Weak convergence of Euler scheme for SDEs with low regular drift

Yongqiang Suo[a)], Chenggui Yuan[a)], Shao-Qin Zhang[b)]

[a)]Mathematic department, Swansea University, Bay Campus, SA1 8EN, UK

Email: 971001@swansea.ac.uk   C.Yuan@swansea.ac.uk

[b)]School of Statistics and Mathematics

Central University of Finance and Economics, Beijing 100081, China

Email: zhangsq@cufe.edu.cn

Abstract

In this paper, we investigate the weak convergence rate of Euler-Maruyama's approximation for stochastic differential equations with low regular drifts. Explicit weak convergence rates are presented if drifts satisfy an integrability condition including discontinuous functions which can be non-piecewise continuous or in some fractional Sobolev space.

AMS Subject Classification (2020): 60H10, 34K26, 65C30

Keywords: low regular coefficients, weak convergence rate, Euler-Maruyama's approximation

1 Introduction

Stochastic differential equations (SDEs for short) with singular coefficients have been extensively studied recently, see [12, 25, 26, 27, 28] and references therein. Meanwhile, in order for one to understand the numerical approximation of SDEs with irregular coefficients, numerical schemes have been established. The strong and weak convergence rates of Euler-Maruyama's (abbreviated as EM's) scheme for irregular SDEs were obtained, see [2, 3, 6, 7, 11, 13, 14, 16, 17, 19, 20, 23] for instance. [5, 8, 9, 15, 18, 10, 22] investigated L^p-approximation of solutions to SDEs with a discontinuous drift and obtained the corresponding L^p-error rates under differential assumptions on coefficients. More precisely, [9] investigated the L^p-error rate at least $1/2$ with $p \in [1, \infty)$ for the scalar SDEs with a piecewise Lipschitz drift and a Lipschitz diffusion coefficient that is non-zero at discontinuity points of the drift coefficient, this result has been extended to the case of scalar jump-diffusion SDEs in [22]. Based on the assumptions in [9, 22], [8, 10] showed the L^p-error rate at least $3/4$ under additional piecewise smoothness assumptions on the coefficients, where they employed a novel technique by studying equations with coupled noise, and also showed that besides the assumptions in [10], the L^p-error rate $\frac{3}{4}$ can not in general be improved even under further piecewise regularity.)???? the L^p-error rate 3/4 can be in general not be improved even when further piecewise regularity were imposed on coefficients of the scalar SDEs additionally to the assumptions in [10]. Under the condition of the Sobolev-Slobodeckij-type regularity of order $\kappa \in (0,1)$, [18] obtained the L^2-error rate $\min\{3/4, (1+\kappa)/2\} - \varepsilon$ (for arbitrarily small $\varepsilon > 0$) of the equidistant EM's scheme for scalar SDEs with irregular drift and additive noise by using an explicit Zvonkin-type transformation and the Girsanov transformation. By using a suitable non-equidistant discretization, [18] also yields the strong convergence order of $\frac{1+\kappa}{2} - \varepsilon$ for the corresponding EM's scheme.

In this paper, we shall investigate the weak error of EM's scheme for the following SDE on $\mathbb{R}^d$

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

where $(W_t)_{t\geq 0}$ is a d-dimensional Brownian motion with respect to a complete filtration probability space $(\Omega, (\mathscr{F}_t)_{t\geq 0}, \mathscr{F}, \mathbb{P})$. The associated EM's scheme reads as follows: for any $\delta \in (0,1)$,

$$\mathrm{d}X_t^{(\delta)} = b(X_{t_\delta}^{(\delta)})\mathrm{d}t + \sigma \mathrm{d}W_t, \ X_0^{(\delta)} = x, \tag{1.2}$$

where $t_\delta = [t/\delta]\delta$ and $[t/\delta]$ denotes the integer part of t/δ. The weak error is concerned with the convergence of the distribution of the EM's scheme. Precisely, it is concerned with the approximation of $\mathbb{E}f(X_t)$ by $\mathbb{E}f(X_t^{(\delta)})$ for a given function f. The weak error has been obtained for some SDEs with discontinuous drifts in [7, 11, 21]. It is worth noting that the test function f in these references is assumed to be Hölder continuous. When the test function f was relaxed to be just measurable and bounded, the result of weak convergence rate of EM's scheme was obtained in [1] for SDEs with smooth coefficients. Recently, [4, 23] investigated the weak convergence rate of EM's scheme for SDEs with irregular coefficients by using Girsanov's transformation, and [3] used an integrability condition to obtain strong convergence rates for multidimensional SDEs with the aid of the Krylov estimate and the heat kernel estimate of Gaussian type process established by the parametrix method in [16]. Inspired by [3] and [4, 23], we shall give a note on the weak error for (1.1) with b satisfying an integrability condition (see (H2) below) which is similar to (A2') in [3], and the given function f being only bounded and measurable on $\mathbb{R}^d$. Functions satisfying (H2) is called the Gaussian-Besov regularity, see comments after (A2') of [3]. Discontinuous functions can also satisfy some kind of Sobolev-Slobodeckij-type regularity which subjects to the Gaussian-Besov regularity indicated by (H2), see examples in subsection 2.2 or [3, Example 4.3]. Thus we say the drift term b is "low regular" instead of irregular here. Moreover, (H2) also allows that the drift term satisfies a sub-linear growth condition (see (H1) below).

The remainder of this paper is organized as follows: The main result is presented in Section 2. All the proofs are given in Section 3.

2 Main Result and Examples

2.1 Assumption and Main result

Let $|\cdot|$ be the Euclidean norm, $\langle\cdot,\cdot\rangle$ be the Euclidean product. $\|\cdot\|$ denotes the operator norm. We denote $\|f\|_\infty = \sup_{x\in\mathbb{R}^d}|f(x)|$ for any bounded and measurable function f on $\mathbb{R}^d$. Throughout this paper, we assume that the coefficients of (1.1) satisfy the following assumptions:

(H1) $b:\mathbb{R}^d\to\mathbb{R}^d$ is measurable and σ is an invertible $d\times d$-matrix. There exist $\beta\in[0,1)$ and nonnegative constants L_1, L_2 such that

$$|b(x)|\leq L_1+L_2|x|^\beta,\ x\in\mathbb{R}^d.$$

(H2) There exist $p_0\geq 2$, $\alpha>0$ and $\phi\in C((0,+\infty);(0,+\infty))$ which is non increasing on $(0,l)$ and $\int_0^l\phi^2(s)\mathrm{d}s<\infty$ for some $l>0$ such that

$$\sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}|b(y)-b(x)|^{p_0}\frac{\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}}{s^{\frac{d}{2}}r^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\leq(\phi(s)r^\alpha)^{p_0},\ s>0,r\in[0,1].$$

It is clear that (1.2) also has a unique strong solution. The index α in (H2) is used to characterize the order of the continuity and the function ϕ is used to characterize the type of the continuity. From examples in the next subsection, it is clear that functions sharing the same order of continuity can have different types of continuity.

We now formulate the main result.

Theorem 2.1. *Assume (H1)-(H2). Then for any $T>0$ and any bounded measurable function f on $\mathbb{R}^d$, there exists a constant $C_{T,p_0,\sigma,x}>0$ such that*

$$|\mathbb{E}f(X_t)-\mathbb{E}f(X_t^{(\delta)})|\leq C_{T,p_0,\sigma,x}\|f\|_\infty\delta^\alpha,\ t\in[0,T],\tag{2.1}$$

where p_0 is defined in (H2). If the growth condition in (H1) is replaced by $|b(x)|\leq L_1+L_2|x|$, then (2.1) also holds for T,L_2,p_0 and σ satisfying

$$TL_2\|\sigma^{-1}\|\|\sigma\|\frac{\sqrt{2(p_0+1)(p_0+3)}}{p_0-1}<1.\tag{2.2}$$

Remark 2.1. By [28, Theorem 1.1], (1.1) has a unique strong solution under (H1). It is also clear that (1.2) has a unique pathwise solution.

