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

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

$x_0 \in \mathbb{R}^d$, does not make sense, excepted if we consider it in a very general *generalized functions* sense. 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 previous equation with a noise perturbation, which is expected to have a regularizing effect, i.e.

$$dX_t = b(t, X_t)dt + dW_t, \quad X_0 = x_0, \tag{2}$$

where W is a standard d-dimensional Brownian motion. Formally speaking, the Kolmogorov equation associated with previous stochastic differential equation is

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

for suitable final conditions f. Equation (3) was studied in the one-dimensional setting for instance by [17] for any b which is derivative of a continuous

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

[2]Department of Mathematics, King's College London, Strand, London, WC2R 2LS, UK Elena.Issoglio@kcl.ac.uk

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

Date: July 17, 2014.

function and in the multidimensional setting by [10], for a class of b of gradient type belonging to a given Sobolev space with negative derivation order. The equation in [10] involves the product of distributions in the sense of *paraproduct*, which is a natural extension of pointwise product for distributions.

The point of view of the present paper is to keep the same interpretation of the product as in [10] 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 23. 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

$$dX_t = dW_t + b_n(t, X_t)dt, \tag{4}$$

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 [14]; 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, 12, 16]. In all these cases solutions were semimartingales. More recently, [5] 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

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt, \tag{5}$$

where σ is a strictly positive continuous function and b is the derivative of a real continuous function was solved and analyzed carefully in [7] and [8], which treated well-posedness of the martingale problem, Itô formula under weak conditions, semimartingale characterization and Lyons-Zheng decomposition. The only supplementary assumption was the existence of the function $\Sigma(x) = 2\int_0^x \frac{b}{\sigma^2} dy$ as limit of appropriate regularizations. Bass and Chen [1] were also interested in (2) and they provided a well-stated framework when σ is γ-Hölder continuous and b is γ-Hölder continuous, $\gamma > \frac{1}{2}$. In [17] the authors have also shown that in some cases the SDE can be considered in the strong (probabilistic) sense, i.e. when the probability space and the Brownian motion are fixed at the beginning.
As far as the multidimensional case is concerned, it seems that the first paper was again of Bass and Chen, see [2]. Those authors have focused (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 [8] 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 [11] has considered other type of transforms to study similar equations. Indeed the transformation introduced by Zvonkin in [20], 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 [6]), have efficiently made use of a (multidimensional) Zvonkin type transform for the study of an SDE with measurable non necessarily bounded drift, which however is still a function. Zvonkin transform consisted there to transform a solution X to (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 $Y_t = \varphi(t, X_t)$ is a solution of an SDE for which one can show pathwise existence and 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 [10], based on paraproducts 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 observing that $Y_t = \varphi(t, X_t)$ where $\varphi(t,x) = x + u(t,x), (t,x) \in [0,T]\times\mathbb{R}^d$ and 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 following 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 $i = 1,\dots,d$. 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 [15]. We recall below a few elements of this theory; in particular, when we say that the paraproduct exists in $\mathcal{S}'$ we mean that the limit (14) 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, $1 \leq p \leq \infty$.

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

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

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*

$$\begin{aligned}\langle u(t),\psi\rangle &= \int_t^T \langle b(r)\cdot\nabla u(r),P(r-t)\psi\rangle\,dr \\ &\quad+\int_t^T \langle b(r)-\lambda u(r),P(r-t)\psi\rangle\,dr.\end{aligned}\tag{7}$$

Here $(P(t))_{t\geq 0}$ denotes the heat semigroup on $\mathcal{S}$ generated by $\frac{1}{2}\Delta - I$, defined for each $\psi\in\mathcal{S}$ as

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

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

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

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

The fractional Sobolev spaces H_q^r 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* $L^1\left(0,T;H_{q_i}^{r_i}\right)$ *for different* $r_i\in\mathbb{R}$, $q_i\geq 1$, $i=1,2,3$, *then (see e.g.* (20)*) there exist common* $r\in\mathbb{R}$, $q\geq 1$ *such that* $b,u,b\cdot\nabla u\in L^1\left([0,T];H_q^r\right)$.

