Hölder Flow and Differentiability for SDEs with Nonregular Drift

E. Fedrizzi[1], F. Flandoli[2]

(1) Université Paris Didérot, Sorbonne Paris Cité, Laboratoire de Probabilités et Modèles Aléatoires, Paris, France

(2) Dipartimento di Matematica Applicata, Università di Pisa, Italia

March 4, 2012

Abstract

We prove the existence of a stochastic flow of Hölder homeomorphisms for solutions of SDEs with singular time dependent drift having only certain integrability properties. We also show that the solution map $x \to X^x$ is differentiable in a weak sense.

1 Introduction

In this work, we consider the d–dimensional stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \mathrm{d}W_t \\ \qquad X_0 = x \end{cases} \tag{1.1}$$

in $[0, T] \times \mathbb{R}^d$ for singular drift coefficients b. Here, W_t is a standard Wiener process in $\mathbb{R}^d$. We only assume an integrability condition on b:

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

for some $p, q \in \mathbb{R}$ such that

$$p \geq 2, \quad q > 2, \quad \frac{d}{p} + \frac{2}{q} < 1. \tag{1.3}$$

Since the vector field b is not regular, we emphasize that solutions of the SDE are supposed to be such that (1.1) makes sense, that is,

$$P\bigg(\int_0^T \big|b(t, X_t)\big|\ \mathrm{d}t < \infty\bigg) = 1.$$

Krylov and Röckner [1] showed the existence and uniqueness of a local strong solution for this SDE, assuming only locally the integrability condition (1.2). This article has been one of the main sources of inspiration for our work. Many variants of this problem have also been studied by Zhang (see e.g., [2]), where the case of a diffusion coefficient different from the identity is studied, with slightly stronger assumptions on b, and following works.

This work has two main objectives: first, we show that under the same kind of hypothesis of [1] it is possible to construct, for the unique strong solution of (1.1), a stochastic flow of α–Hölder continuous homeomorphisms for every $\alpha < 1$; second, we prove that the solution map $x \to X^x$ is differentiable in some weak sense. We state now the precise result.

Definition 1.1. Take $T \in [0, \infty)$, $s \in [0, T]$ and let $b : \mathbb{R}^{d+1} \to \mathbb{R}^d$ be a vector field of class $L^q_p(T)$ with p, q satisfying (1.3). Let a filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ be given, with a Wiener process W_t defined on it. For every $x \in \mathbb{R}^d$, let $X^x_{s,t}$ be the unique continuous, $\{\mathscr{F}_{s,t}\}$-adapted, d-dimensional process defined for $t \in [s,T]$ on the filtered space $(\Omega, \mathscr{F}, \mathscr{F}_{s,t}, P)$ for which

$$P\bigg(\int_0^\infty \left| b(t, X_{s,t}) \right|^2 \, \mathrm{d}t < \infty \bigg) = 1$$

holds, and almost surely for all $t \in [s,T]$

$$X_{s,t} = x + \int_s^t b(r, X_r) \, \mathrm{d}r + W_t - W_s \, .$$

We call the process $X^x_{s,t}$ a (strong) solution of the SDE (1.1).

Theorem 1.2. *Let $X^x_{s,t}$ be the unique strong solution of (1.1). Then, for this solution there exists a stochastic flow of homeomorphisms $\psi_{s,t}(x)(\omega)$ which is α-Hölder continuous for every $\alpha < 1$. Moreover, the solution map $x \to X^x_{s,t}$ is differentiable in the following weak sense: let $e_1, ..., e_d$ be the canonical basis of $\mathbb{R}^d$; then, for every $i = 1, ..., d$ and every $x \in \mathbb{R}^d$, the limit*

$$\lim_{h \to 0} \frac{X_{s,\cdot}(x + he_i) - X_{s,\cdot}(x)}{h} \tag{1.4}$$

exists as a strong limit in $L^2\left(\Omega \times [0,T]\, ; \mathbb{R}^d\right)$.

Remark 1.3. Note that, even if we cannot provide a variational equation satisfied by the process X in the general case, there exists a transformed process $Y_t = \phi_t(X_t)$ (introduced below) which satisfy the variational equation (1.10). This will be shown in theorem 5.12.

Remark 1.4. The limit (1.4) exists as a strong limit also in $L^p\left(\Omega \times [0,T]\, ; \mathbb{R}^d\right)$ for every $p \geq 2$, see Remark 5.16.

These results are issued from a partially new and more quantitative approach to the problem, based on an idea of Flandoli, Gubinelli and Priola [3]. This strategy also allows for a proof of the well–posedness of equation (1.1) which we find to be somewhat easier than the one presented in [1]. Indeed, we do not use a *by–contradiction* argument, but rather exploit an explicit Zvonkin–type transformation [4] (Itô–Tanaka trick) and Gronwall inequality to obtain a better understanding of the dependence of the solution from the initial data. This allows us to show the existence of a flow of Hölder continuous homeomorphisms for the solution. Similar computations leads then to the result on the weak differentiability of the solution map. In a previous work [5], we used a similar approach leading to an iterative and probably even easier proof of the well-posedness of problem (1.1). However, since in the latter the transformed equation was not used explicitly, it was impossible to address the differentiability and much harder to prove the flow property; only a pathwise Hölder dependence on the initial data was proved in this work.

The choice of assuming a global integrability condition on b considerably simplifies the proofs of existence and uniqueness of solutions since no localization process is required; the extension of our proof to the case of a locally integrable b can be obtained by the same localization process used in [1], but we would then need to add specific hypothesis guaranteeing global existence to be able to construct the flow. For examples of conditions assuring the nonexplosion of solutions if b is only taken to be in $L^q_{p\ loc}$ we refer to [6], [1] and references therein.

In order to give a clear idea of the transformation used, we allow ourselves to perform here a few formal computations. Consider the vector-valued ($\mathbb{R}^d$-valued) backward PDE

$$\begin{cases} \frac{\partial U}{\partial t} + \frac{1}{2}\triangle U + b \cdot \nabla U = \lambda U - b \\ U(T,x) = 0, \end{cases} \tag{1.5}$$

which we call the PDE associated to the SDE (1.1) even if it is not the traditional associated Kolmogorov equation, and assume all functions are sufficiently regular. Then, Itô formula gives

$$\begin{aligned} \mathrm{d}U(t,X_t) &= \frac{\partial U}{\partial t}(t,X_t)\,\mathrm{d}t + \nabla U(t,X_t)\cdot\big(b(t,X_t)\,\mathrm{d}t + \mathrm{d}W_t\big) + \frac{1}{2}\triangle U(t,X_t)\,\mathrm{d}t \\ &= \lambda\, U(t,X_t)\,\mathrm{d}t - b(t,X_t)\,\mathrm{d}t + \nabla U(t,X_t)\cdot \mathrm{d}W_t \end{aligned} \tag{1.6}$$

and thus for the new process $Y_t := X_t + U(t,X_t)$ we have

$$\begin{aligned} \mathrm{d}Y_t &= b(t,X_t)\,\mathrm{d}t + \mathrm{d}W_t + \lambda\, U(t,X_t)\,\mathrm{d}t - b(t,X_t)\,\mathrm{d}t + \nabla U(t,X_t)\cdot \mathrm{d}W_t \\ &= \lambda U(t,X_t)\,\mathrm{d}t + \big(I + \nabla U(t,X_t)\big)\cdot \mathrm{d}W_t. \end{aligned}$$

We will show in section 3.1 that for every $t \in [0,T]$, the function

$$x \mapsto \phi_t(x) := x + U(t,x) \tag{1.7}$$

is an isomorphism. Then, the equation $Y_t = \phi_t(X_t)$ is equivalent to $X_t = \phi_t^{-1}(Y_t)$, and we have for $t \in [s,T]$

$$\mathrm{d}Y_t = \lambda U\big(t,\phi_t^{-1}(Y_t)\big)\,\mathrm{d}t + \Big[I + \nabla U\big(t,\phi_t^{-1}(Y_t)\big)\Big]\cdot \mathrm{d}W_t = \widetilde{b}(t,Y_t)\,\mathrm{d}t + \widetilde{\sigma}(t,Y_t)\cdot \mathrm{d}W_t, \tag{1.8}$$

where

$$\widetilde{b}(t,y) := \lambda U\big(t,\phi_t^{-1}(y)\big), \qquad \widetilde{\sigma}(t,y) := I + \nabla U\big(t,\phi_t^{-1}(y)\big). \tag{1.9}$$

The intuitive idea is that this equation has more regular coefficients than the original SDE. Therefore, it is easier to prove existence, uniqueness, the flow property for this solution and the weak differentiability of the solution map. Also, a variational equation for the process Y can be obtained: calling $\big(\xi_t^{x,i}\big)_{t\in[s,T]}$ the continuous adapted modification of the weak derivative of $x \to Y_t^x$ (in the sense of Theorem 1.2), we have that $\xi_t^{x,i}$ is the unique continuous adapted solution of the variational equation

$$\mathrm{d}\xi_t^{x,i} = \nabla\tilde{b}\,(t,Y_t^x)\,\xi_t^{x,i}\,\mathrm{d}t + \sum_{k=1}^{d} \nabla\tilde{\sigma}_k\,(t,Y_t^x)\,\xi_t^{x,i}\,\mathrm{d}W_t^k, \quad \xi_0^{x,i} = e_i. \tag{1.10}$$

The organization of the work is as follows. In section 2, we recall a classical approach based on Girsanov theorem to prove weak existence of solutions of the SDE (1.1). In section 3, we study the PDE (1.5), prove the existence and uniqueness of a fairly regular solution U and study the associated transformation (1.7). Then, in section 4, we prove the strong uniqueness property for the solutions of (1.1) and the transformed SDE (1.8), so that strong existence for the solutions of the two SDEs follows by the classical Yamada–Watanabe principle [7]. Finally, in the last section we prove the flow property for the solutions of the two SDEs, the weak derivability property for the solution map and the variational equation (1.10). The proof of Theorem 1.2 is contained in section 5.2.

Notation

$\mathcal{C}^0$ denotes the space of continuous functions, $\mathcal{C}^k$ and $\mathcal{C}^\infty$ the spaces of functions which are k-times differentiable and smooth functions, respectively. $\mathcal{C}^\alpha$ for $\alpha \in (0,1)$ denotes the space of Hölder continuous functions. The use of subscripts $\mathcal{C}_b$ and $\mathcal{C}_c$ indicates that we are working with a space of bounded or compact-support functions, respectively. $W^{\alpha,p}(\,\cdot\,) = (1-\triangle)^{\alpha/2}\, L^p\,(\,\cdot\,)$ is the standard Sobolev space. If $f(t,x)$ is a function of time and space, for its space norm we will use the short notation $\|f(t)\|$, which is a function of time, so that $\|f(t)\|_{L^p(\mathbb{R}^d)}$ will denote the L^p-norm in space only. When time and space are both involved in the norm, we will use superscripts to characterize the time-part of the norm and subscripts for the space-part: we will have $L_p^q(S,T) = L^q\big(S,T;L^p(R^d)\big)$ or $L_p^q(T) \equiv L_p^q(0,T)$. When working with PDEs (in section 3) we will have to use the spaces $\mathbb{H}^q_{\alpha,p}\,(T) = L^q\,\big(0,T\,;\,W^{\alpha,p}(R^d)\big)$, $\mathbb{H}_p^{\beta,q}\,(T) = W^{\beta,q}\,\big(0,T\,;\,L^p(R^d)\big)$ and especially $H^q_{\alpha,p}(T) = \mathbb{H}^q_{\alpha,p}(T) \cap \mathbb{H}_p^{1,q}(T)$.

2 Weak existence of solutions of the SDE

For self-containedness, we collect here a few known results which we will need to refer to throughout the article. They are taken from [1] and previous works, see for instance [8] and [9], and include weak existence of a solution X by Girsanov theorem, a formula for the density of the law of the solution with respect to Wiener measure, weak uniqueness and the exponential integrability of the process $\left|f(t,X_t)\right|^2$ when $f \in L_p^q(T)$ with p,q satisfying (1.3).