For the bounded and irregular b, there are many results on strong and weak error of EM's scheme, see e.g. [3, 7, 11, 18] and references therein, and the weak error can not be derived from the strong error directly if f is just a bounded and measurable function. We would like to highlight that authors in [18] has obtained the rate of strong convergence for one-dimensional SDEs if b is in $L^1(\mathbb{R})$ and bounded, and satisfies the Sobolev-Slobodeckij-type regularity.

This result is better than the present one in Theorem 2.1. However, results in [18] relied on an Zvonkin-type transformation which can be given explicitly in one dimension, and some favourable properties are lost in high dimensions. Here, only Girsanov's transformation is used, while we allow that the SDE is multi-dimensional and that the drift satisfies a sub-linear growth condition. Moreover, we obtain the same convergence rate when b has linear growth, as long as (2.2) holds. Our assumption (H2) also includes the Sobolev-Slobodeckij-type regularity, see Example 2.4 in the next subsection. To obtain higher convergence rate as in [18], it seems that we need to make a deep investigation on the Zvonkin-type transformation.

In the assumption (H2), if α is a decreasing function of p_0, then we can choose $p_0 = 2$ and obtain the highest convergence rate in (2.1), see Example 2.3.

Remark 2.2. In [3], the strong convergence and the convergence rate are investigated with the drift satisfying an integrability condition and boundedness. Here we obtain the weak convergence rate of EM's scheme, where the drift does not need to be bounded and the test function f in (2.1) is only bounded and measurable. Moreover, the convergence rate is better than the rate obtained in [3, Theorem 1.3].

From examples in the next subsection, one can see that the drift could be discontinuous. This means that we have extended the results in [1] where the coefficients must be smooth. However, our result is not optimal in the smooth case since the classical order of the weak error is $\alpha = 1$ for SDEs with smooth coefficients in [1].

Remark 2.3. In [19, 21], authors considered the weak convergence rate of the EM's scheme for (1.1) with the drift b is of sub-linear growth and $b = b^H + b^A$, where b^H is α-Hölder for some $\alpha \in (0, 1)$ and b^A belongs to a class $\mathcal{A}$ which does not contain any nontrivial Hölder continuous functions. The order of the convergence rate obtained in [21] is $\frac{\alpha}{2} \wedge \frac{1}{4}$, even if $b^A \equiv 0$. However, the order of the convergence rate in Theorem 2.1 comes from the continuity order α in (H2), and it can be greater than $\frac{1}{4}$.

The class $\mathcal{A}$ in [19, 21] is given by $\mathcal{A}$-approximation. In contrast to the $\mathcal{A}$-approximation, our condition (H2) is more explicit. Moreover, for any time independent function ζ in the class $\mathcal{A}$ of [19], it satisfies (H2) with $p_0 = 2, \alpha = \frac{1}{4}$ and $\phi(s) = s^{-\frac{1}{4}}\sqrt{1 + \sqrt{s}}$. In fact, according to [19, Definition 2.1], ζ is bounded and there exists a sequence $\{\zeta_n\}_{n\geq 1}$ such that $\zeta_n \in C^1(\mathbb{R}^d)$ is uniformly bounded and converges to ζ locally in $L^1(\mathbb{R}^d)$, and there exists $K > 0$ such that

$$\sup_{n\geq 1,\ a\in\mathbb{R}^d} \int_{\mathbb{R}^d} \|\nabla\zeta_n(x + a)\| \frac{e^{-\frac{|x|^2}{s}}}{s^{(d-1)/2}} \mathrm{d}x \leq K(1 + \sqrt{s}). \tag{2.3}$$

Noting that

$$\sup_{x\geq 0}(x^{\gamma'} \mathrm{e}^{-\gamma x^2}) = \left(\frac{\gamma'}{2\,\mathrm{e}\,\gamma}\right)^{\gamma'/2}, \qquad \gamma', \gamma > 0, \tag{2.4}$$

we then obtain from (2.3) and (2.4) that

$$\begin{aligned}
&\int_{\mathbb{R}^d\times\mathbb{R}^d} |\zeta(x)-\zeta(y)|^2 \frac{e^{-\frac{|x-z|^2}{s}-\frac{|x-y|^2}{r}}}{(sr)^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\quad\le \|\zeta\|_\infty \lim_{n\to+\infty}\int_{\mathbb{R}^d\times\mathbb{R}^d} |\zeta_n(x)-\zeta_n(y)| \frac{e^{-\frac{|x-z|^2}{s}-\frac{|x-y|^2}{r}}}{(sr)^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\quad= \|\zeta\|_\infty \lim_{n\to+\infty}\int_0^1\int_{\mathbb{R}^d\times\mathbb{R}^d} \|\nabla_{y-x}\zeta_n(x+\theta(y-x))\| \frac{e^{-\frac{|x-z|^2}{s}-\frac{|x-y|^2}{r}}}{(sr)^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\mathrm{d}\theta\\
&\quad\le \|\zeta\|_\infty \lim_{n\to+\infty}\int_0^1 \left(\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\|\nabla\zeta_n(x+\theta h+z)\|\frac{e^{-\frac{|x|^2}{s}}}{s^{d/2}}\mathrm{d}x\right)\frac{|h|e^{\frac{-|h|^2}{r}}}{r^{\frac{d}{2}}}\mathrm{d}h\right)\mathrm{d}\theta\\
&\quad\le \|\zeta\|_\infty \int_{\mathbb{R}^d} Ks^{-\frac{1}{2}}(1+\sqrt{s})\frac{|h|e^{\frac{-|h|^2}{r}}}{r^{\frac{d}{2}}}\mathrm{d}h\\
&\quad\le C\|\zeta\|_\infty s^{-\frac{1}{2}}(1+\sqrt{s})r^{\frac{1}{2}},
\end{aligned}$$

where the constant C is independent of z. The class $\mathcal{A}$ used in [21] allows functions in $\mathcal{A}$ can be just exponentially bounded. However, they assume that the drift is only sublinear growth. There is no example showing that the class $\mathcal{A}$ used in [21] can contain functions which are more irregular than functions in $\mathcal{A}$ of [19].

2.2 Illustrative examples

In this subsection, we give several examples to illustrate the condition (H2) and the order of the convergence rate α. Before our concrete examples, we give some comments on (H2). According to the proof of Theorem 2.1, $X_t^{(\delta)}$ and X_t are weak solutions of the equation $Y_t = X_0 + \sigma W_t$ in suitable probability spaces. By using the Girsanov transformation, the error between $X_t^{(\delta)}$ and X_t mainly comes from the following term

$$\Big|\int_0^T \langle \sigma^{-1}(b(Y_s)-b(Y_{s_\delta})), \mathrm{d}W_s\rangle\Big|.$$

Since Y_t is a Gaussian process, (H2) is convenient to estimate the above stochastic integral, see (3.20) and the proof of Lemma 3.3 for more details. Comparing with the definition of Besov space (see [24, (1.13)]), we call functions satisfying (H2) the Gaussian-Besov class. The exponential terms in the integrand of (H2) allow that b can grow to infinity as $|x|$ increases.

