Rev. Mat. Iberoam. 38 (2022), no. 3, 947–980
DOI 10.4171/RMI/1313

SDEs with random and irregular coefficients

Guohuan Zhao

Abstract. We consider Itô uniformly nondegenerate equations with random coefficients. When the coefficients satisfy some low regularity assumptions with respect to the spatial variables and Malliavin differentiability assumptions on the sample points, the unique solvability of singular SDEs is proved by solving backward stochastic Kolmogorov equations and utilizing a modified Zvonkin type transformation.

1. Introduction

The main purpose of this work is to study the well-posedness of stochastic differential equations (SDEs) with random and irregular coefficients. More precisely, we are concerned with the following SDE in $\mathbb{R}^n$:

$$X_t(\omega) = X_0(\omega) + \int_0^t \sigma_s(X_s,\omega)\,\mathrm{d}W_s(\omega) + \int_0^t b_s(X_s,\omega)\,\mathrm{d}s. \tag{1.1}$$

Here $\{W_t\}_{t\in[0,1]}$ is a d-dimensional Brownian motion defined on a complete filtered probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbf{P})$, where $\mathcal{F}$ and $\mathcal{F}_t$ are generated by $\{W_s\}_{s\in[0,1]}$ and $\{W_s\}_{s\in[0,t]}$, respectively. The coefficients $\sigma:\mathbb{R}^n\times[0,1]\times\Omega\to\mathbb{R}^n\otimes\mathbb{R}^d$ and $b:\mathbb{R}^n\times[0,1]\times\Omega\to\mathbb{R}^n$ are $\mathcal{B}\times\mathcal{P}$-measurable, where $\mathcal{B}$ denotes the Borel algebra on $\mathbb{R}^n$ and $\mathcal{P}$ stands for the collection of all the progressively measurable sets on $[0,1]\times\Omega$.

In the past half century, a great deal of mathematical effort in stochastic analysis has been devoted to the study of the existence, uniqueness and regularity properties of strong solutions to Itô uniformly nondegenerate stochastic equations with *deterministic* and irregular drifts. When $\nabla\sigma\in L^{2d}_{\mathrm{loc}}$ and b is bounded, Veretennikov [19] proved the strong existence and uniqueness of solutions to the SDE (1.1) by developing a original idea proposed by Zvonkin in [26]. In the case where $\sigma=\mathbb{I}$ and $b\in L^q_tL^p_x$ with $n/p+2/q<1$, using Girsanov's transformation and $L^q_tL^p_x$-estimates for parabolic equations, Krylov–Röckner [11] obtained the existence and uniqueness of strong solutions to (1.1). After that, a lot of works investigated properties of the strong solution to (1.1) with singular drifts. Among all, we mention that the Hölder continuity of the stochastic flow was proved by Fedrizzi and Flandoli in [5], provided that the coefficients meet the same condition as

2020 Mathematics Subject Classification: 60H07, 60H10, 60H15.
Keywords: Backward SPDEs, Malliavin calculus, Schauder estimate, singular SDEs.

in [11]. When b is bounded, Menoukeu et al. [15] obtained the weak differentiability of the stochastic flow and the Malliavin differentiability of X_t with respect to the sample ω by using Malliavin's calculus. Zhang [21] extended Veretennikov's unique strong solvability result to the case where $\nabla\sigma, b \in L_t^q L_x^p$ with $n/p + 2/q < 1$. Under similar conditions, the regularity of strong solutions with respect to the initial data and sample point was also shown in [22] and [20]. For more recent results, we refer the reader to [12] and [16]. We also note that martingale problems and stochastic Lagrangian flows corresponding to (1.1) were studied by many researchers, among which we quote [2, 17, 23–25].

The well-posedness and regularity of strong solutions to SDEs with singular coefficients is not only a fundamental theoretical problem, but also has a wide range of applications in many mathematical and physical problems. For instance, in the remarkable paper [7], Flandoli, Gubinelli and Priola studied the following linear stochastic transport equation (see also [6]):

$$\partial_t u + b \cdot \nabla u + \nabla u \circ \frac{\mathrm{d}W_t}{\mathrm{d}t} = 0, \quad u_0 = \varphi, \tag{1.2}$$

where $b: \mathbb{R}^n \times [0,1] \to \mathbb{R}^n$ is *deterministic*. Using the stochastic flow of the corresponding SDE (or stochastic characteristics), they proved the existence and uniqueness for the above equation in the L^∞-setting, provided that the drift b is α-Hölder continuous uniform in t and the divergence of b satisfies some integrability conditions. However, as mentioned in [7], one of the major obstacles to extending the regularization by noise phenomenon to the case where b is random is the fact that, even when b is Hölder continuous in x, the stochastic characteristics corresponding to (1.2) may not uniquely exist. Below is a simple but typical example.

Example 1.1. Let $d = n = 1$. Assume $\sigma = 1$ and

$$b_t(x) = \sqrt{|x - W_t|} \wedge 1, \quad X_0 = 0.$$

Denote $Y_t := X_t - W_t$. Then Y_t satisfies the following random ODE:

$$\mathrm{d}Y_t(\omega) = b_t(Y_t(\omega) + W_t(\omega), \omega)\,\mathrm{d}t = \big(\sqrt{|Y_t(\omega)|} \wedge 1\big)\,\mathrm{d}t, \quad Y_0 = 0.$$

One can verify that $y_t^{(1)} \equiv 0$ and $y_t^{(2)} = t^2/4$ are two solutions of the above ODE, which implies $X_t^{(1)} = W_t$ and $X_t^{(2)} = t^2/4 + W_t$ are two $\mathcal{F}_t$-adapted solutions to the equation

$$X_t = \int_0^t b_s(X_s)\,\mathrm{d}s + W_t, \quad t \in [0,1].$$

The above example proves that the nondegeneracy of the noise and the uniform Hölder continuity of $b_t(\cdot, \omega)$ are insufficient to guarantee the well-posedness of (1.1). To the best of our knowledge, there is no much literature addressing this issue so far. The main work before this paper is Duboscq–Réveillac [4], which studies the stochastic regularization effects of diffusions with random drift coefficients on random functions. After adding some Malliavin differentiability conditions on b and f, the authors extended the boundedness of time average of a deterministic function f depending on a diffusion process X with deterministic drift coefficient b to random mappings f and b by investigating the backward stochastic Kolmogorov equation (1.5) ($a \equiv \mathbb{I}$) in some L^p-type space. However, their work misses some important requirements, in particular because it asks for a specific form of Malliavin derivative for the drift, and in certain situations, also $W^{1,p}$ regularity for the drift with respect to x, which makes the results not so strong. This paper attempts to make some progress in this direction. Roughly speaking, our main result, Theorem 1.2, shows that if the noise is additive and nondegenerate, and if b is Hölder in x, the well-posedness of the Itô equation (1.1) is guaranteed when $Db_t(x)$, the Malliavin derivative of $b_t(x)$, also satisfies a Hölder continuity assumption with respect to x.

With a slight abuse of notation, we shall abbreviate $L^p(\Omega, \mathcal{F}, \mathbf{P}; \mathbb{R}^m)$ as $L^p(\Omega)$; the integer m may take different values in different places. Our main result is:

Theorem 1.2. *Let $\alpha \in (0,1)$, $p > n/\alpha$, $\Lambda > 1$, $\Delta := \{(s,t) \in [0,1]^2 : 0 \leqslant s \leqslant t \leqslant 1\}$ and let D be the Malliavin derivative operator. Assume that σ and b are $\mathcal{B} \times \mathcal{P}$ measurable. Then equation* (1.1) *admits a unique solution if σ and b satisfy the following assumptions*:

(i) *for almost all $\omega \in \Omega$, $\sigma(\omega)$ and $b(\omega)$ are bounded, and for all $x, y \in \mathbb{R}^n$, $t \in [0,1]$,*

$$|b_t(x,\omega) - b_t(y,\omega)| \leqslant \Lambda |x-y|^\alpha, \quad |\sigma_t(x,\omega) - \sigma_t(y,\omega)| \leqslant \Lambda |x-y|;$$

(ii) *for almost all $\omega \in \Omega$ and all $(x,t) \in \mathbb{R}^n \times [0,1]$,*

$$\Lambda^{-1}|\xi|^2 \leqslant \frac{1}{2}\sigma_t^{ik}\sigma_t^{jk}(x,\omega)\xi_i\xi_j \leqslant \Lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^d;$$

(iii) *for each $(x,t) \in \mathbb{R}^n \times [0,1]$, $\sigma_t(x), b_t(x)$ are Malliavin differentiable and the random fields $D_s\sigma_t(x)$ and $D_s b_t(x)$ have continuous versions, as maps from $\mathbb{R}^n \times \Delta$ to $L^{2p}(\Omega)$, such that*

$$\sup_{(s,t)\in\Delta} \big(\|D_s\sigma_t\|_{C^\alpha(\mathbb{R}^n;L^{2p}(\Omega))} + \|D_s b_t\|_{C^\alpha(\mathbb{R}^n;L^{2p}(\Omega))}\big) \leqslant \Lambda. \tag{1.3}$$

We give an example of b meeting the conditions in Theorem 1.2.

Example 1.3. Let $n = d = 1$, $\alpha \in (0,1)$, $p > 1/\alpha$. Assume $\bar{b}\colon [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ is a bounded function satisfying

$$\big(|\bar{b}_t(x,y) - \bar{b}_t(x',y)| + |\partial_y \bar{b}_t(x,y) - \partial_y \bar{b}_t(x',y)|\big) \leqslant C|x-x'|^\alpha$$

for all $x, x', y \in \mathbb{R}^n$ and $t \in [0,1]$, and

$$b_t(x,\omega) := \bar{b}_t\Big(x, \int_0^t h_r(\omega)\,\mathrm{d}W_r(\omega)\Big).$$

Here h is an adapted process satisfying

$$\sup_{s\in[0,1]} \mathbf{E}\Big(|h_s|^{2p} + \int_0^1 |D_s h_r|^{2p}\,\mathrm{d}r\Big) < \infty.$$

Noting that

$$D_s b_t(x) = \partial_y \bar{b}_t\Big(x, \int_0^t h_r\,\mathrm{d}W_r\Big)\Big(\int_s^t D_s h_r\,\mathrm{d}W_r + h_s\Big)\mathbf{1}_\Delta(s,t),$$

by the Burkholder–Davis–Gundy inequality, one sees that

$$\sup_{t\in[0,1];\omega\in\Omega}\|b_t(\cdot,\omega)\|_{C^\alpha(\mathbb{R})}+\sup_{(s,t)\in\Delta}\|D_s b_t\|_{C^\alpha(\mathbb{R};L^{2p}(\Omega))}$$

$$\leqslant C\Big[1+\sup_{s\in[0,1]}\mathbf{E}\Big(|h_s|^{2p}+\int_0^1|D_s h_r|^{2p}\,\mathrm{d}r\Big)^{1/2p}\Big]<\infty,$$

so b satisfies the conditions (i) and (iii) in Theorem 1.2.

Our approach to the study of the well-posedness of (1.1) shall use a modified Zvonkin transformation. Such kind of trick was first proposed in [26] for solving SDEs with deterministic and bounded coefficients. To explain our main idea, let us first give a brief introduction to Zvonkin's idea. Denote

$$a=\frac{1}{2}\sigma\sigma^*,\quad L_t u=a_t^{ij}\,\partial_{ij}u+b_t^i\,\partial_i u.$$

When a and b are deterministic, $a,b\in L_t^\infty C_x^\alpha$ and a is uniformly elliptic, so by Schauder's estimate, the following backward equation:

$$\partial_t u+L_t u=-b,\quad u_T(x)=0,$$

admits a unique solution $u\in L_t^\infty C_x^{2+\alpha}$ with $\partial_t u\in L_t^\infty C_x^\alpha$. Moreover, if T is sufficiently small, the map $x\mapsto\phi_t(x):=x+u_t(x)$ is a C^2-homeomorphism. Assuming that X_t solves (1.1), by Itô's formula, $Y_t:=\phi_t(X_t)$ satisfies a new SDE with Lipschitz continuous coefficients. Thus, the strong uniqueness of the solution to the original equation is given by the one of the new equation. In the case where σ and b are progressive measurable and

$$\operatorname{ess\,sup}_{\omega\in\Omega}\big(\|\sigma(\omega)\|_{L_t^\infty C_x^\alpha}+\|b(\omega)\|_{L_t^\infty C_x^\alpha}\big)<\infty,$$

and thanks to the classic Schauder estimate, one can solve pointwisely the backward equation

$$\partial_t w+L_t w+f=0,\quad w_T(x)=0. \tag{1.4}$$

Moreover, w satisfies

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\big(\|w(\omega)\|_{L_t^\infty C_x^{2+\alpha}}+\|\partial_t w(\omega)\|_{L_t^\infty C_x^\alpha}\big)\leqslant C\operatorname*{ess\,sup}_{\omega\in\Omega}\|f(\omega)\|_{L_t^\infty C_x^\alpha}.$$

However, in this case, for each $x\in\mathbb{R}^d$, the process $w_\cdot(\cdot,x)\colon(t,\omega)\mapsto w_t(x,\omega)$ is non-adapted, so one cannot apply the Itô–Wentzell formula as in the deterministic case. A very natural way to overcome this difficulty is to consider the function $u_t:=\mathbf{E}(w_t|\mathcal{F}_t)$ instead of w_t. Formally, u_t satisfies the following backward stochastic Kolmogorov equation (see Lemma 3.1):

$$\mathrm{d}u_t+(L_t u_t+f_t)\,\mathrm{d}t=v_t\cdot\mathrm{d}W_t,\quad u_T(x)=0. \tag{1.5}$$

Let us mention that a more general class of semi-linear equations including (1.5) was already studied by Du–Qiu–Tang [3] in L^p-spaces and also by Tang–Wei [18] in Hölder spaces. However, the main obstacle for applying their result for our purposes is that one can only expect that the vector field v is in some L^p (or C^α) space, which is far from what is needed to apply the Itô–Wentzell formula (see Lemma A.7). Inspired by [4] and [26], in this paper we prove a $C^{2+\alpha}$ type estimate (Theorem 3.5) for (u,v), provided that the coefficients satisfy some Malliavin differentiability conditions. To achieve this purpose, we first extend the classic Schauder estimate to random PDEs with Banach variables. Such kind of extension gives a $C^{2+\alpha}$ estimate for u, as well as a C^{α} estimate for v (see Lemma 3.1). The main result of this paper is Theorem 3.5, where we give a $C^{2+\alpha}$ estimate for v, provided that the Malliavin derivatives of the coefficients satisfy (1.3). To us, such kind of result is new and intriguing. With such a regularity estimate in hand, we then use a modified Itô–Wentzell's formula and a Zvonkin type transformation to prove the well-posedness of (1.1). We believe our results have the potential to be applied to stochastic transport equations with random coefficients and some other nonlinear stochastic PDEs.

This paper is organized as follows. In Section 2, we investigate a random Banach-valued non-adapted Kolmogorov equation, and prove its well-posedness in some Hölder type spaces. In Section 3, we study the solvability of the backward stochastic Kolmogorov equation (1.5) in some $C^{2+\alpha}$ space. Our main result is proved in Section 4. An Itô–Wenzell type formula and some auxiliary lemmas used in our main proofs are presented in the Appendix.

2. Schauder estimates for random Banach-valued PDEs

In this section, we give a self-contained proof of a Schauder type estimate for random Banach-valued parabolic PDEs by using the Littlewood–Paley decomposition.

Let $T \in (0,1]$, let D be a domain of $\mathbb{R}^n$, let $D_T = D \times [0,T]$, and let $\mathcal{B}$ be a real Banach space. For $\alpha \in (0,1)$ and a strongly continuous function $g\colon D \to \mathcal{B}$, we define

$$\|g\|_{0;D} := \sup_{x\in D} |g(x)|_{\mathcal{B}}, \quad [g]_{\alpha;D} := \sup_{x,y\in D} \frac{|g(x)-g(y)|_{\mathcal{B}}}{|x-y|^{\alpha}}.$$

For $k \in \mathbb{N}$,

$$\|g\|_{C^{k+\alpha}(D;\mathcal{B})} := \sum_{i=0}^{k} \|\nabla^i g\|_{0;D} + [\nabla^k g]_{\alpha;D}.$$

Here and below, all the derivatives of an $\mathcal{B}$-valued function are defined with respect to the spatial variable in the strong sense, namely, ∇g is the unique map from $\mathbb{R}^n$ to $\mathcal{L}(\mathbb{R}^n;\mathcal{B})$ such that $\lim_{|h|\to 0} |g(x+h) - g(x) - \nabla g(x)\cdot h|_{\mathcal{B}} = 0$. For any $\beta \geqslant 0$, the space $C^{\beta,0}_{x,t}(D_T;\mathcal{B})$ consists of all continuous functions $f\colon D_T \to \mathcal{B}$ such that

$$\|f\|_{C^{\beta,0}_{x,t}(D_T;\mathcal{B})} := \sup_{t\in[0,T]} \|f(t)\|_{C^{\beta}(D;\mathcal{B})} < \infty.$$

Below we always denote $Q_T = \mathbb{R}^n \times [0,T]$ and $Q = Q_1$. If there is no confusion with the time parameter T and the underlying Banach space $\mathcal{B}$, we simply write C^{β} and $C^{\beta}_{x,t}$ instead of $C^{\beta}(\mathbb{R}^n;\mathcal{B})$ and $C^{\beta,0}_{x,t}(Q_T;\mathcal{B})$, respectively.