Lemma 2.1. *Let W_t^x be a d–dimensional Wiener process starting from the point x at time 0. Let f be a nonnegative Borel function on $\mathbb{R}^{d+1}$ belonging to $L_{p'}^{q'}$ for some $p',q' \in [1,\infty]$ such that*

$$\frac{d}{p'} + \frac{2}{q'} < 2. \tag{2.1}$$

Then there exist two positive constants N and ε depending only on p',q',d such that for any $t > s \geq 0$

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[\int_s^t f\big(r, W_{r-s}^x\big)\mathrm{d}r\Big] \leq N(t-s)^{\varepsilon}\, \|f\|_{L_{p'}^{q'}}. \tag{2.2}$$

Proof. With $\frac{1}{p}+\frac{1}{p'}=1$, $\frac{1}{q}+\frac{1}{q'}=1$, write explicitly the density of the Wiener process and use Hölder inequality first with respect to the space and then with respect to the time variables to obtain

$$\begin{aligned}
\mathbb{E}\Big[\int_s^t f\big(r, W_{r-s}^x\big)\mathrm{d}r\Big] &\leq \int_s^t \left(\int_{\mathbb{R}^d} f^p(r,y)\,\mathrm{d}y\right)^{1/p} \left(\int_{\mathbb{R}^d} \big(2\pi(r-s)\big)^{-p'd/2}\, e^{\frac{-p'|y|^2}{2(r-s)}}\,\mathrm{d}y\right)^{1/p'} \mathrm{d}r \\
&\leq (p')^{-\frac{d}{2p'}}\, \|f\|_{L_p^q} \left(\int_s^t \big(2\pi(r-s)\big)^{-q'\frac{d}{2}\frac{p'-1}{p'}}\,\mathrm{d}r\right)^{1/q'} \\
&= N\|f\|_{L_p^q}\, \big(t-s\big)^{1-1/q'-d/2p'}.
\end{aligned}$$

□

Remark 2.2. If $g \in L_p^q$ for some p,q satisfying (1.3), then $f=|g|^2 \in L_{p'}^{q'}$ for $p=2p'$, $q=2q'$, so that p',q' satisfy (2.1) and $\|g^2\|_{L_{p'}^{q'}(T)} \leq \|g\|^2_{L_p^q(T)}$. Therefore

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\left[\int_s^t g^2\left(r, x+W_{r-s}\right)\mathrm{d}r\right] \leq N(t-s)^{\varepsilon}\, \|g\|^2_{L_p^q(T)} \leq C_{p,q,d,\varepsilon,T}\|g\|^2_{L_p^q(T)}. \tag{2.3}$$

Corollary 2.3. *Let W_t^x be a d–dimensional Wiener process starting from the point x at time 0. Let $T \in [0,\infty)$ and f be a vector field of class $L_p^q(T)$ for some p,q satisfying (1.3). Then, there exists a constant K_f depending on d,p,q,T and $\|f\|_{L_p^q(T)}$ such that*

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[e^{\int_0^T \left|f(s,W_s^x)\right|^2 \mathrm{d}s}\Big] \leq K_f < \infty\,.$$

Proof. Since the estimate (2.3) is uniform in t,s,x, there exists a $\delta > 0$ such that

$$\sup_{s,x} \mathbb{E}\left[\int_s^{s+\delta} \left|f\left(r, x+W_{r-s}\right)\right|^2 \mathrm{d}r\right] \leq C_{p,q,d,T,\|g\|_{L_p^q(T)}} < 1.$$

By Khas'minskii's Lemma (see [10] or [11, lemma 2.1]),

$$\sup_{s,x} \mathbb{E}\left[\exp\Big(\int_s^{s+\delta} \left|f\left(r, x+W_{r-s}\right)\right|^2 \mathrm{d}r\Big)\right] \leq C'_{p,q,d,T,\|g\|_{L_p^q(T)}}$$

and one completes the proof splitting $[0,T]$ into a union of intervals of length $\leq \delta$ and using the Markov property of the Wiener process. □

Proposition 2.4. *Take $T \in [0,\infty)$ and let b be a vector field of class $L_p^q(T)$ with p,q satisfying (1.3). Let also W_t^x be a d–dimensional Wiener process defined on a filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ and starting from some point $x \in \mathbb{R}^d$ at $t=0$. Then all (positive and negative) moments of*

$$\rho_T := \exp\Big(\int_0^T b(s, W_s^x)\,\mathrm{d}W_s^x - \frac{1}{2}\int_0^T \big|b(s,W_s^x)\big|^2 \mathrm{d}s\Big) \tag{2.4}$$

are finite.

Proof. Applying Corollary 2.3 to $|b|^2$ we get the Novikov condition guaranteeing that

$$\rho_t = \exp\left(\int_0^t b(s, W_s^x)\,\mathrm{d}W_s^x - \frac{1}{2}\int_0^t \big|b(s,W_s^x)\big|^2 \mathrm{d}s\right) \tag{2.5}$$

is an exponential martingale, and in particular $\mathbb{E}[\rho_T]=1$. Take $\alpha \in \mathbb{R}$ and set $\bar b = 2\alpha b$, which is again an element of $L_p^q(T)$. Define the corresponding exponential martingale $\bar\rho_t$ with $\bar b$ in place of b. Then,

$$\begin{aligned}\mathbb{E}\big[\rho_T^\alpha\big] &\le \mathbb{E}\Big[\exp\Big(\int_0^T \bar b(s,W_s^x)\,\mathrm{d}W_s^x - \frac{1}{2}\int_0^t \big|\bar b(s,W_s^x)\big|^2\mathrm{d}s\Big)\Big]^{\frac{1}{2}} \\ &\qquad + \mathbb{E}\Big[\exp\Big((2\alpha^2-\alpha)\int_0^t \big|b(s,W_s^x)\big|^2\mathrm{d}s\Big)\Big]^{\frac{1}{2}} \\ &= \mathbb{E}\big[\ \bar\rho_T\ \big]^{\frac{1}{2}}\mathbb{E}\Big[\exp\Big((2\alpha^2-\alpha)\int_0^t \big|b(s,W_s^x)\big|^2\mathrm{d}s\Big)\Big]^{\frac{1}{2}},\end{aligned}$$

which is finite since $\mathbb{E}\big[\ \bar\rho_T\ \big]=1$ and the second term is finite due to Corollary 2.3 applied to $(2\alpha^2-\alpha)|b|^2$. □

By a classical application of Girsanov Theorem (see [1, lemma 3.2] for details) we have:

Theorem 2.5. *Given $T \in [0,\infty)$, $s \in [0,T]$, a vector field b of class $L_p^q(T)$ with p,q satisfying (1.3) and $x \in \mathbb{R}^d$, there exist processes X_t, W_t defined for $t \in [0,T]$ on a filtered space $(\Omega,\mathscr{F},\mathscr{F}_t,P)$ such that W_t is a d–dimensional $\{\mathscr{F}_t\}$–Wiener process and X_t is an $\{\mathscr{F}_t\}$–adapted, continuous, d–dimensional process for which*

$$P\left(\int_0^T |b(t,X_t)|^2\,\mathrm{d}t < \infty\right) = 1 \tag{2.6}$$

and almost surely, for all $t \in [s,T]$

$$X_t = x + \int_s^t b(r,X_r)\,\mathrm{d}r + W_t - W_s.$$

When both a solution X of equation (1.1) and the Wiener process itself satisfy condition (2.6), we may apply a result of absolutely continuous change of measures, see Liptser–Shiryaev [12, Theorems 7.7 and 7.9]. We know that the Wiener process satisfies this condition, when $b \in L_p^q(T)$, by remark 2.2. We have to impose by assumption the condition (2.6) on solutions.

Corollary 2.6. *Take $b \in L_p^q(T)$ for p,q such that (1.3) holds. Let (X,W) be a (weak) solution of equation (1.1) in the sense of theorem 2.5, in particular with X satisfying condition (2.6). Then, for any non negative Borel function Φ defined on the space $\mathcal{C}^0\big([0,T];\mathbb{R}^d\big)$ we have*

$$\mathbb{E}\left[\Phi(X)\right] = \mathbb{E}\left[\Phi(x+W)\ e^{\int_0^T b(s,x+W_s)\cdot \mathrm{d}W_s - 1/2\int_0^T |b(s,x+W_s)|^2\mathrm{d}s}\right]. \tag{2.7}$$

In particular, weak uniqueness holds for the equation (1.1), in the class of solutions satisfying (2.6). Moreover, if $f \in L_{\tilde p}^{\tilde q}(T)$ where $\tilde p, \tilde q$ are such that $d/\tilde p + 2/\tilde q < 1$, then, for any $k \in \mathbb{R}$ there exists a constant C_f depending on $\|f\|_{L_{\tilde p}^{\tilde q}(T)}$ such that

$$\mathbb{E}\left[e^{k\int_0^T |f(t,X_t)|^2\,\mathrm{d}t}\right] \le C_f. \tag{2.8}$$

Proof. The first part of the corollary depends on the above mentioned results of [12, theorems 7.7 and 7.9]. To prove the exponential integrability of $|f(t,X_t)|^2$, notice that by (2.7) we have

$$\mathbb{E}\left[e^{k\int_0^T|f(t,X_t)|^2\mathrm{d}t}\right]=\mathbb{E}\left[e^{\int_0^T b(s,x+W_s)\cdot\mathrm{d}W_s-1/2\int_0^T|b(s,x+W_s)|^2\,\mathrm{d}s+k\int_0^T|f(t,x+W_t)|^2\,\mathrm{d}t}\right]$$

and thus it is sufficient to repeat the estimates made above to prove that $\mathbb{E}\left[\rho_T^\alpha\right]$ was finite. □

With the same proof, namely

$$\mathbb{E}\Big[\,\big|X_t\big|^p\Big]=\mathbb{E}\Big[\,|x+W_t|^p\;e^{\int_0^T b(s,x+W_s)\cdot\mathrm{d}W_s-1/2\int_0^T|b(s,x+W_s)|^2\,\mathrm{d}s}\Big]$$

followed by Hölder inequality as in the proof made above to prove that $\mathbb{E}\left[\rho_T^\alpha\right]$ was finite, we also have:

Proposition 2.7. *Let (X,W) be a (weak) solution of equation (1.1). Then*

$$\sup_{t\in[0,T]}\mathbb{E}\left[|X_t|^p\right]<\infty$$

for every $p\geq 1$.

3 The associated parabolic problem

In this section, we collect all the analytical material regarding the PDE (1.5) and its solutions. In particular, what we need is a good regularity result for solutions of the PDE and the invertibility of the function $\phi_t(x)$ defined by (1.7). To ease notation, in the present and following sections we will always take the initial time $s=0$.

We will need the following easy modification of the classical Gronwall Lemma (for a proof, see [13]).

Lemma 3.1 (Modified Gronwall)**.** *Let f, g and v be measurable functions defined on $[0,T]$. Assume that f,g are positive and that for any $t\in[0,T]$, $g(s-t)$ and $\big(g(s-t)\,v(s)\big)$ belong to $L^1(t,T)$. If for any $t\in[0,T]$, v satisfies also the integral inequality*

$$v(t)\leq f(t)+\int_t^T g(s-t)\,v(s)\,\mathrm{d}s,$$

then, for all $t\in[0,T]$, v satisfies the Gronwall inequality

$$v(t)\leq f(t)+\int_t^T f(s)\,g(s-t)\,\exp\Big(\int_t^s g(r-t)\,\mathrm{d}r\Big)\,\mathrm{d}s.$$

The following lemma [1, Lemma 10.2] presents a key technical result: it provides the regularity of the functions belonging to the space $H^q_{2,p}(T)$. Much of what follows in the present section relies on this result.

Recall the definition of the following functional spaces:
$\mathbb{H}^q_{\alpha,p}(T)=L^q\big(0,T;W^{\alpha,p}(R^d)\big)$, $\mathbb{H}^{\beta,q}_p(T)=W^{\beta,q}\big(0,T;L^p(R^d)\big)$ and $H^q_{\alpha,p}(T)=\mathbb{H}^q_{\alpha,p}(T)\cap\mathbb{H}^{1,q}_p(T)$.

Lemma 3.2. *Let $p,q\in(1,\infty)$, $T\in(0,\infty)$ and $u\in H^q_{2,p}(T)$. Then we have:*

1. *If $\frac{d}{p}+\frac{2}{q}<2$ then $u(t,x)$ is a bounded Hölder continuous function on $[0,T]\times R^d$. More precisely, for any ε, $\delta\in(0,1]$ satisfying*

$$\varepsilon+\frac{d}{p}+\frac{2}{q}<2,\qquad 2\delta+\frac{d}{p}+\frac{2}{q}<2$$

there exists a constant N, depending only on p,q,ε,δ, such that for all $s,t\in[0,T]$ and $x,y\in R^d$, $x\neq y$ we have

$$|u(t,x)-u(s,x)|\leq N|t-s|^\delta\,\|u\|^{1-1/q-\delta}_{\mathbb{H}^q_{2,p}(T)}\,\|D_t\,u\|^{1/q+\delta}_{L^q_p(T)};\tag{3.1}$$

$$|u(t,x)|+\frac{|u(t,x)-u(t,y)|}{|x-y|^\varepsilon}\leq NT^{-1/q}\Big(\|u\|_{\mathbb{H}^q_{2,p}(T)}+T\|D_t\,u\|_{L^q_p(T)}\Big).\tag{3.2}$$

2. If $\frac{d}{p}+\frac{2}{q}<1$ then $\nabla u(t,x)$ is Hölder continuous in $[0,T]\times R^d$, namely for any $\varepsilon\in(0,1)$ satisfying

$$\varepsilon+\frac{d}{p}+\frac{2}{q}<1 \tag{3.3}$$

there exists a constant N, depending only on p,q,ε, such that for all $s,t\in[0,T]$ and $x,y\in R^d$, $x\neq y$, equations (3.1) and (3.2) holds with ∇u in place of u and $\delta=\varepsilon/2$.

Theorem 3.3 (Main PDE Theorem)**.** *Take p,q such that (1.3) holds and $\lambda>0$. Consider the vector fields $(b,f)(t,x):\mathbb{R}^{d+1}\to\mathbb{R}^d\ \in L^q_p(T)$. Then in $H^q_{2,p}(T)$ there exists a unique solution of the backward parabolic system*

$$\begin{cases} D_t u+\frac{1}{2}\Delta u+b\cdot\nabla u-\lambda u+f=0\\ u(T,x)=0.\end{cases} \tag{3.4}$$

For this solution there exists a finite constant N depending only on d,p,q,T,λ and $\|b\|_{L^q_p(T)}$ such that

$$\|u\|_{H^q_{2,p}(T)}\equiv\|D_t u\|_{L^q_p(T)}+\|u\|_{\mathbb{H}^q_{2,p}(T)}\leq N\|f\|_{L^q_p(T)}. \tag{3.5}$$

Proof. We want to emphasize that, since we are working in $\mathbb{R}^d$, (3.4) is actually a collection of d independent equations. In other words, (3.4) has to be interpreted componentwise. A similar result has already been proved in [1, Theorem 10.3]. For further details we direct the reader there, or to [13, Theorem 4.7].