Example 2.2. *If b is the Hölder continuous with exponent β, i.e.*

$$|b(y)-b(x)| \le L|x-y|^\beta,$$

then (H2) holds with $\alpha = \frac{\beta}{2}$ and a constant function $\phi(s)$. It is clear that b has sublinear growth if $\beta < 1$. Then for any $T > 0$, (2.1) holds with $\alpha = \frac{\beta}{2}$.

Proof. By the Hölder continuity and (2.4), the assertion follows from the following inequality

$$
\begin{aligned}
&\sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}|b(y)-b(x)|^{p_0}\frac{\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}}{s^{\frac{d}{2}}r^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\le L^{p_0}\sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}|y-x|^{\beta p_0}\frac{\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}}{s^{\frac{d}{2}}r^{\frac{d}{2}}}\mathrm{d}x\mathrm{d}y\\
&\le L^{p_0}\frac{1}{s^{\frac{d}{2}}r^{\frac{d}{2}}}\left(\frac{\beta p_0 r}{\mathrm{e}}\right)^{\frac{\beta p_0}{2}}\sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d}\mathrm{e}^{-\frac{|x-z|^2}{s}}\mathrm{e}^{-\frac{|y-x|^2}{2r}}\mathrm{d}x\mathrm{d}y\\
&\le CL^{p_0}\left(\frac{\beta p_0 r}{\mathrm{e}}\right)^{\frac{\beta p_0}{2}}.
\end{aligned}
$$

□

The following example shows that (H2) can hold even if the drift term b is not piecewise continuous.

Example 2.3. *Let A be the Smith-Volterra-Cantor set on $[0,1]$, which is constructed in the following way. The first step, we let $I_{1,1}=\left(\frac{3}{8},\frac{5}{8}\right)$, $J_{1,1}=\left[0,\frac{3}{8}\right]$, $J_{1,2}=[\frac{5}{8},1]$ and remove the open interval $I_{1,1}$ from $[0,1]$. The second step, we remove the middle $\frac{1}{4^2}$ open intervals, denoting by $I_{2,1}$ and $I_{2,2}$, from $J_{1,1}$ and $J_{1,2}$ respectively, i.e. $I_{2,1}=\left(\frac{5}{32},\frac{7}{32}\right)$, $I_{2,2}=\left(\frac{25}{32},\frac{27}{32}\right)$. The intervals left are denoted by $J_{2,1},J_{2,2},J_{2,3},J_{2,4}$, i.e.*

$$
J_{2,1}=\left[0,\frac{5}{32}\right],J_{2,2}=\left[\frac{7}{32},\frac{3}{8}\right],J_{2,3}=\left[\frac{5}{8},\frac{25}{32}\right],J_{2,4}=\left[\frac{27}{32},1\right].
$$

For the n-th step, we remove the middle $\frac{1}{4^n}$ open intervals $I_{n,1},\cdots,I_{n,2^{n-1}}$ from $J_{n-1,1},\cdots,J_{n-1,2^{n-1}}$ respectively, and the intervals left are denoted by $J_{n,1},\cdots,J_{n,2^n}$. Let

$$
A=\bigcap_{n=1}^{\infty}\left(\bigcup_{k=1}^{2^n}J_{n,k}\right).
$$

Then A is a nowhere dense set and the Lebesgue measure of A is $1/2$. Define

$$
\begin{aligned}
b(x)&=\mathbb{1}_{[0,1]}(x)-\sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}}2^{-(n+j)}\mathbb{1}_{I_{n,j}}(x)\\
&=\mathbb{1}_A(x)+\sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}}\left(1-2^{-(n+j)}\right)\mathbb{1}_{I_{n,j}}(x).
\end{aligned}
$$

All of the endpoints of the intervals $\bar{I}_{n,j}$ are the discontinuous points of b, which is dense in A. For any interval $I\subset[0,1]$ such that $I\cap A\neq\emptyset$, it always contains the discontinuous points of b. However, any interval $I\subset[0,1]$ such *that $I\cap A=\emptyset$, it is a subset of some $I_{n,j}$. Hence, b is not a piecewise continuous function. In the following, we shall show that b satisfies condition (H2) with $p_0=2$ and $\alpha=\frac{1}{4}$ and $\phi(s)=Cs^{-\frac{1}{4}}$.*

Proof. For $u>0$ and any interval (a_1,a_2) (it is similar for $[a_1,a_2]$),

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\left|\mathbb{1}_{(a_1,a_2)}(x+u)-\mathbb{1}_{(a_1,a_2)}(x)\right|^2\mathrm{d}x\\
&\quad=\int_{a_1-u}^{a_2-u}\mathbb{1}_{(a_1,a_2)^c}(x)\mathrm{d}x+\int_{a_1}^{a_2}\mathbb{1}_{(a_1-u,a_2-u)^c}(x)\mathrm{d}x\\
&\quad=\int_{a_1-u}^{(a_2-u)\wedge a_1}\mathrm{d}x+\int_{(a_2-u)\vee a_1}^{a_2}\mathrm{d}x\\
&\quad\le 2\left(|u|\wedge(a_2-a_1)\right).
\end{aligned}$$

For $u<0$, we obtain that

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\left|\mathbb{1}_{(a_1,a_2)}(x+u)-\mathbb{1}_{(a_1,a_2)}(x)\right|^2\mathrm{d}x\\
&\quad=\int_{-\infty}^{+\infty}\left|\mathbb{1}_{(a_1,a_2)}(v)-\mathbb{1}_{(a_1,a_2)}(v-u)\right|^2\mathrm{d}v\le 2\left(|u|\wedge(a_2-a_1)\right).
\end{aligned}$$

Hence, it follows from Jensen's inequality that

$$\begin{aligned}
&\int_{-\infty}^{+\infty}|b(x+u)-b(x)|^2\mathrm{d}x\\
&\quad\le\int_{-\infty}^{+\infty}\Bigg(\left|\mathbb{1}_{[0,1]}(x+u)-\mathbb{1}_{[0,1]}(x)\right|\\
&\qquad+\sum_{n=1}^{+\infty}\sum_{j=1}^{2^{n-1}}2^{-(n+j)}\left|\mathbb{1}_{I_{n,j}}(x+z)-\mathbb{1}_{I_{n,j}}(x)\right|\Bigg)^2\mathrm{d}x\\
&\quad\le\left(1+\sum_{n=1}^{+\infty}\sum_{j=1}^{2^{n-1}}2^{-(n+j)}\right)\Bigg\{\int_{-\infty}^{+\infty}\left|\mathbb{1}_{[0,1]}(x+u)-\mathbb{1}_{[0,1]}(x)\right|^2\mathrm{d}x\\
&\qquad+\sum_{n=1}^{+\infty}\sum_{j=1}^{2^{n-1}}2^{-(n+j)}\int_{-\infty}^{+\infty}\left|\mathbb{1}_{I_{n,j}}(x+u)-\mathbb{1}_{I_{n,j}}(x)\right|^2\mathrm{d}x\Bigg\}\\
&\quad\le 2\left(1+\sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}}2^{-(n+j)}\right)^2|u|=4|u|.
\end{aligned}$$

Combining this with (2.4), we obtain that

$$\sup_{z\in\mathbb{R}}\int_{\mathbb{R}\times\mathbb{R}}|b(y)-b(x)|^2\frac{\mathrm{e}^{-\frac{|x-z|^2}{s}}\,\mathrm{e}^{-\frac{|y-x|^2}{r}}}{s^{\frac{1}{2}}r^{\frac{1}{2}}}\mathrm{d}x\mathrm{d}y$$

$$
\begin{aligned}
&\leq \frac{1}{s^{\frac{1}{2}}r^{\frac{1}{2}}}\int_{\mathbb{R}} \mathrm{e}^{-\frac{|u|^2}{r}}\int_{\mathbb{R}} |b(x+u)-b(x)|^2 \mathrm{d}x\mathrm{d}u\\
&\leq \frac{4}{s^{\frac{1}{2}}r^{\frac{1}{2}}}\int_{\mathbb{R}} \mathrm{e}^{-\frac{|u|^2}{r}}|u|\mathrm{d}u = \left(Cs^{-\frac{1}{4}}r^{\frac{1}{4}}\right)^2.
\end{aligned}
$$

□

A general class of functions that satisfies (H2) is the fractional Sobolev space $W^{\beta,p}(\mathbb{R}^d)$, showing as follows.

