambda}_0(t-j\delta))^{d/2}}, \quad x, y \in \mathbb{R}^d, \quad t > j\delta, \quad \delta \in (0,1),
\tag{2.3}
$$

where

$$
\kappa_0 := 4(1+24d)\hat{\lambda}_0, \qquad \Lambda_3 := \Lambda_2 \mathrm{e}^{\frac{\|b\|_\infty^2}{2\hat{\lambda}_0}}\Big(\frac{\kappa_0}{2\check{\lambda}_0}\Big)^{d/2}.
\tag{2.4}
$$

Proof. For fixed $t > 0$, there is an integer $k \ge 0$ such that $[k\delta, (k+1)\delta)$. By a direct calculation, it follows from (1.2) and (1.3) that

$$
p^{(\delta)}(k\delta, t, x, y) \le \frac{\mathrm{e}^{-\frac{|y-x-b(x)(t-k\delta)|^2}{2\hat{\lambda}_0(t-k\delta)}}}{(2\pi\check{\lambda}_0)(t-k\delta)^{d/2}} \le \mathrm{e}^{\frac{\|b\|_\infty^2}{2\hat{\lambda}_0}}\frac{\mathrm{e}^{-\frac{|y-x|^2}{4\hat{\lambda}_0(t-k\delta)}}}{(2\pi\check{\lambda}_0(t-k\delta))^{d/2}},
\tag{2.5}
$$

where in the second inequality we used the basic inequality: $|a-b|^2 \ge \frac{1}{2}|a|^2 - |b|^2, a, b \in \mathbb{R}^d$. Next, by invoking Lemma A.2 below, one has

$$
p^{(\delta)}(j\delta, j'\delta, x, x') \le \frac{\Lambda_2 \mathrm{e}^{-\frac{|x'-x|^2}{\kappa_0(j'\delta-j\delta)}}}{(2\pi\check{\lambda}_0(j'\delta-j\delta))^{d/2}}, \quad j' > j, \;\; x, x' \in \mathbb{R}^d,
\tag{2.6}
$$

where Λ_2, κ_0 were given in (2.2) and (2.4), respectively. Subsequently, (2.3) follows immediately by taking advantage of the Chapman-Kolmogrov equation

$$p^{(\delta)}(j\delta, t, x, y) = \int_{\mathbb{R}^d} p^{(\delta)}(j\delta, \lfloor t/\delta \rfloor \delta, x, u) p^{(\delta)}(\lfloor t/\delta \rfloor \delta, t, u, y) \mathrm{d}u,$$

and the fact that

$$\int_{\mathbb{R}^d} \frac{\mathrm{e}^{-\frac{|u-x|^2}{\kappa_0(k\delta - j\delta)}}}{(2\pi\check{\lambda}_0(k\delta - j\delta))^{d/2}} \frac{\mathrm{e}^{-\frac{|y-u|^2}{4\hat{\lambda}_0(t-k\delta)}}}{(2\pi\check{\lambda}_0(t-k\delta))^{d/2}} \mathrm{d}u \le \Big(\frac{\kappa_0}{2\check{\lambda}_0}\Big)^{d/2} \frac{\mathrm{e}^{-\frac{|y-x|^2}{\kappa_0(t-j\delta)}}}{(2\pi\check{\lambda}_0(t-j\delta))^{d/2}}, \quad k > j.$$

□

Lemma 2.2. *Under* (**A1**)-(**A3**)*, for any* $T > 0$*, there exists a constant* $C > 0$ *such that*

$$\int_0^T \mathbb{E}\big|b(X_t^{(\delta)}) - b(X_{t_\delta}^{(\delta)})\big|^\gamma \mathrm{d}t \le C\delta^{1\wedge\frac{\alpha\gamma}{2}}, \tag{2.7}$$

where $\alpha > 0$ *was introduced in* (**A2**).

Proof. Observe that

$$\begin{aligned}\int_0^T \mathbb{E}\big|b(X_t^{(\delta)}) - b(X_{t_\delta}^{(\delta)})\big|^\gamma \mathrm{d}t &= \int_0^\delta \mathbb{E}\big|b(X_t^{(\delta)}) - b(X_0^{(\delta)})\big|^\gamma \mathrm{d}t \\ &\quad + \sum_{k=1}^{\lfloor T/\delta \rfloor} \int_{k\delta}^{T\wedge(k+1)\delta} \mathbb{E}\big|b(X_t^{(\delta)}) - b(X_{k\delta}^{(\delta)})\big|^\gamma \mathrm{d}t.\end{aligned}$$

By $\|b\|_\infty < \infty$ due to (**A1**), it follows that

$$\int_0^\delta \mathbb{E}\big|b(X_t^{(\delta)}) - b(X_0^{(\delta)})\big|^\gamma \mathrm{d}t \le 2^\gamma \|b\|_\infty^\gamma \delta. \tag{2.8}$$

For $t \in [k\delta, (k+1)\delta)$, by taking the mutual independence between $X_{k\delta}^{(\delta)}$ and $W_t - W_{k\delta}$ into account and employing Lemma 2.1, we derive that

$$\begin{aligned}&\mathbb{E}\big|b(X_t^{(\delta)}) - b(X_{k\delta}^{(\delta)})\big|^\gamma \\ &= \mathbb{E}\big|b(X_{k\delta}^{(\delta)} + b(X_{k\delta}^{(\delta)})(t-k\delta) + \sigma(X_{k\delta}^{(\delta)})(W_t - W_{k\delta})) - b(X_{k\delta}^{(\delta)})\big|^\gamma \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(y+z) - b(y)|^\gamma p^{(\delta)}(0, k\delta, x, y) \\ &\quad \times \frac{\exp(-\frac{1}{2(t-k\delta)}\langle(\sigma^*\sigma)^{-1}(y)(z - b(y)(t-k\delta)), z - b(y)(t-k\delta)\rangle)}{\sqrt{(2\pi)^d \det((t-k\delta)(\sigma\sigma^*)(y))}} \mathrm{d}y\mathrm{d}z \\ &\le \frac{C_1}{(k\delta(t-k\delta))^{d/2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(y+z) - b(y)|^\gamma \mathrm{e}^{-\frac{|z|^2}{4\hat{\lambda}_0(t-k\delta)}} \mathrm{e}^{-\frac{|x-y|^2}{\kappa_0 k\delta}} \mathrm{d}y\mathrm{d}z,\end{aligned} \tag{2.9}$$

for some constant $C_1 > 0$, where κ_0 was given in (2.4). With the aid of the fact that

$$\sup_{x\ge 0}(x^\gamma \mathrm{e}^{-\beta x^2}) = \Big(\frac{\gamma}{2\mathrm{e}\beta}\Big)^{\frac{\gamma}{2}}, \quad \gamma, \beta > 0, \tag{2.10}$$

we infer from (**A2**) and (2.9) that

$$\begin{aligned}\mathbb{E}|b(X_t^{(\delta)})-b(X_{k\delta}^{(\delta)})|^2&\le\frac{C_2\phi_\gamma(\kappa_0k\delta)}{(t-k\delta)^{d/2}}\int_{\mathbb{R}^d}|z|^{\alpha_\gamma}\mathrm{e}^{-\frac{|z|^2}{4\hat\lambda_0(t-k\delta)}}\mathrm{d}z\\&\le\frac{C_3\phi_\gamma(\kappa_0k\delta)\delta^{\frac{\alpha_\gamma}{2}}}{(t-k\delta)^{d/2}}\int_{\mathbb{R}^d}\mathrm{e}^{-\frac{|z|^2}{8\hat\lambda_0(t-k\delta)}}\mathrm{d}z\le C_4\phi_\gamma(\kappa_0k\delta)\delta^{\frac{\alpha_\gamma}{2}}\end{aligned}$$

for some constants $C_2,C_3,C_4>0$. Whence, we arrive at

$$\sum_{k=1}^{\lfloor T/\delta\rfloor}\int_{k\delta}^{T\wedge(k+1)\delta}\mathbb{E}\big|b(X_t^{(\delta)})-b(X_{k\delta}^{(\delta)})\big|^\gamma\mathrm{d}t\le C_4\delta^{\frac{\alpha_\gamma}{2}}\int_\delta^T\phi_\gamma(\kappa_0\lfloor t/\delta\rfloor\delta)\mathrm{d}t. \tag{2.11}$$

Observe that

$$\begin{aligned}\int_\delta^T\phi_\gamma(\kappa_0\lfloor t/\delta\rfloor\delta)\mathrm{d}t&=\sum_{i=1}^{\lfloor T/\delta\rfloor}\int_{(i-1)\delta}^{((1+i)\delta)\wedge T-\delta}\phi_\gamma(\kappa_0 i\delta)\mathrm{d}t\\&\le\sum_{i=1}^{\lfloor T/\delta\rfloor}\int_{(i-1)\delta}^{i\delta}\phi_\gamma(\kappa_0 i\delta)\mathrm{d}t\\&\le\sum_{i=1}^{\lfloor T/\delta\rfloor}\int_{(i-1)\delta}^{i\delta}\phi_\gamma(\kappa_0 t)\mathrm{d}t\le\frac{1}{\kappa_0}\int_0^{\kappa_0T}\phi_\gamma(t)\mathrm{d}t,\end{aligned}$$

where in the second inequality we utilized that $\phi_\gamma:(0,\infty)\to(0,\infty)$ is decreasing. Whence, (2.7) holds true by combining (2.8) with (2.11) and by utilizing $\int_0^{\kappa_0T}\phi_\gamma(t)\mathrm{d}t<\infty$. □

For any $p,q\ge1$ and $0\le S\le T$, let $L_q^p(S,T)=L^q([S,T];L^p(\mathbb{R}^d))$ be the family of all Borel measurable functions $f:[S,T]\times\mathbb{R}^d\to\mathbb{R}^d$ endowed with the norm

$$\|f\|_{L_q^p(S,T)}:=\bigg(\int_S^T\bigg(\int_{\mathbb{R}^d}|f_t(x)|^p\mathrm{d}x\bigg)^{\frac{q}{p}}\mathrm{d}t\bigg)^{\frac1q}<\infty.$$

For simplicity, in the sequel, we shall write $L_q^p(T)$ in place of $L_q^p(0,T)$. Compared with (1.1), in (1.6) we have written the drift term as $b(X_{t_\delta}^{(\delta)})$ in lieu of $b(X_t^{(\delta)})$ so that the classical Krylov estimate (see e.g. [4, 9, 14, 30, 32]) is unapplicable directly. However, the following lemma manifests that $(X_t^{(\delta)})_{t\ge0}$ still satisfies the Khasminskii estimate by employing Gaussian type estimate of heat kernel although the Krylov estimate for $(X_{t_\delta}^{(\delta)})_{t\ge0}$ is invalid as Remark 2.5 below describes.

