Journal of Theoretical Probability (2019) 32:848–871
https://doi.org/10.1007/s10959-018-0854-9

CrossMark

Convergence Rate of Euler–Maruyama Scheme for SDEs with Hölder–Dini Continuous Drifts

Jianhai Bao[2] · Xing Huang[1] · Chenggui Yuan[2]

Received: 6 May 2017 / Revised: 30 March 2018 / Published online: 31 August 2018

Abstract
In this paper, we are concerned with convergence rate of Euler–Maruyama scheme for stochastic differential equations with Hölder–Dini continuous drifts. The key contributions are as follows: (i) by means of regularity of non-degenerate Kolmogrov equation, we investigate convergence rate of Euler–Maruyama scheme for a class of stochastic differential equations which allow the drifts to be Dini continuous and unbounded; (ii) by the aid of regularization properties of degenerate Kolmogrov equation, we discuss convergence rate of Euler–Maruyama scheme for a range of degenerate stochastic differential equations where the drifts are Hölder–Dini continuous of order $\frac{2}{3}$ with respect to the first component and are merely Dini-continuous concerning the second component.

Keywords Euler–Maruyama scheme · Convergence rate · Hölder–Dini continuity · Degenerate stochastic differential equation · Kolmogorov equation

Mathematics Subject Classification (2010) 60H35 · 41A25 · 60H10

1 Introduction and Main Results

In their paper [23], Wang and Zhang studied existence and uniqueness for a class of stochastic differential equations (SDEs) with Hölder–Dini continuous drifts; Wang

Supported in part by NNSFC (11771326, 11831014).

✉ Chenggui Yuan
c.yuan@swansea.ac.uk

Jianhai Bao
jianhai.bao@Swansea.ac.uk

Xing Huang
hxsc19880409@163.com

1 Center of Applied Mathematics, Tianjin University, Tianjin 300072, China

2 Department of Mathematics, Swansea University, Singleton Park SA2 8PP, UK

[22] also investigated the strong Feller property, log-Harnack inequality and gradient estimates for SDEs with Dini-continuous drifts. So far there are no numerical schemes available for SDEs with Hölder–Dini continuous drifts. So the aim of this paper is to prove the convergence of Euler–Maruyama (EM) scheme and obtain the rate of convergence for these equations under reasonable conditions.

It is well-known that convergence rate of EM for SDEs with regular coefficients is one-half, see, e.g., [11]. With regard to convergence rate of EM scheme under various settings, we refer to, e.g., [1] for stochastic differential delay equations (SDDEs) with polynomial growth with respect to (w.r.t.) the delay variables, [4] for SDDEs under local Lipschitz and monotonicity condition, [14] for SDEs with discontinuous coefficients, and [25] for SDEs under log-Lipschitz condition, whereas for SDEs with non-globally Lipschitz continuous coefficients; see, e.g., [2,6–8], to name a few. On the other hand, Hairer et al. [5] have established the first result in the literature that Euler's method converges to the solution of an SDE with smooth coefficients in the strong and numerical weak sense without any arbitrarily small polynomial rate of convergence, and Jentzen et al. [9] have further given a counterexample that no approximation method converges to the true solution in the mean square sense with polynomial rate.

The rate of convergence of EM scheme for SDEs with irregular coefficients has also gained much attention. For instance, adopting the Yamada–Watanabe approximation approach, [3] discussed strong convergence rate in L^p-norm sense; using the Yamada–Watanabe approximation trick and heat kernel estimate, [16] studied strong convergence rate in L^1-norm sense for a class of non-degenerate SDEs, where the bounded drift term satisfies a weak monotonicity and is of bounded variation w.r.t. a Gaussian measure and the diffusion term is Hölder continuous; applying the Zvonkin transformation, [18] discussed strong convergence rate in L^p-norm sense for SDEs with additive noises, where the drift coefficient is bounded and Hölder continuous.

It is worth pointing out that [16,18] focused on convergence rate of EM for SDEs with Hölder continuous and bounded drifts, which rules out Hölder–Dini continuous and unbounded drifts. On the other hand, most of the existing literature on convergence rate of EM scheme is concerned with non-degenerate SDEs. Yet the corresponding issue for degenerate SDEs is scarce, to the best of our knowledge. So, in this work, we will not only investigate the convergence of the EM scheme for SDEs with Hölder–Dini continuous drifts, but will also study the degenerate setup. For well-posedness of SDEs with singular coefficients, we refer to, e.g., [13,22,23,27] for more details.

Throughout the paper, the following notation will be used. Let n, m be positive integers, $(\mathbb{R}^n, \langle\cdot,\cdot\rangle, |\cdot|)$ the n-dimensional Euclidean space, and $\mathbb{R}^n\otimes\mathbb{R}^m$ the family of all $n\times m$ matrices. Let $\|\cdot\|$ and $\|\cdot\|_{\rm HS}$ stand for the usual operator norm and the Hilbert–Schmidt norm, respectively. Fix $T>0$ and set $\|f\|_{T,\infty}:=\sup_{t\in[0,T],x\in\mathbb{R}^m}\|f(t,x)\|$ for an operator-valued map f on $[0,T]\times\mathbb{R}^m$. $C(\mathbb{R}^m;\mathbb{R}^n)$ means the continuous functions $f:\mathbb{R}^m\to\mathbb{R}^n$. Let $C^2(\mathbb{R}^n;\mathbb{R}^n\otimes\mathbb{R}^n)$ be the family of all continuously twice differentiable functions $f:\mathbb{R}^n\to\mathbb{R}^n\otimes\mathbb{R}^n$. Denote $\mathbb{M}^n_{\rm non}$ by the collection of all nonsingular $n\times n$-matrices. Let $\mathscr{S}_0$ be the collection of all slowly varying functions $\phi:\mathbb{R}_+\to\mathbb{R}_+$ at zero in Karamata's sense (i.e., $\lim_{t\to0}\frac{\phi(\lambda t)}{\phi(t)}=1$ for any $\lambda>0$), which are bounded from 0 and ∞ on $[\varepsilon, \infty)$ for any $\varepsilon > 0$. Let $\mathscr{D}_0$ be the family of Dini functions, i.e.,

$$\mathscr{D}_0 := \left\{\phi \middle| \phi : \mathbb{R}_+ \to \mathbb{R}_+ \text{ is increasing and } \int_0^1 \frac{\phi(s)}{s}\mathrm{d}s < \infty\right\}.$$

A function $f : \mathbb{R}^m \to \mathbb{R}^n$ is called Dini continuity if there exists $\phi \in \mathscr{D}_0$ such that $|f(x)-f(y)| \le \phi(|x-y|)$ for any $x, y \in \mathbb{R}^m$. We remark that every Dini-continuous function is continuous and every Lipschitz continuous function is Dini continuous; Moreover, if f is Hölder continuous, then f is Dini continuous. Nevertheless, there are numerous Dini-continuous functions, which are not Hölder continuous at all, see, e.g.,

$$\phi(x) = \begin{cases} \frac{1}{(\log(c+x^{-1}))^{(1+\delta)}}, & x > 0 \\ 0, & x = 0 \end{cases}$$

for some constants $\delta > 0$ and $c \ge \mathrm{e}^{3+2\delta}$. Set

$$\mathscr{D} := \{\phi \in \mathscr{D}_0 | \phi^2 \text{ is concave}\} \quad \text{and} \quad \mathscr{D}^\varepsilon := \{\phi \in \mathscr{D} | \phi^{2(1+\varepsilon)} \text{ is concave}\}$$

for some $\varepsilon \in (0, 1)$ sufficiently small. Clearly, ϕ constructed above belongs to $\mathscr{D}^\varepsilon$. A function $f : \mathbb{R}^m \to \mathbb{R}^n$ is called Hölder–Dini continuity of order $\alpha \in [0, 1)$ if

$$|f(x) - f(y)| \le |x-y|^\alpha \phi(|x-y|), \quad |x-y| \le 1$$

for some $\phi \in \mathscr{D}_0$; see, for instance,

$$f(x) = \begin{cases} \frac{1}{(1+x)^\alpha(\log(c+x^{-1}))^{(1+\delta)}}, & x > 0 \\ 0, & x = 0 \end{cases}$$

for some constants $c, \delta > 0$ and $\alpha \in (0, 1)$.

Before proceeding further, a few words about the notation are in order. Generic constants will be denoted by c; we use the shorthand notation $a \lesssim b$ to mean $a \le c\, b$. If the constant c depends on a parameter p, we shall also write c_p and $a \lesssim_p b$. Throughout the paper, for fixed $T > 0$, $C_T > 0$, dependent on the quantity T, is a generic constant which may change from line to line.

1.1 Non-degenerate SDEs with Bounded Coefficients

In this subsection, we consider an SDE on $(\mathbb{R}^n, \langle \cdot, \cdot \rangle, |\cdot|)$

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t, \quad t > 0, \; X_0 = x, \tag{1.1}$$

where $b : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n \otimes \mathbb{R}^n$, and $(W_t)_{t\ge0}$ is an n-dimensional Brownian motion defined on a complete filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge0}, \mathbb{P})$.

With regard to (1.1), we suppose that there exists $\phi \in \mathscr{D}$ such that, for any $s, t \in [0, T]$ and $x, y \in \mathbb{R}^n$,

(A1) $\sigma_t \in C^2(\mathbb{R}^n; \mathbb{R}^n \otimes \mathbb{R}^n)$, $\sigma_t(x) \in \mathbb{M}^n_{\text{non}}$, and

$$\|b\|_{T,\infty} + \sum_{i=0}^{2} \|\nabla^i \sigma\|_{T,\infty} + \|\nabla \sigma^{-1}\|_{T,\infty} + \|\sigma^{-1}\|_{T,\infty} < \infty, \tag{1.2}$$

where ∇^i means the ith order gradient operator;

(A2) (Regularity of b w.r.t. spatial variables)

$$|b_t(x) - b_t(y)| \le \phi(|x - y|);$$

(A3) (Regularity of b and σ w.r.t. time variables)

$$|b_s(x) - b_t(x)| + \|\sigma_s(x) - \sigma_t(x)\|_{\text{HS}} \le \phi(|s - t|).$$

Without loss of generality, we take an integer $N > 0$ sufficiently large such that the stepsize $\delta := T/N \in (0, 1)$. The continuous-time EM scheme corresponding to (1.1) is

$$\mathrm{d}Y_t = b_{t_\delta}(Y_{t_\delta})\mathrm{d}t + \sigma_{t_\delta}(Y_{t_\delta})\mathrm{d}W_t, \quad t > 0, \; Y_0 = X_0 = x. \tag{1.3}$$

Herein, $t_\delta := \lfloor t/\delta \rfloor \delta$ with $\lfloor t/\delta \rfloor$ the integer part of t/δ.

