Stochastic differential equations with low regularity growing drifts and applications

Jinlong Wei[a], Junhao Hu[b] and Chenggui Yuan[c*]

[a]*School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China*

weijinlong.hust@gmail.com

[b]*School of Mathematics and Statistics, South-Central Minzu University Wuhan 430074, China*

junhaohu74@163.com

[c]*Department of Mathematics, Swansea University, Bay Campus Swansea SA1 8EN, United Kingdom*

C.Yuan@swansea.ac.uk

Abstract By using the Itô–Tanaka trick, we prove the unique strong solvability as well as the gradient estimates for stochastic differential equations with irregular growing drifts in low regularity Lebesgue–Hölder spaces $L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))$ with $\alpha\in(0,1)$ and $q\in(2/(1+\alpha),2)$. As applications, we show an L^2-transportation cost inequality for these stochastic differential equations first, and then establish the unique strong solvability for a class of stochastic transport equations.

Keywords: Low regularity growing drift, Unique strong solvability, Itô–Tanaka trick, Kolmogorov equation, L^2-transportation cost inequality

MSC (2020): 60H10; 60H15; 35K15

1 Introduction

Let $T>0$ be a given real number. We are concerned with the following stochastic differential equation (SDE for short) in $\mathbb{R}^d$:

$$dX_{s,t}=b(t,X_{s,t})dt+\sigma(t)dW_t,\ t\in(s,T],\ X_{s,t}|_{t=s}=x\in\mathbb{R}^d, \tag{1.1}$$

where $\{W_t\}_{0\leqslant t\leqslant T}=\{(W_{1,t},\ldots,W_{d,t})^\top\}_{0\leqslant t\leqslant T}$ is a d-dimensional standard Wiener process defined on a given stochastic basis $(\Omega,\mathcal{F},\mathbb{P},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T})$ and $b:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$, $\sigma:[0,T]\to\mathbb{R}^{d\times d}$ are Borel measurable functions.

*Corresponding author.

When b and σ are bounded, Veretennikov [40] first proved the unique strong solvability for SDE (1.1). Since then, Veretennikov's result was strengthened in different forms, see [7, 29, 44]. When $\sigma = I_{d\times d}$ and b is more regular, i.e., $b \in L^\infty(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ with $\alpha \in (0,1)$, Flandoli, Gubinelli and Priola [11] proved that the unique strong solution forms a $\mathcal{C}^{1,\alpha'}$ ($\alpha' \in (0,\alpha)$) stochastic flow of diffeomorphisms. This result was then generalized by Wei, Duan, Gao and Lv [43] to the case of $b \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ with $q > 2/\alpha$, and by Wang and Zhang [41] to a class of degenerate SDEs.

When $\sigma = I_{d\times d}$ and b is not bounded but only integrable, which is in the Krylov–Röckner class:

$$b \in L^q(0,T;L^p(\mathbb{R}^d;\mathbb{R}^d)) \tag{1.2}$$

with some $p, q \in [2,+\infty]$ such that

$$-\frac{2}{q} - \frac{d}{p} > -1, \tag{1.3}$$

the unique strong solvability for (1.1) was first obtained by Krylov and Röckner [25]. Recently, by using the Itô–Tanaka trick, Fedrizzi and Flandoli [9] proved that the unique strong solution forms a $\mathcal{C}^{\alpha'}$ ($\alpha' \in (0,1)$) stochastic flow of homeomorphisms. Some further extensions for nonconstant diffusion coefficients can be found in Zhang [49, 50], Zhang and Yuan [48]. More recently, Xia, Xie, Zhang and Zhao [47] studied the weak differentiability of the unique strong solution with respect to the starting point, and proved the Bismut–Elworthy–Li derivative formula for the strong solution.

It is known that solutions of Navier–Stokes equations can be analysed by probabilistic representations based on SDEs with irregular coefficient b, see e.g., Rezakhanlou [32], Constantin and Iyer [6]. From the viewpoint of Navier–Stokes equations b can be taken in the critical case, i.e. the greater-than sign in (1.3) is replaced by the following equal sign:

$$-\frac{2}{q} - \frac{d}{p} = -1. \tag{1.4}$$

Therefore, the study of the unique solvability for (1.1), (1.2) and (1.4) is of very high importance. When $\sigma = I_{d\times d}$ and $p < +\infty$, the unique strong solvability for the critical case has been established by Röckner and Zhao [33]. When $\sigma = I_{d\times d}$, $q = 2$ and $p = +\infty$, the local well-posedness was derived by Beck, Flandoli, Gubinelli and Maurell [3]. This result was later extended globally by Wei, Wang, Lv, and Duan [46] under the assumption that b is locally Dini continuous in the spatial variables. More recently, Krylov [23] proved strong uniqueness for Morrey drift and VMO diffusion coefficients, notably covering cases such as $\sigma = I_{d\times d}$, $q = 2$ and $p = +\infty$. However, the problem of strong existence remains open. We also refer to [16, 18, 19, 20, 21, 22, 30, 34, 45] for more details in this direction.

In the following, we shall interpret the above critical condition by using a different philosophy through the scaling transformation. Firstly, let us introduce some notions.

1.1 Lebesgue–Hölder spaces

Let $\alpha \in (0,1)$. We define the Hölder space $\mathcal{C}^\alpha(\mathbb{R}^d)$ as the set consisting of all continuous functions $h : \mathbb{R}^d \to \mathbb{R}$ for which

$$[h]_\alpha = \sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|h(x)-h(y)|}{|x-y|^\alpha} < +\infty.$$

The set $\mathcal{C}^\alpha(\mathbb{R}^d)$ becomes a Banach space with respect to the norm

$$\|h\|_{\mathcal{C}^\alpha(\mathbb{R}^d)} = \sup_{x\in\mathbb{R}^d} \frac{|h(x)|}{1+|x|^\alpha} + [h]_\alpha =: \|(1+|\cdot|^\alpha)^{-1}h(\cdot)\|_0 + [h]_\alpha.$$

We then define $\mathcal{C}_b^\alpha(\mathbb{R}^d)$ as the subset of $\mathcal{C}^\alpha(\mathbb{R}^d)$ consisting of all bounded elements, and for $h \in \mathcal{C}_b^\alpha(\mathbb{R}^d)$, we define

$$\|h\|_{\mathcal{C}_b^\alpha(\mathbb{R}^d)} = \sup_{x\in\mathbb{R}^d} |h(x)| + \sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|h(x)-h(y)|}{|x-y|^\alpha} = \|h\|_0 + [h]_\alpha.$$

For $q \in [1,+\infty]$, we denote by $L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))$ (or $L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$) the set consisting all elements belong to $L^q(0,T)$ as $\mathcal{C}^\alpha(\mathbb{R}^d)$ (or $\mathcal{C}_b^\alpha(\mathbb{R}^d)$)-valued functions. When $f \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))$ and $g \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$, we set

$$\begin{cases} \|f\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))} = \left[\int_0^T \|f(t,\cdot)\|^q_{\mathcal{C}^\alpha(\mathbb{R}^d)} dt\right]^{\frac{1}{q}}, \\ \|g\|_{L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))} = \left[\int_0^T \|g(t,\cdot)\|^q_{\mathcal{C}_b^\alpha(\mathbb{R}^d)} dt\right]^{\frac{1}{q}}, \end{cases}$$

where the integrals are interpreted as the essential supremum when $q = +\infty$. These norms also have equivalent forms. For example, when $g \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$, the equivalent norms can be given by

$$\begin{cases} \|g\|_{L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))} = \left[\|g\|^q_{q,0} + [g]^q_{q,\alpha}\right]^{\frac{1}{q}}, & \text{if } q < +\infty, \\ \|g\|_{L^\infty(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))} = \|g\|_{\infty,0} + [g]_{\infty,\alpha}, \end{cases}$$

where

$$\begin{cases} \|g\|^q_{q,0} = \int_0^T \|g(t,\cdot)\|_0^q dt \quad \text{and} \quad [g]^q_{q,\alpha} = \int_0^T [g(t,\cdot)]^q_\alpha dt, \quad \text{if } q < +\infty, \\ \|g\|_{\infty,0} = \operatorname*{esssup}_{t\in[0,T]} \|g(t,\cdot)\|_0 \quad \text{and} \quad [g]_{\infty,\alpha} = \operatorname*{esssup}_{t\in[0,T]} [g(t,\cdot)]_\alpha, \end{cases}$$

and for simplicity we write $\|g(t,\cdot)\|_0$ as $\|g(t)\|_0$ (and similarly for other terms) throughout this paper.

For $h \in L^\infty(\mathbb{R}^d)$, we define its Poisson integral

$$P_\xi h(x) = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \int_{\mathbb{R}^d} \frac{\xi h(x-z)}{(\xi^2+|z|^2)^{\frac{d+1}{2}}} dz, \quad \forall\, \xi \in \mathbb{R}_+. \tag{1.5}$$

By [36, Proposition 7, p.142], $h \in \mathcal{C}_b^\alpha(\mathbb{R}^d)$ if and only if $h \in \mathcal{C}_b(\mathbb{R}^d)$ and there exists a positive constant A such that

$$\|\partial_\xi P_\xi h\|_0 = \sup_{x\in\mathbb{R}^d} |\partial_\xi P_\xi h(x)| \leqslant A\xi^{-1+\alpha}, \quad \forall\, \xi \in \mathbb{R}_+.$$

Moreover, if $h \in \mathcal{C}_b^\alpha(\mathbb{R}^d)$, then $\|h\|_0 + \sup_{\xi>0}[\xi^{1-\alpha}\|\partial_\xi P_\xi h\|_0]$ and $\|h\|_{\mathcal{C}_b^\alpha(\mathbb{R}^d)}$ are equivalent norms. Thus, for $g \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$, we can define

$$\left\{\begin{array}{l} \|g\|_{L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))} = \Big[\|g\|_{q,0}^q + \int_0^T \sup\limits_{\xi>0} \|\xi^{1-\alpha}\partial_\xi P_\xi g(t)\|_0^q dt\Big]^{\frac{1}{q}} \\ = \Big[\|g\|_{q,0}^q + \int_0^T \sup\limits_{(\xi,x)\in\mathbb{R}_+\times\mathbb{R}^d} |\xi^{1-\alpha}\partial_\xi P_\xi g(t,x)|^q dt\Big]^{\frac{1}{q}}, \ \text{ if } \ q<+\infty, \\ \|g\|_{L^\infty(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))} = \|g\|_{\infty,0} + \operatorname*{esssup}\limits_{t\in[0,T]} \sup\limits_{(\xi,x)\in\mathbb{R}_+\times\mathbb{R}^d} |\xi^{1-\alpha}\partial_\xi P_\xi g(t,x)|. \end{array}\right. \tag{1.6}$$

1.2 Scaling transformations

Suppose $k \in \mathbb{N}_0$ and $p \in [1,+\infty]$. Let $W^{k,p}(\mathbb{R}^d)$ be the Sobolev space consisting of all locally integrable functions $h : \mathbb{R}^d \to \mathbb{R}$ such that for every $0 \leqslant i \leqslant k$, $\nabla^i h$ exists in the weak sense and belongs to $L^p(\mathbb{R}^d)$ $(:= W^{0,p}(\mathbb{R}^d))$. For $h_1 \in W^{k,p}(\mathbb{R}^d)$ and $h_2 \in \mathcal{C}^\alpha(\mathbb{R}^d), \alpha \in (0,1)$, by the scaling transformation, it yields that

$$\|\nabla^k(h_1(l\cdot))\|_{L^p(\mathbb{R}^d)} = l^{k-\frac{d}{p}}\|\nabla^k h_1\|_{L^p(\mathbb{R}^d)} \ \text{ and } \ [h_2(l\cdot)]_\alpha = l^\alpha [h_2]_\alpha, \quad \forall\, l>0.$$

Thus, the scaling indexes of h_1 and h_2 can be defined by $Index(h_1) = k-d/p$ and $Index(h_2) = \alpha$ ($Index$ by I_{nd} for short), respectively. By the localization, we also regard the scaling index of h_1 as $k-d/p$ if $h_1 \in W^{k,p}_{loc}(\mathbb{R}^d)$. Notice that for a second order parabolic equation we can 'trade' the spatial regularity against the temperal regularity at a cost of one temperal derivative for two spatial derivatives, the scaling indexes of f_1 and f_2 are given by $k-2/q-d/p$ and $\alpha-2/q$ if $f_1 \in L^q(0,T;W^{k,p}_{loc}(\mathbb{R}^d))$ and $f_2 \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))$, respectively.

From the scaling transformations, the critical condition (1.4) on the drift b for SDE (1.1) can be restated as $I_{nd}(b) = -1$. On the other hand, by the Cauchy–Lipschitz theorem, if $b \in L^1(0,T;Lip(\mathbb{R}^d;\mathbb{R}^d)) \subset L^1(0,T;W^{1,\infty}_{loc}(\mathbb{R}^d;\mathbb{R}^d))$ ($\sigma \in L^2(0,T)$ is enough), Itô [15] proved that there exists a unique strong solution to (1.1), and when $b \in L^1(0,T;Lip(\mathbb{R}^d;\mathbb{R}^d))$ we also have $I_{nd}(b) = -1$. Hence, we give a unified view for SDEs between the classical Itô theory and the modern Krylov–Röckner theory. However, from the viewpoint of Itô's theory, the drift can be taken in a low regularity Banach space for the temporal variable (such as L^1) if it has 'good' regularity in the spatial variables (such as Lipschitz continuity), and thus we could establish the unique strong solvability for (1.1) if the drift is in this low regularity space. This motivates our study, and our main result can make a theoretical bridge between Itô's framework and the Krylov–Röckner approach. Due to the technical difficulties inherent in the critical case, we restrict our analysis to the subcritical regime, i.e. $I_{nd}(b) > -1$.

Recently, for the sub-critical drift which is square-integrable in the temperal variable, and bounded and Hölder continuous in the spatial variables, Tian, Ding and Wei [39] proved the unique strong solvability for SDE (1.1) for Sobolev differentiable diffusion. More recently, Galeati and Gerencsér [13] studied (1.1) for a general fractional Brownian noise with the Hurst index $H \in \mathbb{R}_+\setminus\mathbb{N}$ in which the drift is in $L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ such that $q \in (1,2]$ and $\alpha \in (1-(q-1)/(qH),1)$, by developing some new stochastic sewing lemmas, they proved the unique strong solvability as well as some other properties for solutions, such as stability, continuous differentiability of the flow and its inverse and Malliavin differentiability. In particular, these results are true for the Wiener process ($H = 1/2$) with $b \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ for $\alpha \in (0,1)$

and $q \in (2/(1+\alpha), 2)$ ($I_{nd}(b) > -1$). However, from Itô's theory, the drift does not need to be bounded for the spatial variables, and this problem has been studied by Flandoli, Gubinelli and Priola in [12] for the time-independent case. There are relatively few works to discuss (1.1) when the drift is only q-th integrable in the temperal variable. It is still unknown whether (1.1) is well-posed or not when $b \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ for $\alpha \in (0,1)$ and $q \in (1,2)$ such that $I_{nd}(b) > -1$. In this paper, by applying Itô–Tanaka's trick and combining the regularity estimates of solutions for Kolmorogov equations, we will establish the strong well-posedness to (1.1) for a class of low regularity growing drifts. Before giving the main result, we need a definition.

Definition 1.1. *([27, p.114]) A stochastic flow of homeomorphisms on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T})$ associated to SDE (1.1) is a map $(s,t,x,\omega) \to X_{s,t}(x,\omega)$, defined for $0 \leqslant s \leqslant t \leqslant T$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ with values in $\mathbb{R}^d$, such that*

(i) the process $\{X_{s,\cdot}(x)\} = \{X_{s,t}(x),\ t \in [s,T]\}$ is a continuous $\{\mathcal{F}_{s,t}\}_{s\leqslant t\leqslant T}$-adapted solution of (1.1) for every $s \in [0,T]$ and $x \in \mathbb{R}^d$;

(ii) $\mathbb{P}$-a.s., $X_{s,t}(\cdot)$ is a homeomorphism, for all $0 \leqslant s \leqslant t \leqslant T$, and the functions $X_{s,t}(x)$ and $X_{s,t}^{-1}(x)$ are continuous in (s,t,x), where $X_{s,t}^{-1}(\cdot)$ is the inverse of $X_{s,t}(\cdot)$;

(iii) $\mathbb{P}$-a.s., $X_{s,t}(x) = X_{\tau,t}(X_{s,\tau}(x))$ for all $0 \leqslant s \leqslant \tau \leqslant t \leqslant T$, $x \in \mathbb{R}^d$ and $X_{s,s}(x) = x$.

Now, let us give our main result.

1.3 Main result

Theorem 1.2. *Let $b \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ with $\alpha \in (0,1)$ and $q \in (2/(1+\alpha), 2)$, and let $\sigma \in L^\infty(0,T;\mathbb{R}^{d\times d})$. We assume further that $(a_{i,j})_{d\times d} = (\sigma_{i,k}\sigma_{j,k})_{d\times d}$ is uniformly elliptic, i.e. for every $t \in [0,T]$, there is a constant $\Theta > 1$ such that*

$$\Theta^{-1}|\vartheta|^2 \leqslant \vartheta^\top a(t)\vartheta \leqslant \Theta|\vartheta|^2, \quad \forall\ \vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_d) \in \mathbb{R}^d. \tag{1.7}$$

Then it holds that:

*(i) (**Stochastic flow of homeomorphisms**) There exists a unique stochastic flow of homeomorphisms $\{X_{s,t}(x),\ t \in [s,T]\}$ to SDE (1.1).*

*(ii) (**Gradient and Hölder estimates**) $X_{s,t}(x)$ and $X_{s,t}^{-1}(x)$ are differentiable in x for every $0 \leqslant s \leqslant t \leqslant T$, and for every $p \geqslant 2$,*

$$\sup_{x\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\mathbb{E}\Big[\sup_{s\leqslant t\leqslant T}\|\nabla X_{s,t}(x)\|^p\Big] + \sup_{x\in\mathbb{R}^d}\sup_{0\leqslant t\leqslant T}\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\|\nabla X_{s,t}^{-1}(x)\|^p\Big] < +\infty. \tag{1.8}$$

Moreover, $\nabla X_{s,t}(x)$ and $\nabla X_{s,t}^{-1}(x)$ have continuous realizations (denoted by themselves), which are locally $(\gamma_1,\gamma_1,\gamma_2)$-Hölder continuous in (s,t,x) with $0 < \gamma_1 < (1-1/q)(1+\alpha-2/q)$ and $0 < \gamma_2 < 1+\alpha-2/q$. Furthermore, for every $\beta \in (0, 1+\alpha-2/q)$, every $p \geqslant 2$ and every $R > 0$,

$$\begin{aligned}&\sup_{0\leqslant s\leqslant T}\mathbb{E}\Big[\sup_{s\leqslant t\leqslant T}\Big(\sup_{x,y\in B_R, x\neq y}\frac{\|\nabla X_{s,t}(x) - \nabla X_{s,t}(y)\|}{|x-y|^\beta}\Big)^p\Big] \\ &+ \sup_{0\leqslant t\leqslant T}\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\Big(\sup_{x,y\in B_R, x\neq y}\frac{\|\nabla X_{s,t}^{-1}(x) - \nabla X_{s,t}^{-1}(y)\|}{|x-y|^\beta}\Big)^p\Big] < +\infty,\end{aligned} \tag{1.9}$$

here a random field $\nabla\tilde{X}_{s,t}(\cdot)$ is called a realization of $\nabla X_{s,t}(\cdot)$ if there exists $\Omega_0 \subset \Omega$ such that $\mathbb{P}(\Omega_0) = 1$ and for each $\omega \in \Omega_0$, $\nabla X_{s,t}(x,\omega) = \nabla\tilde{X}_{s,t}(x,\omega)$ for all $s \in [0,T]$, $t \in [s,T]$ and $x \in \mathbb{R}^d$.

*(iii) (**Stability**) Let ρ be a symmetric regularizing kernel, that is*

$$\rho(x) = \rho(-x) \ \text{ with } \ 0 \leqslant \rho \in \mathcal{C}_0^\infty(\mathbb{R}^d), \ \ \text{supp}(\rho) \subset B_1 \ \text{ and } \ \int_{\mathbb{R}^d} \rho(x)dx = 1. \tag{1.10}$$

For $n \in \mathbb{N}$, we set $\rho_n(x) = n^d\rho(nx)$ and

$$b_n(t,x) = \int_{\mathbb{R}^d} b(t,x-y)\rho_n(y)dy =: b * \rho_n(t,x). \tag{1.11}$$

Let X^n be the stochastic flow corresponding to the vector field b_n and $(X^n)^{-1}$ be its inverse. Then for every $p \geqslant 2$,

$$\begin{aligned} &\lim_{n\to+\infty} \sup_{x\in\mathbb{R}^d} \sup_{0\leqslant s\leqslant T} \mathbb{E}\big[\sup_{s\leqslant t\leqslant T} |X^n_{s,t}(x) - X_{s,t}(x)|^p\big] \\ &+ \lim_{n\to+\infty} \sup_{x\in\mathbb{R}^d} \sup_{s\leqslant t\leqslant T} \mathbb{E}\big[\sup_{0\leqslant s\leqslant t} |(X^n_{s,t})^{-1}(x) - X^{-1}_{s,t}(x)|^p\big] = 0. \end{aligned} \tag{1.12}$$

Remark 1.3. *By (3.111), (3.136) and the Kolmogorov–Chentsov continuity criterion, the random field $Y_{s,t}(y)$ has a continuous realization (denoted by itself), which is (β_1,β_1)-Hölder continuous in (s,t) with $\beta_1 \in (0,1/2)$ and continuously differentiable in y. Let U be the unique strong solution of the Cauchy problem (3.94). We set $\Phi(t,x) = x + U(t,x)$ and $\Psi(t,\cdot) = \Phi^{-1}(t,\cdot)$. Observe that $X_{s,t}(x) = \Psi(t, Y_{s,t}(\Phi(s,x)))$, then for every $s < s'$, $t < t'$ and a.s. $\omega \in \Omega$, by (3.100) we have*

$$\begin{aligned} |X_{s,t}(x) - X_{s',t'}(x)| \leqslant& |\Psi(t, Y_{s,t}(\Phi(s,x))) - \Psi(t', Y_{s,t}(\Phi(s,x)))| \\ &+ |\Psi(t', Y_{s,t}(\Phi(s,x))) - \Psi(t', Y_{s',t'}(\Phi(s',x)))| \\ \leqslant& C\big[|t-t'|^{1-\frac{1}{q}} + |s-s'|^{\beta_1} + |t-t'|^{\beta_1} + |\Phi(s,x) - \Phi(s',x)|\big] \\ \leqslant& C\big[|t-t'|^{1-\frac{1}{q}} + |s-s'|^{1-\frac{1}{q}}\big], \end{aligned}$$

where the last inequality follows from the Sobolev imbedding $W^{1,q}(0,T;\mathcal{C}(\mathbb{R}^d)) \hookrightarrow \mathcal{C}^{1-\frac{1}{q}}([0,T];\mathcal{C}(\mathbb{R}^d))$, with the parameter β_1 choosen in the interval $(1-1/q, 1/2)$. Therefore, $X_{s,t}(x)$ is (β_2,β_2)-Hölder continuous in (s,t) for every $\beta_2 \in (0, 1-1/q]$.

1.4 Key insights into the proof of Theorem 1.2

We now give a concise overview of the proof methodology for our main result. For simplicity, we focus on the case of $\sigma = I_{d\times d}$ and $s = 0$. We employ Itô–Tanaka's trick (or Zvonkin's transformation) to address this issue and the core of the analysis, therefore, involves obtaining suitable a priori estimates for solutions of the following backward nonhomogeneous Kolmogorov equation:

$$\partial_t U(t,x) + \frac{1}{2}\Delta U(t,x) + b(t,x)\cdot\nabla U(t,x) - \lambda U(t,x) = -b(t,x), \ \ (t,x) \in [0,T) \times \mathbb{R}^d,$$

with $U(T,x)=0$. For a sufficiently regular drift b, assume the solution U possesses adequate regularity and satisfies $\sup_{0\leqslant t\leqslant T}\|\nabla U(t)\|_0<1/2$. Then $\Phi(t,x)=x+U(t,x)$ forms a $\mathcal{C}^1$ diffeomorphism uniformly in $t\in[0,T]$. If $X_t(x)$ satisfies SDE (1.1), then applying Itô's formula to the process $Y_t(y):=X_t(x)+U(t,X_t(x))$, we find that $Y_t(y)$ satisfies

$$dY_t=\lambda U(t,\Psi(t,Y_t))dt+[I+\nabla U(t,\Psi(t,Y_t))]dW_t,\ t\in(0,T],\ Y_t|_{t=0}=y, \tag{1.13}$$

and vice versa, where $\Psi(t,\cdot)=\Phi^{-1}(t,\cdot)$. Thus, it suffices to prove conclusions for Y_t.