Lemma 2.3. *Assume* (**A1**) *and* (**A3**). *Then, for $f\in L_q^p(T)$ with $(p,q)\in\mathscr{K}_1$ and $T>0$, the following Khasminskii type estimate*

$$\mathbb{E}\exp\bigg(\lambda\int_0^T|f_t(X_t^{(\delta)})|\mathrm{d}t\bigg)\le2^{1+T(2\lambda\alpha_0\|f\|_{L_q^p(T)})^{\gamma_0}},\qquad\lambda>0 \tag{2.12}$$

holds, where $\gamma_0:=\frac{1}{1-1/q-d/2p}$ and

$$\alpha_0:=\frac{(1-1/p)^{\frac d2(1-1/p)}}{(\check\lambda_0(2\pi)^{\frac1p})^{\frac d2}}\Big\{\hat\lambda_0^{\frac d2(1-1/p)}+\Lambda_3\big(\gamma_0(1-1/q)\big)^{\frac{q-1}{q}}(\kappa_0/2)^{\frac d2(1-1/p)}\Big\}. \tag{2.13}$$

Proof. For $0 \le s \le t \le T$, note that

$$\mathbb{E}\Big(\int_s^t |f_r(X_r^{(\delta)})|\mathrm{d}r\Big|\mathscr{F}_s\Big) = \mathbb{E}\Big(\int_s^{t\wedge(s_\delta+\delta)} |f_r(X_r^{(\delta)})|\mathrm{d}r\Big|\mathscr{F}_s\Big) + \mathbb{E}\Big(\int_{t\wedge(s_\delta+\delta)}^t |f_r(X_r^{(\delta)})|\mathrm{d}r\Big|\mathscr{F}_s\Big)$$
$$=: I_1(s,t) + I_2(s,t).$$

Since

$$X_r^{(\delta)} = X_{s_\delta}^{(\delta)} + b(X_{s_\delta}^{(\delta)})(r-s_\delta) + \sigma(X_{s_\delta}^{(\delta)})(W_s - W_{s_\delta}) + \sigma(X_{s_\delta}^{(\delta)})(W_r - W_s), \quad r \in [s, s_\delta+\delta),$$

we derive from (1.2) and Hölder's inequality that

$$\begin{aligned}
I_1(s,t) &= \int_s^{t\wedge(s_\delta+\delta)} \int_{\mathbb{R}^d} f_r(y_{x,w}+z) \\
&\quad\times \frac{\exp\Big(-\frac{1}{2(r-s)}\langle(\sigma\sigma^*)^{-1}(x)(z-y_{x,w}), z-y_{x,w}\rangle\Big)}{\sqrt{(2\pi(r-s))^d\det((\sigma\sigma^*)(x))}}\mathrm{d}z\Big|_{x=X_{s_\delta}^{(\delta)}}^{w=W_s-W_{s_\delta}}\mathrm{d}r \\
&\le \|f\|_{L_q^p(T)}\Big(\int_s^{t\wedge(s_\delta+\delta)}\Big(\frac{1}{\sqrt{(2\pi(r-s))^d\det((\sigma\sigma^*)(x))}} \\
&\quad\times\Big(\int_{\mathbb{R}^d}\exp\Big(-\frac{p}{2(p-1)(r-s)}\langle(\sigma\sigma^*)^{-1}(x)z,z\rangle\Big)\mathrm{d}z\Big)^{\frac{p-1}{p}}\Big)^{\frac{q}{q-1}}\mathrm{d}r\Big)^{\frac{q-1}{q}}\Big|_{x=X_{s_\delta}^{(\delta)}} \\
&\le (2\pi)^{-\frac{d}{2p}}((p-1)/p)^{\frac{d}{2}(1-\frac{1}{p})}(\hat{\lambda}_0^{1-\frac{1}{p}}/\breve{\lambda}_0)^{\frac{d}{2}}(t-s)^{\frac{1}{\gamma_0}}\|f\|_{L_q^p(T)},
\end{aligned} \tag{2.14}$$

where $y_{x,w} := x + b(x)(r-s_\delta) + \sigma(x)w, x \in \mathbb{R}^d, w \in \mathbb{R}^m$, and $\gamma_0 := \frac{1}{1-1/q-d/2p}$. For $r \ge k\delta$, let $X_{k\delta,r}^{(\delta),x}$ be the EM scheme determined by (1.6) with $X_{k\delta,k\delta}^{(\delta),x} = x$. From the tower property of conditional expectation, one has

$$\begin{aligned}
I_2(s,t) &\le \int_{s_\delta+\delta}^t \mathbb{E}\Big(|f_r(X_r^{(\delta)})|\Big|\mathscr{F}_s\Big)\mathrm{d}r = \int_{s_\delta+\delta}^t \mathbb{E}\Big(\mathbb{E}\Big(|f_r(X_r^{(\delta)})|\Big|\mathscr{F}_{s_\delta+s}\Big)|\mathscr{F}_s\Big)\mathrm{d}r \\
&= \int_{s_\delta+\delta}^t \mathbb{E}\Big(\mathbb{E}|f_r(X_{s_\delta+\delta,r}^{(\delta),x})|\Big|_{x=X_{s_\delta+\delta}^{(\delta)}}\Big|\mathscr{F}_s\Big)\mathrm{d}r.
\end{aligned}$$

In terms of Lemma 2.1, besides Hölder's inequality, one obtains that

$$\begin{aligned}
\mathbb{E}|f_r(X_{s_\delta+\delta,r}^{(\delta),x})| &\le \frac{\Lambda_3}{(2\pi\breve{\lambda}_0(r-s_\delta-\delta))^{d/2}}\int_{\mathbb{R}^d}|f_r(y)|\mathrm{e}^{-\frac{|x-y|^2}{\kappa_0(r-s_\delta-\delta)}}\mathrm{d}y \\
&\le \frac{\Lambda_3}{((2\pi)^{\frac{1}{p}}\breve{\lambda}_0)^{d/2}}\Big(\frac{\kappa_0(p-1)}{2p}\Big)^{\frac{d}{2}(1-1/p)}(r-s_\delta-\delta)^{-\frac{d}{2p}}\Big(\int_{\mathbb{R}^d}|f_r(y)|^p\mathrm{d}y\Big)^{\frac{1}{p}}.
\end{aligned}$$

This further yields by Hölder's inequality that

$$\begin{aligned}
I_2(s,t) &\le \frac{\Lambda_3}{((2\pi)^{\frac{1}{p}}\breve{\lambda}_0)^{d/2}}\Big(\frac{\kappa_0(p-1)}{2p}\Big)^{\frac{d}{2}(1-1/p)}\int_{s_\delta+\delta}^t(r-s_\delta-\delta)^{-\frac{d}{2p}}\Big(\int_{\mathbb{R}^d}|f_r(y)|^p\mathrm{d}y\Big)^{\frac{1}{p}}\mathrm{d}r \\
&= \frac{\Lambda_3\Big(\gamma_0(1-1/q)\Big)^{\frac{q-1}{q}}}{((2\pi)^{\frac{1}{p}}\breve{\lambda}_0)^{d/2}}\Big(\frac{\kappa_0}{2}(1-1/p)\Big)^{\frac{d}{2}(1-1/p)}(t-s)^{\frac{1}{\gamma_0}}\|f\|_{L_q^p(T)}.
\end{aligned} \tag{2.15}$$

Hence, (2.14) and (2.15) imply

$$\mathbb{E}\Big(\int_s^t \big|f_r(X_r^{(\delta)})\big|\mathrm{d}r\Big|\mathscr{F}_s\Big) \le \alpha_0 \,\|f\|_{L_q^p(T)}(t-s)^{\frac{1}{\gamma_0}}, \quad 0\le s\le t\le T, \tag{2.16}$$

in which $\alpha_0>0$ was introduced in (2.13). For each $k\ge 1$, applying inductively (2.16) gives

$$\begin{aligned}
&\mathbb{E}\bigg(\Big(\int_s^t \big|f_r(X_r^{(\delta)})\big|\mathrm{d}r\Big)^k\bigg|\mathscr{F}_s\bigg)\\
&= k!\mathbb{E}\bigg(\int_{\triangle_{k-1}(s,t)} \big|f_{r_1}(X_{r_1}^{(\delta)})\big|\cdots\big|f_{r_{k-1}}(X_{r_{k-1}}^{(\delta)})\big|\mathrm{d}r_1\cdots\mathrm{d}r_{k-1}\\
&\quad\times\mathbb{E}\Big(\int_{r_{k-1}}^t |f_k(X_{r_k}^{(\delta)})|\mathrm{d}r_k\Big|\mathscr{F}_{r_{k-1}}\Big)\bigg|\mathscr{F}_s\bigg)\\
&\le \alpha_0 k!(t-s)^{\frac{1}{\gamma_0}}\|f\|_{L_q^p(T)}\\
&\quad\times\mathbb{E}\bigg(\Big(\int_{\triangle_{k-1}(s,t)} \big|f_{r_1}(X_{r_1}^{(\delta)})\big|\cdots\big|f_{r_{k-1}}(X_{r_{k-1}}^{(\delta)})\big|\mathrm{d}r_1\cdots\mathrm{d}r_{k-1}\Big)\bigg|\mathscr{F}_s\bigg)\\
&\le\cdots\le k!\big(\alpha_0(t-s)^{\frac{1}{\gamma_0}}\|f\|_{L_q^p(T)}\big)^k, \qquad 0\le s\le t\le T,
\end{aligned}\tag{2.17}$$

where

$$\triangle_k(s,t):=\big\{(r_1,\cdots,r_k)\in\mathbb{R}^k: s\le r_1\le\cdots\le r_k\le t\big\}.$$

Taking $\delta_0=(2\alpha_0\lambda\|f\|_{L_q^p(T)})^{-\gamma_0}$, one obviously has $\lambda\alpha_0\delta_0^{\frac{1}{\gamma_0}}\|f\|_{L_q^p(T)}=\frac{1}{2}$. With this and (2.17) in hand, we derive that

$$\mathbb{E}\bigg(\exp\bigg(\lambda\int_{(i-1)\delta_0}^{i\delta_0\wedge T}\big|f_t(X_t^{(\delta)})\big|\mathrm{d}t\bigg)\bigg|\mathscr{F}_{(i-1)\delta_0}\bigg)\le\sum_{k=0}^\infty\frac{1}{2^k}=2,\quad i\ge 1, \tag{2.18}$$

which further implies inductively that

$$\begin{aligned}
\mathbb{E}\exp\bigg(\lambda\int_0^T\big|f_t(X_t^{(\delta)})\big|\mathrm{d}t\bigg) &= \mathbb{E}\bigg(\exp\bigg(\lambda\sum_{i=1}^{\lfloor T/\delta_0\rfloor}\int_{(i-1)\delta_0}^{i\delta_0}\big|f_t(X_t^{(\delta)})\big|\mathrm{d}t\bigg)\\
&\quad\times\mathbb{E}\bigg(\exp\bigg(\lambda\int_{\lfloor T/\delta_0\rfloor\delta_0}^{T}\big|f_t(X_t^{(\delta)})\big|\mathrm{d}t\bigg)\bigg|\mathscr{F}_{\lfloor T/\delta_0\rfloor\delta_0}\bigg)\bigg)\\
&\le 2\,\mathbb{E}\exp\bigg(\lambda\sum_{i=1}^{\lfloor T/\delta_0\rfloor}\int_{(i-1)\delta_0}^{i\delta_0}\big|f_t(X_t^{(\delta)})\big|\mathrm{d}t\bigg)\\
&\le\cdots\le 2^{1+T/\delta_0}.
\end{aligned}\tag{2.19}$$

Therefore, (2.12) is now available by recalling $\delta_0=(2\alpha_0\lambda\|f\|_{L_q^p(T)})^{-\gamma_0}$. □

The following lemma is concerned with Khasminskii's estimate for the solution process $(X_t)_{t\ge0}$, which is more or less standard; see, for instance, [4, 9, 14, 30, 32]. Whereas, we herein state the Khasminskii estimate and provide a sketch of its proof merely for the sake of explicit upper bound.

