Article:

Flandoli, F, Issoglio, E orcid.org/0000-0003-3035-2712 and Russo, F (2017) Multidimensional stochastic differential equations with distributional drift. Transactions of the American Mathematical Society, 369. pp. 1665-1688. ISSN 0002-9947

https://doi.org/10.1090/tran/6729

MULTIDIMENSIONAL STOCHASTIC DIFFERENTIAL EQUATIONS WITH DISTRIBUTIONAL DRIFT

FRANCO FLANDOLI[1], ELENA ISSOGLIO[2], AND FRANCESCO RUSSO[3]

Abstract. This paper investigates a time-dependent multidimensional stochastic differential equation with drift being a distribution in a suitable class of Sobolev spaces with negative derivation order. This is done through a careful analysis of the corresponding Kolmogorov equation whose coefficient is a distribution.

Key words and phrases: Stochastic differential equations; distributional drift; Kolmogorov equation.

AMS-classification: 60H10; 35K10; 60H30; 35B65.

1. Introduction

Let us consider a distribution valued function $b : [0,T] \to \mathcal{S}'(\mathbb{R}^d)$, where $\mathcal{S}'(\mathbb{R}^d)$ is the space of tempered distributions. An ordinary differential equation of the type

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t, \quad X_0 = x_0, \tag{1}$$

$x_0 \in \mathbb{R}^d$, $t \in [0,T]$, does not make sense, except if we consider it in a very general context of *generalized functions*. Even if b is function valued, without a minimum regularity in space, problem (1), is generally not well-posed. A motivation for studying (1) is for instance to consider b as a quenched realization of some (not necessarily Gaussian) random field. In the annealed form, (1) is a singular *passive tracer* type equation.

Let us consider now equation (1) with a noise perturbation, which is expected to have a regularizing effect, i.e.,

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t, \quad X_0 = x_0, \tag{2}$$

for $t \in [0,T]$, where W is a standard d-dimensional Brownian motion. Formally speaking, the Kolmogorov equation associated with the stochastic differential equation (2) is

$$\begin{cases} \partial_t u = b \cdot \nabla u + \frac{1}{2}\Delta u & \text{on } [0,T] \times \mathbb{R}^d, \\ u(T, \cdot) = f & \text{on } \mathbb{R}^d, \end{cases} \tag{3}$$

[1]Dipartimento Matematica, Largo Bruno Pontecorvo 5, C.A.P. 56127, Pisa, Italia flandoli@dma.unipi.it

[2]Department of Mathematics, University of Leeds, Leeds, LS2 9JT, UK E.Issoglio@leeds.ac.uk

[3]Unité de Mathématiques appliquées, ENSTA ParisTech, Université Paris-Saclay, 828, boulevard des Maréchaux, F-91120 Palaiseau, France francesco.russo@ensta-paristech.fr

Date: April 21, 2015.

for suitable final conditions f. Equation (3) was studied in the one-dimensional setting for instance by [23] for any time independent b which is the derivative in the distributional sense of a continuous function and in the multidimensional setting by [13], for a class of b of gradient type belonging to a given Sobolev space with negative derivation order. The equation in [13] involves the *pointwise product* of distributions which in the literature is defined by means of paraproducts.

The point of view of the present paper is to keep the same interpretation of the product as in [13] and to exploit the solution of a PDE of the same nature as (3) in order to give sense and study solutions of (2). A solution X of (2) is often identified as a *diffusion with distributional drift*.
Of course the sense of equation (2) has to be made precise. The type of solution we consider will be called *virtual solution*, see Definition 25. That solution will fulfill in particular the property to be the limit in law, when $n \to \infty$, of solutions to classical stochastic differential equations

$$\mathrm{d}X_t^n = \mathrm{d}W_t + b_n(t, X_t^n)\mathrm{d}t, \quad t \in [0, T], \tag{4}$$

where $b_n = b \star \phi_n$ and (ϕ_n) is a sequence of mollifiers converging to the Dirac measure.

Diffusions in the generalized sense were studied by several authors beginning with, at least in our knowledge [20]; later on, many authors considered special cases of stochastic differential equations with generalized coefficients, it is difficult to quote them all: in particular, we refer to the case when b is a measure, [4, 7, 18, 22]. [4] has even considered the case when b is a not necessarily locally finite signed measure and the process is a possibly exploding semimartingale. In all these cases solutions were semimartingales. In fact, [8] considered special cases of non-semimartingales solving stochastic differential equations with generalized drift; those cases include examples coming from Bessel processes.

The case of time independent SDEs in dimension one of the type

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}W_t + b(X_t)\mathrm{d}t, \quad t \in [0, T], \tag{5}$$

where σ is a strictly positive continuous function and b is the derivative of a real continuous function was solved and analyzed carefully in [10] and [11], which treated well-posedness of the martingale problem, Itô's formula under weak conditions, semimartingale characterization and Lyons-Zheng decomposition. The only supplementary assumption was the existence of the function $\Sigma(x) = 2\int_0^x \frac{b}{\sigma^2}\mathrm{d}y$ as limit of appropriate regularizations. Also in [1] the authors were interested in (2) and they provided a well-stated framework when σ and b are γ-Hölder continuous, $\gamma > \frac{1}{2}$. In [23] the authors have also shown that in some cases strong solutions (namely solutions adapted to the completed Brownian filtration) exist and pathwise uniqueness holds.
As far as the multidimensional case is concerned, it seems that the first paper was [2]. Here the authors have focused on (2) in the case of a time independent drift b which is a measure of Kato class.

Coming back to the one-dimensional case, the main idea of [11] was the so called Zvonkin transform which allows to transform the candidate solution process X into a solution of a stochastic differential equation with continuous non-degenerate coefficients without drift. Recently [16] has considered other types of transforms to study similar equations. Indeed the transformation introduced by Zvonkin in [27], when the drift is a function, is also stated in the multidimensional case. In a series of papers the first named author and coauthors (see for instance [9]), have efficiently made use of a (multidimensional) Zvonkin type transform for the study of an SDE with measurable not necessarily bounded drift, which however is still a function. Zvonkin transform consisted there to transform a solution X of (2) (which makes sense being a classical SDE) through a solution $\varphi : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ of a PDE which is close to the associated Kolmogorov equation (3) with some suitable final condition. The resulting process Y with $Y_t = \varphi(t, X_t)$ for $t \in [0,T]$ is a solution of an SDE for which one can show strong existence and pathwise uniqueness.

Here we have imported that method for the study of our time-dependent multidimensional SDE with distributional drift.

The paper is organized as follows. In Section 2 we adapt the techniques of [13], based on pointwise products for investigating existence and uniqueness for a well chosen PDE of the same type as (3), see (6). In Section 3 we introduce the notion of *virtual solution* of (2). The construction will be based on the transformation $X_t = \psi(t, Y_t)$ for $t \in [0,T]$, where Y is the solution of (34) and $\varphi(t,x) = x + u(t,x), (t,x) \in [0,T] \times \mathbb{R}^d$, with u being the solution of (6). Section 3.3 shows that the virtual solution is indeed the limit of classical solutions of regularized stochastic differential equations.

2. The Kolmogorov PDE

2.1. **Setting and preliminaries.** Let b be a vector field on $[0,T] \times \mathbb{R}^d, d \geq 1$, which is a distribution in space and weakly bounded in time, that is $b \in L^\infty([0,T]; \mathcal{S}'(\mathbb{R}^d; \mathbb{R}^d))$. Let $\lambda > 0$. We consider the parabolic PDE in $[0,T] \times \mathbb{R}^d$

$$\begin{cases} \partial_t u + L^b u - (\lambda+1)u = -b & \text{on } [0,T] \times \mathbb{R}^d, \\ u(T) = 0 & \text{on } \mathbb{R}^d, \end{cases} \tag{6}$$

where $L^b u = \frac{1}{2}\Delta u + b \cdot \nabla u$ has to be interpreted componentwise, that is $(L^b u)_i = \frac{1}{2}\Delta u_i + b \cdot \nabla u_i$ for $i = 1, \ldots, d$. A continuous function $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ will also be considered without any comment as $u : [0,T] \to C(\mathbb{R}^d; \mathbb{R}^d)$. In particular we will write $u(t,x) = u(t)(x)$ for all $(t,x) \in [0,T] \times \mathbb{R}^d$.

Remark 1. *All the results we are going to prove remain valid for the equation*

$$\begin{cases} \partial_t u + L^{b_1} u - (\lambda+1)\, u = -b_2 & \textit{on } [0,T] \times \mathbb{R}^d, \\ u\,(T) = 0 & \textit{on } \mathbb{R}^d, \end{cases}$$

where b_1, b_2 both satisfy the same assumptions as b. We restrict the discussion to the case $b_1 = b_2 = b$ to avoid notational confusion in the subsequent sections.

Clearly we have to specify the meaning of the *product* $b \cdot \nabla u_i$ as b is a distribution. In particular, we are going to make use in an essential way the notion of paraproduct, see [21]. We recall below a few elements of this theory; in particular, when we say that the pointwise product exists in $\mathcal{S}'$ we mean that the limit (7) exists in $\mathcal{S}'$. For shortness we denote by $\mathcal{S}'$ and $\mathcal{S}$ the spaces $\mathcal{S}'(\mathbb{R}^d;\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d;\mathbb{R}^d)$ respectively. Similarly for the L^p-spaces, $1 \le p \le \infty$. We denote by $\langle\cdot,\cdot\rangle$ the dual pairing between an element of $\mathcal{S}'$ and an element of $\mathcal{S}$.

We now recall a definition of a pointwise product between a function and a distribution (see e. g. [21]) and some useful properties.
Suppose we are given $f \in \mathcal{S}'(\mathbb{R}^d)$. Choose a function $\psi \in \mathcal{S}(\mathbb{R}^d)$ such that $0 \le \psi(x) \le 1$ for every $x \in \mathbb{R}^d$, $\psi(x) = 1$ if $|x| \le 1$ and $\psi(x) = 0$ if $|x| \ge \frac{3}{2}$. Then consider the following approximation $S^j f$ of f for each $j \in \mathbb{N}$

$$S^j f(x) := \left(\psi\left(\frac{\xi}{2^j}\right)\hat{f}\right)^{\vee}(x),$$

that is in fact the convolution of f against the smoothing rescaled function ψ_j associated with ψ. This approximation is used to define the product fg of two distributions $f, g \in \mathcal{S}'$ as follows:

$$fg := \lim_{j\to\infty} S^j f S^j g, \tag{7}$$

if the limit exists in $\mathcal{S}'(\mathbb{R}^d)$. The convergence in the case we are interested in is part of the assertion below (see [12] appendix C.4, [21] Theorem 4.4.3/1).