When $b\in L^\infty(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ with $\alpha>0$, the classical Schauder theory implies that $\nabla U\in L^\infty(0,T;\mathcal{C}_b^{1,\alpha}(\mathbb{R}^d;\mathbb{R}^{d\times d}))$. By Kunita's results (see [26, Chap. 2]), $\{Y_t(y),t\in[0,T]\}$ satisfies colclusions (i) and (ii) in Theorem 1.2. In [26], Kunita first differentiated (1.13) with respect to y, to get

$$d\nabla Y_t=\lambda\nabla U(t,\Psi(t,Y_t))\nabla\Psi(t,Y_t)\nabla Y_tdt+\nabla^2U(t,\Psi(t,Y_t))\nabla\Psi(t,Y_t)\nabla Y_tdW_t,$$

and then estimated $\|\nabla Y_t\|^p$ for every $p\geqslant 2$ by

$$\begin{aligned}&\Big\|\int_0^t\lambda\nabla U(\tau,\Psi(\tau,Y_\tau))\nabla\Psi(\tau,Y_\tau)\nabla Y_\tau d\tau+\int_0^t\nabla^2U(\tau,\Psi(\tau,Y_\tau))\nabla\Psi(\tau,Y_\tau)\nabla Y_\tau dW_\tau\Big\|^p\\&\leqslant 2^{p-1}\Big\|\int_0^t\lambda\nabla U(\tau,\Psi(\tau,Y_\tau))\nabla\Psi(\tau,Y_\tau)\nabla Y_\tau d\tau\Big\|^p\\&\quad+2^{p-1}\Big\|\int_0^t\nabla^2U(\tau,\Psi(\tau,Y_\tau))\nabla\Psi(\tau,Y_\tau)\nabla Y_\tau dW_\tau\Big\|^p.\end{aligned}$$

The calculations for the first integral are straightforward, whereas for the second one, Kunita applied the Burkholder–Davis–Gundy (BDG for short) inequality and the Hölder inequality

$$\begin{aligned}&\mathbb{E}\Big\|\int_0^t\nabla^2U(\tau,\Psi(\tau,Y_\tau))\nabla\Psi(\tau,Y_\tau)\nabla Y_\tau dW_\tau\Big\|^p\\&\leqslant C\mathbb{E}\Big[\int_0^t\|\nabla^2U(\tau)\|_0^2\|\nabla Y_\tau\|^2d\tau\Big]^{\frac{p}{2}}\leqslant Ct^{\frac{p-2}{2}}\mathbb{E}\int_0^t\|\nabla^2U(\tau)\|_0^p\|\nabla Y_\tau\|^pd\tau.\end{aligned}\tag{1.14}$$

So he imposed the condition that $\|\nabla^2U(\tau)\|_0$ is bounded with respect to τ. At the same time, by the first inequality in (1.14), $\|\nabla^2U(\tau)\|_0\in L^{\tilde p}(0,T)$ for some $\tilde p\geqslant 2$ is proper. Moreover, when $\tilde p>2$, it is enough to get the Hölder continuity of the solution in t if one modifies the estimate in (1.14) by

$$\mathbb{E}\Big[\int_0^t\|\nabla^2U(\tau)\|_0^2\|\nabla Y_\tau\|^2d\tau\Big]^{\frac{p}{2}}\leqslant t^{\frac{(\tilde p-2)p}{2\tilde p}}\mathbb{E}\Big[\int_0^t\|\nabla^2U(\tau)\|_0^{\tilde p}\|\nabla Y_\tau\|_0^{\tilde p}d\tau\Big]^{\frac{p}{\tilde p}},\quad\forall\ p\geqslant 2.$$

This is our key point. Besides, we also need the Hölder continuity of ∇^2U in the spatial variables to get the same type regularity for $\nabla Y_t(\cdot)$. In conclusion, for some $\alpha_1>0$ and $\tilde p>2$, $\nabla U\in L^{\tilde p}(0,T;\mathcal{C}_b^{1,\alpha_1}(\mathbb{R}^d;\mathbb{R}^{d\times d}))$ may be optimal. However, if $b\in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ which is stated in Theorem 1.2, the classical PDEs techniques may fail to get the ideal regularity. Indeed, if we treat the right hand side $-b\cdot\nabla U-b$ as generated by a smooth function U for the following Kolmogorov equation

$$\partial_tU(t,x)+\frac{1}{2}\Delta U(t,x)=-b(t,x)\cdot\nabla U(t,x)-b(t,x),\ (t,x)\in[0,T)\times\mathbb{R}^d,$$

with zero terminal data, the Schauder estimate yields $U \in L^\infty(\mathcal{C}^\beta)$ for $\beta = \alpha + 2 - 2/q$, leading to $b \cdot \nabla U \in L^q(\mathcal{C}^{\alpha+1-2/q})$. Since $\alpha + 1 - 2/q < \alpha$, an iterative procedure gradually degrades the spatial regularity of U, eventually rendering the product $b \cdot \nabla U$ ill-defined within finitely many steps. New techniques should be introduced.

On the other hand, if b is space independent, i.e. $b(t, x) = b(t)$, along the characteristic lines $-\int_t^T b(\tau)d\tau + constant =: c(t)$, $U(t, x + c(t))$ satisfies a standard nonhomogeneous heat equation and the desired regularity for U can be derived [17]. Thus, we may get the required estimates through a perturbative approach. This idea comes from [4, 35] directly. Precisely, if one lets x_t be a solution of $\dot{x}_t = b(t, x_0 + x_t), x_t|_{t=T} = 0$, $x_0 \in \mathbb{R}^d$, then $\hat{U}(t, x) := U(t, x + x_0 + x_t)$ satisfies

$$\partial_t \hat{U}(t, x) + \frac{1}{2}\Delta \hat{U}(t, x) = -\hat{b}(t, x) \cdot \nabla \hat{U}(t, x) - \tilde{b}(t, x),$$

where $\hat{b}(t, x) = b(t, x + x_0 + x_t) - b(t, x_0 + x_t)$ and $\tilde{b}(t, x) = b(t, x + x_0 + x_t)$.

Observe that $x_0 \in \mathbb{R}^d$ is arbitrary and

$$|\hat{b}(t, x) \cdot \nabla \hat{U}(t, x)| \leqslant [b(t)]_\alpha |x|^\alpha \|\nabla U(t)\|_0 \quad \text{and} \quad |\tilde{b}(t, x) - \tilde{b}(t, 0)| \leqslant [b(t)]_{\alpha-\theta}|x|^{\alpha-\theta},$$

the heat kernel estimate implies

$$\sup_{0\leqslant t\leqslant T} \|\nabla U(t)\|_0 \leqslant C[b]_{q,\alpha} \sup_{0\leqslant t\leqslant T} \|\nabla U(t)\|_0 + C[b]_{q,\alpha-\theta}, \tag{1.15}$$

where $\theta \in [0, 1 + \alpha - 2/q)$. If $[b]_{q,\alpha}$ is sufficiently small, one obtains the gradient estimate for U. Similarly, the $L^2(0, T; L^\infty(\mathbb{R}^d))$ estimate for $\nabla^2 U$ can be calculated as well. For general b, to overcome the difficulties arising from the non-smallness, we add a dissipation term $-\lambda U$ for large enough λ.

However, since these regularity properties are insufficient for our purposes, more refined estimates are required. We note that the Hölder norm of a Borel function can be characterised via the Poisson integral, which inherently regularises local singularities emerging in pointwise descriptions. This global approach significantly simplifies problem analysis by transforming intricate pointwise examinations into unified integral frameworks, thereby avoiding cumbersome local estimations and mitigating technical challenges associated with singularities. Combining this methodology with the a priori estimate (1.15), we derive enhanced space-time Lebesgue–Schauder type estimates. While this approach originates from [36] in the context of time-independent coefficients, our contribution generalises the Poisson characterisation from classical Hölder spaces to mixed Lebesgue–Hölder spaces. This constitutes the primary distinction from prior works, fundamentally extending the functional-analytic framework.

Finally, we note that if the Laplacian in the backward Kolmogorov equation is replaced by a general diffusion operator $\sum_{i,j=1}^d a_{i,j}(t, x)\partial^2_{x_i,x_j}$, the perturbative method cannot be directly adapted. This necessitates novel approaches for handling SDEs with nonconstant diffusion coefficients.

1.5 Applications

As the first application, we are interested in the transportation cost inequalities for SDE (1.1) with the irregular drift. The transportation cost inequality was first established by Talagrand

[37] for the Gaussian measure, and then extended by many researchers to (1.1). When b is time independent and Lipschitzian, the L^2-transportation cost inequality was first proved by Djellout, Guillin and Wu [8]. This result was then generalized by Suo, Yuan and Zhang [38] to the bounded and Dini continuous ones, and by Bahlali, Mouchtabih and Tangpi [2] to the drift which satisfies sub-critical LPS condition (1.2). However, it is still unknown whether the L^2-transportation cost inequality is true or not under the sub-critical low regularity growing drift. In this section, using Theorem 1.2 and a homeomorphic transformation, we will establish an L^2-transportation cost inequality for (1.1) with $b \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$. Before giving the result, we give some notions.

Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T})$ be the classical Wiener space, i.e. $\Omega = \mathcal{C}_0([0,T];\mathbb{R}^d) = \{\phi \in \mathcal{C}([0,T];\mathbb{R}^d); \phi(0) = 0\}$ and $\mathbb{P}$ is the Wiener measure under which the evaluation map at $t \in [0,T]$ is a d-dimensional standard Wiener process. Let $\mathscr{P}(\Omega)$ be the class of all the probability measures on Ω, and let $\mathbb{W}_\Omega$ be the L^2-Wasserstein distance on $\mathscr{P}(\Omega)$ defined by the uniform norm:

$$\mathbb{W}_\Omega^2(\mu,\nu) = \inf_{\pi\in\mathscr{C}(\mu,\nu)} \int_{\Omega\times\Omega} \|\phi_1-\phi_2\|_\Omega^2 d\pi(\phi_1,\phi_2),$$

where $\mathscr{C}(\mu,\nu)$ is the space of all couplings of μ and ν, and

$$\|\phi_1-\phi_2\|_\Omega = \sup_{0\leqslant t\leqslant T} |\phi_1(t)-\phi_2(t)|. \tag{1.16}$$

We have the following L^2-transportation cost inequality.

Theorem 1.4. *Let b be given in Theorem 1.2. Let $\{X_t(x), t\in[0,T]\}$ be the unique strong solution of SDE (1.1) with $s=0$ and let $\tilde{\mathbb{P}}$ be its law on Ω. Then there is a positive constant $C(d,T,\alpha,q,\Theta,[b]_{q,\alpha})$ such that for any $\tilde{\mathbb{Q}} \in \mathscr{P}(\Omega)$,*

$$\mathbb{W}_\Omega^2(\tilde{\mathbb{P}},\tilde{\mathbb{Q}}) \leqslant 4e^{C(d,T,\alpha,q,\Theta,[b]_{q,\alpha})} H(\tilde{\mathbb{Q}}|\tilde{\mathbb{P}}), \tag{1.17}$$

where $H(\tilde{\mathbb{Q}}|\tilde{\mathbb{P}})$ is the relative entropy (or Kullback information) of $\tilde{\mathbb{Q}}$ with respect to $\tilde{\mathbb{P}}$:

$$H(\tilde{\mathbb{Q}}|\tilde{\mathbb{P}}) = \begin{cases} \int_\Omega \log\frac{d\tilde{\mathbb{Q}}}{d\tilde{\mathbb{P}}} d\tilde{\mathbb{Q}}, & \text{if } \tilde{\mathbb{Q}} \ll \tilde{\mathbb{P}}, \\ +\infty, & \text{otherwise}. \end{cases} \tag{1.18}$$

The second application is concerned with the following stochastic transport equation with low regularity growing drift:

$$\begin{cases} \partial_t u(t,x) + b(t,x)\cdot\nabla u(t,x) + \sum\limits_{i=1}^d \partial_{x_i} u(t,x)\circ \dot{W}_{i,t} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad (t,x)\in(0,T]\times\mathbb{R}^d, \\ u(t,x)|_{t=0} = u_0(x), \quad x\in\mathbb{R}^d, \end{cases} \tag{1.19}$$

where the stochastic integral with the notation $\circ$ is interpreted in Stratonovich sense, and others are interpreted in Itô's. The choice of the Stratonovich integral in (1.19) is motivated by the mass conservation. In fact, if b is divergence free, we rewrite (1.19) by

$$\partial_t u(t,x) + \mathrm{div}[(b(t,x)+\dot{W}_t)u(t,x)] = 0, \ u(t,x)|_{t=0} = u_0(x),$$

which implies

$$\int_{\mathbb{R}^d} u(t,x)dx = \int_{\mathbb{R}^d} u_0(x)dx, \quad \forall\, t\in[0,T], \quad \mathbb{P}-a.s..$$

Firstly, we give the definition of strong solutions for the above Cauchy problem.

Definition 1.5. *Let $u_0\in W^{1,\infty}_{loc}(\mathbb{R}^d)$, $b\in L^1(0,T;L^1_{loc}(\mathbb{R}^d;\mathbb{R}^d))$. Let $u\in L^\infty(\Omega\times[0,T]\times\mathbb{R}^d)$ be a random field. We call u a stochastic strong solution of (1.19) if*

$$\sup_{0\leqslant t\leqslant T}\mathbb{E}\|\nabla u(t)\|^p_{L^\infty_{loc}(\mathbb{R}^d)}<+\infty, \quad \forall\, p\geqslant 2, \tag{1.20}$$

and for every $\varphi\in\mathcal{C}_0^\infty(\mathbb{R}^d)$, $\int_{\mathbb{R}^d}u(t,x)\varphi(x)dx$ has a continuous modification which is an $\mathcal{F}_t$-semimartingale and for every $t\in[0,T]$,

$$\begin{aligned}\int_{\mathbb{R}^d}u(t,x)\varphi(x)dx=&\int_{\mathbb{R}^d}u_0(x)\varphi(x)dx-\int_0^t d\tau\int_{\mathbb{R}^d}b(\tau,x)\cdot\nabla u(\tau,x)\varphi(x)dx\\&+\sum_{i=1}^d\int_0^t\circ dW_{i,\tau}\int_{\mathbb{R}^d}u(\tau,x)\partial_{x_i}\varphi(x)dx, \quad \mathbb{P}-a.s..\end{aligned}\tag{1.21}$$

Let us now state our second result.

Theorem 1.6. ***(Existence and uniqueness)*** *Suppose $u_0\in W^{1,\infty}(\mathbb{R}^d)$. Let b be stated in Theorem 1.2, which is divergence free. Then there exists a unique stochastic strong solution to the Cauchy problem (1.19). Moreover, the unique stochastic strong solution can be represented by $u(t,x)=u_0(X_t^{-1}(x))$, where $X_t(x)$ is the unique strong solution of the associated SDE (1.1) with $\sigma=I_{d\times d}$ and $s=0$.*

Remark 1.7. *Instead of $W^{1,\infty}$, if one requires $u_0\in\cap_{r\geqslant1}W^{1,r}(\mathbb{R}^d)$, the existence and uniqueness of $\cap_{r\geqslant1}W^{1,r}_{loc}(\mathbb{R}^d)$-solutions was also obtained by Fedrizzi and Flandoli in [10, Theorem 1] when the drift b is in the Krylov–Röckner class. After the result of [10, Theorem 1], it remains to open the question whether the solution is Lipschitz continuous (or more) when $u_0\in W^{1,\infty}(\mathbb{R}^d)$ (or more) for irregular drift.*

- *When $d=1$, the answer for the above question is positive for certain discontinuous drift b, including for instance $b(x)=sign(x)$, see [1].*
- *When $b\in L^q(0,T;\mathcal{C}^\alpha_b(\mathbb{R}^d;\mathbb{R}^d))$ with $\alpha\in(0,1)$ and $q>2/\alpha$, it is positive as well, see [43] (or see [11] if $q=+\infty$).*
- *When b is in the Krylov–Röckner class, by virtue of [31, Corollary 1.1], we also get the local quasi-Lipschitz estimate for the stochastic strong solution, i.e.*

$$\mathbb{P}\Big\{\sup_{x,y\in B_R,x\neq y}\frac{|u(t,x)-u(t,y)|}{|x-y|\exp\Big(C(d,T,p,q,R)\big(\log\frac{R}{|x-y|}\big)^\zeta\Big)}<+\infty\Big\}=1,$$

where $2\zeta=1+d/p+2/q$. However, it is still unknown whether the solution is (locally) Lipschitz continuous or not.

- *Now, the drift is in the Lebesgue–Hölder space $L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ with $\alpha\in(0,1)$ and $q\in(2/(1+\alpha),2)$, we get $\mathbb{P}(u(t)\in W^{1,\infty}_{loc}(\mathbb{R}^d))=1$ via (1.20) as well. So we give a positive answer for the above question for this low regularity growing drift.*

Notations. The letter C denotes a positive constant, whose values may change in different places. For a parameter or a function $\tilde{\zeta}$, $C(\tilde{\zeta})$ means the constant is only dependent on $\tilde{\zeta}$, and we also write it as C if there is no confusion. We use ∇ to denote the gradient of a function with respect to the spatial variables. As usual, $\mathbb{N}$ is the set of positive natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. $\mathbb{R}_+ = (0, +\infty)$ represents the set of all positive real numbers. a.s. is the abbreviation of almost surely. For every $R > 0$, $B_R := \{x \in \mathbb{R}^d : |x| \leqslant R\}$. For a given $\mathbb{R}^{n\times m}$ matrix-valued function Ξ with $n, m \in \mathbb{N}$, $\Xi^\top$ and $\|\Xi\|$ represent its transposition and Hilbert–Schmidt norm, respectively. If $n = m$, $tr(\Xi)$ stands for the trace of Ξ.

2 Lebesgue–Schauder estimates for Kolmogorov equations

Let $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ be Borel measurable functions. Consider the following Kolmogorov equation for $u : [0, T] \times \mathbb{R}^d \to \mathbb{R}$:

$$\left\{\begin{aligned} &\partial_t u(t,x) = \tfrac{1}{2}\Delta u(t,x) + b(t,x)\cdot\nabla u(t,x) \\ &\qquad\qquad -\lambda u(t,x) + f(t,x),\ (t,x) \in (0,T]\times\mathbb{R}^d, \\ &u(t,x)|_{t=0} = 0,\ x \in \mathbb{R}^d, \end{aligned}\right. \tag{2.22}$$

where $\lambda > 0$ is a real number. If $u \in L^1(0, T; \mathcal{C}^2(\mathbb{R}^d)) \cap W^{1,1}(0, T; \mathcal{C}(\mathbb{R}^d))$ such that (2.22) holds true for almost all $(t, x) \in (0, T) \times \mathbb{R}^d$, then the unknown function u is said to be a strong solution of (2.22). Before proving the strong well-posedness of (2.22), we introduce a useful lemma.

Lemma 2.1. *([14, Theorem 4.5.3]) (Hardy–Littlewood–Sobolev's convolution inequality) Let $\tilde{k} \in \mathbb{N}$, $1 < p_1 < +\infty$ and $\psi(y) = |y|^{-\tilde{k}/p_1}$. Let $1 < p_2 < p_3 < +\infty$ such that $1/p_1 + 1/p_2 = 1 + 1/p_3$. If $h \in L^{p_2}(\mathbb{R}^{\tilde{k}})$, then $h * \psi \in L^{p_3}(\mathbb{R}^{\tilde{k}})$ and there is a positive constant $C(p_1, p_2)$ such that*

$$\|h * \psi\|_{L^{p_3}(\mathbb{R}^{\tilde{k}})} \leqslant C(p_1, p_2)\|h\|_{L^{p_2}(\mathbb{R}^{\tilde{k}})}.$$

Now let us establish the strong well-posedness for (2.22).

Lemma 2.2. *Assume that $b \in L^q(0, T; \mathcal{C}^\alpha(\mathbb{R}^d; \mathbb{R}^d))$ and $f \in L^q(0, T; \mathcal{C}^\alpha(\mathbb{R}^d))$ with $\alpha \in (0, 1)$ and $q \in (2/(1 + \alpha), 2)$.*

*(i) **(Existence and uniqueness)** Then there is a unique strong solution u to (2.22), which also lies in $\mathcal{G}^{\alpha,\theta}_{q,T}$ for every $\theta \in [0, 1 + \alpha - 2/q)$, where*

$$\begin{aligned} \mathcal{G}^{\alpha,\theta}_{q,T} = \{&v \in L^\infty(0,T;\mathcal{C}^1(\mathbb{R}^d)) \cap W^{1,q}(0,T;\mathcal{C}(\mathbb{R}^d)); \nabla v \in L^\infty(0,T;\mathcal{C}^\theta_b(\mathbb{R}^d;\mathbb{R}^d)), \\ &\nabla^2 v \in L^{\frac{2q}{2-q\alpha+q\theta}}(0,T;\mathcal{C}^\theta_b(\mathbb{R}^d;\mathbb{R}^{d\times d})) \cap L^2(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^{d\times d})) \\ &\text{and } \|(1+|\cdot|^\alpha)^{-1}\partial_t v(\cdot,\cdot)\|_{q,0} < +\infty\}. \end{aligned} \tag{2.23}$$

Moreover, $\nabla u \in \mathcal{C}^{\frac{\theta}{2}}([0, T]; \mathcal{C}_b(\mathbb{R}^d; \mathbb{R}^d))$ for every $\theta \in [0, 1+\alpha-2/q)$ and there is a real number $\varepsilon > 0$ such that for large enough $\lambda > 0$,

$$\|\nabla u\|_{\infty,0} = \sup_{0\leqslant t\leqslant T} \|\nabla u(t)\|_0 \leqslant C(d, T, \alpha, q, [b]_{q,\alpha})\lambda^{-\varepsilon}[f]_{q,\alpha}. \tag{2.24}$$

(ii) ***(Stability)*** *Let ρ and b_n be given by (1.10) and (1.11), respectively. We set*

$$f_n(t,x) = (f(t,\cdot) * \rho_n)(x) = \int_{\mathbb{R}^d} f(t,x-y)\rho_n(y)dy. \tag{2.25}$$

Let u_n be the unique strong solution of (2.22) with b and f replaced by b_n and f_n, respectively. Then u_n belongs to $\mathcal{G}^{\alpha,\theta}_{q,T}$ for every $\theta \in [0, 1+\alpha-2/q)$ and satisfies (2.24) uniformly in n. Furthermore, $u_n - u \in \mathcal{C}([0,T]; \mathcal{C}^1_b(\mathbb{R}^d))$ and for every $p \geqslant 2$ we have

$$\lim_{n\to+\infty} \left[\|u_n - u\|_{\infty,0} + \|\nabla u_n - \nabla u\|_{p,0}\right] = 0. \tag{2.26}$$

Proof. Clearly, if $u \in L^1(0,T;\mathcal{C}^2(\mathbb{R}^d)) \cap W^{1,1}(0,T;\mathcal{C}(\mathbb{R}^d))$ solves the Cauchy problem (2.22) for $\lambda \geqslant \lambda_0$ with some sufficiently large real number λ_0, then for all $\lambda > 0$, $\tilde{u}(t,x) = u(t,x)e^{(\lambda_0-\lambda)t} \in L^1(0,T;\mathcal{C}^2(\mathbb{R}^d)) \cap W^{1,1}(0,T;\mathcal{C}(\mathbb{R}^d))$ solves the equation

$$\partial_t \tilde{u}(t,x) = \frac{1}{2}\Delta \tilde{u}(t,x) + b(t,x)\cdot\nabla\tilde{u}(t,x) - \lambda\tilde{u}(t,x) + \tilde{f}(t,x), \quad (t,x) \in (0,T]\times\mathbb{R}^d,$$

with zero initial data, where $\tilde{f}(t,x) = f(t,x)e^{(\lambda_0-\lambda)t}$, and vice versa. So we just need to prove the well-posedness of (2.22) for some sufficiently large λ.