The first contribution in this paper is stated as follows.

Theorem 1.1 *Let* **(A1)**–**(A3)** *hold. Then*

$$\left(\mathbb{E} \left(\sup_{0 \le t \le T} |X_t - Y_t|^2 \right) \right)^{1/2} \lesssim_T \phi(C_T \sqrt{\delta})$$

for some constant $C_T \ge 1$.

Under **(A1)** and **(A2)**, (1.1) admits a unique non-explosive strong solution $(X_t)_{t \in [0,T]}$; see, e.g., [22, Theorem 1.1]. In Theorem 1.1, by taking $\phi(x) = x^\beta$ for $x \ge 0$ and $\beta \in (0, 1]$, and inspecting closely the argument of Theorem 1.1, the concave property of ϕ^2 can be dropped. Moreover, we have

$$\mathbb{E} \left(\sup_{0 \le t \le T} |X_t - Y_t|^2 \right) \lesssim_T \delta^\beta.$$

So, our present result covers [18, Theorem 2.13], where the drift is Hölder continuous. In particular, for the setting $\beta = 1$, it reduces to the classical result on strong convergence of EM scheme for SDEs with regular coefficients; see, e.g., [11] for more details.

1.2 Non-degenerate SDEs with Unbounded Coefficients

As we see, in Theorem 1.1, the coefficients are uniformly bounded, and that the drift term b satisfies the global Dini-continuous condition [see **(A2)** above], which seems to be a little bit stringent. Therefore, concerning the coefficients, it is quite natural to replace uniform boundedness by local boundedness and global Dini continuity by local Dini continuity, respectively.

In lieu of **(A1)**–**(A3)**, as for (1.1) we assume that, for any $s, t \in [0, T]$ and $k \geq 1$,

(A1') $\sigma_t \in C^2(\mathbb{R}^n; \mathbb{R}^n \otimes \mathbb{R}^n)$, for every $x \in \mathbb{R}^n$, $\sigma_t(x) \in \mathbb{M}^n_{\text{non}}$, and

$$|b_t(x)| + \sum_{i=0}^{2} \|\nabla^i \sigma_t(x)\|_{\text{HS}} + \|\nabla \sigma_t^{-1}(x)\|_{\text{HS}} + \|\sigma_t^{-1}(x)\|_{\text{HS}} \leq K_T(1 + |x|), \quad x \in \mathbb{R}^n$$

for some constant $K_T > 0$;

(A2') (Regularity of b w.r.t. spatial variables) There exists $\phi_k \in \mathscr{D}$ such that

$$|b_t(x) - b_t(y)| \leq \phi_k(|x - y|), \quad |x| \vee |y| \leq k;$$

(A3') (Regularity of b and σ w.r.t. time variables) For $\phi_k \in \mathscr{D}$ such that **(A2')**,

$$|b_s(x) - b_t(x)| + \|\sigma_s(x) - \sigma_t(x)\|_{\text{HS}} \leq \phi_k(|s - t|), \quad |x| \leq k.$$

By employing the cutoff approach, Theorem 1.1 can be extended to include SDEs with local Dini-continuous coefficients, which is presented as below.

Theorem 1.2 *Assume* **(A1')**–**(A3')** *hold. Then it holds that*

$$\lim_{\delta \to 0} \mathbb{E}\left(\sup_{0 \leq t \leq T} |X_t - Y_t|^2\right) = 0. \tag{1.4}$$

In particular, if $\phi_k(s) = \mathrm{e}^{\mathrm{e}^{c_0 k^4}} s^\alpha$, $s \geq 0$, *for some* $\alpha \in (0, 1]$ *and* $c_0 > 0$, *then*

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |X_t - Y_t|^2\right) \lesssim \inf_{\varepsilon \in (0,1)} \left\{(\log\log(\delta^{-\alpha\varepsilon}))^{-\frac{1}{4}} + \delta^{\alpha(1-\varepsilon)}\right\}. \tag{1.5}$$

Moreover, if $\sigma_\cdot(\cdot)$ *is uniformly bounded (i.e.,* $\|\sigma\|_{T,\infty} < \infty$*), then*

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |X_t - Y_t|^2\right) \lesssim \inf_{\varepsilon \in (0,1)} \left\{\exp\left(-\frac{1}{C_T \|\sigma\|^2_{T,\infty}} (\log\log(\delta^{-\alpha\varepsilon}))^{\frac{1}{2}}\right) + \delta^{\alpha(1-\varepsilon)}\right\} \tag{1.6}$$

for some constant $C_T > 0$, *where* $\|\sigma\|_{T,\infty} := \sup_{0 \leq t \leq T, x \in \mathbb{R}^n} \|\sigma_t(x)\|_{\text{HS}}$.

Under **(A1')** and **(A2')**, (1.1) enjoys a unique strong solution $(X_t)_{t\in[0,T]}$; see, for instance, [22, Theorem 1.1]. Theorem 1.2 has improved the result in [17] since the drift involved is allowed to be unbounded and local Dini continuous, while the drift in [17] is bounded and Hölder continuous. Furthermore, by comparing (1.5) with (1.6), we infer that the convergence rate of EM scheme is better whenever $\sigma_\cdot(\cdot)$ is uniformly bounded.

Remark 1.3 In fact, in terms of [10, Theorem D], (1.4) holds under **(A1')–(A3')** as well as the pathwise uniqueness of (1.1), whereas in Sect. 4 we provide an alternative proof of (1.4) in order to reveal the convergence rate of the EM scheme.

1.3 Degenerate SDEs

So far, most of the existing literature on convergence of EM scheme for SDEs with irregular coefficients is concerned with non-degenerate SDEs; see, e.g., [16–18] for SDEs driven by Brownian motions, and [18] for SDEs driven by jump processes. The issue for the setup of degenerate SDEs has not yet been considered to date to the best of our knowledge. Nevertheless, in this subsection, we make an attempt to discuss the topic for degenerate SDEs with Hölder–Dini continuous drift.

For notation simplicity, we shall write $\mathbb{R}^{2n}$ instead of $\mathbb{R}^n\times\mathbb{R}^n$. Consider the following degenerate SDE on $\mathbb{R}^{2n}$

$$\begin{cases} dX_t^{(1)} = b_t^{(1)}(X_t^{(1)}, X_t^{(2)})dt, & X_0^{(1)} = x^{(1)} \in \mathbb{R}^n, \\ dX_t^{(2)} = b_t^{(2)}(X_t^{(1)}, X_t^{(2)})dt + \sigma_t(X_t^{(1)}, X_t^{(2)})dW_t, & X_0^{(2)} = x^{(2)} \in \mathbb{R}^n, \end{cases} \tag{1.7}$$

where $b_t^{(1)}, b_t^{(2)} : \mathbb{R}^{2n} \to \mathbb{R}^n$, $\sigma_t : \mathbb{R}^{2n} \to \mathbb{R}^n\otimes\mathbb{R}^n$, and $(W_t)_{t\geq0}$ is an n-dimensional Brownian motion defined on the complete filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq0}, \mathbb{P})$. (1.7) is also called the stochastic Hamiltonian system, which has been investigated extensively in [24,26] on Bismut formulae, in [15] on ergodicity, in [21] on hypercontractivity, and in [23] on wellposedness, to name a few. For applications of the model (1.7), we refer to, e.g., Soize [20].

Write the gradient operator on $\mathbb{R}^{2n}$ as $\nabla = (\nabla^{(1)}, \nabla^{(2)})$, where $\nabla^{(1)}$ and $\nabla^{(2)}$ stand for the gradient operators w.r.t. the first and the second components, respectively.

We assume that there exists $\phi \in \mathscr{D}^{\varepsilon}\cap\mathscr{S}_0$ such that for any $x = (x^{(1)}, x^{(2)})$, $y = (y^{(1)}, y^{(2)}) \in \mathbb{R}^{2n}$ and $s, t \in [0, T]$,

(C1) (Hypoellipticity) $(\nabla^{(2)}b_t^{(1)})(x), \sigma_t(x) \in \mathbb{M}^n_{\rm non}$, and

$$\|b^{(1)}\|_{T,\infty} + \|b^{(2)}\|_{T,\infty} + \|\nabla^{(2)}b^{(1)}\|_{T,\infty} + \left\|(\nabla^{(2)}b^{(1)})^{-1}\right\|_{T,\infty} + \|\sigma\|_{T,\infty} + \|\nabla\sigma\|_{T,\infty} + \|\sigma^{-1}\|_{T,\infty} < \infty;$$

(C2) (Regularity of $b^{(1)}$ w.r.t. spatial variables)

$$|b_t^{(1)}(x) - b_t^{(1)}(y)| \leq |x^{(1)} - y^{(1)}|^{\frac{2}{3}}\phi(|x^{(1)} - y^{(1)}|) \quad \text{if } x^{(2)} = y^{(2)},$$

$$\|(\nabla^{(2)}b_t^{(1)})(x)-(\nabla^{(2)}b_t^{(1)})(y)\|_{\rm HS}\le\phi(|x^{(2)}-y^{(2)}|)\qquad \text{if } x^{(1)}=y^{(1)};$$

(C3) (Regularity of $b^{(2)}$ w.r.t. spatial variables)

$$|b_t^{(2)}(x)-b_t^{(2)}(y)|\le|x^{(1)}-y^{(1)}|^{\frac{2}{3}}\phi(|x^{(1)}-y^{(1)}|)+\phi^{\frac{7}{2}}(|x^{(2)}-y^{(2)}|);$$

(C4) (Regularity of $b^{(1)}$, $b^{(2)}$ and σ w.r.t. time variables)

$$|b_t^{(1)}(x)-b_s^{(1)}(x)|+|b_t^{(2)}(x)-b_s^{(2)}(x)|+\|\sigma_t(x)-\sigma_s(x)\|_{\rm HS}\le\phi(|t-s|).$$

Observe from **(C2)** and **(C3)** that $b^{(1)}(\cdot,x^{(2)})$ and $b^{(2)}(\cdot,x^{(2)})$ with fixed $x^{(2)}$ are locally Hölder–Dini continuous of order $\frac{2}{3}$, and $(\nabla^{(2)}b^{(1)})(x^{(1)},\cdot)$ and $b^{(2)}(x^{(1)},\cdot)$ with fixed $x^{(1)}$ are merely Dini-continuous.

