HAL
open science

Weak regularization by stochastic drift : result and counter example

Paul-Eric Chaudru de Raynal

To cite this version:

Paul-Eric Chaudru de Raynal. Weak regularization by stochastic drift : result and counter example. Discrete and Continuous Dynamical Systems - Series A, 2018, 38 (3), pp.1269-1291. 10.3934/dcds.2018052 . hal-01333206v2

HAL Id: hal-01333206

https://hal.science/hal-01333206v2

Submitted on 13 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WEAK REGULARIZATION BY STOCHASTIC DRIFT : RESULT AND COUNTER EXAMPLE

Paul-Eric Chaudru de Raynal

Univ. Savoie Mont Blanc, CNRS, LAMA
F-73000 Chambéry, France

Abstract. In this paper, weak uniqueness of hypoelliptic stochastic differential equation with Hölder drift is proved when the Hölder exponent is strictly greater than 1/3. This result then "extends" to a weak framework the previous works [4, 23, 10], where strong uniqueness was proved when the regularity index of the drift is strictly greater than 2/3. Part of the result is also shown to be almost sharp thanks to a counter example when the Hölder exponent of the degenerate component is just below 1/3.

The approach is based on martingale problem formulation of Stroock and Varadhan and so on smoothing properties of the associated PDE which is, in the current setting, degenerate.

This is a reprint version of an article published in Discrete & Continuous Dynamical Systems - A, 2018, 38 (3) : 1269-1291

1 Introduction

Let d be a positive integer and $\mathcal{M}_d(\mathbb{R})$ be the set of $d\times d$ matrices with real coefficients. For a given positive T, given measurable functions $F_1, F_2, \sigma : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^d\times\mathbb{R}^d\times\mathcal{M}_d(\mathbb{R})$ and $(B_t, t\geq 0)$ a standard d-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ we consider the following $\mathbb{R}^d\times\mathbb{R}^d$ system for any t in $[0,T]$:

$$\begin{cases} \mathrm{d}X_t^1 = F_1(t, X_t^1, X_t^2)\mathrm{d}t + \sigma(t, X_t^1, X_t^2)\mathrm{d}B_t, & X_0^1 = x_1, \\ \mathrm{d}X_t^2 = F_2(t, X_t^1, X_t^2)\mathrm{d}t, & X_0^2 = x_2, \end{cases} \tag{1.1}$$

where x_1 and x_2 belong to $\mathbb{R}^d$ and where the diffusion matrix $a := \sigma\sigma^*$ is[1] supposed to be uniformly elliptic.

In this work, we aim at proving that this system is well-posed (*i.e.* there exists a unique solution), in the weak sense, when the drift is singular. Indeed, in that case, uniqueness of the associated martingale problem from Stroock and Varadhan's theory [21] fails since the noise of the system degenerates. We nevertheless show that under a suitable Hölder assumption on the drift, Lipschitz condition on the diffusion matrix, and hypoellipticity condition on the system, weak well-posedness holds for (1.1). By suitable, we mean that there exists a threshold for the Hölder-continuity of the drift with respect to (w.r.t.) the degenerate argument. This Hölder-exponent is supposed to be strictly greater than 1/3. We also show that this threshold is almost sharp thanks to a counter-example when the Hölder

1991 *Mathematics Subject Classification.* Primary: 60H10, 34F05; Secondary: 60H30.

Key words and phrases. regularization by noise, martingale problem, Kolmogorov hypoelliptic PDEs, parametrix, gradient estimate.

[1]The notation "$*$" stands for the transpose.

exponent is strictly less than 1/3.

Mathematical background. It may be a real challenge to show well-posedness of a differential system with drift less than Lipschitz continuous (see [9] for a work in that direction). The Peano example is a very good illustration of this phenomenon: for any α in $(0,1)$ the equation

$$\mathrm{d}Y_t = \mathrm{sign}(Y_t)|Y_t|^\alpha \mathrm{d}t,\ Y_0 = 0, \quad t \in [0,T], \tag{1.2}$$

has an infinite number of solutions of the form $\pm c_\alpha (t-t^\star)^{1/(1-\alpha)} \mathbf{1}_{[t^\star;+\infty)},\ t^\star \in [0,T]$. Nevertheless, it has been shown that this equation is well-posed (in a strong and weak sense) as soon as it is infinitesimally perturbed by a Brownian motion. More precisely, the equation

$$\mathrm{d}Y_t = b(Y_t)\mathrm{d}t + \mathrm{d}B_t,\ Y_0 = 0, \quad t \in [0,T], \tag{1.3}$$

admits a unique strong solution (*i.e.* there exists an almost surely unique solution adapted to the filtration generated by the Brownian motion) as soon as the function $b : \mathbb{R}^d \ni x \mapsto b(x) \in \mathbb{R}^d$ is measurable and bounded. This phenomenon is known as *regularization by noise.*