Definition 2. *Let $b, u : [0,T] \to \mathcal{S}'$ be such that*

(i) *the pointwise product $b(t)\cdot\nabla u(t)$ exists in $\mathcal{S}'$ for a.e. $t \in [0,T]$,*
(ii) *there are $r \in \mathbb{R}$, $q \ge 1$ such that $b, u, b\cdot\nabla u \in L^1\left([0,T]; H^r_q\right)$.*

We say that u is a mild solution of equation (6) in $\mathcal{S}'$ if, for every $\psi \in \mathcal{S}$ and $t \in [0,T]$, we have

$$\begin{aligned}\langle u(t), \psi\rangle = &\int_t^T \langle b(r)\cdot\nabla u(r), P(r-t)\psi\rangle \,\mathrm{d}r \\ &+ \int_t^T \langle b(r) - \lambda u(r), P(r-t)\psi\rangle \,\mathrm{d}r.\end{aligned} \tag{8}$$

Here $(P(t))_{t\ge0}$ denotes the heat semigroup on $\mathcal{S}$ generated by $\frac{1}{2}\Delta - I$, defined for each $\psi \in \mathcal{S}$ as

$$(P(t)\psi)(x) = \int_{\mathbb{R}^d} p_t(x-y)\psi(y)\,\mathrm{d}y,$$

where $p_t(x)$ is the heat kernel $p_t(x) = e^{-t}\frac{1}{(2t\pi)^{d/2}}\exp\left(-\frac{|x|_d^2}{2t}\right)$ and $|\cdot|_d$ is the usual Euclidean norm in $\mathbb{R}^d$. The semigroup $(P(t))_{t\ge0}$ extends to $\mathcal{S}'$, where it is defined as

$$(P_{\mathcal{S}'}(t)h)(\psi) = \langle h, \int_{\mathbb{R}^d} p_t(\cdot - y)\psi(y)\mathrm{d}y\rangle,$$

for every $h \in \mathcal{S}'$, $\psi \in \mathcal{S}$.

The fractional Sobolev spaces H^r_q are the so called Bessel potential spaces and will be defined in the sequel.

Remark 3. *If $b, u, b\cdot\nabla u$ a priori belong to spaces $L^1\left([0,T]; H^{r_i}_{q_i}\right)$ for different $r_i \in \mathbb{R}$, $q_i \ge 1$, $i = 1,2,3$, then (see e.g. (21)) there exist common $r \in \mathbb{R}$, $q \ge 1$ such that $b, u, b\cdot\nabla u \in L^1\left([0,T]; H^r_q\right)$.*

The semigroup $(P_{\mathcal{S}'}(t))_{t\geq 0}$ maps any $L^p\left(\mathbb{R}^d\right)$ into itself, for any given $p \in (1,\infty)$; the restriction $(P_p(t))_{t\geq 0}$ to $L^p\left(\mathbb{R}^d\right)$ is a bounded analytic semigroup, with generator $-A_p$, where $A_p = I - \frac{1}{2}\Delta$, see [6, Thm. 1.4.1, 1.4.2]. The fractional powers of A_p of order $s \in \mathbb{R}$ are then well-defined, see [19]. The fractional Sobolev spaces $H_p^s(\mathbb{R}^d)$ of order $s \in \mathbb{R}$ are then $H_p^s(\mathbb{R}^d) := A_p^{-s/2}(L^p(\mathbb{R}^d))$ for all $s \in \mathbb{R}$ and they are Banach spaces when endowed with the norm $\|\cdot\|_{H_p^s} = \|A_p^{s/2}(\cdot)\|_{L^p}$. The domain of $A_p^{s/2}$ is then the Sobolev space of order s, that is $D(A_p^{s/2}) = H_p^s(\mathbb{R}^d)$, for all $s \in \mathbb{R}$. Furthermore, the negative powers $A_p^{-s/2}$ act as isomorphism from $H_p^\gamma(\mathbb{R}^d)$ onto $H_p^{\gamma+s}(\mathbb{R}^d)$ for $\gamma \in \mathbb{R}$.

We have defined so far function spaces and operators in the case of scalar valued functions. The extension to vector valued functions must be understood componentwise. For instance, the space $H_p^s\left(\mathbb{R}^d, \mathbb{R}^d\right)$ is the set of all vector fields $u : \mathbb{R}^d \to \mathbb{R}^d$ such that $u^i \in H_p^s\left(\mathbb{R}^d\right)$ for each component u^i of u; the vector field $P_p(t)\, u : \mathbb{R}^d \to \mathbb{R}^d$ has components $P_p(t)\, u^i$, and so on. Since we use vector fields more often than scalar functions, we shorten some of the notations: we shall write H_p^s for $H_p^s\left(\mathbb{R}^d, \mathbb{R}^d\right)$. Finally, we denote by $H_{p,q}^{-\beta}$ the space $H_p^{-\beta} \cap H_q^{-\beta}$ with the usual norm.

Lemma 4. *Let $1 < p, q < \infty$ and $0 < \beta < \delta$ and assume that $q > p \vee \frac{d}{\delta}$. Then for every $f \in H_p^\delta(\mathbb{R}^d)$ and $g \in H_q^{-\beta}(\mathbb{R}^d)$ we have $fg \in H_p^{-\beta}(\mathbb{R}^d)$ and there exists a positive constant c such that*

$$\|fg\|_{H_p^{-\beta}(\mathbb{R}^d)} \leq c\|f\|_{H_p^\delta(\mathbb{R}^d)} \cdot \|g\|_{H_q^{-\beta}(\mathbb{R}^d)}. \tag{9}$$

For the following the reader can also consult [25, Section 2.7.1]. Let us consider the spaces $C^{0,0}(\mathbb{R}^d;\mathbb{R}^d)$ and $C^{1,0}(\mathbb{R}^d;\mathbb{R}^d)$ defined as the closure of $\mathcal{S}$ with respect to the norm $\|f\|_{C^{0,0}} = \|f\|_{L^\infty}$ and $\|f\|_{C^{1,0}} = \|f\|_{L^\infty} + \|\nabla f\|_{L^\infty}$, respectively. For $0 < \alpha < 1$ we will consider the Banach spaces

$$\begin{aligned} C^{0,\alpha} &= \{f \in C^{0,0}(\mathbb{R}^d;\mathbb{R}^d) : \|f\|_{C^{0,\alpha}} < \infty\}, \\ C^{1,\alpha} &= \{f \in C^{1,0}(\mathbb{R}^d;\mathbb{R}^d) : \|f\|_{C^{1,\alpha}} < \infty\}, \end{aligned}$$

endowed with the norms

$$\begin{aligned} \|f\|_{C^{0,\alpha}} &:= \|f\|_{L^\infty} + \sup_{x\neq y\in\mathbb{R}^d} \frac{|f(x)-f(y)|}{|x-y|^\alpha} \\ \|f\|_{C^{1,\alpha}} &:= \|f\|_{L^\infty} + \|\nabla f\|_{L^\infty} + \sup_{x\neq y\in\mathbb{R}^d} \frac{|\nabla f(x)-\nabla f(y)|}{|x-y|^\alpha}, \end{aligned}$$

respectively.

From now on, we are going to make the following **standing assumption** on the drift b and on the possible choice of parameters:

Assumption 5. *Let $\beta \in \left(0, \frac{1}{2}\right)$, $q \in \left(\frac{d}{1-\beta}, \frac{d}{\beta}\right)$ and set $\tilde{q} := \frac{d}{1-\beta}$. The drift b will always be of the type*

$$b \in L^\infty\left([0,T]; H_{\tilde{q},q}^{-\beta}\right).$$

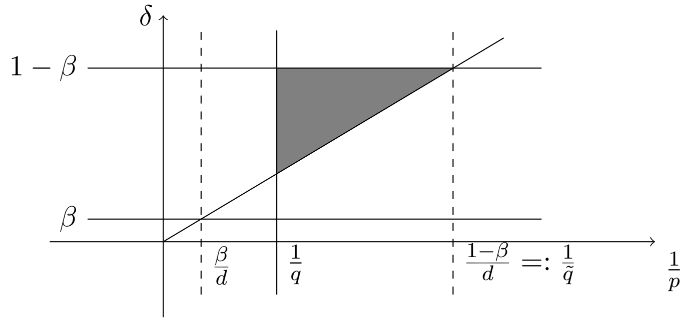

FIGURE 1. The set $K(\beta, q)$.

Remark 6. *The fact that* $b \in L^\infty\left([0,T]; H^{-\beta}_{\tilde{q},q}\right)$ *implies, for each* $p \in [\tilde{q}, q]$, *that* $b \in L^\infty\left([0,T]; H^{-\beta}_p\right)$.

Moreover we consider the set

$$K(\beta, q) := \left\{\kappa = (\delta, p) : \ \beta < \delta < 1-\beta, \ \frac{d}{\delta} < p < q\right\}, \tag{10}$$

which is drawn in Figure 1. Note that $K(\beta, q)$ is nonempty since $\beta < \frac{1}{2}$ and $\frac{d}{1-\beta} < q < \frac{d}{\beta}$.

Remark 7. *As a consequence of Lemma 4, for* $0 < \beta < \delta$ *and* $q > p \vee \frac{d}{\delta}$ *and if* $b \in L^\infty([0,T]; H^{-\beta}_q)$ *and* $u \in C^0([0,T]; H^{1+\delta}_p)$, *then for all* $t \in [0,T]$ *we have* $b(t) \cdot \nabla u(t) \in H^{-\beta}_p$ *and*

$$\|b(t) \cdot \nabla u(t)\|_{H^{-\beta}_p} \leq c \|b\|_{\infty, H^{-\beta}_q} \|u(t)\|_{H^\delta_p},$$

having used the continuity of ∇ *from* $H^{1+\delta}_p$ *to* H^δ_p. *Moreover any choice* $(\delta, p) \in K(\beta, q)$ *satisfies the hypothesis in Lemma 4.*

Definition 8. *Let* $(\delta, p) \in K(\beta, q)$. *We say that* $u \in C\left([0,T]; H^{1+\delta}_p\right)$ *is a mild solution of equation* (6) *in* $H^{1+\delta}_p$ *if*

$$u(t) = \int_t^T P_p(r-t)\, b(r) \cdot \nabla u(r)\, \mathrm{d}r + \int_t^T P_p(r-t)\, (b(r) - \lambda u(r))\, \mathrm{d}r, \tag{11}$$

for every $t \in [0,T]$.

Remark 9. *Notice that* $b \cdot \nabla u \in L^\infty\left([0,T]; H^{-\beta}_p\right)$ *by Remark 7. By Remark 6,* $b \in L^\infty\left([0,T]; H^{-\beta}_p\right)$. *Moreover* $\lambda u \in L^\infty\left([0,T]; H^{-\beta}_p\right)$ *by the embedding* $H^{1+\delta}_p \subset H^{-\beta}_p$. *Therefore the integrals in Definition 8 are meaningful in* $H^{-\beta}_p$.

Note that setting $v(t,x) := u(T-t, x)$, the PDE (6) can be equivalently rewritten as

$$\begin{cases} \partial_t v = L^b v - (\lambda+1)v + b & \text{on } [0,T] \times \mathbb{R}^d, \\ v(0) = 0 & \text{on } \mathbb{R}^d. \end{cases} \tag{12}$$

The notion of mild solutions of equation (12) in $\mathcal{S}'$ and in $H_p^{1+\delta}$ are analogous to Definition 2 and Definition 8, respectively. In particular the mild solution in $H_p^{1+\delta}$ verifies

$$v(t) = \int_0^t P_p(t-r)\left(b(r)\cdot\nabla v(r)\right)\mathrm{d}r + \int_0^t P_p(t-r)(b(r)-\lambda v(r))\mathrm{d}r. \tag{13}$$

Clearly the regularity properties of u and v are the same.
For a Banach space X we denote the usual norm in $L^\infty([0,T];X)$ by $\|f\|_{\infty,X}$ for $f\in L^\infty([0,T];X)$. Moreover, on the Banach space $C([0,T];X)$ with norm $\|f\|_{\infty,X} := \sup_{0\le t\le T}\|f(t)\|_X$ for $f\in C([0,T];X)$, we introduce a family of equivalent norms $\{\|\cdot\|_{\infty,X}^{(\rho)},\ \rho\ge 1\}$ as follows:

$$\|f\|_{\infty,X}^{(\rho)} := \sup_{0\le t\le T} \mathrm{e}^{-\rho t}\|f(t)\|_X. \tag{14}$$

Next we state a mapping property of the heat semigroup $P_p(t)$ on $L^p(\mathbb{R}^d)$: it maps distributions of fractional order $-\beta$ into functions of fractional order $1+\delta$ and the price one has to pay is a singularity in time. The proof is analogous to the one in [13, Prop. 3.2] and is based on the analyticity of the semigroup.