On the other hand, if $u \in \mathcal{G}^{\alpha,\theta}_{q,T}$ for some $\theta \in [0, 1+\alpha-2/q)$ is a strong solution of (2.22), then it has the following equivalent representation (see [39, Lemma 2.1]):

$$\begin{aligned} u(t,x) = &\int_0^t e^{-\lambda(t-\tau)} K(t-\tau,\cdot) * (b(\tau,\cdot)\cdot\nabla u(\tau,\cdot))(x)d\tau \\ &+ \int_0^t e^{-\lambda(t-\tau)} K(t-\tau,\cdot) * f(\tau,\cdot)(x)d\tau, \end{aligned} \tag{2.27}$$

where $K(t-\tau,x) = (2\pi(t-\tau))^{-\frac{d}{2}} e^{-\frac{|x|^2}{2(t-\tau)}}$. Thus, it suffices to show the integral equation (2.27) has a unique strong solution $u \in \mathcal{G}^{\alpha,\theta}_{q,T}$ for every $\theta \in [0, 1+\alpha-2/q)$.

Firstly, let us prove the existence part. To simplify the notation, we set

$$\begin{aligned} \tilde{\mathcal{G}}^{\alpha,\theta}_{q,T} = \{&v \in L^\infty(0,T;\mathcal{C}^1(\mathbb{R}^d)); \quad \nabla v \in L^\infty(0,T;\mathcal{C}^\theta_b(\mathbb{R}^d;\mathbb{R}^d)), \\ &\nabla^2 v \in L^{\frac{2q}{2-q\alpha+q\theta}}(0,T;\mathcal{C}^\theta_b(\mathbb{R}^d;\mathbb{R}^{d\times d})) \cap L^2(0,T;\mathcal{C}^{1+\alpha-\frac{2}{q}}_b(\mathbb{R}^d;\mathbb{R}^{d\times d}))\}. \end{aligned} \tag{2.28}$$

The proof is divided into three steps.

Step 1. We assume that $b \in L^\infty(0,T;\mathcal{C}^\infty_b(\mathbb{R}^d;\mathbb{R}^d))$ and $f \in L^\infty(0,T;\mathcal{C}^\infty_b(\mathbb{R}^d))$. If $b \equiv 0$, there is a unique strong solution $u \in L^\infty(0,T;\mathcal{C}^\infty_b(\mathbb{R}^d))$ of (2.27). For $b \neq 0$, we define a mapping on $L^\infty(0,T;\mathcal{C}^\infty_b(\mathbb{R}^d))$ by

$$\begin{aligned} \mathcal{T}w(t,x) = &\int_0^t e^{-\lambda(t-\tau)} K(t-\tau,\cdot) * (b(\tau,\cdot)\cdot\nabla w(\tau,\cdot))(x)d\tau \\ &+ \int_0^t e^{-\lambda(t-\tau)} K(t-\tau,\cdot) * f(\tau,\cdot)(x)d\tau. \end{aligned} \tag{2.29}$$

From (2.29), for every fixed $k \in \mathbb{N}$ and every $w_1, w_2 \in L^\infty(0,T;\mathcal{C}_b^\infty(\mathbb{R}^d))$, then

$$
\begin{aligned}
&\|\mathcal{T}w_1 - \mathcal{T}w_2\|_{L^\infty(0,T;\mathcal{C}_b^k(\mathbb{R}^d))} \\
=&\Big\| \int_0^t e^{-\lambda(t-\tau)} K(t-\tau,\cdot) * [b(\tau,\cdot)\cdot\nabla(w_1(\tau,\cdot)-w_2(\tau,\cdot))](x)d\tau \Big\|_{\infty,0} \\
&+\sum_{i=1}^k \Big\| \int_0^t e^{-\lambda(t-\tau)} \nabla K(t-\tau,\cdot) * \\
&\quad \nabla^{i-1}[b(\tau,\cdot)\cdot\nabla(w_1(\tau,\cdot)-w_2(\tau,\cdot))](x)d\tau \Big\|_{\infty,0} \\
\leqslant& C\lambda^{-\varepsilon}\|b\|_{L^\infty(0,T;\mathcal{C}_b^{k-1}(\mathbb{R}^d;\mathbb{R}^d))}\|w_1-w_2\|_{L^\infty(0,T;\mathcal{C}_b^k(\mathbb{R}^d))},
\end{aligned}
\tag{2.30}
$$

where $\varepsilon \in (0,1/2)$.

By choosing λ big enough, the mapping $\mathcal{T}$ on $L^\infty(0,T;\mathcal{C}_b^k(\mathbb{R}^d))$ is contractive. With the aid of Banach's contraction mapping principle, there is a unique strong solution $u \in L^\infty(0,T;\mathcal{C}_b^k(\mathbb{R}^d))$ of (2.27). Since k is arbitrary, we get $u \in L^\infty(0,T;\mathcal{C}_b^\infty(\mathbb{R}^d))$.

We now use a perturbative approach, stemming from [4, 35], to obtain more refined estimates for u. Let $x_0 \in \mathbb{R}^d$. Consider the following differential equation:

$$
\dot{x}_t = -b(t, x_0 + x_t), \quad x_t|_{t=0} = 0. \tag{2.31}
$$

There exists a unique solution to (2.31) for $b \in L^\infty(0,T;\mathcal{C}_b^\infty(\mathbb{R}^d;\mathbb{R}^d))$. By setting $\hat{u}(t,x) := u(t,x+x_0+x_t)$, $\hat{b}(t,x) := b(t,x+x_0+x_t) - b(t,x_0+x_t)$ and $\hat{f}(t,x) := f(t,x+x_0+x_t)$, then

$$
\partial_t \hat{u}(t,x) = \frac{1}{2}\Delta\hat{u}(t,x) + \hat{b}(t,x)\cdot\nabla\hat{u}(t,x) - \lambda\hat{u}(t,x) + \hat{f}(t,x), \tag{2.32}
$$

with $\hat{u}(t,x)|_{t=0} = 0$, which also implies

$$
\begin{aligned}
\hat{u}(t,x) =& \int_0^t e^{-\lambda(t-\tau)}d\tau \int_{\mathbb{R}^d} K(t-\tau,x-y)[\hat{b}(\tau,y)\cdot\nabla\hat{u}(\tau,y)]dy \\
&+ \int_0^t e^{-\lambda(t-\tau)}d\tau \int_{\mathbb{R}^d} K(t-\tau,x-y)\hat{f}(\tau,y)dy.
\end{aligned}
\tag{2.33}
$$

Observe that

$$
|\hat{b}(\tau,y)\cdot\nabla\hat{u}(\tau,y)| \leqslant [b(\tau)]_\alpha |y|^\alpha \|\nabla u(\tau)\|_0 \ \text{ and } \ |\hat{f}(\tau,y)-\hat{f}(\tau,0)| \leqslant [f(\tau)]_{\alpha-\theta}|y|^{\alpha-\theta}, \tag{2.34}
$$

where $\theta \in [0, 1+\alpha-2/q]$, then $|\nabla\hat{u}(t,0)|$ can be bounded by

$$
\int_0^t e^{-\lambda(t-\tau)}d\tau \int_{\mathbb{R}^d} |\nabla K(t-\tau,y)|\big([b(\tau)]_\alpha|y|^\alpha\|\nabla u(\tau)\|_0 + [f(\tau)]_{\alpha-\theta}|y|^{\alpha-\theta}\big)dy,
$$

thereby concluding that

$$
|\nabla\hat{u}(t,0)| \leqslant C(d)\int_0^t e^{-\lambda(t-\tau)}\big\{(t-\tau)^{\frac{\alpha-1}{2}}[b(\tau)]_\alpha\|\nabla u(\tau)\|_0 + (t-\tau)^{\frac{\alpha-\theta-1}{2}}[f(\tau)]_{\alpha-\theta}\big\}d\tau. \tag{2.35}
$$

Since $x_0 \in \mathbb{R}^d$ is arbitrary, with the help of Hölder's inequality, we get for every $\theta < 1+\alpha-2/q$ that

$$\begin{aligned}
&\sup_{0\leqslant t\leqslant T}\|\nabla u(t)\|_0 \leqslant C(d)[b]_{q,\alpha}\sup_{0\leqslant \tau\leqslant T}\|\nabla u(\tau)\|_0\Big(\int_0^T e^{-\lambda q'\tau}\tau^{\frac{(\alpha-1)q'}{2}}d\tau\Big)^{\frac{1}{q'}}\\
&\quad + C(d)[f]_{q,\alpha-\theta}\Big(\int_0^T e^{-\lambda q'\tau}\tau^{\frac{(\alpha-\theta-1)q'}{2}}d\tau\Big)^{\frac{1}{q'}}\\
\leqslant & C(d)[b]_{q,\alpha}\sup_{0\leqslant \tau\leqslant T}\|\nabla u(\tau)\|_0\Big(\int_0^T e^{-\frac{2\lambda q'(1-\varepsilon_1)\tau}{2(1-\varepsilon_1)-q'(1-\alpha)}}d\tau\Big)^{\frac{2(1-\varepsilon_1)-q'(1-\alpha)}{2q'(1-\varepsilon_1)}}\\
&\times\Big(\int_0^T \tau^{\varepsilon_1-1}d\tau\Big)^{\frac{1-\alpha}{2(1-\varepsilon_1)}} + C(d)[f]_{q,\alpha-\theta}\Big(\int_0^T \tau^{\varepsilon_2-1}d\tau\Big)^{\frac{1+\theta-\alpha}{2(1-\varepsilon_2)}}\\
&\times\Big(\int_0^T e^{-\frac{2\lambda q'(1-\varepsilon_2)\tau}{2(1-\varepsilon_2)-q'(1+\theta-\alpha)}}d\tau\Big)^{\frac{2(1-\varepsilon_2)-q'(1+\theta-\alpha)}{2q'(1-\varepsilon_2)}}\\
\leqslant & C(d,T,\alpha,q)[b]_{q,\alpha}\sup_{0\leqslant \tau\leqslant T}\|\nabla u(\tau)\|_0\lambda^{-\frac{2(1-\varepsilon_1)-q'(1-\alpha)}{2q'(1-\varepsilon_1)}}\\
&\quad + C(d,T,\alpha,q,\theta)[f]_{q,\alpha-\theta}\lambda^{-\frac{2(1-\varepsilon_2)-q'(1+\theta-\alpha)}{2q'(1-\varepsilon_2)}},
\end{aligned}\tag{2.36}$$

where $q'=q/(q-1)$, $\varepsilon_1\in(0,1-q'(1-\alpha)/2)$ and $\varepsilon_2\in(0,1-q'(1+\theta-\alpha)/2)$.

We take λ large enough such that

$$C(d,T,\alpha,q)[b]_{q,\alpha}\lambda^{-\frac{2(1-\varepsilon_1)-q'(1-\alpha)}{2q'(1-\varepsilon_1)}} < \frac{1}{2}.\tag{2.37}$$

Then by (2.36)–(2.37), it yields that

$$\sup_{0\leqslant t\leqslant T}\|\nabla u(t)\|_0 \leqslant C(d,T,\alpha,q,\theta,[b]_{q,\alpha})\lambda^{-\frac{2(1-\varepsilon_2)-q'(1+\theta-\alpha)}{2q'(1-\varepsilon_2)}}[f]_{q,\alpha-\theta}.\tag{2.38}$$

In particular, if θ vanishes, there exists some $\varepsilon>0$ such that

$$\sup_{0\leqslant t\leqslant T}\|\nabla u(t)\|_0 \leqslant C(d,T,\alpha,q,[b]_{q,\alpha})\lambda^{-\varepsilon}[f]_{q,\alpha}.\tag{2.39}$$

Let

$$b^1(t)=\sup_{x\in\mathbb{R}^d}\frac{|b(t,x)|}{1+|x|^\alpha} \quad\text{and}\quad f^1(t)=\sup_{x\in\mathbb{R}^d}\frac{|f(t,x)|}{1+|x|^\alpha}.$$

Then $b^1, f^1\in L^q(0,T)$.

Combining (2.27) and (2.39), it can be deduced that

$$\begin{aligned}
&\sup_{0\leqslant t\leqslant T}\|(1+|\cdot|^\alpha)^{-1}u(t,\cdot)\|_0\\
\leqslant & \sup_{0\leqslant t\leqslant T}\sup_{x\in\mathbb{R}^d}\frac{1}{1+|x|^\alpha}\Big[\int_0^t e^{-\lambda(t-\tau)}d\tau\int_{\mathbb{R}^d}K(t-\tau,y)b^1(\tau)\big(1+|x|^\alpha+|y|^\alpha\big)\\
&\quad\times\|\nabla u\|_{\infty,0}dy+\int_0^t e^{-\lambda(t-\tau)}d\tau\int_{\mathbb{R}^d}K(t-\tau,y)f^1(\tau)\big(1+|x|^\alpha+|y|^\alpha\big)dy\Big]\\
\leqslant & C(d,T,\alpha,q,\lambda,[b]_{q,\alpha})\big([f]_{q,\alpha}\|b^1\|_q+\|f^1\|_q\big)\\
\leqslant & C(d,T,\alpha,q,\lambda,\|b\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))})\|f\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))},
\end{aligned}\tag{2.40}$$

where $\|b^1\|_q = \|b^1\|_{L^q(0,T)}$ and $\|f^1\|_q = \|f^1\|_{L^q(0,T)}$.

For $\nabla^2 u$, we estimate from (2.31) to (2.34) that

$$\begin{aligned}\|\nabla^2 u(t)\|_0 \leqslant& C(d)\int_0^t e^{-\lambda(t-\tau)}(t-\tau)^{\frac{\alpha}{2}-1}[b(\tau)]_\alpha\|\nabla u(\tau)\|_0 d\tau\\ &+C(d)\int_0^t e^{-\lambda(t-\tau)}(t-\tau)^{\frac{\alpha-\theta}{2}-1}[f(\tau)]_{\alpha-\theta}d\tau.\end{aligned} \tag{2.41}$$

Hence,

$$\begin{aligned}\|\nabla^2 u\|_{\frac{2q}{2-q\alpha+q\theta},0} \leqslant& C[b]_{q,\alpha}\sup_{0\leqslant\tau\leqslant T}\|\nabla u(\tau)\|_0 + C[f]_{q,\alpha-\theta}\\ \leqslant& C(d,T,\alpha,q,\theta,[b]_{q,\alpha})[f]_{q,\alpha-\theta},\end{aligned} \tag{2.42}$$

if one uses Lemma 2.1. In particular, when $\theta=0$, we achieve

$$\|\nabla^2 u\|_{\frac{2q}{2-q\alpha},0} \leqslant C(d,T,\alpha,q,[b]_{q,\alpha})[f]_{q,\alpha}. \tag{2.43}$$

Let P_ξ be given by (1.5). We set $v_\xi(t,x)=\partial_\xi P_\xi u(t,x)$, then

$$\begin{cases}\partial_t v_\xi(t,x) = \frac{1}{2}\Delta v_\xi(t,x) + b(t,x)\cdot\nabla v_\xi(t,x)\\ \qquad\qquad -\lambda v_\xi(t,x) + g_\xi(t,x), \quad (t,x)\in(0,T]\times\mathbb{R}^d,\\ v_\xi(t,x)|_{t=0}=0,\ x\in\mathbb{R}^d,\end{cases} \tag{2.44}$$

where

$$g_\xi(t,x)=\partial_\xi P_\xi f(t,x)+\partial_\xi P_\xi(b(t,x)\cdot\nabla u(t,x)) - b(t,x)\cdot\partial_\xi P_\xi\nabla u(t,x).$$

Owing to (2.38), for every fixed $\theta\in(0,1+\alpha-2/q)$, we conclude

$$\begin{aligned}&\sup_{0\leqslant t\leqslant T}\|\nabla v_\xi(t)\|_0\\ \leqslant& C[g_\xi]_{q,\alpha-\theta} \leqslant C\big([\partial_\xi P_\xi f]_{q,\alpha-\theta} + [\partial_\xi P_\xi(b\cdot\nabla u) - b\cdot\partial_\xi P_\xi\nabla u]_{q,\alpha-\theta}\big).\end{aligned} \tag{2.45}$$

By [4, Lemma 2.1], for every $0<\beta\leqslant\alpha<1$, there exists a positive constant $C(d,\alpha,\beta)$ such that

$$[\partial_\xi P_\xi(h_1h_2) - h_1\partial_\xi P_\xi h_2]_{\alpha-\beta} \leqslant C(d,\alpha,\beta)[h_1]_\alpha\|h_2\|_0\xi^{\beta-1},$$

if $[h_1]_\alpha$ and $\|h_2\|_0$ are finite, thereby establishing that

$$[\partial_\xi P_\xi f(t)]_{\alpha-\theta} \leqslant C(d,\alpha,\theta)[f(t)]_\alpha\xi^{\theta-1} \tag{2.46}$$

and

$$[\partial_\xi P_\xi(b(t)\cdot\nabla u(t)) - b(t)\cdot\partial_\xi P_\xi\nabla u(t)]_{\alpha-\theta} \leqslant C[b(t)]_\alpha\|\nabla u(t)\|_0\xi^{\theta-1}. \tag{2.47}$$

Combining (2.45)–(2.47) and (2.39), we thereby obtain

$$\sup_{0\leqslant t\leqslant T}\|\nabla v_\xi(t)\|_0 \leqslant C\big([f]_{q,\alpha}+[b]_{q,\alpha}\sup_{0\leqslant t\leqslant T}\|\nabla u(t)\|_0\big)\xi^{\theta-1} \leqslant C[f]_{q,\alpha}\xi^{\theta-1}. \tag{2.48}$$

By (2.39), (2.48) and (1.6), then $\nabla u\in L^\infty(0,T;\mathcal{C}_b^\theta(\mathbb{R}^d;\mathbb{R}^d))$.

Similarly, by (2.41), (2.44) and (2.46)–(2.48), we guarantee for every $\theta\in(0,1+\alpha-2/q]$ that

$$\begin{aligned}\|\nabla^2 v_\xi(t)\|_0 \leqslant& C\int_0^t e^{-\lambda(t-\tau)}(t-\tau)^{\frac{\alpha}{2}-1}[b(\tau)]_\alpha \sup_{0\leqslant\tau\leqslant T}\|\nabla v_\xi(\tau)\|_0 d\tau\\ &+C\int_0^t e^{-\lambda(t-\tau)}(t-\tau)^{\frac{\alpha-\theta}{2}-1}\big([\partial_\xi P_\xi f(\tau)]_{\alpha-\theta}\\ &+[\partial_\xi P_\xi(b(\tau)\cdot\nabla u(\tau))-b(\tau)\cdot\partial_\xi P_\xi\nabla u(\tau)]_{\alpha-\theta}\big)d\tau\\ \leqslant& C\int_0^t e^{-\lambda(t-\tau)}[(t-\tau)^{\frac{\alpha}{2}-1}+(t-\tau)^{\frac{\alpha-\theta}{2}-1}]\big([b(\tau)]_\alpha+[f(\tau)]_\alpha\big)d\tau\xi^{\theta-1}.\end{aligned}$$

On account of (2.43) and Lemma 2.1, it follows that

$$\begin{aligned}&\|\nabla^2 u\|_{L^{\frac{2q}{2-q\alpha+q\theta}}(0,T;\mathcal{C}_b^\theta(\mathbb{R}^d))}\\ =&\Big[\|\nabla^2 u\|^{\frac{2q}{2-q\alpha+q\theta}}_{\frac{2q}{2-q\alpha+q\theta},0}+\int_0^T\|\xi^{1-\theta}\nabla^2 v_\xi(t)\|^{\frac{2q}{2-q\alpha+q\theta}}_{L^\infty(\mathbb{R}_+\times\mathbb{R}^d)}dt\Big]^{\frac{2-q\alpha+q\theta}{2q}}\\ \leqslant& C(d,T,\alpha,q,\theta,\lambda,[b]_{q,\alpha})[f]_{q,\alpha}.\end{aligned}\tag{2.49}$$

Summing over (2.39), (2.41), (2.48) and (2.49), then $u\in\tilde{\mathcal{G}}^{\alpha,\theta}_{q,T}$ for every $\theta\in[0,1+\alpha-2/q)$.

Step 2. We assume that $b\in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ and $f\in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$. Firstly, we extend the time interval from $[0,T]$ to $(-\infty,T]$ and define them by $b(t,x)=b(0,x)$, $f(t,x)=f(0,x)$ if $t<0$. Let ρ be given by (1.10) and let ϱ be another nonnegative regularizing kernel

$$0\leqslant\varrho\in\mathcal{C}_0^\infty(\mathbb{R}),\quad \mathrm{supp}(\varrho)\subset[0,1]\quad\text{and}\quad\int_{\mathbb{R}}\varrho(t)dt=1.\tag{2.50}$$

For $n,m\in\mathbb{N}$, we set $\rho_n(x)=n^d\rho(nx)$ and $\varrho_m(t)=m\varrho(mt)$. We then smooth b and f by ρ_n and ϱ_m:

$$b_{n,m}(t,x)=(b(\cdot,\cdot)*\rho_n*\varrho_m)(t,x)=\int_{\mathbb{R}^{d+1}}b(t-\tau,x-y)\rho_n(y)\varrho_m(\tau)dyd\tau$$

and

$$f_{n,m}(t,x)=(f(\cdot,\cdot)*\rho_n*\varrho_m)(t,x)=\int_{\mathbb{R}^{d+1}}f(t-\tau,x-y)\rho_n(y)\varrho_m(\tau)dyd\tau.$$

Then $b_{n,m}\in L^\infty(0,T;\mathcal{C}_b^\infty(\mathbb{R}^d;\mathbb{R}^d))$ and $f_{n,m}\in L^\infty(0,T;\mathcal{C}_b^\infty(\mathbb{R}^d))$. Moreover,

$$\left\{\begin{array}{l}\|(1+|\cdot|^\alpha)^{-1}b_{n,m}\|_{q,0}\leqslant 2\|(1+|\cdot|^\alpha)^{-1}b\|_{q,0},\\ \|(1+|\cdot|^\alpha)^{-1}f_{n,m}\|_{q,0}\leqslant 2\|(1+|\cdot|^\alpha)^{-1}f\|_{q,0},\\ [b_{n,m}]_{q,\alpha}\leqslant[b]_{q,\alpha},\ \ [f_{n,m}]_{q,\alpha}\leqslant[f]_{q,\alpha}.\end{array}\right.\tag{2.51}$$

Notice that $b_{n,m}-b=b_{n,m}-b_n+b_n-b$ with $b_n=b(t,\cdot)*\rho_n(x)$, and

$$|b_n(t,x)-b(t,x)|\leqslant\int_{\mathbb{R}^d}|b(t,x-z)-b(t,x)|\rho_n(z)dz\leqslant[b(t)]_\alpha\int_{\mathbb{R}^d}|z|^\alpha\rho_n(z)dz,\tag{2.52}$$

thus $b_n - b \to 0$ in $L^q(0,T;\mathcal{C}_b(\mathbb{R}^d))$. This fact is also true for $f_n - f$. Therefore,

$$\lim_{n\to+\infty}\lim_{m\to+\infty}\Big[\|b_{n,m}-b\|_{L^q(0,T;\mathcal{C}_b^{\alpha'}(\mathbb{R}^d;\mathbb{R}^d))}+\|f_{n,m}-f\|_{L^q(0,T;\mathcal{C}_b^{\alpha'}(\mathbb{R}^d))}\Big]=0, \tag{2.53}$$

for every $\alpha' \in (0,\alpha)$.