Example 2.4. *If there exist $\beta > 0$ and $p \in [2,\infty)\cap(d,+\infty)$ such that the Gagliardo seminorm of b is finite, i.e.*

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

then (H2) holds for $p_0 = p$ with $\alpha = \frac{\beta}{2}$ and $\phi(s) = C_1 s^{-\frac{d}{2}}[b]^p_{W^{\beta,p}}$. Hence, if b satisfies (H1) and $[b]_{W^{\beta,p}} < \infty$ with $p \in [2,\infty)\cap(d,+\infty)$, then (2.1) *holds.*

Proof. Indeed, by Hölder's inequality and (2.4), it follows that

$$
\begin{aligned}
&\frac{1}{(rs)^{\frac{d}{2}}}\int_{\mathbb{R}^d\times\mathbb{R}^d} |b(y)-b(x)|^p\,\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}\,\mathrm{d}x\mathrm{d}y\\
&= \frac{1}{(rs)^{\frac{d}{2}}}\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|b(x)-b(y)|^p}{|x-y|^{d+\beta p}}\,\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{r}}\,|x-y|^{d+\beta p}\mathrm{d}x\mathrm{d}y\\
&\leq C_1 s^{-\frac{d}{2}} r^{\frac{\beta p}{2}}\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|b(x)-b(y)|^p}{|x-y|^{d+\beta p}}\,\mathrm{e}^{-\frac{|x-z|^2}{s}-\frac{|y-x|^2}{2r}}\,\mathrm{d}x\mathrm{d}y\\
&\leq C_1 s^{-\frac{d}{2}} r^{\frac{\beta p}{2}}\,[b]^p_{W^{\beta,p}}.
\end{aligned}
$$

□

3 Proof of Theorem 2.1

The key point for proving the main result is to construct a reference SDE. By using Girsanov's theorem, the reference SDE provides new representations of (1.1) and its EM's approximation SDE (1.2) under another probability measures.

We denote by $Y_t = x + \sigma W_t$ the reference SDE of (1.1). Then Y_t is a time homogenous Markov process with heat kernel w.r.t. the Lebesgue measure as follows:

$$
p_t(x,y) = \frac{\exp\left\{-\frac{\langle(\sigma\sigma^*)^{-1}(y-x),(y-x)\rangle}{2t}\right\}}{\sqrt{(2t\pi)^d\det(\sigma\sigma^*)}},\quad x,y\in\mathbb{R}^d. \tag{3.1}
$$

To prove Theorem 2.1, we give three auxiliary lemmas.

The first lemma is on the exponential estimate of $|b(Y_t)|$. Here, we use a weaker condition (H1') than (H1).

(H1') There exist $\beta \in [0,1)$, nonnegative constants L_1, L_2 and $F \geq 0$ with $F \in L^{p_1}(\mathbb{R}^d)$ for some $p_1 > d$ such that

$$|b(x)| \leq L_1 + L_2|x|^\beta + F(x). \tag{3.2}$$

Lemma 3.1. *Assume (H1') holds. Then for all $T, \lambda > 0$, it holds that*

$$\mathbb{E}\exp\left\{\lambda\int_0^T |\sigma^{-1}b(Y_s)|^2 \mathrm{d}s\right\} < \infty. \tag{3.3}$$

Proof. Note that for any $\varepsilon > 0$

$$L_1 + L_2|x|^\beta \leq L_1 + (1-\beta)L_2^{\frac{1}{1-\beta}}\big(\frac{\beta}{\varepsilon}\big)^{\frac{\beta}{1-\beta}} + \varepsilon|x| =: L(\varepsilon) + \varepsilon|x|, \tag{3.4}$$

and for any $a, b, c, \varepsilon_1, \varepsilon_2 > 0$

$$(a+b+c)^2 \leq (2 + \frac{1}{\varepsilon_1})a^2 + (1+\varepsilon_1+\varepsilon_2)b^2 + (2+\frac{1}{\varepsilon_2})c^2.$$

Combining these with (3.2) and the Hölder inequality, we have that

$$\begin{aligned}
&\mathbb{E}\exp\left\{\lambda\int_0^T |\sigma^{-1}b(Y_s)|^2 \mathrm{d}s\right\}\\
&\leq \mathbb{E}\exp\left\{\lambda\int_0^T \|\sigma^{-1}\|^2\left(L(\varepsilon) + \varepsilon|Y_s| + F(Y_s)\right)^2 \mathrm{d}s\right\}\\
&\leq \mathbb{E}\exp\left\{\lambda\int_0^T \|\sigma^{-1}\|^2\left((L(\varepsilon) + \varepsilon|x|) + \varepsilon|Y_s - x| + F(Y_s)\right)^2 \mathrm{d}s\right\}\\
&\leq \mathbb{E}\exp\left\{\lambda\int_0^T \|\sigma^{-1}\|^2\Big(\left(2+\varepsilon_1^{-1}\right)\left(L(\varepsilon)+\varepsilon|x|\right)^2\right.\\
&\qquad\qquad \left. + (1+\varepsilon_1+\varepsilon_2)\varepsilon^2|Y_s - x|^2 + (2+\varepsilon_2^{-1})F^2(Y_s)\Big)\mathrm{d}s\right\}\\
&\leq \exp\{\lambda T\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon|x|)^2\left(2+\varepsilon_1^{-1}\right)\}\\
&\qquad \times\left(\mathbb{E}\exp\left\{\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2\int_0^T|Y_s - x|^2\mathrm{d}s\right\}\right)^{\frac{1}{1+\varepsilon_1+\varepsilon_2}}\\
&\qquad \times\left(\mathbb{E}\exp\left\{\frac{\lambda(2+\varepsilon_2^{-1})(1+\varepsilon_1+\varepsilon_2)}{\varepsilon_1+\varepsilon_2}\|\sigma^{-1}\|^2\int_0^T F^2(Y_s)\mathrm{d}s\right\}\right)^{\frac{\varepsilon_1+\varepsilon_2}{1+\varepsilon_1+\varepsilon_2}}.
\end{aligned} \tag{3.5}$$

Let

$$\begin{aligned}
I_{1,T} &= \mathbb{E}\exp\left\{\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2\int_0^T |Y_s - x|^2\mathrm{d}s\right\},\\
I_{2,T} &= \mathbb{E}\exp\left\{\frac{\lambda(2+\varepsilon_2^{-1})(1+\varepsilon_1+\varepsilon_2)}{\varepsilon_1+\varepsilon_2}\|\sigma^{-1}\|^2\int_0^T F^2(Y_s)\mathrm{d}s\right\}.
\end{aligned}$$

Since $F \in L^{p_1}(\mathbb{R}^d)$, for any $0 \le S \le T$ and q satisfying $\frac{d}{p_1} + \frac{1}{q} < 1$, we obtain that (see e.g. [12])

$$\mathbb{E}\left[\int_S^T F^2(Y_s)\mathrm{d}s \Big| \mathscr{F}_S\right] \le (T-S)^{\frac{1}{q}} \|F\|_{L^{p_1}}. \tag{3.6}$$