The continuous-time EM scheme associated with (1.7) is as follows:

$$\begin{cases}\mathrm{d}Y_t^{(1)}=b_{t_\delta}^{(1)}(Y_{t_\delta}^{(1)},Y_{t_\delta}^{(2)})\mathrm{d}t, & X_0^{(1)}=x^{(1)}\in\mathbb{R}^n,\\ \mathrm{d}Y_t^{(2)}=b_{t_\delta}^{(2)}(Y_{t_\delta}^{(1)},Y_{t_\delta}^{(2)})\mathrm{d}t+\sigma_{t_\delta}(Y_{t_\delta}^{(1)},Y_{t_\delta}^{(2)})\mathrm{d}W_t, & X_0^{(2)}=x^{(2)}\in\mathbb{R}^n.\end{cases}\tag{1.8}$$

Another contribution in this paper reads as below.

Theorem 1.4 *Let* **(C1)**–**(C4)** *hold. Then*

$$\left(\mathbb{E}\left(\sup_{0\le t\le T}|X_t-Y_t|^2\right)\right)^{1/2}\lesssim_T\phi(C_T\sqrt{\delta})$$

for some constant $C_T\ge1$, in which

$$X_t:=\begin{pmatrix}X_t^{(1)}\\X_t^{(2)}\end{pmatrix}\ \text{and}\ Y_t:=\begin{pmatrix}Y_t^{(1)}\\Y_t^{(2)}\end{pmatrix}.$$

According to [23, Theorem 1.2], (1.7) admits a unique strong solution under the assumptions **(C1)**–**(C3)**. In fact, (1.7) is wellposed under **(C1)**–**(C3)** with $\phi\in\mathscr{D}_0\cap\mathscr{S}_0$ in lieu of $\phi\in\mathscr{D}^\varepsilon\cap\mathscr{S}_0$. Nevertheless, the requirement $\phi\in\mathscr{D}^\varepsilon\cap\mathscr{S}_0$ is imposed in order to reveal the order of convergence for the EM scheme above. By applying the cutoff approach and refining the argument of [23, Theorem 2.3] (see also Lemma 5.1 below), the boundedness of coefficients can be removed. We herein do not go into details since the corresponding trick is quite similar to the proof of Theorem 1.2.

The outline of this paper is organized as follows: In Sect. 2, we elaborate regularity of non-degenerate Kolmogorov equation, which plays an important role in dealing with convergence rate of EM scheme for non-degenerate SDEs with Hölder–Dini continuous and unbounded drifts; In Sects. 3, 4 and 5, we complete the proofs of Theorems 1.1, 1.2 and 1.4, respectively.

2 Regularity of Non-degenerate Kolmogorov Equation

Let $(e_i)_{i\geq 1}$ be an orthogonal basis of $\mathbb{R}^n$. For any $\lambda > 0$, consider the following $\mathbb{R}^n$-valued parabolic equation:

$$\partial_t u_t^\lambda + L_t u_t^\lambda + b_t + \nabla_{b_t} u_t^\lambda = \lambda u_t^\lambda, \quad u_T^\lambda = \mathbf{0_n}, \tag{2.1}$$

where $\nabla_{b_t} u_t^\lambda$ means the directional derivative along the direction b_t, $\mathbf{0_n}$ is the zero vector in $\mathbb{R}^n$ and

$$L_t := \frac{1}{2}\sum_{i,j}\langle(\sigma_t\sigma_t^*)(\cdot)e_i, e_j\rangle\nabla_{e_i}\nabla_{e_j}$$

with σ_t^* standing for the transpose of σ_t. Let $(P_{s,t}^0)_{0\leq s\leq t}$ be the semigroup generated by $(Z_t^{s,x})_{0\leq s\leq t}$ which solves an SDE below

$$\mathrm{d}Z_t^{s,x} = \sigma_t(Z_t^{s,x})\mathrm{d}W_t, \quad t > s, \ Z_s^{s,x} = x. \tag{2.2}$$

By the chain rule, it follows from (2.1) that

$$\begin{aligned}\partial_t\left(\mathrm{e}^{-\lambda(t-s)}P_{s,t}^0 u_t^\lambda\right) &= \mathrm{e}^{-\lambda(t-s)}\left\{-\lambda P_{s,t}^0 u_t^\lambda + P_{s,t}^0 L_t u_t^\lambda + P_{s,t}^0\partial_t u_t^\lambda\right\}\\ &= -\mathrm{e}^{-\lambda(t-s)}P_{s,t}^0\left\{b_t + \nabla_{b_t}u_t^\lambda\right\}.\end{aligned}$$

Thus, integrating from s to T and taking advantage of $u_T^\lambda = \mathbf{0_n}$, we arrive at

$$u_s^\lambda = \int_s^T \mathrm{e}^{-\lambda(t-s)}P_{s,t}^0\left\{b_t + \nabla_{b_t}u_t^\lambda\right\}\mathrm{d}t. \tag{2.3}$$

For notation simplicity, let

$$\Lambda_{T,\sigma} = \mathrm{e}^{\frac{T}{2}\|\nabla\sigma\|_{T,\infty}^2}\|\sigma^{-1}\|_{T,\infty} \tag{2.4}$$

and

$$\begin{aligned}\tilde{\Lambda}_{T,\sigma} = 48\mathrm{e}^{288\,T^2\|\nabla\sigma\|_{T,\infty}^4}&\Big\{6\sqrt{2}\mathrm{e}^{T\|\nabla\sigma\|_{T,\infty}^2}\|\sigma^{-1}\|_{T,\infty}^4 + T\|\nabla\sigma^{-1}\|_{T,\infty}^2\\ &+2T^2\|\nabla^2\sigma\|_{T,\infty}^2\|\sigma^{-1}\|_{T,\infty}^2\mathrm{e}^{2T\|\nabla\sigma\|_{T,\infty}^2}\Big\}.\end{aligned} \tag{2.5}$$

Moreover, set

$$\Upsilon_{T,\sigma} := \sqrt{\tilde{\Lambda}_{T,\sigma}}\left\{3 + 2\|b\|_{T,\infty} + 28\left(\Lambda_{T,\sigma} + \sqrt{\tilde{\Lambda}_{T,\sigma}}\right)\|b\|_{T,\infty}^2\right\}. \tag{2.6}$$

The lemma below plays a crucial role in investigating error analysis.

Lemma 2.1 *Under* **(A1)** *and* **(A2)**, *for any* $\lambda \geq 9\pi\Lambda_{T,\sigma}^2\|b\|_{T,\infty}^2 + 4(\|b\|_{T,\infty} + \Lambda_{T,\sigma})^2$,

(i) (2.1) (*i.e.,* (2.3)) *enjoys a unique strong solution* $u^\lambda \in C([0,T]; C_b^1(\mathbb{R}^n;\mathbb{R}^n))$*;*
(ii) $\|\nabla u^\lambda\|_{T,\infty} \le \frac{1}{2}$*;*
(iii) $\|\nabla^2 u^\lambda\|_{T,\infty} \le \Upsilon_{T,\sigma} \int_0^T \frac{e^{-\lambda t}}{t}\tilde\phi(\|\sigma\|_{T,\infty}\sqrt{t})dt$*, where* $\tilde\phi(s) := \sqrt{\phi^2(s)+s}$*,* $s \ge 0$.

Proof To show (i)–(iii), it boils down to refine the argument of [22, Lemma 2.1]. (i) holds for any $\lambda \ge 4(\|b\|_{T,\infty} + \Lambda_{T,\sigma})^2$ via the Banach fixed-point theorem.

In what follows, we aim to show (ii) and (iii) hold true, one-by-one. Observe from [12, Theorem 3.1, p.218] that

$$d\nabla_\eta Z_t^{s,x} = \left(\nabla_{\nabla_\eta Z_t^{s,x}}\sigma_t\right)\left(Z_t^{s,x}\right)dW_t, \quad t \ge s, \quad \nabla_\eta Z_s^{s,x} = \eta \in \mathbb{R}^n. \tag{2.7}$$

Using Itô's isometry and Gronwall's inequality, one has

$$\mathbb{E}|\nabla_\eta Z_t^{s,x}|^2 \le |\eta|^2 e^{T\|\nabla\sigma\|_{T,\infty}^2}. \tag{2.8}$$

Utilizing the BDG inequality, we deduce that

$$\mathbb{E}|\nabla_\eta Z_t^{s,x}|^4 \le 8\left\{|\eta|^4 + 36(t-s)\|\nabla\sigma\|_{T,\infty}^4 \int_s^t \mathbb{E}|\nabla_\eta Z_u^{s,x}|^4 du\right\},$$

which, combining with Gronwall's inequality, yields that

$$\mathbb{E}|\nabla_\eta Z_t^{s,x}|^4 \le 8|\eta|^4 e^{288\,T^2\|\nabla\sigma\|_{T,\infty}^4}. \tag{2.9}$$

Recall from [22, (2.8)] that the following Bismut formula

$$\nabla_\eta P_{s,t}^0 f(x) = \mathbb{E}\left(\frac{f\left(Z_t^{s,x}\right)}{t-s}\int_s^t \left\langle \sigma_r^{-1}\left(Z_r^{s,x}\right)\nabla_\eta Z_r^{s,x}, dW_r\right\rangle\right), \quad f \in \mathscr{B}_b(\mathbb{R}^n) \tag{2.10}$$

holds. By the Cauchy–Schwartz inequality, the Itô isometry and (2.8), we obtain that

$$|\nabla_\eta P_{s,t}^0 f|^2(x) \le \frac{\Lambda_{T,\sigma}^2 |\eta|^2 P_{s,t}^0 f^2(x)}{t-s}, \quad f \in \mathscr{B}_b(\mathbb{R}^n), \tag{2.11}$$

where $\Lambda_{T,\sigma} > 0$ is defined in (2.4). So, one infers from (2.3) and (2.11) that

$$\begin{aligned}\|\nabla u_s^\lambda\| &\le \int_s^T e^{-\lambda(t-s)}\|\nabla P_{s,t}^0\{b_t + \nabla_{b_t} u_t^\lambda\}\|dt\\ &\le \Lambda_{T,\sigma}\left(1+\|\nabla u^\lambda\|_{T,\infty}\right)\|b\|_{T,\infty}\int_0^T \frac{e^{-\lambda t}}{\sqrt{t}}dt\\ &\le \lambda^{-\frac{1}{2}}\sqrt{\pi}\Lambda_{T,\sigma}\|b\|_{T,\infty}(1+\|\nabla u^\lambda\|_{T,\infty}).\end{aligned}$$

Thus, (ii) follows by taking $\lambda \ge 9\pi\Lambda_{T,\sigma}^2\|b\|_{T,\infty}^2$.