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 [3, Thm. 1.4.1, 1.4.2]. The fractional powers of A_p of order $\alpha\in\mathbb{R}$ are then well defined, see [13]. 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.

For the following, see [19, 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

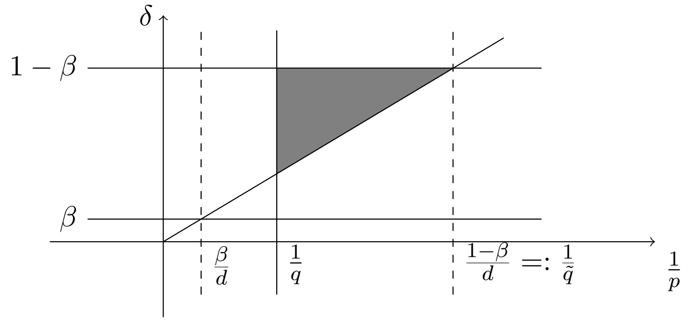

FIGURE 1. The set $K(\beta, q)$.

the norm $\|f\|_{C^{0,0}} = \|f\|_{L^\infty}$ and $\|f\|_{C^{1,0}} = \|f\|_{L^\infty} + \|\nabla f\|_{L^\infty}$, respectively. For $\alpha > 0$ we will consider the Banach spaces

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

endowed with the norms

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

respectively.

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

Assumption 4. *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*

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

Remark 5. *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

$$K(\beta, q) := \left\{\kappa = (\delta, p) : \ \beta < \delta < 1-\beta, \ \frac{d}{\delta} < p < q\right\} \tag{8}$$

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

Definition 6. *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*

$$u(t) = \int_t^T P_p(r-t)\, b(r) \cdot \nabla u(r)\, dr + \int_t^T P_p(r-t)\,(b(r) - \lambda u(r))\, dr, \tag{9}$$

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

Remark 7. *Notice that* $b \cdot \nabla u \in L^\infty\left([0,T]\,;H_p^{-\beta}\right)$ *by Lemma 10. By Remark 5,* $b \in L^\infty\left([0,T]\,;H_p^{-\beta}\right)$. *Moreover* $\lambda u \in L^\infty\left([0,T]\,;H_p^{-\beta}\right)$ *by the embedding* $H_p^{1+\delta} \subset H_p^{-\beta}$. *Therefore the integrals in Definition 6 are meaningful in* $H_p^{-\beta}$.

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

$$\begin{cases} \partial_t v = L^b v - (\lambda+1)v + b, & \text{on } [0,T]\times\mathbb{R}^d \\ v(0) = 0 & \text{on } \mathbb{R}^d\,. \end{cases} \tag{10}$$

The notion of mild solutions in $\mathcal{S}'$ and in $H_p^{1+\delta}$ are analogous to Definition 2 and Definition 6, respectively. In particular the mild solution in $H_p^{1+\delta}$ is given by

$$v(t) = \int_0^t P_p(t-r)\left(b(r)\cdot\nabla v(r)\right)\mathrm{d}r + \int_0^t P_p(t-r)(b(r)-\lambda v(r))\mathrm{d}r. \tag{11}$$

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 $C^0([0,T];X)$ with norm $\|f\|_{0,X} := \sup_{0\le t\le T}\|f(t)\|_X$ for $f \in C^0([0,T];X)$, we introduce a family of equivalent norms $\{\|\cdot\|_{0,X}^{(\rho)}\,,\ \rho\ge 1\}$ as follows:

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

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 [10, Prop. 3.2] and is based on the analyticity of the semigroup.