Lemma 2.4. *Assume* (**A1**) *and* (**A3**). *Then, for* $f \in L_q^p(T)$ *with* $(p,q) \in \mathscr{K}_1$, $\lambda > 0$ *and* $T > 0$,

$$\mathbb{E}\exp\left(\lambda\int_0^T |f_t(X_t)|\mathrm{d}t\right) \le 2^{1+T\left(2\lambda\hat{\alpha}_0\|f\|_{L_q^p(T)}\right)^{\gamma_0}}, \tag{2.20}$$

where

$$\hat{\alpha}_0 := (2\pi)^{-\frac{d}{2p}}\hat{\beta}_T(8(p-1)/p)^{\frac{d}{2}(1-\frac{1}{p})}(\hat{\lambda}_0^{1-\frac{1}{p}}/\check{\lambda}_0)^{\frac{d}{2}}, \qquad \hat{\beta}_T := \mathrm{e}^{\frac{\|b\|_\infty^2 T}{2\check{\lambda}_0}}\sum_{i=0}^{\infty}\frac{\beta_T^i}{\Gamma(1+\frac{i}{2})} \tag{2.21}$$

with β_T *being given in* (A.2) *below.*

Proof. By (A.1) below, it follows from Hölder's inequality and Markov property that

$$\begin{aligned}\mathbb{E}\left(\int_s^t |f_r(X_r)|\mathrm{d}r\bigg|\mathscr{F}_s\right) &= \int_s^t \mathbb{E}|f_r(X_r^{s,x})|\mathrm{d}r\bigg|_{x=X_s}\\ &\le \hat{\beta}_T\int_s^t\int_{\mathbb{R}^d}|f_r(y)|\frac{\mathrm{e}^{-\frac{|y-x|^2}{16\hat{\lambda}_0(r-s)}}}{(2\pi\check{\lambda}_0(r-s))^{d/2}}\mathrm{d}y\mathrm{d}r\bigg|_{x=X_s}\\ &\le \hat{\alpha}_0(t-s)^{1-\frac{d}{2p}-\frac{1}{q}}\|f\|_{L_q^p(T)},\end{aligned} \tag{2.22}$$

where $(X_t^{s,x})_{t\ge s}$ stands for the solution to (1.1) with the initial value $X_s^{s,x} = x$, and $\hat{\beta}_T, \hat{\alpha}_0 > 0$ were introduced in (2.21). Then, (2.20) follows immediately by utilizing (2.22) and by following the exact line to derive (2.19). □

Remark 2.5. In (2.16), Krylov's estimate for $(X_t^{(\delta)})_{t\ge0}$ instead of $(X_{t_\delta}^{(\delta)})_{t\ge0}$ is available. Whereas, the Krylov estimate associated with $(X_{t_\delta}^{(\delta)})_{t\ge0}$ no longer holds true. Indeed, if we take $s,t \in [k\delta,(k+1)\delta)$ for some integer $k \ge 1$, we obviously have

$$\mathbb{E}\left(\int_s^t \left|f_{r_\delta}(X_{r_\delta}^{(\delta)})\right|\mathrm{d}r\bigg|\mathscr{F}_s\right) = |f_{k\delta}(X_{k\delta}^{(\delta)})|(t-s), \quad f \in L_q^p(T),\ (p,q) \in \mathscr{K}_1 \tag{2.23}$$

which is a random variable. Hence, it is impossible to control the quantity on the left hand side of (2.23) by $\|f\|_{L_q^p(T)}$ up to a constant; see also e.g. [26] for more details.

Before we go further, we introduce some additional notation. For $p \ge 1$ and $m \ge 0$, let H_p^m be the usual Sobolev space on $\mathbb{R}^d$ with the norm

$$\|f\|_{H_p^m} := \sum_{k=0}^m \|\nabla^m f\|_{L^p},$$

where ∇^m denotes the m-th order gradient operator. For $q \ge 1$ and $0 \le S \le T$, let

$$\mathbb{H}_p^{m,q}(S,T) = L^q([S,T];H_p^m)$$

and $\mathscr{H}_p^{m,q}(S,T)$ be the collection of all functions $f : [S,T]\times\mathbb{R}^d \to \mathbb{R}$ such that $f \in \mathbb{H}_p^{m,q}(S,T)$ and $\partial_t f \in L_q^p(S,T)$. For a locally integrable function $h : \mathbb{R}^d \to \mathbb{R}$, the Hardy-Littlewood maximal operator $\mathscr{M}h$ is defined as below

$$(\mathscr{M}h)(x) = \sup_{r>0}\frac{1}{|B_r(x)|}\int_{B_r(x)} h(y)\mathrm{d}y, \qquad x \in \mathbb{R}^d,$$

where $B_r(x)$ is the ball with the radius r centered at the point x and $|B_r(x)|$ denotes the d-dimensional Lebesgue measure of $B_r(x)$.

To make the content self-contained, we recall the Hardy-Littlewood maximum theorem, which is stated as the lemma below.

Lemma 2.6. *For any $f \in W^{1,1}_{loc}(\mathbb{R}^d)$, there exists a constant $C > 0$ such that*

$$|f(x) - f(y)| \le C|x - y|\{(\mathscr{M}|\nabla f|)(x) + (\mathscr{M}|\nabla f|)(y)\}, \qquad a.e.\ x, y \in \mathbb{R}^d. \tag{2.24}$$

Moreover, for any $f \in L^p(\mathbb{R}^d), p > 1$, there exists a constant C_p, independent of d, such that

$$\|\mathscr{M} f\|_{L^p} \le C_p \|f\|_{L^p}. \tag{2.25}$$

Remark 2.7. *For the detailed proof of* (2.24)*, please refer to the counterpart of [33, Lemma 3.5]. On the other hand, the inequality in* (2.25) *is called the Hardy-Littlewood maximum inequality, which can be consulted in [28, Theorem 1, p5].*

Now we are in position to complete the

Proof of Theorem 1.3. For any $\lambda > 0$, consider the following PDE for $u^\lambda : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$:

$$\partial_t u^\lambda + \frac{1}{2} \sum_{i,j=1}^{d} \langle \sigma\sigma^* e_i, e_j\rangle \nabla_{e_i} \nabla_{e_j} u^\lambda + b + \nabla_b u^\lambda = \lambda u^\lambda, \tag{2.26}$$

where $(e_j)_{1\le j\le d}$ stipulates the orthogonal basis of $\mathbb{R}^d$ and $(\nabla_b u^\lambda)(x)$ (resp. $(\nabla_{e_j} u^\lambda)(x)$) means the directional derivative of u^λ at the point x along the direction $b(x)$ (resp. e_j). According to [30, Lemma 4.3], (2.26) has a unique solution $u^\lambda \in \mathscr{H}^{2,2q}_{2p}(0, T)$ for the pair $(p, q) \in \mathscr{K}_1$ due to $p > \frac{d}{2}$ satisfying

$$(1 \vee \lambda)^{\frac{1}{2}(1-\frac{d}{2p}-\frac{1}{q})} \|\nabla u^\lambda\|_{T,\infty} + \|\nabla^2 u^\lambda\|_{L^{2p}_{2q}(T)} \le c_1 \||b|^2\|_{L^p} \tag{2.27}$$

for some constant $c_1 > 0$, where $\|\nabla u^\lambda\|_{T,\infty} := \sup_{0\le t\le T, x\in\mathbb{R}^d} \|\nabla u^\lambda_t(x)\|_{\mathrm{HS}}$. With the help of (2.27), there is a constant $\lambda_0 \ge 1$ such that

$$\|\nabla u^\lambda\|_{T,\infty} \le \frac{1}{2}, \quad \lambda \ge \lambda_0. \tag{2.28}$$

For $u^\lambda \in \mathscr{H}^{2,2q}_{2p}(0, T)$, there exists a sequence $u^{\lambda,k} \in C^{1,2}([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ such that

$$\lim_{k\to\infty} \|u^{\lambda,k} - u^\lambda\|_{\mathscr{H}^{2,2q}_{2p}(0,T)} = 0,$$

where

$$\|u\|_{\mathscr{H}^{2,2q}_{2p}(0,T)} := \|\partial . u\|_{L^{2p}_{2q}(0,T)} + \|u\|_{\mathbb{H}^{2,2q}_{2p}(0,T)}.$$

Henceforth, we can apply directly Itô's formula to $u^\lambda \in \mathscr{H}^{2,2q}_{2p}(0, T)$ by adopting a standard approximation approach; see e.g. the arguments of [30, Theorem 2.1] and [32, Lemma 4.3] for more details. Set $\theta_t^\lambda(x) := x + u_t^\lambda(x), x \in \mathbb{R}^d$, and $Z_t^{(\delta)} := X_t - X_t^{(\delta)}$. By Itô's formula, we obtain from (2.26) that

$$
\begin{aligned}
&\mathrm{d}\theta_t^\lambda(X_t) = \lambda u^\lambda(X_t)\mathrm{d}t + \nabla\theta_t^\lambda(X_t)\sigma(X_t)\mathrm{d}W_t \\
&\mathrm{d}\theta_t^\lambda(X_t^{(\delta)}) = \Big\{\lambda u^\lambda(X_t^{(\delta)}) + \nabla\theta_t^\lambda(X_t^{(\delta)})\big(b(X_{t_\delta}^{(\delta)}) - b(X_t^{(\delta)})\big) + \frac{1}{2}\sum_{i,j=1}^d \big\langle\big((\sigma\sigma^*)(X_{t_\delta}^{(\delta)}) \\
&\qquad - (\sigma\sigma^*)(X_t^{(\delta)})\big)e_i, e_j\big\rangle \nabla_{e_i}\nabla_{e_j}u_t^\lambda(X_t^\delta)\Big\}\mathrm{d}t + \nabla\theta_t^\lambda(X_t^{(\delta)})\sigma(X_{t_\delta}^{(\delta)})\mathrm{d}W_t.
\end{aligned}
\tag{2.29}
$$

Let $\Gamma_t = \theta_t^\lambda(X_t) - \theta_t^\lambda(X_t^{(\delta)}), t \geq 0$. From (2.28), it is easy to see that

$$
\frac{1}{2}\big|Z_t^{(\delta)}\big| \leq |\Gamma_t| \leq \frac{3}{2}\big|Z_t^{(\delta)}\big|. \tag{2.30}
$$