Regularization by noise of systems with singular drift has been widely studied in the past few years. Since the pioneering one dimensional work of Zvonkin [26] and its generalization to the multidimensional setting by Veretenikov [22] (where stochastic system with bounded drift and additive noise are handled), several authors extended the result. Krylov and Röckner [17] showed that SDE with additive noise and $\mathbb{L}_p$ drift (where p depends on the dimension of the system) are also well-posed and Zhang [24] proved the case of multiplicative noise with uniformly elliptic and Sobolev diffusion matrix. More recently, Flandoli, Issoglio and Russo [11] studied the case of weak well-posedness of (1.3) for a distributional drift, *i.e.* in the Hölder space C^α where α is strictly greater than $-1/2$[2] and Delarue and Diel [5] for Hölder regularity strictly greater than $-2/3$ in the one dimensional case. This last result has been generalized in any dimension by Cannizzaro and Chouk [2]. Also Catellier and Gubinelli [3] considered systems perturbed by fractional Brownian motion. We refer to the notes of Flandoli [12] for a general account on this topics.

In our case the setting is a bit different since the noise added in the system acts only by means of random drift (*i.e.* the system degenerates). Indeed, taking $F_1 = 0,\ \sigma = I_d,\ F_2(t,x_1,x_2) = x_1 + f_2(x_2)$, the archetypal example of system (1.1) writes

$$\mathrm{d}X_t^2 = \big(B_t + f_2(X_t^2)\big)\mathrm{d}t, \qquad X_0^2 = x_2, \tag{1.4}$$

where the function f_2 is supposed to be only Hölder-continuous. Thus, the system can be seen as a classical ODE whose drift is perturbed by a Brownian motion: the perturbation is then of macroscopic type. We hence consider a *regularization by stochastic drift.*[3]

The first work in that direction is due to Chaudru de Raynal [4] where strong well-posedness of (1.1) is proved when the drift is Hölder continuous with Hölder exponent w.r.t. the degenerate argument strictly greater than 2/3 and where the system is also supposed to be hypoelliptic. Since then, several Authors have studied the strong well-posedness of (1.1) with different approaches and have obtained, with weaker conditions, the same kind of threshold: in [23], the Authors used an approach based

[2] see [14] for a definition of such a space

[3] Note that in our general case, we consider a non-linear perturbation of the drift of the second component by a diffusion process which also depends on the perturbed component.

on gradient estimates on the associated semi-group to show that the system is strongly well-posed when the drift satisfies a Hölder-Dini condition with Hölder exponent of 2/3 w.r.t. the degenerate component; in [10], the Authors used a PDE approach and obtained strong well-posedness as soon as the drift is weakly differentiable in the degenerate direction, with order of differentiation of 2/3.

Hence, many techniques have been used to study the strong well posedness of such a system and all of them end with this particular threshold of 2/3. Thus, this critical value seems to be not an artefact, but something which is deeper and related to the nature of the system and of the well posedness as well. When investigating the sharpness of such a threshold, we obtained a class of counter examples for the weak uniqueness as soon as the Hölder regularity of the drift (of the degenerate component) is less than 1/3. Actually, this counter-example in somehow more general since it allows to obtain thresholds for the weak well-posedness of a large class of perturbation of (1.2) (including our present degenerate setting). This then leads us to investigate what could be expected in that case. Using the Zvonkin theory together with a martingale problem approach, we succeeded in extending the previously known results to this setting and obtain almost sharp weak well posedness result for the drift of the second component.