Lemma 10. *Let $0<\beta<\delta$, $\delta+\beta<1$ and $w\in H_p^{-\beta}(\mathbb{R}^d)$. Then $P_p(t)w\in H_p^{1+\delta}(\mathbb{R}^d)$ for any $t>0$ and moreover there exists a positive constant c such that*

$$\|P_p(t)w\|_{H_p^{1+\delta}(\mathbb{R}^d)} \le c\,\|w\|_{H_p^{-\beta}(\mathbb{R}^d)}\, t^{-\frac{1+\delta+\beta}{2}}. \tag{15}$$

Proposition 11. *Let $f\in L^\infty\left([0,T];H_p^{-\beta}\right)$ and $g:[0,T]\to H_p^{-\beta}$ for $\beta\in\mathbb{R}$ defined as*

$$g(t) = \int_0^t P_p(t-s)f(s)\,\mathrm{d}s.$$

Then $g\in C^\gamma\left([0,T];H_p^{2-2\epsilon-\beta}\right)$ for every $\epsilon>0$ and $\gamma\in(0,\epsilon)$.

Proof. First observe that for $f\in D(A_p^\gamma)$ there exists $C_\gamma>0$ such that

$$\|P_p(t)f-f\|_{L^p} \le C_\gamma t^\gamma \|f\|_{H_p^{2\gamma}}, \tag{16}$$

for all $t\in[0,T]$ (see [19, Thm 6.13, (d)]).
Let $0\le r<t\le T$. We have

$$\begin{aligned} g(t)-g(r) &= \int_0^t P_p(t-s)f(s)\,\mathrm{d}s - \int_0^r P_p(r-s)f(s)\,\mathrm{d}s \\ &= \int_r^t P_p(t-s)f(s)\,\mathrm{d}s + \int_0^r \left(P_p(t-s)-P_p(r-s)\right)f(s)\,\mathrm{d}s \\ &= \int_r^t P_p(t-s)f(s)\,\mathrm{d}s \\ &\quad + \int_0^r A_p^\gamma P_p(r-s)\left(A_p^{-\gamma}P_p(t-r)f(s) - A_p^{-\gamma}f(s)\right)\mathrm{d}s, \end{aligned}$$

so that

$$
\begin{aligned}
&\|g(t)-g(r)\|_{H_p^{2-2\epsilon-\beta}}\\
&\leq \int_r^t \|P_p(t-s)f(s)\|_{H_p^{2-2\epsilon-\beta}}\mathrm{d}s\\
&\quad + \int_0^r \|A_p^{\gamma}P_p(r-s)\left(A_p^{-\gamma}P_p(t-r)f(s)-A_p^{-\gamma}f(s)\right)\|_{H_p^{2-2\epsilon-\beta}}\mathrm{d}s\\
&\leq \int_r^t \|A_p^{1-\epsilon-\beta/2}P_p(t-s)f(s)\|_{L^p}\mathrm{d}s\\
&\quad + \int_0^r \|A_p^{1-\epsilon-\beta/2+\gamma}P_p(r-s)\left(A_p^{-\gamma}P_p(t-r)f(s)-A_p^{-\gamma}f(s)\right)\|_{L^p}\mathrm{d}s\\
&=: \text{(S1)}+\text{(S2)}.
\end{aligned}
$$

Let us consider (S1) first. We have

$$
\begin{aligned}
\text{(S1)} &\leq \int_r^t \|A_p^{1-\epsilon}P_p(t-s)\|_{L^p\to L^p}\|A^{-\beta/2}f(s)\|_{L^p}\mathrm{d}s\\
&\leq \int_r^t C_\epsilon (t-s)^{-1+\epsilon}\|f(s)\|_{H_p^{-\beta}}\mathrm{d}s\\
&\leq C_\epsilon (t-r)^{\epsilon}\|f\|_{\infty,H_p^{-\beta}},
\end{aligned}
$$

having used [19, Thm 6.13, (c)]. Using again the same result, the term (S2), together with (16), gives (with the constant C changing from line to line)

$$
\begin{aligned}
\text{(S2)} &= \int_0^r \left\|A_p^{1-\epsilon+\gamma}P_p(r-s)\left(P_p(t-r)A_p^{-\gamma-\beta/2}f(s)-A_p^{-\gamma-\beta/2}f(s)\right)\right\|_{L^p}\mathrm{d}s\\
&\leq C\int_0^r (r-s)^{-1+\epsilon-\gamma}\left\|P_p(t-r)A_p^{-\gamma-\beta/2}f(s)-A_p^{-\gamma-\beta/2}f(s)\right\|_{L^p}\mathrm{d}s\\
&\leq C\int_0^r (r-s)^{-1+\epsilon-\gamma}(t-r)^{\gamma}\|A_p^{-\gamma-\beta/2}f(s)\|_{H_p^{2\gamma}}\mathrm{d}s\\
&\leq C(t-r)^{\gamma}\int_0^r (r-s)^{-1+\epsilon-\gamma}\|f(s)\|_{H_p^{-\beta}}\mathrm{d}s\\
&\leq C(t-r)^{\gamma}\int_0^r (r-s)^{-1+\epsilon-\gamma}\|f\|_{\infty,H_p^{-\beta}}\mathrm{d}s\\
&\leq C(t-r)^{\gamma}r^{\epsilon-\gamma}\|f\|_{\infty,H_p^{-\beta}}.
\end{aligned}
$$

Therefore we have $g\in C^{\gamma}\left([0,T];H_p^{2-2\epsilon-\beta}\right)$ for each $0<\gamma<\epsilon$ and the proof is complete. □

The following lemma gives integral bounds which will be used later. The proof makes use of the Gamma and the Beta functions together with some basic integral estimates. We recall the definition of the Gamma function:

$$
\Gamma(a)=\int_0^\infty e^{-t}t^{a-1}\mathrm{d}t,
$$

and the integral converges for any $a\in\mathbb{C}$ such that $\mathrm{Re}(a)>0$.

Lemma 12. *If $0 \le s < t \le T < \infty$ and $0 \le \theta < 1$ then for any $\rho \ge 1$ it holds*

$$\int_s^t e^{-\rho r} r^{-\theta} \mathrm{d}r \le \Gamma(1-\theta)\rho^{\theta-1}. \tag{17}$$

Moreover if $\gamma > 0$ is such that $\theta + \gamma < 1$ then for any $\rho \ge 1$ there exists a positive constant C such that

$$\int_0^t e^{-\rho(t-r)} (t-r)^{-\theta} r^{-\gamma} \mathrm{d}r \le C\rho^{\theta-1+\gamma}. \tag{18}$$

Lemma 13. *Let $1 < p, q < \infty$ and $0 < \beta < \delta$ with $q > p \vee \frac{d}{\delta}$ and let $\beta + \delta < 1$. Then for $b \in L^\infty([0,T]; H^{-\beta}_{p,q})$ and $v \in C([0,T]; H^{1+\delta}_p)$ we have*

(i) $\int_0^\cdot P_p(\cdot - r) b(r) \mathrm{d}r \in C([0,T]; H^{1+\delta}_p)$*;*

(ii) $\int_0^\cdot P_p(\cdot - r)\left(b(r) \cdot \nabla v(r)\right) \mathrm{d}r \in C([0,T]; H^{1+\delta}_p)$ *with*

$$\left\| \int_0^\cdot P_p(\cdot - r)\left(b(r)\cdot \nabla v(r)\right) \mathrm{d}r \right\|^{(\rho)}_{\infty, H^{1+\delta}_p} \le c(\rho) \|v\|^{(\rho)}_{\infty, H^{1+\delta}_p};$$

(iii) $\lambda \int_0^\cdot P_p(\cdot - r) v(r) \mathrm{d}r \in C([0,T]; H^{1+\delta}_p)$ *with*

$$\left\| \lambda \int_0^\cdot P_p(\cdot - r) v(r) \mathrm{d}r \right\|^{(\rho)}_{\infty, H^{1+\delta}_p} \le c(\rho) \|v\|^{(\rho)}_{\infty, H^{1+\delta}_p},$$

where the constant $c(\rho)$ is independent of v and tends to zero as ρ tends to infinity.

Observe that $(\delta, p) \in K(\beta, q)$ satisfies the hypothesis in Lemma 13.

Proof. (i) Lemma 10 implies that $P_p(t)b(t) \in H^{1+\delta}_p$ for each $t \in [0,T]$. Choosing $\epsilon = \frac{1-\beta-\delta}{2}$, Proposition 11 implies item (i).

(ii) Similarly to part (i), the first part follows by Proposition 11. Moreover

$$\begin{aligned} &\sup_{0\le t\le T} \mathrm{e}^{-\rho t} \left\| \int_0^t P_p(t-r)\left(b(r)\cdot\nabla v(r)\right)\mathrm{d}r \right\|_{H^{1+\delta}_p} \\ &\le c \sup_{0\le t\le T} \int_0^t \mathrm{e}^{-\rho t} (t-r)^{-\frac{1+\delta+\beta}{2}} \|v(r)\|_{H^{1+\delta}_p} \|b(r)\|_{H^{-\beta}_q} \mathrm{d}r \\ &\le c\|b\|_{\infty, H^{-\beta}_q} \sup_{0\le t\le T} \int_0^t \mathrm{e}^{-\rho r} \|v(r)\|_{H^{1+\delta}_p} \mathrm{e}^{-\rho(t-r)} (t-r)^{-\frac{1+\delta+\beta}{2}} \mathrm{d}r \\ &\le c\|v\|^{(\rho)}_{\infty, H^{1+\delta}_p} \|b\|_{\infty, H^{-\beta}_q} \rho^{\frac{\delta+\beta-1}{2}} < \infty. \end{aligned}$$

Thus $\left\|\int_0^\cdot P_p(\cdot - r)\left(b(r)\cdot\nabla v(r)\right)\mathrm{d}r\right\|^{(\rho)}_{\infty, H^{1+\delta}_p} \le c(\rho)\|v\|^{(\rho)}_{\infty, H^{1+\delta}_p}$.

(iii) Similarly to parts (i) and (ii) the continuity property follows by Proposition 11. Then

$$\begin{aligned} \sup_{0\le t\le T} \mathrm{e}^{-\rho t} \left\| \int_0^t P_p(t-r) v(r) \mathrm{d}r \right\|_{H^{1+\delta}_p} &\le c \sup_{0\le t\le T} \int_0^t \mathrm{e}^{-\rho t} \|v(r)\|_{H^{1+\delta}_p} \mathrm{d}r \\ &\le c\|v\|^{(\rho)}_{\infty, H^{1+\delta}_p} \rho^{-1} < \infty. \end{aligned}$$

□

2.2. **Existence.** Let us now introduce the *integral operator* $I_t(v)$ as the right hand side of (13), that is, given any $v \in C([0,T]; H_p^{1+\delta})$, we define for all $t \in [0,T]$

$$I_t(v) := \int_0^t P_p(t-r)\,(b(r)\cdot\nabla v(r))\,\mathrm{d}r + \int_0^t P_p(t-r)(b(r)-\lambda v(r))\mathrm{d}r. \tag{19}$$

By Lemma 13, the integral operator is well-defined and it is a linear operator on $C([0,T]; H_p^{1+\delta})$.