Whence, by Itô's formula, we derive from (2.29) that for $\gamma \geq 2$ in (**A2**),

$$
\begin{aligned}
|Z_t^{(\delta)}|^\gamma &\leq 2^\gamma\gamma\lambda\int_0^t |\Gamma(s)|^{\gamma-2}\big\langle\Gamma(s), u^\lambda(X_s) - u^\lambda(X_s^{(\delta)})\big\rangle\mathrm{d}s \\
&\quad + 2^\gamma\gamma\int_0^t |\Gamma(s)|^{\gamma-2}\big\langle\Gamma(s), \nabla\theta_s^\lambda(X_t^{(\delta)})\big(b(X_s^{(\delta)}) - b(X_{s_\delta}^{(\delta)})\big)\big\rangle\mathrm{d}s \\
&\quad + 2^{\gamma-1}\gamma\sum_{i,j=1}^d\int_0^t |\Gamma(s)|^{\gamma-2}\big\langle\big((\sigma\sigma^*)(X_{s_\delta}^{(\delta)}) - (\sigma\sigma^*)(X_s^{(\delta)})\big)e_i, e_j\big\rangle\big\langle\Gamma(s), \nabla_{e_i}\nabla_{e_j}u_s^\lambda(X_s^\delta)\big\rangle\mathrm{d}s \\
&\quad + 2^{\gamma-1}\gamma(\gamma-1)\int_0^t |\Gamma(s)|^{\gamma-2}\big\|\nabla\theta_s^\lambda(X_s)\sigma(X_s) - \nabla\theta_s^\lambda(X_s^{(\delta)})\sigma(X_{s_\delta}^{(\delta)})\big\|_{\mathrm{HS}}^2\mathrm{d}s + \mathrm{d}M_t \\
&=: I_{1,\delta}(t) + I_{2,\delta}(t) + I_{3,\delta}(t) + I_{4,\delta}(t) + M_t,
\end{aligned}
\tag{2.31}
$$

where

$$
M_t := 2^\gamma\gamma\int_0^t |\Gamma(s)|^{\gamma-2}\big\langle\Gamma(s), \big((\nabla\theta_s^\lambda\sigma)(X_s) - \nabla\theta_s^\lambda(X_s^{(\delta)})\sigma(X_{s_\delta}^{(\delta)})\big)\mathrm{d}W_s\big\rangle.
$$

By means of (2.28), we have

$$
I_{1,\delta}(t) \leq 3^{\gamma-1}\gamma\lambda\int_0^t \big|Z_s^{(\delta)}\big|^2\mathrm{d}s. \tag{2.32}
$$

Also, by virtue of (2.28), besides (2.30), we find that there exists a constant $c_2 > 0$ such that

$$
I_{2,\delta}(t) \leq c_2\Big\{\int_0^t \big|Z_s^{(\delta)}\big|^\gamma\mathrm{d}s + \int_0^t \big|b(X_s^{(\delta)}) - b(X_{s_\delta}^{(\delta)})\big|^\gamma\mathrm{d}s\Big\}. \tag{2.33}
$$

By means of (1.3) and (1.5), one has

$$
\begin{aligned}
\big\|(\sigma\sigma^*)(x) - (\sigma\sigma^*)(y)\big\|_{\mathrm{HS}} &\leq \big(\|\sigma(x)\|_{\mathrm{op}} + \|\sigma(y)\|_{\mathrm{op}}\big)\big\|\sigma(x) - \sigma(y)\big\|_{\mathrm{HS}} \\
&\leq 2\,L_0\sqrt{\hat{\lambda}_0 d}\,|x-y|, \quad x, y \in \mathbb{R}^d.
\end{aligned}
$$

This, combining (2.30)and using Young's inequality, leads to

$$
\begin{aligned}
I_{3,\delta}(t) &\le c_3 \int_0^t \big|X_s^{(\delta)} - X_{s_\delta}^{(\delta)}\big| \cdot \big|Z_s^{(\delta)}\big|^{\gamma-1} \cdot \big\|\nabla^2 u_s^\lambda(X_s^\delta)\big\|_{\mathrm{HS}} \mathrm{d}s \\
&\le \frac{c_3}{\gamma} \int_0^t \Big\{(\gamma-1)\big|Z_s^{(\delta)}\big|^\gamma \big\|\nabla^2 u_s^\lambda(X_s^\delta)\big\|_{\mathrm{HS}}^{\frac{\gamma}{\gamma-1}} + \big|X_s^{(\delta)} - X_{s_\delta}^{(\delta)}\big|^\gamma\Big\}\mathrm{d}s \\
&\le \frac{c_3}{2} \int_0^t \Big\{\big|Z_s^{(\delta)}\big|^\gamma \big(\gamma\big\|\nabla^2 u_s^\lambda(X_s^\delta)\big\|_{\mathrm{HS}}^2 + \gamma - 2\big) + \big|X_s^{(\delta)} - X_{s_\delta}^{(\delta)}\big|^\gamma\Big\}\mathrm{d}s
\end{aligned} \tag{2.34}
$$

for some constant $c_3 > 0$. Furthermore, thanks to (1.2), (1.5), (2.24), (2.28) as well as (2.30), we derive from Hölder's inequality that

$$
\begin{aligned}
I_{4,\delta}(t) &\le 2^{\gamma+1}\gamma(\gamma-1) \int_0^t |\Gamma(s)|^{\gamma-2}\Big\{\big\|\sigma(X_s) - \sigma(X_{s_\delta}^{(\delta)})\big\|_{\mathrm{HS}}^2 \\
&\quad + \big\|\big(\nabla u_s^\lambda(X_s) - \nabla u_s^\lambda(X_s^{(\delta)})\big)\sigma(X_s)\big\|_{\mathrm{HS}}^2 \\
&\quad + \big\|\nabla u_s^\lambda(X_s^{(\delta)})\big(\sigma(X_s) - \sigma(X_{s_\delta}^{(\delta)})\big\|_{\mathrm{HS}}^2\Big\}\mathrm{d}s \\
&\le c_4 \int_0^t |\Gamma(s)|^{\gamma-2}\Big\{\big\|\sigma(X_s) - \sigma(X_{s_\delta}^{(\delta)})\big\|_{\mathrm{HS}}^2 + \big\|\nabla u_s^\lambda(X_s) - \nabla u_s^\lambda(X_s^{(\delta)})\big\|_{\mathrm{HS}}^2\Big\}\mathrm{d}s \\
&\le c_5 \int_0^t |\Gamma(s)|^{\gamma-2}\big|Z_s^{(\delta)}\big|^2\Big\{\big(\mathscr{M}\|\nabla^2 u_s^\lambda\|_{\mathrm{HS}}^2\big)(X_s) + \big(\mathscr{M}\|\nabla^2 u_s^\lambda\|_{\mathrm{HS}}^2\big)\big(X_s^{(\delta)}\big)\Big\}\mathrm{d}s \\
&\quad + c_5 \int_0^t |\Gamma(s)|^{\gamma-2}\big\{\big|Z_s^{(\delta)}\big|^2 + \big|X_s^{(\delta)} - X_{s_\delta}^{(\delta)}\big|^2\big\}\mathrm{d}s \\
&\le c_6 \int_0^t \big|Z_s^{(\delta)}\big|^\gamma\Big\{\big(\mathscr{M}\|\nabla^2 u_s^\lambda\|_{\mathrm{HS}}^2\big)(X_s) + \big(\mathscr{M}\|\nabla^2 u_s^\lambda\|_{\mathrm{HS}}^2\big)\big(X_s^{(\delta)}\big)\Big\}\mathrm{d}s \\
&\quad + c_6 \int_0^t \big\{\big|Z_s^{(\delta)}\big|^\gamma + \big|X_s^{(\delta)} - X_{s_\delta}^{(\delta)}\big|^\gamma\big\}\mathrm{d}s
\end{aligned} \tag{2.35}
$$

for some constants $c_4, c_5, c_6 > 0$. As a result, plugging (2.32)-(2.35) into (2.31) gives that

$$
\big|Z_t^{(\delta)}\big|^\gamma \le \int_0^t \big|Z_s^{(\delta)}\big|^\gamma \mathrm{d}A_s + \int_0^t \big\{c_2\big|b(X_s^{(\delta)}) - b(X_{s_\delta}^{(\delta)})\big|^\gamma + (c_3/2 + c_6)\big|X_s^{(\delta)} - X_{s_\delta}^{(\delta)}\big|^\gamma\big\}\mathrm{d}s + M_t,
$$

in which, for some constant $\hat{c}_1 > 0$,

$$
A_t := \hat{c}_1 \int_0^t \big\{1 + \big(\mathscr{M}\|\nabla^2 u_s^\lambda\|_{\mathrm{HS}}^2\big)(X_s) + \big(\mathscr{M}\|\nabla^2 u_s^\lambda\|_{\mathrm{HS}}^2\big)(X_s^{(\delta)}) + \big\|\nabla^2 u_s^\lambda\big\|_{\mathrm{HS}}^2\big(X_s^\delta\big)\big\}\mathrm{d}s, \quad t \ge 0.
$$

Consequently, we deduce by stochastic Gronwall's inequality (see e.g. [30, Lemma 3.8]) that, for $0 < \kappa' < \kappa < 1$,

$$
\begin{aligned}
\Big(\mathbb{E}\big\|Z^{(\delta)}\big\|_{t,\infty}^{\kappa'\gamma}\Big)^{1/\kappa'} &\le \Big(\frac{\kappa}{\kappa - \kappa'}\Big)^{1/\kappa'}\Big(\mathbb{E}\,\mathrm{e}^{\kappa A_t/(1-\kappa)}\Big)^{(1-\kappa)/\kappa} \\
&\quad \times \int_0^t \big\{c_2\mathbb{E}\big|b(X_s^{(\delta)}) - b(X_{s_\delta}^{(\delta)})\big|^\gamma + (c_3/2 + c_6)\mathbb{E}\big|X_s^{(\delta)} - X_{s_\delta}^{(\delta)}\big|^\gamma\big\}\mathrm{d}s,
\end{aligned}
$$

where $\|f\|_{t,\infty} := \sup_{0 \le s \le t}|f(s)|$ for a continuous function $f : \mathbb{R}_+ \to \mathbb{R}^d$. The estimate above, together with Lemma 2.2 and the fact that

$$
\sup_{0 \le t \le T} \mathbb{E}\big|X_t^{(\delta)} - X_{t_\delta}^{(\delta)}\big|^\gamma \le \hat{c}_2 \delta^{\frac{\gamma}{2}}
$$

for some constant $\hat{c}_2 > 0$, leads to

$$\Big(\mathbb{E}\big\|Z^{(\delta)}\big\|_{t,\infty}^{\kappa'\gamma}\Big)^{1/\kappa'} \le \hat{c}_3\big(\mathbb{E}\,\mathrm{e}^{\kappa A_t/(1-\kappa)}\big)^{\frac{1}{\kappa}-1}\big(\delta^{\frac{\gamma}{2}}+\delta^{\frac{\alpha\gamma}{2}}\big) \tag{2.36}$$

for some constant $\hat{c}_3 > 0$. By Hölder's inequality, we deuce for some constant $\hat{c}_4 > 0$,

$$\begin{aligned}\mathbb{E}\mathrm{e}^{\frac{\kappa A_t}{1-\kappa}} &\le \mathrm{e}^{\frac{\kappa\,\hat{c}_1 t}{1-\kappa}}\bigg(\mathbb{E}\exp\bigg(\hat{c}_4\int_0^t\big(\mathscr{M}\big\|\nabla^2 u_s^\lambda\big\|_{\mathrm{HS}}^2\big)(X_s)\mathrm{d}s\bigg)\bigg)^{1/2}\\ &\quad\times\bigg(\mathbb{E}\exp\bigg(\hat{c}_4\int_0^t\big(\mathscr{M}\big\|\nabla^2 u_s^\lambda\big\|_{\mathrm{HS}}^2\big)\big(X_s^{(\delta)}\big)\mathrm{d}s\bigg)\bigg)^{1/4}\\ &\quad\times\bigg(\mathbb{E}\exp\bigg(\hat{c}_4\int_0^t\big\|\nabla^2 u_s^\lambda\big\|_{\mathrm{HS}}^2\big(X_s^\delta\big)\mathrm{d}s\bigg)\bigg)^{1/4}.\end{aligned}$$