2.1. Littlewood–Paley decomposition

Let $\mathcal{S}(\mathbb{R}^n)$ be the Schwartz space of all rapidly decreasing complex valued functions on $\mathbb{R}^n$, and let $\mathcal{S}'(\mathbb{R}^n)$ be the dual space of $\mathcal{S}(\mathbb{R}^n)$ (the tempered distributions space). Given

$f \in \mathcal{S}(\mathbb{R}^n)$, the Fourier transform and the inverse Fourier transform of f are defined by

$$\mathcal{F}(f)(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x, \quad \mathcal{F}^{-1} f(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x.$$

Let $\chi : \mathbb{R}^n \to [0, 1]$ be a smooth radial function such that

$$\chi(\xi) = 1, \ |\xi| \leqslant 1; \quad \chi(\xi) = 0, \quad |\xi| \geqslant 3/2.$$

Define

$$\varphi(\xi) := \chi(\xi) - \chi(2\xi), \quad \varphi_{-1}(\cdot) := \chi(2\cdot), \quad \varphi_j(\cdot) := \varphi(2^{-j}\cdot)\,(j = 0, 1, 2, \dots).$$

It is easy to see that $\varphi \geqslant 0$ and $\operatorname{supp}\varphi \subset B_{3/2} \setminus B_{1/2}$, and formally,

$$\sum_{j=-1}^{k} \varphi_j(\xi) = \chi(2^{-k}\xi) \xrightarrow{k\uparrow\infty} 1.$$

In particular, if $|j - j'| \geqslant 2$, then

$$\operatorname{supp}\varphi(2^{-j}\cdot) \cap \operatorname{supp}\varphi(2^{-j'}\cdot) = \varnothing.$$

Let $\widetilde{\varphi}$ be another smooth radial function such that $\operatorname{supp}\widetilde{\varphi} \in B_{7/4}\backslash B_{1/4}$ and $\widetilde{\varphi}(x) = 1$ for all $x \in B_{3/2}\backslash B_{1/2}$. Denote

$$h_j := \mathcal{F}^{-1}(\varphi_j), \quad \widetilde{h}_j := \mathcal{F}^{-1}(\widetilde{\varphi}_j).$$

For any $f \in L^1(\mathbb{R}^n; \mathcal{B}) + L^\infty(\mathbb{R}^n; \mathcal{B})$, define

$$\Delta_j f := \int_{\mathbb{R}^n} h_j(x-y) f(y)\,\mathrm{d}y, \quad \widetilde{\Delta}_j f := \int_{\mathbb{R}^n} \widetilde{h}_j(x-y) f(y)\,\mathrm{d}y.$$

2.2. A basic a priori estimate

Assume $(\Omega, \mathcal{F}, \mathbf{P})$ is a complete probability space, $\mathcal{H}$ is a real Hilbert spaces and $\mathcal{B} = L^p(\Omega, \mathcal{F}, \mathbf{P}; \mathcal{H})$ for some $p \geqslant 2$. Let a^{ij}, b^i and c be real-valued measurable functions on $Q \times \Omega$ and define

$$L_t := a_t^{ij}\partial_{ij} + b_t^i \partial_i + c_t.$$

Fix $T \in (0, 1]$. We first give a precise definition of solutions to the following $\mathcal{B}$-valued PDE:

$$\begin{cases} \partial_t w + L_t w + f = 0 & \text{in } Q_T^o, \\ w_T = 0 & \text{on } \mathbb{R}^n. \end{cases} \tag{2.1}$$

Definition 2.1. A function $w: Q_T \to \mathcal{B}$ is called a solution of (2.1) if

(1) for each $t \in [0, T]$, $w(t, \cdot)$ is a twice strongly differentiable function from $\mathbb{R}^n$ to $\mathcal{B}$;

(2) for each $x \in \mathbb{R}^n$, the process $w(\cdot, x)$ is absolutely continuous from $[0, T]$ to $\mathcal{B}$, and satisfies

$$w_t(x) = \int_t^T (L_s w_s + f_s)(x)\,\mathrm{d}s.$$

In order to study the solvability of (2.1), we need the following.

Assumption 2.2. *The map* $(x,t,\omega)\mapsto(a_t(x,\omega),b_t(x,\omega),c_t(x,\omega),f_t(x,\omega))$ *is* $\mathcal{B}(Q)\times\mathcal{F}$ *measurable and there are constants* $\alpha\in(0,1)$ *and* $\Lambda>1$ *such that for almost all* $\omega\in\Omega$,

$$\|a^{ij}(\omega)\|_{C^{\alpha,0}_{x,t}}+\|b^i(\omega)\|_{C^{\alpha,0}_{x,t}}+\|c(\omega)\|_{C^{\alpha,0}_{x,t}}\leqslant\Lambda, \tag{$\mathbf{H}_1$}$$

and

$$\Lambda^{-1}|\xi|^2\leqslant a^{ij}(\omega)\xi_i\xi_j\leqslant\Lambda|\xi|^2. \tag{$\mathbf{H}_2$}$$

Our main result in this section is the following.

Theorem 2.3. *Under Assumption 2.2, for any* $f\in C^\alpha_{x,t}$, *equation* (2.1) *admits a unique solution* w *in* $C^{2+\alpha}_{x,t}$. *Moreover,*

$$\|\partial_t w\|_{C^\alpha_{x,t}}+\|w\|_{C^{2+\alpha}_{x,t}}+T^{-1}\|w\|_{C^0_{x,t}}\leqslant C\|f\|_{C^\alpha_{x,t}}, \tag{2.2}$$

where C *only depends on* n, p, α *and* Λ.

As in the proof for the classic Schauder estimate, we first consider the case $a_t(x,\omega)=a_t(\omega)$ and $b=c=0$. Define

$$A_{t,s}:=\int_t^s a(r)\,\mathrm{d}r,\qquad p^a_{t,s}(x):=(\det 4\pi A_{t,s})^{-1/2}\exp(-\langle x,A^{-1}_{t,s}x\rangle)$$

and

$$P^a_{t,s}f(x):=\int_{\mathbb{R}^n}p^a_{t,s}(x-y)f(y)\,\mathrm{d}y.$$

Lemma 2.4. *Let* $T\in(0,1]$, $\alpha\in(0,1)$. *Assume* a *is* x-*independent and satisfies* ($\mathbf{H}_2$). *For any* $f\in C^\alpha_{x,t}$, *the function* $w_t(x)=\int_t^T P^a_{t,s}f_s(x)\,\mathrm{d}s$ *is the unique function in* $C^{2+\alpha}_{x,t}$ *satisfying*

$$w_t=\int_t^T(a^{ij}_s\partial_{ij}w_s+f_s)\,\mathrm{d}s. \tag{2.3}$$

Moreover, there is a constant C, *that only depends on* n, α, p *and* Λ, *such that*

$$\|\partial_t w\|_{C^\alpha_{x,t}}+\|w\|_{C^{2+\alpha}_{x,t}}+T^{-1}\|w\|_{C^\alpha_{x,t}}\leqslant C\|f\|_{C^\alpha_{x,t}}. \tag{2.4}$$

Proof. We first prove that the map w defined above satisfies (2.4) by using Littlewood–Paley decompositions. Recall that $\mathcal{B}=L^p(\Omega,\mathcal{F},\mathbf{P};\mathcal{H})$. For any $g\in L^1(\mathbb{R}^n;\mathcal{B})+L^\infty(\mathbb{R}^n;\mathcal{B})$, by Minkowski's inequality, we have

$$\begin{aligned}\|(\Delta_j P^a_{t,s}g)(x)\|_{\mathcal{B}}&=\big(\mathbf{E}|(\Delta_j P^a_{t,s}g)(x)|^p_{\mathcal{H}}\big)^{1/p}=\big(\mathbf{E}|(P^a_{t,s}\widetilde{\Delta}_j\Delta_j g)(x)|^p_{\mathcal{H}}\big)^{1/p}\\&=\bigg[\int_\Omega\bigg|\int_{\mathbb{R}^n}(p^{a(\omega)}_{t,s}*\widetilde{h}_j)(y)\cdot\Delta_j g(x-y,\omega)\,\mathrm{d}y\bigg|^p_{\mathcal{H}}\mathbf{P}(\mathrm{d}\omega)\bigg]^{1/p}\\&\leqslant\int_{\mathbb{R}^n}\mathrm{d}y\bigg[\int_\Omega|p^{a(\omega)}_{t,s}*\widetilde{h}_j(y)|^p\cdot|\Delta_j g(x-y,\omega)|^p_{\mathcal{H}}\,\mathbf{P}(\mathrm{d}\omega)\bigg]^{1/p}\\&\leqslant\|\Delta_j g\|_0\int_{\mathbb{R}^n}\big[\operatorname{ess\,sup}_{\omega\in\Omega}|p^{a(\omega)}_{t,s}*\widetilde{h}_j(y)|\big]\,\mathrm{d}y.\end{aligned} \tag{2.5}$$

By ($\mathbf{H}_2$),

$$\int_{\mathbb{R}^n}\left[\operatorname{ess\,sup}_{\omega\in\Omega}|p_{t,s}^{a(\omega)}*\widetilde{h}_j(x)|\right]\mathrm{d}x\leqslant\Big\|\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}|p_{t,s}^{a}*\widetilde{h}_j(x)|\Big\|_{L_x^1}$$
$$=\int_{\mathbb{R}^n}\mathrm{d}x\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}\left|\int_{\mathbb{R}^n}p_{t,s}^{a}(x-y)\,2^{jn}\,\widetilde{h}_0(2^jy)\,\mathrm{d}y\right|$$
$$=\int_{\mathbb{R}^n}\mathrm{d}x\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}\left|\int_{\mathbb{R}^n}2^{jn}\,p_{t,s}^{2^{2j}a}(2^jx-z)\,\widetilde{h}_0(z)\,\mathrm{d}z\right|$$
$$=\int_{\mathbb{R}^n}\mathrm{d}x\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}\left|\int_{\mathbb{R}^n}p_{t,s}^{2^{2j}a}(x-z)\,\widetilde{h}_0(z)\,\mathrm{d}z\right|.$$

Noting that

$$\|f\|_{L^1}\leqslant C_{n,N}\|(1+|x|^{2N})f(x)\|_{L^\infty},\quad\forall N>n/2,$$

and

$$\mathcal{F}^{-1}(p_{t,s}^{a})(\xi)=\exp(-\langle\xi,A_{t,s}\,\xi\rangle),$$

we obtain

$$\int_{\mathbb{R}^n}\left[\operatorname{ess\,sup}_{\omega\in\Omega}|p_{t,s}^{a(\omega)}*\widetilde{h}_j(x)|\right]\mathrm{d}x\leqslant\int_{\mathbb{R}^n}\mathrm{d}x\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}\left|\int_{\mathbb{R}^n}p_{t,s}^{2^{2j}a}(x-z)\,\widetilde{h}_0(z)\,\mathrm{d}z\right|$$
$$\leqslant C\Big\|(1+|x|^{2N})\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}|p_{t,s}^{2^{2j}a}*\widetilde{h}_0|(x)\Big\|_{L_x^\infty}$$
$$=C\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}\Big\|(1+|x|^{2N})|p_{t,s}^{2^{2j}a}*\widetilde{h}_0|(x)\Big\|_{L_x^\infty}$$
$$\leqslant C\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}\Big\|(1+\Delta^N)[\mathcal{F}^{-1}(p_{t,s}^{2^{2j}a})\cdot\mathcal{F}^{-1}(\widetilde{h}_0)](\xi)\Big\|_{L_\xi^1}$$
$$=C\sup_{\mathbb{I}/\Lambda\leqslant a\leqslant\Lambda\mathbb{I}}\int_{B_{7/4}\backslash B_{1/4}}\left|(1+\Delta^N)[\exp(-2^{2j}\langle\xi,A_{t,s}\xi\rangle)\cdot\widetilde{\varphi}](\xi)\right|\,\mathrm{d}\xi.$$

Since $\sup_{|\alpha|=k}\partial^\alpha(\mathrm{e}^{a|\xi|^2})\leqslant C(1+|a|)^k(1+|\xi|)^k\mathrm{e}^{a|\xi|^2}$, we get

$$\int_{\mathbb{R}^n}\left[\operatorname{ess\,sup}_{\omega\in\Omega}|p_{t,s}^{a(\omega)}*\widetilde{h}_j(x)|\right]\mathrm{d}x\leqslant C\int_{1/4\leqslant|\xi|\leqslant7/4}[1+(\Lambda 2^{2j}(s-t))^{2N}]\exp[-2^{2j}(s-t)|\xi|^2/\Lambda]\,\mathrm{d}\xi.\tag{2.6}$$

Denote $\Lambda_j:=\Lambda 2^{2j}(s-t)$ and $\lambda_j:=\frac{1}{16}\Lambda^{-1}2^{2j}(s-t)$. Combining (2.5) and (2.6), we get

$$\|\Delta_jP_{t,s}^{a}g\|_0=\sup_{x\in\mathbb{R}^n}\|(\Delta_jP_{t,s}^{a}g)(x)\|_{\mathcal{B}}\leqslant C(1+\Lambda_j^{2N})\,\mathrm{e}^{-\lambda_j}\left|B_{7/4}\backslash B_{1/4}\right|\|\Delta_jg\|_0.$$

By Lemma A.1 and the elementary inequality

$$(1+\Lambda_j^{2N})\,\mathrm{e}^{-\lambda_j}\leqslant C_k(1\wedge[2^{2j}\cdot(s-t)]^{-k})\quad(\forall k\in\mathbb{N}),$$

we get

$$\|\Delta_j P^a_{t,s} g\|_0 \leqslant C\,2^{-j\alpha}\|g\|_\alpha(1+\Lambda_j^{2N})\,\mathrm{e}^{-\lambda_j} \leqslant C_k\,2^{-j\alpha}(1\wedge[2^{2j}\cdot(s-t)]^{-k})\|g\|_\alpha.$$

This yields

$$\|\Delta_j w_t\|_0 = \left\|\Delta_j \int_t^T P^a_{t,s} f_s \;\mathrm{d}s\right\|_0 \leqslant C\,2^{-j\alpha}\|f\|_{C^\alpha_{x,t}} \int_0^{T-t} (1\wedge 2^{-2jk} r^{-k})\,\mathrm{d}r.$$

If $t \geqslant T-2^{-2j}$, then

$$\|\Delta_j w_t\|_0 \leqslant C\,2^{-j\alpha}\|f\|_{C^\alpha_{x,t}}\cdot(T-t) \leqslant C\,2^{-j(2+\alpha)}\|f\|_{C^\alpha_{x,t}};$$

if $t < T-2^{-2j}$, by choosing $k=2$, then

$$\|\Delta_j w_t\|_0 \leqslant C\,2^{-j\alpha}\|f\|_{C^\alpha_{x,t}}\cdot\left(2^{-2j}+2^{-4j}\int_{2^{-2j}}^{T-t} s^{-2}\,\mathrm{d}s\right) \leqslant C\,2^{-j(2+\alpha)}\|f\|_{C^\alpha_{x,t}}.$$

Again using Lemma A.1, one sees that

$$\|w\|_{C^{2+\alpha}_{x,t}} \leqslant C \sup_{\substack{t\in[0,T]\\ j\geqslant -1}} \left(2^{-j(2+\alpha)}\|\Delta_j w_t\|_0\right) \leqslant C\|f\|_{C^\alpha_{x,t}}.$$

This completes the proof of (2.4). By basic calculations, one can verify that w satisfies (2.3). It remains to show that the w defined above is the unique solution to (2.1) in $C^{2+\alpha}_{x,t}$. Assume $\widetilde{w}\in C^{2+\alpha}_{x,t}$ is another function satisfying (2.3). Let $0\leqslant\varrho\in C^\infty_c(\mathbb{R}^n)$ be such that $\int\varrho=1$ and write $\varrho_\varepsilon(x)=\varepsilon^{-n}\varrho(x/\varepsilon)$. Define $v:=w-\widetilde{w}$ and $v^\varepsilon:=v*\varrho_\varepsilon$. For any $k>n/p$, $N>1$ and $\varepsilon\in(0,1)$, by the Sobolev embedding and Hölder's inequality,

$$\begin{aligned}
&\mathbf{E}\|v^\varepsilon_{t_1}-v^\varepsilon_{t_2}\|^p_{L^\infty(B_N;\mathcal{H})} = \mathbf{E}\sup_{\|h\|_{\mathcal{H}}=1}\|\langle v^\varepsilon_{t_1}-v^\varepsilon_{t_2},h\rangle\|^p_{L^\infty(B_N)}\\
&\leqslant CN^{kp-n}\,\mathbf{E}\sup_{\|h\|_{\mathcal{H}}=1}\|\langle v^\varepsilon_{t_1}-v^\varepsilon_{t_2},h\rangle\|^p_{W^{k,p}(B_N)}\\
&\leqslant CN^{kp-n}\,\mathbf{E}\sum_{i=0}^k\int_{B_N}\left|\nabla^i\int_{t_1}^{t_2}(a^{ij}\partial_{ij}v^\varepsilon_s)(x)\,\mathrm{d}s\right|^p_{\mathcal{H}}\mathrm{d}x\\
&\leqslant CN^{kp-n}|t_2-t_1|^{p-1}\sum_{i=2}^{k+2}\int_{B_N}\int_{t_1}^{t_2}\mathbf{E}\left|\int_{B_{N+1}}v_s(y)\nabla^i\varrho_\varepsilon(x-y)\,\mathrm{d}y\right|^p_{\mathcal{H}}\mathrm{d}s\;\mathrm{d}x\\
&\leqslant C_\varepsilon N^{kp+np-n}|t_2-t_1|^{p-1}\int_{t_1}^{t_2}\int_{B_{N+1}}\mathbf{E}|v_s(y)|^p_{\mathcal{H}}\,\mathrm{d}y \leqslant C_\varepsilon N^{(k+n)p}|t_2-t_1|^p\|v\|^p_{C^0_{x,t}}.
\end{aligned}$$