Let us remark that Definition 8 is in fact meaningful under the assumptions of Lemma 13, which are more general than the ones of Definition 8 (see Remark 15).

Theorem 14. *Let $1 < p, q < \infty$ and $0 < \beta < \delta$ with $q > p \vee \frac{d}{\delta}$ and let $\beta + \delta < 1$. Then for $b \in L^\infty([0,T]; H_{p,q}^{-\beta})$ there exists a unique mild solution v to the PDE (13) in $H_p^{1+\delta}$. Moreover for any $0 < \gamma < 1-\delta-\beta$ the solution v is in $C^\gamma([0,T]; H_p^{1+\delta})$.*

Proof. By Lemma 13 the integral operator is a contraction for some ρ large enough, thus by the Banach fixed point theorem there exists a unique mild solution $v \in C([0,T]; H_p^{1+\delta})$ to the PDE (13). For this solution we obtain Hölder continuity in time of order γ for each $0 < \gamma < 1-\delta-\beta$. In fact each term on the right-hand side of (19) is γ-Hölder continuous by Proposition 11 as $b, b\cdot\nabla v, v \in L^\infty([0,T]; H_p^{-\beta})$. □

Remark 15. *By Theorem 14 and by the definition of $K(\beta, q)$, for each $(\delta, p) \in K(\beta, q)$ there exists a unique mild solution in $H_p^{1+\delta}$. However notice that the assumptions of Theorem 14 are slightly more general than those of Assumption 5 and of the set $K(\beta, q)$. Indeed, the following conditions are not required for the existence of the solution to the PDE (Lemma 13 and Theorem 14):*

- *the condition $\frac{d}{\delta} < p$ appearing in the definition of the region $K(\beta, q)$ is only needed in order to embed the fractional Sobolev space $H_p^{1+\delta}$ into $C^{1,\alpha}$ (Theorem 16).*
- *the condition $q < \frac{d}{\beta}$ appearing in Assumption 5 is only needed in Theorem 19 in order to show uniqueness for the solution u, independently of the choice of $(\delta, p) \in K(\beta, q)$.*

The following embedding theorem describes how to compare fractional Sobolev spaces with different orders and provides a generalisation of the Morrey inequality to fractional Sobolev spaces. For the proof we refer to [25, Thm. 2.8.1, Remark 2].

Theorem 16. Fractional Morrey inequality. *Let $0 < \delta < 1$ and $d/\delta < p < \infty$. If $f \in H_p^{1+\delta}(\mathbb{R}^d)$ then there exists a unique version of f (which we denote again by f) such that f is differentiable. Moreover $f \in C^{1,\alpha}(\mathbb{R}^d)$ with $\alpha = \delta - d/p$ and*

$$\|f\|_{C^{1,\alpha}} \le c\|f\|_{H_p^{1+\delta}}, \quad \|\nabla f\|_{C^{0,\alpha}} \le c\|\nabla f\|_{H_p^\delta}, \tag{20}$$

where $c = c(\delta, p, d)$ is a universal constant.
Embedding property. *For $1 < p \le q < \infty$ and $s - \frac{d}{p} \ge t - \frac{d}{q}$ we have*

$$H_p^s(\mathbb{R}^d) \subset H_q^t(\mathbb{R}^d). \tag{21}$$

Remark 17. *According to the fractional Morrey inequality, if $u(t) \in H_p^{1+\delta}$ then $\nabla u(t) \in C^{0,\alpha}$ for $\alpha = \delta - d/p$ if $p > d/\delta$. In this case the condition on the pointwise product $q > p \vee d/\delta$ reduces to $q > p$.*

2.3. **Uniqueness.** In this section we show that the solution u is unique, independently of the choice of $(\delta, p) \in K(\beta, q)$.

Lemma 18. *Let u be a mild solution in $\mathcal{S}'$ such that $u \in C\left([0,T]\,; H_p^{1+\delta}\right)$ for some $(\delta, p) \in K(\beta, q)$. Then u is a mild solution of (6) in $H_p^{1+\delta}$.*

Proof. As explained in Remark 9, $b \cdot \nabla u, b, \lambda u \in L^\infty\left([0,T]\,; H_p^{-\beta}\right)$. Given $\psi \in \mathcal{S}$ and $h \in H_p^{-\beta}$, we have

$$\langle h, P(s)\psi\rangle = \langle P_p(s)\,h, \psi\rangle\,, \tag{22}$$

for all $s \ge 0$. Indeed, $P_p(s)\,h = P(s)\,h$ when $h \in \mathcal{S}$ and $\langle P(s)\,h, \psi\rangle = \langle h, P(s)\,\psi\rangle$ when $h, \psi \in \mathcal{S}$, hence (22) holds for all $h, \psi \in \mathcal{S}$, therefore for all $h \in H_p^{-\beta}$ by density. Hence, from identity (8) we get

$$\begin{aligned}\langle u(t), \psi\rangle &= \int_t^T \langle P_p(r-t)\, b(r) \cdot \nabla u(r), \psi\rangle\, \mathrm{d}r \\ &\quad + \int_t^T \langle P_p(r-t)\,(b(r) - \lambda u(r)), \psi\rangle\, \mathrm{d}r.\end{aligned}$$

This implies (11). □

Theorem 19. *The solution u of (6) is unique, in the sense that for each $\kappa_1, \kappa_2 \in K(\beta, q)$, given two mild solutions $u^{\kappa_1}, u^{\kappa_2}$ of (6), there exists $\kappa_0 = (\delta_0, p_0) \in K(\beta, q)$ such that $u^{\kappa_1}, u^{\kappa_2} \in C([0,T]; H_{p_0}^{1+\delta_0})$ and the two solutions coincide in this bigger space.*

Proof. In order to find a suitable κ_0 we proceed in two steps.

Step 1. Assume first that $p_1 = p_2 =: p$. Then $H_{p_i}^{\delta_i} \subset H_p^{\delta_1 \wedge \delta_2}$. The intuition in Figure 1 is that we move downwards along the vertical line passing from $\frac{1}{p}$.

Step 2. If, on the contrary, $\frac{1}{p_1} < \frac{1}{p_2}$ (the opposite case is analogous) we may reduce ourselves to Step 1 in the following way: $H_{p_2}^{\delta_2} \subset H_{p_1}^x$ for $x = \delta_2 - \frac{d}{p_2} + \frac{d}{p_1}$ (using Theorem 16, equation (21)). Now $H_{p_1}^x$ and $H_{p_1}^{\delta_1}$ can be compared as in Step 1. The intuition in Figure 1 is that we move the rightmost point to the left along the line with slope d.

By Theorem 14 we have a unique mild solution u^{κ_i} in $C([0,T]; H_{p_i}^{1+\delta_i})$ for each set of parameters $\kappa_i = (\delta_i, p_i) \in K(\beta, q)$, $i = 0, 1, 2$. By Steps 1 and 2, the space with $i = 0$ includes the other two, thus $u^{\kappa_i} \in C([0,T]; H_{p_0}^{1+\delta_0})$ for each $i = 0, 1, 2$ and moreover u^{κ_i} are mild solutions in $\mathcal{S}'$. Lemma 18 concludes the proof. □

2.4. **Further regularity properties.** We derive now stronger regularity properties for the mild solution v of (13). Since $v(t,x) = u(T-t,x)$ the same properties hold for the mild solution u of (11).

In the following lemma we show that the mild solution v is differentiable in space and its gradient can be bounded by $\frac{1}{2}$ for some λ big enough. For this reason here we stress the dependence of the solution v on the parameter λ by writing v_λ.

Lemma 20. *Let $(\delta, p) \in K(\beta, q)$ and let v_λ be the mild solution to (12) in $H_p^{1+\delta}$. Fix ρ such that the integral operator (19) is a contraction on $C([0,T]; H_p^{1+\delta})$ with the norm (14) and let $\lambda > \rho$. Then $v_\lambda(t) \in C^{1,\alpha}$ with $\alpha = \delta - d/p$ for each fixed t and*

$$\sup_{0\le t\le T}\left(\sup_{x\in\mathbb{R}^d}|v_\lambda(t,x)|\right)\le C, \tag{23}$$

$$\sup_{0\le t\le T}\left(\sup_{x\in\mathbb{R}^d}|\nabla v_\lambda(t,x)|\right)\le \frac{c\|b\|_{\infty,H_p^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}}{1-c'\|b\|_{\infty,H_q^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}}, \tag{24}$$

for some universal constants C, c, c'. In particular,

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|\nabla v_\lambda(t,x)|\to 0,$$

as $\lambda\to\infty$.

Proof. By Theorem 16 and the definition of the set $K(\beta,q)$ we have that $v_\lambda(t) \in C^{1,\alpha}$ and (23) holds using the definition of the norms in $C([0,T]; H_p^{1+\delta})$ and $C^{1,\alpha}$.
Lemma 10 ensures that $P_t w \in H_p^{1+\delta}$ for $w \in H_p^{-\beta}$ and so $\nabla P_t w \in H_p^\delta$. By the fractional Morrey inequality (Theorem 16) we have $P_t w \in C^{1,\alpha}(\mathbb{R}^d)$ and for each $t>0$

$$\sup_{x\in\mathbb{R}^d}|\left(\nabla P_t w\right)(x)|\le c\|\nabla P_t w\|_{H_p^\delta}\le c\|P_t w\|_{H_p^{1+\delta}}\le ct^{-\frac{1+\delta+\beta}{2}}\|w\|_{H_p^{-\beta}}, \tag{25}$$

having used (15) in the latter inequality. Notice that the constant c depends only on δ, p and d.