This, in addition to (2.12), (2.20), (2.25) as well as (2.27), implies that

$$\begin{aligned}\mathbb{E}\mathrm{e}^{\frac{\kappa A_t}{1-\kappa}} &\le \exp\Big(\hat{c}_5\Big(1+\big\|\|\nabla^2 u^\lambda\|_{\mathrm{HS}}^2\big\|_{L_q^p(T)}^{\gamma_0}+\big\|\mathscr{M}\|\nabla^2 u^\lambda\|_{\mathrm{HS}}^2\big\|_{L_q^p(T)}^{\gamma_0}\Big)\Big)\\ &\le \exp\Big(\hat{c}_5\Big(1+\big\|\nabla^2 u^\lambda\big\|_{L_{2q}^{2p}(T)}^{2\gamma_0}\Big)\Big)\\ &\le \exp\Big(\hat{c}_5\Big(1+\big\||b|^2\big\|_{L^p}^{\gamma_0}\Big)\Big)\end{aligned} \tag{2.37}$$

for some constants $\hat{c}_5, \hat{c}_6, \hat{c}_7 > 0$. Substituting (2.37) back into (2.36), we find constants $\hat{c}_8, \hat{c}_9 > 0$ such that

$$\mathbb{E}\big\|Z^{(\delta)}\big\|_{t,\infty}^{\kappa'\gamma} \le \hat{c}_8\exp\Big(\hat{c}_9\Big(1+\big\||b|^2\big\|_{L^p}^{\gamma_0}\Big)\Big)\big(\delta^{\frac{\gamma}{2}}+\delta^{\frac{\alpha\gamma}{2}}\big)^{\kappa'}$$

so that we have

$$\mathbb{E}\big\|Z^{(\delta)}\big\|_{t,\infty}^{\beta} \le \hat{c}_8\exp\Big(\hat{c}_9\Big(1+\big\||b|^2\big\|_{L^p}^{\gamma_0}\Big)\Big)\Big(\delta^{\frac{\beta}{2}}+\delta^{\frac{\alpha\gamma\beta}{2\gamma}}\Big), \quad \beta\in(0,\gamma).$$

We therefore complete the proof. □

3 Proof of Theorem 1.4

In this section, we aim to complete the proof of Theorem 1.4 by carrying out a truncation approach; see, for example, [1, 23] for further details.

Let $\psi:\mathbb{R}_+\to[0,1]$ be a smooth function such that

$$\psi(r)=1, \quad r\in[0,1], \qquad \psi(r)\equiv 0, \quad r\ge 2.$$

For each integer $k\ge 1$, let $b_k(x)=b(x)\psi(|x|/k)$, $x\in\mathbb{R}^d$, be the truncation function associated with the drift b. A direct calculation shows that

$$\|b_k\|_\infty \le \|b\|_\infty \qquad\text{and}\qquad \big\||b_k|^2\big\|_{L^p} \le \bigg(\frac{2^d\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}\bigg)^{1/p}k^{\frac{d}{p}}\|b\|_\infty^2. \tag{3.1}$$

Consider the following truncated SDE corresponding to (1.1)

$$\mathrm{d}X_t^k = b_k(X_t^k)\mathrm{d}t+\sigma(X_t^k)\mathrm{d}W_t, \quad t\ge 0, \quad X_0^k = X_0. \tag{3.2}$$

The EM scheme associated with (3.2) is given by

$$\mathrm{d}X_t^{k,(\delta)} = b_k(X_{t_\delta}^{k,(\delta)})\mathrm{d}t + \sigma(X_{t_\delta}^{k,(\delta)})\mathrm{d}W_t, \quad t \geq 0, \quad X_0^{k,(\delta)} = X_0^{(k)}.$$

Observe that for $\beta \in (0,\gamma)$

$$\begin{aligned}\mathbb{E}\big\|X - X^{(\delta)}\big\|_{T,\infty}^{\beta} &\leq 3^{0\vee(\beta-1)}\{\mathbb{E}\big\|X - X^{k}\big\|_{T,\infty}^{2} + \mathbb{E}\big\|X^{(\delta)} - X^{k,(\delta)}\big\|_{T,\infty}^{2} \\ &\quad + \mathbb{E}\big\|X_t^{k} - X^{k,(\delta)}\big\|_{T,\infty}^{2}\} \\ &=: 3^{0\vee(\beta-1)}\{I_1 + I_2 + I_3\},\end{aligned} \tag{3.3}$$

where, for a map $f : [0,T] \to \mathbb{R}^d$, we set $\|f\|_{T,\infty} := \sup_{0\leq t\leq T}|f(t)|$. Via Hölder's inequality and the fact

$$\{X_t \neq X_t^k, 0 \leq t \leq T\} \subseteq \{\|X\|_{T,\infty} \geq k\},$$

it follows that

$$I_1 = \mathbb{E}\Big(\big\|X - X^k\big\|_{T,\infty}^{\beta}\mathbf{1}_{\{\|X\|_{T,\infty}\geq k\}}\Big) \leq \Big(\mathbb{E}\big\|X - X^k\big\|_{T,\infty}^{2\beta}\Big)^{1/2}\Big(\mathbb{P}\big(\big\|X\big\|_{T,\infty} \geq k\big)\Big)^{1/2}.$$

Since

$$\big\|X\big\|_{T,\infty} \leq |x| + \|b\|_\infty T + |M|_{T,\infty},$$

in which

$$M_t := \int_0^t \sigma(X_s)\mathrm{d}W_s, \qquad t \geq 0,$$

with the quadratic variation $\langle M\rangle_T \leq d\hat{\lambda}_0 T$, we derive from [27, Proposition 6.8, p147] that

$$\begin{aligned}\mathbb{P}(\|X\|_{T,\infty} \geq k) &\leq \mathbb{P}\Big(\|M\|_{T,\infty} \geq k - |x| - \|b\|_\infty T, \langle M\rangle_T \leq d\hat{\lambda}_0 T\Big) \\ &\leq 2\,d\exp\Big(-\frac{(k - |x| - \|b\|_\infty T)^2}{4d^2\hat{\lambda}_0 T}\Big) \\ &\leq 2\,d\exp\Big(\frac{(|x| + \|b\|_\infty T)^2}{4d^2\hat{\lambda}_0 T}\Big)\mathrm{e}^{-\frac{k^2}{8d^2\hat{\lambda}_0 T}},\end{aligned} \tag{3.4}$$

where in the last display we used the inequality: $(a-b)^2 \geq a^2/2 - b^2, a, b \in \mathbb{R}$. Thus, (3.4), besides

$$\mathbb{E}\|X\|_{T,\infty}^{2\beta} + \mathbb{E}\|X^k\|_{T,\infty}^{2\beta} \leq C_1$$

for some constant C_1, yields

$$I_1 \leq C_2\exp\Big(\frac{(|x| + \|b\|_\infty T)^2}{8d^2\hat{\lambda}_0 T}\Big)\mathrm{e}^{-\frac{k^2}{16d^2\hat{\lambda}_0 T}} \tag{3.5}$$

for some constant $C_2 > 0$. Following a similar procedure to derive (3.5), we also derive that

$$I_2 \leq C_3\exp\Big(\frac{(|x| + \|b\|_\infty T)^2}{8d^2\hat{\lambda}_0 T}\Big)\mathrm{e}^{-\frac{k^2}{16d^2\hat{\lambda}_0 T}} \tag{3.6}$$

for some constant $C_3>0$. Moreover, for $(p,q)\in\mathscr{K}_2$, according to Theorem 1.3, there exist constants $C_4, C_5>0$ such that

$$\mathbb{E}\|X^k-X^{k,(\delta)}\|_{T,\infty}^2\le C_4\mathrm{e}^{C_5\big\||b_k|^2\big\|_{L^p}^{\gamma_0}}\Big(\delta^{\frac{\beta}{2}}+\delta^{\frac{\alpha_\gamma\beta}{2\gamma}}\Big).$$

This, together with (3.1), implies

$$\mathbb{E}\|X^k-X^{k,(\delta)}\|_{T,\infty}^2\le C_4\mathrm{e}^{C_6\|b\|_\infty^{2\gamma_0}k^{\frac{d\gamma_0}{p}}}\Big(\delta^{\frac{\beta}{2}}+\delta^{\frac{\alpha_\gamma\beta}{2\gamma}}\Big)\tag{3.7}$$

for some constant $C_6>0$. As a consequence, from (3.5), (3.6), and (3.7), we arrive at

$$\mathbb{E}\|X-X^{(\delta)}\|_{T,\infty}^2\le C_8\Big\{\mathrm{e}^{-\frac{k^2}{16d^2\hat{\lambda}_0T}}+\mathrm{e}^{C_7k^{\frac{d\gamma_0}{p}}}\delta^{\frac{\beta}{2}(1\wedge\frac{\alpha_\gamma}{\gamma})}\Big\}$$

for some constants $C_7, C_8>0$. Thereby, the desired assertion (1.8) follows by taking

$$k=\Big(-8\beta d^2\hat{\lambda}_0T\Big(1\wedge\frac{\alpha_\gamma}{\gamma}\Big)\log\delta\Big)^{\frac{1}{2}}.$$

4 Illustrative examples

In this section, we intend to give examples to demonstrate that the assumption imposed on drift term holds true.

Example 4.1. Let $b(x)=\mathbf{1}_{[a_1,a_2]}(x), x\in\mathbb{R}$, for some constants $a_1<a_2$. Apparently, b is not continuous at all but $b^2\in L^p$ for any $p\ge1$. Observe that

$$\lim_{\varepsilon\downarrow0}\frac{-\varepsilon(b(a_1-\varepsilon)-b(a_1))}{\varepsilon^2}=\lim_{\varepsilon\downarrow0}\frac{1}{\varepsilon}=\infty$$

so that b does not obey the one-sided Lipschitz condition. Next we aim to show that b given above satisfies (**A2**). By a direct calculation, for any $s>0, \gamma\ge2$ and $y\in\mathbb{R}$,

$$\begin{aligned}\int_{-\infty}^{\infty}|b(x+y+z)-b(x+y)|^\kappa\mathrm{e}^{-\frac{x^2}{s}}\mathrm{d}x&\le\int_{-\infty}^{\infty}|b(x+z)-b(x)|^\kappa\mathrm{d}x\\&=\int_{a_1-z}^{a_2-z}\mathbf{1}_{[a_1,a_2]^c}(x)\mathrm{d}x+\int_{a_1}^{a_2}\mathbf{1}_{[a_1-z,a_2-z]^c}(x)\mathrm{d}x\\&=:I_1(z)+I_2(z).\end{aligned}$$

If $z\ge0$, then

$$I_1(z)=\int_{a_1-z}^{(a_2-z)\wedge a_1}\mathrm{d}x\le|z|\quad\text{and}\quad I_2(z)=\int_{(a_2-z)\vee a_1}^{a_2}\mathrm{d}x\le|z|.$$

On the other hand, for $z<0$, we have

$$I_1(z)=\int_{(a_1-z)\vee a_2}^{a_2-z}\mathrm{d}x\le|z|\quad\text{and}\quad I_2(z)=\int_{a_1}^{a_2\wedge(a_1-z)}\mathrm{d}x\le|z|.$$

So, (**A2**) holds true with $\alpha=1$ and $\phi(s)=s^{-\frac{1}{2}}$, $s>0$.