This yields the following Khasminskii's estimate (see e.g. [27, Lemma 3.5]): for any $C > 0$

$$\mathbb{E} \exp\left\{C \int_0^T F^2(Y_s)\mathrm{d}s\right\} < \infty. \tag{3.7}$$

Thus, for any $\lambda, \varepsilon_1, \varepsilon_2 > 0$, one has

$$I_{2,T} < \infty. \tag{3.8}$$

For $I_{1,T}$. Since ε, ε_1 and ε_2 are arbitrary, for any $T > 0$, we can choose them sufficiently small such that

$$1 - 2T^2(1+\varepsilon_1+\varepsilon_2)^2\lambda\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2 =: \hat{\lambda} > 0.$$

This, together with the Jensen inequality and the heat kernel (3.1), yields that

$$\begin{aligned}
I_{1,T} &= \mathbb{E} \exp\left\{\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2 \int_0^T |Y_s - x|^2 \mathrm{d}s\right\} \\
&\le \frac{1}{T}\int_0^T \mathbb{E}\exp\left\{T\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2|Y_s-x|^2\right\}\mathrm{d}s \\
&= \int_0^T \int_{\mathbb{R}^d} \frac{\exp\left\{T\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2|y|^2 - \frac{|\sigma^{-1}y|^2}{2s}\right\}}{T\sqrt{(2s\pi)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s \\
&\le \int_0^T \int_{\mathbb{R}^d} \frac{\exp\left\{-\left(\frac{1-2sT\lambda(1+\varepsilon_1+\varepsilon_2)^2\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2}{2s}\right)|\sigma^{-1}y|^2\right\}}{T\sqrt{(2s\pi)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s \\
&\le \int_0^T \int_{\mathbb{R}^d} \frac{\exp\left\{-\left(\frac{\hat{\lambda}}{2s}\right)|\sigma^{-1}y|^2\right\}}{T\sqrt{(2s\pi)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s \\
&= \hat{\lambda}^{-\frac{d}{2}} < \infty.
\end{aligned} \tag{3.9}$$

Plugging (3.9) and (3.8) into (3.5), then (3.3) follows. □

The following lemma deals with the exponential estimate of $|b(Y_{t_\delta})|$, where $\{Y_{t_\delta}\}_{t\in[0,T]}$ denotes the solution to the discrete-time EM's scheme. The Krylov estimate (3.6) fails for Y_{s_δ}, (see [3, Remark 2.5] or [23]). Hence, we use (H1) in Lemma 3.2 instead of (H1').

Lemma 3.2. *Assume (H1). Then for all $T > 0, \lambda > 0$, we have*

$$\sup_{0<\delta<1\wedge T} \mathbb{E}\exp\left\{\lambda \int_0^T |\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\} < \infty. \tag{3.10}$$

Proof. Splitting the interval $[0,T]$ and applying (3.4), it follows from the elementary inequality that

$$
\begin{aligned}
&\mathbb{E}\exp\left\{\lambda\int_0^T|\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\}\\
&=\mathbb{E}\Big\{\exp\left\{\lambda\int_0^\delta|\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\}\exp\left\{\lambda\int_\delta^T|\sigma^{-1}b(Y_{s_\delta})|^2\mathrm{d}s\right\}\Big\}\\
&\le\exp\left\{\lambda\delta\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon x)^2\right\}\mathbb{E}\exp\left\{\lambda\int_\delta^T\|\sigma^{-1}\|^2|L(\varepsilon)+\varepsilon x+\varepsilon(Y_{s_\delta}-x)|^2\mathrm{d}s\right\}\\
&\le\exp\{\lambda\delta\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon x)^2\}\exp\{\lambda(T-\delta)\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon|x|)^2\left(1+\varepsilon_1^{-1}\right)\}\\
&\qquad\times\mathbb{E}\exp\left\{\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\int_\delta^T|Y_{s_\delta}-x|^2\mathrm{d}s\right\}\\
&\le\exp\{\lambda T\|\sigma^{-1}\|^2(L(\varepsilon)+\varepsilon x)^2\left(1+\varepsilon_1^{-1}\right)\}\\
&\qquad\times\mathbb{E}\exp\left\{\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\int_\delta^T|Y_{s_\delta}-x|^2\mathrm{d}s\right\}.
\end{aligned}\tag{3.11}
$$

For any $T,\lambda>0$, we choose ε and ε_1 sufficiently small such that

$$1-2T^2\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2=:\breve{\lambda}>0.$$

This, together with the Jensen inequality and (3.1), yields that

$$
\begin{aligned}
&\mathbb{E}\exp\left\{\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\int_\delta^T|Y_{s_\delta}-x|^2\mathrm{d}s\right\}\\
&\le\frac{1}{T-\delta}\int_\delta^T\mathbb{E}\exp\{(T-\delta)\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2|Y_{s_\delta}-x|^2\}\mathrm{d}s\\
&\le\int_\delta^T\int_{\mathbb{R}^d}\frac{\exp\{(T-\delta)\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2|y|^2-\frac{\langle(\sigma\sigma^*)^{-1}y,y\rangle}{2s_\delta}\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&\le\int_\delta^T\int_{\mathbb{R}^d}\frac{\exp\{(T-\delta)\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2|\sigma^{-1}y|^2-\frac{|\sigma^{-1}y|^2}{2s_\delta}\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&\le\int_\delta^T\int_{\mathbb{R}^d}\frac{\exp\{-\frac{(1-2(T-\delta)^2\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2)}{2s_\delta}|\sigma^{-1}y|^2\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&\le\int_\delta^T\int_{\mathbb{R}^d}\frac{\exp\{-\frac{(1-2T^2\lambda(1+\varepsilon_1)\varepsilon^2\|\sigma^{-1}\|^2\|\sigma\|^2)}{2s_\delta}|\sigma^{-1}y|^2\}}{(T-\delta)\sqrt{(2\pi s_\delta)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}s\\
&=\breve{\lambda}^{-\frac{d}{2}}<\infty.
\end{aligned}\tag{3.12}
$$

Combining this with (3.11), we have that (3.10) holds. □

Remark 3.1. According to the proofs of Lemma 3.1 and Lemma 3.2 (see especially (3.9), (3.12), and the definitions of $\hat{\lambda}$ and $\breve{\lambda}$), we have that $\epsilon=O(T^{-1})$ as $T\to+\infty$. Then the constant $TL^2(\epsilon)$ in (3.5) and (3.11) is of the order

$(1-\beta)^2(L_2^{\frac{2}{1+\beta}}T)^{\frac{1+\beta}{1-\beta}}$. Hence, for larger $L_2^{\frac{2}{1+\beta}}T$, the closer β to 1, the greater upper bound of (3.3) and (3.10).

Lemma 3.1 and Lemma 3.2 serve for using the Novikov condition in the proof of Theorem 2.1. For the case that $\beta < 1$, the constant λ in both lemmas can be arbitrary. For the case that $\beta = 1$, with $\epsilon = L_2$ and $L(\epsilon) = L_1$ in (3.4), according to (3.9) and (3.12), one can see from the definitions of $\hat{\lambda}$ and $\check{\lambda}$ that (3.3) and (3.10) hold for $\lambda > 0$ and $T > 0$ satisfying the following condition

$$2T^2\lambda L_2^2\|\sigma^{-1}\|^2\|\sigma\|^2 < 1, \tag{3.13}$$

and sufficiently small ε_1 and ε_2.