This theorem for $b=0$, $\lambda=0$ was proved in [1, Theorem 10.3]. Therefore the *method of continuity* is applicable and we only need to prove the *a-priori* Schauder estimate (3.5) assuming that the solution already exists. In the following, K will indicate different constants depending only on d,p,q,T. Set $\tilde f:=f+b\cdot\nabla u-\lambda u$; [14, theorem 1.2] gives for $S\in[0,T]$

$$I(S):=\|D_t u\|^q_{L^q_p(S,T)}+\|u\|^q_{\mathbb{H}^q_{2,p}(S,T)}\leq K\Big(\|f\|^q_{L^q_p(S,T)}+\|b\cdot\nabla u\|^q_{L^q_p(S,T)}+\lambda^q\|u\|^q_{L^q_p(S,T)}\Big).$$

By Lemma 3.2 (we use that ∇u is Hölder continuous in time), $|\nabla u(t,x)|=\big|\nabla u(t,x)-\nabla u(T,x)\big|\leq KI^{\frac{1}{q}}(t)$. Furthermore

$$\begin{aligned}\|b\cdot\nabla u\|^q_{L^q_p(S,T)}&\leq\int_S^T\sup_x|\nabla u(t,x)|^q\|b(t,\cdot)\|^q_{L^p(\mathbb{R}^d)}\,\mathrm{d}t\leq K\int_S^T I(t)\|b(t,\cdot)\|^q_{L^p(\mathbb{R}^d)}\,\mathrm{d}t;\\ \|u\|^q_{L^q_p(S,T)}&\leq K\int_S^T\Big(\int_t^T\|D_s u(s)\|_{L^p(\mathbb{R}^d)}\,\mathrm{d}s\Big)^q\mathrm{d}t\leq K\int_S^T\int_t^T\|D_s u(s)\|^q_{L^p(\mathbb{R}^d)}\,\mathrm{d}s\,\mathrm{d}t\\ &\leq K\int_S^T I(t)\,\mathrm{d}t.\end{aligned}$$

Combining the above equations we get

$$I(S)\leq K\|f\|^q_{L^q_p(S,T)}+K\int_S^T I(t)\Big(\|b(t,\cdot)\|^q_{L^p(\mathbb{R}^d)}+\lambda^q\Big)\,\mathrm{d}t.$$

Finally, we estimate $I(0)$ by means of Gronwall inequality, and (3.5) follows. □

3.1 An invertibility result

To emphasize the dependence on λ, just for the rest of this section we will denote solutions of the PDE (3.4) as u_λ. As remarked in [5], the control of ∇u_λ (proven here in Lemma 3.4) is the key point to obtain the invertibility of the function ϕ_t of (1.7), which we will prove in Lemma 3.5.

Lemma 3.4 (Estimate for ∇u_λ)**.** *Let u_λ be the solution of (3.4). Then*

$$\sup_{t\in[0,T]}\|\nabla u_\lambda(t)\|_{\mathcal{C}^0(\mathbb{R}^d)}\to 0\qquad as\ \ \lambda\to+\infty.$$

Proof. Rewrite the PDE (3.4) as

$$D_t\, u_\lambda + \frac{1}{2}\triangle\, u_\lambda - \lambda\, u_\lambda = -(f + b\cdot\nabla u_\lambda) = g. \tag{3.6}$$

Denoting by P_t the heat semigroup, we have the well-known estimate

$$\big\|\nabla^\alpha P_t g\big\|_{L^p(\mathbb{R}^d)} \le C t^{-\frac{\alpha}{2}}\ \big\|g\big\|_{L^p(\mathbb{R}^d)}. \tag{3.7}$$

Extend the functions b and f to $[0,\infty)\times\mathbb{R}^d$ defining them as zero for $t > T$ (this does not change their L^q_p–norm) and keep the same notation for the extended functions. The solution u_λ can then be written as a convolution (the integrand is zero after time T):

$$u_\lambda(t,x) = \int_t^T e^{-\lambda(r-t)}\, P_{r-t}\, g(r,\cdot)(x)\ \mathrm{d}r.$$

Differentiating in the above formula we get

$$\big\|\nabla u_\lambda(t)\big\|_{L^\infty(\mathbb{R}^d)} \le \int_t^T e^{-\lambda(r-t)}\, \big\|\nabla P_{r-t} g(r)\big\|_{L^\infty(\mathbb{R}^d)}\ \mathrm{d}r. \tag{3.8}$$

By Sobolev embedding theorem $\big(W^{s,p}(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$ for $s \ge d/p\big)$ and (3.7), we get

$$\begin{aligned}\big\|\nabla P_{r-t} g(r)\big\|_{L^\infty(\mathbb{R}^d)} &\le C\Big(\|\nabla P_{r-t} g(r)\|_{L^p(\mathbb{R}^d)} + \|\nabla^s \nabla P_{r-t} g(r)\|_{L^p(\mathbb{R}^d)}\Big)\\ &\le \frac{C\|g(r)\|_{L^p(\mathbb{R}^d)}}{(r-t)^{1/2}} + \frac{C\|g(r)\|_{L^p(\mathbb{R}^d)}}{(r-t)^{(1+s)/2}}\\ &\le K_{p,d,T}\ t^{-(1+s)/2}\|g(r)\|_{L^p(\mathbb{R}^d)}.\end{aligned}$$

For $d/p < s < 1 - 2/q$ we have that $(1+s)/2 < 1 - 1/q = 1/q'$, implying that the positive function $\gamma(t) := K_{p,d,T}\ t^{-(1+s)/2}$ belongs to $L^{q'}(\mathbb{R}_+)$. Equation (3.6) itself gives $\|g(t)\|_{L^p(\mathbb{R}^d)} \in L^q(\mathbb{R}_+)$, so that by (3.8) and Hölder inequality we get

$$\begin{aligned}\big\|\nabla u_\lambda(t)\big\|_{L^\infty(\mathbb{R}^d)} \le &\int_t^T e^{-\lambda(r-t)}\, \gamma(r-t)\, \|f(r)\|_{L^p(\mathbb{R}^d)}\ \mathrm{d}r\\ &+ \int_t^T e^{-\lambda(r-t)}\, \gamma(r-t)\, \|b(r)\|_{L^p(\mathbb{R}^d)}\, \|\nabla u_\lambda(r)\|_{L^\infty(\mathbb{R}^d)}\ \mathrm{d}r.\end{aligned} \tag{3.9}$$

Since by Lemma 3.2 u_λ is bounded, for any $t \in [0,T]$ the functions $\gamma(r-t)\|b(r)\|_{L^p(\mathbb{R}^d)}$ and $u(r)\gamma(r-t)\|b(r)\|_{L^p(\mathbb{R}^d)}$, seen as functions of r, belong to $L^1(t,T)$. We can, therefore, apply the modified Gronwall lemma 3.1 to obtain from (3.9)

$$\big\|\nabla u_\lambda(t)\big\|_{L^\infty(\mathbb{R}^d)} \le \alpha(t) + \int_t^T \alpha(s)\, \gamma(s-t)\|b(s)\|_{L^p(\mathbb{R}^d)}\ \exp\Big(\int_t^s \gamma(r-t)\|b(r)\|_{L^p(\mathbb{R}^d)}\ \mathrm{d}r\Big)\,\mathrm{d}s, \tag{3.10}$$

where

$$\alpha(t) = \int_t^T e^{-\lambda(r-t)}\, \gamma(r-t)\, \|f(r)\|_{L^p(\mathbb{R}^d)}\ \mathrm{d}r \ge 0.$$

We can easily control the function α: keeping in mind the definition of γ, for any given $\varepsilon > 0$ it is possible to find a $\delta \in (0, T-t)$ such that $\|\gamma\|_{L^{q'}(0,\delta)} < \varepsilon$. Then, split the integral defining α on the interval $[t, t+\delta]$ and the complement of it, where the exponential term $e^{-\lambda\delta}$ is arbitrarily small (say, less than ε) for λ large enough:

$$\begin{aligned}\alpha(t) &\le \Big(\int_0^\delta \gamma(r)^{q'}\,\mathrm{d}r\Big)^{1/q'}\|f\|_{L^q_p(T)} + \Big(\int_\delta^{T-t} e^{-q'\lambda r}\,\gamma(r)^{q'}\,\mathrm{d}r\Big)^{1/q'}\|f(r)\|_{L^q_p(T)}\\ &\le \varepsilon\Big(1 + \|\gamma\|_{L^{q'}(0,T)}\Big)\|f\|_{L^q_p(T)}.\end{aligned} \tag{3.11}$$

Note that the bound on α we have obtained is uniform in t. Set

$$K_{T,d,p,q,s,\|b\|_{L^q_p(T)}} := \|\gamma\|_{L^{q'}(T)} \|b\|_{L^q_p(T)} \geq \int_t^s \gamma(r-t) \|b(r)\|_{L^p(\mathbb{R}^d)} \, \mathrm{d}r$$

and obtain from (3.10) and (3.11)

$$\big\|\nabla u_\lambda(t)\big\|_{L^\infty(\mathbb{R}^d)} \leq \varepsilon \|f\|_{L^q_p(T)} \Big(1 + \|\gamma\|_{L^{q'}(0,T)}\Big)\Big(1 + Ke^K\Big).$$

Since by lemma 3.2 ∇u_λ is a continuous function, the lemma is proved. ☐

Lemma 3.5 (Invertibility)**.** *Define*

$$\phi_\lambda(t,x) := x + u_\lambda(t,x) \tag{3.12}$$

and use the short notation $\phi^{-1}(t,\cdot)$ for the inverse function $\big(\phi(t,\cdot)\big)^{-1}(y)$. For λ large enough, such that

$$\sup_{t\geq 0} \|\nabla u_\lambda(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)} < 1/2, \tag{3.13}$$

the following statements hold:

1. *uniformly in $t \in [0,T]$, $\phi_\lambda(t,\cdot)$ has bounded first derivatives, globally Hölder continuous;*
2. *$\phi_\lambda(t,\cdot)$ is a $\mathcal{C}^1$–diffeomorphism for every $t \in [0,T]$;*
3. *$\phi_\lambda^{-1}(t,\cdot)$ has bounded first spatial derivatives, uniformly in t;*
4. *ϕ_λ and $\phi_\lambda(t,\cdot)^{-1}$ are continuous in (t,x).*

Proof. Note that the existence of a λ for which the estimate (3.13) holds is guaranteed by Lemma 3.4. The results of this lemma are mainly borrowed from [3, Lemma 6].

1. The uniform bound for the first derivatives of $\phi_\lambda(t,\cdot)$ follows by assumption (3.13); the Hölder continuity property follows from Lemma 3.2 because $u_\lambda \in H^q_{2,p}(T)$.

2. Recall the classical Hadamard theorem (see, e.g., [15, Theorem V.59, pag 330]): let $g : \mathbb{R}^d \to \mathbb{R}^d$ be of class $\mathcal{C}^1$; suppose that $\lim_{|x|\to\infty} |g(x)| = \infty$ and that the Jacobian matrix $\nabla g(x)$ is an isomorphism of $\mathbb{R}^d$ for all $x \in \mathbb{R}^d$. Then g is a $\mathcal{C}^1$ diffeomorphism of $\mathbb{R}^d$. Applying this result to $u_\lambda(t,\cdot)$ we get the assertion.

3. We know now that $\phi_\lambda^{-1}(t,\cdot)$ is of class $\mathcal{C}^1(\mathbb{R}^d)$, so that for all $y \in \mathbb{R}^d$

$$\nabla\phi_\lambda^{-1}(t,y) = \Big[\nabla\phi_\lambda\big(t,\phi_\lambda^{-1}(t,y)\big)\Big]^{-1} = \Big[I + \nabla u_\lambda\big(t,\phi_\lambda^{-1}(t,y)\big)\Big]^{-1} = \sum_{k\geq 0}\Big[-\nabla u_\lambda\big(t,\phi_\lambda^{-1}(t,y)\big)\Big]^k.$$

It follows that

$$\sup_{t\geq 0}\big\|\nabla\phi_\lambda^{-1}(t,\cdot)\big\|_{\mathcal{C}^0(\mathbb{R}^d)} \leq \sum_{k\geq 0}\Big[\sup_{t\geq 0}\big\|\nabla u_\lambda(t,\cdot)\big\|_{\mathcal{C}^0(\mathbb{R}^d)}\Big]^k < \infty.$$

4. ϕ_λ is continuous by definition. To see that also the inverse function $\phi_\lambda(t,\cdot)$ is continuous, assume by contradiction that there exists a sequence $\{t_n, y_n\}_n \subset [0,T]\times\mathbb{R}^d$ converging to (t,y) and such that $x_n := \phi_\lambda^{-1}(t_n,\cdot)(y_n) \nrightarrow x := \phi_\lambda^{-1}(t,\cdot)(y)$. If the sequence $\{x_n\}_n$ is bounded in $\mathbb{R}^d$, using the injectivity and the continuity of the function ϕ_λ on a convergent subsequence $x_{n_k} \to x' \neq x$, we get that

$$y_{n_k} := \phi_\lambda(t_{n_k}, x_{n_k}) \to \phi_\lambda(t,x') \neq \phi_\lambda(t,x) = y,$$

contradicting the assumed convergence of $y_{n_k} \to y$. If, instead, $|x_n| \to \infty$, we see from (3.12) that $|\phi_\lambda(t_n,x_n)| \to \infty$, because $|u_\lambda|$ is bounded on $[0,T]\times\mathbb{R}^d$. But this contradicts the fact that we had chosen $\{t_n, y_n\} = \{t_n, \phi_\lambda(t_n,x_n)\}$ to be convergent. ☐

4 Strong uniqueness

In this section, we prove strong uniqueness for the SDE (1.1) and Lemma 4.5, which is crucial whenever dealing with increments of functions in $\mathbb{H}^q_{1,p}(T)$ (as in the case of $\widetilde{\sigma}$). The main result reads.