Lemma 8. *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*

$$\|P_p(t)w\|_{H_p^{1+\delta}(\mathbb{R}^d)} \le c\,\|w\|_{H_p^{-\beta}(\mathbb{R}^d)}\, t^{-\frac{1+\delta+\beta}{2}}. \tag{12}$$

Proposition 9. *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 $f\in D(A_p^\gamma)$ then there exists $C_\gamma>0$ such that

$$\|P_p(t)f - f\|_{L^p} \le C_\gamma t^\gamma \|f\|_{H_p^{2\gamma}} \tag{13}$$

for all $t\in[0,T]$ (see [13, 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

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

Let us consider (S1) first. We have

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

having used [13, Thm 6.13, (c)]. Moreover, the term (S2), together with (13), gives

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

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. □

We now recall a definition of a paraproduct between a function and a distribution (see e. g. [15]) 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 function ψ. This approximation is used to define the product fg of two distributions as follows:

$$fg := \lim_{j\to\infty} S^j f S^j g \tag{14}$$

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 [9] appendix C.4, [15] Theorem 4.4.3/1).

Lemma 10. *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*

$$\|fg\|_{H_p^{-\beta}(\mathbb{R}^d)} \le c\|f\|_{H_p^{\delta}(\mathbb{R}^d)} \cdot \|g\|_{H_q^{-\beta}(\mathbb{R}^d)}. \tag{15}$$

As a consequence of this lemma, for $0<\beta<\delta$ and $q > p \vee \frac{d}{\delta}$ and if $b \in L^{\infty}([0,T];H_q^{-\beta})$ and $u \in C^0([0,T];H_p^{1+\delta})$, then for all $t \in [0,T]$ we have $b(t)\cdot\nabla u(t) \in H_p^{-\beta}$ and

$$\|b(t)\cdot\nabla u(t)\|_{H_p^{-\beta}} \le c\|b\|_{\infty,H_q^{-\beta}}\|u(t)\|_{H_p^{\delta}}$$

having used the continuity of ∇ from $H_p^{1+\delta}$ to H_p^{δ}. Moreover any choice $(\delta,p) \in K(\beta,q)$ satisfies the hypothesis in Lemma 10.

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

$$\int_s^t e^{-\rho r}r^{-\theta}\mathrm{d}r \le \Gamma(1-\theta)\rho^{\theta-1}. \tag{16}$$

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

$$\int_0^t e^{-\rho(t-r)}(t-r)^{-\theta}r^{-\gamma}\mathrm{d}r \le C\rho^{\theta-1+\gamma}. \tag{17}$$

Lemma 12. *Let $1 < p, q < \infty$ and $0 < \beta < \delta$ with $q > p \vee \frac{d}{\delta}$ and let $\beta + \delta < 1$. Then for $b \in L^\infty([0,T]; H_{p,q}^{-\beta})$ and $v \in C^0([0,T]; H_p^{1+\delta})$ we have*

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

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

$$\left\|\int_0^\cdot P_p(\cdot - r)\left(b(r) \cdot \nabla v(r)\right)\mathrm{d}r\right\|_{0,H_p^{1+\delta}}^{(\rho)} \leq c(\rho)\|v\|_{0,H_p^{1+\delta}}^{(\rho)};$$

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

$$\left\|\lambda\int_0^\cdot P_p(\cdot - r)v(r)\mathrm{d}r\right\|_{0,H_p^{1+\delta}}^{(\rho)} \leq c(\rho)\|v\|_{0,H_p^{1+\delta}}^{(\rho)},$$

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 12.

Proof. (i) By Lemma 8 we have that $P_p(t)b(t) \in H_p^{1+\delta}$ and

$$\begin{aligned}
\left\|\int_0^\cdot P_p(\cdot - r)b(r)\mathrm{d}r\right\|_{0,H_p^{1+\delta}}^{(\rho)} &= \sup_{0\leq t\leq T} \mathrm{e}^{-\rho t}\left\|\int_0^t P_p(t-r)b(r)\mathrm{d}r\right\|_{H_p^{1+\delta}} \\
&\leq \sup_{0\leq t\leq T}\int_0^t \mathrm{e}^{-\rho t}(t-r)^{-\frac{1+\delta+\beta}{2}}\|b(r)\|_{H_p^{-\beta}}\mathrm{d}r \\
&\leq \|b\|_{\infty,H_p^{-\beta}}\sup_{0\leq t\leq T}\int_0^t \mathrm{e}^{-\rho t}(t-r)^{-\frac{1+\delta+\beta}{2}}\mathrm{d}r \\
&\leq c\|b\|_{\infty,H_p^{-\beta}}\rho^{\frac{\delta+\beta-1}{2}} < \infty,
\end{aligned}$$