Consider the following Kolmogorov equation:

$$\begin{cases} \partial_t u_{n,m}(t,x) = \frac{1}{2}\Delta u_{n,m}(t,x) + b_{n,m}(t,x)\cdot\nabla u_{n,m}(t,x) \\ \qquad\qquad -\lambda u_{n,m}(t,x) + f_{n,m}(t,x),\ (t,x)\in(0,T]\times\mathbb{R}^d, \\ u_{n,m}(t,x)|_{t=0}=0,\ x\in\mathbb{R}^d. \end{cases} \tag{2.54}$$

By **Step 1**, there exists a unique strong solution $u_{n,m}$ to the Cauchy problem (2.54). Moreover, $u_{n,m} \in \tilde{\mathcal{G}}^{\alpha,\theta}_{q,T}$ for every $\theta \in [0, 1+\alpha-2/q)$, and by (2.39), (2.40), (2.43), (2.48) and (2.49), for large enough $\lambda > 0$ we have

$$\begin{aligned} &\sup_{0\leqslant t\leqslant T}\|(1+|\cdot|^\alpha)^{-1}u_{n,m}(t,\cdot)\|_0 \\ \leqslant & C(d,T,\alpha,q,\lambda,\|b_{n,m}\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))})\|f_{n,m}\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))} \\ \leqslant & C(d,T,\alpha,q,\lambda,\|b\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))})\|f\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))} \end{aligned} \tag{2.55}$$

and

$$\begin{cases} \sup\limits_{0\leqslant t\leqslant T}\|\nabla u_{n,m}(t)\|_0 \leqslant C(d,T,\alpha,q,[b]_{q,\alpha})\lambda^{-\varepsilon}[f]_{q,\alpha}, \\ \sup\limits_{0\leqslant t\leqslant T}[\nabla u_{n,m}(t)]_\theta \leqslant C(d,T,\alpha,q,\theta,[b]_{q,\alpha})[f]_{q,\alpha},\ \forall\ \theta\in(0,1+\alpha-\frac{2}{q}), \\ \|\nabla^2 u_{n,m}\|_{L^{\frac{2q}{2-q\alpha+q\theta}}(0,T;\mathcal{C}_b^\theta(\mathbb{R}^d))} \leqslant C(d,T,\alpha,q,\theta,[b]_{q,\alpha})[f]_{q,\alpha}, \\ \qquad\qquad\qquad\qquad \forall\ \theta\in[0,1+\alpha-\frac{2}{q}]. \end{cases} \tag{2.56}$$

By (2.54)–(2.56), then $u_{n,m} \in \mathcal{G}^{\alpha,\theta}_{q,T}$ and there is a positive constant C which depends on $d,T,\alpha,q,\theta,\lambda$ and $\|b\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))}$ such that

$$\begin{aligned} &\|(1+|\cdot|^\alpha)^{-1}\partial_t u_{n,m}(t,\cdot)\|_{q,0} \\ \leqslant & \Big[\frac{1}{2}\|\nabla^2 u_{n,m}\|_{2,0}+\|\nabla u_{n,m}\|_{\infty,0}\|(1+|\cdot|^\alpha)^{-1}b_{n,m}(t,\cdot)\|_{q,0} \\ & +\lambda\|(1+|\cdot|^\alpha)^{-1}u_{n,m}(t,\cdot)\|_{\infty,0}+\|(1+|\cdot|^\alpha)^{-1}f_{n,m}(t,\cdot)\|_{q,0}\Big] \\ \leqslant & C\big[\|b\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))}+\|f\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))}\big]. \end{aligned} \tag{2.57}$$

On account of (2.55)–(2.57), there exists a (unlabelled) subsequence $u_{n,m}$ and a measurable function $u \in \mathcal{G}^{\alpha,\theta}_{q,T}$ with $\theta \in [0, 1+\alpha-2/q)$ such that $u_{n,m}(t,x) \to u(t,x)$ for a.e. $(t,x) \in [0,T]\times\mathbb{R}^d$ as m and n tend to infinity in turn. Furthermore, (2.55)–(2.57) hold true for u. Since $u_{n,m}$ satisfies (2.54), u satisfies (2.22) as well.

Step 3. For $b \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ and $f \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))$, we define $b_R(t,x) = b(t,x\chi_R(x))$ and $f_R(t,x) = f(t,x\chi_R(x))$, where $R>0$, $\chi_R(x)=\chi(x/R)$ and

$$\chi\in\mathcal{C}_0^\infty(\mathbb{R}^d),\quad 0\leqslant\chi\leqslant 1,\quad \chi'\leqslant 2\ \text{ and } \chi(x)=\begin{cases} 1, & \text{if } x\in B_1, \\ 0, & \text{if } x\in\mathbb{R}^d\setminus B_2. \end{cases} \tag{2.58}$$

Then $b_R \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ and $f_R \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$. Additionally,

$$\lim_{R\to+\infty}\big[|b_R(t,x)-b(t,x)|+|f_R(t,x)-f(t,x)|\big]=0,\quad \forall\ (t,x)\in[0,T]\times\mathbb{R}^d. \tag{2.59}$$

On the other hand, we have

$$\begin{aligned}[b_R]_{q,\alpha}\leqslant&[b]_{q,\alpha}\sup_{x,y\in\mathbb{R}^d,x\neq y}\frac{|x\chi_R(x)-y\chi_R(y)|^\alpha}{|x-y|^\alpha}\\ \leqslant&[b]_{q,\alpha}\sup_{x,y\in\mathbb{R}^d,x\neq y,\tau\in[0,1]}\big[\chi_R^\alpha(x)+|\chi_R'(\tau x+(1-\tau)y)|^\alpha\big]\leqslant 3[b]_{q,\alpha}\end{aligned} \tag{2.60}$$

and

$$[f_R]_{q,\alpha}\leqslant[f]_{q,\alpha}\sup_{x,y\in\mathbb{R}^d,x\neq y}\frac{|x\chi_R(x)-y\chi_R(y)|^\alpha}{|x-y|^\alpha}\leqslant 3[f]_{q,\alpha}. \tag{2.61}$$

By **Step 2**, there is a unique $u_R\in\mathcal{G}_{q,T}^{\alpha,\theta}$ with $\theta\in[0,1+\alpha-2/q)$ and $u_R(t,x)|_{t=0}=0$, which solves the following equation

$$\partial_t u_R(t,x)=\frac{1}{2}\Delta u_R(t,x)+b_R(t,x)\cdot\nabla u_R(t,x)-\lambda u_R(t,x)+f_R(t,x),\ (t,x)\in(0,T]\times\mathbb{R}^d. \tag{2.62}$$

In addition,

$$\begin{aligned}&\sup_{0\leqslant t\leqslant T}\|(1+|\cdot|^\alpha)^{-1}u_R(t,\cdot)\|_0+\|(1+|\cdot|)^{-1}\partial_t u_{n,m}(t,\cdot)\|_{q,0}\\ \leqslant&C(d,T,\alpha,q,\lambda,\|b_R\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))})\|f_R\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))}\\ \leqslant&C(d,T,\alpha,q,\lambda,\|b\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))})\|f\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))}\end{aligned} \tag{2.63}$$

and

$$\left\{\begin{aligned}&\sup_{0\leqslant t\leqslant T}\|\nabla u_R(t)\|_0\leqslant C(d,T,\alpha,q,[b]_{q,\alpha})\lambda^{-\varepsilon}[f]_{q,\alpha},\\ &\sup_{0\leqslant t\leqslant T}[\nabla u_R(t)]_\theta\leqslant C(d,T,\alpha,q,\lambda,\theta,[b]_{q,\alpha})[f]_{q,\alpha},\quad \forall\ \theta\in(0,1+\alpha-\tfrac{2}{q}),\\ &\|\nabla^2 u_R\|_{L^{\frac{2q}{2-q\alpha+q\theta}}(0,T;\mathcal{C}_b^\theta(\mathbb{R}^d))}\leqslant C(d,T,\alpha,q,\lambda,\theta,[b]_{q,\alpha})[f]_{q,\alpha},\\ &\qquad\qquad\qquad\qquad\forall\ \theta\in[0,1+\alpha-\tfrac{2}{q}].\end{aligned}\right. \tag{2.64}$$

In view of (2.63) and (2.64), by letting R tend to infinity in (2.62) we get the existence of solutions in $\mathcal{G}_{q,T}^{\alpha,\theta}$. It remains to check $\nabla u\in\mathcal{C}^{\frac{\theta}{2}}([0,T];\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^d))$ for every $\theta\in[0,1+\alpha-2/q)$.

By (2.33) for every $0\leqslant t_1<t_2\leqslant T$, then

$$\begin{aligned}&\nabla\hat{u}(t_2,0)-\nabla\hat{u}(t_1,0)\\ =&\Big[\int_0^{t_2}e^{-\lambda(t_2-\tau)}d\tau\int_{\mathbb{R}^d}\nabla K(t_2-\tau,y)[\hat{b}(\tau,y)\cdot\nabla\hat{u}(\tau,y)]dy\\ &-\int_0^{t_1}e^{-\lambda(t_1-\tau)}d\tau\int_{\mathbb{R}^d}\nabla K(t_1-\tau,y)[\hat{b}(\tau,y)\cdot\nabla\hat{u}(\tau,y)]dy\Big]\\ &+\Big[\int_0^{t_2}e^{-\lambda(t_2-\tau)}d\tau\int_{\mathbb{R}^d}\nabla K(t_2-\tau,y)[\hat{f}(\tau,y)-\hat{f}(\tau,0)]dy\\ &-\int_0^{t_1}e^{-\lambda(t_2-\tau)}d\tau\int_{\mathbb{R}^d}\nabla K(t_1-\tau,y)[\hat{f}(\tau,y)-\hat{f}(\tau,0)]dy\Big]\\ =:&\ J_1(t_1,t_2,x_0)+J_2(t_1,t_2,x_0).\end{aligned} \tag{2.65}$$

For J_1, we find

$$
\begin{aligned}
&\sup_{x_0\in\mathbb{R}^d}|J_1(t_1,t_2,x_0) = \sup_{x_0\in\mathbb{R}^d}\Big|\int_{t_1}^{t_2} e^{-\lambda(t_2-\tau)}d\tau\int_{\mathbb{R}^d}\nabla K(t_2-\tau,y)[\hat{b}(\tau,y)\cdot\nabla\hat{u}(\tau,y)]dy\\
&+\int_0^{t_1}\big[e^{-\lambda(t_2-\tau)}-e^{-\lambda(t_1-\tau)}\big]d\tau\int_{\mathbb{R}^d}\nabla K(t_2-\tau,y)[\hat{b}(\tau,y)\cdot\nabla\hat{u}(\tau,y)]dy\\
&+\int_0^{t_1} e^{-\lambda(t_1-\tau)}d\tau\int_{\mathbb{R}^d}[\nabla K(t_2-\tau,y)-\nabla K(t_1-\tau,y)][\hat{b}(\tau,y)\cdot\nabla\hat{u}(\tau,y)]dy\Big| \qquad (2.66)\\
\leqslant& C\sup_{0\leqslant t\leqslant T}\|\nabla u(\tau)\|_0\Big[\int_{t_1}^{t_2}(t_2-\tau)^{\frac{\alpha-1}{2}}[b(\tau)]_\alpha d\tau+(t_2-t_1)\int_0^{t_1}(t_2-\tau)^{\frac{\alpha-1}{2}}[b(\tau)]_\alpha d\tau\\
&+\int_0^{t_1}[b(\tau)]_\alpha d\tau\int_{\mathbb{R}^d}|\nabla K(t_2-\tau,y)-\nabla K(t_1-\tau,y)||y|^\alpha dy\Big].
\end{aligned}
$$

Thanks to Hölder's inequality,

$$
\begin{aligned}
&\int_{t_1}^{t_2}(t_2-\tau)^{\frac{\alpha-1}{2}}[b(\tau)]_\alpha d\tau+(t_2-t_1)\int_0^{t_1}(t_2-\tau)^{\frac{\alpha-1}{2}}[b(\tau)]_\alpha d\tau\\
\leqslant&[b]_{q,\alpha}\bigg\{\Big[\int_{t_1}^{t_2}(t_2-\tau)^{\frac{q(\alpha-1)}{2(q-1)}}d\tau\Big]^{\frac{q-1}{q}}+(t_2-t_1)\Big[\int_0^{t_1}(t_2-\tau)^{\frac{q(\alpha-1)}{2(q-1)}}d\tau\Big]^{\frac{q-1}{q}}\bigg\} \qquad (2.67)\\
\leqslant& C\big[(t_2-t_1)^{\frac{1+\alpha}{2}-\frac{1}{q}}+(t_2-t_1)\big]\leqslant C(t_2-t_1)^{\frac{1+\alpha}{2}-\frac{1}{q}}.
\end{aligned}
$$

Let $\epsilon_1>0$ be small enough. In view of the mean value theorem, we have the following interpolation relation

$$
\begin{aligned}
&\big|\nabla K(t_2-\tau,y)-\nabla K(t_1-\tau,y)\big|\\
\leqslant&\big[|\nabla K(t_2-\tau,y)|+|\nabla K(t_1-\tau,y)|\big]^{\frac{1-\alpha}{2}+\frac{1}{q}+\epsilon_1}\\
&\times\big|\nabla K(t_2-\tau,y)-\nabla K(t_1-\tau,y)\big|^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1}\\
\leqslant& C\Big[(t_2-\tau)^{-\frac{d+1}{2}}e^{-\frac{|y|^2}{4(t_2-\tau)}}+(t_1-\tau)^{-\frac{d+1}{2}}e^{-\frac{|y|^2}{4(t_1-\tau)}}\Big]^{\frac{1-\alpha}{2}+\frac{1}{q}+\epsilon_1}\\
&\times\Big[(t_1+\upsilon(t_2-t_1)-\tau)^{-\frac{d+3}{2}}e^{-\frac{|y|^2}{4(t_1+\upsilon(t_2-t_1)-\tau)}}\Big]^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1}\big|t_2-t_1\big|^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1} \qquad (2.68)\\
\leqslant& C\Big[(t_2-\tau)^{-\frac{d+1}{2}}e^{-\frac{|y|^2}{4(t_2-\tau)}}+(t_1-\tau)^{-\frac{d+1}{2}}e^{-\frac{|y|^2}{4(t_1-\tau)}}\Big]^{\frac{1-\alpha}{2}+\frac{1}{q}+\epsilon_1}\\
&\times|y|^{-(d+3)(\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1)}(t_2-t_1)^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1},
\end{aligned}
$$

where $\upsilon\in(0,1)$ and in the last inequality we have used the following fact

$$
\sup_{t\in\mathbb{R}_+,y\in\mathbb{R}^d}\Big(\frac{|y|^2}{t}\Big)^{\frac{d+3}{2}}e^{-\frac{|y|^2}{4t}}\leqslant C<+\infty.
$$

By (2.67) and (2.68), we derive

$$\begin{aligned}
&\int_0^{t_1}[b(\tau)]_\alpha d\tau\int_{\mathbb{R}^d}|\nabla K(t_2-\tau,y)-\nabla K(t_1-\tau,y)||y|^\alpha dy\\
\leqslant & C|t_2-t_1|^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1}\int_0^{t_1}\Big[(t_2-\tau)^{-1+\frac{1}{q}+\epsilon_1}+(t_1-\tau)^{-1+\frac{1}{q}+\epsilon_1}\Big][b(\tau)]_\alpha d\tau\\
\leqslant & C[b]_{q,\alpha}(t_2-t_1)^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1}.
\end{aligned}\tag{2.69}$$

Summing over (2.66), (2.67) and (2.68), it can be inferred that

$$\sup_{x_0\in\mathbb{R}^d}|J_1(t_1,t_2,x_0)|\leqslant C(t_2-t_1)^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1},\tag{2.70}$$

for every small enough $\epsilon_1>0$.

Analogue calculations also implies

$$\sup_{x_0\in\mathbb{R}^d}|J_2(t_1,t_2,x_0)|\leqslant C(t_2-t_1)^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1}.\tag{2.71}$$

Combining (2.70) and (2.71), we deduce

$$\sup_{x\in\mathbb{R}^d}|\nabla u(t_2,x)-\nabla u(t_1,x)|=\sup_{x_0\in\mathbb{R}^d}|\nabla\hat{u}(t_2,0)-\nabla\hat{u}(t_1,0)|\leqslant C(t_2-t_1)^{\frac{1+\alpha}{2}-\frac{1}{q}-\epsilon_1},$$

leading to the conclusion that $\nabla u\in\mathcal{C}^{\frac{\theta}{2}}([0,T];\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^d))$ for every $\theta\in[0,1+\alpha-2/q)$.

(ii) Let u_n be the unique strong solution of (2.22) with b and f replaced by b_n and f_n, respectively. Then u_n lies in $\mathcal{G}^{\alpha,\theta}_{q,T}$ for every $\theta\in[0,1+\alpha-2/q)$ and satisfies (2.24) uniformly in n. It remains to check $u_n-u\in\mathcal{C}([0,T];\mathcal{C}_b(\mathbb{R}^d))$ and (2.26) holds.

We set $v_n=u_n-u$, then v_n satisfies

$$\left\{\begin{array}{l}\partial_t v_n(t,x)=\frac{1}{2}\Delta v_n(t,x)+b_n(t,x)\cdot\nabla v_n(t,x)\\ \qquad\qquad -\lambda v_n(t,x)+F_n(t,x),\ (t,x)\in(0,T]\times\mathbb{R}^d,\\ v_n(t,x)|_{t=0}=0,\ x\in\mathbb{R}^d,\end{array}\right.\tag{2.72}$$

where $F_n(t,x)=f_n(t,x)-f(t,x)+(b_n(t,x)-b(t,x))\cdot\nabla u(t,x)$.

Let $x_0\in\mathbb{R}^d$. Consider the following differential equation:

$$\dot{x}^n_t=-b_n(t,x_0+x^n_t),\quad x^n_t|_{t=0}=0.\tag{2.73}$$

There exists a unique solution to (2.73). By setting $\hat{v}_n(t,x):=v_n(t,x+x_0+x^n_t)$, $\hat{b}_n(t,x):=b_n(t,x+x_0+x^n_t)-b_n(t,x_0+x^n_t)$ and $\hat{F}_n(t,x):=F_n(t,x+x_0+x^n_t)$, then

$$\left\{\begin{array}{l}\partial_t \hat{v}_n(t,x)=\frac{1}{2}\Delta \hat{v}_n(t,x)+\hat{b}_n(t,x)\cdot\nabla \hat{v}_n(t,x)\\ \qquad\qquad -\lambda \hat{v}_n(t,x)+\hat{F}_n(t,x),\ (t,x)\in(0,T]\times\mathbb{R}^d,\\ \hat{v}_n(t,x)|_{t=0}=0,\ x\in\mathbb{R}^d.\end{array}\right.$$

Thus,

$$\begin{aligned}\hat{v}_n(t,x) =& \int_0^t e^{-\lambda(t-\tau)} d\tau \int_{\mathbb{R}^d} K(t-\tau, x-y)[\hat{b}_n(\tau,y)\cdot\nabla\hat{v}_n(\tau,y)]dy \\ &+ \int_0^t e^{-\lambda(t-\tau)} d\tau \int_{\mathbb{R}^d} K(t-\tau,x-y)\hat{F}_n(\tau,y)dy.\end{aligned} \tag{2.74}$$

Observing that

$$|\hat{b}_n(\tau,y)\cdot\nabla\hat{v}_n(\tau,y)| \leqslant [b_n(\tau)]_\alpha|y|^\alpha\|\nabla v_n(\tau)\|_0 \leqslant [b(\tau)]_\alpha|y|^\alpha \sup_{0\leqslant\tau\leqslant T}\|\nabla v_n(\tau)\|_0 \tag{2.75}$$

and

$$\begin{aligned}|\hat{F}_n(\tau,y)| \leqslant& |f_n(\tau,y+x_0+x_\tau^n) - f(\tau,y+x_0+x_\tau^n)| \\ &+ |(b_n(\tau,y+x_0+x_\tau^n) - b(\tau,y+x_0+x_\tau^n))\cdot\nabla u(\tau,y+x_0+x_\tau^n)| \\ \leqslant& \int_{\mathbb{R}^d} |f(\tau,y+x_0+x_\tau^n-z) - f(\tau,y+x_0+x_\tau^n)|\rho_n(z)dz \\ &+ \int_{\mathbb{R}^d} |b(\tau,y+x_0+x_\tau^n-z) - b(\tau,y+x_0+x_\tau^n)|\rho_n(z)dz\|\nabla u\|_{\infty,0} \\ \leqslant& \big([f(\tau)]_\alpha + [b(\tau)]_\alpha\|\nabla u\|_{\infty,0}\big)\int_{\mathbb{R}^d}|z|^\alpha\rho_n(z)dz \\ \leqslant& C\big([f(\tau)]_\alpha + [b(\tau)]_\alpha\big)\int_{\mathbb{R}^d}|z|^\alpha\rho_n(z)dz,\end{aligned} \tag{2.76}$$

we obtain

$$\begin{aligned}\sup_{0\leqslant t\leqslant T}\|v_n(t)\|_0 =& \sup_{0\leqslant t\leqslant T}\sup_{x_0\in\mathbb{R}^d}|\hat{v}_n(t,0)| \\ &\leqslant C([b]_{q,\alpha} + [f]_{q,\alpha})\Big(\|\nabla v_n\|_{\infty,0} + \int_{\mathbb{R}^d}|z|^\alpha\rho_n(z)dz\Big),\end{aligned} \tag{2.77}$$

which suggests $v_n \in \mathcal{C}([0,T];\mathcal{C}_b(\mathbb{R}^d))$. Moreover, we get an analogue of (2.64) for v_n, and in particular

$$\sup_{n\geqslant 1}\|v_n\|_{L^\infty(0,T;\mathcal{C}_b^1(\mathbb{R}^d))} \leqslant C\|f\|_{L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))}. \tag{2.78}$$

It remains to check (2.26). By (2.78) and the interpolation inequality, it suffices to show

$$\lim_{n\to+\infty}\Big[\sup_{0\leqslant t\leqslant T}\|v_n(t)\|_0 + \|\nabla v_n\|_{p,0}\Big] = 0, \tag{2.79}$$

for some large enough λ and some fixed $p > 2$.

Let $\epsilon_2 > 0$ be sufficiently small. It follows from (2.74) that

$$\begin{aligned}|\nabla\hat{v}_n(t,0)| \leqslant& C\int_0^t e^{-\lambda(t-\tau)}(t-\tau)^{\frac{\alpha-1}{2}}[b_n(\tau)]_\alpha\|\nabla v_n(\tau)\|_0 d\tau \\ &+ C\int_0^t e^{-\lambda(t-\tau)}(t-\tau)^{\frac{\alpha-\epsilon_2-1}{2}}[f_n(\tau)-f(\tau)]_{\alpha-\epsilon_2}d\tau \\ &+ C\int_0^t e^{-\lambda(t-\tau)}(t-\tau)^{\frac{\epsilon_2}{2}+\frac{1}{q}-1}[b_n(\tau)-b(\tau)]_{\frac{2}{q}-1+\epsilon_2}[\nabla u(\tau)]_{\frac{2}{q}-1+\epsilon_2}d\tau \\ =:& \tilde{J}_1(t) + \tilde{J}_2(t) + \tilde{J}_3(t).\end{aligned} \tag{2.80}$$

Let $p > 2$, which will be chosen later. As a consequence of Hausdorff–Young's convolution inequality and Hölder's inequality, it leads to

$$\begin{aligned}\|\tilde{J}_1\|_{L^p(0,T)} \leqslant & C[b_n]_{q,\alpha}\|\nabla v_n\|_{p,0}\Big(\int_0^T e^{-\lambda q'\tau}\tau^{\frac{(\alpha-1)q'}{2}}d\tau\Big)^{\frac{1}{q'}} \\ \leqslant & C[b]_{q,\alpha}\|\nabla v_n\|_{p,0}\Big(\int_0^T e^{-\lambda q'\tau}\tau^{\frac{(\alpha-1)q'}{2}}d\tau\Big)^{\frac{1}{q'}} \leqslant \frac{1}{2}\|\nabla v_n\|_{p,0},\end{aligned} \tag{2.81}$$

where in the last inequality we have chosen λ large enough and used the fact $(1-\alpha)q' < 2$ for $1+\alpha > 2/q$.

A straightforward computation for $J_2(t)$ reveals

$$\|\tilde{J}_2\|_{L^p(0,T)} \leqslant C[f_n - f]_{q,\alpha-\epsilon_2}\Big(\int_0^T e^{-\lambda q'\tau}\tau^{\frac{(\alpha-\epsilon_2-1)q'}{2}}d\tau\Big)^{\frac{2}{q'}} \leqslant C[f_n - f]_{q,\alpha-\epsilon_2}. \tag{2.82}$$

We then use Lemma 2.1 first, Hölder's inequality next, to check $J_3(t)$ that

$$\|\tilde{J}_3\|_{L^p(0,T)} \leqslant C[b_n - b]_{q,\frac{2}{q}-1+\epsilon_2}[\nabla u]_{p,\frac{2}{q}-1+\epsilon_2}. \tag{2.83}$$

If $q \in (4/(2+\alpha), 2)$, then $1+\alpha-2/q > 2/q-1$. Thus, for small enough ϵ_2, $[\nabla u(\cdot)]_{\frac{2}{q}-1+\epsilon_2} \in L^\infty(0,T)$. Otherwise,

$$\begin{aligned}\sup_{x,y\in\mathbb{R}^d, x\neq y}\frac{|\nabla u(\cdot,x)-\nabla u(\cdot,y)|}{|x-y|^{\frac{2}{q}-1+\epsilon_2}} \leqslant & 2\|\nabla u\|_{\infty,0}^{2-\frac{2}{q}-\epsilon_2}\|\nabla^2 u(\cdot)\|_0^{\frac{2}{q}-1+\epsilon_2} \\ \in & L^{\frac{2q^2}{(2-q\alpha)(2-q+q\epsilon_2)}}(0,T),\end{aligned} \tag{2.84}$$

since $\|\nabla^2 u(\cdot)\|_0 \in L^{\frac{2q}{2-q\alpha}}(0,T)$.