If we assume for a moment that the mild solution v_λ of (12) is also a solution of

$$\begin{aligned} v_\lambda = &\int_0^t \mathrm{e}^{-\lambda(t-r)}P_p(t-r)\left(b(r)\cdot\nabla v_\lambda(r)\right)\mathrm{d}r \\ &+\int_0^t \mathrm{e}^{-\lambda(t-r)}P_p(t-r)b(r)\mathrm{d}r, \end{aligned} \tag{26}$$

then differentiating in x we get

$$\begin{aligned} \nabla v_\lambda(t,\cdot) = &\int_0^t \mathrm{e}^{-\lambda(t-r)}\nabla P_p(t-r)\left(b(r)\cdot\nabla v_\lambda(r)\right)\mathrm{d}r \\ &+\int_0^t \mathrm{e}^{-\lambda(t-r)}\nabla P_p(t-r)b(r)\mathrm{d}r. \end{aligned}$$

We take the H_p^δ-norm and use (25) with Lemma 4 to obtain

$$\begin{aligned}\|\nabla v_\lambda(t)\|_{H_p^\delta} \le & c\int_0^t \mathrm{e}^{-\lambda(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}}\|b(r)\|_{H_q^{-\beta}}\|\nabla v_\lambda(r)\|_{H_p^\delta}\mathrm{d}r \\ &+ c\int_0^t \mathrm{e}^{-\lambda(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}}\|b(r)\|_{H_p^{-\beta}}\mathrm{d}r \\ \le & c'\|b\|_{\infty,H_q^{-\beta}}\sup_{0\le r\le t}\|\nabla v_\lambda(r)\|_{H_p^\delta}\int_0^t \mathrm{e}^{-\lambda(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}}\mathrm{d}r \\ &+ c\|b\|_{\infty,H_p^{-\beta}}\int_0^t \mathrm{e}^{-\lambda(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}}\mathrm{d}r,\end{aligned}$$

so that by Lemma 12 we get

$$\begin{aligned}\sup_{0\le t\le T}\|\nabla v_\lambda(t)\|_{H_p^\delta} \le & c'\|b\|_{\infty,H_q^{-\beta}}\sup_{0\le t\le T}\|\nabla v_\lambda(t)\|_{H_p^\delta}\lambda^{\frac{\delta+\beta-1}{2}} \\ &+ c\|b\|_{\infty,H_p^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}.\end{aligned}$$

Choosing $\lambda > \lambda^* := \left(\frac{1}{c'\|b\|_{\infty,H_q^{-\beta}}}\right)^{\frac{2}{\delta+\beta-1}}$ yields

$$\sup_{0\le t\le T}\|\nabla v_\lambda(t)\|_{H_p^\delta} \le \frac{c\|b\|_{\infty,H_p^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}}{1-c'\|b\|_{\infty,H_q^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}},$$

which tends to zero as $\lambda\to\infty$. The fractional Morrey inequality (20) together with the latter bound gives

$$\begin{aligned}\sup_{0\le t\le T}\left(\sup_{x\in\mathbb{R}^d}|\nabla v_\lambda(t,x)|\right) &\le \sup_{0\le t\le T} c\|\nabla v_\lambda(t)\|_{H_p^\delta} \\ &\le \frac{c\|b\|_{\infty,H_p^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}}{1-c'\|b\|_{\infty,H_q^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}},\end{aligned}$$

which tends to zero as $\lambda\to\infty$.

It is left to prove that a solution of (13) in $H_p^{1+\delta}$ it is also a solution of (26). There are several proofs of this fact, let us see one of them. Computing each term against a test function $\psi\in\mathcal{S}$ we get the mild formulation

$$\begin{aligned}\langle v(t),\psi\rangle =& \int_0^t \langle b(r)\cdot\nabla v(r), P(t-r)\psi\rangle\,\mathrm{d}r \\ &+\int_0^t \langle b(r)-\lambda v(r), P(t-r)\psi\rangle\,\mathrm{d}r,\end{aligned}$$

used in the definition of mild solution in $\mathcal{S}'$. Let us choose in particular $\psi=\psi_k$ where $\psi_k(x)=e^{ix\cdot k}$, for a generic $k\in\mathbb{R}^d$, and let us write $v_k(t)=\left\langle v(t), e^{ix\cdot k}\right\rangle$ (the fact that ψ_k is complex-valued makes no difference, it is sufficient to treat separately the real and imaginary part). Using the explicit formula for $P(t)$, it is not difficult to check that

$$P(t)\psi_k = e^{-\left(|k|^2+1\right)t}\psi_k \tag{27}$$

and therefore

$$v_k(t) = \int_0^t e^{-\left(|k|^2+1\right)(t-r)} g_k(r)\,\mathrm{d}r - \lambda \int_0^t e^{-\left(|k|^2+1\right)(t-r)} v_k(r)\,\mathrm{d}r,$$

where $g_k(r) = \langle b(r) \cdot \nabla v(r) + b(r), \psi_k \rangle$. At the level of this scalar equation it is an easy manipulation to differentiate and rewrite it as

$$v_k(t) = \int_0^t e^{-\left(|k|^2+1+\lambda\right)(t-r)} g_k(r)\,\mathrm{d}r.$$

This identity, using again (27), can be rewritten as

$$\langle v(t), \psi_k \rangle = \int_0^t e^{-\lambda(t-r)} \langle b(r) \cdot \nabla v(r) + b(r), P(t-r)\psi_k \rangle\, dr$$

and then we deduce (26) as we did in the proof of Lemma 18. □

Lemma 21. *Let $v = v_\lambda$ for λ as in Lemma 20. Then v and ∇v are jointly continuous in (t, x).*

Proof. It is sufficient to prove the claim for ∇v. Let $(t, x), (s, y) \in [0, T] \times \mathbb{R}^d$. We have

$$\begin{aligned} |\nabla v(t,x) - \nabla v(s,y)| \leq & |\nabla v(t,x) - \nabla v(s,x)| + |\nabla v(s,x) - \nabla v(s,y)| \\ \leq & \sup_{x\in\mathbb{R}^d} |\nabla v(t,x) - \nabla v(s,x)| + |\nabla v(s,x) - \nabla v(s,y)| \\ \leq & \|v(t,\cdot) - v(s,\cdot)\|_{C^{1,\alpha}} + \|v(s,\cdot)\|_{C^{1,\alpha}} |x-y|^\alpha \\ \leq & \|v(t,\cdot) - v(s,\cdot)\|_{H_p^{1+\delta}} + \|v(s,\cdot)\|_{H_p^{1+\delta}} |x-y|^\alpha \\ \leq & \|v(t,\cdot) - v(s,\cdot)\|_{H_p^{1+\delta}} + \|v\|_{C^\gamma([0,T];H_p^{1+\delta})} |x-y|^\alpha \\ \leq & \|v\|_{C^\gamma([0,T];H_p^{1+\delta})} (|t-s|^\gamma + |x-y|^\alpha), \end{aligned}$$

having used the embedding property (20) with $\alpha = \delta - d/p$ and the Hölder property of v from Lemma 20. □

Lemma 22. *For λ large enough the function $x \mapsto \varphi(t, x)$ defined as $\varphi(t, x) = x + u(t, x)$ is invertible for each fixed $t \in [0, T]$ and, denoting its inverse by $\psi(t, \cdot)$ the function $(t, y) \mapsto \psi(t, y)$ is jointly continuous. Moreover $\psi(t, \cdot)$ is Lipschitz with Lipschitz constant $k = 2$, for every $t \in [0, T]$.*

We will sometimes use the shorthand notation φ_t for $\varphi(t, \cdot)$ and analogously for its inverse.

Proof. **Step 1** (invertibility of φ_t). Let t be fixed and $x_1, x_2 \in \mathbb{R}$. Recall that by Lemma 20 for λ large enough we have

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla u(t,x)| \leq \frac{1}{2}, \tag{28}$$

so that

$$|u(t,x_2) - u(t,x_1)| \leq \int_0^1 |\nabla u(t, ax_2 + (1-a)x_1)||x_1 - x_2|\mathrm{d}a \leq \frac{1}{2}|x_1 - x_2|.$$

Then the map $x \mapsto y - u(t, x)$ is a contraction for each $y \in \mathbb{R}^d$ and therefore for each $y \in \mathbb{R}^d$ there exists a unique $x \in \mathbb{R}^d$ such that $x = y - u(t, x)$ that is

$y = \varphi(t,x)$. Thus $\varphi(t,\cdot)$ is invertible for each $t \in [0,T]$ with inverse denoted by ψ_t.

Step 2 (Lipschitz character of ψ_t, uniformly in t). To show that ψ_t is Lipschitz with constant k we can equivalently show that for each $x_1, x_2 \in \mathbb{R}^d$ it holds $|\varphi_t(x_1) - \varphi_t(x_2)| \geq \frac{1}{k}|x_1 - x_2|$. We have

$$|\varphi_t(x_1) - \varphi_t(x_2)| \geq \inf_{x\in\mathbb{R}^d} |\nabla\varphi(t,x)||x_1 - x_2| = \frac{1}{2}|x_1 - x_2|,$$

because of (28) together with $\nabla\varphi = \mathrm{I}_d + \nabla u$.

Step 3 (continuity of $s \mapsto \psi(s,y)$). Let us fix $y \in \mathbb{R}^d$ and take $t_1, t_2 \in [0,T]$. Denote by $x_1 = \psi(t_1,y)$ and $x_2 = \psi(t_2,y)$ so that $y = \varphi(t_1,x_1) = x_1 + u(t_1,x_1)$ and $y = \varphi(t_2,x_2) = x_2 + u(t_2,x_2)$. We have

$$\begin{aligned} |\psi(t_1,y) - \psi(t_2,y)| &= |x_1 - x_2| \\ &= |u(t_1,x_1) - u(t_2,x_2)| \\ &\leq |u(t_1,x_1) - u(t_1,x_2)| + |u(t_1,x_2) - u(t_2,x_2)| \\ &\leq \frac{1}{2}|x_1 - x_2| + |u(t_1,x_2) - u(t_2,x_2)|. \end{aligned} \tag{29}$$

Let us denote by $w(x) := u(t_1,x) - u(t_2,x)$. Clearly $w \in H^{1+\delta}_p$ for each t_1, t_2 and by Theorem 16 (Morrey inequality) we have that w is continuous, bounded and

$$|u(t_1,x_2) - u(t_2,x_2)| \leq \sup_{x\in\mathbb{R}^d} |w(x)| \leq c\|w\|_{H^{1+\delta}_p}.$$

By Theorem 14 $u \in C^\gamma([0,T]; H^{1+\delta}_p)$ and so $\|w\|_{H^{1+\delta}_p} \leq c|t_1 - t_2|^\gamma$. Using this result together with (29) we obtain

$$\frac{1}{2}|x_1 - x_2| = \frac{1}{2}|\psi(t_1,y) - \psi(t_2,y)| \leq c|t_1 - t_2|^\gamma,$$

which shows the claim.

Continuity of $(t,y) \mapsto \psi(t,y)$ now follows. □

Lemma 23. *If $b_n \to b$ in $L^\infty\left([0,T]; H^{-\beta}_{\tilde{q},q}\right)$ then $v_n \to v$ in $C([0,T]; H^{1+\delta}_p)$.*

Proof. Let $\lambda > 0$ be fixed. We consider the integral equation (13) on $H^{1+\delta}_p$ so the semigroup will be denoted by P_p. Observe that by Lemma 10 we have

$$\begin{aligned} &\|P_p(t-r)\left(b_n(r)\cdot\nabla v_n(r) - b(r)\cdot\nabla v(r)\right)\|_{H^{1+\delta}_p} \\ &\quad\leq c(t-r)^{-\frac{1+\delta+\beta}{2}} \|b_n(r)\cdot\nabla v_n(r) - b(r)\cdot\nabla v(r)\|_{H^{-\beta}_p} \\ &\quad\leq c(t-r)^{-\frac{1+\delta+\beta}{2}} \Big(\|b_n(r)\|_{H^{-\beta}_q}\|v_n(r) - v(r)\|_{H^{1+\delta}_p} \\ &\qquad + \|b_n(r) - b(r)\|_{H^{-\beta}_q}\|v(r)\|_{H^{1+\delta}_p} \Big) \\ &\quad\leq c(t-r)^{-\frac{1+\delta+\beta}{2}} \Big(\|b_n\|_{\infty,H^{-\beta}_q}\|v_n(r) - v(r)\|_{H^{1+\delta}_p} \\ &\qquad + \|b_n - b\|_{\infty,H^{-\beta}_q}\|v(r)\|_{H^{1+\delta}_p} \Big), \end{aligned}$$

where the second to last line is bounded through Lemma 4. Thus, by (13)