Theorem 4.1. *Equation (1.1) has a unique strong solution.*

This result has already been proved in [1, theorem 3.7]. However, we find the proof given here a little easier: it follows more classical principles (Gronwall Lemma, for instance), and is more quantitative, with the advantage to allow the proof, in the next section, of additional results on the continuous dependence on initial conditions.

We have already noticed in Section 2 that there exist weak solutions of the SDE (1.1). By the classical Yamada-Watanabe theorem, or by the construction given by Gyongy-Krylov [16], we, therefore, have also the existence of strong solutions, due to strong uniqueness.

Let $X^{(1)}$ and $X^{(2)}$ be two solutions of the SDE (1.1), defined on the same filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ with the same Wiener process W_t. For uniqueness we assume that they have the same initial condition. However, for later use, let us develop some computations in the case of two possibly different initial conditions $x^{(1)}$ and $x^{(2)}$. Let $u_\lambda \in H^q_{2,p}(T)$ be the solution of the parabolic equation (1.5), corresponding to a choice of λ such that we have the invertibility results listed in the previous section. Thus, let ϕ_t be the invertible transformation $\phi_t(x) = u_\lambda(t,x) + x$, analyzed in Lemma 3.5.

Set $Y^{(i)}_t = \phi_t(X^{(i)}_t)$, $i = 1, 2$. These are solutions of the transformed SDE (1.8). The computations presented in the introduction (1.5)–(1.8), which show this fact, are rigorous, because of an extension of Itô formula to the class $H^q_{2,p}(T)$, proved by [1, theorem 3.7].

Theorem 4.2 (Itô formula)**.** *Let $p, q \in (1,\infty)$ be real numbers satisfying Condition (1.3) and $u \in H^q_{2,p}$. Let (X, W) be a solution of the SDE (1.1) provided by Theorem 2.5. Then, with probability one, (1.6) holds for every $t \in [0,T]$.*

It is useful to summarize the regularity of the coefficients $\widetilde{b}$ and $\widetilde{\sigma}$ of equation (1.8), defined by (1.9):

$$\widetilde{b}(t,x) = \lambda u\left(t, \phi_t^{-1}(x)\right); \qquad \widetilde{\sigma}(t,x) = I + \nabla u\left(t, \phi_t^{-1}(x)\right).$$

We have dropped the notation λ in u_λ since from now on λ is fixed, such that the invertibility properties discussed above hold true. The reader will notice that the main properties of these new fields depend on the uniform boundedness of ∇u, key property due to Assumption (1.3). Recall that in this section we still assume $s = 0$.

Proposition 4.3. *We have*

$$\nabla\widetilde{b} \in \mathcal{C}^0\big([0,T]; \mathcal{C}^0_b(\mathbb{R}^d)\big); \quad \widetilde{\sigma} \in \mathcal{C}^0\big([0,T]; \mathcal{C}^0_b(\mathbb{R}^d)\big); \tag{4.1}$$

$$\widetilde{\sigma} \in L^q\left(0,T; W^{1,p}(\mathbb{R}^d)\right). \tag{4.2}$$

Proof. By Lemma 3.2, $\nabla u \in \mathcal{C}^\varepsilon_b\left([0,T]\times\mathbb{R}^d\right)$ for some $\varepsilon \in (0,1)$. By definition of ϕ_t and the invertibility lemma 3.5, $\nabla\phi_t^{-1}(x)$ is bounded continuous. This gives us (4.1). Moreover, since $\nabla\phi_t^{-1}(x)$ is bounded, we have

$$\left|\,\partial_k\widetilde{\sigma}(t)\,\right| \le C\left|\nabla^2 u\big(t, \phi_t^{-1}(x)\big)\right|.$$

Hence

$$\int_{\mathbb{R}^d} \left|\partial_k\widetilde{\sigma}(t)\right|^p \mathrm{d}x \le C\int_{\mathbb{R}^d} \left|\nabla^2 u\big(t,\phi_t^{-1}(x)\big)\right|^p \mathrm{d}x \le C'\int_{\mathbb{R}^d} \left|\nabla^2 u(t,y)\right|^p \mathrm{d}y$$

because the Jacobian determinant of ϕ_t is bounded, because $\nabla\phi_t(x)$ is bounded. We easily get (4.2) and the proof is complete. □

The regularity of $\widetilde{\sigma}$ is not sufficient to apply standard results on SDEs. However, one can deal with increments of $\widetilde{\sigma}$ on different solutions by means of a process A_t introduced by Veretennikov [9]. Denote by $\frac{\left|\widetilde{\sigma}(s,Y_s^{(1)})-\widetilde{\sigma}(s,Y_s^{(2)})\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2} 1_{\left\{Y_s^{(1)}\neq Y_s^{(2)}\right\}}$ the non negative function equal to $\frac{\left|\widetilde{\sigma}(s,Y_s^{(1)})-\widetilde{\sigma}(s,Y_s^{(2)})\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2}$ when $Y_s^{(1)} \neq Y_s^{(2)}$ and equal to zero otherwise.

Set

$$A_t := \int_0^t \frac{\left|\widetilde{\sigma}\big(s,Y_s^{(1)}\big)-\widetilde{\sigma}\big(s,Y_s^{(2)}\big)\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2} 1_{\left\{Y_s^{(1)}\neq Y_s^{(2)}\right\}} \mathrm{d}s,$$

which a priori may be infinite.

The following lemma is proved in [1, lemma 5.4].

Lemma 4.4. *A_t is a real valued continuous, adapted, increasing process, such that $\mathbb{E}\left[A_T\right] < \infty$, and for every $t \in [0,T]$*

$$\int_0^t \left|\widetilde{\sigma}\left(s,Y_s^{(1)}\right)-\widetilde{\sigma}\left(s,Y_s^{(2)}\right)\right|^2 \mathrm{d}s = \int_0^t \left|Y_s^{(1)}-Y_s^{(2)}\right|^2 \mathrm{d}A_s. \tag{4.3}$$

This lemma admits an "exponential" version. We state and prove it here, to have in particular a self-contained proof of the previous lemma, but the exponential character will only be used only *after* the proof of strong uniqueness, in the next section.

Lemma 4.5. *For any $k \in \mathbb{R}$*

$$\mathbb{E}\left[e^{kA_T}\right] < \infty\,. \tag{4.4}$$

Proof. For $Y_s^{(1)} \neq Y_s^{(2)}$, we also have $X_s^{(1)} \neq X_s^{(2)}$ (and vice versa, so the functions $1_{\left\{Y_s^{(1)}\neq Y_s^{(2)}\right\}}$ and $1_{\left\{X_s^{(1)}\neq X_s^{(2)}\right\}}$ coincide) so, by the definitions of $\widetilde{\sigma}$ and $Y_s^{(1)}$, we may also write

$$\begin{aligned}\frac{\left|\widetilde{\sigma}\left(s,Y_s^{(1)}\right)-\widetilde{\sigma}\left(s,Y_s^{(2)}\right)\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2} &= \frac{\left|\nabla u\left(s,\phi_s^{-1}(Y_s^{(1)})\right)-\nabla u\left(s,\phi_s^{-1}(Y_s^{(2)})\right)\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2} \\ &= \frac{\left|\nabla u\left(s,X_s^{(1)}\right)-\nabla u\left(s,X_s^{(2)}\right)\right|^2}{\left|X_s^{(1)}-X_s^{(2)}\right|^2} \frac{\left|X_s^{(1)}-X_s^{(2)}\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2}.\end{aligned}$$

Notice that $\frac{\left|X_s^{(1)}-X_s^{(2)}\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2}$ is uniformly bounded by a constant C, because $X_t^{(i)} = \phi_t^{-1}(Y_t^{(i)})$, $i = 1, 2$, and $\nabla\phi_t^{-1}$ is uniformly bounded, by definition of ϕ_t and the invertibility Lemma 3.5. Hence

$$\frac{\left|\widetilde{\sigma}\left(s,Y_s^{(1)}\right)-\widetilde{\sigma}\left(s,Y_s^{(2)}\right)\right|^2}{\left|Y_s^{(1)}-Y_s^{(2)}\right|^2} \leq C\,\frac{\left|\nabla u\left(s,X_s^{(1)}\right)-\nabla u\left(s,X_s^{(2)}\right)\right|^2}{\left|X_s^{(1)}-X_s^{(2)}\right|^2}.$$

Thus, it is sufficient to prove that

$$\mathbb{E}\left[\exp\left(k\int_0^T \frac{\left|\nabla u\left(s,X_s^{(1)}\right)-\nabla u\left(s,X_s^{(2)}\right)\right|^2}{\left|X_s^{(1)}-X_s^{(2)}\right|^2} 1_{\left\{X_s^{(1)}\neq X_s^{(2)}\right\}} \mathrm{d}s\right)\right] \leq N_u < \infty \tag{4.5}$$

where N_u is a constant depending on $\|u\|_{H^q_{2,p}(T)}$. The space $\mathcal{C}_c^\infty\left([0,T]\times\mathbb{R}^d\right)$ is dense in $H^q_{2,p}(T)$, and if $u_n \to u$ in $H^q_{2,p}(T)$, then $\nabla u_n \to \nabla u$ uniformly in $[0,T]\times\mathbb{R}^d$, by Lemma 3.2. Thus, if we show (4.5) for $u \in \mathcal{C}_c^\infty\left([0,T]\times\mathbb{R}^d\right)$ with a constant N_u depending only on $\|u\|_{H^q_{2,p}(T)}$, the proof will be complete.

For smooth functions we have

$$\frac{\left|\nabla u\left(s, X_s^{(1)}\right) - \nabla u\left(s, X_s^{(2)}\right)\right|^2}{\left|X_s^{(1)} - X_s^{(2)}\right|^2} 1_{\left\{X_s^{(1)} \neq X_s^{(2)}\right\}} \leq \int_0^1 \left|\nabla^2 u\left(s, rX_s^{(1)} + (1-r)X_s^{(2)}\right)\right|^2 \mathrm{d}r.$$

Using the convexity of the exponential function, we obtain that the left-hand side of (4.5) is less than a constant times

$$\int_0^1 \mathbb{E}\left[\exp\left(k\int_0^T \left|\nabla^2 u\left(s, rX_s^{(1)} + (1-r)X_s^{(2)}\right)\right|^2 \mathrm{d}s\right)\right] \mathrm{d}r. \tag{4.6}$$

For any function $f(t,x) : [0,T]\times\mathbb{R}^d \to \mathbb{R}^k$, define a function $f^{(r)}(t,x,y) : [0,T]\times\mathbb{R}^{2d} \to \mathbb{R}^{2k}$ as

$$f^{(r)}(t,x,y) := \big(r\, f(t,x), (1-r) f(t,y)\big)$$

and use the notations

$$X_s^{(r)} = \big(rX_s^{(1)}, (1-r)X_s^{(2)}\big), \quad x^{(r)} = \big(rx^{(1)}, (1-r)x^{(2)}\big), \quad W_s^{(r)} = \big(rW_s, (1-r)W_s\big),$$

where $X_s^{(r)}, x^{(r)}, W_s^{(r)} \in \mathbb{R}^{2d}$. We have

$$X_t^{(r)} = x^{(r)} + \int_0^t b_s^{(r)}\,\mathrm{d}s + W_t^{(r)}.$$

We also define

$$\rho_r := \exp\left(-\int_0^T b_t^{(r)}\cdot \mathrm{d}W_t^{(r)} - \frac{1}{2}\int_0^T |\bar{b}_t^{(r)}|^2\,\mathrm{d}t\right)$$

and obtain

$$\mathbb{E}\left[\exp\left(\lambda\int_0^T \left|b_t^{(r)}\right|^2 \mathrm{d}t\right)\right] = \mathbb{E}\left[\exp\left(\lambda\, r^2 \int_0^T \left|b\left(t, X_t^{(1)}\right)\right|^2 \mathrm{d}t\right)\, \exp\left(\lambda\,(1-r)^2 \int_0^T \left|b\left(t, X_t^{(2)}\right)\right|^2 \mathrm{d}t\right)\right],$$

which is finite (by Hölder inequality) using the exponential estimates on solutions of equation (1.1) proved in Corollary 2.6. Hence, Novikov condition is fulfilled; by Girsanov theorem, $X_t^{(r)}$ is a Wiener process from $x^{(r)}$, on $\big(\Omega, \mathscr{F}, \mathscr{F}_t, Q^{(r)}\big)$ with

$$\left.\frac{\mathrm{d}Q^{(r)}}{\mathrm{d}P}\right|_{\mathscr{F}_T} = \rho_T^{(r)} := \exp\Big(-\int_0^T b_t^{(r)}\;\cdot \mathrm{d}W_t^{(r)} - \frac{1}{2}\int_0^T |b_t^{(r)}|^2\,\mathrm{d}t\Big).$$

Therefore we obtain (we indicate by superscripts the measure used in the expected values)

$$\begin{aligned}
\mathbb{E}^P\left[\exp\left(k\int_0^T \left|(\nabla^2 u)^{(r)}(s, X_s^{(r)})\right|^2 \mathrm{d}s\right)\right] &= \mathbb{E}^P\left[\rho_r^{-1/2}\rho_r^{1/2}\exp\left(k\int_0^T \left|(\nabla^2 u)^{(r)}(s, X_s^{(r)})\right|^2 \mathrm{d}s\right)\right] \\
&\leq C\,\mathbb{E}^P\left[\rho_r \exp\left(2k\int_0^T \left|(\nabla^2 u)^{(r)}(s, X_s^{(r)})\right|^2 \mathrm{d}s\right)\right]^{1/2} \\
&= C\,\mathbb{E}^{Q^{(r)}}\left[\exp\left(2k\int_0^T \left|\nabla^2 u(s, W_s)\right|^2 \mathrm{d}s\right)\right]^{1/2}.
\end{aligned}$$

This is bounded by a constant depending only on the L_p^q norm of $\nabla^2 u$, hence by the $H^q_{2,p}$-norm of u, and k. The proof is complete. □

Proof of theorem 4.1. The property $X^{(1)} = X^{(2)}$ (when $x^{(1)} = x^{(2)}$) is equivalent to $Y^{(1)} = Y^{(2)}$. Let us prove the latter identity.