Lemma 3.3. *Assume (H2). Then there exists a constant $C_\sigma > 0$ depending on σ only such that for all $0 < s \le t \le T$ we have*

$$\mathbb{E}|b(Y_t) - b(Y_s)|^{p_0} \le C_\sigma(\phi(2s\|\sigma\|^2)(2(t-s)\|\sigma\|^2)^\alpha)^{p_0}. \tag{3.14}$$

Proof. By the definition of reference SDE, it is easy to see that

$$\mathbb{E}|b(Y_t) - b(Y_s)|^{p_0} = \mathbb{E}|b(x+\sigma W_t) - b(x+\sigma W_s)|^{p_0}.$$

Noting that $W_t - W_s$ and W_s are mutually independent, we obtain from (3.1) and (H2) that

$$\begin{aligned}
&\mathbb{E}|b(x+\sigma W_t) - b(x+\sigma W_s)|^{p_0} \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(x+y) - b(x+z)|^{p_0} p_{t-s}(x+z, x+y)p_s(x, x+z)\mathrm{d}y\mathrm{d}z \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(x+y) - b(x+z)|^{p_0} \frac{\mathrm{e}^{-\frac{\langle(\sigma\sigma^*)^{-1}(y-z),(y-z)\rangle}{2(t-s)}}}{\sqrt{(2\pi(t-s))^d\det(\sigma\sigma^*)}}\frac{\mathrm{e}^{-\frac{\langle(\sigma\sigma^*)^{-1}z,z\rangle}{2s}}}{\sqrt{(2\pi s)^d\det(\sigma\sigma^*)}}\mathrm{d}y\mathrm{d}z \\
&\le \frac{\|\sigma\|^{2d}}{\pi^d\det(\sigma\sigma^*)}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(x+y) - b(x+z)|^{p_0}\frac{\mathrm{e}^{-\frac{|y-z|^2}{2\|\sigma\|^2(t-s)}}\mathrm{e}^{-\frac{|z|^2}{2\|\sigma\|^2 s}}}{(2(t-s)\|\sigma\|^2)^{d/2}(2s\|\sigma\|^2)^{d/2}}\mathrm{d}y\mathrm{d}z \\
&= \frac{\|\sigma\|^{2d}}{\pi^d\det(\sigma\sigma^*)}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(u) - b(v)|^{p_0}\frac{\mathrm{e}^{-\frac{|u-v|^2}{2\|\sigma\|^2(t-s)}}\mathrm{e}^{-\frac{|v-x|^2}{2\|\sigma\|^2 s}}}{(2(t-s)\|\sigma\|^2)^{d/2}(2s\|\sigma\|^2)^{d/2}}\mathrm{d}u\mathrm{d}v \\
&\le \sup_{x\in\mathbb{R}^d}\frac{\|\sigma\|^{2d}}{\pi^d\det(\sigma\sigma^*)}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |b(u) - b(v)|^{p_0}\frac{\mathrm{e}^{-\frac{|u-v|^2}{2\|\sigma\|^2(t-s)}}\mathrm{e}^{-\frac{|v-x|^2}{2\|\sigma\|^2 s}}}{(2(t-s)\|\sigma\|^2)^{d/2}(2s\|\sigma\|^2)^{d/2}}\mathrm{d}u\mathrm{d}v \\
&\le \frac{\|\sigma\|^{2d}}{\pi^d\det(\sigma\sigma^*)}(\phi(2s\|\sigma\|^2)(2(t-s)\|\sigma\|^2)^\alpha)^{p_0},
\end{aligned}$$

which implies that (3.14) holds by taking $C_\sigma = \frac{\|\sigma\|^{2d}}{\pi^d\det(\sigma\sigma^*)}$. □

Now, we are in position to finish the Proof of Theorem 2.1.

Proof of Theorem 2.1. Let

$$\hat{W}_t = W_t - \int_0^t \sigma^{-1} b(Y_s)\mathrm{d}s, \quad \tilde{W}_t = W_t - \int_0^t \sigma^{-1} b(Y_{s_\delta})\mathrm{d}s,$$

$$R_{1,T} = \exp\Big\{ \int_0^T \langle \sigma^{-1} b(Y_s), \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^T |\sigma^{-1} b(Y_s)|^2 \mathrm{d}s \Big\},$$

$$R_{2,T} = \exp\Big\{ \int_0^T \langle \sigma^{-1} b(Y_{s_\delta}), \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^T \left|\sigma^{-1} b(Y_{s_\delta})\right|^2 \mathrm{d}s \Big\}.$$

The proof is divided into two steps:

Step (i), we shall prove that the assertion holds under (H1) and (H2).

We first show that $\{\hat{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_1 := R_{1,T}\mathbb{P}$, and $\{\tilde{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_2 := R_{2,T}\mathbb{P}$. In view of Lemma 3.1, the Girsanov theorem implies that $\{R_{1,t}\}_{t\in[0,T]}$ is a martingale and $\{\hat{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_1$. Similarly, it follows from Lemma 3.2 and Novikov's condition that $\{\tilde{W}_t\}_{t\in[0,T]}$ is a Brownian motion under $\mathbb{Q}_2$.

Then, we can reformulate $Y_t = x + \sigma W_t$ as follows:

$$Y_t = x + \int_0^t b(Y_s)\mathrm{d}s + \sigma \hat{W}_t,$$

which means that $(Y_t, \hat{W}_t)$ under $\mathbb{Q}_1$ is a weak solution of (1.1). Hence, Y_t under $\mathbb{Q}_1$ has the same law as X_t under $\mathbb{P}$ due to the pathwise uniqueness of the solutions to (1.1) (see Remark 2.1). Similarly, reformulating $Y_t = x + \sigma W_t$ as follows:

$$Y_t = x + \int_0^t b(Y_{s_\delta})\mathrm{d}s + \sigma \tilde{W}_t. \tag{3.15}$$

Then $(Y_t, \tilde{W}_t)$ under $\mathbb{Q}_2$ is also a weak solution of (1.2). Hence Y_t under $\mathbb{Q}_2$ has the same law as $X_t^{(\delta)}$ under $\mathbb{P}$ due to the pathwise uniqueness of solutions to the equation (1.2).

From these equivalence relations, we obtain that for any bounded and measurable function f on $\mathbb{R}^d$

$$\begin{aligned} &|\mathbb{E} f(X_t) - \mathbb{E} f(X_t^{(\delta)})| = |\mathbb{E}_{\mathbb{Q}_1} f(Y_t) - \mathbb{E}_{\mathbb{Q}_2} f(Y_t)| \\ &= \mathbb{E}|(R_{1,T} - R_{2,T}) f(Y_t)| \le \|f\|_\infty \mathbb{E}|R_{1,T} - R_{2,T}|. \end{aligned}$$