$$
\begin{aligned}
\|v-v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} &= \sup_{0\le t\le T} \mathrm{e}^{-\rho t}\|v(t)-v_n(t)\|_{H^{1+\delta}_p}\\
\le& \sup_{0\le t\le T}\mathrm{e}^{-\rho t}\Big(\int_0^t \|P_p(t-r)\left(b_n(r)\cdot\nabla v_n(r)-b(r)\cdot\nabla v(r)\right)\|_{H^{1+\delta}_p}\,\mathrm{d}r\\
&+\int_0^t \big\|P_p(t-r)\big(b_n(r)-b(r)+\lambda(v(r)-v_n(r))\big)\big\|_{H^{1+\delta}_p}\,\mathrm{d}r\Big)\\
\le& \sup_{0\le t\le T}\mathrm{e}^{-\rho t}\Big(c\|b_n\|_{\infty,H^{-\beta}_q}\int_0^t (t-r)^{-\frac{1+\delta+\beta}{2}}\|v_n(r)-v(r)\|_{H^{1+\delta}_p}\mathrm{d}r\\
&+c\|b_n-b\|_{\infty,H^{-\beta}_q}\int_0^t (t-r)^{-\frac{1+\delta+\beta}{2}}\|v(r)\|_{H^{1+\delta}_p}\mathrm{d}r\\
&+c\|b_n-b\|_{\infty,H^{-\beta}_q}\int_0^t (t-r)^{-\frac{1+\delta+\beta}{2}}\mathrm{d}r + c\lambda\int_0^t\|v(r)-v_n(r)\|_{H^{1+\delta}_p}\mathrm{d}r\Big)\\
\le& c\|b_n\|_{\infty,H^{-\beta}_q}\sup_{0\le t\le T}\int_0^t \mathrm{e}^{-\rho(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}}\mathrm{e}^{-\rho r}\|v_n(r)-v(r)\|_{H^{1+\delta}_p}\mathrm{d}r\\
&+c\|b_n-b\|_{\infty,H^{-\beta}_q}\cdot\\
&\cdot\sup_{0\le t\le T}\int_0^t \mathrm{e}^{-\rho(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}}\mathrm{e}^{-\rho r}\left(\|v(r)\|_{H^{1+\delta}_p}+1\right)\mathrm{d}r\\
&+c\lambda\sup_{0\le t\le T}\int_0^t \mathrm{e}^{-\rho(t-r)}\mathrm{e}^{-\rho r}\|v_n(r)-v(r)\|_{H^{1+\delta}_p}\mathrm{d}r,
\end{aligned}
$$

where we have used again Lemma 10. Consequently

$$
\begin{aligned}
\|v-v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} \le& c\|b_n\|_{\infty,H^{-\beta}_q}\|v_n-v\|^{(\rho)}_{\infty,H^{1+\delta}_p}\rho^{\frac{\delta+\beta-1}{2}}\\
&+c\|b_n-b\|_{\infty,H^{-\beta}_q}\left(\|v\|^{(\rho)}_{\infty,H^{1+\delta}_p}+1\right)\rho^{\frac{\delta+\beta-1}{2}}\\
&+c\lambda\|v_n-v\|^{(\rho)}_{\infty,H^{1+\delta}_p}\rho^{-1}.
\end{aligned}
$$

The last bound is due to Lemma 12. Since $\|b_n\|_{\infty,H^{-\beta}_q}\to\|b\|_{\infty,H^{-\beta}_q}$ then there exists $n_0\in\mathbb{N}$ such that $\|b_n\|_{\infty,H^{-\beta}_q}\le 2\|b\|_{\infty,H^{-\beta}_q}$ for all $n\ge n_0$. Choose now ρ big enough in order to have

$$1-c\left(\|b\|_{\infty,H^{-\beta}_q}\rho^{\frac{\delta+\beta-1}{2}}+\lambda\rho^{-1}\right)>0$$

and then we have for each $n\ge n_0$

$$\|v-v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p}\le c\frac{\left(\|v\|^{(\rho)}_{\infty,H^{1+\delta}_p}+1\right)\rho^{\frac{\delta+\beta-1}{2}}}{1-c\left(\|b\|_{\infty,H^{-\beta}_q}\rho^{\frac{\delta+\beta-1}{2}}+\lambda\rho^{-1}\right)}\|b_n-b\|_{\infty,H^{-\beta}_q},$$

which concludes the proof. □

Lemma 24. (i) *Let $\|b_n\|_{\infty,H^{-\beta}_{\tilde q,q}}\le c\|b\|_{\infty,H^{-\beta}_{\tilde q,q}}$ for a constant c not depending on n. Then there exists a constant $C>0$ such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left(|u_n(t,x)|+|\nabla u_n(t,x)|\right)\le C,$$

for every $n \in \mathbb{N}$.

(ii) *There exists $\lambda \geq 0$ such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla u_n(t,x)| \leq \frac{1}{2} \tag{30}$$

and

$$\sup_{(t,y)\in[0,T]\times\mathbb{R}^d} |\nabla \psi_n(t,y)| \leq 2,$$

for every $n \in \mathbb{N}$ and similarly for ∇u and $\nabla \psi$.

(iii) *If $b_n \to b$ in $L^\infty([0,T]; H^{-\beta}_{\tilde{q},q})$, then we have $u_n \to u$, $\nabla u_n \to \nabla u$, $\varphi_n \to \varphi$ and $\psi_n \to \psi$ uniformly on $[0,T] \times \mathbb{R}^d$.*

Proof. (i) The proof has the same structure as the proof of Lemma 23, but slightly simplified as the difference $v_n - v$ is replaced with v_n. In the following bounds the constant c may change from line to line and one gets

$$\begin{aligned}
\|v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} \leq& c\|b_n\|_{\infty,H^{-\beta}_{\tilde{q},q}} \|v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} \rho^{\frac{\delta+\beta-1}{2}} \\
&+ c\|b_n\|_{\infty,H^{-\beta}_q} \rho^{\frac{\delta+\beta-1}{2}} + c\lambda \|v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} \rho^{-1} \\
\leq& c\|b\|_{\infty,H^{-\beta}_{\tilde{q},q}} \|v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} \rho^{\frac{\delta+\beta-1}{2}} \\
&+ c\|b\|_{\infty,H^{-\beta}_{\tilde{q},q}} \rho^{\frac{\delta+\beta-1}{2}} + c\lambda \|v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} \rho^{-1},
\end{aligned}$$

where the latter bound holds thanks to the assumption on the b_n's. Now we choose ρ large enough such that

$$1 - c\left(\|b\|_{\infty,H^{-\beta}_{\tilde{q},q}} \rho^{\frac{\delta+\beta-1}{2}} + \rho^{-1} \right) > 0$$

and get for every $n \in \mathbb{N}$

$$\|v_n\|^{(\rho)}_{\infty,H^{1+\delta}_p} \leq \frac{c\rho^{\frac{\delta+\beta-1}{2}}}{1 - c\left(\|b\|_{\infty,H^{-\beta}_{\tilde{q},q}} \rho^{\frac{\delta+\beta-1}{2}} + \rho^{-1} \right)} \|b\|_{\infty,H^{-\beta}_{\tilde{q},q}} =: C.$$

(ii) The uniform bound (30) on ∇u_n is obtained simply applying Lemma 20 to u_n in place of u_λ. For what concerns the second bound involving $\nabla \psi_n$ we observe that $\nabla \varphi_n(t,x)$ is non-degenerate uniformly in t, x, n since for each $\xi \in \mathbb{R}^d$ we have

$$|\nabla \varphi_n(t,x) \cdot \xi| \geq |\xi| - |\nabla u_n(t,x) \cdot \xi| \geq \frac{1}{2}|\xi|,$$

having used (30) for the latter inequality. This implies that $\nabla \psi_n(t,x)$ is well-defined for each (t,x). Further note that by Lemma 22 we have that $\psi_n(t,\cdot)$ is Lipschitz with constant $k = 2$, uniformly in t and n, and this now implies the claim.

(iii) We know that $u_n \to u$ in $C\left([0,T]; H^{1+\delta}_p\right)$, namely

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|u_n(t) - u(t)\|_{H^{1+\delta}_p} = 0.$$

By Sobolev embedding theorem, there is a constant $C > 0$ such that

$$\begin{aligned}&\sup_{t\in[0,T]}\left(\sup_{x\in\mathbb{R}^d}|u_n(t,x)-u(t,x)|+\sup_{x\in\mathbb{R}^n}|\nabla u_n(t,x)-\nabla u(t,x)|\right)\\&\leq C\sup_{t\in[0,T]}\|u_n(t)-u(t)\|_{H_p^{1+\delta}}.\end{aligned}$$

Hence $u_n \to u$ and $\nabla u_n \to \nabla u$, uniformly on $[0,T]\times\mathbb{R}^d$. Since $\varphi_n - \varphi = u_n - u$, we also have that $\varphi_n \to \varphi$ uniformly on $[0,T]\times\mathbb{R}^d$. Let us prove the uniform convergence of ψ_n to ψ.

Given $y \in \mathbb{R}^d$, we know that for every $t \in [0,T]$ and $n \in \mathbb{N}$ there exist $x(t), x_n(t) \in \mathbb{R}^d$ such that

$$\begin{aligned}x(t)+u(t,x(t))&=y\\x_n(t)+u_n(t,x_n(t))&=y\end{aligned}$$

and we have called $x(t)$ and $x_n(t)$ by $\psi(t,y)$ and $\psi_n(t,y)$ respectively. Then

$$\begin{aligned}|x_n(t)-x(t)|&=|u_n(t,x_n(t))-u(t,x(t))|\\&\leq|u_n(t,x_n(t))-u_n(t,x(t))|+|u_n(t,x(t))-u(t,x(t))|\\&\leq\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|\nabla u_n(t,x)|\,|x_n(t)-x(t)|\\&\quad+\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u_n(t,x)-u(t,x)|.\end{aligned}$$

Since $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|\nabla u_n(t,x)| \leq \frac{1}{2}$, we deduce

$$|x_n(t)-x(t)|\leq 2\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u_n(t,x)-u(t,x)|,$$

namely

$$|\psi_n(t,y)-\psi(t,y)|\leq 2\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u_n(t,x)-u(t,x)|$$

which implies that $\psi_n \to \psi$ uniformly on $[0,T]\times\mathbb{R}^d$. □

3. The virtual solution

From now on, we fix λ and ρ big enough so that Theorem 14 and Lemma 22 hold true. As usual, the drift b is chosen according to Assumption 5.

3.1. **Heuristics and motivation.** We consider the following d-dimensional SDE

$$\mathrm{d}X_t = b(t,X_t)\mathrm{d}t+\mathrm{d}W_t,\quad t\in[0,T], \tag{31}$$

with initial condition $X_0 = x$ where b is a distribution. Formally, the integral form is

$$X_t = x+\int_0^t b(s,X_s)\mathrm{d}s+W_t,\quad t\in[0,T], \tag{32}$$

but the integral appearing on the right hand side is not well-defined, a priori. We aim to give a meaning to this equation by introducing a suitable notion of solution to the SDE (31). Let u be a mild solution to the PDE (6): we shall make use of u to define a notion of solution to the SDE (31).

By stochastic basis we mean a pentuple $(\Omega, \mathcal{F}, \mathbb{F}, P, W)$ where $(\Omega, \mathcal{F}, P)$ is a complete probability space with a completed filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ and W is a d-dimensional $\mathbb{F}$-Brownian motion. In the spirit of weak solutions, we cannot assume that $\mathbb{F}$ is the completed filtration associated to W.