In the sequel, we intend to verify (iii). Set $\gamma_{s,t} := \nabla_\eta \nabla_{\eta'} Z_t^{s,x}$ for any $\eta, \eta' \in \mathbb{R}^n$. Notice from (2.7) that

$$\mathrm{d}\gamma_{s,t} = \left\{ \left(\nabla_{\gamma_{s,t}} \sigma_t\right) \left(Z_t^{s,x}\right) + \left(\nabla_{\nabla_\eta Z_t^{s,x}} \nabla_{\nabla_{\eta'} Z_t^{s,x}} \sigma_t\right) \left(Z_t^{s,x}\right) \right\} \mathrm{d}W_t, \quad t \geq s, \ \gamma_{s,s} = \mathbf{0_n}.$$

By the Doob submartingale inequality and the Itô isometry, besides the Gronwall inequality and (2.8), we derive that

$$\sup_{s \leq t \leq T} \mathbb{E}|\gamma_{s,t}|^2 \leq 16T \|\nabla^2 \sigma\|_{T,\infty}^2 \mathrm{e}^{288T^2 \|\nabla\sigma\|_{T,\infty}^4 + 2T \|\nabla\sigma\|_{T,\infty}^2} |\eta|^2 |\eta'|^2. \tag{2.12}$$

From (2.10) and the Markov property, we have

$$\nabla_\eta P_{s,t}^0 f(x) = \mathbb{E} \left(\frac{\left(P^0_{\frac{t+s}{2},t} f\right) \left(Z^{s,x}_{\frac{t+s}{2}}\right)}{(t-s)/2} \int_s^{\frac{t+s}{2}} \left\langle \sigma_r^{-1} \left(Z_r^{s,x}\right) \nabla_\eta Z_r^{s,x}, \mathrm{d}W_r \right\rangle \right).$$

This further gives that

$$\begin{aligned}
&\frac{1}{2} \left(\nabla_{\eta'} \nabla_\eta P_{s,t}^0 f\right)(x) \\
&= \mathbb{E} \left(\frac{\left(\nabla_{\nabla_{\eta'} Z^{s,x}_{\frac{t+s}{2}}} P^0_{\frac{t+s}{2},t} f\right) \left(Z^{s,x}_{\frac{t+s}{2}}\right)}{t-s} \int_s^{\frac{t+s}{2}} \left\langle \sigma_r^{-1} \left(Z_r^{s,x}\right) \nabla_\eta Z_r^{s,x}, \mathrm{d}W_r \right\rangle \right) \\
&+ \mathbb{E} \left(\frac{\left(P^0_{\frac{t+s}{2},t} f\right) \left(Z^{s,x}_{\frac{t+s}{2}}\right)}{t-s} \int_s^{\frac{t+s}{2}} \left\langle \left(\nabla_{\nabla_{\eta'} Z_r^{s,x}} \sigma_r^{-1}\right) \left(Z_r^{s,x}\right) \nabla_\eta Z_r^{s,x}, \mathrm{d}W_r \right\rangle \right) \\
&+ \mathbb{E} \left(\frac{\left(P^0_{\frac{t+s}{2},t} f\right) \left(Z^{s,x}_{\frac{t+s}{2}}\right)}{t-s} \int_s^{\frac{t+s}{2}} \left\langle \sigma_r^{-1} \left(Z_r^{s,x}\right) \nabla_{\eta'} \nabla_\eta Z_r^{s,x}, \mathrm{d}W_r \right\rangle \right).
\end{aligned}$$

Thus, applying Cauchy–Schwartz's inequality and Itô's isometry and taking (2.9), (2.11) and (2.12) into consideration, we derive that

$$\begin{aligned}
&|\nabla_{\eta'} \nabla_\eta P_{s,t}^0 f|^2(x) \\
&\leq 12 \left\{ 6 \|\sigma^{-1}\|_{T,\infty}^2 \frac{\mathbb{E} \left| \nabla P^0_{\frac{t+s}{2},t} f \right|^2 \left(Z^{s,x}_{\frac{t+s}{2}}\right)}{(t-s)^{5/2}} \right. \\
&\quad \times \left(\mathbb{E} \left| \nabla_{\eta'} Z^{s,x}_{\frac{t+s}{2}} \right|^4 \right)^{1/2} \left(\int_s^{\frac{t+s}{2}} \mathbb{E} \left| \nabla_\eta Z_r^{s,x} \right|^4 \mathrm{d}r \right)^{1/2}
\end{aligned}$$

$$+\frac{P_{s,t}^0 f^2(x)}{(t-s)^2}\|\nabla\sigma^{-1}\|_{T,\infty}^2\int_s^{\frac{t+s}{2}}\left(\mathbb{E}|\nabla_{\eta'}Z_r^{s,x}|^4\right)^{1/2}\left(\mathbb{E}|\nabla_\eta Z_r^{s,x}|^4\right)^{1/2}\mathrm{d}r$$
$$+\frac{P_{s,t}^0 f^2(x)}{(t-s)^2}\|\sigma^{-1}\|_{T,\infty}^2\int_s^{\frac{t+s}{2}}\mathbb{E}\left|\nabla_{\eta'}\nabla_\eta Z_r^{s,x}\right|^2\mathrm{d}r\Bigg\}$$
$$\leq \tilde{\Lambda}_{T,\sigma}|\eta|^2|\eta'|^2\frac{P_{s,t}^0 f^2(x)}{(t-s)^2}, \tag{2.13}$$

where $\tilde{\Lambda}_{T,\sigma}>0$ is defined as in (2.5).

Set $\tilde{f}(\cdot):=f(\cdot)-f(x)$ for fixed $x\in\mathbb{R}^n$ and $f\in\mathscr{B}_b(\mathbb{R}^n)$ which verifies

$$|f(x)-f(y)|\leq\phi(|x-y|),\quad x,y\in\mathbb{R}^n \tag{2.14}$$

for some $\phi\in\mathscr{D}$. For $f\in\mathscr{B}_b(\mathbb{R}^n)$ such that (2.14), (2.13) implies that

$$\begin{aligned}|\nabla_{\eta'}\nabla_\eta P_{s,t}^0 f|^2(x)=|\nabla_{\eta'}\nabla_\eta P_{s,t}^0\tilde{f}|^2(x)&\leq\frac{\tilde{\Lambda}_{T,\sigma}|\eta|^2|\eta'|^2}{(t-s)^2}\mathbb{E}|f(Z_t^{s,x})-f(x)|^2\\&\leq\frac{\tilde{\Lambda}_{T,\sigma}|\eta|^2|\eta'|^2}{(t-s)^2}\phi^2(\|\sigma\|_{T,\infty}(t-s)^{1/2}),\end{aligned} \tag{2.15}$$

where in the second display we have used that

$$Z_t^{s,x}-x=\int_s^t\sigma_r(Z_r^{s,x})\mathrm{d}W_r,$$

and utilized Jensen's inequality as well as Itô's isometry.

Let $f_t=b_t+\nabla_{b_t}u_t^\lambda$. For any $\lambda\geq 9\pi\Lambda_{T,\sigma}^2\|b\|_{T,\infty}^2+4(\|b\|_{T,\infty}+\Lambda_{T,\sigma})^2$, note from (ii), (2.11) and (2.13) that

$$\begin{aligned}|f_t(x)-f_t(y)|&\leq(1+\|\nabla u^\lambda\|_{T,\infty})\phi(|x-y|)\\&\quad+\|b\|_{T,\infty}\|\nabla u_t^\lambda(x)-\nabla u_t(y)\|\mathbf{1}_{\{|x-y|\geq1\}}\\&\quad+\|b\|_{T,\infty}\|\nabla u_t^\lambda(x)-\nabla u_t(y)\|\mathbf{1}_{\{|x-y|\leq1\}}\\&\leq\frac{3}{2}\phi(|x-y|)+\|b\|_{T,\infty}\sqrt{|x-y|}\mathbf{1}_{\{|x-y|\geq1\}}\\&\quad+10\left(\Lambda_{T,\sigma}+\sqrt{\tilde{\Lambda}_{T,\sigma}}\right)\|b\|_{T,\infty}^2\sqrt{|x-y|}\sqrt{|x-y|}\\&\quad\times\log\left(\mathrm{e}+\frac{1}{|x-y|}\right)\mathbf{1}_{\{|x-y|\leq1\}}\\&\leq\left\{3+2\|b\|_{T,\infty}+28\left(\Lambda_{T,\sigma}+\sqrt{\tilde{\Lambda}_{T,\sigma}}\right)\|b\|_{T,\infty}^2\right\}\tilde{\phi}(|x-y|)\end{aligned}$$

with $\tilde{\phi}(s):=\sqrt{\phi^2(s)+s}$, $s\geq0$, where in the second inequality we have used [22, Lemma 2.2 (1)], and the fact that the function $[0,1]\ni x\mapsto\sqrt{x}\log(\mathrm{e}+\frac{1}{x})$ is non-decreasing. As a result, (iii) follows from (2.15). □

3 Proof of Theorem 1.1

With Lemma 2.1 in hand, we now in the position to complete the

Proof of Theorem 1.1 Throughout the whole proof, we assume $\lambda \geq 9\pi\Lambda_{T,\sigma}^2\|b\|_{T,\infty}^2 + 4(\|b\|_{T,\infty} + \Lambda_{T,\sigma})^2$ so that (i)–(iii) in Lemma 2.1 hold. For any $t \in [0, T]$, applying Itô's formula to $x + u_t^\lambda(x)$, $x \in \mathbb{R}^n$, we deduce from (2.1) that

$$X_t + u_t^\lambda(X_t) = x + u_0^\lambda(x) + \lambda \int_0^t u_s^\lambda(X_s)\mathrm{d}s + \int_0^t \{\mathbf{I_{n\times n}} + (\nabla u_s^\lambda)(\cdot)\}(X_s)\sigma_s(X_s)\mathrm{d}W_s, \tag{3.1}$$

where $\mathbf{I_{n\times n}}$ is an $n \times n$ identity matrix, and that

$$\begin{aligned} Y_t + u_t^\lambda(Y_t) &= x + u_0^\lambda(x) + \lambda \int_0^t u_s^\lambda(Y_s)\mathrm{d}s + \int_0^t \{\mathbf{I_{n\times n}} + (\nabla u_s^\lambda)(\cdot)\}(Y_s)\sigma_{s_\delta}(Y_{s_\delta})\mathrm{d}W_s \\ &\quad + \int_0^t \{\mathbf{I_{n\times n}} + (\nabla u_s^\lambda)(\cdot)\}(Y_s)\{b_{s_\delta}(Y_{s_\delta}) - b_s(Y_s)\}\mathrm{d}s \\ &\quad + \frac{1}{2}\int_0^t \sum_{k,j} \big\langle \{(\sigma_{s_\delta}\sigma_{s_\delta}^*)(Y_{s_\delta}) - (\sigma_s\sigma_s^*)(Y_s)\}e_k, e_j\big\rangle (\nabla_{e_k}\nabla_{e_j}u_s^\lambda)(Y_s)\mathrm{d}s. \end{aligned} \tag{3.2}$$