As we will need them in the next section, we will develop the computations in a setting slightly more general than necessary for the present proof, namely we do not assume for now that $x^1 = x^2$ and allow a to be any real number larger or equal to 2.

From the equations for $Y^{(i)}$ and Itô formula, we get

$$\frac{1}{a}\,\mathrm{d}\left|Y_s^{(1)} - Y_s^{(2)}\right|^a = \left\langle \widetilde{b}\left(s, Y_s^{(1)}\right) - \widetilde{b}\left(s, Y_s^{(2)}\right),\ \left(Y_s^{(1)} - Y_s^{(2)}\right)^{a-1}\right\rangle \mathrm{d}s$$
$$+ \left\langle \left(\widetilde{\sigma}\left(s, Y_s^{(1)}\right) - \widetilde{\sigma}\left(s, Y_s^{(2)}\right)\right)\cdot \mathrm{d}W_s,\ \left(Y_s^{(1)} - Y_s^{(2)}\right)^{a-1}\right\rangle$$
$$+ \frac{a-1}{2}\mathrm{Trace}\left(\left[\widetilde{\sigma}\left(s, Y_s^{(1)}\right) - \widetilde{\sigma}\left(s, Y_s^{(2)}\right)\right]\left[\widetilde{\sigma}\left(s, Y_s^{(1)}\right) - \widetilde{\sigma}\left(s, Y_s^{(2)}\right)\right]^T\right)\left(Y_s^{(1)} - Y_s^{(2)}\right)^{a-2}\mathrm{d}s.$$

Using equations (1.8) and (4.1) we see that Y_t^{a-1} is a square integrable process. Thus, using the boundedness of $\nabla\widetilde{b}$ for the first term, the boundedness of $\widetilde{\sigma}$ to know that the second term is of the form $\mathrm{d}M_s$ where M_s is a martingale, and the definition of A_t for the last term, we get

$$\mathrm{d}\left|Y_s^{(1)} - Y_s^{(2)}\right|^a \le C\left|Y_s^{(1)} - Y_s^{(2)}\right|^a \mathrm{d}s + a\,\mathrm{d}M_s + \frac{a(a-1)}{2}\left|Y_s^{(1)} - Y_s^{(2)}\right|^a \mathrm{d}A_s.$$

Set $k := \frac{a(a-1)}{2} \ge 1$. Then

$$\mathrm{d}\left(e^{-kA_s}\left|Y_s^{(1)} - Y_s^{(2)}\right|^a\right) \le C\,e^{-kA_s}\left|Y_s^{(1)} - Y_s^{(2)}\right|^a \mathrm{d}s + a\,e^{-kA_s}\mathrm{d}M_s,$$

which leads to

$$\mathbb{E}\left[e^{-kA_t}\left|Y_t^{(1)} - Y_t^{(2)}\right|^a\right] \le \left|Y_0^{(1)} - Y_0^{(2)}\right|^a + \int_0^t C\,\mathbb{E}\left[e^{-kA_s}\left|Y_s^{(1)} - Y_s^{(2)}\right|^a\right]\mathrm{d}s$$

using the fact that $\int_0^t e^{-kA_s}\mathrm{d}M_s$ is a martingale (M is a martingale and e^{-kA_s} is bounded by 1). By Grönwall lemma this gives us

$$\mathbb{E}\left[e^{-kA_t}\left|Y_t^{(1)} - Y_t^{(2)}\right|^a\right] \le \left|Y_0^{(1)} - Y_0^{(2)}\right|^a e^{CT}, \tag{4.7}$$

hence $P\left(\left|Y_t^{(1)} - Y_t^{(2)}\right|^2 = 0\right) = 1$ if $x^1 = x^2$ because $Y_0^{(1)} = Y_0^{(2)}$ and $e^{-kA_t} > 0$ almost surely. This implies $Y^{(1)} = Y^{(2)}$ (by the continuity of trajectories). The proof is complete. □

For completeness, let us state the well posedness also for the auxiliary equation (1.8). This is almost obvious but remind that the previous proof started from solutions of (1.1), not (1.8).

Theorem 4.6. *Equation (1.8) has a unique strong solution.*

Proof. Strong existence is true since $Y_t = \phi_t(X_t)$ is a strong solution of (1.8), if X is a strong solution of (1.1). Let $Y^{(1)}$ and $Y^{(2)}$ be two solutions on the same filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ with the same Wiener process W_t. With tedious computations one can show that $X_t^{(i)} = \phi_t^{-1}(Y_t^{(i)})$, $i = 1, 2$, are solutions of (1.1), hence they coincide. Then $Y_t^{(i)}$ coincide. □

5 Flow property for solutions of SDEs

In this section, we show the existence of a stochastic flow of Hölder-continuous homeomorphisms for the SDE (1.1) and the weak differentiability of the solution map $x \to X_{s,t}^x$, thus completing the proof of Theorem 1.2.

Recall the following relevant definition from [17].

Definition 5.1. A *stochastic flow of homeomorphisms* (or simply a *flow*) on the filtered space with a Wiener process $(\Omega, \mathscr{F}, \mathscr{F}_t, P, W_t)$ associated with the SDE (1.1) is a function $(s,t,x,\omega) \mapsto \varphi_{s,t}(x)(\omega)$, defined for $0 \le s \le t \le T$, $x \in \mathbb{R}^d$ and $\omega \in \Omega$ with values in $\mathbb{R}^d$, such that for any $s \in [0,T]$

1. for any $x \in \mathbb{R}^d$, the process $X^{s,x} = \{X^x_{s,t} : t \in [s,T]\}$ defined as $X^x_{s,t} := \varphi_{s,t}(x)$ is a continuous $\{\mathscr{F}_{s,t}\}$-adapted solution of (1.1);
2. P-almost surely, $\varphi_{s,t}(x) : \mathbb{R}^d \to \mathbb{R}^d$ is a homeomorphism: $\varphi_{s,t}$ and $\varphi^{-1}_{s,t}$ are one to one and onto for every $t \in [s,T]$, and they are continuous in (t,x);
3. P-almost surely, $\varphi_{s,t}(x) = \varphi_{u,t}\big(\varphi_{s,u}(x)\big)$ for all $s \le u \le t \le T$ and $x \in \mathbb{R}^d$, and $\varphi_{s,s}(x) = x$.

To obtain the existence of a flow for (1.1) we need to prove it first for the SDE (1.8) (see Theorem 5.10 below). We have two natural candidates for the flows associated to the SDEs considered: the function

$$\psi_{s,t}(x)(\omega) := X^x_{s,t}(\omega) = \phi^{-1}_{s,t}\big(Y^{\phi_{s,s}(x)}_{s,t}\big) \tag{5.1}$$

for (1.1) and the function

$$\varphi_{s,t}(x)(\omega) := Y_{s,t}(x)(\omega) \tag{5.2}$$

for the transformed SDE (1.8). Note that points 1 and 3 of Definition 5.1 follow directly from the above definitions. Therefore, to prove that ψ and φ indeed define stochastic flows of homeomorphisms for the two SDEs we only need to show that they are homeomorphisms. Many of the proofs of this section are adapted from [17]; further details can also be found in [13]. Here $\widetilde{\sigma}$ is not Lipschitz continuous, so that we will have to deal with increments of $\widetilde{\sigma}$ on different solutions using again the process A_t introduced in Lemma 4.4.

5.1 Existence of a Flow of Homeomorphisms for the New SDE

5.1.1 Continuity

Solutions of (1.8) depend continuously on the initial data; this is a key point to construct the flow.

Proposition 5.2 (Continuity Estimate 1)**.** *Let $s \in [0,T]$. For $t \in [s,T]$ let Y^s be the unique strong solution of the transformed equation (1.8) starting at time s. Then, for any $a \ge 2$ there exists a positive constant C_a such that for any $r, t \in [s,T]$ and $x, y \in \mathbb{R}^d$,*

$$\mathbb{E}\Big[\big|Y^x_{s,t} - Y^y_{s,r}\big|^a\Big] \le C_a\Big(|x-y|^a + |t-r|^{\frac{a}{2}}\Big). \tag{5.3}$$

Proof. Consider the case $t \ge r$ and note that it suffices to show that

$$\begin{aligned} \mathbb{E}\Big[\big|Y^x_{s,r} - Y^y_{s,r}\big|^a\Big] &\le C|x-y|^a; \\ \mathbb{E}\Big[\big|Y^x_{s,t} - Y^x_{s,r}\big|^a\Big] &\le C_a|t-r|^{\frac{a}{2}}. \end{aligned} \tag{5.4}$$

The first inequality has already been proved in (4.7) and the second one easily follows from the boundedness of $\widetilde{b}$ and $\widetilde{\sigma}$ provided by (4.1). □

Remark 5.3. Simply applying Kolmogorov regularity theorem, one can prove the existence of a modification $\widetilde{Y}^x_{s,t}$ of the solution $Y^x_{s,t}$ such that the function $\varphi_s(\omega)(\cdot,\cdot) : [s,T]\times\mathbb{R}^d \to \mathbb{R}^d$ mapping $(t,x) \mapsto \widetilde{Y}^x_{s,t}$ is continuous for almost every $\omega \in \Omega$. It is also immediate to verify that $\varphi_s(\omega)$ is (α,β)-Hölder continuous in (t,x), almost surely in ω, for any $\alpha < \frac{1}{2}$ and $\beta < 1$.
In what follows, we will always work with this continuous modification, which we will still denote by Y.

5.1.2 Injectivity

Lemma 5.4. *(See [17, lemma II.2.4]) Let a be any real number and $\varepsilon > 0$. Then there is a positive constant C_a (independent of ε) such that*

$$\mathbb{E}\Big[\Big(\varepsilon + \big|Y^x_{s,t} - Y^y_{s,t}\big|^2\Big)^a\Big] \le C_a\left(\varepsilon + |x-y|^2\right)^a \tag{5.5}$$

holds for any $s \in [0,T]$, $t \in [s,T]$ and $x, y \in \mathbb{R}^d$.