Combining (2.80)–(2.84), for every fixed $q \in (2/(1+\alpha), 2)$, if one chooses $p = 2q^2/[(2-q\alpha)(2-q+q\epsilon_2)]$, then

$$\|\nabla v_n\|_{p,0} \leqslant C\big([f_n - f]_{q,\alpha-\epsilon_2} + [b_n - b]_{q,\frac{2}{q}-1+\epsilon_2}\big). \tag{2.85}$$

This, together with the interpolation inequality and the following fact

$$\begin{cases}[f_n - f]_{q,\alpha} \leqslant 2[f]_{q,\alpha}, \ [b_n - b]_{q,\alpha} \leqslant 2[b]_{q,\alpha}, \\ \|f_n - f\|_{q,0} \leqslant [f]_{q,\alpha}\int_{\mathbb{R}^d}|z|^\alpha\rho_n(z)dz, \ \|b_n - b\|_{q,0} \leqslant [b]_{q,\alpha}\int_{\mathbb{R}^d}|z|^\alpha\rho_n(z)dz,\end{cases} \tag{2.86}$$

yields that

$$\lim_{n\to+\infty}\|\nabla v_n\|_{p,0} = 0, \quad \forall\ p \geqslant 2. \tag{2.87}$$

Similar calculations used for ∇v_n is adapted to v_n, we gain

$$\sup_{0\leqslant t\leqslant T}\|v_n(t)\| \leqslant C\big(\|b_n\|_{q,0}\|\nabla v_n\|_{q',0} + \|f_n - f\|_{q,0} + [b_n - b]_{q,\frac{2}{q}-1+\epsilon_2}\big).$$

This, in conjunction with (2.86) and (2.87), indicates the estimate (2.26). $\square$

We now generalize the constant coefficients equation (2.22) to variable coefficients and establish a corresponding analogue of Lemma 2.1. To be precise, let us consider the following Kolmogorov equation:

$$\begin{cases} \partial_t u(t,x) = \frac{1}{2}\sum\limits_{i,j=1}^{d} a_{i,j}(t)\partial^2_{x_i,x_j}u(t,x) + b(t,x)\cdot\nabla u(t,x) \\ \qquad\qquad -\lambda u(t,x) + f(t,x), \quad (t,x)\in(0,T]\times\mathbb{R}^d, \\ u(t,x)|_{t=0} = 0, \ x\in\mathbb{R}^d, \end{cases} \tag{2.88}$$

where $a_{i,j}(t)$ are Borel bounded measurable functions, which satisfies condition (1.7).

For $0\leqslant s<t\leqslant T$, we set

$$A_{s,t} := \int_s^t a(\tau)d\tau, \ B_{s,t} = A_{s,t}^{-1}.$$

For every $\vartheta\in\mathbb{R}^d$, it is obvious that

$$\Theta^{-1}(t-s)|\vartheta|^2 \leqslant \vartheta^\top\mathcal{A}_{s,t}\vartheta \leqslant \Theta(t-s)|\vartheta|^2, \quad \mathcal{A}_{s,t} = A_{s,t} \ \text{ or } \ B_{s,t}.$$

Let

$$\hat{K}(s,t,x) = (2\pi)^{-\frac{d}{2}}\det(B_{s,t})^{\frac{1}{2}}\exp\Big\{-\frac{(B_{s,t}x,x)}{2}\Big\}. \tag{2.89}$$

Then for every $0\leqslant s<t\leqslant T$ and $x\in\mathbb{R}^d$, there exist positive constants $C(\Theta)$ and $\tilde{C}(\Theta)$ such that

$$|\nabla^k\hat{K}(s,t,x)| \leqslant C(t-s)^{-\frac{d+k}{2}}e^{-\frac{\tilde{C}|x-y|^2}{t-s}}, \ k=0,1,2. \tag{2.90}$$

By (2.90), all calculations used in Lemma 2.2 are applicable for (2.88). We obtain

Theorem 2.3. *Suppose $b\in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ and $f\in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))$ with $\alpha\in(0,1)$ and $q\in(2/(1+\alpha),2)$. Let $(a_{i,j})_{d\times d}$ be a symmetric $d\times d$ matrix-valued bounded function, which satisfies condition (1.7).*

(i) ***(Existence and uniqueness)*** *Then there is a unique strong solution u to (2.88), which also lies in $\mathcal{G}^{\alpha,\theta}_{q,T}$ for every $\theta\in[0,1+\alpha-2/q)$. Moreover, $\nabla u\in\mathcal{C}^{\frac{\theta}{2}}([0,T];\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^d))$ for every $\theta\in[0,1+\alpha-2/q)$ and there is a real number $\varepsilon>0$ such that for large enough $\lambda>0$,*

$$\|\nabla u\|_{\infty,0} \leqslant C(d,T,\alpha,q,\Theta,[b]_{q,\alpha})\lambda^{-\varepsilon}[f]_{q,\alpha}. \tag{2.91}$$

(ii) ***(Stability)*** *Let u_n be the unique strong solution of (2.88) with b and f replaced by $b_n = b*\rho_n$ and $f_n = f*\rho_n$, respectively. Then u_n belongs to $\mathcal{G}^{\alpha,\theta}_{q,T}$ for every $\theta\in[0,1+\alpha-2/q)$ and satisfies (2.24) uniformly in n. Furthermore, $u_n-u\in\mathcal{C}([0,T];\mathcal{C}^1_b(\mathbb{R}^d))$ and for every $p\geqslant 2$ we have*

$$\lim_{n\to+\infty}\big[\|u_n-u\|_{\infty,0} + \|\nabla u_n-\nabla u\|_{p,0}\big] = 0. \tag{2.92}$$

Remark 2.4. *Let $u \in \mathcal{G}^{\alpha,\theta}_{q,T}$ be the unique strong solution of (2.88). For every small enough $\epsilon_3 > 0$, we have*

$$\begin{aligned}
&\sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|\nabla u(t,x) - \nabla u(t,y)|}{|x-y|^{1+\alpha-\frac{2}{q}}} \\
=&\left[\sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|\nabla u(t,x) - \nabla u(t,y)|}{|x-y|^{\frac{1+\alpha-\frac{2}{q}-\epsilon_3}{1-\epsilon_3}}}\right]^{1-\epsilon_3} \left[\sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|\nabla u(t,x) - \nabla u(t,y)|}{|x-y|}\right]^{\epsilon_3}.
\end{aligned}$$

Observe that $1+\alpha-2/q-\epsilon_3 < (1-\epsilon_3)(1+\alpha-2/q)$ and $\nabla^2 u \in L^{\frac{2q}{2-q\alpha}}(0,T;\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^{d\times d}))$, then

$$\sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|\nabla u(\cdot,x) - \nabla u(\cdot,y)|}{|x-y|^{1+\alpha-\frac{2}{q}}} \leqslant [\nabla u]^{1-\epsilon_3}_{\infty,\frac{1+\alpha-\frac{2}{q}-\epsilon_3}{1-\epsilon_3}} \|\nabla^2 u(\cdot)\|_0^{\epsilon_3} \in L^{\frac{2q}{\epsilon_3(2-q\alpha)}}(0,T).$$

Therefore, $\nabla u \in \cap_{p\geqslant 2} L^p(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^d))$ and for every $p \geqslant 2$ there is a positive constant $C(d,T,\alpha,q,\lambda,p,[b]_{q,\alpha})$ such that

$$\|\nabla u\|_{L^p(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^d))} \leqslant C(d,T,\alpha,q,\lambda,p,[b]_{q,\alpha})[f]_{q,\alpha}. \tag{2.93}$$

Remark 2.5. *For the Cauchy problem (2.88), when $f \in L^q(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$ for $q \in (1,+\infty]$ and $b \in L^\infty(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$, Krylov [17] proved that it is uniquely solvable in $L^q(0,T;\mathcal{C}_b^{2+\alpha}(\mathbb{R}^d)) \cap W^{1,q}(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d))$, and when f has polynomially or exponentially growth with Hölder norms, the well-posedness was also established by Lorenzi [28]. Recently, Tian, Ding and Wei [39] generalized Krylov and Lorenzi's results to the case of $b \in L^2(0,T;\mathcal{C}_b^\alpha(\mathbb{R}^d;\mathbb{R}^d))$. When the coefficients are bounded in the temporal variable and locally α-Hölder continuous in the spatial variables, the $L^\infty(\mathcal{C}^{2+\alpha})$-Schauder estimate for solutions was derived by Krylov and Priola [24]. This result also applies to the following parabolic PDE (also see [5] for fractional PDE):*

$$\begin{cases}
\partial_t u(t,x) + \sum\limits_{i,j=1}^d a_{i,j}(t,x)\partial^2_{x_i,x_j} u(t,x) + b(t,x)\cdot\nabla u(t,x) \\
\qquad\qquad -c(t,x)u(t,x) = f(t,x), \quad (t,x) \in (T,S)\times\mathbb{R}^d, \\
u(t,x)|_{t=S} = g(x), \quad |f(t,x)| \leqslant F_0 c(t,x).
\end{cases}$$

Here, we only assume that $b \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ and $f \in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d))$ with $q \in (2/(1+\alpha),2)$, so we extend the existing results. This result plays a key role in proving the strong well-posedness for SDE (1.1).

3 Proof of Theorem 1.2

Let $\lambda > 0$ be a large enough real number. Consider the following vector-valued Cauchy problem:

$$\begin{cases}
\partial_t U(t,x) + \frac{1}{2}\sum\limits_{i,j=1}^d a_{i,j}(t)\partial^2_{x_i,x_j} U(t,x) + b(t,x)\cdot\nabla U(t,x) \\
\qquad\qquad = \lambda U(t,x) - b(t,x), \quad (t,x) \in [0,T)\times\mathbb{R}^d, \\
U(t,x)|_{t=T} = 0, \ x \in \mathbb{R}^d,
\end{cases} \tag{3.94}$$

where $(a_{i,j})_{d\times d}=\sigma\sigma^\top$. Since $b\in L^q(0,T;\mathcal{C}^\alpha(\mathbb{R}^d;\mathbb{R}^d))$ and $\sigma\in L^\infty(0,T)$, in view of Theorem 2.3, there is a unique $U\in(\mathcal{G}^{\alpha,\theta}_{q,T})^d$ (see (2.23)) solving the Cauchy problem (3.94) for every $\theta\in[0,1+\alpha-2/q)$. What is more, by (2.91) there is a real number $\varepsilon>0$ such that for large enough $\lambda>0$,

$$\sup_{0\leqslant t\leqslant T}\|\nabla U(t)\|_0\leqslant C(d,T,\alpha,q,\Theta,[b]_{q,\alpha})\lambda^{-\varepsilon}[b]_{q,\alpha}<\frac{1}{2}. \tag{3.95}$$

We set $\Phi(t,x)=x+U(t,x)$, then Φ forms a non-singular diffeomorphism of class $\mathcal{C}^{1,\theta}$ uniformly in $t\in[0,T]$ via (3.95) and

$$\frac{1}{2}<\sup_{0\leqslant t\leqslant T}\|\nabla\Phi(t)\|_0<\frac{3}{2},\quad \frac{2}{3}<\sup_{0\leqslant t\leqslant T}\|\nabla\Psi(t)\|_0<2, \tag{3.96}$$

where $\Psi(t,\cdot)=\Phi^{-1}(t,\cdot)$.

By Theorem 2.3 and Remark 2.4, one has

$$\nabla\Phi=I+\nabla U\in\bigcap_{p\geqslant 2}L^p(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^{d\times d}))\cap\bigcap_{0\leqslant\theta<1+\alpha-\frac{2}{q}}\mathcal{C}^{\frac{\theta}{2}}([0,T];\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^{d\times d})). \tag{3.97}$$

Notice that

$$\nabla\Psi(t,x)=[\nabla\Phi(t,\Psi(t,x))]^{-1}, \tag{3.98}$$

then

$$\begin{aligned}
&\sup_{x,y\in\mathbb{R}^d,x\neq y}\frac{\|\nabla\Psi(\cdot,x)-\nabla\Psi(\cdot,y)\|}{|x-y|^{1+\alpha-\frac{2}{q}}}\\
=&\sup_{x,y\in\mathbb{R}^d,x\neq y}\frac{\|[\nabla\Phi(\cdot,\Psi(\cdot,x))]^{-1}[\nabla\Phi(\cdot,\Psi(\cdot,y))-\nabla\Phi(\cdot,\Psi(\cdot,x))][\nabla\Phi(\cdot,\Psi(\cdot,y))]^{-1}\|}{|x-y|^{1+\alpha-\frac{2}{q}}}\\
\leqslant&4\sup_{x,y\in\mathbb{R}^d,x\neq y}\frac{\|\nabla\Phi(\cdot,\Psi(\cdot,y))-\nabla\Phi(\cdot,\Psi(\cdot,x))\|}{|x-y|^{1+\alpha-\frac{2}{q}}}\leqslant 8[\nabla\Phi(\cdot)]_{1+\alpha-\frac{2}{q}}\in\bigcap_{p\geqslant 2}L^p(0,T),
\end{aligned} \tag{3.99}$$

where in the last two inequalities we have used (3.96).

On the other hand, by (3.96) and Theorem 2.3, for every $0\leqslant t_1<t_2\leqslant T$,

$$\begin{aligned}
&|\Psi(t_2,x)-\Psi(t_1,x)|\leqslant 2|\Phi(t_2,\Psi(t_1,x))-\Phi(t_2,\Psi(t_2,x))|\\
=&2|\Phi(t_2,\Psi(t_1,x))-\Phi(t_1,\Psi(t_1,x))|\\
=&2|U(t_2,\Psi(t_1,x))-U(t_1,\Psi(t_1,x))|\leqslant C(1+|x|^\alpha)(t_2-t_1)^{1-\frac{1}{q}},
\end{aligned} \tag{3.100}$$

where the last inequality stems from the Sobolev imbedding $W^{1,q}(0,T)\hookrightarrow\mathcal{C}^{1-\frac{1}{q}}([0,T])$.

Similar to (3.99), one gets

$$\begin{aligned}
&\sup_{t_1,t_2\in[0,T],t_1\neq t_2}\frac{\|\nabla\Psi(t_2,x)-\nabla\Psi(t_1,x)\|}{|t_2-t_1|^{\tilde{\theta}}}\\
\leqslant&4\sup_{t_1,t_2\in[0,T],t_1\neq t_2}\frac{\|\nabla\Phi(t_2,\Psi(t_2,x))-\nabla\Phi(t_1,\Psi(t_1,x))\|}{|t_2-t_1|^{\tilde{\theta}}}\\
\leqslant&4\sup_{t_1,t_2\in[0,T],t_1\neq t_2}\sup_{x\in\mathbb{R}^d}\frac{\|\nabla\Phi(t_2,x)-\nabla\Phi(t_1,x)\|}{|t_2-t_1|^{\tilde{\theta}}}\\
&+4[\nabla\Phi]_{\infty,1+\alpha-\frac{2}{q}-\tilde{\epsilon}_1}\sup_{t_1,t_2\in[0,T],t_1\neq t_2}\frac{|\Psi(t_2,x)-\Psi(t_1,x)|^{1+\alpha-\frac{2}{q}-\tilde{\epsilon}_1}}{|t_2-t_1|^{\tilde{\theta}}},
\end{aligned}$$

which can be bounded by

$$C\Big[\sup_{t_1,t_2\in[0,T],t_1\neq t_2}\sup_{x\in\mathbb{R}^d}\frac{\|\nabla\Phi(t_2,x)-\nabla\Phi(t_1,x)\|}{|t_2-t_1|^{\tilde{\theta}}}+(1+|x|^\alpha)T^{(1+\alpha-\frac{2}{q}-\tilde{\epsilon}_1)(1-\frac{1}{q})-\tilde{\theta}}\Big]<+\infty,$$

for every $0\leqslant\tilde{\theta}<(1+\alpha-2/q)(1-1/q)$ if one chooses $\tilde{\epsilon}_1>0$ sufficiently small. Thus,

$$\nabla\Psi\in\bigcap_{0\leqslant\tilde{\theta}<(1+\alpha-\frac{2}{q})(1-\frac{1}{q})}\mathcal{C}^{\tilde{\theta}}([0,T];\mathcal{C}(\mathbb{R}^d;\mathbb{R}^{d\times d}))\tag{3.101}$$

and

$$\sup_{t_1,t_2\in[0,T],t_1\neq t_2}\frac{\|\nabla\Psi(t_2,x)-\nabla\Psi(t_1,x)\|}{|t_2-t_1|^{\tilde{\theta}}}\leqslant C(1+|x|^\alpha).\tag{3.102}$$

For $0<\tilde{\epsilon}_2<1$ and $t\in[0,T]$, define

$$U_{\tilde{\epsilon}_2}(t,x)=\frac{1}{\tilde{\epsilon}_2}\int_t^{t+\tilde{\epsilon}_2}U(\tau,x)d\tau=\int_0^1U(t+\tau\tilde{\epsilon}_2,x)d\tau$$

and $\Phi_{\tilde{\epsilon}_2}(t,x)=x+U_{\tilde{\epsilon}_2}(t,x)$, where $U(\tau,x):=U(T,x)=0$ when $\tau>T$. Then $\Phi_{\tilde{\epsilon}_2}\in\mathcal{C}^1([0,T];\mathcal{C}(\mathbb{R}^d))\cap\mathcal{C}([0,T];\mathcal{C}^2(\mathbb{R}^d))$. Let $X_{s,t}$ be a strong solution of SDE (1.1). In light of Itô's formula, one verifies that

$$\begin{aligned}&\Phi_{\tilde{\epsilon}_2}(t,X_{s,t}(x))=\Phi_{\tilde{\epsilon}_2}(s,x)+\int_s^t\partial_\tau U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x))d\tau\\&+\int_s^t b(\tau,X_{s,\tau}(x))\cdot\nabla U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x))d\tau+\int_s^t b(\tau,X_{s,\tau}(x))d\tau\\&+\frac{1}{2}\int_s^t tr(\nabla^2U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x))\sigma(\tau)\sigma^\top(\tau))d\tau+\int_s^t[I+\nabla U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x))]\sigma(\tau)dW_\tau.\end{aligned}\tag{3.103}$$

Since $U\in(\mathcal{G}^{\alpha,\theta}_{q,T})^d$, if one lets $\tilde{\epsilon}_2$ tend to 0, we obtain

$$\begin{cases}U_{\tilde{\epsilon}_2}(t,X_{s,t}(x))\longrightarrow U(t,X_{s,t}(x)), & \forall\ t\in[s,T],\ \mathbb{P}-a.s.,\\ \nabla U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x))\longrightarrow\nabla U(\tau,X_{s,\tau}(x)), & \forall\ \tau\in[s,T],\ \mathbb{P}-a.s.,\\ \partial_\tau U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x))\longrightarrow\partial_\tau U(\tau,X_{s,\tau}(x)), & a.e.\ \tau\in[s,T],\ \mathbb{P}-a.s.,\\ \nabla^2U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x))\longrightarrow\nabla^2U_{\tilde{\epsilon}_2}(\tau,X_{s,\tau}(x)), & a.e.\ \tau\in[s,T],\ \mathbb{P}-a.s..\end{cases}\tag{3.104}$$

By combining (3.103) and (3.104), along with the fact that $U(t,x)$ solves the Cauchy problem (3.94), an application of Lebesgue's dominated convergence theorem yields

$$\Phi(t,X_{s,t}(x))=\Phi(s,x)+\lambda\int_s^tU(\tau,X_{s,\tau}(x))d\tau+\int_s^t[I+\nabla U(\tau,X_{s,\tau}(x))]\sigma(\tau)dW_\tau.\tag{3.105}$$

Denote $Y_{s,t}=\Phi(t,X_{s,t})$, it follows from (3.105) that

$$\begin{cases}dY_{s,t}=\lambda U(t,\Psi(t,Y_{s,t}))dt+[I+\nabla U(t,\Psi(t,Y_{s,t}))]\sigma(t)dW_t\\ \qquad\quad=:\tilde{b}(t,Y_{s,t})dt+\tilde{\sigma}(t,Y_{s,t})dW_t,\ t\in(s,T],\\ Y_{s,t}|_{t=s}=y=\Phi(s,x).\end{cases}\tag{3.106}$$

Conversely, if $Y_{s,t}$ is a strong solution of SDE (3.106), then $X_{s,t} = \Psi(t, Y_{s,t})$ satisfies SDE (1.1). Therefore SDEs (1.1) and (3.106) are equivalent. Observing that $\partial_t U \in L^q(0,T;\mathcal{C}(\mathbb{R}^d;\mathbb{R}^d))$, one has $U \in \mathcal{C}^{1-\frac{1}{q}}([0,T];\mathcal{C}(\mathbb{R}^d;\mathbb{R}^d))$ and $\Phi \in \mathcal{C}^{1-\frac{1}{q}}([0,T];\mathcal{C}(\mathbb{R}^d;\mathbb{R}^d))$ by using the Sobolev imbedding. This, together with the fact $\Phi(t,\cdot)$ forms a non-singular diffeomorphism of class $\mathcal{C}^{1,\theta}$ uniformly in t, implies that it suffices to establish conclusions (i) and (1.8) directly for $Y_{s,t}$ and its inverse. For assertions (1.9) and (iii), the strategy proceeds in two stages: first proving them for $Y_{s,t}$ and $Y_{s,t}^{-1}$, then extending the results to $X_{s,t}$ and $X_{s,t}^{-1}$ via pushforward/pullback operations induced by Φ.

(i) We divide the proof of stochastic flow of homeomorphisms into two parts.

• **The unique strong solvability.** By the regularity of U and assumptions on σ, we have $\nabla\tilde{b} \in \mathcal{C}([0,T];\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^{d\times d}))$ and $\tilde{\sigma} \in L^{\frac{2q}{2-q\alpha}}(0,T;\mathcal{C}_b^1(\mathbb{R}^d;\mathbb{R}^{d\times d})) \cap \mathcal{C}([0,T];\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^{d\times d}))$. Owing to Cauchy–Lipschitz's theorem, there exists a unique strong solution $Y_{s,t}(y)$ to (3.106). An application of the Itô formula to $|Y_{s,t}|^p$ yields that

$$\begin{aligned} d|Y_{s,t}(y)|^p \leqslant & p|Y_{s,t}(y)|^{p-1}|\tilde{b}(t,Y_{s,t}(y))|dt + \frac{p(p-1)}{2}|Y_{s,t}(y)|^{p-2}\|\tilde{\sigma}(t,Y_{s,t}(y))\|^2 dt \\ & + p|Y_{s,t}(y)|^{p-2}\langle Y_{s,t}(y), \tilde{\sigma}(t,Y_{s,t}(y))dW_t\rangle \\ \leqslant & C[1+|Y_{s,t}(y)|^p]dt + p|Y_{s,t}(y)|^{p-2}\langle Y_{s,t}(y), \tilde{\sigma}(t,Y_{s,t}(y))dW_t\rangle. \end{aligned}$$

For every $t > s$, $\int_s^t |Y_{s,\tau}(y)|^{p-2}\langle Y_{s,\tau}(y), \tilde{\sigma}(\tau, Y_{s,\tau}(y))dW_\tau\rangle$ is a martingale. Then, as a result of Grönwall's inequality, to get

$$\sup_{s\leqslant t\leqslant T} \mathbb{E}|Y_{s,t}(y)|^p \leqslant C(1+|y|^p). \tag{3.107}$$

• **Stochastic flow of homeomorphisms.** Due to [26, Lemmas II.2.4, II.4.1 and II.4.2], we should prove that: for every $x, x', y, y' \in \mathbb{R}^d$ ($x \neq y$, $x' \neq y'$) and every $s, t, s', t' \in [0,T]$ ($s < t$, $s' < t'$),

$$\sup_{s\leqslant t\leqslant T} \mathbb{E}|Y_{s,t}(x) - Y_{s,t}(y)|^{2\varsigma} \leqslant C|x-y|^{2\varsigma}, \quad \forall\, \varsigma < 0, \tag{3.108}$$

and

$$\begin{aligned} & \mathbb{E}|\eta_{s,t}(x,x') - \eta_{s',t'}(y,y')|^p \\ \leqslant & C\tilde{\delta}^{-2p}\Big\{\big(1+|x|^p+|x'|^p+|y|^p+|y'|^p\big)\big[|s-s'|^{\frac{p}{2}} + |t-t'|^{\frac{p}{2}}\big] \\ & + |x-y|^p + |x'-y'|^p\Big\}, \ \forall\, p \geqslant 2, \quad |x-x'| \geqslant \tilde{\delta} > 0, \quad |y-y'| \geqslant \tilde{\delta} > 0, \end{aligned} \tag{3.109}$$

and

$$\mathbb{E}|\eta_{s,t}(\hat{x}) - \eta_{s',t'}(\hat{y})|^p \leqslant C\big[|\hat{x}-\hat{y}|^p + |s-s'|^{\frac{p}{2}} + |t-t'|^{\frac{p}{2}}\big], \quad \forall\, p > 0, \tag{3.110}$$

where

$$\eta_{s,t}(x,x') = \frac{1}{|Y_{s,t}(x) - Y_{s,t}(x')|} \quad \text{and} \quad \eta_{s,t}(\hat{x}) = \begin{cases} \frac{1}{1+|Y_{s,t}(x)|}, & \text{if } \hat{x} = |x|^{-2}x \in \mathbb{R}^d, \\ 0, & \text{if } \hat{x} = |x|^{-2}x = \infty. \end{cases}$$

For every $p \geqslant 2$,

$$\begin{aligned} & |\eta_{s,t}(x,x') - \eta_{s',t'}(y,y')|^p \\ \leqslant & 2^{p-1}|\eta_{s,t}(x,y)|^p|\eta_{s',t'}(x',y')|^p\big[|Y_{s,t}(x) - Y_{s',t'}(y)|^p + |Y_{s,t}(x') - Y_{s',t'}(y')|^p\big] \end{aligned}$$

and

$$|\eta_{s,t}(\hat{x})-\eta_{s',t'}(\hat{y})|^p\leqslant|\eta_{s,t}(\hat{x})|^p|\eta_{s',t'}(\hat{y})|^p|Y_{s,t}(x)-Y_{s',t'}(y)|^p.$$

These, together with the element inequality $|a_1+b_1|^{\tilde{p}}\leqslant\max\{2^{\tilde{p}-1},1\}[a_1^{\tilde{p}}+b_1^{\tilde{p}}]$ $(a_1,b_1,\tilde{p}\in\mathbb{R}_+)$, Hölder's inequality, (3.102), and the estimate

$$\mathbb{E}\sup_{s\leqslant t\leqslant T}(1+|Y_{s,t}(y)|)^{\varsigma}\leqslant C(1+|y|)^{\varsigma},\quad\forall\,\varsigma<0,$$

which can be proved clearly since $\tilde{b}$ is Lipschitz continuous and $\tilde{\sigma}$ is bounded, implies the estimates (3.109) and (3.110) if one shows the following inequality

$$\mathbb{E}|Y_{s,t}(x)-Y_{s',t'}(y)|^p\leqslant\Big\{|x-y|^p+(1+|y|^p)\big[|s-s'|^{\frac{p}{2}}+|t-t'|^{\frac{p}{2}}\big]\Big\},\ \forall\ p\geqslant 2.\qquad(3.111)$$

Thus, it suffices to show (3.108) and (3.111).