having used Lemma 11 for the last inequality.
(ii) Similarly to part (i) we have

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

(iii) Similarly to parts (i) and (ii) we get

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

□

2.2. **Existence.** Let us now introduce the *integral operator* $I_t(v)$ as the right hand side of (11), that is, given any $v \in C^0([0,T]; H_p^{1+\delta})$, we define for all $t \in [0,T]$

$$I_t(v) := \int_0^t P_p(t-r)\left(b(r)\cdot\nabla v(r)\right)\mathrm{d}r + \int_0^t P_p(t-r)(b(r) - \lambda v(r))\mathrm{d}r. \tag{18}$$

By Lemma 12, the integral operator is well defined and it is a linear operator on $C^0([0,T]; H_p^{1+\delta})$.

Let us remark that Definition 6 is in fact meaningful under the assumptions of Lemma 12, which are more general than the one of Definition 6 (see Remark 14).

Theorem 13. *Under the condition of Lemma 12, there exists a unique mild solution v to the PDE (11) 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 12 the integral operator is a contraction for some ρ large enough, thus by the Banach fixed point theorem there exists a unique mild solution $v \in C^0([0,T]; H_p^{1+\delta})$ to the PDE (11). 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 (18) is γ-Hölder continuous by Proposition 9 as $b, b\cdot\nabla v, v \in L^\infty([0,T]; H_p^{-\beta})$. □

Remark 14. *By Theorem 13 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 13 are slightly more general than those of Assumption 4 and of the set $K(\beta,q)$. Indeed, the following conditions are not required for the existence of the solution to the PDE (Lemma 12 and Theorem 13):*

- *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 15).*
- *the condition $q < \frac{d}{\beta}$ appearing in Assumption 4 is only needed in Theorem 18 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 Morrey inequality to fractional Sobolev spaces. For the proof we refer to [19, Thm. 2.8.1, Remark 2].

Theorem 15. 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 f) such that f is differentiable. Moreover $f \in C^{1,\alpha}(\mathbb{R}^d)$ with $\alpha = \delta - d/p$ and*

$$\|f\|_{C^{1,\alpha}} \le c\|f\|_{H_p^{1+\delta}}, \quad \|\nabla f\|_{C^{0,\alpha}} \le c\|\nabla f\|_{H_p^\delta}, \tag{19}$$

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*

$$H_p^s(\mathbb{R}^d) \subset H_q^t(\mathbb{R}^d). \tag{20}$$

Remark 16. *According to the fractional Morrey inequality, for $u(t) \in H_p^{1+\delta}$ then $\nabla u(t) \in C^{0,\alpha}$ for $\alpha = \delta - d/p$ if $p > d/\delta$. In this case the condition on the paraproduct $q > \max\{p, d/\delta\}$ 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 17. *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 solution in $H_p^{1+\delta}$.*

Proof. As explained in Remark 7, $b \cdot \nabla u, b, \lambda u \in L^\infty\left([0,T]\,; H_p^{-\beta}\right)$. Given $\psi \in \mathcal{S}$ and $h \in H_p^{-\beta}$, we have

$$\langle h, P\left(s\right)\psi\rangle = \langle P_p\left(s\right)h, \psi\rangle \tag{21}$$

for all $s \ge 0$. Indeed, $P_p\left(s\right)h = P\left(s\right)h$ when $h \in \mathcal{S}$ and $\langle P\left(s\right)h, \psi\rangle = \langle h, P\left(s\right)\psi\rangle$ when $h, \psi \in \mathcal{S}$, hence (21) holds for all $h, \psi \in \mathcal{S}$, therefore for all $h \in H_p^{-\beta}$ by density. Hence, from identity (7) we get