For notation simplicity, set

$$M_t^\lambda := X_t - Y_t + u_t^\lambda(X_t) - u_t^\lambda(Y_t). \tag{3.3}$$

Using the elementary inequality: $(a+b)^2 \leq (1+\varepsilon)(a^2+\varepsilon^{-1}b^2)$ for arbitrary $\varepsilon, a, b > 0$, we derive from (ii) that

$$\begin{aligned} |X_t - Y_t|^2 &\leq (1+\varepsilon)(|M_t^\lambda|^2 + \varepsilon^{-1}|u_t^\lambda(X_t) - u_t^\lambda(Y_t)|^2) \\ &\leq (1+\varepsilon)\left(|M_t^\lambda|^2 + \frac{\varepsilon^{-1}}{4}|X_t - Y_t|^2\right). \end{aligned}$$

In particular, taking $\varepsilon = 1$ leads to

$$|X_t - Y_t|^2 \leq \frac{1}{2}|X_t - Y_t|^2 + 2|M_t^\lambda|^2.$$

As a consequence,

$$\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s - Y_s|^2\right) \leq 4\mathbb{E}\left(\sup_{0\leq s\leq t}|M_s^\lambda|^2\right). \tag{3.4}$$

In what follows, our goal is to estimate the term on the right-hand side of (3.4). Observe from the definition of the Hilbert–Schmidt norm that

$$\int_0^t \mathbb{E}\left|\sum_{k,j}\langle\{(\sigma_{s_\delta}\sigma_{s_\delta}^*)(Y_{s_\delta})-(\sigma_s\sigma_s^*)(Y_s)\}e_k,e_j\rangle(\nabla_{e_k}\nabla_{e_j}u_s^\lambda)(Y_s)\right|^2\mathrm{d}s$$
$$\lesssim_T \|\nabla^2u^\lambda\|_{T,\infty}^2\int_0^t\mathbb{E}\|(\sigma_{s_\delta}\sigma_{s_\delta}^*)(Y_{s_\delta})-(\sigma_s\sigma_s^*)(Y_s)\|_{\mathrm{HS}}^2\mathrm{d}s. \tag{3.5}$$

Thus, by Hölder's inequality, Doob's submartingale inequality and Itô's isometry, it follows from (3.1), (3.2) and (3.5) that

$$\begin{aligned}
\mathbb{E}\left(\sup_{0\le s\le t}|M_s^\lambda|^2\right)\le C_T\Big\{&\lambda^2\int_0^t\mathbb{E}|u_s^\lambda(X_s)-u_s^\lambda(Y_s)|^2\mathrm{d}s\\
&+(1+\|\nabla u\|_{T,\infty}^2)\int_0^t\mathbb{E}|b_{s_\delta}(Y_s)-b_{s_\delta}(Y_{s_\delta})|^2\mathrm{d}s\\
&+(1+\|\nabla u\|_{T,\infty}^2)\int_0^t\mathbb{E}|b_s(Y_s)-b_{s_\delta}(Y_s)|^2\mathrm{d}s\\
&+\int_0^t\mathbb{E}\|\{(\nabla u_s^\lambda)(X_s)-(\nabla u_s^\lambda)(Y_s)\}\sigma_s(X_s)\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+(1+\|\nabla u\|_{T,\infty}^2)\int_0^t\mathbb{E}\|\sigma_{s_\delta}(X_s)-\sigma_{s_\delta}(Y_{s_\delta})\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+\|\nabla^2u^\lambda\|_{T,\infty}^2\int_0^t\mathbb{E}\|\{\sigma_{s_\delta}(Y_s)-\sigma_{s_\delta}(Y_{s_\delta})\}\sigma_{s_\delta}^*(Y_{s_\delta})\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+\|\nabla^2u^\lambda\|_{T,\infty}^2\int_0^t\mathbb{E}\|\sigma_s(Y_s)\{\sigma_{s_\delta}^*(Y_s)-\sigma_{s_\delta}^*(Y_{s_\delta})\}\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+(1+\|\nabla u\|_{T,\infty}^2)\int_0^t\mathbb{E}\|\sigma_s(X_s)-\sigma_{s_\delta}(X_s)\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+\|\nabla^2u^\lambda\|_{T,\infty}^2\int_0^t\mathbb{E}\|\sigma_s(Y_s)\{\sigma_s^*(Y_s)-\sigma_{s_\delta}^*(Y_s)\}\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+\|\nabla^2u^\lambda\|_{T,\infty}^2\int_0^t\mathbb{E}\|\{\sigma_s(Y_s)-\sigma_{s_\delta}(Y_s)\}\sigma_{s_\delta}^*(Y_{s_\delta})\|_{\mathrm{HS}}^2\mathrm{d}s\Big\}\\
=:C_T&\left(\sum_{i=1}^{10}I_i(t)\right)
\end{aligned}$$

for some constant $C_T>0$. Also, applying Hölder's inequality and Itô's isometry, we deduce from (**A1**) that

$$\mathbb{E}|Y_t-Y_{t_\delta}|^2\le\beta_T\delta \tag{3.6}$$

for some constant $\beta_T\ge1$. By Taylor's expansion, it is obvious to see that

$$I_1(t)+I_4(t)\lesssim\{\lambda^2\|\nabla u^\lambda\|_{T,\infty}^2+\|\nabla^2u^\lambda\|_{T,\infty}^2\|\sigma\|_{T,\infty}^2\}\int_0^t\mathbb{E}|X_s-Y_s|^2\mathrm{d}s. \tag{3.7}$$

From (**A3**) and due to the fact that $\phi(\cdot)$ is increasing and $\delta \in (0, 1)$, one has

$$I_3(t) + \sum_{i=8}^{10} I_i(t) \lesssim_T \{1 + \|\nabla u^\lambda\|_{T,\infty}^2 + \|\nabla^2 u^\lambda\|_{T,\infty}^2 \|\sigma\|_{T,\infty}^2\}\phi^2(\sqrt{\delta}). \tag{3.8}$$

In view of (**A2**), we derive that

$$\begin{aligned} I_2(t) &+ \sum_{i=5}^{7} I_i(t) \\ &\lesssim \{1 + \|\nabla u^\lambda\|_{T,\infty}^2\} \int_0^t \mathbb{E}\phi(|Y_s - Y_{s_\delta}|)^2 \mathrm{d}s \\ &+ \{1 + \|\nabla u^\lambda\|_{T,\infty}^2\} \|\nabla\sigma\|_{T,\infty}^2 \int_0^t \mathbb{E}|X_s - Y_s|^2 \mathrm{d}s \\ &+ \{1 + \|\nabla u^\lambda\|_{T,\infty}^2 + \|\nabla^2 u^\lambda\|_{T,\infty}^2 \|\sigma\|_{T,\infty}^2\} \|\nabla\sigma\|_{T,\infty}^2 \int_0^t \mathbb{E}|Y_s - Y_{s_\delta}|^2 \mathrm{d}s. \end{aligned} \tag{3.9}$$

Thus, taking (3.6)–(3.9) into account and applying Jensen's inequality gives that

$$\mathbb{E}\left(\sup_{0\le s\le t} |M_s^\lambda|^2\right) \lesssim_T C_{T,\sigma,\lambda}\{\delta + \phi^2(\beta_T\sqrt{\delta})\} + C_{T,\sigma,\lambda} \int_0^t \mathbb{E}|X_s - Y_s|^2 \mathrm{d}s,$$

where

$$C_{T,\sigma,\lambda} := \{1 + \|\nabla\sigma\|_{T,\infty}^2\} \left\{\frac{5}{4} + (1 + \lambda^2)\|\nabla^2 u^\lambda\|_{T,\infty}^2 \|\sigma\|_{T,\infty}^2\right\}. \tag{3.10}$$

Owing to $\phi \in \mathscr{D}$, we conclude that $\phi(0) = 0$, $\phi' > 0$ and $\phi'' < 0$ so that, for any $c > 0$ and $\delta \in (0, 1)$,

$$\phi(c\delta) = \phi(0) + \phi'(\xi)c\delta \ge \phi'(c)c\delta,$$

where $\xi \in (0, c\delta)$. This further implies that

$$\mathbb{E}\left(\sup_{0\le s\le t} |M_s^\lambda|^2\right) \lesssim_T C_{T,\sigma,\lambda}\phi^2(\beta_T\sqrt{\delta}) + C_{T,\sigma,\lambda} \int_0^t \mathbb{E}|X_s - Y_s|^2 \mathrm{d}s.$$

Substituting this into (3.4) gives that

$$\mathbb{E}\left(\sup_{0\le s\le t} |X_s - Y_s|^2\right) \lesssim_T C_{T,\sigma,\lambda}\phi^2(\beta_T\sqrt{\delta}) + C_{T,\sigma,\lambda} \int_0^t \mathbb{E}|X_s - Y_s|^2 \mathrm{d}s.$$

Thus, Gronwall's inequality implies that there exists $\tilde{C}_T > 0$ such that

$$\mathbb{E}\left(\sup_{0\le s\le t}|X_s - Y_s|^2\right) \le \tilde{C}_T C_{T,\sigma,\lambda} e^{\tilde{C}_T C_{T,\sigma,\lambda}} \phi^2(\beta_T\sqrt{\delta}). \tag{3.11}$$

So the desired assertion holds immediately. □

4 Proof of Theorem 1.2

We shall adopt the cutoff approach to finish the

Proof of Theorem 1.2 Take $\psi \in C_b^\infty(\mathbb{R}_+)$ such that $0 \le \psi \le 1$, $\psi(r) = 1$ for $r \in [0, 1]$ and $\psi(r) = 0$ for $r \ge 2$. For any $t \in [0, T]$ and $k \ge 1$, define the cutoff functions

$$b_t^{(k)}(x) = b_t(x)\psi(|x|/k) \quad \text{and} \quad \sigma_t^{(k)}(x) = \sigma_t(\psi(|x|/k)x), \quad x \in \mathbb{R}^n.$$

It is easy to see that $b^{(k)}$ and $\sigma^{(k)}$ satisfy **(A1)**. For fixed $k \ge 1$, consider the following SDE

$$\mathrm{d}X_t^{(k)} = b_t^{(k)}(X_t^{(k)})\mathrm{d}t + \sigma_t^{(k)}(X_t^{(k)})\mathrm{d}W_t, \quad t > 0, \ X_0^{(k)} = X_0 = x. \tag{4.1}$$