Proof. Fix any $s \in [0,T]$, $t \in [s,T]$, $x \in \mathbb{R}^d$ and set $F(x) := f^a(x)$, $f(x) := (\varepsilon + |x|^2)$. Set also $\eta_t := Y^x_{s,t} - Y^y_{s,t}$, so that applying Itô formula we obtain

$$\begin{aligned}F(\eta_t) - F(\eta_s) =& 2a\int_s^t f^{a-1}(\eta_r)\eta_r\Big[\widetilde{b}\big(r, Y^x_{s,r}\big) - \widetilde{b}\big(r, Y^y_{s,r}\big)\Big]\mathrm{d}r + 2a\int_s^t f^{a-1}(\eta_r)\eta_r\Big[\widetilde{\sigma}\big(r, Y^x_{s,r}\big) - \widetilde{\sigma}\big(r, Y^y_{s,r}\big)\Big]\mathrm{d}W_r \\ &+ a\sum_{i,j}\int_s^t f^{a-2}(\eta_r)\Big[f(\eta_r)\delta_{i,j} + 2(a-1)\,\eta^i_r\eta^j_r\Big] \\ &\qquad\times\Big[\Big(\widetilde{\sigma}\big(r, Y^x_{s,r}\big) - \widetilde{\sigma}\big(r, Y^y_{s,r}\big)\Big)\Big(\widetilde{\sigma}\big(r, Y^x_{s,r}\big) - \widetilde{\sigma}\big(r, Y^y_{s,r}\big)\Big)^t\Big]^{i,j}\mathrm{d}r.\end{aligned}$$

Recall that $|x| \le f^{1/2}(x)$ and that the coefficient $\widetilde{b}$ is Lipschitz continuous: $|\widetilde{b}(t,x) - |\widetilde{b}(t,y)| \le L f^{1/2}(x-y)$. Then

$$\begin{aligned}F(\eta_t) - F(\eta_s) \le& 2|a|\int_s^t f^a(\eta_r)\,\mathrm{d}r + 2\,a\int_s^t f^{a-1}(\eta_r)\eta_r\Big[\widetilde{\sigma}\big(r, Y^x_{s,r}\big) - \widetilde{\sigma}\big(r, Y^y_{s,r}\big)\Big]\,\mathrm{d}W_r \\ &+ |a|\,C\int_s^t f^{a-1}(\eta_r)\big(\delta_{i,j} + 2|a-1|\big)\,|\eta_r|^2\,\mathrm{d}A_r \\ \le& C_a\bigg(\int_s^t f^a(\eta_r)\,\mathrm{d}r + \int_s^t f^{a-1}(\eta_r)\eta_r\Big[\widetilde{\sigma}\big(r, Y^x_{s,r}\big) - \widetilde{\sigma}\big(r, Y^y_{s,r}\big)\Big]\,\mathrm{d}W_r + \int_s^t f^a(\eta_r)\,\mathrm{d}A_r\bigg).\end{aligned}$$

Here, A_t is the process introduced in Lemma 4.4. Note that the central term of the last line above is a martingale because, as remarked in the proof of Theorem 4.1, all moments of Y are finite. Since $e^{-A_t} \le 1$, Grönwall inequality gives

$$\mathbb{E}\Big[e^{-A_t}\Big(\varepsilon + \big|Y^x_{s,t} - Y^y_{s,t}\big|^2\Big)^a\Big] = \mathbb{E}\Big[e^{-A_t}F(\eta_t)\Big] \le C_{a,T}\,F(x-y) = C_{a,T}\left(\varepsilon + |x-y|^2\right)^a.$$

To complete the proof of the lemma it suffices to apply Hölder inequality and to estimate $\mathbb{E}[e^{2A_T}]$ using Lemma 4.5:

$$\mathbb{E}\Big[F(\eta_t)\Big]^2 \le \mathbb{E}\Big[e^{2A_t}\Big]\mathbb{E}\Big[e^{-2A_t}F^2(\eta_t)\Big] \le \mathbb{E}\Big[e^{2A_T}\Big]C_{a,T}\,F^2(x-y) = C_{a,T}\,F^2(x-y).$$

□

Corollary 5.5. *Let ε tend to zero in lemma 5.4. Then, by monotone convergence, we have:*

$$\mathbb{E}\Big[\big|Y^x_{s,t} - Y^y_{s,t}\big|^a\Big] \le C_{a,T}\,|x-y|^a. \tag{5.6}$$

From this inequality in the case $a < 0$ we get that $x \ne y \Rightarrow Y^x_{s,t} \ne Y^y_{s,t}$ almost surely for any $s < t$. Kunita calls this property "weak injectivity".

Lemma 5.6. *(See [17, lemma II.4.1]) For every fixed $s \in [0,T]$, set*

$$\eta_t(x,y) := \frac{1}{|Y^x_{s,t} - Y^y_{s,t}|}$$

for $t \in [s,T]$. Then for any $a > 2$ there exists a constant C_a such that for any $\delta > 0$

$$\mathbb{E}\Big[\big|\eta_t(x,y) - \eta_{t'}(x',y')\big|^a\Big] \le C_a\, \delta^{-2a}\big[|x-x'|^a + |y-y'|^a + |t-t'|^{a/2}\big] \tag{5.7}$$

holds for any $t, t' \in [s,T]$ and $|x - x'| \ge \delta$, $|y - y'| \ge \delta$.

Proof. To ease notation, write η and η' for $\eta_t(x,y)$ and $\eta_{t'}(x',y')$ respectively and set $\xi = \eta^{-1}$ and $\xi' = \eta'^{-1}$. Simple algebraic computations yield

$$\Big|\eta_t(x,y) - \eta_{t'}(x',y')\Big|^a \le C_a|\eta|^a|\eta'|^a\left(|Y^{x'}_{t'} - Y^x_t|^a + |Y^{y'}_{t'} - Y^y_t|^a\right). \tag{5.8}$$

To complete the proof, we take expectations in (5.8) and use Hölder inequality, Corollary 5.5 and Proposition 5.2 to estimate the different terms:

$$\begin{aligned}
&\mathbb{E}\Big[\big|\eta_t(x,y) - \eta_{t'}(x',y')\big|^a\Big] \\
&\quad\le C_a\,\mathbb{E}\big[|\eta|^{4a}\big]^{1/4}\mathbb{E}\big[|\eta'|^{4a}\big]^{1/4}\bigg(\mathbb{E}\Big[|Y^{x'}_{s,t'} - Y^x_{s,t}|^{2a}\Big]^{1/2} + \mathbb{E}\Big[|Y^{y'}_{s,t'} - Y^y_{s,t}|^{2a}\Big]^{1/2}\bigg) \\
&\quad\le C_a|x-y|^{-a}|x'-y'|^{-a}\Big(|x-x'|^a + |y-y'|^a + 2(t-t')^{a/2}\Big).
\end{aligned}$$

□

Theorem 5.7. *The map $Y_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ is one to one for any $t \in [s,T]$ almost surely.*

Proof. Take $a/2 > 2(d+1)$ in Lemma 5.6. Kolmogorov theorem states that $\eta_t(x,y)$ is continuous in (t,x,y) in the domain $\{(t,x,y)|\, t \in [s,T], |x-y| \ge \delta\}$. Since δ is arbitrary, it is also continuous in the domain $\mathcal{D} := \{(t,x,y)|\, t \in [s,T], x \ne y\}$. Note that $\mathcal{D}$ has at most two connected components, every one of which intersects the hyperplane $\{t = s\}$. Then, since η_s is finite, η_t must be finite on all of $\mathcal{D}$. Therefore, if $x \ne y$, $Y^x_{s,t} \ne Y^y_{s,t}$, and the theorem is proved. □

5.1.3 Surjectivity

The original proof of the following lemma, due to Kunita [17, lemma II.4.2], requires no modification, so we will only report a sketch of it for completeness.

Lemma 5.8. *Fix any $s \in [0,T]$ and let $\widehat{\mathbb{R}}^d$ be the one point compactification of $\mathbb{R}^d$; set $\widehat{x} := x/|x|^2$ for $x \in \mathbb{R}^d \backslash \{0\}$ and $\widehat{x} := \infty$ for $x = 0$. Define for every $t \in [s,T]$*

$$\eta_t(\widehat{x}) := \begin{cases} \dfrac{1}{1+|Y^x_{s,t}|} & if\ x \in \mathbb{R}^d \\ 0 & if\ \widehat{x} = 0 \end{cases}.$$

Then, for any $a \in (0,\infty)$ there exists a constant C_a such that

$$\mathbb{E}\Big[\big|\eta_t(\widehat{x}) - \eta_{t'}(\widehat{y})\big|^a\Big] \le C_a\Big(|\widehat{x} - \widehat{y}|^a + |t-t'|^{a/2}\Big).$$

Sketch of proof. Since $\big|\eta_t(\widehat{x}) - \eta_{t'}(\widehat{y})\big|^a \le \eta^a_t(\widehat{x})\,\eta^a_{t'}(\widehat{y})\big|Y^x_{s,t} - Y^y_{s,t'}\big|^a$, by Hölder inequality

$$\mathbb{E}\Big[\big|\eta_t(\widehat{x}) - \eta_{t'}(\widehat{y})\big|^a\Big] \le \mathbb{E}\Big[\big|\eta_t(\widehat{x})\big|^{4a}\Big]^{\frac14}\mathbb{E}\Big[\big|\eta_{t'}(\widehat{y})\big|^{4a}\Big]^{\frac14}\mathbb{E}\Big[\big|Y^x_{s,t} - Y^y_{s,t'}\big|^{2a}\Big]^{\frac12}.$$

If $\widehat{x}, \widehat{y} \ne 0$, to complete the proof one can use Lemma 5.4, Proposition 5.2 and the inequality $(1+|x|)^{-1}(1+|y|)^{-1}|x-y| \le |\widehat{x} - \widehat{y}|$. The other cases are easier. □

Theorem 5.9. *The map $Y_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ is onto for any $t \in [s,T]$ almost surely.*

Proof. Take $a > 2(d+3)$ in lemma 5.8. Then, by Kolmogorov Theorem, $\eta_t(\widehat{x})$ is continuous at $\widehat{x} = 0$. Therefore, $Y^x_{s,t}$ can be extended to a continuous map from $\widehat{\mathbb{R}^d}$ into itself for any $t \in [s,T]$ almost surely and the extension $\widehat{Y}^x_{s,t}(\omega)$ is continuous in (t,x) almost surely. For all ω such that $\widehat{Y}$ is continuous, the map $\widehat{Y}_{s,t}(\omega) : \widehat{\mathbb{R}}^d \to \widehat{\mathbb{R}}^d$ is homotopically equivalent to the identity map $\widehat{Y}_{s,s}(\omega)$. Proceeding by contradiction, assume that $\widehat{Y}_{s,t}(\omega)$ is not surjective. Then it takes values in $\widehat{\mathbb{R}^d}$ without one point, which is a contractible space, so that it must be homotopically equivalent to a constant. This implies that also the map $Id_{\widehat{\mathbb{R}^d}} = \widehat{Y}_{s,s}(\omega)$ is homotopically equivalent to a constant, and the space $\widehat{\mathbb{R}^d}$ would be contractible, which is absurd (because, e.g., $\pi_d(\widehat{\mathbb{R}^d}) = \mathbb{Z}$). The contradiction found shows that the function $\widehat{Y}_{s,t}(\omega)$ needs to be an onto map. Since $\widehat{Y}^\infty_{s,t}(\omega) = \infty$, the restriction of $\widehat{Y}_{s,t}(\omega)$ to $\mathbb{R}^d$ is again onto. The theorem is proved. □

5.1.4 Existence of the Flow

The following theorem resumes the results obtained so far for the flow associated with the transformed SDE (1.8).

Theorem 5.10. *Let $Y^x_{s,t}(\omega)$ be the continuous modification of the unique global solution of the transformed SDE (1.8) provided by Corollary 5.3. Then for almost every ω, $\varphi_{s,t}(x)(\omega) := Y_{s,t}(x)(\omega) : \mathbb{R}^d \to \mathbb{R}^d$ defines a stochastic flow of homeomorphisms.*

Proof. The continuous function $\varphi_{s,t}(x)(\omega)$ is one to one by Theorem 5.7 and it is onto by Theorem 5.9. Hence, the inverse map $\big(\varphi_{s,t}(\omega)\big)^{-1} = \big(Y_{s,t}(\omega)\big)^{-1}$ is well defined, one to one and onto. We claim that it is also continuous. Indeed, since the map $\widehat{Y}_{s,t}(\omega)$ is one to one and continuous from the compact space $\widehat{\mathbb{R}^d}$ into itself, it is a closed map. Hence the inverse map $\big(\widehat{Y}_{s,t}(\omega)\big)^{-1}$ is continuous, and so is its restriction to $\mathbb{R}^d$. □

5.2 The flow of the original SDE

Proof of theorem 1.2. Theorem 2.5 provides the weak existence of a solution X and Theorem 4.1 the strong uniqueness property for this solution, so that by the Yamada–Watanabe principle the solution is a strong solution. We construct the flow. Consider the associated "transformed SDE" (1.8), for which in Section 4 we have shown the existence and uniqueness of a strong solution. We know that for every $t \in [s,T]$ and $x \in \mathbb{R}^d$

$$Y^y_{s,t} = \phi_t\big(X^x_{s,t}\big), \qquad X^x_{s,t} = \phi_t^{-1}\big(Y^y_{s,t}\big), \tag{5.9}$$

where ϕ_t is the function studied in lemma 3.5 and $y := \phi_s(x)$. Theorem 5.10 provides the flow for (1.8):

$$Y^y_{s,t}(\omega) = \varphi_{s,t}(y)(\omega). \tag{5.10}$$

Define

$$\psi_{s,t}(x)(\omega) := \phi_t^{-1}\Big[\varphi_{s,t}\big(\phi_s(x)\big)(\omega)\Big]. \tag{5.11}$$

We claim that ψ is a flow for the SDE (1.1).

First, it follows directly from (5.9), (5.10) and Definition (5.11) that

$$P\Big(X^x_{s,t} = \psi_{s,t}(x)\Big) = 1. \tag{5.12}$$

Recall that, by Lemma 3.5, the functions ϕ_t and ϕ_t^{-1} are continuous in (t,x) and, for every fixed $t \in [0,T]$, they are bijective. Since φ is a flow, it follows from Definition (5.11) that ψ is almost surely continuous and bijective. Note that

$$\psi^{-1}_{s,t}(x)(\omega) := \phi_s^{-1}\Big[\varphi^{-1}_{s,t}\big(\phi_t(x)\big)(\omega)\Big] \tag{5.13}$$

is the inverse function of (5.11) and is again continuous in (t,x). This is the second point of Definition 5.1.