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

This implies (9). □

Theorem 18. *The solution u of (6) is unique, in the sense that for each $\kappa_1, \kappa_2 \in K(\beta, q)$ there exists $\kappa_0 = (\delta_0, p_0) \in K(\beta, q)$ such that $u^{\kappa_1}, u^{\kappa_2} \in C^0([0,T]; H_{p_0}^{1+\delta_0})$ and the two solutions coincide in this bigger space.*

Proof. In order to find a suitable κ_0 we proceed in 2 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 15, equation (20)). 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 13 we have a unique mild solution u^{κ_i} in $C^0([0,T]; H_{p_i}^{1+\delta_i})$ 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 $u^{\kappa_i} \in C^0([0,T]; H_{p_0}^{1+\delta_0})$ for each $i=0,1,2$ and moreover u^{κ_i} are mild solutions in $\mathcal{S}'$. Lemma 17 concludes the proof. □

2.4. **Further regularity properties.** We derive now stronger regularity properties for the mild solution v of (11). Since $v(t,x) = u(T-t,x)$ the same properties hold for the mild solution u of (9).

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 19. *Let $(\delta, p) \in K(\beta, q)$ and let v_λ be the mild solution to* (11) *in $H_p^{1+\delta}$. Fix ρ such that the integral operator is a contraction and let $\lambda > \rho$. Then $v_\lambda(t) \in C^{1,\alpha}$ with $\alpha = \delta - d/p$ for each fixed t and*

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

where the choice of λ depends only on $\delta, \beta, \|b\|_{\infty, H_p^{-\beta}}$ and $\|b\|_{\infty, H_q^{-\beta}}$.

Proof. Lemma 8 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 15) we have that $P_t w \in C^{1,\alpha}(\mathbb{R}^d)$ and for each $t > 0$

$$\sup_{x\in\mathbb{R}^d} |\left(\nabla P_t w\right)(x)| \leq c \|\nabla P_t w\|_{H_p^\delta} \leq c \|P_t w\|_{H_p^{1+\delta}} \leq c t^{-\frac{1+\delta+\beta}{2}} \|w\|_{H_p^{-\beta}}, \tag{22}$$

having used (12) 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 (11) is also a solution of

$$\begin{aligned} v_\lambda = &\int_0^t \mathrm{e}^{-\lambda(t-r)} P_p(t-r) \left(b(r)\cdot\nabla v_\lambda(r)\right) \mathrm{d}r \\ &+ \int_0^t \mathrm{e}^{-\lambda(t-r)} P_p(t-r) b(r) \mathrm{d}r, \end{aligned} \tag{23}$$

then differentiating in x we get

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

Take the H_p^δ-norm and use (22) with Lemma 10 to obtain

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

so that by Lemma 11 we get

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

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

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

which tends to zero as $\lambda\to\infty$. The fractional Morrey inequality (19) together with the latter bound gives

$$\begin{aligned}\sup_{0\leq t\leq T}\left(\sup_{x\in\mathbb{R}^d}|\nabla v_\lambda(t,x)|\right) & \leq \sup_{0<t<T} c\|\nabla v(t)\|_{H_p^\delta} \\ & \leq \frac{c\|b\|_{\infty,H_p^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}}{1-c'\|b\|_{\infty,H_q^{-\beta}}\lambda^{\frac{\delta+\beta-1}{2}}} \qquad (24)\end{aligned}$$

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

It is left to prove that a solution of (11) in $H_p^{1+\delta}$ it is also a solution of (23). 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

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

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

$$P(t)\psi_k = e^{-\left(|k|^2+1\right)t}\psi_k \qquad (25)$$

and therefore

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

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

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

This identity, using again (25), 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 (23) as we did in the proof of Lemma 17. □

Lemma 20. *Let $v = v_\lambda$ for λ as in Lemma 19. 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

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

having used the embedding property (19) with $\alpha = \delta - d/p$ and the Hölder property of v from Lemma 19. □

Lemma 21. *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 the inverse $(t,y) \mapsto \varphi^{-1}(t,y)$ is jointly continuous. Moreover φ^{-1} is Lipschitz with Lipschitz constant $k = 2$.*

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 19 for λ large enough we have

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla u(t,x)| \le \frac{1}{2} \tag{26}$$

so that

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

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 φ_t^{-1}.