The corresponding continuous-time EM of (4.1) is defined by

$$\mathrm{d}Y_t^{(k)} = b_{t_\delta}^{(k)}(Y_{t_\delta}^{(k)})\mathrm{d}t + \sigma_{t_\delta}^{(k)}(Y_{t_\delta}^{(k)})\mathrm{d}W_t, \quad t > 0, \ Y_0^{(k)} = X_0 = x. \tag{4.2}$$

Applying BDG's inequality, Hölder's inequality and Gronwall's inequality, we deduce from **(A1')** that

$$\mathbb{E}\left(\sup_{0\le t\le T}|X_t|^4\right) + \mathbb{E}\left(\sup_{0\le t\le T}|Y_t|^4\right) + \mathbb{E}\left(\sup_{0\le t\le T}|X_t^{(k)}|^4\right) + \mathbb{E}\left(\sup_{0\le t\le T}|Y_t^{(k)}|^4\right) \le C_T \tag{4.3}$$

for some constant $C_T > 0$. Note that

$$\begin{aligned}\mathbb{E}\left(\sup_{0\le t\le T}|X_t - Y_t|^2\right) &\le 3\,\mathbb{E}\left(\sup_{0\le t\le T}|X_t - X_t^{(k)}|^2\right) + 3\,\mathbb{E}\left(\sup_{0\le t\le T}|X_t^{(k)} - Y_t^{(k)}|^2\right)\\ &\quad + 3\,\mathbb{E}\left(\sup_{0\le t\le T}|Y_t - Y_t^{(k)}|^2\right)\\ &=: I_1 + I_2 + I_3.\end{aligned}$$

For the terms I_1 and I_3, in terms of the Chebyshev inequality we find from (4.3) that

$$\begin{aligned} I_1 + I_3 &\lesssim \mathbb{E}\left(\sup_{0\le t\le T}|X_t - X_t^{(k)}|^2\mathbf{1}_{\{\sup_{0\le t\le T}|X_t|\ge k\}}\right)\\ &\quad + \mathbb{E}\left(\sup_{0\le t\le T}|Y_t - Y_t^{(k)}|^2\mathbf{1}_{\{\sup_{0\le t\le T}|Y_t|\ge k\}}\right)\\ &\lesssim \sqrt{\mathbb{E}\left(\sup_{0\le t\le T}|X_t|^4\right)+\mathbb{E}\left(\sup_{0\le t\le T}|X_t^{(k)}|^4\right)}\frac{\sqrt{\mathbb{E}\left(\sup_{0\le t\le T}|X_t|^2\right)}}{k}\\ &\quad + \sqrt{\mathbb{E}\left(\sup_{0\le t\le T}|Y_t|^4\right)+\mathbb{E}\left(\sup_{0\le t\le T}|Y_t^{(k)}|^4\right)}\frac{\sqrt{\mathbb{E}\left(\sup_{0\le t\le T}|Y_t|^2\right)}}{k}\\ &\lesssim_T \frac{1}{k}, \end{aligned} \tag{4.4}$$

where in the first display we have used the facts that $\{X_t \neq X_t^{(k)}\} \subset \{\sup_{0\le s\le t}|X_s| \ge k\}$ and $\{Y_t \neq Y_t^{(k)}\} \subset \{\sup_{0\le s\le t}|Y_s| \ge k\}$. Observe from **(A1')** that $9\pi\Lambda^2_{T,\sigma^{(k)}}\|b^{(k)}\|^2_{T,\infty} + 4(\|b^{(k)}\|_{T,\infty} + \Lambda_{T,\sigma^{(k)}})^2 \le \mathrm{e}^{c\,k^2}$ for some $c > 0$. Next, according to (3.11), by taking $\lambda = \mathrm{e}^{c\,k^2}$ there exits $C_T > 0$ such that

$$I_2 \le \mathrm{e}^{C_T C_{T,\sigma^{(k)},\lambda}}\phi_k^2(\beta_T\sqrt{\delta}). \tag{4.5}$$

Herein, $C_{T,\sigma^{(k)},\lambda} > 0$ is defined as in (3.10) with σ and u^λ replaced by $\sigma^{(k)}$ and $u^{\lambda,k}$, respectively, where $u^{\lambda,k}$ solves (2.3) by writing $b^{(k)}$ instead of b. Consequently, we conclude that

$$\mathbb{E}\left(\sup_{0\le t\le T}|X_t - Y_t|^2\right) \le \frac{\bar{c}_0}{k} + \bar{c}_0\mathrm{e}^{C_T C_{T,\sigma^{(k)},\lambda}}\phi_k^2(\beta_T\sqrt{\delta}) \tag{4.6}$$

for some $\bar{c}_0 > 0$. For any $\varepsilon > 0$, taking $k = \lfloor 2\bar{c}_0/\varepsilon\rfloor$ and letting δ go to zero implies that

$$\lim_{\delta\to 0}\mathbb{E}\left(\sup_{0\le t\le T}|X_t - Y_t|^2\right) \le \varepsilon.$$

Thus, (1.4) follows due to the arbitrariness of ε.

For $\phi_k(s) = \mathrm{e}^{\mathrm{e}^{c_0 k^4}}s^\alpha$, $s \ge 0$, with $\alpha \in (0, 1]$, we deduce from Lemma 2.1 (iii) that

$$\|\nabla^2 u^{\lambda,k}\|_{T,\infty} \le \frac{1}{2} \tag{4.7}$$

whenever

$$\lambda \geq \left\{2\Upsilon_{T,\sigma^{(k)}}\left(\mathrm{e}^{\mathrm{e}^{c_0k^4}}\|\sigma^{(k)}\|_{T,\infty}^{\alpha}\Gamma(\alpha/2)+\|\sigma^{(k)}\|_{T,\infty}^{1/2}\Gamma(1/4)\right)\right\}^{2/\alpha} + 9\pi(\Lambda_{T,\sigma^{(k)}})^2\|b^{(k)}\|_{T,\infty}^2+4(\|b^{(k)}\|_{T,\infty}+\Lambda_{T,\sigma^{(k)}})^2. \tag{4.8}$$

Since the right-hand side of (4.8) can be bounded by $\mathrm{e}^{\mathrm{e}^{\bar{C}_T k^4}}$ for some constant $\bar{C}_T > 0$ due to **(A1')**, we can take $\lambda = \mathrm{e}^{\mathrm{e}^{\bar{C}_T k^4}}$ so that (4.7) holds. Thus, (4.6), together with (4.7) and **(A1')**, yields that

$$\mathbb{E}\left(\sup_{0\leq t\leq T}|X_t-Y_t|^2\right)\leq \frac{\hat{C}_T}{k}+\hat{C}_T\mathrm{e}^{\mathrm{e}^{\tilde{C}_Tk^4}}\delta^{\alpha}$$

for some constants $\hat{C}_T, \tilde{C}_T > 0$. Thus, (1.5) follows immediately by taking

$$k=\left\lfloor\left(\frac{1}{\tilde{C}_T}\log\log\delta^{-\alpha\varepsilon}\right)^{\frac{1}{4}}\right\rfloor. \tag{4.9}$$

Next, we aim to show that (1.6) holds true. In view of (4.3) and (4.4), it follows from Hölder's inequality that

$$\begin{aligned} I_1+I_3 &\lesssim \sqrt{\mathbb{E}\left(\sup_{0\leq t\leq T}|X_t-X_t^{(k)}|^4\right)}\sqrt{\mathbb{P}\left(\sup_{0\leq t\leq T}|X_t|\geq k\right)} \\ &\quad+\sqrt{\mathbb{E}\left(\sup_{0\leq t\leq T}|Y_t-Y_t^{(k)}|^4\right)}\sqrt{\mathbb{P}\left(\sup_{0\leq t\leq T}|Y_t|\geq k\right)} \\ &\lesssim_T \sqrt{\mathbb{P}\left(\sup_{0\leq t\leq T}|X_t|\geq k\right)}+\sqrt{\mathbb{P}\left(\sup_{0\leq t\leq T}|Y_t|\geq k\right)}. \end{aligned} \tag{4.10}$$

By **(A1')**, we infer that

$$\begin{aligned} \sup_{0\leq s\leq t}|Y_s| &\leq |x|+K_TT+\sup_{0\leq s\leq t}|N_t|+K_T\int_0^t|Y_{s_\delta}|\mathrm{d}s \\ &\leq |x|+K_TT+\sup_{0\leq s\leq t}|N_t|+K_T\int_0^t\sup_{0\leq r\leq s}|Y_r|\mathrm{d}s \end{aligned} \tag{4.11}$$

where

$$N_t := \int_0^t \sigma_{s_\delta}(Y_{s_\delta})\mathrm{d}W_s.$$

Thus, Gronwall's inequality enables us to get that

$$\sup_{0\le s\le t}|Y_s| \le (|x|+K_T T)\mathrm{e}^{K_T T} + \mathrm{e}^{K_T T}\sup_{0\le s\le t}|N_s|.$$

For any integer $k \ge 1$ such that

$$\rho := k\mathrm{e}^{-K_T T} - |x| - K_T T > 0,$$

we derive from (4.11) that

$$\mathbb{P}\left(\sup_{0\le t\le T}|Y_t| \ge k\right) = \mathbb{P}\left(\sup_{0\le t\le T}|N_t| \ge \rho\right).$$

This, by taking advantage of [19, Proposition 6.8], yields that

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le T}|Y_t| \ge k\right) &= \mathbb{P}\left(\langle N\rangle_T \le \|\sigma\|_\infty^2 T,\ \sup_{0\le t\le T}|N_t| \ge \rho\right)\\ &\le 2n\exp\left(-\frac{\rho^2}{2n\|\sigma\|_{T,\infty}^2 T}\right),\end{aligned} \tag{4.12}$$

where $\langle N\rangle_t$ stands for the quadratic variation process of N_t. Next, by using the inequality: $(a-b)^2 \ge \frac{1}{2}a^2 - b^2$, $a, b \in \mathbb{R}$, we deduce from (4.12) that

$$\mathbb{P}\left(\sup_{0\le t\le T}|Y_t| \ge k\right) \le 2n\exp\left(\frac{(|x|+K_T T)^2}{2n\|\sigma\|_{T,\infty}^2 T}\right)\exp\left(-\frac{k^2}{4n\|\sigma\|_{T,\infty}^2 T\mathrm{e}^{2K_T T}}\right). \tag{4.13}$$

Similarly, one can obtain that

$$\mathbb{P}\left(\sup_{0\le t\le T}|X_t| \ge k\right) \le 2n\exp\left(\frac{(|x|+K_T T)^2}{2n\|\sigma\|_{T,\infty}^2 T}\right)\exp\left(-\frac{k^2}{4n\|\sigma\|_{T,\infty}^2 T\mathrm{e}^{2K_T T}}\right). \tag{4.14}$$