Strategy of proof. Our strategy relies on the martingale problem approach of Stroock and Varadhan [21]. We indeed know that under our setting (coefficients with at most linear growth) the system (1.1) admits at least a weak solution. We then show that this solution is unique. To do so, we investigate the regularity of the (mild) solution of the associated PDE. Namely, denoting by $\mathrm{Tr}(a)$ the trace of the matrix a, "$\cdot$" the standard Euclidean inner product on $\mathbb{R}^d$ and $\mathcal{L}$ the generator of (1.1):

$$\mathcal{L} \;:=\; \frac{1}{2}\mathrm{Tr}(a(t,x_1,x_2)D^2_{x_1}) + [F_1(t,x_1,x_2)]\cdot[D_{x_1}] + [F_2(t,x_1,x_2)]\cdot[D_{x_2}]\,, \tag{1.5}$$

we exhibit a "good" theory for the PDE

$$(\partial_t + \mathcal{L})u = f \tag{1.6}$$

set on the cylinder $[0,T)\times\mathbb{R}^{2d}$ with terminal condition 0 at time T and where the function f belongs to a certain class of functions $\mathcal{F}$.

By "good", we mean that we can consider a sequence of classical solutions $(u^n)_{n\geq 0}$ and associated derivative in the non-degenerate direction $(D_{x_1}u^n)_{n\geq 0}$ along a sequence of mollified coefficients $(F_1^n, F_2^n, a^n)_{n\geq 0}$ that satisfy *a priori* estimates depending only on the regularity of (F_1, F_2, a). By using Arzelá-Ascoli Theorem, this allows to extract a converging subsequence to the mild solution of (1.6) on every compact subset of $[0,T]\times\mathbb{R}^{2d}$.

Hence, thanks to Itô's Formula, one can show that the quantity

$$\Big(u(t,X_t^1,X_t^2) - \int_0^t f(s,X_s^1,X_s^2)\mathrm{d}s\Big)_{0\leq t\leq T},$$

is a martingale. By letting the class of function $\mathcal{F}$ be sufficiently rich, this allows us to prove uniqueness of the marginals of the weak solution of (1.1) and then of the law itself.

Here, the crucial point is that the operator is not uniformly parabolic: the second order differentiation operator in $\mathcal{L}$ only acts in the first (and non-degenerate) direction "x_1". Therefore, we expect a loss of the regularization effect w.r.t. the degenerate component of (1.1). Nevertheless, we show

that the noise still regularizes, even in the degenerate direction, by means of the random drift: we can benefit from the hypoellipticity of the system.

The system (1.4) indeed relies on the so-called Kolmogorov example [16], which is also the archetypal example of hypoelliptic system without elliptic diffusion matrix. In our setting, the hypoellipticity assumption translates as a non-degeneracy assumption on the derivative of the drift function F_2 w.r.t. the first component. Together with the Hölder assumption, this can be seen as a weak Hörmander condition, in reference to the work of Hörmander [15] on degenerate operators of divergent form.

Let us emphasize that, from this Hörmander's framework point of view, the term of *regularization by stochastic drift* lies into the fact that one needs the vector field of the drift of the second component to be (uniformly) non degenerate w.r.t. the first spatial direction in order to make the family of Lie bracket associated with the vector fields of $\mathcal{L}$ spans the whole space, even the degenerate direction.

Our system appears as a non-linear generalization of Kolmogorov's example. Degenerate operators of this form have been studied by many authors see *e.g.* the works of Di Francesco and Polidoro [8], and Delarue and Menozzi [7]. We also emphasize that, in [19], Menozzi proved the weak well-posedness of a generalization of (1.1) with Lipschitz drift and Hölder diffusion matrix.

Nevertheless, to the best of our knowledge a "good" theory, in the sense mentioned above, for the PDE (1.6) has not been exhibited yet. We here prove the aforementioned estimates by using a first order parametrix (see [13]) expansion of the operator $\mathcal{L}$ defined by (1.5). This parametrix expansion is based on the knowledge of the related linearized and frozen version of (1.1) coming essentially from the previous work of Delarue and Menozzi [7].