Let us prove (3.108) first. For $\tilde{\epsilon}_2>0$, we choose $F_{\tilde{\epsilon}_2}(x)=f^{\varsigma}_{\tilde{\epsilon}_2}(x)=(\tilde{\epsilon}_2+|x|^2)^{\varsigma}$ and $Y_{s,t}(x,y):=Y_{s,t}(x)-Y_{s,t}(y)$. Thanks to the Itô formula, then

$$\begin{aligned}F_{\tilde{\epsilon}_2}(Y_{s,t}(x,y))\leqslant&F_{\tilde{\epsilon}_2}(x-y)+C|\varsigma|\int_s^t F_{\tilde{\epsilon}_2}(Y_{s,\tau}(x,y))d\tau\\&+C|\varsigma(\varsigma-1)|\int_s^t\kappa^2(\tau)F_{\tilde{\epsilon}_2}(Y_{s,\tau}(x,y))d\tau\\&+2\varsigma\int_s^t f^{\varsigma-1}_{\tilde{\epsilon}_2}(Y_{s,\tau}(x,y))\langle Y_{s,\tau}(x,y),(\tilde{\sigma}(\tau,Y_{s,\tau}(x))-\tilde{\sigma}(\tau,Y_{s,\tau}(y)))dW_\tau\rangle,\end{aligned}\qquad(3.112)$$

where $\kappa(\tau)=\|\nabla^2U(\tau)\|_0\in L^{\frac{2q}{2-q\alpha}}(0,T)$ for $\nabla U\in L^{\frac{2q}{2-q\alpha}}(0,T;\mathcal{C}^1_b(\mathbb{R}^d;\mathbb{R}^{d\times d}))$ and

$$Y_{s,\tau}=(Y^1_{s,\tau},Y^2_{s,\tau},\ldots,Y^d_{s,\tau}),\quad\delta_{i,j}=\left\{\begin{array}{ll}1,&\text{if}\ \ i=j,\\0,&\text{if}\ \ i\neq j.\end{array}\right.$$

By (3.112) and the Grönwall inequality, we derive

$$\sup_{s\leqslant t\leqslant T}\mathbb{E}\big[\tilde{\epsilon}_2+|Y_{s,t}(x)-Y_{s,t}(y)|^2\big]^{\varsigma}\leqslant C\big[\tilde{\epsilon}_2+|x-y|^2\big]^{\varsigma}.$$

By letting $\tilde{\epsilon}_2\downarrow 0$, then (3.108) holds.

Let $x,y,\in\mathbb{R}^d$ $(x\neq y)$ and $s,t,s',t'\in[0,T]$ $(s<t,\ s'<t')$. Without loss of generality, we assume $s<s'<t<t'$. For every $p\geqslant 2$, then

$$\begin{aligned}&|Y_{s,t}(x)-Y_{s',t'}(y)|^p\\\leqslant&3^{p-1}\big[|Y_{s,t}(x)-Y_{s,t}(y)|^p+|Y_{s,t}(y)-Y_{s',t}(y)|^p+|Y_{s',t}(y)-Y_{s',t'}(y)|^p\big].\end{aligned}\qquad(3.113)$$

Applying the Itô formula to $|Y_{s,t}(x)-Y_{s,t}(y)|^p$ to get

$$\begin{aligned}&\mathbb{E}|Y_{s,t}(x)-Y_{s,t}(y)|^p\\\leqslant&p(p-1)\Big\{\mathbb{E}\int_s^t|Y_{s,\tau}(x)-Y_{s,\tau}(y)|^{p-1}|\tilde{b}(\tau,Y_{s,\tau}(x))-\tilde{b}(\tau,Y_{s,\tau}(y))|d\tau\\&+\mathbb{E}\int_s^t|Y_{s,\tau}(x)-Y_{s,\tau}(y)|^{p-2}\|\tilde{\sigma}(\tau,Y_{s,\tau}(x))-\tilde{\sigma}(\tau,Y_{s,\tau}(y))\|^2d\tau\Big\}+|x-y|^p\\\leqslant&|x-y|^p+C\int_s^t[1+\kappa^2(\tau)]\mathbb{E}|Y_{s,\tau}(x)-Y_{s,\tau}(y)|^pd\tau,\end{aligned}\qquad(3.114)$$

which suggests

$$\sup_{s\leqslant t\leqslant T}\mathbb{E}|Y_{s,t}(x)-Y_{s,t}(y)|^p\leqslant C|x-y|^p. \tag{3.115}$$

For $|Y_{s,t}(y)-Y_{s',t}(y)|^p$, analogue calculations imply

$$\begin{aligned}
&\mathbb{E}|Y_{s,t}(y)-Y_{s',t}(y)|^p\\
\leqslant&\mathbb{E}|Y_{s,s'}(y)-y|^p+C\mathbb{E}\int_{s'}^t|Y_{s,\tau}(y)-Y_{s',\tau}(y)|^{p-1}|\tilde{b}(\tau,Y_{s,\tau}(y))-\tilde{b}(\tau,Y_{s',\tau}(y))|d\tau\\
&+C\mathbb{E}\int_{s'}^t|Y_{s,\tau}(y)-Y_{s',\tau}(y)|^{p-2}\|\tilde{\sigma}(\tau,Y_{s,\tau}(y))-\tilde{\sigma}(\tau,Y_{s,\tau}(y))\|^2d\tau\\
\leqslant&\mathbb{E}|Y_{s,s'}(y)-y|^p+C\mathbb{E}\int_{s'}^t[1+\kappa^2(\tau)]|Y_{s,\tau}(y)-Y_{s',\tau}(y)|^pd\tau,
\end{aligned}$$

where κ is given in (3.112). This, along with the Grönwall, Minkowski and BGD inequalities, leads to

$$\begin{aligned}
&\mathbb{E}|Y_{s,t}(y)-Y_{s',t}(y)|^p\leqslant C\mathbb{E}|Y_{s,s'}(y)-y|^p\\
=&C\mathbb{E}\Big|\int_s^{s'}\tilde{b}(\tau,Y_{s,\tau}(y))d\tau+\int_s^{s'}\tilde{\sigma}(\tau,Y_{s,\tau}(y))dW_\tau\Big|^p\\
\leqslant&C\Big|\int_s^{s'}\big[\mathbb{E}|\tilde{b}(\tau,Y_{s,\tau}(y))|^p\big]^{\frac{1}{p}}d\tau\Big|^p+C\mathbb{E}\Big[\int_s^{s'}\|\tilde{\sigma}(\tau,Y_{s,\tau}(y))\|^2d\tau\Big]^{\frac{p}{2}}\\
\leqslant&C\big[1+\sup_{s\leqslant\tau\leqslant T}\mathbb{E}|Y_{s,\tau}(y)|^p\big]|s-s'|^p+C|s-s'|^{\frac{p}{2}}\\
\leqslant&C\big[(1+|y|^p)|s-s'|^p+|s-s'|^{\frac{p}{2}}\big]\leqslant C(1+|y|^p)|s-s'|^{\frac{p}{2}},
\end{aligned} \tag{3.116}$$

where in the fourth line we have used the fact $\tilde{b}$ is Lipschitzian uniformly in temperal variable and $\tilde{\sigma}$ is bounded, and in the fifth line we have used (3.107).

For the term $|Y_{s',t}(y)-Y_{s',t'}(y)|^p$, one deduces that

$$\begin{aligned}
\mathbb{E}|Y_{s',t}(y)-Y_{s',t'}(y)|^p=&\mathbb{E}\Big|\int_t^{t'}\tilde{b}(\tau,Y_{s',\tau}(y))d\tau+\int_t^{t'}\tilde{\sigma}(\tau,Y_{s',\tau}(y))dW_\tau\Big|^p\\
\leqslant&C\big(1+|y|^p\big)|t-t'|^{\frac{p}{2}}.
\end{aligned} \tag{3.117}$$

Summing over (3.113), (3.115)–(3.117), we obtain (3.111). Thus $Y_{s,t}(\cdot)$ forms a homeomorphism. Since $Y_{s,t}$ satisfies equation (3.106), we have

$$Y_{s,t}(Y_{s,t}^{-1}(y))=Y_{s,t}^{-1}(y)+\int_s^t\tilde{b}(\tau,Y_{s,\tau}(Y_{s,t}^{-1}(y)))d\tau+\int_s^t\tilde{\sigma}(\tau,Y_{s,\tau}(Y_{s,t}^{-1}(y)))dW_\tau.$$

Note that $Y_{s,\tau}(Y_{s,t}^{-1}(y))=Y_{\tau,t}^{-1}(y)$, thus

$$Y_{s,t}^{-1}(y)=y-\int_s^t\tilde{b}(\tau,Y_{\tau,t}^{-1}(y))d\tau-\int_s^t\tilde{\sigma}(\tau,Y_{\tau,t}^{-1}(y))dW_\tau. \tag{3.118}$$

Hence $Y_{s,t}^{-1}(y)$ is continuous in (s,t,y), a.s. $\omega\in\Omega$, and $\{Y_{s,t}(x),\ t\in[s,T]\}$ forms a stochastic flow of homeomorphisms to SDE (3.106).

(ii) We now turn to show the differentiability, gradient and Hölder estimates. Observing that the inverse flow $Y_{s,t}^{-1}$ satisfies SDE (3.118), which has the same form as the original one (3.106) (only the drift and diffusion have opposite signs), the proof of the differentiability, gradient and Hölder estimates for $Y_{s,t}^{-1}$ are similar to that of $Y_{s,t}$ after taking into consideration the backward character of the equation. For the differentiability and gradient estimates, it suffices to present detailed proofs solely for $Y_{s,t}$. Hölder estimates follow first for $Y_{s,t}$, then for $X_{s,t}$ via Ψ.

Let $e \in \mathbb{R}^d$ with $|e| = 1$. For $\delta \in \mathbb{R}\backslash\{0\}$, we set $Y_{s,t}^{\delta}(y) := [Y_{s,t}(y+\delta e) - Y_{s,t}(y)]/\delta$. By (3.106), we have

$$\begin{aligned} Y_{s,t}^{\delta}(y) =& e + \int_s^t \Big(\int_0^1 \nabla\tilde{b}(\tau, \iota Y_{s,\tau}(y+\delta e) + (1-\iota)Y_{s,\tau}(y)) Y_{s,\tau}^{\delta}(y) d\iota \Big) d\tau \\ & + \int_s^t \Big(\int_0^1 \nabla\tilde{\sigma}(\tau, \iota Y_{s,\tau}(y+\delta e) + (1-\iota)Y_{s,\tau}(y)) Y_{s,\tau}^{\delta}(y) d\iota \Big) dW_\tau. \end{aligned} \tag{3.119}$$

Notice that

$$\nabla\tilde{b}(t,x) = \lambda \nabla U(t, \Psi(t,x)) \nabla\Psi(t,x), \quad \nabla\tilde{\sigma}(t,x) = \nabla^2 U(t,\Psi(t,x))\nabla\Psi(t,x)\sigma(t)$$

and

$$\begin{cases} \nabla U \in \bigcap\limits_{p\geqslant 2} L^p(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^{d\times d})) \cap \bigcap\limits_{0\leqslant\theta<1+\alpha-\frac{2}{q}} L^\infty(0,T;\mathcal{C}_b^{\theta}(\mathbb{R}^d;\mathbb{R}^{d\times d})), \\ \nabla^2 U \in L^{\frac{2q}{2-q\alpha}}(0,T;\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^{d^3})) \cap L^2(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^{d^3})), \end{cases}$$

by (3.96), (3.98) and (3.99), then

$$\begin{cases} \nabla\tilde{b} \in \bigcap\limits_{p\geqslant 2} L^p(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^{d\times d})) \cap \bigcap\limits_{0\leqslant\theta<1+\alpha-\frac{2}{q}} L^\infty(0,T;\mathcal{C}_b^{\theta}(\mathbb{R}^d;\mathbb{R}^{d\times d})), \\ \nabla\tilde{\sigma} \in \bigcap\limits_{2\leqslant p<\frac{2q}{2-q\alpha}} L^p(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^{d^3})) \subset L^2(0,T;\mathcal{C}_b^{1+\alpha-\frac{2}{q}}(\mathbb{R}^d;\mathbb{R}^{d^3})). \end{cases} \tag{3.120}$$

If one sets

$$\tilde{\kappa}(\tau) = \|\nabla\tilde{\sigma}(\tau)\|_0, \quad \bar{\kappa}(\tau) = [\nabla\tilde{\sigma}(\tau)]_{1+\alpha-\frac{2}{q}} \ \text{ and } \ \hat{\kappa}(\tau) = [\nabla\tilde{b}(\tau)]_{1+\alpha-\frac{2}{q}}, \tag{3.121}$$

then $\tilde{\kappa}, \bar{\kappa}, \hat{\kappa} \in L^2(0,T)$.

By Itô's formula, Hölder's inequality and (3.120), we conclude

$$\mathbb{E}|Y_{s,t}^{\delta}(y)|^p \leqslant 1 + C\mathbb{E}\int_s^t \big[1+\tilde{\kappa}^2(\tau)\big]|Y_{s,\tau}^{\delta}(y)|^p d\tau,$$

which suggests that

$$\sup_{s\leqslant t\leqslant T} \mathbb{E}|Y_{s,t}^{\delta}(y)|^p \leqslant C(p), \quad \forall\, p\geqslant 2.$$

This, together with Itô's formula and BDG's inequality, leads to

$$\mathbb{E}\sup_{s\leqslant t\leqslant T} |Y_{s,t}^{\delta}(y)|^p \leqslant C(p), \quad \forall\, p\geqslant 2. \tag{3.122}$$

We rewrite (3.119) by

$$dY_{s,t}^{\delta}(y) = \hat{b}_\delta(t, Y_{s,t}(y,\delta e))Y_{s,t}^{\delta}(y)dt + \hat{\sigma}_\delta(t, Y_{s,t}(y,\delta e))Y_{s,t}^{\delta}(y)dW_t,\ t \in (s,T], \tag{3.123}$$

where

$$\begin{aligned}\hat{b}_\delta(t, Y_{s,t}(y,\delta e)) &= \int_0^1 \nabla\tilde{b}(t, \iota Y_{s,t}(y+\delta e) + (1-\iota)Y_{s,t}(y))d\iota,\\ \hat{\sigma}_\delta(t, Y_{s,t}(y,\delta e)) &= \int_0^1 \nabla\tilde{\sigma}(t, \iota Y_{s,t}(y+\delta e) + (1-\iota)Y_{s,t}(y))d\iota.\end{aligned}$$

Let $\delta, \delta' \in \mathbb{R}$ ($\delta \neq \delta'$), $x, y \in \mathbb{R}^d$ ($x \neq y$) and $s, t, s', t' \in [0,T]$ ($s \leqslant t$, $s' \leqslant t'$). Without loss of generality, we assume $s < s' < t < t'$. For every $p \geqslant 2$, then

$$\begin{aligned}&|Y_{s,t}^{\delta}(x) - Y_{s',t'}^{\delta'}(y)|^p\\ \leqslant&3^{p-1}\big[|Y_{s,t}^{\delta}(x) - Y_{s,t}^{\delta'}(y)|^p + |Y_{s,t}^{\delta'}(y) - Y_{s',t}^{\delta'}(y)|^p + |Y_{s',t}^{\delta'}(y) - Y_{s',t'}^{\delta'}(y)|^p\big].\end{aligned} \tag{3.124}$$

Invoking Itô's formula and Hölder's inequality, we guarantee

$$\begin{aligned}&\mathbb{E}|Y_{s,t}^{\delta}(x) - Y_{s,t}^{\delta'}(y)|^p \leqslant C\Big\{\mathbb{E}\int_s^t |Y_{s,\tau}^{\delta}(x) - Y_{s,\tau}^{\delta'}(y)|^{p-1}|Y_{s,\tau}^{\delta'}(y)|\\ &\times|\hat{b}_\delta(\tau, Y_{s,\tau}(x,\delta e)) - \hat{b}_{\delta'}(\tau, Y_{s,\tau}(y,\delta' e))|d\tau + \mathbb{E}\int_s^t |Y_{s,\tau}^{\delta}(x) - Y_{s,\tau}^{\delta'}(y)|^{p-2}\\ &\times|Y_{s,\tau}^{\delta'}(y)|^2\|\hat{\sigma}_\delta(\tau, Y_{s,\tau}(x,\delta e)) - \hat{\sigma}_{\delta'}(\tau, Y_{s,\tau}(y,\delta' e))\|^2 d\tau\Big\}\\ &+ C\mathbb{E}\int_s^t [1+\tilde{\kappa}^2(\tau)]|Y_{s,\tau}^{\delta}(x) - Y_{s,\tau}^{\delta'}(y)|^p d\tau.\end{aligned} \tag{3.125}$$

Let $\bar{\kappa}$ and $\hat{\kappa}$ be given in (3.121). By (3.120), we get

$$\begin{aligned}&\|\hat{b}_\delta(\tau, Y_{s,\tau}(x,\delta e)) - \hat{b}_{\delta'}(\tau, Y_{s,\tau}(y,\delta' e))\|\\ \leqslant&\int_0^1 \|\nabla\tilde{b}(\tau, \iota Y_{s,\tau}(x+\delta e) + (1-\iota)Y_{s,\tau}(x))\\ &- \nabla\tilde{b}(\tau, \iota Y_{s,\tau}(y+\delta' e) + (1-\iota)Y_{s,\tau}(y))\|d\iota\\ \leqslant&C\hat{\kappa}(\tau)\Big[|Y_{s,\tau}(x+\delta e) - Y_{s,\tau}(y+\delta' e)|^{1+\alpha-\frac{2}{q}} + |Y_{s,\tau}(x) - Y_{s,\tau}(y)|^{1+\alpha-\frac{2}{q}}\Big]\end{aligned} \tag{3.126}$$

and

$$\begin{aligned}&\|\hat{\sigma}_\delta(\tau, Y_{s,\tau}(x,\delta e)) - \hat{\sigma}_{\delta'}(\tau, Y_{s,\tau}(y,\delta' e))\|\\ \leqslant&\int_0^1 \|\nabla\tilde{\sigma}(\tau, \iota Y_{s,\tau}(x+\delta e) + (1-\iota)Y_{s,\tau}(x))\\ &- \nabla\tilde{\sigma}(\tau, \iota Y_{s,\tau}(y+\delta' e) + (1-\iota)Y_{s,\tau}(y))\|d\iota\\ \leqslant&C\bar{\kappa}(\tau)\Big[|Y_{s,\tau}(x+\delta e) - Y_{s,\tau}(y+\delta' e)|^{1+\alpha-\frac{2}{q}} + |Y_{s,\tau}(x) - Y_{s,\tau}(y)|^{1+\alpha-\frac{2}{q}}\Big].\end{aligned} \tag{3.127}$$

Combining (3.125)–(3.127), it follows that

$$
\begin{aligned}
&\mathbb{E}|Y^{\delta}_{s,t}(x)-Y^{\delta'}_{s,t}(y)|^p\\
\leqslant& C\mathbb{E}\int_s^t[1+\tilde{\kappa}^2(\tau)+\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)]|Y^{\delta}_{s,\tau}(x)-Y^{\delta'}_{s,\tau}(y)|^p d\tau\\
&+C\int_s^t\big[\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)\big]\big[\mathbb{E}|Y^{\delta'}_{s,\tau}(y)|^{2p}\big]^{\frac{1}{2}}\Big\{\mathbb{E}\big[|Y_{s,\tau}(x+\delta e)-Y_{s,\tau}(y+\delta' e)|^{2p(1+\alpha-\frac{2}{q})}\\
&\quad+|Y_{s,\tau}(y)-Y_{s,\tau}(x)|^{2p(1+\alpha-\frac{2}{q})}\big]\Big\}^{\frac{1}{2}}d\tau.
\end{aligned}
$$

By (3.111) and (3.122), then

$$
\begin{aligned}
&\mathbb{E}|Y^{\delta}_{s,t}(x)-Y^{\delta'}_{s,t}(y)|^p\\
\leqslant& C\mathbb{E}\int_s^t\big[1+\tilde{\kappa}^2(\tau)+\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)\big]|Y^{\delta}_{s,\tau}(x)-Y^{\delta'}_{s,\tau}(y)|^p d\tau\\
&+C\int_s^t\big[\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)\big]d\tau\big[|x-y|^{p(1+\alpha-\frac{2}{q})}+|\delta-\delta'|^{p(1+\alpha-\frac{2}{q})}\big],
\end{aligned}
\tag{3.128}
$$

which also implies

$$
\mathbb{E}|Y^{\delta}_{s,t}(x)-Y^{\delta'}_{s,t}(y)|^p\leqslant C\big[|x-y|^{p(1+\alpha-\frac{2}{q})}+|\delta-\delta'|^{p(1+\alpha-\frac{2}{q})}\big],\tag{3.129}
$$

if one uses Grönwall's inequality.