Due to Kolmogorov's criterion, for almost all $\omega\in\Omega$ and all $\varepsilon\in(0,1)$, $(x,t)\in Q_T$,

$$\|v^\varepsilon_t(x,\omega)\|_{\mathcal{H}} \leqslant C_\varepsilon(\omega)(1+|x|)^{k+n},$$

which means $v^\varepsilon_t(\cdot,\omega)$ satisfies a certain growth condition at infinity. On the other hand, by definition, for almost all $\omega\in\Omega$ and each $h\in\mathcal{H}$, the real valued function $\langle v^\varepsilon_t(\omega),h\rangle$ satisfies

$$\partial_t\langle v^\varepsilon_t(\omega),h\rangle + a^{ij}_t(\omega)\partial_{ij}\langle v^\varepsilon_t(\omega),h\rangle = 0,\quad \langle v^\varepsilon_T(\omega),h\rangle = 0.$$

Thus, we have $\langle v_t^\varepsilon(\omega), h\rangle \equiv 0$ (see [8], Chapter 7, p. 176), i.e., $w * \varrho_\varepsilon = \widetilde{w} * \varrho_\varepsilon$ a.s. So

$$\begin{aligned}\|w_t(x) - \widetilde{w}_t(x)\|_{\mathcal{B}} &\leqslant \lim_{\varepsilon\to 0}\|w_t(x) - (w * \varrho_\varepsilon)_t(x)\|_{\mathcal{B}} + \lim_{\varepsilon\to 0}\|\widetilde{w}_t(x) - (\widetilde{w} * \varrho_\varepsilon)_t(x)\|_{\mathcal{B}} \\ &= 0.\end{aligned}$$

This completes our proof. ∎

Proof of Theorem 2.3. Thanks to Lemma 2.4 and the method of continuity, we only need to prove the a priori estimate (2.2). Assume $w \in C^{2+\alpha}_{x,t}$ is a solution to (2.1). Let $\chi \in C_c^\infty(\mathbb{R}^d)$ be so that $\chi(x) = 1$ if $|x| \leqslant 1$ and $\chi(x) = 0$ if $|x| \geqslant 2$. Fix a number $\delta > 0$, which will be determined later. Define $\chi_\delta^z = \chi((x-z)/\delta)$. Then

$$\partial_t(w\chi_\delta^z) + L_t^z(w\chi_\delta^z) + (f\chi_\delta^z) + [\chi_\delta^z L_t w - L_t^z(w\chi_\delta^z)] = 0,$$

where $L_t^z w_t(x) := a_t^{ij}(z)\partial_{ij} w_t(x)$. Using ($\mathbf{H}_1$) and noting that

$$\chi_\delta^z L_t w - L_t^z(w\chi_\delta^z) = \chi_\delta^z(a^{ij} - a_z^{ij})\partial_{ij} w + (b^i\chi_\delta^z - 2a_z^{ij}\partial_j\chi_\delta^z)\partial_i w + (c\chi_\delta^z - a_z^{ij}\partial_{ij}\chi_\delta^z)w,$$

we have

$$\begin{aligned}\|[\chi_\delta^z L_t w - L_t^z(w\chi_\delta^z)]\|_{C^\alpha_{x,t}} &\leqslant C\delta^\alpha\|\nabla^2 w\|_{C^{\alpha,0}_{x,t}(B_{2\delta}(z)\times[0,T];\mathcal{B}))} \\ &\quad + C\big(\delta^{-\alpha}\|\nabla^2 w\|_{C^0_{x,t}} + \delta^{-1-\alpha}\|\nabla w\|_{C^\alpha_{x,t}} + \delta^{-2-\alpha}\|w\|_{C^\alpha_{x,t}}\big).\end{aligned} \tag{2.7}$$

Combining Lemma 2.4 and equation (2.7), we obtain that for any $\delta > 0$,

$$\begin{aligned}\sup_{z\in\mathbb{R}^n}\|w\|_{C^{2+\alpha,0}_{x,t}(B_{2\delta}(z)\times[0,T];\mathcal{B})} &\leqslant C_n\sup_{z\in\mathbb{R}^n}\|w\|_{C^{2+\alpha,0}_{x,t}(B_\delta(z)\times[0,T];\mathcal{B})} \\ &\leqslant C\sup_{z\in\mathbb{R}^n}\|w\chi_\delta^z\|_{C^{2+\alpha}_{x,t}} \leqslant C\sup_{z\in\mathbb{R}^n}\|f\chi_\delta^z + [\chi_\delta^z L_t w - L_t^z(w\chi_\delta^z)]\|_{C^\alpha_{x,t}} \\ &\leqslant C\delta^\alpha\sup_{z\in\mathbb{R}^n}\|w\|_{C^{2+\alpha,0}_{x,t}(B_{2\delta}(z)\times[0,T];\mathcal{B})} \\ &\quad + C\big(\delta^{-\alpha}\|\nabla^2 w\|_{C^0_{x,t}} + \delta^{-1-\alpha}\|\nabla w\|_{C^\alpha_{x,t}} + \delta^{-2-\alpha}\|w\|_{C^\alpha_{x,t}} + \delta^{-\alpha}\|f\|_{C^\alpha_{x,t}}\big).\end{aligned}$$

By choosing $\delta \in (0,1)$ sufficiently small such that $C\delta^\alpha \leqslant 1/2$, we obtain

$$\sup_{z\in\mathbb{R}^n}\|w\|_{C^{2+\alpha,0}_{x,t}(B_{2\delta}(z)\times[0,T];\mathcal{B})} \leqslant C_\delta\big(\|w\|_{C^2_{x,t}} + \|f\|_{C^\alpha_{x,t}}\big).$$

Using interpolation, we get

$$\begin{aligned}\|w\|_{C^{2+\alpha}_{x,t}} &\leqslant C_\delta\sup_{z\in\mathbb{R}^n}\|w\|_{C^{2+\alpha,0}_{x,t}(B_{2\delta}(z)\times[0,T];\mathcal{B})} \\ &\leqslant \varepsilon C_\delta\|w\|_{C^{2+\alpha}_{x,t}} + C_{\delta,\varepsilon}\big(\|w\|_{C^0_{x,t}} + \|f\|_{C^\alpha_{x,t}}\big), \quad \forall\varepsilon \in (0,1).\end{aligned}$$

By choosing ε small such that $\varepsilon C_\delta \leqslant 1/2$, we get

$$\|w\|_{C^{2+\alpha}_{x,t}} \leqslant C\big(\|w\|_{C^0_{x,t}} + \|f\|_{C^\alpha_{x,t}}\big). \tag{2.8}$$

It remains to show that $\|w\|_{C^0_{x,t}}$ can be controlled by $\|f\|_{C^\alpha_{x,t}}$. By Minkowski's inequality, for any $t \in [0, T]$,

$$\Big(\mathbf{E}\int_{B_r(x)} |w_t(y)|^p_{\mathcal{H}}\,\mathrm{d}y\Big)^{1/p} = \Big(\mathbf{E}\int_{B_r(x)} \Big|\int_t^T \partial_s w_s(y)\,\mathrm{d}s\Big|^p_{\mathcal{H}}\,\mathrm{d}y\Big)^{1/p}$$

$$= \Big(\mathbf{E}\int_{B_r(x)} \Big|\int_t^T (L_s w_s + f_s)(y)\,\mathrm{d}s\Big|^p_{\mathcal{H}}\,\mathrm{d}y\Big)^{1/p}$$

$$\leqslant C\int_t^T \Big(\mathbf{E}\int_{B_r(x)} |L_s w_s + f_s|^p_{\mathcal{H}}(y)\,\mathrm{d}y\Big)^{1/p}\mathrm{d}s \leqslant CTr^{n/p}(\|w\|_{C^2_{x,t}} + \|f\|_{C^0_{x,t}}). \tag{2.9}$$

One the other hand, by Hölder's inequality,

$$|w_t(x)|_{\mathcal{B}} \leqslant \fint_{B_r(x)} |w_t(x) - w_t(y)|_{\mathcal{B}}\,\mathrm{d}y + \fint_{B_r(x)} |w_t(y)|_{\mathcal{B}}\,\mathrm{d}y \tag{2.10}$$

$$\leqslant \|\nabla w\|_{C^0_{x,t}} \fint_{B_r(x)} |x-y|\,\mathrm{d}y + \fint_{B_r(x)} \Big(\mathbf{E}\int_{B_r(x)} |w_t(y)|^p_{\mathcal{H}}\Big)^{1/p}\,\mathrm{d}y$$

$$\leqslant r\|\nabla w\|_{C^0_{x,t}} + r^{-n/p}\Big(\mathbf{E}\int_{B_r(x)} |w_t(y)|^p_{\mathcal{H}}\,\mathrm{d}y\Big)^{1/p}.$$

Combining (2.9) and (2.10), we obtain

$$\|w\|_{C^0_{x,t}} \leqslant r\|\nabla w\|_{C^0_{x,t}} + CT(\|w\|_{C^2_{x,t}} + \|f\|_{C^0_{x,t}}).$$

Due to (2.8),

$$\|w\|_{C^2_{x,t}} \leqslant C(\|f\|_{C^\alpha_{x,t}} + \|w\|_{C^0_{x,t}}).$$

Combining the above two inequalities and letting $r \to 0$, we get

$$\|w\|_{C^0_{x,t}} \leqslant CT(\|w\|_{C^0_{x,t}} + \|f\|_{C^\alpha_{x,t}}).$$

By choosing T sufficiently small such that $CT \leqslant 1/2$, we get

$$\|w\|_{C^0_{x,t}} \leqslant CT\|f\|_{C^\alpha_{x,t}}.$$

This together with (2.8) implies that (2.2) holds for some small $T > 0$. The same estimate for arbitrary $T \in (0, 1]$ can be obtained by induction. ■

Remark 2.5. If f satisfies

$$\operatorname*{ess\,sup}_{\omega\in\Omega} \|f(\omega)\|_{C^{\alpha,0}_{x,t}(Q_T;\mathbb{R})} < \infty,$$

then (2.1) can be solved pointwisely, and by the classic Schauder estimate, it holds that

$$\operatorname*{ess\,sup}_{\omega\in\Omega} \big(\|\partial_t w(\omega)\|_{C^{\alpha,0}_{x,t}(Q_T;\mathbb{R})} + \|w(\omega)\|_{C^{2+\alpha,0}_{x,t}(Q_T;\mathbb{R})} + T^{-1}\|w(\omega)\|_{C^{\alpha,0}_{x,t}(Q_T;\mathbb{R})}\big)$$

$$\leqslant C \operatorname*{ess\,sup}_{\omega\in\Omega} \|f(\omega)\|_{C^{\alpha,0}_{x,t}(Q_T;\mathbb{R})}.$$

3. A Schauder estimate for a backward SPDE

In this section, we prove the solvability of (1.5) in the $C^{2+\alpha}_{x,t} \times C^{2+\alpha}_{x,t}$ space. Recall that W_t is a d-dimensional Brownian motion on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$, $\mathcal{F}_t = \sigma\{W_s : s \leqslant t\} \bigvee \mathcal{N}$ and $\mathcal{F} = \mathcal{F}_1$. For any $t \in [0,1]$ and $X \in \mathcal{F}$, we denote $\mathbf{E}^t X := \mathbf{E}(X|\mathcal{F}_t)$. Throughout this section, we always assume $T \in (0,1]$, and that $\mathscr{H}$ is a real Hilbert space, $\mathscr{B} = L^p(\Omega;\mathscr{H})$ for some $p \geqslant 2$ and $H = L^2([0,1];\mathbb{R}^d)$. With a slight abuse of notation, we write $L^p(\Omega) = L^p(\Omega;\mathbb{R}^m)$ for some integer $m \geqslant 1$ that can change in different places.

Lemma 3.1. *Let $\mathscr{H} = \mathbb{R}$. Assume that a, b, c are $\mathcal{B} \times \mathcal{P}$ measurable and satisfy Assumption 2.2. Then the BSPDE*

$$u_t(x) = \int_t^T (L_s u_s + f_s)(x)\,\mathrm{d}s - \int_t^T v_s(x)\cdot \mathrm{d}W_s$$

has an $\mathcal{F}_t$-adapted solution (u, v) in $C^{2+\alpha}_{x,t} \times C^\alpha(\mathbb{R}^n; L^p(\Omega; H)$ and $u_t = \mathbf{E}^t w_t$, where w is the solution to (2.1). Moreover,

$$\|u\|_{C^{2+\alpha}_{x,t}} + T^{-1}\|u\|_{C^0_{x,t}} + \|v\|_{C^\alpha(\mathbb{R}^n;L^p(\Omega;H))} \leqslant C\|f\|_{C^\alpha_{x,t}},$$

where C only depends on n, d, p, α and Λ.

Proof. Let w be the solution of (2.1). Define $u_t(x) = \mathbf{E}^t w_t(x)$. By Theorem 2.3 and Lemma A.4,

$$\|u\|_{C^{2+\alpha}_{x,t}} + T^{-1}\|u\|_{C^0_{x,t}} \leqslant C\|f\|_{C^\alpha_{x,t}}.$$

Since $a_t(x), b_t(x) \in \mathcal{F}_t$, by the definition of u, we have

$$\begin{aligned}
u_t(x) &= \mathbf{E}^t\Big\{\int_t^T [(L_s w_s + f_s)(x)]\,\mathrm{d}s\Big\}\\
&= \int_t^T \mathbf{E}^s[(L_s w_s + f_s)(x)]\,\mathrm{d}s + \Big\{\int_t^T \mathbf{E}^t[(L_s w_s + f_s)(x)]\,\mathrm{d}s - \int_t^T \mathbf{E}^s[(L_s w_s + f_s)(x)]\,\mathrm{d}s\Big\}\\
&= \int_t^T (L_s u_s + f_s)(x)\,\mathrm{d}s + m_t(x) - m_T(x).
\end{aligned}$$

Here,

$$m_t(x) = \int_t^T \mathbf{E}^t[(L_s w_s + f_s)(x)]\,\mathrm{d}s + \int_0^t \mathbf{E}^s[(L_s w_s + f_s)(x)]\,\mathrm{d}s \in \mathcal{F}_t. \tag{3.1}$$

For any $t \in [0,T]$, noting that

$$\begin{aligned}
\mathbf{E}^t m_T(x) &= \mathbf{E}^t \int_0^T \mathbf{E}^s[(L_s w_s + f_s)(x)]\,\mathrm{d}s\\
&= \mathbf{E}^t \int_0^t \mathbf{E}^s[(L_s w_s + f_s)(x)]\,\mathrm{d}s + \mathbf{E}^t \int_t^T \mathbf{E}^s[(L_s w_s + f_s)(x)]\,\mathrm{d}s\\
&= \int_0^t \mathbf{E}^s[(L_s w_s + f_s)(x)]\,\mathrm{d}s + \int_t^T \mathbf{E}^t[(L_s w_s + f_s)(x)]\,\mathrm{d}s = m_t(x),
\end{aligned}$$

we deduce that $m_\cdot(x)$ is a $\mathcal{F}_t$-martingale. By Theorem 2.3, (3.1) and Lemma A.4, one can see that $m \in C^\alpha_{x,t}$. Thanks to the martingale representation, there is an $\mathcal{F}_t$-adapted process $v_\cdot(x)$ such that

$$m_t(x) - m_0(x) = \int_0^t v_s(x) \cdot \mathrm{d}W_s.$$

Hence, we have

$$u_t(x) = \int_t^T (L_s u_s + f_s)(x)\,\mathrm{d}s - \int_t^T v_s(x) \cdot \mathrm{d}W_s,$$

i.e.,

$$u_t(x) = u_0(x) - \int_0^t (Lu_s + f_s)(x)\,\mathrm{d}s + \int_0^t v_s(x) \cdot \mathrm{d}W_s.$$

By (3.1) and the Burkholder–Davis–Gundy inequality, we obtain

$$\begin{aligned}
\mathbf{E}\Big[\Big(\int_0^T |v_t(x) - v_t(y)|^2\,\mathrm{d}t\Big)^{p/2}\Big] &= \mathbf{E}\langle m(x) - m(y)\rangle_T^{p/2} \leqslant C\,\mathbf{E}|m_T(x) - m_T(y)|^p \\
&= C\,\mathbf{E}\Big|\int_0^T \mathbf{E}^s[(L_s w_s + f_s)(x) - (L_s w_s + f_s)(y)]\,\mathrm{d}s\Big|^p \\
&\leqslant C \int_0^T \mathbf{E}\Big|\mathbf{E}^s\Big[(L_s w_s + f_s)(x) - (L_s w_s + f_s)(y)\Big]\Big|^p\,\mathrm{d}s \\
&\leqslant C \int_0^T \mathbf{E}|(L_s w_s + f_s)(x) - (L_s w_s + f_s)(y)|^p\,\mathrm{d}s \\
&\leqslant C|x - y|^{\alpha p}\left(\|w\|^p_{C^{2+\alpha}_{x,t}} + \|f\|^p_{C^\alpha_{x,t}}\right) \leqslant C|x - y|^{\alpha p}\|f\|^p_{C^\alpha_{x,t}},
\end{aligned}$$

which yields

$$\|v\|_{C^\alpha(\mathbb{R}^n;L^p(\Omega;H))} \leqslant C\|f\|_{C^\alpha_{x,t}},$$

and completes the proof. ∎

As we mentioned in the introduction, Zvonkin type transforms are an effective way to prove the well-posedness of SDEs with singular coefficients. However, the C^α-regularity of v in the spatial variable is not enough to apply this trick. So we need to get a better regularity estimate for v under some mild conditions. To achieve this goal, we start with some definitions and lemmas. Let $\mathcal{S}_b$ be a set of random variables of the form

$$F = f(\langle h_1, W\rangle, \dots, \langle h_m, W\rangle),$$

where $f \in C^\infty_b(\mathbb{R}^m)$, $h_i \in H$ and $\langle h_i, W\rangle := \int_0^1 h_s\,\mathrm{d}W_s$. We define the operator D on $\mathcal{S}_b$, with values in the set of H-valued random variables, by

$$DF = \sum_{i=1}^m \partial_i f(\langle h_1, W\rangle, \dots, \langle h_m, W\rangle) h_i.$$

For any $p \in [1, \infty)$, $\mathbb{D}^{1,p}$ is the closure of the set $\mathcal{S}_b$ with respect to the norm $\|F\|_{\mathbb{D}^{1,p}} := \|F\|_p + \|DF\|_{L^p(\Omega;H)}$.