Using the inequality $|\,\mathrm{e}^x - \mathrm{e}^y\,| \le (\mathrm{e}^x \vee \mathrm{e}^y)|x-y|$, Hölder's inequality and Minkowski's inequality, we derive from definitions of $R_{1,T}$ and $R_{2,T}$ that

$$\begin{aligned} &\mathbb{E}|R_{1,T} - R_{2,T}| \\ &\le \mathbb{E}\Big\{ (R_{1,T} \vee R_{2,T}) \Big| \int_0^T \langle \sigma^{-1}(b(Y_s) - b(Y_{s_\delta})), \mathrm{d}W_s\rangle \\ &\qquad + \frac{1}{2}\int_0^T \Big(|\sigma^{-1} b(Y_{s_\delta})|^2 - |\sigma^{-1} b(Y_s)|^2 \Big)\mathrm{d}s \Big| \Big\} \end{aligned}$$

$$
\begin{aligned}
&\le \mathbb{E}\Big[(R_{1,T}\vee R_{2,T})\Big|\int_0^T \langle \sigma^{-1}(b(Y_s)-b(Y_{s_\delta})), \mathrm{d}W_s\rangle\Big|\Big]\\
&\quad+\frac{1}{2}\mathbb{E}\Big[(R_{1,T}\vee R_{2,T})\Big|\int_0^T \Big(|\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\Big)\mathrm{d}s\Big|\Big]\\
&\le \Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0}{p_0-1}}\Big)^{\frac{p_0-1}{p_0}}\Big(\mathbb{E}\Big|\int_0^T \langle \sigma^{-1}(b(Y_s)-b(Y_{s_\delta})), \mathrm{d}W_s\rangle\Big|^{p_0}\Big)^{\frac{1}{p_0}}\\
&\quad+\frac{1}{2}\Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0+1}{p_0-1}}\Big)^{\frac{p_0-1}{p_0+1}}\Big(\mathbb{E}\Big|\int_0^T (|\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2)\mathrm{d}s\Big|^{\frac{p_0+1}{2}}\Big)^{\frac{2}{p_0+1}}\\
&\le \Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0}{p_0-1}}\Big)^{\frac{p_0-1}{p_0}}\Big(\mathbb{E}\Big|\int_0^T \langle \sigma^{-1}(b(Y_s)-b(Y_{s_\delta})), \mathrm{d}W_s\rangle\Big|^{p_0}\Big)^{\frac{1}{p_0}}\\
&\quad+\frac{1}{2}\Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0+1}{p_0-1}}\Big)^{\frac{p_0-1}{p_0+1}}\int_0^T \Big(\mathbb{E}\Big||\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\Big|^{\frac{p_0+1}{2}}\Big)^{\frac{2}{p_0+1}}\mathrm{d}s\\
&=: \Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0}{p_0-1}}\Big)^{\frac{p_0-1}{p_0}}G_{1,T}+\frac{1}{2}\Big(\mathbb{E}(R_{1,T}\vee R_{2,T})^{\frac{p_0+1}{p_0-1}}\Big)^{\frac{p_0-1}{p_0+1}}G_{2,T}. \qquad (3.16)
\end{aligned}
$$

Let

$$M_t=\int_0^t \langle \sigma^{-1}b(Y_s), \mathrm{d}W_s\rangle \quad\text{and}\quad \hat{M}_t(q)=\mathrm{e}^{2qM_t-2q^2\langle M_\cdot\rangle_t},\ q>0.$$

By Lemma 3.1, for any $q>1$, $\hat{M}_t(q)$ is an exponential martingale. Then, the Hölder inequality implies that

$$
\begin{aligned}
\mathbb{E}R_{1,T}^{\frac{p_0}{p_0-1}} &= \mathbb{E}\exp\Big\{\frac{p_0}{p_0-1}\int_0^T\langle\sigma^{-1}b(Y_s),\mathrm{d}W_s\rangle\\
&\qquad\qquad -\frac{p_0}{2(p_0-1)}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\}\\
&\le \Big(\mathbb{E}\hat{M}_T(\frac{p_0}{p_0-1})\Big)^{1/2}\Big(\mathbb{E}\exp\Big\{\frac{p_0(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\}\Big)^{1/2}\\
&= \Big(\mathbb{E}\exp\Big\{\frac{p_0(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\}\Big)^{1/2}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}R_{1,T}^{\frac{p_0+1}{p_0-1}} &\le \Big(\mathbb{E}\hat{M}_T(\frac{p_0+1}{p_0-1})\Big)^{1/2}\Big(\mathbb{E}\exp\Big\{\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\}\Big)^{1/2}\\
&= \Big(\mathbb{E}\exp\Big\{\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}\int_0^T|\sigma^{-1}b(Y_s)|^2\mathrm{d}s\Big\}\Big)^{1/2}.
\end{aligned}
$$

Then, it follows from Lemma 3.1 again that

$$\mathbb{E}\left(R_{1,T}^{\frac{p_0+1}{p_0-1}}+R_{1,T}^{\frac{p_0}{p_0-1}}\right)<\infty. \qquad (3.17)$$

Similarly, we can prove by using Hölder's inequality and Lemma 3.2 that

$$\mathbb{E}\left(R_{2,T}^{\frac{p_0+1}{p_0-1}}+R_{2,T}^{\frac{p_0}{p_0-1}}\right)<\infty. \tag{3.18}$$

Since $\phi\in C((0,+\infty),(0,+\infty))$, there is $C>0$ depending on l,T,σ such that

$$\phi^2(r)\le C\phi^2(s),\ l\le s\le r\le 2\|\sigma\|^2T.$$

Combining this with that ϕ is non increasing on $(0,l)$ and $\int_0^l\phi^2(s)\mathrm{d}s<\infty$, which yields $\int_0^{2\|\sigma\|^2T}\phi^2(s)\mathrm{d}s<\infty$ since $\phi\in C((0,+\infty),(0,+\infty))$, we obtain that

$$\begin{aligned}\sum_{k=1}^{[T/\delta]}\phi^2(2k\delta\|\sigma\|^2)\delta&\le\sum_{k=1}^{[T/\delta]}\int_{((k-1)\delta)\wedge\frac{l}{2\|\sigma\|^2}}^{(k\delta)\wedge\frac{l}{2\|\sigma\|^2}}\phi^2(2\|\sigma\|^2r)\mathrm{d}r+C\int_{\frac{l}{2\|\sigma\|^2}}^{T}\phi^2(2\|\sigma\|^2r)\mathrm{d}r\\&=\int_0^{\frac{l}{2\|\sigma\|^2}}\phi^2(2\|\sigma\|^2r)\mathrm{d}r+C\int_{\frac{l}{2\|\sigma\|^2}}^{T}\phi^2(2\|\sigma\|^2r)\mathrm{d}r\\&=\frac{1\vee C}{2\|\sigma\|^2}\int_0^{2\|\sigma\|^2T}\phi^2(s)\mathrm{d}s<\infty.\end{aligned} \tag{3.19}$$

This, together with the B-D-G inequality and Lemma 3.3, yields that for $p_0\ge 2$

$$\begin{aligned}&G_{1,T}\\&=\left(\mathbb{E}\Big|\int_0^T\langle\sigma^{-1}(b(Y_s)-b(Y_{s_\delta})),\mathrm{d}W_s\rangle\Big|^{p_0}\right)^{1/p_0}\\&\le\left(\frac{p_0}{p_0-1}\right)^{\frac{p_0}{2}}\left(\frac{p_0(p_0-1)}{2}\right)^{\frac12}\|\sigma^{-1}\|\left(\int_0^T\left(\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{p_0}\right)^{\frac{2}{p_0}}\mathrm{d}s\right)^{\frac12}\\&\le\delta^\alpha\frac{2^\alpha\|\sigma\|^{\frac{2d}{p_0}+2\alpha}\|\sigma^{-1}\|}{(\pi^d\det(\sigma\sigma^*))^{\frac{1}{p_0}}}\left(\frac{p_0}{p_0-1}\right)^{\frac{p_0}{2}}\left(\frac{p_0(p_0-1)}{2}\right)^{\frac12}\Big(\int_0^T\phi^2(2s_\delta\|\sigma\|^2)\mathrm{d}s\Big)^{\frac12}\\&\le\delta^\alpha\frac{\sqrt{1\vee C}2^{\alpha-\frac12}\|\sigma\|^{\frac{2d}{p_0}+2\alpha-1}\|\sigma^{-1}\|}{(\pi^d\det(\sigma\sigma^*))^{\frac{1}{p_0}}}\left(\frac{p_0}{p_0-1}\right)^{\frac{p_0}{2}}\left(\frac{p_0(p_0-1)}{2}\right)^{\frac12}\Big(\int_0^{2\|\sigma\|^2T}\phi^2(s)\mathrm{d}s\Big)^{\frac12}\\&=C_{T,p_0,\sigma,\alpha,\phi}\delta^\alpha.\end{aligned} \tag{3.20}$$