For $|Y^{\delta'}_{s,t}(y)-Y^{\delta'}_{s',t}(y)|^p$, we estimate that

$$
\begin{aligned}
&\mathbb{E}|Y^{\delta'}_{s,t}(y)-Y^{\delta'}_{s',t}(y)|^p\\
\leqslant&\mathbb{E}|Y^{\delta'}_{s,s'}(y)-e|^p+C\mathbb{E}\int_{s'}^t\big[1+\tilde{\kappa}^2(\tau)+\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)\big]|Y^{\delta'}_{s,\tau}(y)-Y^{\delta'}_{s',\tau}(y)|^p d\tau\\
&+C\int_{s'}^t\big[\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)\big]\big[\mathbb{E}|Y^{\delta'}_{s,\tau}(y)|^{2p}\big]^{\frac{1}{2}}\\
&\quad\times\Big\{\mathbb{E}\big[|Y_{s,\tau}(y+\delta'e)-Y_{s',\tau}(y+\delta'e)|^{2p(1+\alpha-\frac{2}{q})}+|Y_{s,\tau}(y)-Y_{s',\tau}(y)|^{2p(1+\alpha-\frac{2}{q})}\big]\Big\}^{\frac{1}{2}}d\tau\\
\leqslant&\mathbb{E}|Y^{\delta'}_{s,s'}(y)-e|^p+C\mathbb{E}\int_{s'}^t\big[1+\tilde{\kappa}^2(\tau)+\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)\big]|Y^{\delta'}_{s,\tau}(y)-Y^{\delta'}_{s',\tau}(y)|^p d\tau\\
&+C\int_{s'}^t\big[\hat{\kappa}(\tau)+\bar{\kappa}^2(\tau)\big]d\tau\big[1+|y|^p+|\delta'|^p\big]|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}.
\end{aligned}
$$

This results in

$$
\mathbb{E}|Y^{\delta'}_{s,t}(y)-Y^{\delta'}_{s',t}(y)|^p\leqslant\mathbb{E}|Y^{\delta'}_{s,s'}(y)-e|^p+C[1+|y|^p+|\delta'|^p]|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}.\tag{3.130}
$$

Notice that

$$\begin{aligned}
&\mathbb{E}|Y^{\delta'}_{s,s'}(y)-e|^p\\
=&\mathbb{E}\Big|\int_s^{s'}\hat{b}_{\delta'}(\tau,Y_{s,\tau}(y,\delta' e))Y^{\delta'}_{s,\tau}(y)d\tau+\int_s^{s'}\hat{\sigma}_{\delta'}(\tau,Y_{s,\tau}(y,\delta' e))Y^{\delta'}_{s,\tau}(y)dW_\tau\Big|^p\\
\leqslant&C\Big|\int_s^{s'}\big[\mathbb{E}|Y^{\delta'}_{s,\tau}(y)|^p\big]^{\frac{1}{p}}d\tau\Big|^p+C\mathbb{E}\Big[\int_s^{s'}\kappa^2(\tau)|Y^{\delta'}_{s,\tau}(y)|^2d\tau\Big]^{\frac{p}{2}}\\
\leqslant&C|s-s'|^p+C\mathbb{E}\sup_{s\leqslant\tau\leqslant T}|Y^{\delta'}_{s,\tau}(y)|^p\Big[\int_s^{s'}\kappa^{\frac{2q}{2-q\alpha}}(\tau)d\tau\Big]^{\frac{p(2-q\alpha)}{2q}}|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}\\
\leqslant&C\big[|s-s'|^p+|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}\big],
\end{aligned}$$

where κ is given in (3.112), we conclude

$$\mathbb{E}|Y^{\delta'}_{s,t}(y)-Y^{\delta'}_{s',t}(y)|^p\leqslant C\big[1+|y|^p+|\delta'|^p\big]|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}. \tag{3.131}$$

For the term $|Y^{\delta'}_{s',t}(y)-Y^{\delta'}_{s',t'}(y)|^p$, it yields that

$$\begin{aligned}
&\mathbb{E}|Y^{\delta'}_{s',t}(y)-Y^{\delta'}_{s',t'}(y)|^p\\
=&\mathbb{E}\Big|\int_t^{t'}\hat{b}_{\delta'}(\tau,Y_{s',\tau}(y,\delta' e))Y^{\delta'}_{s',\tau}(y)d\tau+\int_t^{t'}\hat{\sigma}_{\delta'}(\tau,Y_{s',\tau}(y,\delta' e))Y^{\delta'}_{s',\tau}(y)dW_\tau\Big|^p\\
\leqslant&C|t-t'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}.
\end{aligned} \tag{3.132}$$

Summing over (3.124), (3.129), (3.131) and (3.132), one ends up with

$$\begin{aligned}
\mathbb{E}|Y^{\delta}_{s,t}(x)-Y^{\delta'}_{s',t'}(y)|^p\leqslant C\Big\{&|x-y|^{p(1+\alpha-\frac{2}{q})}+|\delta-\delta'|^{p(1+\alpha-\frac{2}{q})}\\
&+[1+|y|^p+|\delta'|^p]|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}+|t-t'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}\Big\}.
\end{aligned} \tag{3.133}$$

By Kolmogorov's extension theorem $Y^\delta_{s,t}(y)$ has a continuous extension at $\delta=0$ for $0\leqslant s<t\leqslant T$ and $y\in\mathbb{R}^d$, a.s.. This means $Y_{s,t}(\cdot)$ is continuously differentiable in the domain $\{(s,t,y)\mid 0\leqslant s\leqslant t\leqslant T, y\in\mathbb{R}^d\}$. Moreover, $\nabla Y_{s,t}(y)$ satisfies

$$\left\{\begin{aligned}
&d\nabla Y_{s,t}(y)=\lambda\nabla U(t,\Psi(t,Y_{s,t}(y)))\nabla\Psi(t,Y_{s,t}(y))\nabla Y_{s,t}(y)dt\\
&\quad+\nabla^2U(t,\Psi(t,Y_{s,t}(y)))\nabla\Psi(t,Y_{s,t}(y))\nabla Y_{s,t}(y)\sigma(t)dW_t,\quad t\in(s,T],\\
&\nabla Y_{s,t}(y)|_{t=s}=I_{d\times d}.
\end{aligned}\right. \tag{3.134}$$

Furthermore, for every $p\geqslant 2$, $x,y\in\mathbb{R}^d$ $(x\neq y)$ and $s,t,s',t'\in[0,T]$ $(s\leqslant t,\ s'\leqslant t')$, we have

$$\mathbb{E}\sup_{s\leqslant t\leqslant T}\|\nabla Y_{s,t}(y)\|^p\leqslant C(p) \tag{3.135}$$

and

$$\begin{aligned}
&\mathbb{E}\|\nabla Y_{s,t}(x)-\nabla Y_{s',t'}(y)\|^p\\
\leqslant&C\Big\{|x-y|^{p(1+\alpha-\frac{2}{q})}+\big[1+|y|^p\big]|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}+|t-t'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}\Big\}.
\end{aligned} \tag{3.136}$$

By (3.136) and the Kolmogorov–Chentsov continuity criterion, $\nabla Y_{s,t}(y)$ has a continuous realization (denoted by itself), which is locally $(\beta/2, \beta/2, \beta)$-Hölder continuous in (s,t,y), for every $\beta \in (0, 1+\alpha-2/q)$, i.e., there exists $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that $\nabla Y_{s,t}(y)$ is locally $(\beta/2, \beta/2, \beta)$-Hölder continuous in (s,t,y) for all $\omega \in \Omega_0$. In addition, by (3.120), Itô's formula and BDG's inequality, we also get an analogue of (3.136) that

$$\sup_{0\leqslant s\leqslant T} \mathbb{E} \sup_{s\leqslant t\leqslant T} \|\nabla Y_{s,t}(x) - \nabla Y_{s,t}(y)\|^p \leqslant C|x-y|^{p(1+\alpha-\frac{2}{q})}, \quad \forall\, p \geqslant 2. \tag{3.137}$$

By [42, Theorem 1.3], for every $\beta \in (0, 1+\alpha-2/q)$ and every $R > 0$, then

$$\sup_{0\leqslant s\leqslant T} \mathbb{E}\Big[\sup_{s\leqslant t\leqslant T} \Big(\sup_{x,y\in B_R, x\neq y} \frac{\|\nabla Y_{s,t}(x) - \nabla Y_{s,t}(y)\|}{|x-y|^\beta}\Big)^p\Big] < +\infty, \quad \forall\, p \geqslant 2. \tag{3.138}$$

Since $X_{s,t}(x) = \Psi(t, Y_{s,t}(\Phi(s,x)))$, we have

$$\nabla X_{s,t}(x) = \nabla\Psi(t, Y_{s,t}(\Phi(s,x)))\nabla Y_{s,t}(\Phi(s,x))\nabla\Phi(s,x). \tag{3.139}$$

By (3.101), (3.102) and (3.139),

$$\begin{aligned}
&\|\nabla X_{s,t}(x) - \nabla X_{s',t'}(y)\| \\
\leqslant& C|t-t'|^{\tilde{\theta}}[1 + |Y_{s,t}(\Phi(s,x))|^\alpha]\|\nabla Y_{s,t}(\Phi(s,x))\| \\
&+ C|Y_{s,t}(\Phi(s,x)) - Y_{s',t'}(\Phi(s',y))|^\theta \|\nabla Y_{s,t}(\Phi(s,x))\| \\
&+ C\|\nabla Y_{s,t}(\Phi(s,x)) - \nabla Y_{s',t'}(\Phi(s',y))\| \\
&+ C\|\nabla Y_{s',t'}(\Phi(s',y))\|\|\nabla\Phi(s,x) - \nabla\Phi(s',y)\|,
\end{aligned} \tag{3.140}$$

where $x, y \in \mathbb{R}^d$, $s,t,s',t' \in [0,T]$ $(s \leqslant t,\ s' \leqslant t')$, $0 < \tilde{\theta} < (1+\alpha-2/q)(1-1/q)$ and $\theta \in (0, 1+\alpha-2/q)$.

On account of (3.97), (3.107), (3.111), (3.135) and (3.136), for every $p \geqslant 2$,

$$\begin{aligned}
&\mathbb{E}\|\nabla X_{s,t}(x) - \nabla X_{s',t'}(y)\|^p \leqslant C|t-t'|^{p\tilde{\theta}}\big[1 + |\Phi(s,x)|^{p\alpha}\big] \\
&\quad + C\Big\{|\Phi(s,x) - \Phi(s',y)|^{p\theta} + \big[1 + |\Phi(s',y)|^p\big]\big[|s-s'|^{\frac{p\theta}{2}} + |t-t'|^{\frac{p\theta}{2}}\big] \\
&\quad + |\Phi(s,x) - \Phi(s',y)|^{p(1+\alpha-\frac{2}{q})} + \big[1 + |\Phi(s',y)|^p\big]|s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})} \\
&\quad + |t-t'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}\Big\} + C|s-s'|^{\frac{p\theta}{2}} + C|x-y|^{p\theta} \\
\leqslant& C\big(1 + |x|^p + |y|^p\big)\Big[|t-t'|^{p\tilde{\theta}} + |s-s'|^{(1-\frac{1}{q})p\theta} + |s-s'|^{\frac{p\theta}{2}} + |t-t'|^{\frac{p\theta}{2}} \\
&\quad + |s-s'|^{p(1+\alpha-\frac{2}{q})(1-\frac{1}{q})} + |s-s'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})} + |t-t'|^{\frac{p}{2}(1+\alpha-\frac{2}{q})}\Big] \\
&\quad + C\Big[|x-y|^{p\theta} + |x-y|^{p(1+\alpha-\frac{2}{q})}\Big] \\
\leqslant& C\Big[|x-y|^{p\theta} + \big(1 + |x|^p + |y|^p\big)\big(|t-t'|^{p\tilde{\theta}} + |s-s'|^{(1-\frac{1}{q})p\theta}\big)\Big],
\end{aligned} \tag{3.141}$$

where $\theta > 2\tilde{\theta}$ and $x, y \in B_R$ $(R > 0)$.

Using the Kolmogorov–Chentsov continuity criterion again, $\nabla X_{s,t}(x)$ has a continuous realization (denoted by itself), which is locally $(\gamma_1, \gamma_1, \gamma_2)$-Hölder continuous in (s, t, x) for every $0 < \gamma_1 < (1-1/q)(1+\alpha-2/q)$ and $0 < \gamma_2 < 1+\alpha-2/q$. In particular, (3.141) holds true for $s = s'$ and $t = t'$. We then apply Itô's formula and BDG's inequality, to get

$$\sup_{0\leqslant s\leqslant T} \mathbb{E} \sup_{s\leqslant t\leqslant T} \|\nabla X_{s,t}(x) - \nabla X_{s,t}(y)\|^p \leqslant C|x-y|^{p\theta}, \quad \forall\, x, y \in B_R.$$

This, together with [42, Theorem 1.3], yields the estimate (1.9) for $\nabla X_{s,t}(x)$.

(iii) Clearly, it is sufficient to prove the stability for $X_{s,t}$. Let U_n be the unique strong solution of (3.94) with b replaced by b_n. By Theorem 2.3, U_n belongs to $(\mathcal{G}^{\alpha,\theta}_{q,T})^d$ for every $\theta \in [0, 1+\alpha-2/q)$ and satisfies (2.91) uniformly in n. Moreover, $U_n - U \in \mathcal{C}([0,T]; \mathcal{C}^1_b(\mathbb{R}^d; \mathbb{R}^d))$ and

$$\lim_{n\to+\infty} \left[\|U_n - U\|_{\infty,0} + \|\nabla U_n - \nabla U\|_{p,0}\right] = 0. \tag{3.142}$$

Let $\Phi_n(t,x) = x + U_n(t,x)$, then $\{\Phi_n\}_{n\geqslant 1}$ form non-singular diffeomorphisms of class $\mathcal{C}^{1,\theta}$ uniformly in $(t,n) \in [0,T]\times\mathbb{N}$, and

$$\frac{1}{2} < \sup_{0\leqslant t\leqslant T} \|\nabla\Phi_n(t)\|_0 < \frac{3}{2}, \quad \frac{2}{3} < \sup_{0\leqslant t\leqslant T} \|\nabla\Psi_n(t)\|_0 < 2,$$

where $\Psi_n(t,\cdot) = \Phi_n^{-1}(t,\cdot)$. Notice that

$$\begin{aligned} |\Psi_n(t,x) - \Psi(t,x)| \leqslant& \sup_{0\leqslant t\leqslant T} \|[\nabla\Phi_n(t)]^{-1}\|_0 |\Phi_n(t,\Psi_n(t,x)) - \Phi_n(t,\Psi(t,x))| \\ \leqslant& 2|\Phi_n(t,\Psi_n(t,x)) - \Phi_n(t,\Psi(t,x))| \\ =& 2|\Phi(t,\Psi(t,x)) - \Phi_n(t,\Psi(t,x))| \leqslant 2\sup_{0\leqslant t\leqslant T} \|U_n(t) - U(t)\|_0, \end{aligned}$$

then

$$\lim_{n\to+\infty} \|\Psi_n - \Psi\|_{\infty,0} = 0. \tag{3.143}$$

Let $Y^n_{s,t} = \Phi_n(t, X^n_{s,t})$. It satisfies the following equation:

$$\begin{cases} dY^n_{s,t} = \lambda U_n(t,\Psi_n(t,Y^n_{s,t}))dt + [I + \nabla U_n(t,\Psi_n(t,Y^n_{s,t}))]\sigma(t)dW_t \\ \qquad =: \tilde{b}_n(t,Y^n_{s,t})dt + \tilde{\sigma}_n(t,Y^n_{s,t})dW_t, \ t \in (s,T], \\ Y^n_{s,t}|_{t=s} = y = \Phi_n(s,x). \end{cases} \tag{3.144}$$

Observe that

$$\begin{aligned} &|X^n_{s,t}(x) - X_{s,t}(x)| = |\Psi^n(t, Y^n_{s,t}(\Phi^n(s,x))) - \Psi(t, Y_{s,t}(\Phi(s,x)))| \\ \leqslant& \|\Psi_n - \Psi\|_{\infty,0} + \|\nabla\Psi\|_{\infty,0} |Y^n_{s,t}(\Phi^n(s,x)) - Y_{s,t}(\Phi^n(s,x))| \\ &+ \|\nabla\Psi\|_{\infty,0} |Y_{s,t}(\Phi^n(s,x)) - Y_{s,t}(\Phi(s,x))|, \end{aligned}$$

then

$$
\begin{aligned}
&\lim_{n\to+\infty}\sup_{x\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\mathbb{E}\sup_{s\leqslant t\leqslant T}|X^n_{s,t}(x)-X_{s,t}(x)|^p\\
\leqslant& C\lim_{n\to+\infty}\big[\|\Psi_n-\Psi\|^p_{\infty,0}\\
&+\sup_{x\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\mathbb{E}\sup_{s\leqslant t\leqslant T}|Y^n_{s,t}(\Phi^n(s,x))-Y_{s,t}(\Phi^n(s,x))|^p\\
&+\sup_{x\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\mathbb{E}\sup_{s\leqslant t\leqslant T}|Y_{s,t}(\Phi^n(s,x))-Y_{s,t}(\Phi(s,x))|^p\big]\\
\leqslant& C\lim_{n\to+\infty}\big[\sup_{y\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\mathbb{E}\sup_{s\leqslant t\leqslant T}|Y^n_{s,t}(y)-Y_{s,t}(y)|^p+\|\Phi^n-\Phi\|^p_{\infty,0}\big]\\
=&C\lim_{n\to+\infty}\sup_{y\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\mathbb{E}\sup_{s\leqslant t\leqslant T}|Y^n_{s,t}(y)-Y_{s,t}(y)|^p,
\end{aligned}
\tag{3.145}
$$

thus it suffices to show the stability for $Y_{s,t}$.

For every $\tilde{p}\geqslant 2$, applying Itô's formula for $|Y^n_{s,t}-Y_{s,t}|^{\tilde{p}}$, to get

$$
\begin{aligned}
d|Y^n_{s,t}-Y_{s,t}|^{\tilde{p}}\leqslant& C(\tilde{p})|Y^n_{s,t}-Y_{s,t}|^{\tilde{p}-1}\big[\|U_n-U\|_{\infty,0}+\|\nabla U\|_{\infty,0}\|\Psi_n-\Psi\|_{\infty,0}\\
&+\|\nabla U\|_{\infty,0}\|\nabla\Psi\|_{\infty,0}|Y^n_{s,t}-Y_{s,t}|\big]dt+C(\tilde{p})|Y^n_{s,t}-Y_{s,t}|^{\tilde{p}-2}\\
&\times\big[\|\nabla U_n(t)-\nabla U(t)\|^2_0+\|\nabla^2U(t)\|^2_0\|\Psi_n-\Psi\|^2_{\infty,0}\\
&+\|\nabla^2U(t)\|^2_0\|\nabla\Psi\|^2_{\infty,0}|Y^n_{s,t}-Y_{s,t}|^2\big]dt+\tilde{p}|Y^n_{s,t}-Y_{s,t}|^{\tilde{p}-2}\\
&\times\langle Y^n_{s,t}-Y_{s,t},[\nabla U_n(t,\Psi_n(t,Y^n_{s,t}))-\nabla U(t,\Psi(t,Y_{s,t}))]\sigma(t)dW_t\rangle.
\end{aligned}
\tag{3.146}
$$

By using Hölder's inequality, we deduce from (3.146) that

$$
\begin{aligned}
\mathbb{E}|Y^n_{s,t}-Y_{s,t}|^{\tilde{p}}\leqslant& C(\tilde{p})\mathbb{E}\int_s^t[1+\kappa^2(\tau)]|Y^n_{s,\tau}-Y_{s,\tau}|^{\tilde{p}}d\tau\\
&+C(\tilde{p})\big[\|U_n-U\|^{\tilde{p}}_{\infty,0}+\|\Psi_n-\Psi\|^{\tilde{p}}_{\infty,0}+\|\nabla U_n-\nabla U\|^{\tilde{p}}_{\tilde{p},0}\big],
\end{aligned}
$$

where κ is given in (3.112).

This, together with the Gronwall's inequality, (3.142) and (3.143), implies

$$
\begin{aligned}
&\lim_{n\to+\infty}\sup_{y\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\sup_{s\leqslant t\leqslant T}\mathbb{E}\big[|Y^n_{s,t}(y)-Y_{s,t}(y)|^{\tilde{p}}\big]\\
\leqslant& C\lim_{n\to+\infty}\big[\|U_n-U\|^{\tilde{p}}_{\infty,0}+\|\Psi_n-\Psi\|^{\tilde{p}}_{\infty,0}+\|\nabla U_n-\nabla U\|^{\tilde{p}}_{\tilde{p},0}\big]=0.
\end{aligned}
\tag{3.147}
$$

By (3.146), (3.147) and BDG's inequality to get for every $p\geqslant 2$ that

$$
\begin{aligned}
&\lim_{n\to+\infty}\sup_{y\in\mathbb{R}^d}\sup_{0\leqslant s\leqslant T}\mathbb{E}\sup_{s\leqslant t\leqslant T}|Y^n_{s,t}(y)-Y_{s,t}(y)|^p\\
\leqslant& C\lim_{n\to+\infty}\big[\|U_n-U\|^p_{\infty,0}+\|\nabla U_n-\nabla U\|^p_{p,0}+\|\Psi_n-\Psi\|^p_{\infty,0}\big]\\
&+C\lim_{n\to+\infty}\Big[\sup_{0\leqslant s\leqslant T}\mathbb{E}\int_s^T\|Y^n_{s,t}(y)-Y_{s,t}(y)|^{2p-2}\\
&\qquad\times\|\nabla U_n(t,\Psi_n(t,Y^n_{s,t}))-\nabla U(t,\Psi(t,Y_{s,t}))\|^2dt\Big]^{\frac{1}{2}}\\
\leqslant& C\lim_{n\to+\infty}\Big[\sup_{0\leqslant s\leqslant T}\mathbb{E}\int_s^T\|Y^n_{s,t}(y)-Y_{s,t}(y)|^{2p-2}dt\Big]^{\frac{1}{2}}=0.
\end{aligned}
\tag{3.148}
$$

From (3.145) and (3.148), we complete the proof. □

4 Proof of Theorem 1.4

Before giving the proof of L^2-transportation cost inequality (1.17), we need a useful lemma.

Lemma 4.1. *([38, Lemma 2.1]) Let $(\mathscr{E}, \tilde{\rho})$ be a Polish space and Φ be a homeomorphism on $\mathscr{E}$. Let $\mathscr{P}(\mathscr{E})$ be the class of all probability measures on $\mathscr{E}$. For $\mu, \nu \in \mathscr{P}(\mathscr{E})$, we define the L^2-Wasserstein distance between μ and ν by*

$$\mathbb{W}_{\mathscr{E}}^{\tilde{\rho}}(\mu,\nu) = \inf_{\pi \in \mathscr{C}(\mu,\nu)} \Big(\int_{\mathscr{E}\times\mathscr{E}} \tilde{\rho}^2(\phi_1,\phi_2) d\pi(\phi_1,\phi_2) \Big)^{\frac{1}{2}}.$$

The relative entropy (or Kullback information) of ν with respect to μ is given by (1.18). Then we have

(i) $\mathbb{W}_{\mathscr{E}}^{\tilde{\rho}}(\mu,\nu) = \mathbb{W}_{\mathscr{E}}^{\tilde{\rho}\circ\Phi^{-1}}(\mu\circ\Phi^{-1}, \nu\circ\Phi^{-1})$. Moreover, if there are positive constants c_1 and c_2 such that

$$c_1\tilde{\rho}(\phi_1,\phi_2) \leqslant \tilde{\rho}(\Phi(\phi_1),\Phi(\phi_2)) \leqslant c_2\tilde{\rho}(\phi_1,\phi_2), \quad \phi_1,\phi_2 \in \mathscr{E},$$

then

$$c_1\mathbb{W}_{\mathscr{E}}^{\tilde{\rho}}(\mu,\nu) \leqslant \mathbb{W}_{\mathscr{E}}^{\tilde{\rho}}(\mu\circ\Phi^{-1},\nu\circ\Phi^{-1}) \leqslant c_2\mathbb{W}_{\mathscr{E}}^{\tilde{\rho}}(\mu,\nu).$$

(ii) $H(\nu|\mu) = H(\nu\circ\Phi^{-1}|\mu\circ\Phi^{-1})$.

Now, let us give the proof details. Let U be the unique strong solution of (3.94). Then $U \in (\mathcal{G}_{q,T}^{\alpha,\theta})^d$ for every $\theta \in [0, 1+\alpha-2/q)$ and satisfies (3.95) for some large enough λ. We set $\Phi(t,x) = x + U(t,x)$ and define $\Phi, \Phi^{-1} : \Omega \to \Omega$ by

$$\Phi(\phi)(t) = \Phi(t,\phi(t)) \text{ and } \Phi^{-1}(\phi)(t) = \Phi^{-1}(t,\phi(t)),\ \phi \in \Omega,\ t \in [0,T], \tag{4.149}$$

where $\Omega = \mathcal{C}_0([0,T];\mathbb{R}^d)$

By (3.96), Φ forms a homeomorphism on Ω and

$$\frac{1}{2}\|\phi_1-\phi_2\|_\Omega \leqslant \|\Phi(\phi_1)-\Phi(\phi_2)\|_\Omega \leqslant \frac{3}{2}\|\phi_1-\phi_2\|_\Omega, \quad \phi_1,\phi_2 \in \Omega, \tag{4.150}$$

where $\|\cdot\|_\Omega$ is the uniform norm on Ω given by (1.16).