Lemma 3.2. *Suppose* $\{y_t\}_{t\in[0,1]}$ *is a process* (*may not be adapted*) *on* $(\Omega, \mathbf{P}, \mathcal{F})$ *such that*

$$y_t = y_0 + \int_0^t \dot{y}_r \, \mathrm{d}r,$$

with $y_0 \in \mathbb{D}^{1,2}$ *and* $\dot{y} \in L^2([0,1]; \mathbb{D}^{1,2})$. *Then there exists a random field* $\{y_{s,t}\}_{(s,t)\in[0,1]^2}$ *such that for each* $t \in [0,1]$, $y_{\cdot,t} = D_\cdot y_t$ *in* $L^2(\Omega; H)$; *for each* $s \in [0,1]$, *the map* $[0,1] \ni t \mapsto y_{s,t} \in L^2(\Omega; \mathbb{R}^d)$ *is absolutely continuous, and*

$$\mathbf{E}^t y_t = \mathbf{E} y_0 + \int_0^t \mathbf{E}^s \dot{y}_s \, \mathrm{d}s + \int_0^t \mathbf{E}^s y_{s,s} \, \mathrm{d}W_s. \tag{3.2}$$

Proof. By our condition that $y_0 \in \mathbb{D}^{1,2}$ and $\dot{y} \in L^2([0,1]; \mathbb{D}^{1,2})$, we have that $D y_0 \in L^2([0,1] \times \Omega; \mathbb{R}^d)$ and the map $(s,t,\omega) \mapsto D_s \dot{y}_t(\omega)$ is an element of $L^2([0,1]^2 \times \Omega; \mathbb{R}^d)$. By Fubini's theorem, there is a Lebesgue null set $\mathcal{N} \subseteq [0,1]$ such that for each $s \notin \mathcal{N}$, the map $t \mapsto D_s \dot{y}_t$ is an element of $L^2([0,1]; L^2(\Omega))$ and $D_s y_0 \in L^2(\Omega)$. For any $s \in [0,1]$, define

$$y_{s,t} = \begin{cases} D_s y_0 + \int_0^t D_s \dot{y}_r \, \mathrm{d}r & \text{if } s \notin \mathcal{N}, t \in [0,1], \\ 0 & \text{if } s \in \mathcal{N}, t \in [0,1]. \end{cases}$$

Obviously, for each $s \in [0,1]$, the map $[0,1] \ni t \mapsto y_{t,s} \in L^2(\Omega)$ is absolutely continuous. By our assumptions,

$$\int_0^1 \|\dot{y}_r\|_{\mathbb{D}^{1,2}} \, \mathrm{d}r \leqslant \left(\int_0^1 \|\dot{y}_r\|_{\mathbb{D}^{1,2}}^2 \, \mathrm{d}r \right)^{1/2} < \infty,$$

i.e., $\dot{y} \colon [0,1] \to \mathbb{D}^{1,2}$ is Bochner integrable. Since D is a continuous operator from $\mathbb{D}^{1,2}$ to $L^2(\Omega)$, we get

$$D y_t = D y_0 + D \int_0^t \dot{y}_r \, \mathrm{d}r = D y_0 + \int_0^t D \dot{y}_r \, \mathrm{d}r.$$

Combining this with the definition of $y_{s,t}$, we get $y_{\cdot,t} = D_\cdot y_t$ in $L^2(\Omega; H)$ for all $t \in [0,1]$. Moreover, by our assumptions,

$$\begin{aligned} \mathbf{E} \int_0^1 |y_{s,s}|^2 \, \mathrm{d}s &\leqslant \mathbf{E} \int_0^1 |D_s y_0|^2 \, \mathrm{d}s + \mathbf{E} \int_0^1 \left| \int_0^s D_s \dot{y}_r \, \mathrm{d}r \right|^2 \mathrm{d}s \\ &\leqslant \|D y_0\|_2 + \int_0^T \|D \dot{y}_r\|_2^2 \, \mathrm{d}r < \infty, \end{aligned}$$

which means that $y_{s,s}$ is an element of $L^2([0,1] \times \Omega; \mathbb{R}^d)$. By Lemma A.6, we have

$$\begin{aligned} \mathbf{E}^t y_t &= \mathbf{E} y_t + \int_0^t \mathbf{E}^s D_s y_t \cdot \mathrm{d}W_s = \mathbf{E} y_t + \int_0^t \mathbf{E}^s y_{s,t} \cdot \mathrm{d}W_s \\ &= \mathbf{E} y_t + \int_0^t \mathbf{E}^s y_{s,s} \cdot \mathrm{d}W_s + \int_0^t \mathbf{E}^s (y_{s,t} - y_{s,s}) \cdot \mathrm{d}W_s. \end{aligned} \tag{3.3}$$

Note that for any $s \notin \mathcal{N}$, $t \in [0,1]$,

$$y_{s,t} - y_{s,s} = \int_s^t D_s \dot{y}_r \, \mathrm{d}r \, .$$

By the stochastic Fubini theorem,

$$\begin{aligned}\int_0^t \mathbf{E}^s(y_{s,t}-y_{s,s})\cdot \mathrm{d}W_s &= \int_0^t \mathbf{E}^s\Big(\int_s^t D_s\dot{y}_r\,\mathrm{d}r\Big)\cdot \mathrm{d}W_s\\ &= \int_0^t\Big(\int_s^t \mathbf{E}^s D_s\dot{y}_r\,\mathrm{d}r\Big)\cdot \mathrm{d}W_s = \int_0^t \mathrm{d}r\int_0^r \mathbf{E}^s D_s\dot{y}_r\cdot \mathrm{d}W_s\\ &\overset{(A.5)}{=} \int_0^t (\mathbf{E}^r\dot{y}_r-\mathbf{E}\dot{y}_r)\,\mathrm{d}r = \int_0^t \mathbf{E}^r\dot{y}_r\,\mathrm{d}r+\mathbf{E}y_0-\mathbf{E}y_t.\end{aligned}$$

Plugging this into (3.3), we obtain (3.2). ∎

For any $F\in\mathcal{F}$ and $h\in H$, denote

$$\tau_{\varepsilon h}F(\omega):=F\Big(\omega+\varepsilon\int_0^\cdot h_s\,\mathrm{d}s\Big),\quad D_\varepsilon^h F:=\frac{(\tau_{\varepsilon h}F-F)}{\varepsilon}. \tag{3.4}$$

The next lemma is taken from [14], and gives a characterization of the space $\mathbb{D}^{1,p}$ in terms of differentiability properties.

Lemma 3.3. *Let $p\in(1,\infty)$ and $F\in L^p(\Omega)$. The following properties are equivalent.*

(1) $F\in\mathbb{D}^{1,p}$.

(2) *There is $\mathcal{D}F\in L^p(\Omega;H)$ such that for any $h\in H$ and $q\in[1,p)$,*

$$\lim_{\varepsilon\to 0}\mathbf{E}|D_\varepsilon^h F-\langle\mathcal{D}F,h\rangle_H|^q=0.$$

(3) *There is $\mathcal{D}F\in L^p(\Omega;H)$ and some $q\in[1,p)$ such that for any $h\in H$,*

$$\lim_{\varepsilon\to 0}\mathbf{E}|D_\varepsilon^h F-\langle\mathcal{D}F,h\rangle_H|^q=0.$$

Moreover, in that case, $DF=\mathcal{D}F$.

Denote $\Delta_T=\{(s,t):0\leqslant s\leqslant t\leqslant T\}$, $\Delta=\Delta_1$. We need the following.

Assumption 3.4. *For each $(x,t)\in Q$, $a_t(x)$, $b_t(x)$ and $c_t(x)$ are Malliavin differentiable, and each of the random fields $D_s a_t(x)$, $D_s b_t(x)$ and $D_s c_t(x)$ has a continuous version as a map from $\mathbb{R}^n\times\Delta$ to $L^{2p}(\Omega)$ such that*

$$\begin{aligned}\sup_{(s,t)\in\Delta}\big(&\|D_s a_t\|_{C^\alpha(\mathbb{R}^n;L^{2p}(\Omega))}+\|D_s b_t\|_{C^\alpha(\mathbb{R}^n;L^{2p}(\Omega))}\\ &+\|D_s c_t\|_{C^\alpha(\mathbb{R}^n;L^{2p}(\Omega))}\big)\leqslant\Lambda'<\infty.\end{aligned} \tag{$\mathbf{H}_3$}$$

The next theorem is the key to the main purpose of this paper.

Theorem 3.5. *Let $T\in(0,1]$, $q>2p\geqslant 4$ and $C_{x,t}^\beta=C_{x,t}^{\beta,0}(Q_T;L^p(\Omega))$. Under Assumptions 2.2 and 3.4, the following BSPDE:*

$$u_t(x)=\int_t^T(L_s u_s+f_s)(x)\,\mathrm{d}s-\int_t^T v_s(x)\cdot\mathrm{d}W_s, \tag{3.5}$$

has an $\mathcal{F}_t$*-adapted solution* $(u, v) \in C^{2+\alpha}_{x,t} \times C^{2+\alpha}_{x,t}$*, provided that* $f \in C^{\alpha,0}_{x,t}(Q_T; L^q(\Omega))$ *and* $Df \in C^{\alpha,0}_{x,t}(Q_T; L^{2p}(\Omega; H))$*. Moreover, there is a constant* C*, that only depends on* n*,* d*,* p*,* q*,* α*,* Λ *and* Λ'*, such that*

$$\|u\|_{C^{2+\alpha}_{x,t}} + \|v\|_{C^{2+\alpha}_{x,t}} \leqslant C\Big(\|f\|_{C^{\alpha,0}_{x,t}(Q_T;L^q(\Omega))} + \sup_{(s,t)\in\Delta_T} \|D_s f_t\|_{C^\alpha(\mathbb{R}^n;L^{2p}(\Omega;\mathbb{R}^d))}\Big).$$

Proof. We divide the proof into four steps.

Step 1. Let

$$\Lambda_f := \|f\|_{C^{\alpha,0}_{x,t}(Q_T;L^q(\Omega))} + \sup_{(s,t)\in\Delta_T} \|D_s f_t\|_{C^\alpha(\mathbb{R}^n;L^{2p}(\Omega;\mathbb{R}^d))},$$

and let w be the unique solution to equation (2.1) in $C^{2+\alpha,0}_{x,t}(Q_T; L^q(\Omega))$. Below we show that for each (x, t), $w_t(x)$ is Malliavin differentiable, and that Dw satisfies the following $L^p(\Omega; H)$-valued equation:

$$Dw_t = \int_t^T (L_r Dw_r + G_r)\, \mathrm{d}r, \tag{3.6}$$

where $G_r = Df_r + (\partial_{ij} w_r Da^{ij}_r + \partial_i w^i_r Db^i_r + w_r \cdot Dc_r)$. To do this, we consider the following $L^p(\Omega; H)$-valued PDE:

$$\mathcal{D}w_t = \int_t^T L_r(\mathcal{D}w_r)\, \mathrm{d}r + \int_t^T G_r\ \mathrm{d}r = 0. \tag{3.7}$$

By Assumptions 2.2 and 3.4, and Theorem 2.3, we get

$$\|w\|_{C^{2+\alpha,0}_{x,t}(Q_T;L^q(\Omega))} \leqslant C\|f\|_{C^{\alpha,0}_{x,t}(Q_T;L^q(\Omega))} \leqslant C\Lambda_f,$$

and

$$\sum_{i,j} \|Da^{ij}\|_{C^{\alpha,0}_{x,t}(Q_T;L^{2p}(\Omega;H))} + \sum_i \|Db^i\|_{C^{\alpha,0}_{x,t}(Q_T;L^{2p}(\Omega;H))} + \|Dc\|_{C^{\alpha,0}_{x,t}(Q_T;L^{2p}(\Omega;H))} < \infty.$$

Recalling that $q > 2p \geq 4$, Hölder's inequality yields

$$\|G\|_{C^{\alpha,0}_{x,t}(Q_T;L^p(\Omega;H))} \leqslant C\Lambda_f.$$

Due to Theorem 2.3 (with $\mathcal{H} = H$ therein), there exists a unique solution $\mathcal{D}w \in C^{2+\alpha,0}_{x,t}(Q_T; L^p(\Omega; H))$ of (3.7). Thus, for any $h \in H$, $\mathcal{D}^h w_t := \langle \mathcal{D}w_t, h\rangle$ satisfies

$$\mathcal{D}^h w_t - \int_t^T L_r(\mathcal{D}^h w_r)\, \mathrm{d}r = \int_t^T \langle G_r, h\rangle\ \mathrm{d}r, \tag{3.8}$$

$$\|\mathcal{D}^h w\|_{C^{2+\alpha}_{x,t}} + \|\mathcal{D}^h \partial_t w\|_{C^\alpha_{x,t}} \leqslant C|h|_H \Lambda_f. \tag{3.9}$$

Next we show that $D^h_\varepsilon w_t(x)$ (see (3.4) for the definition) converges to $\mathcal{D}^h w_t(x)$ in $L^p(\Omega)$, and as a consequence, we have $\mathcal{D}w_t(x) = Dw_t(x)$. By the definition of $D^h_\varepsilon w$, one sees

$$
\begin{aligned}
&D^h_\varepsilon w_t - \int_t^T \left[\tau_{\varepsilon h} a^{ij}_r\, \partial_{ij} D^h_\varepsilon w_r + \tau_{\varepsilon h} b^i_r\, \partial_i D^h_\varepsilon w_r + \tau_{\varepsilon h} c_r\, D^h_\varepsilon w_r\right] \mathrm{d}r \\
&\qquad = \int_t^T \left[D^h_\varepsilon f_r + D^h_\varepsilon a^{ij}_r\, \partial_{ij} w_r + D^h_\varepsilon b^i_r \partial_i w_r + D^h_\varepsilon\, c_r w_r\right] \mathrm{d}r.
\end{aligned}
\tag{3.10}
$$

Noting that for any $F \in \mathbb{D}^{1,p}$ and $h \in H$,

$$
D^h_\varepsilon F = \frac{(\tau_{\varepsilon h} F - F)}{\varepsilon} = \varepsilon^{-1} \int_0^\varepsilon \tau_{\theta h}\, D^h F\, \mathrm{d}\theta,
\tag{3.11}
$$

we get that for any $q' \in [p, 2p)$,

$$
\begin{aligned}
\mathbf{E}|D^h_\varepsilon f_r(x) - D^h_\varepsilon f_r(y)|^{q'} &= \left\| \varepsilon^{-1} \int_0^\varepsilon \tau_{\theta h}\, [D^h f_r(x) - D^h f_r(y)]\, \mathrm{d}\theta \right\|^{q'}_{L^{q'}(\Omega)} \\
&\leqslant \sup_{0 \leqslant \theta \leqslant \varepsilon} \|\tau_{\theta h}(D^h f_r(x) - D^h f_r(y))\|^{q'}_{L^{q'}(\Omega)}.
\end{aligned}
$$

Due to Girsanov's theorem,

$$
\frac{\mathrm{d}\mathbf{P} \circ \tau^{-1}_{\theta h}}{\mathrm{d}\mathbf{P}} = \mathcal{E}(\theta h) := \exp\left(\theta \int_0^T h_r\, \mathrm{d}W_r - \frac{\theta^2}{2} \int_0^T |h_r|^2\, \mathrm{d}r\right).
$$

Hence,

$$
\begin{aligned}
\mathbf{E}|D^h_\varepsilon f_r(x) - D^h_\varepsilon f_r(y)|^{q'} &\leqslant \sup_{0 \leqslant \theta \leqslant \varepsilon} \mathbf{E}[|D^h f_r(x) - D^h f_r(y)|^{q'} \mathcal{E}(\theta h)] \\
&\leqslant \sup_{0 \leqslant \theta \leqslant \varepsilon} \mathbf{E}[|D^h f_r(x) - D^h f_r(y)|^{2p}]^{\frac{q'}{2p}} \cdot \mathbf{E}[\mathcal{E}^{\frac{q}{2p-q'}}(\theta h)]^{1-\frac{q'}{2p}} \\
&\leqslant C \|Df_r\|^{q'}_{C^\alpha(\mathbb{R}^n; L^{2p}(\Omega; H))} |h|^{q'}_H\, |x-y|^{\alpha q'},
\end{aligned}
$$

where we have used the following fact in the last inequality:

$$
\mathbf{E}\,\mathcal{E}^\kappa(\theta h) = \mathbf{E}\,\mathcal{E}(\kappa\theta h) \exp\left(\frac{\kappa^2 - \kappa}{2} |h|^2_H\right) \leqslant C_\kappa.
$$

Thus,

$$
\sup_{\varepsilon \in (0,1)} \|D^h_\varepsilon f\|_{C^{\alpha,0}_{x,t}(Q_T; L^{q'}(\Omega))} \leqslant C |h|_H \|Df\|_{C^{\alpha,0}_{x,t}(Q_T; L^{2p}(\Omega; H))}.
$$