Step 2 (Lipschitz character of φ_t^{-1}). To show that φ_t^{-1} 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 (26) together with $\nabla\varphi = \mathrm{I} + \nabla u$.

Step 3 (continuity of $s \mapsto \varphi^{-1}(s, y)$). Let us fix $y \in \mathbb{R}^d$ and take $t_1, t_2 \in [0, T]$. Denote by $x_1 = \varphi^{-1}(t_1, y)$ and $x_2 = \varphi^{-1}(t_2, y)$ 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

$$\begin{aligned} |\varphi^{-1}(t_1, y) - \varphi^{-1}(t_2, y)| &= |x_1 - x_2| \\ &= |u(t_1, x_1) - u(t_2, x_2)| \\ &\leq |u(t_1, x_1) - u(t_1, x_2)| + |u(t_1, x_2) - u(t_2, x_2)| \\ &\leq \frac{1}{2}|x_1 - x_2| + |u(t_1, x_2) - u(t_2, x_2)|. \end{aligned} \tag{27}$$

Let us denote by $w(x) := u(t_1, x) - u(t_2, x)$. Clearly $w \in H_p^{1+\delta}$ for each t_1, t_2 and by Theorem 15 (Morrey inequality) we have that w is continuous, bounded and

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

By Theorem 13 $u \in C^\gamma([0, T]; H_p^{1+\delta})$ and so $\|w\|_{H_p^{1+\delta}} \leq c|t_1 - t_2|^\gamma$. Using this result together with (27) we obtain

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

which shows the claim.
Continuity of $(t, y) \mapsto \varphi^{-1}(t, y)$ now follows. ☐

Lemma 22. *If $b_n \to b$ in $L^\infty\left([0, T]; H_{\tilde{q},q}^{-\beta}\right)$ then $v_n \to v$ in $C^0([0, T]; H_p^{1+\delta})$.*

Proof. Let $\lambda > 0$ be fixed. We consider the integral equation (11) on $H_p^{1+\delta}$ so the semigroup will be denoted by P_p. Observe that by Lemma 8 we have

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

where the second to last line is bounded through Lemma 10. Thus, by (11)

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

where we have used again Lemma 8. Consequently

$$
\begin{aligned}
\|v-v_n\|^{(\rho)}_{0,H_p^{1+\delta}} \le & c\|b_n\|_{\infty,H_q^{-\beta}}\|v_n-v\|^{(\rho)}_{0,H_p^{1+\delta}}\rho^{\frac{\delta+\beta-1}{2}}\\
&+c\|b_n-b\|_{\infty,H_q^{-\beta}}\left(\|v\|^{(\rho)}_{0,H_p^{1+\delta}}+1\right)\rho^{\frac{\delta+\beta-1}{2}}\\
&+c\lambda\|v_n-v\|^{(\rho)}_{0,H_p^{1+\delta}}\rho^{-1}.
\end{aligned}
$$

The last bound is due to Lemma 11. Since $\|b_n\|_{\infty,H_q^{-\beta}}\to\|b\|_{\infty,H_q^{-\beta}}$ then there exists $n_0\in\mathbb{N}$ such that $\|b_n\|_{\infty,H_q^{-\beta}}\le 2\|b\|_{\infty,H_q^{-\beta}}$ for all $n\ge n_0$. Choose now ρ big enough in order to have

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

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

$$
\|v-v_n\|^{(\rho)}_{0,H_p^{1+\delta}}\le c\frac{\left(\|v\|^{(\rho)}_{0,H_p^{1+\delta}}+1\right)\rho^{\frac{\delta+\beta-1}{2}}}{1-c\left(\|b\|_{\infty,H_q^{-\beta}}\rho^{\frac{\delta+\beta-1}{2}}+\lambda\rho^{-1}\right)}\|b_n-b\|_{\infty,H_q^{-\beta}}
$$

which concludes the proof. □

3. The virtual solution

From now on, we fix λ and ρ big enough so that Theorem 13 and Lemma 21 hold true. As usual, the drift b is chosen according to Assumption 4.