Definition 25. *Given $x \in \mathbb{R}^d$, a* virtual solution *to the SDE* (31) *with initial value x is a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P, W)$ and a continuous stochastic process $X := (X_t)_{t\in[0,T]}$ on it, $\mathbb{F}$-adapted, such that the integral equation*
(33)
$$X_t = x+u(0,x)-u(t,X_t)+(\lambda+1)\int_0^t u(s,X_s)\mathrm{d}s+\int_0^t (\nabla u(s,X_s)+\mathrm{I}_d)\mathrm{d}W_s,$$
holds for all $t \in [0,T]$, with probability one. Here I_d denotes the $d\times d$ identity matrix and u is the unique mild solution to the PDE (6). *We shorten the notation and say that $(X, \mathbb{F})$ is a virtual solution when the previous objects exist with the required properties.*

The motivation for this definition comes from two facts: i) the not-properly-defined expression $\int_0^t b(s,X_s)\mathrm{d}s$ does not appear in the formulation; ii) when b is a function with reasonable regularity, classical solutions of the SDE (31) are also virtual solutions; this is the content of Proposition 26, where we will illustrate this fact by considering two examples, one of which is the class of drifts investigated by [17]. Similar arguments can be developed for the bounded measurable drift considered by [26].

3.2. Existence and uniqueness of the virtual solution.

To find a virtual solution $(X, \mathbb{F})$ to (31) we first make the following observation. Let us assume that $(X, \mathbb{F})$ is a virtual solution of (31) with initial value $X_0 = x$ and let us introduce the transformation $\varphi(t,x) := x+u(t,x)$ and set $Y_t = \varphi(t,X_t)$ for $t \in [0,T]$. From (33) we obtain
$$\varphi(t,X_t) = x + u(0,x) + (\lambda+1)\int_0^t u(s,X_s)\mathrm{d}s + \int_0^t (\nabla u(s,X_s)+\mathrm{I}_d)\mathrm{d}W_s.$$
Since the function $\varphi(t,\cdot)$ is invertible for all $t \in [0,T]$, we can consider the SDE
$$(34)\quad Y_t = y + (\lambda+1)\int_0^t u(s,\psi(s,Y_s))\mathrm{d}s + \int_0^t (\nabla u(s,\psi(s,Y_s))+\mathrm{I}_d)\mathrm{d}W_s,$$
for $t \in [0,T]$, where $y = x+u(0,x)$. Hence $(Y, \mathbb{F})$ where $Y := (Y_t)_{t\in[0,T]}$, is a solution of (34) with initial value $y \in \mathbb{R}$. Conversely, if $(Y, \mathbb{F})$ is the solution of (34) with initial value $y \in \mathbb{R}$, then $(X, \mathbb{F})$ defined by
$$X_t = \psi(t,Y_t), \quad t \in [0,T],$$
will give us the virtual solution of the SDE (31) with distributional drift and with initial value $x = \psi(0,y)$.

As mentioned above, to gain a better understanding of the concept of virtual solution, we first compare it to some classical solutions. For example let us consider the class of drifts investigated by [17]. Let b be a measurable function $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ such that
$$\int_0^T \left(\int_{\mathbb{R}^d} |b\,(t,x)|^p\,dx\right)^{q/p} \mathrm{d}t < \infty,$$

(we say that $b \in L_t^q(L_x^p)$) for some $p, q \geq 2$ such that

$$\frac{d}{p} + \frac{2}{q} < 1.$$

Under this assumption, there exists a strong solution $(X, \mathbb{F})$ to the SDE (31) and it is pathwise unique, see [17].

Proposition 26. *Suppose that one of the following conditions holds:*

(i) $b \in C\left([0,T]; C_b^1\left(\mathbb{R}^d; \mathbb{R}^d\right)\right)$ *(bounded with bounded first derivatives);*
(ii) $b \in L_t^q(L_x^p)$.

Then the classical solution $(X, \mathbb{F})$ *to the SDE* (31) *is also a virtual solution.*

Proof. Suppose condition (i) holds. Let u be the unique classical solution of equation (6); u is (at least) of class $C^{1,2}\left([0,T] \times \mathbb{R}^d; \mathbb{R}^d\right)$. Since $\varphi(t,x) = x + u(t,x)$ then $\varphi \in C^{1,2}\left([0,T] \times \mathbb{R}^d; \mathbb{R}^d\right)$ as well. Let X be the unique strong solution of equation (31) and let $Y_t = \varphi(t, X_t)$. By Itô's formula, Y satisfies equation (34) and thus $X_t = \psi(t, Y_t)$ is also a virtual solution.

Suppose now that condition (ii) holds. The solution u of the PDE (6), when b is of class $L_t^q(L_x^p)$ with the assumed constraints on (q,p), belongs to $L_t^q(L_x^p)$ with its first and second spatial derivatives, the first spatial derivatives are continuous and bounded, and other regularity properties hold; see [17]. In particular, it is proved there that Itô's formula extends to such functions u and we get

$$\begin{aligned} \mathrm{d}u(t,X_t) = & \left(\frac{\partial u}{\partial t}(t,X_t) + \frac{1}{2}\Delta u(t,X_t) + \nabla u(t,X_t) b(t,X_t)\right)\mathrm{d}t \\ & + \nabla u(t,X_t)\mathrm{d}W_t \\ = & (\lambda+1)u(t,X_t)\mathrm{d}t - b(t,X_t)\mathrm{d}t + \nabla u(t,X_t)\mathrm{d}W_t. \end{aligned}$$

The integral form of the last equation

$$u(t,X_t) = u(0,x) + (\lambda+1)\int_0^t u(s,X_s)\mathrm{d}s - \int_0^t b(s,X_s)\mathrm{d}s + \int_0^t \nabla u(s,X_s)\mathrm{d}W_s,$$

allows us to evaluate the singular term $\int_0^t b(s,X_s)\mathrm{d}s$ as

$$\int_0^t b(s,X_s)\mathrm{d}s = u(0,x) - u(t,X_t) + (\lambda+1)\int_0^t u(s,X_s)\mathrm{d}s + \int_0^t \nabla u(s,X_s)\mathrm{d}W_s.$$

This proves identity (33) and thus $(X, \mathbb{F})$ is a virtual solution. □

Proposition 27. *For every initial condition* $y \in \mathbb{R}$ *there exists a unique weak solution* $(Y, \mathbb{F})$ *to the SDE* (34) *with initial value* y.

Proof. We know that $u, \nabla u$ and ψ are jointly continuous in time and space by Lemma 21 and Lemma 22. This implies that the drift of Y

$$\mu(t,y) := (\lambda+1)u(t,\psi(t,y))$$

and the diffusion coefficient

$$\sigma(t,y) := \nabla u(t,\psi(t,y)) + \mathrm{I}_d = \nabla\varphi(t,\psi(t,y))$$

are continuous. Since by Lemma 20 the function u and its gradient are uniformly bounded, we also have that μ and σ are uniformly bounded. Moreover σ is uniformly non-degenerate since for all $x, \xi \in \mathbb{R}^d$ and $t \in [0,T]$

$$\begin{aligned}|\sigma^T(t,x)\xi| &= |\xi + \xi \cdot \nabla u(t, \psi(t,y))| \\ &\geq |\xi| - |\xi \cdot \nabla u(t,\psi(t,y))| \geq \frac{1}{2}|\xi|\end{aligned}$$

by (28). Thus Theorem 10.2.2 in [24] yields existence and uniqueness of a weak solution of SDE (34) for every initial value $y \in \mathbb{R}$. □

Theorem 28. *For every $x \in \mathbb{R}$ there exists a unique in law virtual solution $(X, \mathbb{F})$ to the SDE (31) with initial value $X_0 = x$ given by $X_t = \psi(t, Y_t), t \in [0,T]$, where $(Y, \mathbb{F})$ is the solution with initial value $y = x + u(0,x)$ given in Proposition 27.*

Proof. We shorten the notation in the proof and write X and Y in place of $(X, \mathbb{F})$ and $(Y, \mathbb{F})$ respectively.
Existence. Let us fix the initial condition $x \in \mathbb{R}$. By Proposition 27 there exists a unique solution Y to the SDE (34) with initial value $y = x + u(0,x)$. Let X be defined by $X_t = \psi(t, Y_t), t \in [0,T]$. By construction X is a virtual solution of (31) with initial condition $X_0 = x$.
Uniqueness. Suppose that $Z := (Z_t)_{t\geq 0}$ is another virtual solution to (31). Then $\tilde{Y}$ defined by $\tilde{Y}_t = \varphi(t, Z_t), t \in [0,T]$ is a solution to (34) with initial value $\tilde{Y}_0 = \varphi(0,x) = x + u(0,x)$. Since equation (34) admits uniqueness in law, the law of $\tilde{Y}$ coincides with the law of Y and by the invertibility of $\varphi(t,\cdot)$ for each $t \in [0,T]$ we get that the laws of X and Z coincide. □

3.3. **Virtual solution as limit of classical solutions.** The concept of *virtual solution* is very convenient in order to prove weak existence and uniqueness; however, it may look a bit artificial. Moreover, a priori, the virtual solution may depend on the parameter λ. These problems are solved by the next proposition which identifies the virtual solution (for any λ) as the limit of classical solutions. This result relates also to the concept of solution introduced in [2].

Proposition 29. *Let $b_n : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be vector fields such that*

(i) $b_n \in C\left([0,T]; C_b^1\left(\mathbb{R}^d; \mathbb{R}^d\right)\right)$ *(bounded with bounded first derivatives);*
(ii) $b_n \to b$ *in* $L^\infty\left([0,T]; H^{-\beta}_{\tilde{q},q}\right)$.

Then the unique strong solution to the equation

$$dX_t^n = \mathrm{d}W_t + b_n(t, X_t^n)\,\mathrm{d}t, \qquad X_0 = x, \tag{35}$$

converges in law to the virtual solution $(X, \mathbb{F})$ of equation (31).

Proof. We shorten the notation in the proof and write X and Y in place of $(X, \mathbb{F})$ and $(Y, \mathbb{F})$ respectively. Recall that λ has been chosen big enough at the beginning of the section.

Step 1 (X^n as virtual solutions). Let u_n be the unique classical solution of equation (6) replacing b with b_n. By Proposition 26, part (i), we have that the unique strong solutions X^n of equations (35) are also virtual solutions. Here $Y_t^n = \varphi_n(t, X_t^n)$ satisfies equation (34) with b_n replacing b. Let us denote by $\tilde{b}_n$ and $\tilde{\sigma}_n$ (respectively $\tilde{b}$ and $\tilde{\sigma}$) the drift and diffusion coefficient of the equation for Y^n (respectively Y), that is

$$\begin{aligned}\tilde{b}_n(t,x) &:= (\lambda+1)u_n(t,\psi_n(t,x)),\\ \tilde{\sigma}_n(t,x) &:= \nabla u_n(t,\psi_n(t,x)) + \mathrm{I}_d.\end{aligned}$$

Step 2 (Upper bounds on u_n and ∇u_n, uniformly in n). Since b_n converges to b in $L^\infty([0,T]; H^{-\beta}_{\tilde{q},q})$ there exists n_0 such that for each $n \geq n_0$ we have $\|b_n\|_{\infty, H^{-\beta}_{\tilde{q},q}} \leq 2\|b\|_{\infty, H^{-\beta}_{\tilde{q},q}}$ and so we can apply Lemma 24 (i) and find a constant $C_1 > 0$ such that

$$\begin{aligned}|u_n(r,z)| &\leq C_1\\ |\nabla u_n(r,z) + \mathrm{I}_d| &\leq C_1,\end{aligned}$$

for all $(r,z) \in [0,T] \times \mathbb{R}^d$.