Similarly, for any $q'' \in (1, 2p)$,

$$
\sup_{\varepsilon \in (0,1)} \left[\|D^h_\varepsilon a\|_{C^{\alpha,0}_{x,t}(Q_T; L^{q''}(\Omega))} + \|D^h_\varepsilon b\|_{C^{\alpha,0}_{x,t}(Q_T; L^{q''}(\Omega))} + \|D^h_\varepsilon\, c_r\|_{C^{\alpha,0}_{x,t}(Q_T; L^{q''}(\Omega))}\right] \leqslant C.
$$

Choosing $q' = p$ and $q'' = \frac{pq}{q-p} \in (p, 2p)$, and noticing that $\|w\|_{C^{2+\alpha,0}_{x,t}(Q_T; L^q(\Omega))} \leqslant C\Lambda_f$, by Hölder's inequality, we get

$$
\sup_{\varepsilon \in (0,1)} \left\| D^h_\varepsilon f + D^h_\varepsilon a^{ij}\, \partial_{ij} w + D^h_\varepsilon b^i \partial_i w + D^h_\varepsilon\, cw \right\|_{C^{\alpha,0}_{x,t}(Q_T; L^p(\Omega))} \leqslant C |h|_H \Lambda_f.
\tag{3.12}
$$

Since $\tau_{\varepsilon h}a, \tau_{\varepsilon h}b, \tau_{\varepsilon h}c$ satisfy (**H**$_1$) and (**H**$_2$), by (3.10), (3.12) and Theorem 2.3, we have

$$\sup_{\varepsilon\in(0,1)} \left(\|D^h_\varepsilon w\|_{C^{2+\alpha}_{x,t}} + \|D^h_\varepsilon \partial_t w\|_{C^\alpha_{x,t}}\right) \leqslant C|h|_H \Lambda_f. \tag{3.13}$$

Let $\delta^h_\varepsilon w := D^h_\varepsilon w - \mathcal{D}^h w$. Next we want to prove that $\delta^h_\varepsilon w_t(x) \to 0$ in $L^p(\Omega)$, for each $(x,t) \in Q_T$. By definition,

$$\begin{aligned} \partial_t \delta^h_\varepsilon w + L_t \delta^h_\varepsilon w &= -(D^h_\varepsilon f - D^h f) \\ &\quad - \left[(D^h_\varepsilon a^{ij} - D^h a^{ij})\partial_{ij} w + (D^h_\varepsilon b^i - D^h b^i)\partial_i w + (D^h_\varepsilon c - D^h c)w\right] \\ &\quad - \varepsilon\left(D^h_\varepsilon a^{ij}\,\partial_{ij} D^h_\varepsilon w + D^h_\varepsilon b^i\,\partial_i D^h_\varepsilon w + D^h_\varepsilon c\, D^h_\varepsilon w\right) =: -\sum_{i=1}^3 F^{\varepsilon,i}, \end{aligned} \tag{3.14}$$

i.e., $\delta^h_\varepsilon w$ is a $L^p(\Omega)$-valued solution to (2.1) with f replaced by $F^\varepsilon_t := \sum_{i=1}^3 F^{\varepsilon,i}_t$. Estimates (3.9) and (3.13) yield

$$\sup_{\varepsilon\in(0,1)} \left(\|\delta^h_\varepsilon w\|_{C^{2+\alpha}_{x,t}} + \|\partial_t \delta^h_\varepsilon w\|_{C^\alpha_{x,t}}\right) \leqslant C|h|_H \Lambda_f. \tag{3.15}$$

By (3.14), for each $R > 0$, we have

$$\begin{aligned} &\partial_t(\delta^h_\varepsilon w\chi_R) + L_t(\delta^h_\varepsilon w\chi_R) + F^\varepsilon\chi_R \\ &\qquad - (2a^{ij}\partial_i\delta^h_\varepsilon w\partial_j\chi_R + \delta^h_\varepsilon w a^{ij}_t\partial_{ij}\chi_R + \delta^h_\varepsilon w b^i_t\partial_i\chi_R) = 0, \end{aligned}$$

where $\chi_R(x) = \chi(x/R)$. Due to our assumptions and (3.15),

$$\left\|(2a^{ij}\partial_i\delta^h_\varepsilon w\partial_j\chi_R + \delta^h_\varepsilon w a^{ij}_t\partial_{ij}\chi_R + \delta^h_\varepsilon w b^i_t\partial_i\chi_R)\right\|_{C^\alpha_{x,t}} \leqslant C|h|_H\Lambda_f/R.$$

So by Theorem 2.3, for any $\alpha' \in (0,\alpha)$,

$$\|\delta^h_\varepsilon w\chi_R\|_{C^{2+\alpha'}_{x,t}} \leqslant C\|F^\varepsilon\chi_R\|_{C^{\alpha'}_{x,t}} + C|h|_H\Lambda_f/R. \tag{3.16}$$

Thanks to Lemma 3.3, for each $(x,t) \in Q_T$, $F^{\varepsilon,1}_t(x) = D^h_\varepsilon f_t(x) - D^h f_t(x) \to 0$ in $L^{2p}(\Omega)$. By (3.11) and the continuity of $Df : Q_T \mapsto L^{2p}(\Omega; H)$, one can verify that the map $Q_T \ni (x,t) \mapsto D^h_\varepsilon f_t(x) \in L^p(\Omega)$ is equivalent continuous. So by the Arzelà–Ascoli theorem, for any sequence $\varepsilon_n \to 0$ $(n\to\infty)$, there exists a subsequence $\varepsilon_{n_k} \to 0$ $(k \to \infty)$ such that for all $R > 0$, $F^{\varepsilon_{n_k},1}\chi_R \to 0$ in $C^{\alpha'}_{x,t}$ with some $\alpha' \in (0,\alpha)$. Similarly, we have $F^{\varepsilon_{n_k},2}\chi_R \to 0$ and $F^{\varepsilon_{n_k},3}\chi_R \to 0$ in $C^{\alpha'}_{x,t}$ as $k \to \infty$. Thus, $\limsup_{\varepsilon\to 0}\|F^\varepsilon\chi_R\|_{C^{\alpha'}_{x,t}} = 0$. So by (3.16), for any $R_0 > 0$,

$$\limsup_{\varepsilon\to 0}\|\delta^h_\varepsilon w\chi_{R_0}\|_{C^{2+\alpha'}_{x,t}} \leqslant \lim_{R\to\infty}\limsup_{\varepsilon\to 0}\|\delta^h_\varepsilon w\chi_R\|_{C^{2+\alpha'}_{x,t}} \leqslant \lim_{R\to\infty} C/R = 0,$$

which implies $D^h_\varepsilon w_t(x) - \mathcal{D}^h w_t(x) \to 0$ in $L^p(\Omega)$. Again by Lemma 3.3, for each $(x,t) \in Q_T$, we have $w_t(x) \in \mathbb{D}^{1,p}$ and $Dw_t(x) = \mathcal{D}w_t(x) \in C^{2+\alpha,0}_{x,t}(Q_T; L^p(\Omega; H))$. Estimate (3.6) follows by the definition of $\mathcal{D}w$.

Step 2. For any $(s,t) \in \Delta_T$, let $w_t^s(x)$ be the solution to the following equation:

$$w_t^s = \int_t^T (L_r w_r^s + g_r^s)\,\mathrm{d}r,$$

where $g_r^s := (D_s a_r^{ij})\partial_{ij} w_r + (D_s b_r^i)\partial_i w_r + (D_s c_r) w_r + D_s f_r$. By Hölder's inequality,

$$\begin{aligned}\|g^s\|_{C^\alpha_{x,t}} &\leqslant \|D_s f\|_{C^{\alpha,0}_{x,t}(Q_T;L^p(\Omega))} + \|w\|_{C^{2+\alpha,0}_{x,t}(Q_T;L^{2p}(\Omega))}\Big(\sum_{ij}\|D_s a^{ij}\|_{C^{\alpha,0}_{x,t}(Q_T;L^{2p}(\Omega))} \\ &\quad + \sum_i \|D_s b^i\|_{C^{\alpha,0}_{x,t}(Q_T;L^{2p}(\Omega))} + \|D_s c\|_{C^{\alpha,0}_{x,t}(Q_T;L^{2p}(\Omega))}\Big) \\ &\leqslant C\|f\|_{C^{\alpha,0}_{x,t}(Q_T;L^q(\Omega))} + C\sup_{(s,t)\in\Delta_T}\|D_s f_t\|_{C^\alpha(\mathbb{R}^n;L^p(\Omega))} \leqslant C\Lambda_f.\end{aligned}$$

Theorem 2.3 yields

$$\sup_{s\in[0,T]}\big(\|\partial_t w^s\|_{C^\alpha_{x,t}} + \|w^s\|_{C^{2+\alpha}_{x,t}}\big) \leqslant C\|g^s\|_{C^\alpha_{x,t}} \leqslant C\Lambda_f. \tag{3.17}$$

Step 3. In this step, we prove that the function $w_t^s(x)$ constructed in Step 2 is a version of $D_s w_t(x)$. Let

$$\mathcal{A}^\alpha = \{w : w \in C^{2+\alpha}_{x,t}, \partial_t w \in C^\alpha_{x,t}\}, \quad \|w\|_{\mathcal{A}^\alpha} := \|w\|_{C^{2+\alpha}_{x,t}} + \|\partial_t w\|_{C^\alpha_{x,t}}.$$

By linearity and Theorem 2.3, the solution map of (2.1),

$$\mathcal{T} : C^\alpha_{x,t} \ni f \mapsto w \in \mathcal{A}^\alpha,$$

is Lipschitz continuous. Since $[0,T] \ni s \mapsto g^s \in C^\alpha_{x,t}$ is measurable, $s \mapsto w^s$ is measurable from $[0,T]$ to $\mathcal{A}^\alpha$. For any $\varphi \in C_c^\infty((0,T);\mathbb{R}^d)$, define

$$w^\varphi = \int_0^T \varphi(s)\cdot w^s\,\mathrm{d}s, \quad g^\varphi = \int_0^T \varphi(s)\cdot g^s\,\mathrm{d}s.$$

Then, one sees that w^φ satisfies

$$w_t^\varphi = \int_t^T (L_r w_r^\varphi + g_r^\varphi)\,\mathrm{d}r.$$

On the other hand, noticing that Dw is the unique solution to (3.6), we have

$$\langle\varphi, Dw_t\rangle_H = \int_t^T (L_r\langle\varphi, Dw_r\rangle_H + \langle\varphi, g_r\rangle_H)\,\mathrm{d}r = \int_t^T (L_r\langle\varphi, Dw_r\rangle_H + g_r^\varphi)\,\mathrm{d}r.$$

So $w^\varphi = \langle\varphi, Dw\rangle$, which implies $s \mapsto w^s$ is a version of Dw.

Step 4. In this step, we prove the $C^{2+\alpha}$ regularity estimate for v. Define $u_t(x) = \mathbf{E}^t w_t(x)$. Theorem 2.3 and Lemma A.4 yield

$$\|u\|_{C^{2+\alpha}_{x,t}} \leqslant \|w\|_{C^{2+\alpha}_{x,t}} \leqslant C\|f\|_{C^\alpha_{x,t}} \leqslant C\Lambda_f.$$

Let $\dot{w}_t(x) := -[L_t w_t(x) + f_t(x)]$. By Step 1, $\dot{w} \in C^{\alpha,0}_{x,t}(Q_T;\mathbb{D}^{1,p})$. Note that

$$w_t(x) = w_0(x) + \int_0^t \dot{w}_s(x)\,\mathrm{d}s.$$

Thanks to Lemma 3.2, for each $(x,t) \in Q_T$,

$$u_t(x) = \mathbf{E}^t w_t(x) = \mathbf{E}w_0(x) + \int_0^t \mathbf{E}^s \dot{w}_s(x)\,\mathrm{d}s + \int_0^t \mathbf{E}^s \mathcal{W}_{s,s}(x)\cdot \mathrm{d}W_s,$$

where $\mathcal{W}_{s,t}(x) = D_s w_0(x) + \int_0^t D_s \dot{w}_r(x)\,\mathrm{d}r$ for all $(x,t) \in Q_T$ and $s \in [0,T]$ a.e. Since

$$\begin{aligned}\mathcal{W}_{s,s}(x) = D_s w_0(x) + \int_0^s D_s\dot{w}_r(x)\,\mathrm{d}r &= \int_s^T D_s[L_r w_r + f_r](x)\,\mathrm{d}r\\ &= \int_s^T [L_r D_s w_r + g^s_r](x)\,\mathrm{d}r = \int_s^T [L_r w^s_r + g^s_r](x) = w^s_s(x), \quad s \in [0,T] \text{ a.e.,}\end{aligned}$$

we get

$$\begin{aligned}u_t(x) &= u_0(x) - \int_0^t \mathbf{E}^s(L_s w_s + f_s)(x)\,\mathrm{d}s + \int_0^t \mathbf{E}^s w^s_s(x)\cdot \mathrm{d}W_s\\ &= u_0(x) - \int_0^t (L_s u_s + f_s)(x)\,\mathrm{d}s + \int_0^t \mathbf{E}^s w^s_s(x)\cdot \mathrm{d}W_s.\end{aligned}$$

Since $u_T(x) = 0$, we have

$$u_0(x) = \int_0^T (L_s u_s + f_s)(x)\,\mathrm{d}s - \int_0^T \mathbf{E}^s w^s_s(x)\cdot \mathrm{d}W_s.$$

Combining the above two equations, we obtain

$$u_t(x) = \int_t^T (L_s u_s + f_s)(x)\,\mathrm{d}s - \int_t^T \mathbf{E}^s w^s_s(x)\,\mathrm{d}W_s.$$

Let $v_s(x) = w^s_s(x)$. Then the above identity implies that $(u_t, v_t) = (\mathbf{E}^t w_t, \mathbf{E}^t w^t_t)$ is a solution to (3.5). Moreover,

$$\|v\|_{C^{2+\alpha}_{x,t}} = \sup_{0\leqslant t\leqslant T} \|\mathbf{E}^t w^t_t\|_{C^{2+\alpha}(\mathbb{R}^n;L^p(\Omega))} \leqslant \sup_{s\in[0,T]} \|w^s\|_{C^{2+\alpha}_{x,t}} \overset{(3.17)}{\leqslant} C\Lambda_f < \infty.$$

So we complete our proof. ∎

Let $\varrho \in C^\infty_c(\mathbb{R}^n)$ be such that $\int \varrho = 1$, and $\varrho_m(x) := m^n \varrho(mx)$. For any function $g: \mathbb{R}^n \to \mathbb{R}^m$, set $g^m := g * \rho_m$.

The following corollary of Theorem 3.5 is standard.

Corollary 3.6 (Stability). *Assume a, b and c satisfy Assumptions 2.2 and 3.4. Let w^m_t (respectively (u^m, v^m)) be the solution to (2.1) (respectively (3.5)) in $C^{2+\alpha}_{x,t}$ (respectively $C^{2+\alpha}_{x,t} \times C^{2+\alpha}_{x,t}$) with a, b, c, f replaced by a^m, b^m, c^m, f^m. Then for any $\beta \in (0,\alpha)$, it holds that*

$$\begin{aligned}\|\partial_t(w - w^m)\|_{C^\beta_{x,t}} + \|w - w^m\|_{C^{2+\beta}_{x,t}} + T^{-1}\|w - w^m\|_{C^0_{x,t}} \to 0 \quad &\text{as } m \to \infty,\\ \|u - u^m\|_{C^{2+\beta}_{x,t}} + \|v - v^m\|_{C^{2+\beta}_{x,t}} \to 0 \quad &\text{as } m \to \infty.\end{aligned}$$

4. SDEs with random singular coefficients

In this section, we give the proof for our main result.

Proof of Theorem 1.2. We first point out that it is enough to prove the well-posedness of (1.1) for $t \in [0, T/2]$, where T is a universal constant depending only on n, α, Λ, p.