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

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t, \tag{28}$$

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

$$X_t = x + \int_0^t b(s, X_s)\mathrm{d}s + W_t, \tag{29}$$

and the integral appearing on the right hand side needs to be defined. We aim to give a meaning to this equation, and in particular to the singular term $\int_0^t b(s, X_s)\mathrm{d}s$, by introducing a suitable notion of solution to the SDE (28).

Let $u(t, x)$ be a mild solution to the PDE (6) and X_t the process solution to (28). Formally applying the Itô formula to $u(t, X_t)$ 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 formally 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 motivates the following definition.

Definition 23. *A process $X := (X_t)_{t\geq 0}$ is called* virtual solution *to the SDE* (28) *if it satisfies the integral equation*

$$X_t = x + u(0, x) - u(t, X_t) + (\lambda+1)\int_0^t u(s, X_s)\mathrm{d}s + \int_0^t \nabla u(s, X_s)\mathrm{d}W_s + W_t, \tag{30}$$

for all $t \in [0, T]$, where u is the unique solution to the PDE (6).

3.2. **Existence and uniqueness of the virtual solution.** To find a virtual solution X to (28), let us introduce the transformation $\varphi(t, x) := x + u(t, x)$ and set $Y_t = \varphi(t, X_t)$. From (30) we obtain

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

Since the function $\varphi(t, \cdot)$ is invertible for all t, we can consider the SDE

$$Y_t = y_0 + (\lambda+1)\int_0^t u(s, \varphi^{-1}(s, Y_s))\mathrm{d}s + \int_0^t \nabla u(s, \varphi^{-1}(s, Y_s))\mathrm{d}W_s + W_t, \tag{31}$$

where $y_0 = x + u(0, x)$. If $Y := (Y_t)_{t\geq 0}$ is the solution of (31) then

$$X_t = \varphi^{-1}(t, Y_t)$$

will give us the virtual solution of the SDE with distributional drift (28).

Proposition 24. *There exists a unique weak solution Y to the SDE* (31)*.*

Proof. We know that $u, \nabla u$ and φ^{-1} are jointly continuous in time and space and u and φ^{-1} are Lipschitz continuous in space by Lemma 20 and Lemma 21. This implies that the drift of Y

$$\mu(t,y) := (\lambda+1)u(t,\varphi^{-1}(t,y))$$

is continuous and with linear growth and the diffusion coefficient

$$\sigma(t,y) := \nabla u(t,\varphi^{-1}(t,y)) + \mathrm{I} = \nabla\varphi(t,\varphi^{-1}(t,y))$$

is continuous, where I denotes the $(d\times d)$-identity matrix. Since by Lemma 19 the gradient of u is uniformly bounded we also have that σ is uniformly bounded. Moreover σ is uniformly non-degenerate since for all $x,\xi\in\mathbb{R}^d$ and $t\in[0,T]$

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

by (26). Thus Theorem 10.2.2 in [18] yields existence and uniqueness of a weak solution. □

Theorem 25. *There exists a unique in law virtual solution X to the SDE* (28) *given by $X_t = \varphi^{-1}(t,Y_t)$, where Y is the process given in Proposition 24.*

Proof. The SDE (31) and the transformation $\varphi(t,x) = x + u(t,x)$ imply that the unique weak solution Y yields a virtual solution X. Suppose that there exists another virtual solution $Z := (Z_t)_{t\geq 0}$ to (30). Then $\varphi(t,Z_t)$ is a solution to (31). Since equation (31) admits uniqueness in law, the law of Y coincides with the law of $\varphi(t,Z)$ and by the invertibility of φ 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 by Bass and Chen [2].