Noting that for any $p\ge1$, one has

$$\mathbb{E}|Y_t|^p\le 2^{p-1}\left(|x|^p+(\sqrt{t}\|\sigma\|)^p\mathbb{E}|W_1|^p\right), \tag{3.21}$$

we derive from (3.4) and (3.19) that

$$\left(\mathbb{E}|b(Y_s)+b(Y_{s_\delta})|^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}$$

$$
\begin{aligned}
&\le \left(\mathbb{E}\left(2L(\varepsilon)+\varepsilon(|Y_s|+|Y_{s_\delta}|)\right)^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}\\
&\le 2\left\{L(\varepsilon)+2^{\frac{p_0^2+1}{p_0(p_0+1)}}\varepsilon\left(|x|+\sqrt{T}\|\sigma\|\left(\mathbb{E}|W_1|^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}\right)\right\}\\
&=: C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,x}.
\end{aligned}
$$

Combining this with Lemma 3.3, (3.19) and Hölder's inequality, we obtain

$$
\begin{aligned}
G_{2,T} &= \frac{1}{2}\int_0^T \left(\mathbb{E}\left||\sigma^{-1}b(Y_{s_\delta})|^2-|\sigma^{-1}b(Y_s)|^2\right|^{\frac{p_0+1}{2}}\right)^{\frac{2}{p_0+1}}\mathrm{d}s\\
&\le \frac{\|\sigma^{-1}\|^2}{2}\int_0^T\left(\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{\frac{p_0+1}{2}}|b(Y_s)+b(Y_{s_\delta})|^{\frac{p_0+1}{2}}\right)^{\frac{2}{p_0+1}}\mathrm{d}s\\
&\le \frac{\|\sigma^{-1}\|^2}{2}\int_0^T\left(\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{p_0}\right)^{\frac{1}{p_0}}\left(\mathbb{E}|b(Y_s)+b(Y_{s_\delta})|^{\frac{p_0(p_0+1)}{p_0-1}}\right)^{\frac{p_0-1}{p_0(p_0+1)}}\mathrm{d}s\\
&\le \frac{\|\sigma^{-1}\|^2}{2}C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,x}\int_0^T\left(\mathbb{E}|b(Y_s)-b(Y_{s_\delta})|^{p_0}\right)^{\frac{1}{p_0}}\mathrm{d}s\\
&\le C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,\phi,x}\delta^\alpha, \qquad (3.22)
\end{aligned}
$$

where

$$
C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,\phi,x}=\frac{1\vee C2^{\alpha-2}\|\sigma\|^{\frac{2d}{p_0}+2\alpha-2}\|\sigma^{-1}\|^2C_{T,p_0,\sigma,L(\varepsilon),\varepsilon,x}}{(\pi^d\det(\sigma\sigma^*))^{\frac{1}{p_0}}}\int_0^{2\|\sigma\|^2T}\phi(s)\mathrm{d}s.
$$

The desired assertion (2.1) is proved by substituting (3.17), (3.18), (3.20) and (3.22) into (3.16). Therefore, the conclusion holds under (H1) and (H2).

Step (ii), we prove that if b satisfies the linear growth condition, then the conclusion (2.1) holds for T satisfying (2.2).

By Remark 3.1, we have that the conclusions of Lemma 3.1 and Lemma 3.2 hold for any λ, T satisfying (3.13). By (2.2) and $\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}>\frac{p_0(p_0+1)}{(p_0-1)^2}>\frac{1}{2}$, we can choose $\lambda=\frac{(p_0+3)(p_0+1)}{(p_0-1)^2}$ in Lemma 3.1 and Lemma 3.2. Then, by checking step (i), we arrive at (3.16). Moreover, (3.17) and (3.18) hold by the same argument together with the stopping time technique. Then, we can conclude the second conclusion from (3.20) and (3.22). The proof is therefore complete. □

Remark 3.2. According to the proof of this theorem, the key point for that f in (2.1) can only be bounded and measurable is that the distributions of $X_t^{(\delta)}$ and X_t come from the same process $Y_t = x+\sigma W_t$. This fails for the multiplicative noise case.

Acknowledgements

The authors would like to thank the associate editor and referee for their helpful comments and suggestions. The third author was supported by the National Natural Science Foundation of China (Grant No. 11901604, 11771326).

References

[1] V. Bally, D. Talay, The law of the Euler scheme for stochastic differential equations, I. Convergence rate of the distribution function, *Probab. Theory Related Fields.* **104** (1) (1996), 43–60.

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

[3] J. Bao, X. Huang, S.-Q. Zhang, Convergence rate of EM algorithm for SDEs under integrability condition, Arxiv:2009.04781v1

[4] J. Bao, J. Shao, Weak convergence of path-dependent SDEs with irregular coefficients, Arxiv:1809.03088

[5] K. Dareiotis, M. Gerencsér, On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift, *Electron. J. Probab.* **25** (2020), 1–18.

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

[7] A. Kohatsu-Higa, A. Lejay, K. Yasuda, Weak rate of convergence of the Euler-Maruyama scheme for stochastic differential equations with non-regular drift, *J. Comp. Appl. Math.* **326** (2017), 138–158.

[8] T. Müller-Gronbach, L. Yaroslavtseva, Sharp lower error bounds for strong approximation of SDEs with discontinuous drift coefficient by coupling of noise, arXiv:2010.00915

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

[10] T. Müller-Gronbach, L. Yaroslavtseva, A strong order 3/4 method for SDEs with discontinuous drift coefficient, *IMA Journal of Numerical Analysis.* **00** (2020), 1–31.

[11] V. Konakov, S. Menozzi, Weak error for the Euler scheme approximation of diffusions with non-smooth coefficients, *Elect. J. Probab.* **22** (2017),1–47.

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

[13] G. Leobacher, M. Szölgyenyi, A numerical method for SDEs with discontinuous drift, *BIT Numer Math.* **56** (2016),151–162.

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

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

[16] V. Lemaire and S. Menozzi, On Some non Asymptotic Bounds for the Euler Scheme, *Electron. J. Probab.* **15** (2010), no 53, 1645–1681.

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

[18] A. Neuenkirch, M. Szölgyenyi, The Euler-Maruyama Scheme for SDEs with Irregular Drift: Convergence Rates via Reduction to a Quadrature Problem, arXiv:1904.07784

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

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

[21] H. L. Ngo, D. Taguchi, Approximation for non-smooth functions of stochastic differential equations with irregular drift. *J. Math. Anal. Appl.* **457** (2018), 361–388.

[22] P. Przybyłowicz, M. Szölgyenyi, Existence, uniqueness, and approximation of solutions of jump-diffusion SDEs with discontinuous drift, arXiv:1912.04215

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

[24] H. Triebel, Theory of Function spaces III. Birkhäuser Verlag, Basel-Boston-Berlin, 2000.

[25] F.-Y. Wang, Estimates for invariant probability measures of degenerate SPDEs with singular and path-dependent drifts. *Proba. Theory* & *Related Fields* (2018), 1–34.

[26] L. Xie, X. Zhang, Sobolev differentiable flows of SDEs with local Sobolev and super-linear growth coefficients, *Ann. Probab.* **44** (2016) 3661–3687.

[27] L. Xie, X. Zhang, Ergodicity of stochastic differential equations with jumps and singular coefficients, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques.* **56** (2020) 175–229.

[28] X. Zhang, Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients, *Stochastic Process. Appl.* **115** (2015),1805–1818.