(a) *Pathwise uniqueness*. Assume that X_t is a solution to (1.1). We prove the uniqueness by using a Zvonkin type transformation. With a slight abuse of notation, we denote $C^{\beta}_{x,t} = C^{\beta,0}_{x,t}(Q_T; L^p(\Omega; \mathbb{R}^m))$, where m is an integer that can change in different places. Recall that $L_t = a^{ij}_t \partial_{ij} + b^i_t \partial_i$. We consider the following BSPDE:

$$\mathrm{d}u_t + (L_t u_t + b_t)\,\mathrm{d}t = v_t \cdot \mathrm{d}W_t, \quad u_T(x) = 0. \tag{4.1}$$

By our assumptions and Theorem 3.5, (4.1) has an $\mathcal{F}_t$-adapted solution (u_t, v_t) and

$$\|u\|_{C^{2+\alpha}_{x,t}} + \|v\|_{C^{2+\alpha}_{x,t}} < \infty. \tag{4.2}$$

Since $u_t = \mathbf{E}^t w_t$, w_t solves

$$\partial_t w + L_t w + b = 0, \quad w_T(x) = 0$$

and

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\Big(\sup_{t\in[0,T]}\|b_t(\cdot,\omega)\|_{C^\alpha} + \sup_{(s,t)\in\Delta_T}\|D_s b_t(\cdot,\omega)\|_{C^\alpha}\Big) < \infty.$$

By Remark 2.5, we have

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\sup_{t\in[0,T]}\big(\|w_t(\cdot,\omega)\|_{C^{2+\alpha}} + T^{-1}\|w_t(\cdot,\omega)\|_{C^\alpha}\big) \leqslant C \operatorname*{ess\,sup}_{\omega\in\Omega}\sup_{t\in[0,T]}\|b_t(\cdot,\omega)\|_{C^\alpha}.$$

Interpolation inequality and the above estimate yield

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\sup_{t\in[0,T]}\|u_t(\cdot,\omega)\|_{C^1} \leqslant \operatorname*{ess\,sup}_{\omega\in\Omega}\sup_{t\in[0,T]}\|w_t(\cdot,\omega)\|_{C^1} \leqslant C_T,$$

where $C_T \to 0$ as $T \to 0$. Below we fix $T = T(n, \alpha, \Lambda, p) > 0$ so that

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\sup_{t\in[0,T]}\|u_t(\cdot,\omega)\|_{C^1} \leqslant \frac{1}{2}.$$

Let $\phi_t(x) = x + u_t(x)$. Then

$$\frac{1}{2} \leqslant \operatorname*{ess\,sup}_{\omega\in\Omega}\sup_{0\leqslant t\leqslant T}\|\nabla\phi_t(x,\omega)\|_{L^\infty} \leqslant \frac{3}{2}. \tag{4.3}$$

So, for almost all $\omega \in \Omega$, $\phi_t(\cdot, \omega)$ is a stochastic $C^{2+\alpha}$-differential homeomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$. By the definition of ϕ,

$$\mathrm{d}\phi_t(x) = -(L_t u_t(x) + b_t(x))\,\mathrm{d}t + v_t(x)\cdot\mathrm{d}W_t = \mathrm{d}u_t(x) = \mathrm{d}g_t(x) + \mathrm{d}m_t(x),$$

where

$$g_t(x) = -\int_0^t (L_s u_s(x) + b_s(x))\,\mathrm{d}s, \quad m_t(x) := \int_0^t v_s(x)\,\mathrm{d}W_s. \tag{4.4}$$

We want to show that ϕ, g, u, v and X are regular enough to apply the Itô–Wenzell formula (see Lemma A.7). Since $\|v\|_{C^{2+\alpha}_{x,t}} < \infty$, we have

$$\sup_{t\in[0,T];x\neq y} \frac{\mathbf{E}|\nabla^2 v_t(x) - \nabla^2 v_t(y)|^p}{|x-y|^{\alpha p}} < \infty.$$

Note that $p > n/\alpha$, so for any $\beta \in (n/p, \alpha)$ and $N > 0$, by Garsia–Rademich–Rumsey's inequality,

$$\begin{aligned}
&\sup_{t\in[0,T]} \mathbf{E}\Big(\sup_{x,y\in B_N} \frac{|\nabla v_t^2(x) - \nabla^2 v_t(y)|}{|x-y|^{\beta-n/p}}\Big)^p \\
&\quad \leqslant C_N \sup_{t\in[0,T]} \mathbf{E}\Big(\int_{B_N}\int_{B_N} \frac{|\nabla^2 v_t(x) - \nabla^2 v_t(y)|^p}{|x-y|^{d+\beta p}}\,\mathrm{d}x\,\mathrm{d}y\Big) \\
&\quad \leqslant C_N \int_{B_N}\int_{B_N} |x-y|^{-d+(\alpha-\beta)p} \leqslant C_N.
\end{aligned}$$

Combining this and the fact that $\sup_{t\in[0,T]} \mathbf{E}|\nabla^2 v_t(0)|^p < \infty$, we get

$$\sup_{t\in[0,T]} \mathbf{E}\Big(\sup_{x\in B_N} |\nabla^2 v_t(x)|^p\Big) < \infty, \quad \forall N > 0.$$

Moreover, one can also prove

$$\sup_{t\in[0,T]} \mathbf{E}\|v_t\|^p_{C^2(B_N)} < \infty, \quad \forall N > 0. \tag{4.5}$$

Recall that $g_t(x)$ and $m_t(x)$ are defined in (4.4). Let

$$\eta_t(x) := \int_0^t g_s(x)\,\mathrm{d}s \overset{(4.1)}{=} u_t(x) - u_0(x) - m_t(x).$$

By Burkholder–Davis–Gundy's inequality, for each $k = 0, 1, 2$,

$$\begin{aligned}
\mathbf{E}\big|\nabla^k m_t(x) - \nabla^k m_t(y)\big|^p &\leqslant C\mathbf{E}\Big[\int_0^t |\nabla^k v_s(x) - \nabla^k v_s(y)|^k\,\mathrm{d}s\Big]^{p/2} \\
&\leqslant C\mathbf{E}\int_0^t |\nabla^k v_s(x) - \nabla^k v_s(y)|^p\,\mathrm{d}s \leqslant C|x-y|^{\alpha p}\|\nabla^k v\|^p_{C^\alpha_{x,t}},
\end{aligned}$$

which together with (4.2) implies

$$\|\eta\|_{C^{2+\alpha}_{x,t}} \leqslant C\big(\|u\|_{C^{2+\alpha}_{x,t}} + \|v\|_{C^{2+\alpha}_{x,t}}\big).$$

By the definition of η,

$$\|\partial_t\eta\|_{C^\alpha_{x,t}} = \|g\|_{C^\alpha_{x,t}} \leqslant \|L_t u + b\|_{C^\alpha_{x,t}} \leqslant C\big(\|u\|_{C^{2+\alpha}_{x,t}} + \|b\|_{C^\alpha_{x,t}}\big).$$

Thanks to Lemma A.3, for any $\beta \in (n/p, \alpha)$ and $\theta = 1/2 + (\alpha - \beta)/2 \in (1/2, 1)$, we have

$$\|\eta\|_{C^\theta_t C^{1+\beta}_x} \leqslant C\|\partial_t\eta\|^\theta_{C^\alpha_{x,t}}\|\eta\|^{1-\theta}_{C^{2+\alpha}_{x,t}}.$$

By the same procedure used for proving (4.5), we have

$$\Big[\mathbf{E}\Big\|\int_{t_1}^{t_2} g_s\,\mathrm{d}s\Big\|^p_{C^1(B_N)}\Big]^{1/p} = \big[\mathbf{E}\|\eta_{t_1} - \eta_{t_2}\|^p_{C^1(B_N)}\big]^{1/p} \leqslant C_N|t_1 - t_2|^\theta, \quad \theta \in (1/2, 1).$$

On the other hand, $\mathbf{E}|X_{t_1} - X_{t_2}|^{p'} \leqslant C|t_1 - t_2|^{p'/2}$, where $p' = p/(p-1)$. So ϕ, g, v, X satisfy all the conditions in Lemma A.7. Using (A.7), we get

$$\begin{aligned}
\mathrm{d}\phi_t(X_t) &= -L_t u_t(X_t) - b_t(X_t)\,\mathrm{d}t + v^k_t(X_t)\,\mathrm{d}W^k_t \\
&\qquad + [b^i_t(X_t)\partial_i\phi_t(X_t) + a^{ij}_t(X_t)\partial_{ij}\phi_t(X_t) + \partial_i v^k_t(X_t)\sigma^{ik}_t(X_t)]\,\mathrm{d}t \\
&\qquad + \partial_i\phi_t(X_t)\sigma^{ik}_t(X_t)\,\mathrm{d}W^k_t \\
&= \partial_i v^k_t(X_t)\sigma^{ik}_t(X_t)\,\mathrm{d}t + \partial_i\phi_t(X_t)\sigma^{ik}_t(X_t)\,\mathrm{d}W^k_t + v^k_t(X_t)\,\mathrm{d}W^k_t.
\end{aligned}$$

Set

$$Y_t = \phi_t(X_t), \quad \widetilde{b}_t(y) = \partial_i v^k_t\sigma^{ik}_t \circ \phi^{-1}_t(y) \quad \text{and} \quad \widetilde{\sigma}_t(y) = [\nabla\phi_t\sigma_t + v_t] \circ \phi^{-1}_t(y).$$

By the above calculations, one sees that

$$Y_t = Y_0 + \int_0^t \widetilde{b}_s(Y_s)\,\mathrm{d}s + \int_0^t \widetilde{\sigma}_s(Y_s)\,\mathrm{d}W_s. \tag{4.6}$$

Thanks to Lemma A.2, $\widetilde{b}$ and $\widetilde{\sigma}$ are $\mathcal{B} \times \mathcal{P}$-measurable. For any $x, y \in B_N$ and $t \in [0, T]$, by the definitions of $\widetilde{b}$ and $\widetilde{\sigma}$, we have

$$\begin{aligned}
|\widetilde{b}_t(0)| + |\widetilde{\sigma}_t(0)| &\leqslant CK^N_t, \\
|\widetilde{b}_t(x) - \widetilde{b}_t(y)| + |\widetilde{\sigma}_t(x) - \widetilde{\sigma}_t(y)| &\leqslant CK^N_t|x - y|,
\end{aligned}$$

where $K^N_t := \|u_t\|_{C^2(B_N)} + \|v_t\|_{C^2(B_N)}$. It is not hard to see that K^N_t is progressive measurable and satisfies

$$\mathbf{E}\int_0^T K^N_t\,\mathrm{d}t \leqslant T\sup_{t\in[0,T]}\mathbf{E}K^N_t \overset{(4.5)}{<} \infty.$$

Thanks to Theorem 1.2 of [9], equation (4.6) admits a unique solution, which implies X_t is unique up to indistinguishability.

(b) *Existence*. Let $b^m_t = b_t * \varrho_m$ and let X^m be the solution to

$$X^m_t = X_0 + \int_0^t b^m_s(X^m_s)\,\mathrm{d}s + \int_0^t \sigma_s(X^m_s)\,\mathrm{d}W_s, \quad t \in [0, T]. \tag{4.7}$$

We claim that X_t^m converges to a process X_t uniformly on compacts in probability (ucp convergence, in short). Let (u^m, v^m) be the pair of functions constructed in Theorem 3.5 satisfying

$$\mathrm{d}u_t^m + \big[a_t^{ij}\partial_{ij}u_t^m + (b_t^m)^i\partial_i u_t^m + b_t^m\big]\,\mathrm{d}t = v_t^m \cdot \mathrm{d}W_t.$$

Like before, we can find a uniform constant $T = T(n, \alpha, \Lambda, p) > 0$ such that $\|\nabla u_t^m\|_{L^\infty} \leqslant 1/2$. Define $\phi_t^m(x) := x + u_t^m(x),\ Y_t^m := \phi_t^m(X_t^m)$ and $Z_t^{m,m'} := Y_t^m - Y_t^{m'}$. Again by Itô–Wentzell's formula, we have

$$Z_t^{m,m'} = Y_t^m - Y_t^{m'} = u_0^m(X_0) - u_0^{m'}(X_0) + \int_0^t [\widetilde{b}_s^m(X_s^m) - \widetilde{b}_s^{m'}(X_s^{m'})]\,\mathrm{d}s$$
$$+ \int_0^t [\widetilde{\sigma}_s^m(X_s^m) - \widetilde{\sigma}_s^{m'}(X_s^{m'})]\,\mathrm{d}W_s,$$

where

$$\widetilde{b}_t^m := [\partial_i v_t^{m,k}\sigma_t^{ik}] \circ (\phi_t^m)^{-1}, \quad \widetilde{\sigma}_t^m := [(\nabla\phi_t^m)\sigma_t + v_t^m] \circ (\phi_t^m)^{-1}.$$

By Itô's formula, for any stopping time $\tau \leqslant T$,

$$\big|Z_{t\wedge\tau}^{m,m'}\big|^2 = |u_0^m(X_0) - u_0^{m'}(X_0)|^2 + 2\int_0^{t\wedge\tau} Z_s^{m,m'} \cdot \big[\widetilde{b}_s^m(Y_s^m) - \widetilde{b}_s^{m'}(Y_s^{m'})\big]\,\mathrm{d}s$$
$$+ \int_0^{t\wedge\tau} \mathrm{tr}\big[\widetilde{\sigma}_s^m(Y_s^m) - \widetilde{\sigma}_s^{m'}(Y_s^{m'})\big]\big[\widetilde{\sigma}_s^m(Y_s^m) - \widetilde{\sigma}_s^{m'}(Y_s^{m'})\big]^*\,\mathrm{d}s + m_{t\wedge\tau}, \tag{4.8}$$

where

$$m_t = 2\int_0^t Z_s^{m,m'} \cdot \big[\widetilde{\sigma}_s^m(Y_s^m) - \widetilde{\sigma}_s^{m'}(Y_s^{m'})\big]\,\mathrm{d}W_s.$$

For any $N, k \in \mathbb{N}$, let $K_t^{m,N} := \|u_t^m\|_{C^2(B_N)} + \|v_t^m\|_{C^2(B_N)}$,

$$\tau^{N,k} = \inf_m \inf\Big\{t \geqslant 0 : \int_0^t (K_s^{m,N})^2\,\mathrm{d}s \geqslant k\Big\} \wedge T,$$

and

$$\sigma^N = \inf_m \inf\big\{t \geqslant 0 : |Y_t^m| > N/2\big\} \wedge T, \quad \sigma^{N,k} := \sigma^N \wedge \tau^{N,k}.$$

For all $x, y \in B_{N/2}$ and $t \in [0, \sigma^{N,k}]$, we have

$$\sup_{m\in\mathbb{N}} \big(|\widetilde{b}_t^m(x) - \widetilde{b}_t^m(y)| + |\widetilde{\sigma}_t^m(x) - \widetilde{\sigma}_t^m(y)|\big) \leqslant C_k|x - y|. \tag{4.9}$$

Since for each $(x, t) \in B_{N/2} \times [0, T]$, $(\phi_t^m)^{-1}(x) \in B_N$, we obtain that for any $x \in B_{N/2}$ and $t \in [0, \tau^{N,k}]$,

$$\begin{aligned}
|\widetilde{b}_t^m(x) - \widetilde{b}_t^{m'}(x)| &\leqslant \big|[\partial_i v_t^{m,k}\sigma_t^{ik}] \circ (\phi_t^m)^{-1}(x) - [\partial_i v_t^{m',k}\sigma_t^{ik}] \circ (\phi_t^m)^{-1}(x)\big| \\
&\quad + \big|[\partial_i v_t^{m',k}\sigma_t^{ik}] \circ (\phi_t^m)^{-1}(x) - [\partial_i v_t^{m',k}\sigma_t^{ik}] \circ (\phi_t^{m'})^{-1}(x)\big| \\
&\leqslant C\|\nabla v_t^m - \nabla v_t^{m'}\|_{L^\infty(B_N)} + C\|v_t^{m'}\|_{C^2(B_N)}|(\phi_t^m)^{-1}(x) - (\phi_t^{m'})^{-1}(x)| \\
&\leqslant C\|v_t^m - v_t^{m'}\|_{C^2(B_N)} + C\|v_t^{m'}\|_{C^2(B_N)} \sup_{y\in B_N} |\phi_t^{m'}(y) - \phi_t^m(y)| \\
&\leqslant C_k\big(\|u_t^m - u_t^{m'}\|_{C^2(B_N)} + \|v_t^m - v_t^{m'}\|_{C^2(B_N)}\big).
\end{aligned}$$

Similarly, for each $x \in B_{N/2}$ and $t \in [0, \tau^{N,k}]$,

$$|\widetilde{\sigma}_t^m(x) - \widetilde{\sigma}_t^{m'}(x)| \leqslant C_k\big(\|u_t^m - u_t^{m'}\|_{C^2(B_N)} + \|v_t^m - v_t^{m'}\|_{C^2(B_N)}\big).$$

By Theorem 3.5, Corollary 3.6 and the same procedure used for proving (4.5), we have

$$\sup_{t\in[0,T],m\in\mathbb{N}} \mathbf{E}|K_t^{m,N}|^p = \sup_{t\in[0,T],m\in\mathbb{N}} \mathbf{E}\big(\|u_t^m\|_{C^2(B_N)} + \|v_t^m\|_{C^2(B_N)}\big)^p < \infty$$

and

$$\lim_{m\to\infty} \sup_{t\in[0,T]} \mathbf{E}\big(\|u_t - u_t^m\|_{C^2(B_N)} + \|v_t - v_t^m\|_{C^2(B_N)}\big)^p = 0.$$

Thus,

$$\lim_{k\to\infty} \tau^{N,k} = T, \qquad \lim_{N\to\infty} \sigma^N = T \tag{4.10}$$

and

$$\lim_{m,m'\to\infty} \mathbf{E}\big(\|\widetilde{\sigma}_t^m - \widetilde{\sigma}_t^{m'}\|_{L^\infty(B_{N/2})}^p + \|\widetilde{b}_t^m - \widetilde{b}_t^{m'}\|_{L^\infty(B_{N/2})}^p \mathbf{1}_{[0,\sigma^{N,k}]}(t)\big) = 0. \tag{4.11}$$

Let $\tau = \sigma^{N,k}$ in (4.8). Using (4.9), we have

$$\begin{aligned}
\big|Z_{t\wedge\sigma^{N,k}}^{m,m'}\big|^2 &\overset{(4.9)}{\leqslant} \big|u_0^m(X_0) - u_0^{m'}(X_0)\big|^2 \\
&\quad + C_k \int_0^{t\wedge\sigma^{N,k}} |Z_s^{m,m'}| \big(|Z_s^{m,m'}| + \|\widetilde{b}_s^m - \widetilde{b}_s^{m'}\|_{L^\infty(B_{N/2})}\big)\, \mathrm{d}s \\
&\quad + C_k \int_0^{t\wedge\sigma^{N,k}} \big(|Z_s^{m,m'}| + \|\widetilde{\sigma}_s^m - \widetilde{\sigma}_s^{m'}\|_{L^\infty(B_{N/2})}\big)^2\, \mathrm{d}s + m_{t\wedge\sigma^{N,k}} \\
&\leqslant \|u_0^m - u_0^{m'}\|_{L^\infty}^2 + C_k \int_0^t \big|Z_{s\wedge\sigma^{N,k}}^{m,m'}\big|^2\, \mathrm{d}s + m_{t\wedge\sigma^{N,k}} \\
&\quad + C_k \int_0^{t\wedge\sigma^{N,k}} \big(\|\widetilde{b}_s^m - \widetilde{b}_s^{m'}\|_{L^\infty(B_{N/2})}^2 + \|\widetilde{\sigma}_s^m - \widetilde{\sigma}_s^{m'}\|_{L^\infty(B_{N/2})}^2\big)\, \mathrm{d}s.
\end{aligned}$$