Inserting (4.13) and (4.14) back into (4.10) leads to

$$I_1 + I_3 \lesssim_T \exp\left(-\frac{k^2}{2n\|\sigma\|_{T,\infty}^2 T\mathrm{e}^{2K_T T}}\right).$$

This, together with (4.5), (4.7) and **(A1')**, gives that

$$\mathbb{E}\left(\sup_{0\le t\le T}|X_t - Y_t|^2\right) \le \hat{C}_T\exp\left(-\frac{k^2}{2n\|\sigma\|_{T,\infty}^2 T\mathrm{e}^{2K_T T}}\right) + \hat{C}_T\mathrm{e}^{\mathrm{e}^{\tilde{C}_T k^4}}\delta^\alpha$$

for some constants $\hat{C}_T, \tilde{C}_T > 0$. As a consequence, (1.6) follows by taking k given in (4.9). □

5 Proof of Theorem 1.4

For simplicity, for any $f : \mathbb{R}^{m_1} \to \mathbb{R}^{m_2}$, let

$$[f]_L = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|}, \quad \|f\|_\infty = \sup_{x \in \mathbb{R}^{m_1}} |f(x)|.$$

The proof of Theorem 1.4 relies on regularization properties of the following $\mathbb{R}^{2n}$-valued degenerate parabolic equation

$$\partial_t u_t^\lambda + \mathscr{L}_t^{b,\sigma} u_t^\lambda + b_t = \lambda u_t^\lambda, \quad u_T^\lambda = \mathbf{0_{2n}}, \quad t \in [0,T], \quad \lambda > 0, \tag{5.1}$$

where $\mathbf{0_{2n}}$ is the zero vector in $\mathbb{R}^{2n}$,

$$b_t := \begin{pmatrix} b_t^{(1)} \\ b_t^{(2)} \end{pmatrix} \text{ and } \mathscr{L}_t^{b,\sigma} u^\lambda := \frac{1}{2} \sum_{i,j=1}^n \langle (\sigma_t \sigma_t^*)(\cdot) e_i, e_j \rangle \nabla_{e_i}^{(2)} \nabla_{e_j}^{(2)} u^\lambda + \nabla_{b_t^{(1)}}^{(1)} u^\lambda + \nabla_{b_t^{(2)}}^{(2)} u^\lambda.$$

The following lemma on regularity estimate of solution to (5.1) is taken from [23, Theorem 3.10, (4.4)] and is an essential ingredient in analyzing numerical approximation.

Lemma 5.1 *Under* **(C1)–(C3)**, (5.1) *has a unique solution* $u^\lambda \in C([0,T]; C_b^1(\mathbb{R}^{2n}; \mathbb{R}^{2n}))$ *such that for all* $t \in [0,T]$,

$$\|\nabla u_t^\lambda\|_\infty + \|\nabla^{(2)} \nabla^{(2)} u_t^\lambda\|_\infty + [\nabla^{(2)} u_t]_L \leq C \int_0^T \mathrm{e}^{-\lambda t} \frac{\phi(t^{\frac{1}{2}})}{t} \mathrm{d}t, \tag{5.2}$$

where $C > 0$ *is a constant.*

From now on, we move forward to complete the

Proof of Theorem 1.4 For notation simplicity, set

$$X_t := \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \end{pmatrix}, \quad Y_t := \begin{pmatrix} Y_t^{(1)} \\ Y_t^{(2)} \end{pmatrix} \text{ and } \quad b_t(x) := \begin{pmatrix} b_t^{(1)}(x) \\ b_t^{(2)}(x) \end{pmatrix}, \quad x \in \mathbb{R}^{2n}.$$

Then (1.7) and (1.8) can be reformulated, respectively, as

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \begin{pmatrix} \mathbf{0_{n\times n}} \\ \sigma_t \end{pmatrix} (X_t)\mathrm{d}W_t, \quad t > 0, \quad X_0 = x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^{2n},$$

where $\mathbf{0_{n\times n}}$ is an $n \times n$ zero matrix, and

$$\mathrm{d}Y_t = b_{t_\delta}(Y_{t_\delta})\mathrm{d}t + \begin{pmatrix}\mathbf{0_{n\times n}}\\ \sigma_{t_\delta}\end{pmatrix}(Y_{t_\delta})\mathrm{d}W_t, \quad t > 0, \quad Y_0 = x \in \mathbb{R}^{2n}.$$

Note from (5.2) that there exists $\lambda_0 > 0$ sufficiently large such that for any $t \in [0, T]$,

$$\|\nabla u_t^\lambda\|_\infty + \|\nabla^{(2)}\nabla^{(2)} u_t^\lambda\|_\infty + [\nabla^{(2)} u_t^\lambda]_L \le \frac{1}{2}, \quad \lambda \ge \lambda_0. \tag{5.3}$$

Applying Itô's formula to $x + u_t^\lambda(x)$ for any $x \in \mathbb{R}^{2n}$, we deduce that

$$\begin{aligned} X_t + u_t^\lambda(X_t) &= x + u_0^\lambda(x) + \lambda\int_0^t u_s^\lambda(X_s)\mathrm{d}s + \int_0^t \begin{pmatrix}\mathbf{0_{n\times n}}\\ \sigma_s\end{pmatrix}(X_s)\mathrm{d}W_s \\ &\quad + \int_0^t \left(\nabla^{(2)}_{\sigma_s \mathrm{d}W_s} u_s^\lambda\right)(X_s), \end{aligned} \tag{5.4}$$

and that

$$\begin{aligned} Y_t + u_t^\lambda(Y_t) &= x + u_0^\lambda(x) + \lambda\int_0^t u_s^\lambda(Y_s)\mathrm{d}s \\ &\quad + \int_0^t \{\mathbf{I_{2n\times 2n}} + (\nabla u_s)(\cdot)\}(Y_s)\{b_{s_\delta}(Y_{s_\delta}) - b_s(Y_s)\}\mathrm{d}s \\ &\quad + \int_0^t \begin{pmatrix}\mathbf{0_{n\times n}}\\ \sigma_{s_\delta}\end{pmatrix}(Y_{s_\delta})\mathrm{d}W_s + \int_0^t \left(\nabla^{(2)}_{\sigma_{s_\delta}(Y_{s_\delta})\mathrm{d}W_s} u_s^\lambda\right)(Y_s) \\ &\quad + \frac{1}{2}\int_0^t \sum_{k,j=1}^n \langle\left\{\left(\sigma_{s_\delta}\sigma_{s_\delta}^*\right)(Y_{s_\delta}) - \left(\sigma_s\sigma_s^*\right)(Y_s)\right\} e_k, e_j\rangle \left(\nabla^{(2)}_{e_k}\nabla^{(2)}_{e_j} u_s^\lambda\right)(Y_s)\mathrm{d}s, \end{aligned} \tag{5.5}$$

where $\mathbf{I_{2n\times 2n}}$ is an $2n \times 2n$ identity matrix. Thus, using Hölder's inequality, Doob's submartingale inequality and Itô's isometry and taking (3.5) into consideration gives that

$$\begin{aligned} \mathbb{E}\left(\sup_{0\le s\le t} |M_s^\lambda|^2\right) \le C_{0,T}\Bigg\{ &\int_0^t \mathbb{E}|u_s^\lambda(X_s) - u_s^\lambda(Y_s)|^2\mathrm{d}s \\ &+ \left(1 + \|\nabla u^\lambda\|_{T,\infty}^2\right)\int_0^t \mathbb{E}|b_{s_\delta}(Y_s) - b_{s_\delta}(Y_{s_\delta})|^2\mathrm{d}s \\ &+ \left(1 + \|\nabla u^\lambda\|_{T,\infty}^2\right)\int_0^t \mathbb{E}|b_s(Y_s) - b_{s_\delta}(Y_s)|^2\mathrm{d}s \\ &+ \int_0^t \mathbb{E}\|\{(\nabla^{(2)} u_s^\lambda)(X_s) - \nabla^{(2)} u_s^\lambda(Y_s)\}\sigma_s(X_s)\|_{\mathrm{HS}}^2\mathrm{d}s \\ &+ \left(1 + \|\nabla^{(2)} u^\lambda\|_{T,\infty}^2\right)\int_0^t \mathbb{E}\|\sigma_{s_\delta}(X_s) - \sigma_{s_\delta}(Y_{s_\delta})\|_{\mathrm{HS}}^2\mathrm{d}s \\ &+ (1 + \|\nabla^{(2)} u^\lambda\|_{T,\infty}^2)\int_0^t \mathbb{E}\|\sigma_s(X_s) - \sigma_{s_\delta}(X_s)\|_{\mathrm{HS}}^2\mathrm{d}s \end{aligned}$$

$$
\begin{aligned}
&+\|\nabla^{(2)}\nabla^{(2)}u^\lambda\|_{T,\infty}^2\int_0^t \mathbb{E}\|\{\sigma_{s_\delta}(Y_s)-\sigma_{s_\delta}(Y_{s_\delta})\}\sigma_{s_\delta}^*(Y_{s_\delta})\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+\|\nabla^{(2)}\nabla^{(2)}u^\lambda\|_{T,\infty}^2\int_0^t \mathbb{E}\|\sigma_s(Y_s)\{\sigma_{s_\delta}^*(Y_s)-\sigma_{s_\delta}^*(Y_{s_\delta})\}\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+\|\nabla^{(2)}\nabla^{(2)}u^\lambda\|_{T,\infty}^2\int_0^t \mathbb{E}\|\sigma_s(Y_s)\{\sigma_s^*(Y_s)-\sigma_{s_\delta}^*(Y_s)\}\|_{\mathrm{HS}}^2\mathrm{d}s\\
&+\|\nabla^{(2)}\nabla^{(2)}u^\lambda\|_{T,\infty}^2\int_0^t \mathbb{E}\|\{\sigma_s(Y_s)-\sigma_{s_\delta}(Y_s)\}\sigma_{s_\delta}^*(Y_{s_\delta})\|_{\mathrm{HS}}^2\mathrm{d}s\Big\}\\
&=: C_{0,T}\left(\sum_{i=1}^{10}J_i(t)\right)
\end{aligned}
$$

for some constant $C_{0,T}>0$, where M_t^λ is defined as in (3.3). By using Hölder's inequality and the BDG inequality, **(C1)** implies that

$$
\mathbb{E}|Y_t-Y_{t_\delta}|^p\lesssim\delta^{\frac{p}{2}},\quad p\geq 1. \tag{5.6}
$$