Example 4.2. For $\theta > 0$ and $p \in [2,\infty) \cap (d,\infty)$, if the Gagliardo seminorm

$$[b]_{W^{p,\theta}} := \Big(\int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{|b(x)-b(y)|^p}{|x-y|^{d+p\theta}} \mathrm{d}x\mathrm{d}y \Big)^{\frac{1}{p}} < \infty,$$

then $b \in GB^2_{1-\frac{d}{p},\theta}(\mathbb{R}^d)$. Indeed, by Hölder's inequality and (2.10), it follows that

$$\begin{aligned}
&\frac{1}{(rs)^{d/2}} \int_{\mathbb{R}^d\times\mathbb{R}^d} |b(x)-b(y)|^2 \mathrm{e}^{-\frac{|x-z|^2}{s}} \mathrm{e}^{-\frac{|y-x|^2}{r}} \mathrm{d}y\mathrm{d}x \\
&= \frac{1}{(rs)^{d/2}} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|b(x)-b(y)|^2}{|x-y|^{\frac{2d}{p}+2\theta}} \mathrm{e}^{-\frac{|x-z|^2}{s}} \mathrm{e}^{-\frac{|y-x|^2}{r}} |x-y|^{\frac{2d}{p}+2\theta} \mathrm{d}y\mathrm{d}x \\
&\le C_1 \frac{[b]^{\frac{2}{p}}_{W^{p,\theta}}}{(rs)^{d/2}} \Big(\int_{\mathbb{R}^d\times\mathbb{R}^d} \mathrm{e}^{-\frac{p|x-z|^2}{(p-2)s}} \mathrm{e}^{-\frac{p|x-y|^2}{(p-2)r}} |x-y|^{\frac{2(d+p\theta)}{p-2}} \mathrm{d}y\mathrm{d}x \Big)^{\frac{p-2}{p}} \\
&\le C_2 [b]^{\frac{2}{p}}_{W^{p,\theta}} \frac{r^{\frac{d}{p}+\theta}}{(rs)^{d/2}} \Big(\int_{\mathbb{R}^d\times\mathbb{R}^d} \mathrm{e}^{-\frac{p|x-z|^2}{(p-2)s}} \mathrm{e}^{-\frac{p|x-y|^2}{2(p-2)r}} \mathrm{d}y\mathrm{d}x \Big)^{\frac{p-2}{p}} \\
&\le C_3 [b]^{\frac{2}{p}}_{W^{p,\theta}} \Big(\frac{p-2}{p} \Big)^{\frac{d(p-2)}{p}} s^{-\frac{d}{p}} r^{\theta}, \quad r,s>0, z\in\mathbb{R}^d, p>2
\end{aligned} \tag{4.1}$$

for some constants $C_1, C_2, C_3 > 0$. On the other hand, if $d = 1$ and $p = 2$, we deduce from (4.1) that $b \in GB^2_{1/2,\theta}(\mathbb{R}^d)$ due to $\lim_{x\to 0} x^x = 1$.

Example 4.3. For $0 < a < b < \infty$, $f(\cdot) := \mathbb{1}_{[a,b]}(\cdot) \in GB^2_{\frac{1}{2},\frac{1}{2}}(\mathbb{R})$ whereas $f \notin W^{\frac{1}{2},2}(\mathbb{R})$. In fact, it is easy to see that

$$f \in \cap_{0\le\theta<\frac{1}{2}} W^{\theta,2}, \qquad \lim_{\theta\uparrow\frac{1}{2}} [f]_{W^{\theta,2}} = \infty,$$

which yields $f \notin W^{\frac{1}{2},2}(\mathbb{R})$. On the other hand, since

$$\frac{1}{(rs)^{d/2}} \int_{\mathbb{R}^2} |f(x)-f(y)|^2 \mathrm{e}^{-\frac{|x-z|^2}{s}} \mathrm{e}^{-\frac{|y-x|^2}{r}} \mathrm{d}y\mathrm{d}x \le C s^{-\frac{1}{2}} r^{\frac{1}{2}}, \quad r,s>0, z\in\mathbb{R}$$

for some constant $C > 0$, we arrive at $f \in GB^2_{\frac{1}{2},\frac{1}{2}}(\mathbb{R})$.

A Appendix

The lemma below provides explicit estimates of the parameters concerning Gaussian type estimate of transition density for the diffusion process $(X_t)_{t\ge0}$ solving (1.1).

Lemma A.1. *Under $\|b\|_\infty < \infty$ and (**A3**), the transition density p of $(X_t)_{t\ge s}$ satisfies*

$$p(s,t,x,x') \le \mathrm{e}^{\frac{\|b\|^2_\infty T}{2\check{\lambda}_0}} \sum_{i=0}^{\infty} \frac{\beta_T^i}{\Gamma(1+\frac{i}{2})} p_0(t-s,x,x'), \quad 0\le s<t\le T, x,x'\in\mathbb{R}^d, \tag{A.1}$$

where $\Gamma(\cdot)$ is the Gamma function, and

$$\beta_T := 2^{3d+1} \Big(\frac{\hat{\lambda}_0}{\check{\lambda}_0}\Big)^{d+1} (\pi T)^{\frac{1}{2}} \Big\{ \frac{\|b\|_\infty}{\sqrt{\hat{\lambda}_0}} + L_0(d+2\sqrt{d}) \Big\} \mathrm{e}^{\frac{\|b\|^2_\infty T}{4\hat{\lambda}_0}}, p_0(t,x,x') := \frac{\mathrm{e}^{-\frac{|x-x'|^2}{16\hat{\lambda}_0 t}}}{(2\pi\check{\lambda}_0 t)^{d/2}}. \tag{A.2}$$

Proof. The proof of Lemma A.1 is based on the parametrix method [13, 15]. To complete the proof of Lemma A.1, it suffices to refine the argument of [13, Lemma 3.2]; see also e.g. [15, p1660-1662] for further details. Under $\|b\|_\infty < \infty$ and (**A3**), X_t admits a smooth transition density $p(s,t,x,y)$ at the point y, given $X_s = x$, such that

$$\begin{aligned}\partial_t p(s,t,x,y) &= L^* p(s,t,x,y), \quad p(s,t,x,\cdot) = \delta_x(\cdot), \quad t \downarrow s,\\ \partial_s p(s,t,x,y) &= -Lp(s,t,x,y), \quad p(s,t,\cdot,y) = \delta_y(\cdot), \quad s \uparrow t,\end{aligned} \tag{A.3}$$

where L is the infinitesimal generator of (1.1) and L^* is its adjoint operator. For $t > s$ and $x, x' \in \mathbb{R}^d$, let $\widetilde{X}_t^{s,x,x'}$ solve the frozen SDE

$$\mathrm{d}\widetilde{X}_t^{s,x,x'} = b(x')\mathrm{d}t + \sigma(x')\mathrm{d}W_t, \quad t > s, \ \ \widetilde{X}_s^{s,x,x'} = x \in \mathbb{R}^d \tag{A.4}$$

and $\widetilde{p}^{x'}(s,t,x,x')$ stand for its transition density at x', given $\widetilde{X}_s^{s,x,x'} = x$. Apparently, $\widetilde{p}^{x'}$ admits the explicit form

$$\widetilde{p}^{x'}(s,t,x,x') = \frac{\mathrm{e}^{-\frac{1}{2(t-s)}\langle(\sigma\sigma^*)^{-1}(x')(x'-x-b(x')(t-s)), x'-x-b(x')(t-s)\rangle}}{\sqrt{(2\pi(t-s))^d \det((\sigma\sigma^*)(x'))}}.$$

A direct calculation yields

$$\partial_s \widetilde{p}^{x'}(s,t,x,x') = -\widetilde{L}^{x'}\widetilde{p}^{x'}(s,t,x,x'), \quad t > s, \quad \widetilde{p}^{x'}(s,t,\cdot,x') \to \delta_{x'}(\cdot), \quad s \uparrow t, \tag{A.5}$$

where $\widetilde{L}^{x'}$ is the infinitesimal generator of (A.4). By (A.3) and (A.4), we derive from [13, (3.8)] that

$$p(s,t,x,x') = \widetilde{p}^{x'}(s,t,x,x') + \int_s^t \int_{\mathbb{R}^d} p(s,u,x,z) H(u,t,z,x')\mathrm{d}z\mathrm{d}u, \tag{A.6}$$

where

$$\begin{aligned}H(s,t,x,x') &:= (L - \widetilde{L}^{x'})\widetilde{p}^{x'}(s,t,x,x')\\ &= \langle b(x) - b(x'), \nabla \widetilde{p}^{x'}(s,t,x,x')\rangle + \frac{1}{2}\langle(\sigma\sigma^*)(x) - (\sigma\sigma^*)(x'), \nabla^2 \widetilde{p}^{x'}(s,t,x,x')\rangle_{\mathrm{HS}}.\end{aligned} \tag{A.7}$$

In (A.6), iterating for $p(s,u,x,z)$ gives

$$p(s,t,x,x') = \sum_{i=0}^{\infty} (\widetilde{p}^{x'} \otimes H^{(i)})(s,t,x,x'), \tag{A.8}$$

where $\widetilde{p} \otimes H^{(0)} := \widetilde{p}$ and $\widetilde{p}^{x'} \otimes H^{(i)} := (\widetilde{p}^{x'} \otimes H^{(i-1)}) \otimes H, i \geq 1$, with

$$(f \otimes g)(s,t,x,x') := \int_s^t \int_{\mathbb{R}^d} f(s,u,x,z) g(u,t,z,y)\mathrm{d}u\mathrm{d}z.$$

If we can claim that

$$|\widetilde{p} \otimes H^{(i)}|(s,t,x,x') \leq \frac{\mathrm{e}^{\frac{\|b\|_\infty^2 T}{2\hat{\lambda}_0}} \beta_T^i}{\Gamma(1+\frac{i}{2})} p_0(t-s,x,x'), \tag{A.9}$$

in which β_T, p_0 were introduced in (A.2), then (A.1) follows from (A.8) and (A.9). Below it suffices to show that (A.9) holds true. By means of (2.10) and $|a-b|^2 \geq \frac{1}{2}|a|^2 - |b|^2, a, b \in \mathbb{R}^d$, it follows from (1.2) and $\|b\|_\infty < \infty$ that

$$
\begin{aligned}
|\nabla \widetilde{p}|(s,t,x,x') &\leq \frac{\sqrt{\hat{\lambda}_0} \mathrm{e}^{\frac{\|b\|_\infty^2 T}{4\hat{\lambda}_0}}}{\check{\lambda}_0 \sqrt{t-s}} p_0(t-s,x,x') \\
\|\nabla^2 \widetilde{p}\|_{\mathrm{HS}}(s,t,x,x') &\leq \frac{(\sqrt{d}+\frac{4}{\mathrm{e}})\mathrm{e}^{\frac{\|b\|_\infty^2 T}{4\hat{\lambda}_0}}}{\check{\lambda}_0(t-s)} \frac{\mathrm{e}^{-\frac{|x'-x|^2}{8\hat{\lambda}_0(t-s)}}}{(2\pi\check{\lambda}_0(t-s))^{d/2}}.
\end{aligned} \tag{A.10}
$$