By Gronwall's inequality and (4.11), we get

$$\begin{aligned}
\mathbf{E}\big|(Z^{m,m'})_{T\wedge\sigma^{N,k}}^*\big|^2 &\leqslant C_k\|u_0^m - u_0^{m'}\|_{L^\infty}^2 \\
&\quad + C_k\mathbf{E}\int_0^T \big(\|\widetilde{\sigma}_s^m - \widetilde{\sigma}_s^{m'}\|_{L^\infty(B_{N/2})}^2 + \|\widetilde{b}_s^m - \widetilde{b}_s^{m'}\|_{L^\infty(B_{N/2})}^2\big)\mathbf{1}_{[0,\sigma^{N,k}]}(s)\, \mathrm{d}s \\
&\overset{(4.11)}{\longrightarrow} 0 \quad \text{as } m, m' \to \infty.
\end{aligned} \tag{4.12}$$

On the other hand,

$$\begin{aligned}
|X_t^m - X_t^{m'}| &= |(\phi_t^m)^{-1}(\phi_t^m(X_t^m)) - (\phi_t^m)^{-1}(\phi_t^m(X_t^{m'}))| \leqslant 2|\phi_t^m(X_t^m) - \phi_t^m(X_t^{m'})| \\
&\leqslant 2|\phi_t^m(X_t^m) - \phi_t^{m'}(X_t^{m'})| + 2|\phi_t^{m'}(X_t^{m'}) - \phi_t^m(X_t^{m'})| \\
&\leqslant 2\|u_t^m - u_t^{m'}\|_{L^\infty} + 2|Y_t^m - Y_t^{m'}|.
\end{aligned} \tag{4.13}$$

Combining (4.12) and (4.13), we get

$$\lim_{m,m'\to\infty} \mathbf{E} \sup_{t\in[0,T]} |X^m_{t\wedge\sigma^{N,k}} - X^{m'}_{t\wedge\sigma^{N,k}}| \leqslant 2 \lim_{m,m'\to\infty} \mathbf{E} \sup_{t\in[0,T]} \|u^m_t - u^{m'}_t\|_{L^\infty} + 2 \lim_{m,m'\to\infty} \mathbf{E}(Z^{m,m'})^*_{T\wedge\sigma^{N,k}} = 0.$$

Noting that

$$\lim_{N\to\infty}\lim_{k\to\infty} \sigma^{N,k} = \lim_{N\to\infty} \sigma^N \overset{(4.10)}{=} T,$$

we obtain

$$\lim_{m,m'\to\infty} \mathbf{P}\Big(\sup_{t\in[0,T/2]} |X^m_t - X^{m'}_t| > \varepsilon)\Big) = 0, \quad \forall \varepsilon > 0.$$

This implies that there is a continuous process $\{X_t\}_{t\in[0,T/2]}$ such that $X^m \to X$ in the sense of ucp. Hence,

$$\int_0^t \sigma_s(X^m_s)\,\mathrm{d}W_s \xrightarrow{\mathbf{P}} \int_0^t \sigma_s(X_s)\,\mathrm{d}W_s, \quad \forall t \in [0, T/2],$$

and for each $t \in [0, T/2]$ and $\varepsilon > 0$,

$$\begin{aligned}
&\mathbf{P}\Big(\Big|\int_0^t b^m_s(X^m_s)\,\mathrm{d}s - \int_0^t b_s(X_s)\,\mathrm{d}s\Big| > \varepsilon\Big)\\
&\leqslant \mathbf{P}\Big(\sup_{t\in[0,T/2]} |b^m(X^m_t) - b(X^m_t)| > \frac{\varepsilon}{2}\Big) + \mathbf{P}\Big(\sup_{t\in[0,T/2]} |b_t(X^m_t) - b_t(X_t)|\,\mathrm{d}s > \frac{\varepsilon}{2}\Big)\\
&\leqslant \mathbf{P}\Big(\|b^m - b\|_{L^\infty(Q_T)} > \frac{\varepsilon}{2}\Big) + \mathbf{P}\Big(\sup_{t\in[0,T/2]} |X^m_t - X_t|^\alpha > \frac{\varepsilon}{2\Lambda}\Big) \to 0, \quad \text{as } m \to \infty.
\end{aligned}$$

Taking limits on both sides of (4.7), one sees that X is a solution to (1.1). ■

A. Appendix

In this section, we give some lemmas used in the previous sections. The following basic result is useful.

Lemma A.1. *Let $f \in L^1(\mathbb{R}^n; \mathcal{B}) + L^\infty(\mathbb{R}^n; \mathcal{B})$.*

(1) (*Bernstein's inequality*) *For any $k = 0, 1, 2, \ldots$, there is a constant $C = C(n, k) > 0$ such that for all $j = -1, 0, 1, \ldots$,*

$$\|\nabla^k \Delta_j f\|_0 \leqslant C\, 2^{kj} \|\Delta_j f\|_0;$$

(2) *For any $\alpha \in (0, 1)$, there is a constant $C = C(\alpha, n) > 1$ such that*

$$C^{-1} \sup_{j\geqslant -1} 2^{j\alpha} \|\Delta_j f\|_0 \leqslant \|f\|_{C^\alpha} \leqslant C \sup_{j\geqslant -1} 2^{j\alpha}\|\Delta_j f\|_0.$$

One can find the proof of above lemma in [1] for $\mathcal{B} = \mathbb{R}$. We present its Banach-valued version below for the reader's convenience.

Proof. For any $j = 0, 1, 2, \ldots$, we have $\int_{\mathbb{R}^n} h_j(z)\,\mathrm{d}z = \varphi_j(0) = 0$, so

$$\begin{aligned}
\|\Delta_j f(x)\|_{\mathcal{B}} &= \Big\| \int_{\mathbb{R}^n} h_j(x-y)[f(y) - f(x)]\,\mathrm{d}y \Big\|_{\mathcal{B}} \\
&= \Big\| \int_{\mathbb{R}^n} 2^{jn} h(2^j(x-y))[f(y) - f(x)]\,\mathrm{d}y \Big\|_{\mathcal{B}} \\
&\leqslant C \|f\|_{C^\alpha} \int_{\mathbb{R}^n} 2^{jn} h(2^j z)|z|^\alpha\,\mathrm{d}z = C\,2^{-j\alpha}\|f\|_{C^\alpha},
\end{aligned}$$

which implies

$$\sup_{j \geqslant -1} 2^{j\alpha}\|\Delta_j f\|_0 \leqslant C_\alpha \|f\|_{C^\alpha}.$$

On the other hand,

$$\begin{aligned}
\Big\| f(x) - \sum_{j=-1}^{k} \Delta_j f(x) \Big\|_{\mathcal{B}} &= \Big\| \int_{\mathbb{R}^n} \mathcal{F}^{-1}(\chi(2^{-k}\cdot))(y)\,[f(x) - f(x-y)]\,\mathrm{d}y \Big\|_{\mathcal{B}} \\
&= \Big\| \int_{B_{2^k\varepsilon}} \mathcal{F}^{-1}(\chi)(z)\,[f(x) - f(x - 2^{-k}z)]\,\mathrm{d}z \Big\|_{\mathcal{B}} \\
&\quad + \Big\| \int_{B^c_{2^k\varepsilon}} \mathcal{F}^{-1}(\chi)(z)\,[f(x) - f(x - 2^{-k}z)]\,\mathrm{d}z \Big\|_{\mathcal{B}} \\
&\leqslant \mathrm{osc}_{B_\varepsilon(x)} f \cdot \int_{\mathbb{R}^n} |\mathcal{F}^{-1}(\chi)(y)|\,\mathrm{d}y + 2\|f\|_0 \int_{B^c_{2^k\varepsilon}} |\mathcal{F}^{-1}(\chi)(y)|\,\mathrm{d}y.
\end{aligned}$$

Letting $k \to \infty$ and then $\varepsilon \to 0$, we obtain that for each $f \in C_b(\mathbb{R}^n; \mathcal{B})$ and $x \in \mathbb{R}^n$, $f(x) = \sum_{j \geqslant -1} \Delta_j f(x)$. Thus, for any $K > 0$,

$$\begin{aligned}
|f(x) - f(y)|_{\mathcal{B}} &\leqslant \sum_{j \geqslant -1} |\Delta_j f(x) - \Delta_j f(y)|_{\mathcal{B}} \leqslant |x-y| \sum_{-1 \leqslant j \leqslant K} \|\nabla \Delta_j f\|_0 + 2 \sum_{j > K} \|\Delta_j f\|_0 \\
&\leqslant C_\alpha \big(|x-y|\,2^{(1-\alpha)K} + C\,2^{-\alpha K}\big) \sup_{j \geqslant -1} 2^{\alpha j}\|\Delta_j f\|_0.
\end{aligned}$$

For any $|x-y| < 1$, by choosing $K = -\log_2(|x-y|)$, we obtain

$$|f(x) - f(y)|_{\mathcal{B}} \leqslant C_\alpha |x-y|^\alpha \sup_{j \geqslant -1} 2^{\alpha j}\|\Delta_j f\|_0,$$

completing the proof. ∎

Suppose $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a continuous homeomorphism on $\mathbb{R}^n$; its inverse map is denoted by f^{-1}. Our next auxiliary lemma is used in the proof of Theorem 1.2.

Lemma A.2. *Suppose $(S; \mathcal{S})$ is a measurable space. Let $F\colon (S \times \mathbb{R}^n; \mathcal{S} \times \mathcal{B}) \to (\mathbb{R}^n; \mathcal{B})$.*

(1) *Assume X is another measurable map from $(S; \mathcal{S})$ to $(\mathbb{R}^n; \mathcal{B})$. Then the map $a \mapsto F(a, X(a))$ is measurable from $(S; \mathcal{S})$ to $(\mathbb{R}^n; \mathcal{B})$.*

(2) *For any $L > 0$, define*

$$H_L := \{f : \mathbb{R}^n \to \mathbb{R}^n | f \text{ is a continuous homeomorphism and } L^{-1}|x-y| \leqslant |f(x) - f(y)| \leqslant L|x-y|\}.$$

If $F : (S \times \mathbb{R}^n; \mathcal{S} \times \mathcal{B}) \to (\mathbb{R}^n; \mathcal{B})$ and for each $a \in S$, $F(a, \cdot) \in H_L$, then the map

$$F^{-1} : S \times \mathbb{R}^n \ni (a, x) \mapsto [F^{-1}(a, \cdot)](x) \in \mathbb{R}^n$$

is $\mathcal{S} \times \mathcal{B}/\mathcal{B}$ measurable.

Proof. (1) This conclusion is trivial since the map $a \mapsto (a, X(a))$ is $\mathcal{S}/\mathcal{S} \times \mathcal{B}$ measurable.

(2) Define

$$d(f, g) := \sup_{x \in \mathbb{R}^n} \frac{|f(x) - g(x)|}{1 + |x|}, \quad \forall f, g \in H_L.$$

It is easy to verify that H_L is a metric space equipped with the metric d. For any $f \in H_L$ and $\varepsilon > 0$, by the continuity of $x \mapsto F(a, x)$, we get

$$\{a : d(F(a, \cdot), f) < \varepsilon\} = \bigcap_{\substack{q \in \mathbb{Q}^n; \\ r \in \mathbb{Q} \cap [0,1)}} \left\{a : \frac{|F(a, q) - f(q)|}{1 + |q|} < r\varepsilon\right\} \in \mathcal{S}.$$

So the map $\overline{F} : (S, \mathcal{S}) \to (H_L, \mathcal{B}(H_L; d))$ is measurable. Obviously, the map

$$\text{Inv} : H_L \ni f \mapsto f^{-1} \in H_L,$$

is well-defined. Now assume $d(f_n, f) \to 0$. Given $x \in \mathbb{R}^n$, call $y = f^{-1}(x)$. Then

$$\begin{aligned} |f_n^{-1}(x) - f^{-1}(x)| &= |f_n^{-1} \circ f(y) - f_n^{-1} \circ f_n(y)| \\ &\leqslant L|f(y) - f_n(y)| \leqslant L(1 + |y|)d(f_n, f). \end{aligned}$$

By definition of H_L,

$$|f(y) - f(0)| \geqslant L^{-1}|y|,$$

which implies

$$|x| = |f(y)| \geqslant L^{-1}|y| - |f(0)|.$$

So

$$|f_n^{-1}(x) - f^{-1}(x)| \leqslant L(1 + Lf(0) + L|x|)d(f_n, f) \leqslant C_{f,L}(1 + |x|)d(f_n, f),$$

which implies $d(f_n^{-1}, f^{-1}) \leqslant C_{f,L} d(f_n, f) \to 0$. Thus, the map Inv: $H_L \to H_L$ is continuous. Hence, the map $\overline{F}^{-1} := \text{Inv} \circ \overline{F}$ from $(S, \mathcal{S})$ to $(H_L, \mathcal{B}(H_L))$ is also measurable. As a consequence, the map

$$F^{-1} : S \times \mathbb{R}^n \ni (a, x) \mapsto [\text{Inv} \circ F(a, \cdot)](x) \in \mathbb{R}^n$$

is $\mathcal{S} \times \mathcal{B}/\mathcal{B}$ measurable. ∎

Roughly speaking, the above lemma shows that if $(a, x) \mapsto F(a, x)$ is measurable, then $(a, x) \mapsto F^{-1}(a, \cdot)(x)$ is also measurable.

The following interpolation lemma is used several times in our paper.

Lemma A.3. *Let $0 \leqslant \gamma_0 < \gamma_1 < \gamma_2$ with $\gamma_1 \notin \mathbb{N}$ and let $\theta := (\gamma_2 - \gamma_1)/(\gamma_2 - \gamma_0) \in (0, 1)$. Write $Q_T = \mathbb{R}^n \times [0, T]$ and let $\mathcal{B}$ be a Banach space. Then there is a constant $C > 0$ such that, for all $f \in C^{\gamma_2}_{x,t}$ with $\partial_t f \in C^{\gamma_0}_{x,t}$,*

$$\|f_{t_1} - f_{t_2}\|_{C^{\gamma_1}} \leqslant C\,|t_1 - t_2|^{\theta}\,\|\partial_t f\|^{\theta}_{C^{\gamma_0}_{x,t}}\,\|f\|^{1-\theta}_{C^{\gamma_2}_{x,t}}. \tag{A.1}$$

Proof. First of all, for any $t \in [0, 1]$, we have

$$\|f_t\|_{C^{\gamma_1}} \leqslant C\,\|f_t\|^{\theta}_{C^{\gamma_0}}\,\|f_t\|^{1-\theta}_{C^{\gamma_2}}.$$

For any $0 \leqslant t_0 < t_1 \leqslant T$, $\beta \in (0, \theta)$ and $q > 1/\theta$, by the Garsia–Rademich–Rumsey inequality, we have

$$\begin{aligned}
\frac{\|f_{t_1} - f_{t_0}\|^q_{C^{\gamma_1}}}{|t_1 - t_0|^{\beta q - 1}} &\leqslant C \int_{t_0}^{t_1}\int_{t_0}^{t_1} \frac{\|f_t - f_s\|^q_{C^{\gamma_1}}}{|t - s|^{1+\beta q}}\,\mathrm{d}s\,\mathrm{d}t \\
&\leqslant C \int_{t_0}^{t_1}\int_{t_0}^{t_1} \|f_t - f_s\|^{\theta q}_{C^{\gamma_0}}\,\|f_t - f_s\|^{(1-\theta)q}_{C^{\gamma_2}}\,|t - s|^{-1-\beta q}\,\mathrm{d}s\,\mathrm{d}t \\
&\leqslant C\Big(\int_{t_0}^{t_1}\int_{t_0}^{t_1} \frac{|t - s|^{\theta q}}{|t - s|^{1+\beta q}}\,\mathrm{d}s\,\mathrm{d}t\Big)\,\|\partial_t f\|^{\theta q}_{C^{\gamma_0}_{x,t}}\,\|u\|^{(1-\theta)q}_{C^{\gamma_2}_{x,t}} \\
&= C\,|t_1 - t_0|^{\theta q - \beta q + 1}\,\|\partial_t f\|^{\theta q}_{C^{\gamma_0}_{x,t}}\,\|u\|^{(1-\theta)q}_{C^{\gamma_2}_{x,t}},
\end{aligned}$$

which gives (A.1). ■

Lemma A.4. *Suppose $\beta \geqslant 0$, $\mathcal{H}$ is a real Hilbert space and $C^\beta_x = C^\beta(\mathbb{R}^n; L^p(\Omega; \mathcal{H}))$. Assume $\mathcal{G}$ is a subalgebra of $\mathcal{F}$. Then*

$$\|\mathbf{E}(X|\mathcal{G})\|_{C^\beta_x} \leqslant \|X\|_{C^\beta_x}. \tag{A.2}$$

Moreover, for any $k \in \mathbb{N}$ with $k \leqslant \beta$,

$$\nabla^k \mathbf{E}(X(x)|\mathcal{G}) = \mathbf{E}(\nabla^k X(x)|\mathcal{G}). \tag{A.3}$$