We choose $(\mathscr{E},\tilde{\rho})$ by $(\Omega, \|\cdot\|_\Omega)$ and $\mathbb{W}_{\mathscr{E}}^{\tilde{\rho}}$ by $\mathbb{W}_\Omega$. By (4.150) and Lemma 4.1, it is sufficient to show

$$\mathbb{W}_\Omega^2(\tilde{\mathbb{Q}}\circ\Phi^{-1},\tilde{\mathbb{P}}\circ\Phi^{-1}) \leqslant e^{C(d,T,\alpha,q,\Theta,[b]_{q,\alpha})} H(\tilde{\mathbb{Q}}\circ\Phi^{-1}|\tilde{\mathbb{P}}\circ\Phi^{-1}). \tag{4.151}$$

Let $\{X_t(x), t\in[0,T]\}$ be the unique strong solution of SDE (1.1) with $s=0$ on the given stochastic basis $(\Omega,\mathcal{F},\mathbb{P},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T})$ with $\tilde{\mathbb{P}}$ as its law on Ω. Then $\tilde{\mathbb{P}}\circ\Phi^{-1}$ is the law of $Y_t(y) = \Phi(t, X_t(x))$ on Ω, where $Y_t(y)$ satisfies

$$\begin{cases} dY_t(y) = \lambda U(t,\Phi^{-1}(t,Y_t(y)))dt + [I + \nabla U(t,\Phi^{-1}(t,Y_t(y)))]\sigma(t)dW_t \\ \qquad\quad =: \tilde{b}(t,Y_t(y)dt + \tilde{\sigma}(t,Y_t(y))dW_t,\ t\in(s,T], \\ Y_t|_{t=0} = y = \Phi(0,x), \end{cases} \tag{4.152}$$

with $\tilde{b} \in \mathcal{C}([0,T];\mathcal{C}^1(\mathbb{R}^d;\mathbb{R}^d))$ and $\tilde{\sigma} \in L^2(0,T;\mathcal{C}_b^1(\mathbb{R}^d;\mathbb{R}^{d\times d})) \cap \mathcal{C}([0,T];\mathcal{C}_b(\mathbb{R}^d;\mathbb{R}^{d\times d}))$, which satisfy

$$\|\nabla\tilde{b}\|_{\infty,0} \leqslant 1, \quad \|\tilde{\sigma}\|_{\infty,0} \leqslant \frac{3\sqrt{d\Theta}}{2} \quad \text{and} \quad \|\nabla\tilde{\sigma}\|_{2,0} \leqslant C(d,T,\alpha,q,\Theta,[b]_{q,\alpha}). \tag{4.153}$$

Let $\tilde{\mathbb{Q}} \in \mathscr{P}(\Omega)$. If $\tilde{\mathbb{Q}} \circ \Phi^{-1}$ is singular with respect to $\tilde{\mathbb{P}} \circ \Phi^{-1}$ or $\tilde{\mathbb{Q}} \circ \Phi^{-1} \ll \tilde{\mathbb{P}} \circ \Phi^{-1}$ with $H(\tilde{\mathbb{Q}} \circ \Phi^{-1}|\tilde{\mathbb{P}} \circ \Phi^{-1}) = +\infty$, then (4.151) holds mutatis mutandis. We now assume that $\tilde{\mathbb{Q}} \circ \Phi^{-1} \ll \tilde{\mathbb{P}} \circ \Phi^{-1}$ and $H(\tilde{\mathbb{Q}} \circ \Phi^{-1}|\tilde{\mathbb{P}} \circ \Phi^{-1}) < +\infty$. Set

$$\mathbb{Q} := \frac{d\tilde{\mathbb{Q}} \circ \Phi^{-1}}{d\tilde{\mathbb{P}} \circ \Phi^{-1}}(Y_{\cdot}(y))\mathbb{P},$$

then

$$\int_\Omega \frac{d\tilde{\mathbb{Q}} \circ \Phi^{-1}}{d\tilde{\mathbb{P}} \circ \Phi^{-1}}(Y_{\cdot}(y))d\mathbb{P} = \int_\Omega \frac{d\tilde{\mathbb{Q}} \circ \Phi^{-1}}{d\tilde{\mathbb{P}} \circ \Phi^{-1}}(\omega)d\tilde{\mathbb{P}} \circ \Phi^{-1}(\omega) = 1.$$

Thus $\mathbb{Q} \in \mathscr{P}(\Omega)$. Under the probability measure $\mathbb{Q}$, there exists a Brownian motion $\{B_t\}_{0\leqslant t\leqslant T}$ and a predictable process $\{\gamma_t\}_{0\leqslant t\leqslant T}$ such that the coordinates system $\{w_t\}_{0\leqslant t\leqslant T}$ of $\mathcal{C}_0([0,T];\mathbb{R}^d)$ verifies

$$w_t = B_t - \int_0^t \gamma_\tau(w)d\tau. \tag{4.154}$$

A further consequence is that (see [8, Section 5])

$$H(\tilde{\mathbb{Q}} \circ \Phi^{-1}|\tilde{\mathbb{P}} \circ \Phi^{-1}) = \frac{1}{2}\mathbb{E}_{\mathbb{Q}} \int_0^T |\gamma_t(w)|^2 dt. \tag{4.155}$$

Let Y_t^γ be the unique strong solution of

$$dY_t^\gamma(y) = \tilde{b}(t, Y_t^\gamma(y))dt + \tilde{\sigma}(t, Y_t^\gamma(y))dB_t,\ t \in (0,T],\ Y_t^\gamma(y)|_{t=0} = y = \Phi(0,x). \tag{4.156}$$

Then the law of $Y_{\cdot}^\gamma$ under $\mathbb{Q}$ is equal to $\tilde{\mathbb{P}} \circ \Phi^{-1}$. By (4.152) and (4.154), Y_t satisfies

$$\begin{cases} dY_t(y) = \tilde{b}(t, Y_t(y))dt + \tilde{\sigma}(t, Y_t(y))dB_t - \tilde{\sigma}(t, Y_t(y))\gamma_t dt,\ t \in (0,T], \\ Y_t(y)|_{t=0} = y = \Phi(0,x), \end{cases} \tag{4.157}$$

under $\mathbb{Q}$. Therefore, (Y^γ, Y) under $\mathbb{Q}$ is a coupling of $(\tilde{\mathbb{P}} \circ \Phi^{-1}, \tilde{\mathbb{Q}} \circ \Phi^{-1})$. By (4.156) and (4.157),

$$\begin{aligned} |Y_t(y) - Y_t^\gamma(y)|^2 =& \Big| \int_0^t [\tilde{b}(\tau, Y_\tau(y)) - \tilde{b}(\tau, Y_\tau^\gamma(y))]d\tau \\ &+ \int_0^t [\tilde{\sigma}(\tau, Y_\tau(y)) - \tilde{\sigma}(\tau, Y_\tau^\gamma(y))]dB_\tau + \int_0^t \tilde{\sigma}(\tau, Y_\tau^\gamma(y))\gamma_\tau d\tau \Big|^2 \\ \leqslant& 3\Big| \int_0^t [\tilde{b}(\tau, Y_\tau(y)) - \tilde{b}(\tau, Y_\tau^\gamma(y))]d\tau \Big|^2 + 3\Big| \int_0^t [\tilde{\sigma}(\tau, Y_\tau(y)) - \tilde{\sigma}(\tau, Y_\tau^\gamma(y))]dB_\tau \Big|^2 \\ &+ 3\Big| \int_0^t \tilde{\sigma}(\tau, Y_\tau^\gamma(y))\gamma_\tau d\tau \Big|^2. \end{aligned}$$

In view of the (4.153), Doob's and Hölder's inequalities, we get

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \sup_{0\leqslant t\leqslant \tau} |Y_t(y) - Y_t^\gamma(y)|^2 \leqslant& 3\tau\mathbb{E}_{\mathbb{Q}} \int_0^\tau \|\nabla\tilde{b}\|_{\infty,0}^2 |Y_\tau(y) - Y_\tau^\gamma(y)|^2 d\tau \\ &+ 12\mathbb{E}_{\mathbb{Q}} \int_0^\tau \|\tilde{\sigma}(\tau, Y_\tau(y)) - \tilde{\sigma}(\tau, Y_\tau^\gamma(y))\|^2 d\tau + 3\|\tilde{\sigma}\|_{\infty,0}^2 \tau\mathbb{E}_{\mathbb{Q}} \int_0^\tau |\gamma_\tau|^2 d\tau \\ \leqslant& 3\mathbb{E}_{\mathbb{Q}} \int_0^\tau \big[T + 4\|\nabla\tilde{\sigma}(\tau)\|_0^2\big]|Y_\tau(y) - Y_\tau^\gamma(y)|^2 d\tau + \frac{27d\Theta T}{4}\mathbb{E}_{\mathbb{Q}} \int_0^\tau |\gamma_\tau|^2 d\tau. \end{aligned}$$

This, together with Grönwall's inequality and (4.155), yields

$$\mathbb{E}_{\mathbb{Q}} \sup_{0\leqslant t\leqslant T} |Y_t(y) - Y_t^{\gamma}(y)|^2 \leqslant e^{C(d,T,\alpha,q,\Theta,[b]_{q,\alpha})} H(\tilde{\mathbb{Q}}\circ\Phi^{-1}|\tilde{\mathbb{P}}\circ\Phi^{-1}). \tag{4.158}$$

Observing that

$$\mathbb{W}_{\Omega}^2(\tilde{\mathbb{Q}}\circ\Phi^{-1},\tilde{\mathbb{P}}\circ\Phi^{-1}) \leqslant \mathbb{E}_{\mathbb{Q}} \sup_{0\leqslant t\leqslant T} |Y_t(y) - Y_t^{\gamma}(y)|^2,$$

we conclude the estimate (4.151). □

5 Proof of Theorem 1.6

Firstly, we prove that $u(t,x) = u_0(X_t^{-1}(x))$ is a stochastic strong solution of (1.19). Given that $u_0 \in L^\infty(\mathbb{R}^d)$, we clearly have $u \in L^\infty(\Omega\times[0,T]\times\mathbb{R}^d)$. From the fact that $X_t^{-1}(x)$ is differentiable in x, we have the following chain rule

$$\nabla_x u_0(X_t^{-1}(x)) = \nabla u_0(X_t^{-1}(x))\nabla X_t^{-1}(x). \tag{5.159}$$

For every $p\geqslant 2$ and every $R>0$, by (1.8) and (1.9), then

$$\sup_{0\leqslant t\leqslant T}\mathbb{E}\sup_{x\in B_R}\|\nabla_x u_0(X_t^{-1}(x))\|^p \leqslant \|\nabla u_0\|^p_{L^\infty(\mathbb{R}^d)} \sup_{0\leqslant t\leqslant T}\mathbb{E}\sup_{x\in B_R}\|\nabla X_t^{-1}(x)\|^p < +\infty, \tag{5.160}$$

i.e. (1.20) holds true.

Given that $\mathrm{div} b = 0$, for every $\varphi\in\mathcal{C}_0^\infty(\mathbb{R}^d)$ and every $t\in[0,T]$, by virtue of Euler's identity: $\det(\nabla X_t(x)) = \exp(\int_0^t \mathrm{div} b(\tau, X_\tau(x))d\tau)$, then

$$\int_{\mathbb{R}^d} u_0(X_t^{-1}(x))\varphi(x)dx = \int_{\mathbb{R}^d} u_0(x)\varphi(X_t(x))dx. \tag{5.161}$$

In light of Itô's formula, it follows that

$$\begin{aligned}\int_{\mathbb{R}^d} u_0(x)\varphi(X_t(x))dx &= \int_{\mathbb{R}^d} u_0(x)\varphi(x)dx + \int_0^t d\tau\int_{\mathbb{R}^d} u_0(x)\nabla\varphi(X_\tau(x))\cdot b(\tau,X_\tau(x))dx\\ &\quad + \sum_{i=1}^d\int_0^t \circ dW_{i,\tau}\int_{\mathbb{R}^d} u_0(x)\partial_{x_i}\varphi(X_\tau(x))dx\\ &= \int_{\mathbb{R}^d} u_0(x)\varphi(x)dx + \int_0^t d\tau\int_{\mathbb{R}^d} u(\tau,x)b(\tau,x)\cdot\nabla\varphi(x)dx\\ &\quad + \sum_{i=1}^d\int_0^t \circ dW_{i,\tau}\int_{\mathbb{R}^d} u(\tau,x)\partial_{x_i}\varphi(x)dx, \quad \mathbb{P}-a.s..\end{aligned} \tag{5.162}$$

If one uses the integration by parts formula for the second integral in the second identity of (5.162), we derive (1.21). By (1.21), $\int_{\mathbb{R}^d} u(t,x)\varphi(x)dx$ has a continuous modification which is an $\mathcal{F}_t$-semimartingale. Thus $u(t,x) = u_0(X_t^{-1}(x))$ is a stochastic strong solution of (1.19).

It remains to show the uniqueness. As the equation is linear, it suffices to prove $u \equiv 0$, a.s. if the initial data vanishes. Let b_n, X^n and $(X^n)^{-1}$ be given in (iii) of Theorem 1.2 with $s = 0$. With the help of (1.12), then

$$\lim_{n\rightarrow+\infty} \sup_{x\in\mathbb{R}^d} \sup_{0\leqslant t\leqslant T} \mathbb{E}|(X_t^n)^{-1}(x) - X_t^{-1}(x)|^p = 0, \quad \forall\, p \geqslant 2. \tag{5.163}$$

Since b_n is smooth with respect to the spatial variables, by virtue of the characteristic lines and Itô's formula, for almost everywhere $(t,x) \in (0,T]\times\mathbb{R}^d$ and every $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $\varphi((X_t^n)^{-1}(x))$ satisfies

$$d\varphi((X_t^n)^{-1}(x)) = -b_n(t,x)\cdot\nabla_x\varphi((X_t^n)^{-1}(x))dt - \sum_{i=1}^d \partial_{x_i}\varphi((X_t^n)^{-1}(x))\circ dW_{i,t}, \quad \mathbb{P}-a.s.. \tag{5.164}$$

As $(X_t^n)^{-1}(x)$ is continuous in (t,x) and (5.163) holds, there is a (unlabelled) subsequence $\varphi((X_t^n)^{-1})$ such that $\varphi((X_t^n)^{-1}(\cdot,\omega)) \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ uniformly in $t\in[0,T]$ and $n \geqslant 1$ for a.s. $\omega\in\Omega$. By (1.21), (5.164) and Itô's formula, for every $t\in[0,T]$, then

$$\begin{aligned}
&\int_{\mathbb{R}^d} u(t,X_t^n(x))\varphi(x)dx = \int_{\mathbb{R}^d} u(t,x)\varphi((X_t^n)^{-1}(x))dx \\
&= -\int_0^t d\tau \int_{\mathbb{R}^d} u(\tau,x)b_n(\tau,x)\cdot\nabla_x\varphi((X_\tau^n)^{-1}(x))dx \\
&\quad -\int_0^t d\tau \int_{\mathbb{R}^d} b(\tau,x)\cdot\nabla u(\tau,x)\varphi((X_\tau^n)^{-1}(x))dx \\
&= \int_0^t d\tau \int_{\mathbb{R}^d} [b_n(\tau,x)-b(\tau,x)]\cdot\nabla u(\tau,x)\varphi((X_\tau^n)^{-1}(x))dx, \quad \mathbb{P}-a.s.,
\end{aligned} \tag{5.165}$$

where in the first identity we have used $\mathrm{div} b_n = 0$, in the third identity we used $\mathrm{div} b = 0$ and the integration by parts formula.

By (2.86), (5.163) and the Lebesgue dominated convergence theorem, if one takes the limit as $n\rightarrow+\infty$ in (5.165), it leads to

$$\int_{\mathbb{R}^d} u(t,X_t(x))\varphi(x)dx = 0, \quad \mathbb{P}-a.s.,$$

which implies $u(t,X_t(x)) = 0$ for almost everywhere $x\in\mathbb{R}^d$ and almost all $\omega\in\Omega$. Since $X_t(x)$ is a stochastic flow of homeomorphisms associated of (1.1) with $s = 0$, we have $u(t,x) = 0$ for almost everywhere $x\in\mathbb{R}^d$ and a.s. $\omega\in\Omega$. □

Conflicts of Interest

The authors declare that they have no competing interests.

Data Availability Statements

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

[1] S. Attanasio, Stochastic flows of diffeomorphisms for one-dimensional SDE with discontinuous drift, Electron. Commun. Probab. **15** (13) (2010) 213–226.

[2] K. Bahlali, S. Mouchtabih, L. Tangpi, Quadratic transportation inequalities for SDEs with measurable drift, Proc. Amer. Math. Soc. **149** (2021) 3583–3596.

[3] L. Beck, F. Flandoli, M. Gubinelli, M. Maurelli, Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness, Electron. J. Probab. **24** (136) (2019) 1–72.

[4] Z.-Q. Chen, R. Song, X. Zhang, Stochastic flows for Lévy processes with Hölder drifts, Rev. Mat. Iberoam. **34** (4) (2018) 1755–1788.

[5] P.E. Chaudru de Raynal, S. Menozzi, E. Priola, Schauder estimates for drifted fractional operators in the supercritical case, J. Funct. Anal. **278** (8) (2020) 1–57.

[6] P. Constantin, G. Iyer, A stochastic Lagrangian representation of the three-dimensional incompressible Navier–Stokes equations, Comm. Pure Appl. Math. **61** (3) (2008) 330–345.

[7] A.M. Davie, Uniqueness of solutions of stochastic differential equations, Int. Math. Res. Not. **2007** (2007) 1–26.

[8] H. Djellout, A. Guilin, L. Wu, Transportation cost-information inequalities for random dynamical systems and diffusions, Ann. Probab. **32** (3) (2004) 2702–2732.

[9] E. Fedrizzi, F. Flandoli, Hölder flow and differentiability for SDEs with nonregular drift, Stoch. Anal. Appl. **31** (4) (2013) 708–736.

[10] E. Fedrizzi, F. Flandoli, Noise prevents singularities in linear transport equations, J. Funct. Anal. **264** (6) (2013) 1329–1354.

[11] F. Flandoli, M. Gubinelli, E. Priola, Well-posedness of the transport equation by stochastic perturbation, Invent. Math. **180** (1) (2010) 1–53.

[12] F. Flandoli, M. Gubinelli, E. Priola, Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift, Bull. Sci. Math. **134(4)** (2010) 405–422.

[13] L. Galeati, M. Gerencsér, Solution theory of fractional SDEs in complete subcritical regimes, Forum Math. Sigma **13** (12) (2025) 1–66.

[14] L. Hörmander, The Analysis of Linear Partial Differential Operators I, Grundlehren der mathematisschen Wissenschaften 256, Springer-Verlag, New York Berlin Heidelberg, 1989.

[15] K. Itô, On stochastic differential equation, Mem. Amer. Math. Soc. **4** (1951) 1–51.

[16] D. Kinzebulatov, K.R. Madou, Stochastic equations with time-dependent singular drift, J. Differ. Equ. **337** (2022) 255–293.

[17] N.V. Krylov, The Calderón–Zygmund theorem and parabolic equations in $L_p(\mathbb{R};\mathcal{C}^{2+\alpha})$, Ann. Scuola Norm. Sup. Pisa CI. Sci. **1** (4) (2002) 799–820.

[18] N.V. Krylov, Elliptic equations with VMO $a, b \in L_d$ and $c \in L_{d/2}$, Trans. Amer. Math. Soc. **374** (4) (2021) 2805–2822.

[19] N.V. Krylov, On stochastic equations with drift in L_d, Ann. Probab. **49** (5) (2021) 2371–2398.

[20] N.V. Krylov, On strong solutions of Ito's equations with $a \in W^1_d$ and $b \in L_d$, Ann. Probab. **49** (6) (2021) 3142–3167.

[21] N.V. Krylov, Elliptic equations in Sobolev spaces with Morrey drift and the zeroth-order coefficients, Trans. Amer. Math. Soc. **376** (10) (2023) 7329–7351.

[22] N.V. Krylov, On strong solutions of Ito's equations with $D\sigma$ and b in Morrey classes containing L_d, Ann. Probab. **51** (5) (2023) 1729–1751.

[23] N.V. Krylov, On weak and strong solutions of time inhomogeneous Itô's equations with VMO diffusion and Morrey drift, Stochastic Process. Appl. 179 (2025) 104505.

[24] N.V. Krylov, E. Priola, Elliptic and parabolic second-order PDEs with growing coefficients, Commun. Partial Differ. Equ. **35** (1) (2010) 1–22.

[25] N.V. Krylov, M. Röckner, Strong solutions to stochastic equations with singular time dependent drift, Probab. Theory Relat. Fields **131** (2) (2005) 154–196.

[26] H. Kunita, Stochastic Differential equations and Stochastic Flows of Diffeomorphisms, In: Ecole d'été de Probabilités de Saint-Flour, XII-1982. Lecture Notes in Math. **1097**, 143–303, Springer, Berlin, 1984.

[27] K. Kunita, Stochastic Flows and Stochastic Differential equations, Cambridge Studies in Advanced Mathematics, **24**, Cambridge University Press, Cambridge, 1990.

[28] L. Lorenzi, Optimal Schauder estimates for parabolic problems with data measurable with respect time, SIAM J. Math. Anal. **32** (3) (2000) 588–615.

[29] A. Mohammed, K. Nilssen, N. Proske, Sobolev differentiable stochastic flows for SDEs with singular coefficients: applications to the transport equation, Ann. Probab. **43** (1) (2015) 1535–1576.

[30] K. Nam, Stochastic differential equations with critical drifts, Stochastic Process. Appl. **130** (9) (2020) 5366–5393.

[31] F. Rezakhanlou, Regular flows for diffusions with rough drifts, arXiv:1405.5856v1, 2014.

[32] F. Rezakhanlou, Stochastically symplectic maps and their applications to the Navier–Stokes equation, Ann. Inst. H. Poincaré Anal. Non Linéaire **33** (1) (2016) 1–22.

[33] M. Röckner, G.H. Zhao, SDEs with critical time dependent drifts: strong solutions, arXiv:2012.04161, 2021.

[34] M. Röckner, G.H. Zhao, SDEs with critical time dependent drifts: weak solutions, Bernoulli **29** (1) (2023) 757–784.

[35] L. Silvestre, Hölder estimates for advection fractional-diffusion equations, Ann. Sc. Norm. Super. Pisa Cl. Sci. **11** (4) (2012) 843–855.

[36] E. Stein, Singular Integrals and Differentiaility Properties of Functions, Princeton University Press, Princeton, 1970.

[37] M. Talagrand, Transportation cost for Gaussian and other product measures, Geom. Funct. Anal. **6** (1996) 587–600.

[38] Y.Q. Suo, C.G. Yuan, S.Q. Zhang, Transportation cost inequalities for SDEs with irregular drifts, Stochastic Process. Appl. **144** (2022) 288–311.

[39] R.R. Tian, L. Ding, J.L. Wei, Strong solutions of stochastic differential equations with square integrable drift, Bull. Sci. Math. **174** (2022) 1–31.

[40] A.J. Veretennikov, On the strong solutions of stochastic differential equations, Theory Probab. Appl. **24** (2) (1979) 354–366.

[41] F.Y. Wang, X.C. Zhang, Degenerate SDE with Hölder–Dini drift and non-Lipschitz noise coefficient, SIAM J. Math. Anal. **48** (3) (2016) 2189–2226.

[42] J.L. Wei, G.Y. Lv, A Kolmogorov type theorem for stochastic fields, Stoch. Anal. Appl. **39** (6) (2021) 1009–1024.

[43] J.L. Wei, J.Q. Duan, H.J. Gao, G.Y. Lv, Stochastic regularization for transport equations, Stoch. Partial Differ. Equ., Anal. Computat. **9** (1) (2021) 105–141.

[44] J.L. Wei, G.Y. Lv, W. Wang, Stochastic transport equation with bounded and Dini continuous drift, J. Differ. Equ. **323** (5) (2022) 359–403.

[45] J.L. Wei, G.Y. Lv, J.-L. Wu, Stochastic differential equations with critically irregular drift coefficients, J. Differ. Equ. **371** (2023) 1–30.

[46] J.L. Wei, W. Wang, G.Y. Lv, J.Q. Duan, A new maximal regularity for parabolic equations and an application, arXiv:2411.13266.

[47] P. Xia, L. Xie, X. Zhang, G. Zhao, $L^q(L^p)$-theory of stochastic differential equations, Stochastic Process. Appl. **130** (8) (2020) 5188–5211.

[48] S.Q. Zhang, C.G. Yuan, A Zvonkin's transformation for stochastic differential equations with singular drift and applications, J. Differ. Equ. **297** (2021) 277–319.

[49] X. Zhang, Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients, Stochastic Process. Appl. **115** (11) (2005) 1805–1818.

[50] X. Zhang, Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients, Electron. J. Probab. **16** (13) (2011) 1096–1116.