Thus, combining (2.10) with (A.10), besides $\|b\|_\infty < \infty$ and (1.3), enables us to obtain

$$
|H|(s,t,x,x') \leq \frac{2\hat{\lambda}_0 \Big\{ \|b\|_\infty/\sqrt{\hat{\lambda}_0} + L_0(d+2\sqrt{d}) \Big\} \mathrm{e}^{\frac{\|b\|_\infty^2 T}{4\hat{\lambda}_0}}}{\check{\lambda}_0\sqrt{t-s}} p_0(t-s,x,x'). \tag{A.11}
$$

By $\int_s^t (t-u)^{-\frac{1}{2}}(u-s)^\alpha \mathrm{d}u = (t-s)^{\alpha+\frac{1}{2}} B(1+\alpha, 1/2), t > s, \alpha > -1$, we have

$$
\Lambda_i(s,t) := \int_s^t \cdots \int_s^{u_{i-1}} (t-u_1)^{-\frac{1}{2}} \cdots (u_{i-1}-u_i)^{-\frac{1}{2}} \mathrm{d}u_i \cdots \mathrm{d}u_1 = \frac{(\pi(t-s))^{\frac{i}{2}}}{\Gamma(1+\frac{i}{2})}, \quad i \geq 1.
$$

Whence, taking advantage of $\|b\|_\infty < \infty$, (1.2), (A.11) as well as

$$
\int_{\mathbb{R}^d} p_0(u-s,x,z) p_0(t-u,y,z) \mathrm{d}z = \Big(\frac{8\hat{\lambda}_0}{\check{\lambda}_0} \Big)^d p_0(t-s,x,x'), \quad s < u < t
$$

yields (A.9). ☐

For $x, x' \in \mathbb{R}^d$ and $j \geq 0$, let $(\widetilde{X}_{i\delta}^{(\delta),j,x,x'})_{i \geq j}$ solve the following frozen EM scheme associated with (1.1)

$$
\widetilde{X}_{(i+1)\delta}^{(\delta),j,x,x'} = \widetilde{X}_{i\delta}^{(\delta),j,x,x'} + b(x')\delta + \sigma(x')(W_{(i+1)\delta} - W_{i\delta}), \quad i \geq j, \quad \widetilde{X}_{j\delta}^{(\delta),j,x,x'} = x.
$$

Write $\widetilde{p}^{(\delta),x'}(j\delta, j'\delta, x, y)$ by the transition density of $\widetilde{X}_{j'\delta}^{(\delta),j,x,x'}$ at the point y, given $\widetilde{X}_{j\delta}^{(\delta),j,x,x'} = x$.

The following lemma reveals explicit upper bounds of coefficients with regard to Gaussian bound of the discrete-time EM scheme.

Lemma A.2. *Under $\|b\|_\infty < \infty$ and* (**A3**), *for any $0 \leq j < j' \leq \lfloor T/\delta \rfloor$*

$$
p^{(\delta)}(j\delta, j'\delta, x, x') \leq \mathrm{e}^{\frac{\|b\|_\infty T}{2\hat{\lambda}_0}} \sum_{k=0}^\infty \frac{\Big(\sqrt{\pi T}\hat{C}_T((1+24d)\hat{\lambda}_0/\check{\lambda}_0)^d\Big)^k}{\Gamma(1+\frac{k}{2})} \frac{\mathrm{e}^{-\frac{|x'-x|^2}{4(1+24d)\hat{\lambda}_0(j'-j)\delta}}}{(2\pi\check{\lambda}_0(j'-j)\delta)^{d/2}}. \tag{A.12}
$$

Proof. To obtain (A.12), we refine the proof of [15, Lemma 4.1]. For $\psi \in C^2(\mathbb{R}^d;\mathbb{R})$ and $j \geq 0$, set

$$
(\mathscr{L}_{j\delta}^{(\delta)}\psi)(x) := \delta^{-1}\{\mathbb{E}(\psi(X_{(j+1)\delta}^{(\delta)})|X_{j\delta}^{(\delta)} = x) - \psi(x)\}, (\hat{\mathscr{L}}_{j\delta}^{(\delta)}\psi)(x) := \delta^{-1}\{\mathbb{E}\psi(\widetilde{X}_{(j+1)\delta}^{(\delta),j,x,x'}) - \psi(x)\}
$$

and

$$H^{(\delta)}(j\delta, j'\delta, x, x') := (\mathscr{L}_{j\delta}^{(\delta)} - \hat{\mathscr{L}}_{j\delta}^{(\delta)})\widetilde{p}^{(\delta),x'}((j+1)\delta, j'\delta, x, x'), \ j' \geq j+1.$$

In what follows, let $0 \leq j < j' \leq \lfloor T/\delta \rfloor$. According to [13, Lemma 3.6], we have

$$p^{(\delta)}(j\delta, j'\delta, x, x') = \sum_{k=0}^{j'-j} (\widetilde{p}^{(\delta),x'} \otimes_\delta H^{(\delta),(k)})(j\delta, j'\delta, x, x'), \tag{A.13}$$

where $\widetilde{p}^{(\delta),x'} \otimes_\delta H^{(\delta),(0)}) = \widetilde{p}^{(\delta),x'}$, $H^{(\delta),(k)} = H^{(\delta)} \otimes_\delta H^{(\delta),(k-1)}$ with $\otimes_\delta$ being the convolution type binary operation defined by

$$(f \otimes_\delta g)(j\delta, j'\delta, x, x') = \delta \sum_{k=j}^{j'-1} \int_{\mathbb{R}^d} f(j\delta, k\delta, x, u) g(k\delta, j'\delta, u, x') \mathrm{d}u.$$

If the assertion

$$H^{(\delta)}(j\delta, j'\delta, x, x') \leq \frac{\hat{C}_T}{\sqrt{(j'-j)\delta}} \frac{\mathrm{e}^{-\frac{|x'-x|^2}{4(1+24d)\hat{\lambda}_0(j'-j)\delta}}}{(2\pi\check{\lambda}_0(j'-j)\delta)^{d/2}} \tag{A.14}$$

holds true, where $\hat{C}_T$ was given in (2.1), then (A.12) follows due to (A.13) by an induction argument. So, in order to complete the proof of Lemma A.2, it remains to verify (A.14). First of all, we show (A.14) for $j' = j+1$. By the definition of $H^{(\delta)}$, observe from (1.2) that

$$\begin{aligned}
|H^{(\delta)}|(j\delta, (j+1)\delta, x, x') &= \frac{1}{\delta}|p^{(\delta)} - \widetilde{p}^{(\delta),x'}|(j\delta, (j+1)\delta, x, x') \\
&\leq \frac{1}{\delta(2\pi\check{\lambda}_0\delta)^{d/2}} \bigg\{ \Big| \mathrm{e}^{-\frac{1}{2\delta}|(\sigma\sigma^*)^{-\frac{1}{2}}(x)(x'-x-b(x)\delta)|^2} - \mathrm{e}^{-\frac{1}{2\delta}|(\sigma\sigma^*)^{-\frac{1}{2}}(x)(x'-x-b(x')\delta)|^2} \Big| \\
&\quad + \Big| \mathrm{e}^{-\frac{1}{2\delta}\langle(\sigma\sigma^*)^{-1}(x)(x'-x-b(x')\delta), x'-x-b(x')\delta\rangle} - \mathrm{e}^{-\frac{1}{2\delta}\langle(\sigma\sigma^*)^{-1}(x')(x'-x-b(x')\delta), x'-x-b(x')\delta\rangle} \Big| \\
&\quad + \frac{1}{2\check{\lambda}_0^d} \mathrm{e}^{-\frac{1}{2\delta}|(\sigma\sigma^*)^{-\frac{1}{2}}(x')(x'-x-b(x')\delta)|^2} |\det((\sigma\sigma^*)(x')) - \det((\sigma\sigma^*)(x))| \bigg\} \\
&=: \frac{1}{\delta(2\pi\check{\lambda}_0\delta)^{d/2}} \{\Lambda_1 + \Lambda_2 + \Lambda_3\}.
\end{aligned}$$

Next, we aim to estimate $\Lambda_1, \Lambda_2, \Lambda_3$, one-by-one. By $\|b\|_\infty < \infty$, (1.2) and (2.10), it follows from the first fundamental theorem of calculus that

$$|\Lambda_1| \leq 2\sqrt{\delta/\check{\lambda}_0}\|b\|_\infty \mathrm{e}^{\frac{\|b\|_\infty^2\delta}{\check{\lambda}_0}} \mathrm{e}^{-\frac{|x-x'|^2}{8\hat{\lambda}_0\delta}}. \tag{A.15}$$

(1.2) and (1.3) imply

$$\|(\sigma\sigma^*)^{-1}(x) - (\sigma\sigma^*)^{-1}(x')\|_{\mathrm{HS}} \leq 2\check{\lambda}_0^{-2}\sqrt{d\hat{\lambda}_0} L_0 |x-x'|.$$

This, by invoking $|\mathrm{e}^a - \mathrm{e}^b| \leq \mathrm{e}^{a\vee b}|a-b|$, $a, b \in \mathbb{R}$, and utilizing $\|b\|_\infty < \infty$, (1.2) and (2.10), yields

$$|\Lambda_2| \leq 4\sqrt{d\delta} L_0 (\hat{\lambda}_0/\check{\lambda}_0)^2 \mathrm{e}^{\frac{\|b\|_\infty^2\delta}{4\check{\lambda}_0}} \mathrm{e}^{-\frac{|x-x'|^2}{16\hat{\lambda}_0\delta}}. \tag{A.16}$$

Also, making use of $\|b\|_\infty<\infty$, (1.2) and (2.10), in addition to

$$|\det((\sigma\sigma^*)(x))-\det((\sigma\sigma^*)(x'))|\le 2d^{\frac{d}{2}+1}d!\hat{\lambda}_0^{d-\frac12}L_0|x-x'|,$$

due to (1.2) and (1.3), we arrive at

$$|\Lambda_3|\le\sqrt{2}d^{\frac{d}{2}+1}d!(\hat{\lambda}_0/\breve{\lambda}_0)^dL_0\sqrt{\delta}\mathrm{e}^{\frac{\|b\|_\infty^2\delta}{2\hat{\lambda}_0}}\mathrm{e}^{-\frac{|x'-x|^2}{8\hat{\lambda}_0\delta}}. \tag{A.17}$$