Proof. We only prove (A.2) when $\beta \in (0, 1)$. Denote $\mathbf{E}^{\mathcal{G}} X(\cdot) := \mathbf{E}(X(\cdot)|\mathcal{G})$. By Jensen's inequality,

$$\begin{aligned}
\mathbf{E}\big|\mathbf{E}^{\mathcal{G}} X(x) - \mathbf{E}^{\mathcal{G}} X(y)\big|^p_{\mathcal{H}} &\leqslant \mathbf{E}\big[\mathbf{E}^{\mathcal{G}}|X(x) - X(y)|_{\mathcal{H}}\big]^p \leqslant \mathbf{E}\big[\mathbf{E}^{\mathcal{G}}\big|X(x) - X(y)\big|^p_{\mathcal{H}}\big] \\
&= \mathbf{E}|X(x) - X(y)|^p_{\mathcal{H}} \leqslant |x - y|^{\beta p}\,\|X\|^p_{C^\beta_x},
\end{aligned}$$

which yields

$$\|\mathbf{E}^{\mathcal{G}} X\|_{C^\beta_x} = \sup_{x,y\in\mathbb{R}^d} \frac{[\mathbf{E}|\mathbf{E}^{\mathcal{G}} X(x) - \mathbf{E}^{\mathcal{G}} X(y)|^p_{\mathcal{H}}]^{1/p}}{|x - y|^\beta} \leqslant \|X\|_{C^\beta_x}.$$

For (A.3), we only give the proof for $k = 1$. Again by Jensen's inequality,

$$\left|\mathbf{E}^{\mathcal{G}} X(x+h) - \mathbf{E}^{\mathcal{G}} X(x) - [\mathbf{E}^{\mathcal{G}} \nabla X(x)] \cdot h\right|_{\mathcal{H}} \leqslant \mathbf{E}^{\mathcal{G}} \left|X(x+h) - X(x) - \nabla X(x) \cdot h\right|_{\mathcal{H}}.$$

Thus,

$$\begin{aligned}\mathbf{E}\left|\mathbf{E}^{\mathcal{G}} X(x+h) - \mathbf{E}^{\mathcal{G}} X(x) - [\mathbf{E}^{\mathcal{G}} \nabla X(x)] \cdot h\right|_{\mathcal{H}}^p & \\ &\leqslant \mathbf{E}\,\mathbf{E}^{\mathcal{G}}\left(\left|X(x+h) - X(x) - \nabla X(x) \cdot h\right|_{\mathcal{H}}^p\right) \\ &= \left|X(x+h) - X(x) - \nabla X(x) \cdot h\right|_{\mathcal{B}}^p \to 0 \quad \text{as } h \to \infty,\end{aligned}$$

which gives the desired result. ■

Lemma A.5. *Suppose* $f : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$ *is* $\mathcal{B} \times \mathcal{F}$ *measurable and* $f \in C^1(\mathbb{R}^n; \mathbb{D}^{1,p})$. *Then* $\nabla f \in C(\mathbb{R}^n; \mathbb{D}^{1,p})$ *and*

$$\nabla D f = D \nabla f.$$

Proof. We assume $n = m = 1$ for simplicity. For any $x \in \mathbb{R}$, by definition,

$$\partial_{x,\theta} f(x) := \frac{f(x+\theta) - f(x)}{\theta} \stackrel{L^p(\Omega)}{\longrightarrow} \partial_x f(x) \ (\theta \to 0).$$

On the other hand, since $Df \in C^1(\mathbb{R}; L^p(\Omega, H))$, we have

$$D\partial_{x,\theta} f(x) = \frac{Df(x+\theta) - Df(x)}{\theta} \stackrel{L^p(\Omega;H)}{\longrightarrow} \partial_x(Df)(x) \ (\theta \to 0).$$

By the closability of the Malliavin derivative, we get $D\partial_x f(x) = \partial_x Df(x) \in \mathbb{D}^{1,p}$ and

$$\|\partial_x f(x)\|_{\mathbb{D}^{1,p}} = \liminf_{|\theta| \to 0} \left\| \frac{f(x+\theta) - f(x)}{\theta} \right\|_{\mathbb{D}^{1,p}} \leqslant \|f\|_{C^1(\mathbb{R};\mathbb{D}^{1,p})}.$$ ■

For any $F \in \mathbb{D}^{1,2}$, we have the following remarkable Clark–Ocone formula:

$$F = \mathbf{E}(F) + \int_0^1 \mathbf{E}^t D_t F \cdot dW_t := \mathbf{E}(F) + \sum_{k=1}^d \int_0^1 \mathbf{E}(D_t^k F | \mathcal{F}_t)\, dW_t^k. \tag{A.4}$$

The identity (A.4) implies the following simple lemma.

Lemma A.6. *Suppose* $F \in \mathbb{D}^{1,2}$. *Then, for each* $t \in [0, 1]$,

$$\mathbf{E}^t F = \mathbf{E} F + \int_0^t \mathbf{E}^s D_s F \cdot \mathrm{d}W_s. \tag{A.5}$$

Proof. By Clark–Ocone's formula,

$$m_t = \mathbf{E} F + \int_0^t \mathbf{E}^s D_s F \cdot \mathrm{d}W_s$$

is a $\mathcal{F}_t$-martingale with $m_1 = F$. Thus,

$$\mathbf{E}^t F = \mathbf{E}^t m_1 = m_t = \mathbf{E} F + \int_0^t \mathbf{E}^s D_s F \cdot \mathrm{d}W_s.$$ ■

The following lemma, which is a modification of Theorem 1.1 in [13], is needed in our proof of main result. A similar result for distributional valued processes can be found in [10].

Lemma A.7 (Itô–Wentzell's formula). *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ be a standard filtered probability space satisfying the common conditions. Let $p, p' \in [1, \infty]$ with $1/p + 1/p' = 1$ and $\alpha_1, \alpha_2 \in (0,1)$ with $\alpha_1 + \alpha_2 > 1$. Suppose $X_t = (X_t^1, \dots, X_t^n)$ are continuous semimartingales and let $\phi_t(x)$ be a random field continuous in $(x,t) \in Q$ almost surely. Assume ϕ and X satisfy*

(1) *for each $t \in [0,1]$, $\mathbb{R}^n \ni x \mapsto \phi_t(x) \in \mathbb{R}$ is C^2 continuous a.s.,*

(2) *for each $x \in \mathbb{R}^n$, $t \mapsto \phi_t(x)$ is a continuous $\mathcal{F}_t$-semimartingale represented as*

$$\phi_t(x) = \phi_0(x) + \int_0^t g_s(x)\,\mathrm{d}s + \int_0^t v_s^k(x)\,\mathrm{d}m_s^k, \tag{A.6}$$

where $m^1, \dots, m^d$ are continuous martingales, and the random fields g and v are locally bounded and

(a) *for each $x \in \mathbb{R}^n$, $t \mapsto g_t(x)$ and $t \mapsto v_t(x)$ are $\mathcal{F}_t$-adapted processes;*

(b) *for each $t \in [0,1]$, $x \mapsto v_t(x)$ is C^1 a.s.;*

(c) *for each $t \in [0,1]$, $x \mapsto g_t(x)$ is continuous, and*

$$\mathbf{E} \sup_{x \in B_N} \left| \nabla \int_{t_1}^{t_2} g_s(x)\,\mathrm{d}s \right|^p \lesssim_{p,N} |t_1 - t_2|^{\alpha_1 p},$$
$$\mathbf{E}|X_{t_1 \wedge \tau_N} - X_{t_2 \wedge \tau_N}|^{p'} \lesssim_{p'N} |t_1 - t_2|^{\alpha_2 p'},$$

where $\tau_N = \inf\{t > 0 : |X_t| > N\}$.

Then we have

$$\begin{aligned} d\phi_t(X_t) &= g_t(X_t)\,\mathrm{d}t + v_t^k(X_t)\,\mathrm{d}m_t^k + \partial_i \phi_t(X_t)\,\mathrm{d}X_t^i \\ &\quad + \frac{1}{2}\partial_{ij}\phi_t(X_t)\,\mathrm{d}\langle X^i, X^j\rangle_t + \partial_i v_t^k(X_t)\,\mathrm{d}\langle m^k, X^i\rangle_t. \end{aligned} \tag{A.7}$$

Proof. The proof is similar to that of Theorem 1.1 in [13]. Without loss of generality, we can assume $|X_t|$ is bounded by a constant N. For any $t > 0$, let $t_l = lt/n, l = 0, \dots, n$. Define $s(n) := t[sn/t]/n$ and $X_s^n := X_{s(n)}$. Then,

$$\begin{aligned} \phi_t(X_t) - \phi_0(X_0) &= \sum_{l=0}^{n-1} [\phi_{t_{l+1}}(X_{t_l}) - \phi_{t_l}(X_{t_l})] + \sum_{l=0}^{n-1} [\phi_{t_{l+1}}(X_{t_{l+1}}) - \phi_{t_{l+1}}(X_{t_l})] \\ &=: I_1^n + I_2^n. \end{aligned}$$

By (A.6) and the definition of X^n,

$$I_1^n = \sum_{l=0}^{n-1} \int_{t_l}^{t_{l+1}} g_s(X_{t_l})\,\mathrm{d}s + \sum_{l=0}^{n-1} \int_{t_l}^{t_{l+1}} v_s^k(X_{t_l})\,\mathrm{d}m_s^k = \int_0^t g_s(X_s^n)\,\mathrm{d}s + \int_0^t v_s^k(X_s^n)\,\mathrm{d}m_s^k.$$

Since $g_s(X^n_s) \to g_s(X_s)$, $v_s(X^n_s) \to v_s(X_s)$ a.s., and g and v are uniformly bounded in $[0,1] \times B_N$, we obtain

$$I^n_1 \xrightarrow{\mathbf{P}} \int_0^t g_s(X_s)\,\mathrm{d}s + \int_0^t v^k_s(X_s)\,\mathrm{d}m^k_s \quad \text{as } n \to \infty.$$

By Taylor expansion,

$$\begin{aligned} I^n_2 &= \sum_{l=0}^{n-1} \partial_i \phi_{t_{l+1}}(X_{t_l})(X^i_{t_{l+1}} - X^i_{t_l}) + \frac{1}{2}\sum_{l=0}^{n-1} \partial_{ij}\phi_{t_{l+1}}(\xi_l)(X^i_{t_{l+1}} - X^i_{t_l})(X^j_{t_{l+1}} - X^j_{t_l}) \\ &=: I^n_{21} + I^n_{22}, \end{aligned}$$

where ξ_l are some random variables between X_{t_l} and $X_{t_{l+1}}$. It is standard to show that

$$I^n_{22} \xrightarrow{\mathbf{P}} \frac{1}{2}\int_0^t \partial_{ij}\phi_s(X_s)\,\mathrm{d}\langle X^i, X^j\rangle_s \quad \text{as } n \to \infty.$$

For I^n_{21}, we rewrite it as

$$\begin{aligned} I^n_{21} &= \sum_{i=0}^{n-1} \partial_i \phi_{t_l}(X_{t_l})(X^i_{t_{l+1}} - X^i_{t_l}) + \sum_{l=0}^{n-1} [\partial_i \phi_{t_{l+1}}(X_{t_l}) - \partial_i \phi_{t_l}(X_{t_l})](X^i_{t_{l+1}} - X^i_{t_l}) \\ &=: I^n_{211} + I^n_{212}. \end{aligned}$$

Like before,

$$I^n_{211} \xrightarrow{\mathbf{P}} \int_0^t \partial_i \phi_s(X_s)\,\mathrm{d}X^i_s, \quad \text{as } n \to \infty.$$

Again by (A.6),

$$\begin{aligned} I^n_{212} &= \sum_{l=0}^{n-1} \partial_i \Big(\int_{t_l}^{t_{l+1}} g_s(X_{t_l})\,\mathrm{d}s\Big)(X^i_{t_{l+1}} - X^i_{t_l}) \\ &\quad + \sum_{l=0}^{n-1} \Big(\int_{t_l}^{t_{l+1}} \partial_i v^k_s(X_{t_l})\,\mathrm{d}m^k_s\Big)(X^i_{t_{l+1}} - X^i_{t_l}) =: I^n_{2121} + I^n_{2122}. \end{aligned}$$

By our assumption (c) and Hölder's inequality,

$$\begin{aligned} \mathbf{E}|I^n_{2121}| &\lesssim \sum_{l=0}^{n-1} \Big[\mathbf{E} \sup_{x\in B_N} \Big|\nabla \int_{t_{l+1}}^{t_l} g_s(x)\,\mathrm{d}s\Big|^p\Big]^{1/p} \Big[\mathbf{E}\big|X_{t_{l+1}} - X_{t_l}\big|^{p'}\Big]^{1/p'} \\ &\lesssim \sum_{l=0}^{n-1} |t_{l+1} - t_l|^{\alpha_1+\alpha_2} \lesssim n^{-\alpha_1-\alpha_2+1} \to 0 \quad \text{as } k \to \infty. \end{aligned}$$

It is standard to show

$$I^n_{2122} \xrightarrow{\mathbf{P}} \int_0^t \partial_i v^k(X_s)\,\mathrm{d}\langle m^k, X^i\rangle_s \quad \text{as } k \to \infty.$$

Combining all the above calculations, we obtain (A.7). ∎

Acknowledgements. The author would like to thank Professor Luo Dejun for raising this problem to him and also for many useful discussions.

Funding. Research of G. Zhao is supported by the German Research Foundation (DFG) through the Collaborative Research Centre(CRC) 1283 “Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

References

[1] Bahouri, H., Chemin, J.-Y. and Danchin, R.: *Fourier analysis and nonlinear partial differential equations*. Grundlehren der mathematischen Wissenschaften 343, Springer, Heidelberg, 2011.

[2] Beck, L., Flandoli, F., Gubinelli, M. and Maurelli, M.: Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electron. J. Probab.* **24** (2019), Paper no. 136, 72 pp.

[3] Du, K., Qiu, J. and Tang, S.: L^p theory for super-parabolic backward stochastic partial differential equations in the whole space. *Appl. Math. Optim.* **65** (2012), no. 2, 175–219.

[4] Duboscq, R. and Réveillac, A.: Stochastic regularization effects of semi-martingales on random functions. *J. Math. Pures Appl. (9)* **106** (2016), no. 6, 1141–1173.

[5] Fedrizzi, E. and Flandoli, F.: Pathwise uniqueness and continuous dependence for SDEs with non-regular drift. *Stochastics* **83** (2011), no. 3, 241–257.

[6] Flandoli, F.: *Random perturbation of PDEs and fluid dynamic models*. École d’Été de Probabilités de Saint-Flour, 2010. Lecture Notes in Mathematics 2015, Springer, Heidelberg, 2011.

[7] Flandoli, F., Gubinelli, M. and Priola, E.: Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* **180** (2010), no. 1, 1–53.

[8] John, F.: *Partial differential equations*. Third edition. Applied Mathematical Sciences 1, Springer-Verlag, New York-Berlin, 1978.

[9] Krylov, N. V.: On Kolmogorov’s equations for finite dimensional diffusions. In *Stochastic PDE’s and Kolmogorov equations in infinite dimensions*, pp. 1-63. Lecture Notes in Math. 1715, Springer, Berlin, 1999.

[10] Krylov, N. V.: On the Itô–Wentzell formula for distribution-valued processes and related topics. *Probab. Theory Related Fields* **150** (2011), no. 1-2, 295–319.

[11] Krylov, N. V. and Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** (2005), no. 2, 154–196.

[12] Krylov, N. V.: On strong solutions of Itô’s equations with $A \in W^1_d$ and $b \in L_d$. *Ann. Probab.* **49** (2021), no. 6, 3142–3167.

[13] Kunita, H.: Some extensions of Itô’s formula. In *Séminaire de Probabilités XV (Univ. Strasbourg, Strasbourg, 1979/1980)*, pp. 118–141. Lecture Notes in Math. 850, Springer, 1981.

[14] Mastrolia, T., Possamaï, D. and Réveillac, A.: On the Malliavin differentiability of BSDEs. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017), no. 1, 464–492.

[15] Menoukeu-Pamen, O., Meyer-Brandis, T., Nilssen, T., Proske, F. and Zhang, T.: A variational approach to the construction and Malliavin differentiability of strong solutions of SDE’s. *Math. Annal.* **357** (2013), no. 2, 761–799.

[16] Röckner, M. and Zhao, G.: Heat kernel and ergodicity of SDEs with distributional drifts: strong solutions. Preprint 2021, arXiv: 2103.05803.

[17] Stroock, D. W. and Varadhan, S. R.: Diffusion processes with continuous coefficients, I. *Comm. Pure Appl. Math.* **22** (1969), no. 3, 345–400.

[18] Tang, S. and Wei, W.: On the Cauchy problem for backward stochastic partial differential equations in Hölder spaces. *Ann. Probab.* **44** (2016), no. 1, 360–398.

[19] Veretennikov, A. Y.: Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S.)* **153** (1980), no. 3, 434–452.

[20] Xia, P., Xie, L., Zhang, X. and Zhao, G.: $L^q(L^p)$-theory of stochastic differential equations. *Stochastic Process. Appl.* **130** (2020), no. 8, 5188–5211.

[21] Zhang, X.: Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** (2011), no. 38, 1096–1116.

[22] Zhang, X.: Stochastic differential equations with Sobolev diffusion and singular drift and applications. *Ann. Appl. Probab.* **26** (2016), no. 5, 2697–2732.

[23] Zhang, X. and Zhao, G.: Heat kernel and ergodicity of SDEs with distributional drifts. Preprint 2017, arXiv: 1710.10537.

[24] Zhang, X. and Zhao, G.: Stochastic Lagrangian path for Leray's solutions of 3D Navier–Stokes equations. *Comm. Math. Phys.* **381**(2021), no. 2, 491–525.

[25] Zhao, G.: Stochastic Lagrangian flows for SDEs with rough coefficients. Preprint 2019, arXiv: 1911.05562.

[26] Zvonkin, A. K.: A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sb. (N.S.)* **22** (1974), no. 1, 129.

Received September 16, 2020. Published online December 4, 2021.

Guohuan Zhao
Faculty of Mathematics, Bielefeld University, 33615 Bielefeld, Germany;
zhaoguohuan@gmail.com