The composition property follows from the same property for the flow φ. Indeed, we have that almost surely

$$\begin{aligned}\psi_{s,t}(x) &= \phi_t^{-1}\Big[\varphi_{s,t}\big(\phi_s(x)\big)\Big] = \phi_t^{-1}\Big[\varphi_{u,t}\big(\varphi_{s,u}\big(\phi_s(x)\big)\big)\Big] \\ &= \phi_t^{-1}\bigg[\varphi_{u,t}\Big(\phi_u\Big[\phi_u^{-1}\big(\varphi_{s,u}\big(\phi_s(x)\big)\big)\Big]\Big)\bigg] = \psi_{u,t}\Big(\psi_{s,u}(x)\Big).\end{aligned}$$

The following lemma 5.11 shows that the function $\psi_{s,t}(x)$ is Hölder continuous. To obtain that also the inverse function $\psi_{s,t}(x)$ is Hölder continuous one just needs to revers time in (1.1) and apply all the above results to the backward SDE.

Finally, to complete the proof of the theorem we only need to address the differentiability of the solution map $x \to X_t^x$, which is provided by Corollary 5.13 below. The proof is complete. ☐

Lemma 5.11. *For every $t \in [s,T]$, there exists a modification of the function $\psi_{s,t}$ which is α-Hölder continuous for every $\alpha < 1$.*

Proof. Recall that the functions ϕ_t and ϕ_t^{-1} are Lipschitz continuous uniformly in time. From Corollary 5.3 we obtain that the function $\varphi_{s,t}$ defining the flow associated to the transformed SDE (1.8) is α-Hölder continuous for every $t \in [s,T]$ and $\alpha < 1$. Then it follows from (5.11) that

$$\begin{aligned}|\psi_{s,t}(x) - \psi_{s,t}(y)| &\le L\Big|\Big[\varphi_{s,t}\big(\phi_s(x)\big)(\omega)\Big] - \Big[\varphi_{s,t}\big(\phi_s(y)\big)(\omega)\Big]\Big| \\ &\le L\big|\phi_s(x) - \phi_s(y)\big|^\alpha \le L^{1+\alpha}|x-y|^\alpha.\end{aligned}$$

The lemma is proved. ☐

5.3 Weak Differentiability

To ease notation, in the following we will take again $s = 0$. With computations similar to the ones presented above, a weak differentiability result (in the sense of Theorem 1.2) can be proven for the map $x \mapsto Y_t^x$. This result can be easily transferred to the process $x \mapsto X_t^x$ by means of the correspondence provided by (1.7): $X_t = \phi_t^{-1}(Y_t)$. With the notation introduced above, we have

Theorem 5.12. *Given $t \in [0,T]$, the map $x \mapsto Y_t^x$ from $\mathbb{R}^d$ to $L^2\left(\Omega \times [0,T]; \mathbb{R}^d\right)$ is differentiable, in the following weak sense: for every $i = 1, ..., d$ and every $x \in \mathbb{R}^d$, the limit*

$$\lim_{h\to 0} \frac{Y_t^{x+he_i} - Y_t^x}{h}$$

exists as a strong limit in $L^2\left(\Omega \times [0,T]; \mathbb{R}^d\right)$. Denote by $\xi_t^{x,i}$ the limit process. The process $\left(\xi_t^{x,i}\right)_{t\in[0,T]}$ has a continuous adapted modification which is the unique continuous adapted solution of the variational equation

$$\mathrm{d}\xi_t^{x,i} = \nabla\tilde{b}\left(t, Y_t^x\right)\xi_t^{x,i}\,\mathrm{d}t + \sum_{k=1}^d \nabla\tilde{\sigma}_k\left(t, Y_t^x\right)\xi_t^{x,i}\,\mathrm{d}W_t^k, \quad \xi_0^{x,i} = e_i. \tag{5.14}$$

Corollary 5.13. *The map $x \mapsto X_t^x$ is weakly differentiable in the sense of Theorem 1.2.*

Proof. We have

$$\frac{X_t^{x+he_i} - X_t^x}{h} = \int_0^1 \nabla\phi_t\Big(rY_t^{\phi_0(x+he_1)} + (1-r)Y_t^{\phi_0(x)}\Big)\mathrm{d}r \cdot \frac{Y_t^{\phi_0(x+he_1)} - Y_t^{\phi_0(x)}}{h}.$$

The family of r.v. $\int_0^1 \nabla\phi_t\big(rY_t^{\phi_0(x+he_1)} + (1-r)Y_t^{\phi_0(x)}\big)\mathrm{d}r$ is equibounded and a.s. convergent to $\nabla\phi_t\big(Y_t^{\phi_0(x)}\big)$ for $h \to 0$. Since ϕ_0 is a bijection of $\mathbb{R}^d$ (see Lemma 3.5), we can rewrite $\phi_0(x + he_1) = \phi_0(x) + k(h)e_i'$, where $\{e_i'\}_{i=1,\dots,d}$ is a new basis for $\mathbb{R}^d$ and $k(h) \to 0$ for $h \to 0$. Therefore, by Theorem 5.12, $\big(Y_t^{\phi_0(x+he_1)} - Y_t^{\phi_0(x)}\big)/h$ strongly converges in $L^2\left(\Omega \times [0,T]; \mathbb{R}^d\right)$. ☐

Remark 5.14. The stochastic integral appearing in (5.14) is meaningful because $\xi^{x,i}_{\cdot} \in L^2\left(\Omega \times [0,T];\mathbb{R}^d\right)$ and $\nabla\tilde{\sigma}_k \in L^q_p(T)$ with $q > 2$.

Remark 5.15. A natural question is whether $\xi^{x,i}_t$ has a continuous version in x. This property would be related to the differentiability of the flow $\varphi_t(\omega)$ associated with equation (1.8). We think that this problem is very difficult.

Note that the variational equation satisfied by $x \mapsto X^x_t$ would contain the term ∇b, which is only distributional. In dimension 1, one could give a meaning to this equation thanks to the properties of the local time, see for instance [18]. However, in higher dimension the problem is still open.

Proof of theorem 5.12. **Step 1** (Preparation) It is sufficient to prove the theorem for $i = 1$. We omit to write i, and set $e = e_1$. Moreover, since $x \in \mathbb{R}^d$ is given, we omit it, together with $\phi_0^{-1}(x)$, where possible.

Introduce, for every $h \neq 0$, the stochastic processes

$$\xi^h_t := \frac{Y^{x+he}_t - Y^x_t}{h}\ , \qquad \theta^h_t = \frac{X^{\phi_0^{-1}(x+he)}_t - X^{\phi_0^{-1}(x)}_t}{h} = \left[\int_0^1 \nabla\phi_t^{-1}(Y^{h,(r)}_t)dr\right]\xi^h_t\ ,$$

where $Y^{h,(r)}_t := r\,Y^x_t + (1-r)\,Y^{x+he}_t$. Let us set $X^{h,(r)}_t = rX^{\phi_0^{-1}(x+h\varepsilon)}_t + (1-r)\,X^{\phi_0^{-1}(x)}_t$. For every $t \geq 0$, $x \in \mathbb{R}^d$, $h \neq 0$, $k = 1, ..., d$, the expression $\int_0^1 D^2U_k\left(t, X^{h,(r)}_t\right)dr$ is meaningful and defines a random $d \times d$ matrix, see Remark 5.17. Therefore, by the approximation argument used in the proof of Lemma 4.5 we get that ξ^h_t is a solution of

$$\mathrm{d}\xi^h_t = \left[\int_0^1 \lambda\nabla U\left(t, X^{h,(r)}_t\right)\mathrm{d}r\right]\theta^h_t\mathrm{d}t + \sum_{k=1}^d \left[\int_0^1 D^2U_k\left(t, X^{h,(r)}_t\right)\mathrm{d}r\right]\theta^h_t\mathrm{d}W^k_t. \tag{5.15}$$

Step 2 (Uniqueness for equation (5.14)) At intuitive level, it is clear that we may control well the finite differences of equation (5.15) if we have a good strategy to control the solution of equation (5.14). Therefore, it looks natural to start by proving uniqueness for equation (5.14) (we do it in the class of adapted continuous processes).

We claim that the difference η_t of two continuous adapted solutions is identically zero (note that the equation satisfied by η_t is linear). Introduce

$$B_t := \int_0^t \sum_{k=1}^d \left|D^2U_k\left(s, X_s\right)\right|^2\left|\nabla\phi_s^{-1}(Y_s)\right|^2\mathrm{d}s + 2\int_0^t \left|\lambda\nabla U\left(s, X_s\right)\right|\left|\nabla\phi_s^{-1}(Y_s)\right|\mathrm{d}s.$$

An analogous of Lemma 4.5 can be proven for this process, which results to be well defined, increasing and with $\mathbb{E}\left[e^{\lambda B_t}\right] < \infty$ for all $\lambda > 0$. The process B_t is chosen so that when we apply Itô formula all terms but one simplifies. We have

$$\frac{1}{2}\mathrm{d}\left(e^{-B_t}\left|\eta_t\right|^2\right) \leq \sum_{k=1}^d e^{-B_t}\eta_t^T\left[D^2U_k\left(t, X_t\right)\nabla\phi_t^{-1}(Y_t)\right]\eta_t\,\mathrm{d}W^k_t.$$

Let τ_n be the first time η_t becomes greater than n, or $\tau_n = T$ if this never happens. Then $\tau_n > 0$, $\lim_{n\to\infty}\tau_n = T$ and

$$e^{-B_{t\wedge\tau_n}}\left|\eta_{t\wedge\tau_n}\right|^2 \leq 2\sum_{k=1}^d \int_0^t e^{-B_s}\eta_s^T\left[D^2U_k\left(s, X_s\right)\nabla\phi_s^{-1}(Y_s)\right]\eta_s 1_{\{s\leq\tau_n\}}\,\mathrm{d}W^k_s.$$

As a by product of the above-mentioned result on B_t, we get that $D^2U_k\left(\cdot, X_\cdot\right)$ is square integrable, so that the stochastic integrals are martingales. Thus $\mathbb{E}\left[e^{-B_{t\wedge\tau_n}}|\eta_{t\wedge\tau_n}|^2\right] \leq 0$. Since $\lim_{n\to\infty}\tau_n = T$, by Fatou lemma we get $\mathbb{E}\left[e^{-B_t}|\eta_t|^2\right] \leq 0$ and (since $e^{-B_t} > 0$ a.s.) $|\eta_t|^2 = 0$ a.s., which implies $\eta = 0$.

Step 3 (Estimates for ξ_t^h) Observe that $\left|\theta_t^h\right| \leq C_\phi \left|\xi_t^h\right|$. Introduce

$$C_t^h = 18C_\phi^2 \int_0^t \sum_{k=1}^d \int_0^1 \left|D^2 U_k\left(s, X_s^{h,(r)}\right)\right|^2 \mathrm{d}r\mathrm{d}s + 4C_\phi \int_0^t \int_0^1 \left|\lambda \nabla U\big(t, X_s^{h,(r)}\big)\right| \mathrm{d}r\mathrm{d}s.$$

Again, this process has the same properties of the process B_t, and in particular

$$\sup_{|h|\leq 1} \mathbb{E}\Big[e^{\lambda C_t^h}\Big] \leq K_T < \infty \tag{5.16}$$

for all $\lambda > 0$ and some constant K_T independent of t (see Remark 5.17). Apply Itô formula with $f(x) = |x|^4$ to the solution ξ_t^h of Equation (5.15). For the corrector we have the bound

$$\frac{1}{2} \sum_{i,j=1}^d \partial_i \partial_j f\big(\xi_t^h\big)\ \mathrm{d}\left[\xi_\cdot^{h,i}, \xi_\cdot^{h,j}\right]_t \leq 18C_\phi^2 \left|\xi_t^h\right|^4 \sum_{i,k=1}^d \int_0^1 \left|\partial_i DU_k\left(t, X_t^{h,(r)}\right)\right|^2 \mathrm{d}r\mathrm{d}t,$$

so that we get

$$\mathrm{d}\left(e^{-C_t^h}\left|\xi_t^h\right|^4\right) \leq 4\sum_{k=1}^d e^{-C_t^h}\left|\xi_t^h\right|^2 \left(\xi_t^h\right)^T \left[\int_0^1 D^2 U_k\left(t, X_t^{h,(r)}\right) dr\right] \theta_t^h\, \mathrm{d}W_t^k.$$

Using again $\left|\theta_t^h\right| \leq C_\phi \left|\xi_t^h\right|$ and arguing as above with stopping times associated to ξ_t^h, we finally get $\mathbb{E}\big[e^{-C_t^h}|\xi_t^h|^4\big] \leq 1$ (because $|\xi_0^h|^4 = 1$). But $\mathbb{E}\big[|\xi_t^h|^2\big] \leq \mathbb{E}\big[e^{C_t^h}\big]^{1/2}\,\mathbb{E}\big[e^{-C_t^h}|\xi_t^h|^4\big]^{1/2}$, so that by (5.16), we have

$$\sup_{|h|\leq 1} \mathbb{E}\left[\left|\xi_t^h\right|^2\right] \leq K_T < \infty.$$

Note that all the above computations can be performed using $f = |x|^{2p}$ for any $p \in [2, \infty)$, to get

$$\sup_{|h|\leq 1} \mathbb{E}\left[\left|\xi_t^h\right|^p\right] \leq K_T < \infty.$$

Step 4 (Weak passage to the limit) From the bound above, there exists a sequence $\xi_t^{h_n}$ which converges weakly to a process ξ_t in $L^p\left(\Omega \times [0,T]; \mathbb{R}^d\right)$, which is progressively measurable because the space of non anticipative processes is a closed subspace of $L^p\left(\Omega \times [0,T]; \mathbb{R}^d\right)$. The r.v. $Y_t^{h_n,(r)}$ and $X_t^{h_n,(r)}$ converge pointwise (recall also the flow property) to Y_t and X_t respectively. The coefficient $\lambda\nabla U$ is bounded continuous, and we have Lemma 5.18 to deal with the diffusion term. Therefore, for every $p \geq 2$, $\int_0^1 \nabla U\big(t, X_t^{h_n,(r)}\big)\mathrm{d}r$ and $\int_0^1 D^2 U_k\big(t, X_t^{h_n,(r)}\big)\mathrm{d}r$ converge strongly to $\nabla U\big(t, X_t\big)$ and $D^2 U_k\big(t, X_t\big)$ respectively, in $L^p\big(\Omega \times [0,T]; \mathbb{R}^d\big)$. It follows that

$$\left[\int_0^1 \lambda \nabla U\left(t, X_t^{h_n,(r)}\right) \mathrm{d}r\right]\left[\int_0^1 \nabla\phi_t^{-1}(Y_t^{h_n,(r)})\mathrm{d}r\right]\xi_t^{h_n} \rightharpoonup \lambda \nabla U\left(t, X_t\right) \nabla\phi_t^{-1}(Y_t)\xi_t$$
$$\left[\int_0^1 D^2 U_k\left(t, X_t^{h_n,(r)}\right) \mathrm{d}r\right]\left[\int_0^1 \nabla\phi_t^{-1}(Y_t^{h_n,(r)})\mathrm{d}r\right]\xi_t^{h_n} \rightharpoonup D^2 U_k\left(t, X_t\right) \nabla\phi_t^{-1}(Y_t)\xi_t$$

(weakly) in $L^{p'}\left(\Omega \times [0,T]; \mathbb{R}^d\right)$, for every $p' \geq 1$. By a classical argument [19, Chap. 3] (repeated for instance in the proof of Theorem 15, [3]), we may pass to the limit in Equation (5.15) and prove that ξ_t solves equation (5.14). A fortiori, ξ_t has a continuous modification in t. In step 2 we have proved that this equation has a unique solution; then the full family $\xi_{t,x}^h$ weakly converges to ξ_t^x.