We therefore conclude that (A.14) holds with $j'=j+1$ by taking (A.15)-(A.17) into account. In the sequel, we are going to show that (A.14) is still available for $j'>j+1$. According to the notion of $H^{(\delta)}$,

$$\begin{aligned}&H^{(\delta)}(j\delta,j'\delta,x,x')\\&=\frac{1}{\delta(2\pi)^{m/2}}\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|z|^2}{2}}\Big\{\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x+\Gamma_z(x),x')-\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x,x')\Big\}\mathrm{d}z\\&\quad-\frac{1}{\delta(2\pi)^{m/2}}\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|z|^2}{2}}\Big\{\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x+\Gamma_z(x'),x')-\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x,x')\Big\}\mathrm{d}z,\end{aligned}$$

where $\Gamma_z(x):=b(x)\delta+\sqrt{\delta}\sigma(x)z, x\in\mathbb{R}^d, z\in\mathbb{R}^m$. By Taylor's expansion, we further have

$$\begin{aligned}&H^{(\delta)}(j\delta,j'\delta,x,x')\\&=\frac{1}{\delta(2\pi)^{m/2}}\bigg\{\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|z|^2}{2}}\langle\nabla\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x,x'),\Gamma_z(x)-\Gamma_z(x')\rangle\mathrm{d}z\\&\quad+\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|z|^2}{2}}\langle\nabla^2\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x,x'),(\Gamma_z\Gamma_z^*)(x)-(\Gamma_z\Gamma_z^*)(x')\rangle_{\mathrm{HS}}\mathrm{d}z\bigg\}\\&\quad+\frac{1}{2\delta(2\pi)^{m/2}}\int_{\mathbb{R}^m}\int_0^1(1-\theta)^2\mathrm{e}^{-\frac{|z|^2}{2}}\Big\{\nabla^3_{\Gamma_z(x)}\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x+\theta\Gamma_z(x),x')\\&\quad-\nabla^3_{\Gamma_z(x')}\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x+\theta\Gamma_z(x'),x')\Big\}\mathrm{d}\theta\mathrm{d}z\\&=:\Pi_1+\Pi_2+\Pi_3,\end{aligned}$$

where ∇^i means the i-th order gradient operator. Employing

$$\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|z|^2}{2}}\mathrm{trace}(A\sigma(x)zz^*\sigma(x)\mathrm{d}z=\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|z|^2}{2}}z^*\sigma^*(x)A\sigma(x)z\mathrm{d}z=(2\pi)^{m/2}\mathrm{trace}(\sigma^*(x)A\sigma(x))$$

for a symmetric $d\times d$-matrix and $\int_{\mathbb{R}^m}\mathrm{e}^{-\frac{|z|^2}{2}}z\mathrm{d}z=\mathbf{0}$ gives

$$\Pi_1+\Pi_2=H((j+1)\delta,j'\delta,x,x')+\frac{\delta}{2}\langle\nabla^2\widetilde{p}^{(\delta),x'}((j+1)\delta,j'\delta,x,x'),(bb^*)(x)-(bb^*)(x')\rangle_{\mathrm{HS}},$$

where H was defined as in (A.7) with $p^{x'}$ replaced by $\widetilde{p}^{(\delta),x'}$. (A.10) and (A.11) enable us to obtain

$$|\Pi_1|+|\Pi_2|\le\frac{2^{\frac{d+1}{2}}\mathrm{e}^{\frac{\|b\|_\infty^2T}{4\hat{\lambda}_0}}}{\breve{\lambda}_0}\bigg\{2\sqrt{\hat{\lambda}_0}\|b\|_\infty+(\|b\|_\infty^2+2\hat{\lambda}_0L_0\sqrt{d})(\sqrt{d}+2)\bigg\}\frac{p_0((j'-j)\delta,x,x')}{\sqrt{(j'-j)\delta}}. \tag{A.18}$$

Note that Π_3 can be reformulated as below

$$\Pi_3 = \frac{1}{2\delta(2\pi)^{m/2}}\int_{\mathbb{R}^m}\int_0^1 (1-\theta)^2 \mathrm{e}^{-\frac{|z|^2}{2}}\Big\{\nabla^3_{\Gamma_z(x)}\widetilde{p}^{(\delta),x'}((j+1)\delta, j'\delta, x+\theta\Gamma_z(x'), x')$$
$$-\nabla^3_{\Gamma_z(x')}\widetilde{p}^{(\delta),x'}((j+1)\delta, j'\delta, x+\theta\Gamma_z(x'), x')\Big\}\mathrm{d}\theta\mathrm{d}z$$
$$+\frac{1}{2\delta(2\pi)^{m/2}}\int_{\mathbb{R}^m}\int_0^1 (1-\theta)^2 \mathrm{e}^{-\frac{|z|^2}{2}}\Big\{\nabla^3_{\Gamma_z(x)}\widetilde{p}^{(\delta),x'}((j+1)\delta, j'\delta, x+\theta\Gamma_z(x), x')$$
$$-\nabla^3_{\Gamma_z(x)}\widetilde{p}^{(\delta),x'}((j+1)\delta, j'\delta, x+\theta\Gamma_z(x'), x')\Big\}\mathrm{d}\theta\mathrm{d}z =: \Pi_{31}+\Pi_{32}.$$

By means of (1.2), (1.3) and (2.10), it follows that

$$|\Pi_{31}| \le \frac{2^{m+\frac{d+21}{2}}(L_0+2\|b\|_\infty)(\|b\|_\infty^2+d\hat{\lambda}_0)\Big(1+\sqrt{2(1+4d)\hat{\lambda}_0}\Big)\mathrm{e}^{\frac{3\|b\|_\infty^2 T}{8d\hat{\lambda}_0}}}{\check{\lambda}_0^{\frac{3}{2}}((j'-j)\delta)^{\frac{1}{2}}} \times \frac{\mathrm{e}^{-\frac{|x'-x|^2}{8(1+4d)\hat{\lambda}_0(j'-j)\delta}}}{(2\pi\check{\lambda}_0(j'-j)\delta)^{d/2}}, \tag{A.19}$$

Also, by exploiting (1.2), and (2.10), we infer from Taylor expansion

$$|\Pi_{32}| \le \frac{2^{m+\frac{d+23}{2}}(L_0+2\|b\|_\infty)(\|b\|_\infty^3+(d\hat{\lambda}_0)^{\frac{3}{2}})\Big(1+\sqrt{2(1+24d)\hat{\lambda}_0}\Big)\mathrm{e}^{\frac{(6\|b\|_\infty^2+\|b\|_\infty)T}{24d\hat{\lambda}_0}}}{\check{\lambda}_0^{2}((j'-j)\delta)^{\frac{1}{2}}} \times \frac{\mathrm{e}^{-\frac{|x'-x|^2}{4(1+24d)\hat{\lambda}_0(j'-j)\delta}}}{(2\pi\check{\lambda}_0(j'-j)\delta)^{d/2}}. \tag{A.20}$$

Consequently, (A.14) follows from (A.18), (A.19), and (A.20).

Acknowledgment

We are indebted to the referees for their valuable comments which have greatly improved our earlier version. The third author is supported by the Emerging Interdisciplinary Project of CUFE.

□

References

[1] Bao, J., Huang, X., Yuan, C., Convergence rate of Euler–Maruyama Scheme for SDEs with Hölder–dini continuous drifts, *J. Theoret. Probab.*, **32** (2019), 848–871.

[2] Dareiotis, K., Kumar, C., Sabanis, S., On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations, *SIAM J. Numer. Anal.*, **54** (2016), 1840–1872.

[3] Flandoli, M., Gubinelli, M., Priola, E., Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift, *Bull. Sci. Math.*, **134** (2010), 405–422.

[4] Gyöngy, I., Martinez, T., On stochastic differential equations with locally unbounded drift, *Czechoslovak Math.J.*, **51** (2001), 763–783.

[5] Gyöngy, I., Rásonyi, M., A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients, *Stoch. Process. Appl.*, **121** (2011), 2189–2200.

[6] Halidias, N., Kloeden, P. E., A note on the Euler-Maruyama scheme for stochastic differential equations with a discontinuous monotone drift coefficient, *BIT*, **48** (2008), 51–59.

[7] Higham, D. J., Mao, X., Stuart, A. M., Strong convergence of Euler-type methods for nonlinear stochastic differential equations, *SIAM J. Numer. Anal.*, **40** (2002), 1041–1063.

[8] Higham, D. J., Mao, X., Yuan, C., Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations, *SIAM J. Numer. Anal.*, **45** (2007), 592–609.

[9] Huang, X., Wang, F.-Y., Distribution Dependent SDEs with Singular Coefficients, *Stoch. Process. Appl.*, **129** (2019), 4747–4770.

[10] Hutzenthaler, M., Jentzen, A., Kloeden, P. E., Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **467** (2011), 1563–1576.

[11] Hutzenthaler, M., Jentzen, A., Kloeden, P. E., Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients, *Ann. Appl. Probab.*, **22** (2012), 1611–1641.

[12] Kloeden, P. E., Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.

[13] Konakov, V., Mammen, E., Local limits theorems for transition densities of Markov chains converging to diffusions, *Probab. Theory Relat. Fields*, **117** (2000), 551-587.

[14] Krylov, N. V., Röckner, M., Strong solutions of stochastic equations with singular time dependent drift, *Probab. Theory Related Fields*, **131** (2005), 154–196.

[15] Lemaire, V., Menozzi, S., On some non asymptotic bounds for the Euler scheme, *Electron. J. Probab.*, **15** (2010), 1645–1681.

[16] Leobacher, G., Szölgyenyi, M., A strong order 1/2 method for multidimensional SDEs with discontinuous drift, *Ann. Appl. Probab.*, **27** (2017), 2383–2418.

[17] Leobacher, G., Szölgyenyi, M. , Convergence of the Euler-Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient, *Numer. Math.*, **138** (2018), 219–239.

[18] Mao, X., The truncated Euler-Maruyama method for stochastic differential equations, *J. Comput. Appl. Math.*, **290** (2015), 370–384.

[19] Mao, X., Yuan, C., *Stochastic differential equations with Markovian switching*, Imperial College Press, London, 2006.

[20] Müller-Gronbach, T., Yaroslavtseva, L., On the performance of the Euler-Maruyama scheme for SDEs with discontinuous drift coefficient, *Ann. Inst. H. Poincaré Probab. Statist.*, **56** (2020), 1162–1178.

[21] Neuenkirch, A., Szölgyenyi, M., Szpruch, L., An adaptive Euler-Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis, *SIAM J. Numer. Anal.*, **57** (2019), 378–403.

[22] Ngo, H-L., Taguchi, D., Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients, *Math. Comp.*, **85** (2016), 1793–1819.

[23] Ngo, H.-L., Taguchi, D., On the Euler–Maruyama approximation for one-dimensional stochastic differential equations with irregular coefficients, *IMA J. Numer. Anal.*, **37** (2017), 1864–1883.

[24] Pamen, O.M., Taguchi, D., Strong rate of convergence for the Euler–Maruyama approximation of SDEs with Hölder continuous drift coefficient. *Stoch. Process. Appl.*, **127** (2017), 2542–2559.

[25] Röckner, M., Zhang, X., Well-posedness of distribution dependent SDEs with singular drifts, arXiv:1809.02216.

[26] Shao, J., Weak convergence of Euler-Maruyama's approximation for SDEs under integrability condition, arXiv:1808.07250.

[27] Shigekawa, I.: Stochastic Analysis, Translations of Mathematical Monographs, 224, Iwanami Series in Modern Mathematics. American Mathematical Society, Providence, 2004.

[28] Stein, Elias M., *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, 1970.

[29] Suo, Y., Yuan, C., Zhang, S.-Q., Weak convergence of Euler scheme for SDEs with singular drift, arXiv:2005.04631, 2020.

[30] Xie, L., Zhang, X., Ergodicity of stochastic differential equations with jumps and singular coefficients, *Ann. Inst. H. Poincaré Probab. Statist.*, **56** (2020), 175–229.

[31] Yan, L., The Euler scheme with irregular coefficients, *Ann. Probab.*, **30** (2002), 1172–1194.

[32] Zhang, X., Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients, *Electron. J. Probab.*, **16** (2011), 1096–1116.

[33] Zhang, X., Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients, *Rev. Mat. Iberoam.*, **29** (2013), 25–52.

[34] Zvonkin, A. K., A transformation of the phase space of a diffusion process that removes the drift, *Math. Sb.*, **93** (1974), 129-149.