Step 3 (Tightness of the laws of Y^n). For $0 \leq s \leq t \leq T$ we have

$$\begin{aligned}|Y_t^n - Y_s^n|^4 \leq{}& 8(\lambda+1)^4 \left(\int_s^t |u_n(r,\psi_n(r,Y_r^n))| \,\mathrm{d}r\right)^4\\ &+ 8\left|\int_s^t [\nabla u_n(r,\psi_n(r,Y_r^n)) + \mathrm{I}_d]\,\mathrm{d}W_r\right|^4.\end{aligned}$$

By the result in Problem 3.29 and Remark 3.30 of [15] (see also Theorem 4.36 in [5] given even in Hilbert spaces) there is a constant $C_2 > 0$ such that

$$\begin{aligned}E\left[|Y_t^n - Y_s^n|^4\right] \leq{}& 8(\lambda+1)^4 C_1^4 (t-s)^4\\ &+ 8C_2 E\left[\left(\int_s^t |\nabla u_n(r,\psi_n(r,Y_r^n)) + \mathrm{I}_d|^2 \,\mathrm{d}r\right)^2\right]\\ \leq{}& 8(\lambda+1)^4 C_1^4 (t-s)^4 + 8C_1^4 C_2 (t-s)^2.\end{aligned}$$

This obviously implies

$$E\left[|Y_t^n - Y_s^n|^4\right] \leq C\,|t-s|^2,$$

for some constant $C > 0$ depending on T and independent of n (recall that λ is given). Moreover the initial condition y_0 is real and independent of n, thus a tightness criterion (see Corollary 16.9, in [14]) implies the tightness of the laws of Y^n in $C\left([0,T];\mathbb{R}^d\right)$.

Step 4 (Weak convergence of Y^n to a solution Y). Let Y be the process defined by $Y_t = \varphi(t, X_t)$, $t \in [0,T]$, where X is the virtual solution in the statement of the proposition. We want to prove that the laws of Y^n converge weakly to the law of Y. Denote by μ^n and μ the laws of Y^n and Y, respectively. To show weak convergence of μ^n to μ it is enough to prove that, for any subsequence μ^{n_k}, there exists a further subsequence $\mu^{n_{k_j}}$ which converges weakly to μ. Given μ^{n_k}, a converging subsequence $\mu^{n_{k_j}}$ exists, since $\{\mu^n\}$ is tight by Step 3, so $\{\mu^{n_k}\}$ is also tight. Denote by μ' the weak limit of $\mu^{n_{k_j}}$ as $j \to \infty$. If we prove that $\mu' = \mu$, for any choice of the subsequence μ^{n_k}, then we have that the whole sequence μ^n converges to μ.

Just to simplify notations, we shall denote $\mu^{n_{k_j}}$ by μ^n and assume that μ^n converges weakly, as $n \to \infty$, to μ'.

Since $\mu^n \rightharpoonup \mu'$, by Skorokhod's representation theorem there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and random variables $\tilde{Y}^n$ (resp. $\tilde{Y}$) taking values in $C([0,T];\mathbb{R}^d)$ endowed with the Borel σ-field, with laws μ^n (resp. μ'), such that $\tilde{Y}^n \to \tilde{Y}$ in $C([0,T];\mathbb{R}^d)$ a.s. If we prove that $\tilde{Y}$ is a weak solution of equation (34), since uniqueness in law holds for that equation and Y is another weak solution, we get $\mu' = \mu$. To prove that $\tilde{Y}$ is a weak solution of equation (34), it is sufficient to prove (see Theorem 18.7 in [14]) that $\tilde{Y}$ solves the following martingale problem: for every $g \in C^\infty(\mathbb{R}^d;\mathbb{R})$ with compact support the process

$$\tilde{M}_t := g(\tilde{Y}_t) - g(\tilde{Y}_0) - \int_0^t L_r g(\tilde{Y}_r)\mathrm{d}r$$

is a martingale, where

$$L_r := \frac{1}{2}\sum_{i,j=1}^d \tilde{a}_{i,j}(r,\cdot)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d \tilde{b}_i(r,\cdot)\frac{\partial}{\partial x_i}$$

and $\tilde{a}(r,x) := \tilde{\sigma}(r,x)\tilde{\sigma}^*(r,x)$. Let $0 \le s < t \le T$ and let $F : C([0,s];\mathbb{R}^d) \to \mathbb{R}$ be a bounded continuous functional. To prove that $\tilde{M}$ is a martingale it is enough to show that

$$E\left[\left(\tilde{M}_t - \tilde{M}_s\right) F(\tilde{Y}_r; r \le s)\right] = 0 \tag{36}$$

on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

Recall that Y^n satisfies equation (34) (with b replaced by b_n). Hence, by Itô's formula,

$$M_t^n := g(Y_t^n) - g(Y_0^n) - \int_0^t L_r^n g(Y_r^n)\mathrm{d}r$$

is a martingale, where

$$L_r^n := \frac{1}{2}\sum_{i,j=1}^d \tilde{a}_{i,j}^n(r,\cdot)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d \left(\tilde{b}_n\right)_i(r,\cdot)\frac{\partial}{\partial x_i}$$

and $\tilde{a}^n(r,x) := \tilde{\sigma}_n(r,x)\tilde{\sigma}_n^*(r,x)$. Therefore

$$E\left[(M_t^n - M_s^n)\, F(Y_r^n; r \le s)\right] = 0.$$

This identity depends only on the law of Y^n, hence

$$E\left[\left(\tilde{M}_t^n - \tilde{M}_s^n\right) F(\tilde{Y}_t^n; r \le s)\right] = 0. \tag{37}$$

Let us denote by $\tilde{Z}_{s,t}^n$ (resp. $\tilde{Z}_{s,t}$) the random variable inside the expectation in (37) (resp. (36)). Each factor in $\tilde{Z}_{s,t}^n$ (resp. $\tilde{Z}_{s,t}$) is uniformly bounded, by the boundedness of F, g and its derivatives, $\tilde{b}_n$ and $\tilde{\sigma}_n$ (due to boundedness of u and ∇u). Moreover $\tilde{Z}_{s,t}^n \to \tilde{Z}_{s,t}$ a.s. since $\tilde{Y}^n(\omega) \to \tilde{Y}(\omega)$ for almost every ω uniformly on compact sets, F, g and its derivatives are continuous, $\tilde{b}_n \to \tilde{b}$ and $\tilde{\sigma}_n \to \tilde{\sigma}$ uniformly on compact sets (remind that $\psi_n \to \psi, u^n \to u, \nabla u^n \to \nabla u$ uniformly by Lemma 24 (iii)). Therefore, by Lebesgue's dominated convergence theorem we conclude that (36) holds.

Step 5 (Weak convergence of X^n to X). The final step consists in showing that X^n converges to X in law. Recall that $X^n = \psi_n(\cdot, Y^n)$. By Lemma 22, ψ is uniformly continuous on compact sets and $Y^n \to Y$ in law by Step 4; it is then easy to deduce that $\psi(\cdot, Y^n) \to \psi(\cdot, Y)$ in law. Indeed given a continuous and bounded functional $F : C([0,T];\mathbb{R}^d) \to \mathbb{R}$, the functional $\eta \mapsto F(\psi(\cdot, \eta(\cdot)))$ is still a continuous bounded functional. Moreover $\psi_n(\cdot, Y^n) - \psi(\cdot, Y^n) \to 0$ in $C([0,T];\mathbb{R}^d)$ P-a.s. since $\psi_n \to \psi$ uniformly by Lemma 24 (iii); hence $\psi_n(\cdot, Y^n) - \psi(\cdot, Y^n) \to 0$ in probability. Then by Theorem 4.1 in [3, Section 4, Chapter 1] we have that $\psi_n(\cdot, Y^n) \to \psi(\cdot, Y)$ weakly, which concludes the proof. □

Examples of b_n which verify (ii) in Proposition 29 are easily obtained by convolutions of b against a sequence of mollifiers converging to a Dirac measure.

Acknowledgments. The authors wish to thank the anonymous referee for her/his careful revision which helped to clarify and improve considerably the initial version of this paper.

The third named author benefited from the support of the "FMJH Program Gaspard Monge in optimization and operation research" (Project 2014-1607H). The second and third named authors were also partially supported by the ANR Project MASTERIE 2010 BLAN 0121 01.

References

[1] R. F. Bass and Z.-Q. Chen. Stochastic differential equations for Dirichlet processes. *Probab. Theory Related Fields*, 121(3):422–446, 2001.

[2] R. F. Bass and Z.-Q. Chen. Brownian motion with singular drift. *Ann. Probab.*, 31(2):791–817, 2003.

[3] P. Billingsley. *Convergence of Probability measures*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, second edition, 1968.

[4] S. Blei and H. J. Engelbert. One-dimensional stochastic differential equations with generalized and singular drift. *Stochastic Process. Appl.*, 123(12):4337–4372, 2013.

[5] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

[6] E. B. Davies. *Heat kernels and spectral theory*, volume 92 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.

[7] H.-J. Engelbert and W. Schmidt. On one-dimensional stochastic differential equations with generalized drift. In *Stochastic differential systems (Marseille-Luminy, 1984)*, volume 69 of *Lecture Notes in Control and Inform. Sci.*, pages 143–155. Springer, Berlin, 1985.

[8] H.-J. Engelbert and J. Wolf. Strong Markov local Dirichlet processes and stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 43(2):331–348, 1998.

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

[10] F. Flandoli, F. Russo, and J. Wolf. Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.*, 40(2):493–542, 2003.

[11] F. Flandoli, F. Russo, and J. Wolf. Some SDEs with distributional drift. II. Lyons-Zheng structure, Itô's formula and semimartingale characterization. *Random Oper. Stochastic Equations*, 12(2):145–184, 2004.

[12] M. Hinz and M. Zähle. Gradient type noises. II. Systems of stochastic partial differential equations. *J. Funct. Anal.*, 256(10):3192–3235, 2009.

[13] E. Issoglio. Transport equations with fractal noise—existence, uniqueness and regularity of the solution. *Z. Anal. Anwend.*, 32(1):37–53, 2013.

[14] O. Kallenberg. *Foundation of Modern Probability.* Springer, second edition, 2002.

[15] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1991.

[16] J. Karatzas and I. Ruf. Pathwise solvability of stochastic integral equationswith generalized drift and non-smooth dispersion functions. *preprint*, 2013. arXiv:1312.7257v1 [math.PR].

[17] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[18] Y. Ouknine. Le "Skew-Brownian motion" et les processus qui en dérivent. *Teor. Veroyatnost. i Primenen.*, 35(1):173–179, 1990.

[19] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1983.

[20] N. I. Portenko. *Generalized diffusion processes*, volume 83 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, RI, 1990. Translated from the Russian by H. H. McFaden.

[21] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *de Gruyter Series in Nonlinear Analysis and Applications.* Walter de Gruyter & Co., Berlin, 1996.

[22] F. Russo and G. Trutnau. About a construction and some analysis of time inhomogeneous diffusions on monotonely moving domains. *J. Funct. Anal.*, 221(1):37–82, 2005.

[23] F. Russo and G. Trutnau. Some parabolic PDEs whose drift is an irregular random noise in space. *Ann. Probab.*, 35(6):2213–2262, 2007.

[24] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1979.

[25] H. Triebel. *Interpolation theory, function spaces, differential operators*, volume 18 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam, 1978.

[26] A. Y. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Math. USSR Sb.*, 39:387–403, 1981.

[27] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.