Proposition 26. *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* $L^\infty\left([0,T]; H^{-\beta}_{\tilde{q},q}\right)$*.*

The unique strong solutions to the equations

$$dX_t^n = dW_t + b_n(t,X_t^n)\,dt, \qquad X_0 = x_0 \tag{32}$$

converge in law to the virtual solution X of equation (28)*.*

Proof. **Step 1** (X^n are virtual solutions). Let u_n be the unique classical solution of equation (6) replacing b with b_n; u_n is (at least) of class $C^{1,2}\left([0,T]\times\mathbb{R}^d;\mathbb{R}^d\right)$. Let $\varphi_n(t,x) = x + u_n(t,x)$ so that again $\varphi_n \in C^{1,2}\left([0,T]\times\mathbb{R}^d;\mathbb{R}^d\right)$. Let X^n be the unique strong solutions of equations

(32) and let $Y_t^n = \varphi_n(t, X_t^n)$. By Itô formula, Y^n satisfies equation (30)(with b_n replacing b) and thus $X_t^n = \varphi_n^{-1}(t, Y_t^n)$ is also a virtual solution. Let us call $\widetilde{b}_n$ and $\widetilde{\sigma}_n$ the drift and diffusion coefficients of the equation satisfied by Y^n.

Step 2 (Uniformity of constants of $\widetilde{b}_n$ and $\widetilde{\sigma}_n$). From assumption (ii) it follows that the norm of b_n in $L^\infty\left([0,T]\,;H_{\tilde{q},q}^{-\beta}\right)$ converges to the norm of b as $n \to \infty$; hence the bound (24) does not depend on n, for n large enough. This allows us to choose λ independently of n (for n large enough). Moreover, u_n and u are Lipschitz with constant not depending on n, since $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla u_n(t,x)|$ and $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla u(t,x)|$ are bounded by a constant not depending on n.

Then, by Lemma 21 we know that φ_n^{-1} are Lipschitz (and have therefore linear growth) with the same constant $k = 2$. The same holds for φ^{-1}.

It follows that the vector fields $\widetilde{b}_n$ have linear growth with constants independent of n, and similarly that the vector fields $\widetilde{\sigma}_n$ are bounded above and below by uniform constants.

Step 3 (Tightness and convergence). The family of the laws of Y^n is tight in $C\left([0,T]\,;\mathbb{R}^d\right)$. Indeed, by the uniform linear growth of $\widetilde{b}_n$ and boundedness of $\widetilde{\sigma}_n$, $\sup_{t\in[0,T]} |Y_t^n|$ has all moments, independent of n. Then we have that

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

for some constant $C > 0$ independent of n. By Kolmogorov theorem, this implies the tightness of the laws of Y^n.

Since $u_n \circ \varphi_n^{-1} \to u \circ \varphi^{-1}$ and $\nabla u_n \circ \varphi_n^{-1} \to \nabla u \circ \varphi^{-1}$ pointwise, it is not difficult to show that every converging subsequence of Y^n converges in law to a solution of (30). Since (30) admits uniqueness in law, the full sequence Y^n converges in law to the unique solution Y of (31).

Step 4 (Back to X^n). The final step consists in showing that X^n converges to X in law. This follows by Skorohod theorem, which allows to reduce the convergence in law to an ucp convergence and from the fact that $\varphi_n^{-1} \to \varphi^{-1}$ pointwise. The proof is complete. □

Examples of b_n which verify (ii) in Proposition 26 are easily obtained by convolutions of b against a sequence of mollifiers converging to a Dirac measure.

Acknowledgements: The second and third named authors were 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] E. B. Davies. *Heat kernels and spectral theory*, volume 92 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.

[4] 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.

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

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

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

[8] 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.

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

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

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

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

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

[14] 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.

[15] 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.

[16] 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.

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

[18] 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.

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

[20] 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.