Step 5 (Strong passage to the limit) Since the approximate equation (5.15) is linear and has strongly convergent coefficients, it is natural to expect the convergence $\xi_t^h \to \xi_t$ to be strong. Indeed, set $\eta_t^h := \xi_t^h - \xi_t$,

$b(s) := \int_0^1 \lambda\nabla U\big(s, X_s^x\big)\mathrm{d}r \int_0^1 \nabla\phi_s^{-1}(Y_s)\mathrm{d}r$, $b^h(s) := \int_0^1 \lambda\nabla U\big(s, X_s^{h,(r)}\big)\mathrm{d}r \int_0^1 \nabla\phi_t^{-1}(Y_t^{h_n,(r)})\mathrm{d}r$ and similarly for the diffusion coefficients, which we call $\sigma(s)$ and $\sigma^h(s)$. Write

$$\eta_t^h = \int_0^t b(s)\eta_s^h \,\mathrm{d}s + \sum_{k=1}^d \int_0^t \sigma_k(s)\eta_s^h \,\mathrm{d}W_s^k + R_t^h;$$

$$R_t^h := \int_0^t \big(b^h(s) - b(s)\big)\xi_s^h \,\mathrm{d}s + \sum_{k=1}^d \int_0^t \big(\sigma_k^h(s) - \sigma_k(s)\big)\xi_s^h \,\mathrm{d}W_s^k.$$

Proceeding again as in Step 3 above, with

$$C_t = 18\int_0^t \left|D^2 U\left(s, X_s\right)\right|^2 \left|\phi_s^{-1}\left(Y_s\right)\right|^2 \mathrm{d}s + 4\lambda \int_0^t \left|\nabla U\left(s, X_s\right)\right| \left|\phi_s^{-1}\left(Y_s\right)\right| \mathrm{d}s$$

we get

$$\frac{1}{2}\,\mathrm{d}\left(e^{-C_s}\left|\eta_s\right|^4\right) \le 4\sum_{k=1}^d e^{-C_s}|\eta_s|^2\eta_s^T\sigma_k(s)\eta_s \,\mathrm{d}W_s^k + 4e^{-C_s}|\eta_s|^2\eta_s \,\mathrm{d}R_s^h + e^{-C_s}|\eta_s|^2 \,\mathrm{d}\big[R^h, \eta\big]_s. \tag{5.17}$$

The second and third terms can be bounded by

$$4e^{-C_s}|\eta_s|^2\eta_s\big(b^h(s) - b(s)\big)\xi_s^h \,\mathrm{d}s + 4\sum_{k=1}^d e^{-C_s}|\eta_s|^2\eta_s\big(\sigma_k^h(s) - \sigma_k(s)\big)\xi_s^h \,\mathrm{d}W_s^k, \tag{5.18}$$

$$18e^{-C_s}|\eta_s|^2 \sum_{k=1}^d \Big(\sigma_k(s)\eta_s^h + \big(\sigma_k^h(s) - \sigma_k(s)\big)\xi_s^h\Big)\big(\sigma_k^h(s) - \sigma_k(s)\big)\xi_s^h \,\mathrm{d}s.$$

We can deal with the first term on the right hand side of (5.17) and the second term of (5.18) with a stopping time for η and ξ, just as in Step 2 or 3. The remaining two terms have weakly convergent elements (and therefore bounded) in $L^p\left(\Omega \times [0,T]; \mathbb{R}^d\right)$ for every $p \ge 2$ and one which is strongly convergent to zero. They are, therefore, strongly convergent to zero in $L^{p'}\left(\Omega \times [0,T]; \mathbb{R}^d\right)$ for every $p' \ge 1$. The strong convergence of $\xi_{t,x}^h$ to $\xi_{t,x}$ in $L^2\left(\Omega \times [0,T]; \mathbb{R}^d\right)$ follows. □

Remark 5.16. From the above proof, one also obtains that $\xi_{t,x}^h$ strongly converges to $\xi_{t,x}$ in $L^p\big(\Omega\times[0,T];\mathbb{R}^d\big)$ for every $p \ge 2$.

5.3.1 Technical lemmas

The following remark follows from the proof of Lemma 4.5.

Remark 5.17. Let X_t^x be the solution of equation (1.1). Let f be a vector field of class $L_p^q(T)$ for p, q as in (1.3). Then, for every $r \in [0,1]$ and $x, y \in \mathbb{R}^d$, the composition

$$f\left(t, X_t^{x,y,r}\right), \quad X_t^{x,y,r} := rX_t^x + (1-r)Y_t^y$$

is uniquely defined (it does not depend on the element in the equivalence class of f) and

$$\sup_{x,y\in\mathbb{R}^d} \sup_{r\in[0,1]} \mathbb{E}\left[\exp\left(\lambda\int_0^T f^2\left(t, X_t^{x,y,r}\right)\mathrm{d}t\right)\right] < \infty \tag{5.19}$$

for all $\lambda > 0$. In particular, $f(t, X_t^{x,y,r}) \in L^a(\Omega \times [0,T])$ for every r and $a \ge 2$.

Lemma 5.18. *Under the assumptions of Remark 5.17, we have that*

$$\lim_{y'\to x'} f\left(t, X_t^{x',y',r}\right) = f\left(t, X_t^{x'}\right) \tag{5.20}$$

in $L^p\big(\Omega \times [0,T]\big)$ *for* $p \geq 2$.

Sketch of Proof. The proof is similar to the proof of [20, Lemma 6.3]. Write (all objects will be defined below)

$$\begin{aligned}\mathbb{E}\Big[\int_0^T |f(t,X_t^{x,y,r}) - f(t,X_t^x)|^p \mathrm{d}t\Big] \leq C\mathbb{E}\Big[\int_0^T &|f(t,X_t^{x,y,r}) - f^k(t,X_t^{x,y,r})|^p \\ &+ |f^k(t,X_t^{x,y,r}) - f^k(t,X_t^x)|^p + |f^k(t,X_t^x) - f(t,X_t^x)|^p \mathrm{d}t\, \mathbf{1}_{A_R}\Big] \\ &+ C\mathbb{E}\Big[\int_0^T |f(t,X_t^{x,y,r}) - f(t,X_t^x)|^p \mathrm{d}t\, \mathbf{1}_{A_R^c}\Big]\end{aligned}$$

We claim that for any $\varepsilon > 0$ there exist R, f^k, δ s.t. for $|y - x| \leq \delta$ the right hand side above is smaller than ε.

Set $\rho := |x| + |y|$. For $R \in \mathbb{R}$, set $A_R := \{X_t^x \in B_R\, \forall t \in [0,T]\} \cap \{X_t^{x,y,r} \in B_R\, \forall t \in [0,T], r \in [0,1]\} \subset \Omega$. We have $\mathbb{E}[|Y_t^y|^a] \leq C(1+|y|^a)$, $|X_t| = |\phi_t^{-1}(Y_t)| \leq C(1+|Y_t|)$, $|Y_0| \leq C(1+|X_0|)$. Therefore, for all $x \in B_\rho$, $E[|X_t|^a] \leq C(1+\rho^a)$ and $P(|X_t^x| > R/2) \to 0$ for $R \to \infty$. From the Hölder continuity in time of the trajectories and $|X^{x,y,r}| \leq (|X^x| \vee |X^y|)$ we get $P(A_R) \to 1$ for $R \to \infty$. The last term above is small for large R due to remark 5.17 and the absolute continuity of the integral. Take R so that it is smaller than $\varepsilon/2$.

Let $f^k \in \mathcal{C}_b^0$ be a sequence of continuous and uniformly bounded functions converging to f on the compact set $[0,T] \times B_R$. When $y \to x$, also the three terms in the first integral are small for large k: two of them due to the uniform convergence $f^k \to f$, and the central one by Vitali convergence theorem because the convergence in probability of $X_t^{x,y,r} \to X_t^x$ passes through continuous functions. We can therefore find f^k, δ so that the first integral is smaller than $\varepsilon/2$. The claim is proved, and the lemma follows. □

References

[1] Krylov, N.V., Röckner, M. 2005. Strong solutions to stochastic equations with singular time dependent drift. *Probab. Theory Relat. Fields* 131: 154-196.

[2] Zhang, X. 2005. Strong solutions of SDES with singular drift and Sobolev diffusion coefficients. *Stoch. Proc. Appl.* 115: 1805-1818.

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

[4] Zvonkin, A.K. 1974. A transformation of the phase space of a diffusion process that removes the drift. *Matematicheskiĭ Sbornik. Novaya Seriya* 93 (135): 129-149 .

[5] Fedrizzi, E., Flandoli, F. 2011. Pathwise uniqueness and continuous dependence for SDEs with nonregular drift. *Stochastics An Internationa Journal of Probability and Stochastic Processes* 83(3): 241-257.

[6] Albeverio, S., Kondratiev, Y.G., Röckner, M. 2003. Strong Feller properties for distorted Brownian motion and applications to finite particle systems with singular interactions. In, *Finite and Infinite Dimensional Analysis in Honor of Leonard Gross* (New Orleans, LA, 2001): 1535. Eds. Contemporary Mathematics, 317. Amer. Math. Soc., Providence, RI.

[7] Yamada, T., Watanabe, S. 1971. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto* 11: 155-167.

[8] Portenko, N.I. 1982. *Generalized Diffusion Processes.* Nauka, Moscow (in Russian); English translation: 1990, Amer. Math. Soc., Providence, Rhode Island.

[9] Veretennikov, A.Yu. 1980. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Matematicheskiĭ Sbornik. Novaya Seriya* 111(153): 434-452.

[10] Khas'minskii, R.Z. 1959. On positive solutions of the equation $Au + Vu = 0$. *Theoret. Probab. Appl.* 4: 309-318.

[11] Sznitman., A.-S., 1998. *Brownian motion, obstacles, and random media.* Springer-Verlag, Berlin and New York.

[12] Liptser, R.S., Shiryaev, A.N. 1974. *Statistics of random processes.* Nauka, Moscow (in Russian); English translation: 1977, Springer-Verlag, Berlin and New York.

[13] Fedrizzi, E. 2009. Uniqueness and flow theorems for solutions of SDEs with low regularity of the drift. Master thesis, Università di Pisa. Aviable on http://etd.adm.unipi.it.

[14] Krylov, N.V. 2001. The heat equation in $L_q((0,T), L_p)-$spaces with weights. *Siam J. Math. Anal.* 32(5): 1117-1141.

[15] Protter, P.E. 2004. *Stochastic integration and differential equations.* Springer Berlin, 2$^{\text{nd}}$ edition.

[16] Gyöngy, I., Krylov, N.V. 1996. Existence of strong solutions for Itô stochastic equations via approximations. *Probab. Theory Relat. Fields* 105: 143-158.

[17] Kunita, H. 1984. *Stochastic differential equations and stochastic flows of diffeomorphisms.* XII École d'été de probabilités de Saint-Flour - 1982. Lecture Notes in Math. 1097: 143-303; Springer: Berlin.

[18] Attanasio, S. 2010. Stochastic flows of diffeomorphisms for one-dimensional SDE with discontinuous drift. *Electronic Communications in Probability* 15 : 213-226.

[19] Pardoux, E. 1975. Equations aux dérivées partielles stochastiques non linéaires monotones. Etude de solutions fortes de type Itô. PhD Thesis, Universitè Paris Sud.

[20] I.Gyöngy 1998. Existence and uniqueness results for semilinear stochastic partial differential equations. *Stochastic Processes and their Applications* 73: 271-299.