Utilizing Taylor's expansion, one gets from (3.6), (5.3) and (5.6) that

$$
\begin{aligned}
J_1(t)+J_4(t)+J_5(t)&\lesssim\Big\{1+\|\nabla u^\lambda\|_{T,\infty}^2+\|\nabla\nabla^{(2)}u^\lambda\|_{T,\infty}^2\|\sigma\|_{T,\infty}^2\Big\}\int_0^t\mathbb{E}|X_s-Y_s|^2\mathrm{d}s\\
&\quad+\{1+\|\nabla^{(2)}u^\lambda\|_{T,\infty}^2\}\int_0^t\mathbb{E}|Y_s-Y_{s_\delta}|^2\mathrm{d}s\\
&\lesssim\delta+\int_0^t\mathbb{E}|X_s-Y_s|^2\mathrm{d}s.
\end{aligned}
$$

Next, **(C1)**, **(C5)** and (5.3) yield that

$$
J_3(t)+J_6(t)+J_9(t)+J_{10}(t)\lesssim\phi^2(\sqrt{\delta}),
$$

where we have also used that $\phi(\cdot)$ is increasing and $\delta\in(0,1)$. Additionally, by virtue of **(C1)**, **(C2)**, and (5.3), we infer from **(C3)** that

$$
\begin{aligned}
J_2(t)+J_7(t)+J_8(t)&\lesssim\delta+\int_0^t\mathbb{E}|b_{s_\delta}\left(Y_s^{(1)},Y_s^{(2)}\right)-b_{s_\delta}\left(Y_{s_\delta}^{(1)},Y_s^{(2)}\right)|^2\mathrm{d}s\\
&\quad+\int_0^t\mathbb{E}\left|b_{s_\delta}\left(Y_{s_\delta}^{(1)},Y_s^{(2)}\right)-b_{s_\delta}\left(Y_{s_\delta}^{(1)},Y_{s_\delta}^{(2)}\right)\right|^2\mathrm{d}s\\
&\leq C_{1,T}\Big\{\delta+\int_0^t\mathbb{E}\left|b_{s_\delta}^{(1)}(Y_s^{(1)},Y_s^{(2)})-b_{s_\delta}^{(1)}(Y_{s_\delta}^{(1)},Y_s^{(2)})\right|^2\mathrm{d}s\\
&\quad+\int_0^t\mathbb{E}\left|b_{s_\delta}^{(2)}(Y_s^{(1)},Y_s^{(2)})-b_{s_\delta}^{(2)}(Y_{s_\delta}^{(1)},Y_s^{(2)})\right|^2\mathrm{d}s\\
&\quad+\int_0^t\mathbb{E}\left|b_{s_\delta}^{(1)}(Y_{s_\delta}^{(1)},Y_s^{(2)})-b_{s_\delta}^{(1)}(Y_{s_\delta}^{(1)},Y_{s_\delta}^{(2)})\right|^2\mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^t \mathbb{E}\left|b^{(2)}_{s_\delta}(Y^{(1)}_{s_\delta}, Y^{(2)}_s) - b^{(2)}_{s_\delta}(Y^{(1)}_{s_\delta}, Y^{(2)}_{s_\delta})\right|^2 \mathrm{d}s\Bigg\} \\
&=: C_{1,T}\left(\delta + \sum_{i=1}^4 \Lambda_i(t)\right)
\end{aligned}
$$

for some constant $C_{1,T} > 0$. From **(C2)**, **(C3)**, (5.6) and $\phi \in \mathscr{D}^\varepsilon$, we derive from Hölder's inequality and Jensen's inequality that

$$
\begin{aligned}
\Lambda_1(t) + \Lambda_2(t) &\lesssim \sum_{i=1}^2 \int_0^t \mathbb{E}\left(\frac{\left|b^{(i)}_{s_\delta}(Y^{(1)}_s, Y^{(2)}_s) - b^{(i)}_{s_\delta}(Y^{(1)}_{s_\delta}, Y^{(2)}_s)\right|}{\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|^{\frac{2}{3}} \phi\left(\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|\right)} \mathbf{1}_{\{\mathbf{Y}^{(1)}_{\mathbf{s}} \neq \mathbf{Y}^{(1)}_{\mathbf{s}}\}}\right. \\
&\qquad \left. \times \left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|^{\frac{2}{3}} \phi\left(\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|\right)\right)^2 \mathrm{d}s \\
&\lesssim \int_0^t \mathbb{E}\left(\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|^{\frac{2}{3}} \phi\left(\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|\right)\right)^2 \mathrm{d}s \\
&\lesssim \int_0^t \left(\mathbb{E}\phi\left(\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|\right)^{2(1+\varepsilon)}\right)^{\frac{1}{1+\varepsilon}} \left(\mathbb{E}\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|^{\frac{4(1+\varepsilon)}{3\varepsilon}}\right)^{\frac{\varepsilon}{1+\varepsilon}} \mathrm{d}s \\
&\lesssim \delta^{\frac{2}{3}} \phi^2(C_{2,T}\sqrt{\delta})
\end{aligned} \tag{5.7}
$$

for some constant $C_{2,T} > 0$. With regard to the term $\Lambda_3(t)$, **(C1)** and (5.6) lead to

$$
\Lambda_3(t) \lesssim \|\nabla^{(2)} b^{(1)}\|^2_{T,\infty} \int_0^t \mathbb{E}\left|Y^{(1)}_s - Y^{(1)}_{s_\delta}\right|^2 \mathrm{d}s \lesssim \delta. \tag{5.8}
$$

Due to **(C3)**, observe from Jensen's inequality and (5.6) that

$$
\begin{aligned}
\Lambda_4(t) &\lesssim \int_0^t \mathbb{E}\left(\frac{|b^{(2)}_{s_\delta}(Y^{(1)}_{s_\delta}, Y^{(2)}_s) - b^{(2)}_{s_\delta}(Y^{(1)}_{s_\delta}, Y^{(2)}_{s_\delta})|}{\phi(|Y^{(2)}_s - Y^{(2)}_{s_\delta}|)} \mathbf{1}_{\{Y^{(2)}_s \neq Y^{(2)}_{s_\delta}\}} \times \phi(|Y^{(2)}_s - Y^{(2)}_{s_\delta}|)\right)^2 \mathrm{d}s \\
&\lesssim \int_0^t \mathbb{E}\phi(|Y^{(2)}_s - Y^{(2)}_{s_\delta}|)^2 \mathrm{d}s \\
&\lesssim \phi^2(C_{3,T}\sqrt{\delta})
\end{aligned}
$$

for some constant $C_{3,T} > 0$. Consequently, we arrive at

$$
\mathbb{E}\left(\sup_{0\le s\le t} |X_s - Y_s|^2\right) \lesssim_T \phi^2(C_{4,T}\sqrt{\delta}) + \int_0^t \mathbb{E} \sup_{0\le r\le s} |X_r - Y_r|^2 \mathrm{d}s
$$

for some constant $C_{4,T} \ge 1$. Thus, the desired assertion follows from the Gronwall inequality. □

Acknowledgements We are indebted to the referee, the associate editor and Professor Feng-Yu Wang for valuable comments and suggestions which have greatly improved our paper.

References

1. Bao, J., Yuan, C.: Convergence rate of EM scheme for SDDEs. Proc. Am. Math. Soc. **141**, 3231–3243 (2013)
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**, 1840–1872 (2016)
3. Gyöngy, I., Rásonyi, M.: A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. Stoch. Process. Appl. **121**, 2189–2200 (2011)
4. Gyöngy, I., Sabanis, S.: A note on Euler approximations for stochastic differential equations with delay. Appl. Math. Optim. **68**, 391–412 (2013)
5. Hairer, M., Hutzenthaler, M., Jentzen, A.: Loss of regularity for Kolmogorov equations. Ann. Probab. **43**, 468–527 (2015)
6. 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**, 592–609 (2007)
7. Hutzenthaler, M., Jentzen, A.: Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. Mem. Am. Math. Soc. **236**, v+99 (2015)
8. 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**, 1611–1641 (2012)
9. Jentzen, A., Müller-Gronbach, T., Yaroslavtseva, L.: On stochastic differential equations with arbitrary slow convergence rates for strong approximation. Commun. Math. Sci. **14**, 1477–1500 (2016)
10. Kaneko, H., Nakao, S.: A note on approximation for stochasitc differential equations. Sémin. Probab. Strasbg. **22**, 155–162 (1988)
11. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin (1992)
12. Kunita, H.: Stochastic differential equations and stochastic flows of diffeomorphisms. In: École d'été de probabilités de Saint-Flour, XII-1982, Lecture Notes in Math., 1097, pp. 143–303, Springer, Berlin (1984)
13. Krylov, N.V., Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. Probab. Theory Relat. Fields **131**, 154–196 (2005)
14. Leobacher, G., Szögyenyi, M.: Convergence of the Euler–Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. arXiv:1610.07047v2
15. Mattingly, J.C., Stuart, A.M., Higham, D.J.: Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. Stoch. Process. Appl. **101**, 185–232 (2002)
16. Ngo, H.-L., Taguchi, D.: Strong rate of convergence for the Euler–Maruyama approximation of stochastic differential equations with irregular coefficients. Math. Comput. **85**, 1793–1819 (2016)
17. Ngo, H.-L., Taguchi, D.: On the Euler–Maruyama approximation for one-dimensional stochastic differential equations with irregular coefficients. arXiv:1509.06532v1
18. Pamen, O.M., Taguchi, D.: Strong rate of convergence for the Euler–Maruyama approximation of SDEs with Hölder continuous drift coefficient. arXiv: 1508.07513v1
19. Shigekawa, I.: Stochastic Analysis, Translations of Mathematical Monographs, 224, Iwanami Series in Modern Mathematics. American Mathematical Society, Providence (2004)
20. Soize, C.: The Fokker–Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solutions. World Scientific Publishing Co., Inc, River Edge (1994)
21. Wang, F.-Y.: Hypercontractivity for Stochastic Hamiltonian Systems. arXiv:1409.1995
22. Wang, F.-Y.: Gradient estimates and applications for SDEs in Hilbert space with multiplicative noise and Dini continuous drift. J. Differ. Equ. **260**, 2792–2829 (2016)

23. Wang, F.-Y., Zhang, X.: Degenerate SDE with Hölder-Dini Drift and Non-Lipschitz Noise Coefficient. SIAM J. Math. Anal. **48**, 2189–2226 (2016)
24. Wang, F.-Y., Zhang, X.: Derivative formula and applications for degenerate diffusion semigroups. J. Math. Pures Appl. **99**, 726–740 (2013)
25. Yuan, C., Mao, X.: A note on the rate of convergence of the Euler–Maruyama method for stochastic differential equations. Stoch. Anal. Appl. **26**, 325–333 (2008)
26. Zhang, X.: Stochastic flows and Bismut formulas for stochastic Hamiltonian systems. Stoch. Proc. Appl. **120**, 1929–1949 (2010)
27. Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. Math. Sb. **93**, 129–149 (1974)