Minimal setting to restore uniqueness. Obviously, all the aforementioned works, as well as this one, lead to the question of the minimal assumption that could be done on the drift in order to restore well-posedness. Having in mind that most of these works use a PDE approach, it seems clear that the assumption on the drift relies on the regularization properties of the semi-group generated by the solution. In comparison with the previous works, the threshold of $1/3$ can be seen as the price to pay to balance the degeneracy of the system: the smoothing effect of the semi-group associated to a degenerate Gaussian process is less efficient than the one of a non-degenerate Gaussian process. We prove that our assumptions are (almost) minimal by giving a counter-example in the case where the drift F_2 is Hölder continuous with Hölder exponent just below $1/3$.

Although this example concerns our degenerate case, we feel that the method could be adapted in order to obtain the optimal threshold (for the weak well posedness) in other settings. This is the reason why we wrote it in a general form. Let us briefly explain why and expose the heuristic rule behind our counter example.

It relies on the work of Delarue and Flandoli [6]. In this paper, the Peano example is investigated: namely, the system of interest is

$$\mathrm{d}Y_t = \mathrm{sign}(Y_t)|Y_t|^{\alpha}\mathrm{d}t + \epsilon \mathrm{d}B_t,\ Y_0 = 0, \quad \epsilon > 0,\ 0 < \alpha < 1. \tag{1.7}$$

The Authors studied the zero-noise limit of the system ($\epsilon \to 0$) pathwise. When doing so, they put in evidence the following crucial phenomenon: in small time there is a competition between the irregularity of the drift and the fluctuations of the noise. The fluctuations of the noise allow the

solution to leave the singularity while the irregularity of the drift (possibly) captures the solution in the singularity. Thus, the more singular the drift is, the more irregular the noise has to be.

This competition can be made explicit. In order to regularize the equation, the noise has to dominate the system in small time. This means that there must exists a time $0 < t_\epsilon < 1$ such that, below this instant, the noise dominates the system and pushes the solution far enough from the singularity, while above, the drift dominates the system and constrains the solution to fluctuate around one of the extreme solution of the deterministic Peano equation. A good way to see how the instant t_ϵ looks like is to compare the fluctuations of the extreme solution ($\pm t^{1/(1-\alpha)}$) with the fluctuations of the noise. Denoting by γ the order of the fluctuations of the noise this leads to the equation

$$\epsilon t_\epsilon^\gamma = t_\epsilon^{1/(1-\alpha)},$$

which gives $t_\epsilon = \epsilon^{(1-\alpha)/(1-\gamma(1-\alpha))}$ and leads to the condition:

$$\alpha > 1 - 1/\gamma. \tag{1.8}$$

The counter example, which also especially compares the fluctuations of the noise with the extreme solution, leads to the same threshold and says that weak uniqueness fails below this threshold.

Obviously, cases where $\alpha < 0$ have to be considered carefully. But if we formally consider the case of a Brownian perturbation, we get $\gamma = 1/2$ and so $\alpha > -1$, which is the sharp threshold exhibited in the recent work of Beck, Flandoli, Gubinelli and Maurelli [1].

In our setting, as suggested by the example (1.4), the noise added in (4.1) can be seen as the integral of a Brownian path, which gives $\gamma = 3/2$. We deduce from equation (1.8) that the threshold for the Hölder-regularity of the drift is $1/3$. We finally emphasize that this heuristic rule gives another (pathwise) interpretation for our threshold in comparison with the one obtained in the non-degenerate cases. Since the noise added in our system degenerates, the fluctuations (which are typically of order $3/2$) are not strong enough to push the solution far enough from the singularity when the drift is too singular (say less than $C^{1/3}$).

Organization of this paper. This paper is organized as follows. In Section 2, we give our main results: weak existence and uniqueness holds for (1.1) and almost sharpness of part of this result. Smoothing properties of PDE (1.6) are given in Section 3 as well as the proof of our main result. Then, we discuss on the almost sharpness of the result in Section 4 by giving a counter-example. Finally, the regularization properties of the PDE (1.6) are proved in Section 5.

2 Notations, assumptions and main results

Notations. In order to simplify the notations, we adopt the following convention: x, y, z, ξ, *etc.* denote the $2d-$dimensional real variables $(x_1, x_2), (y_1, y_2), (z_1, z_2), (\xi_1, \xi_2)$, *etc.*. We denote by $g(t, x) = g(t, x_1, x_2)$ any function g from $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^N$, $N \in \mathbb{N}$ evaluated at point (t, x_1, x_2). Below we sometimes write $X_t = (X_t^1, X_t^2)$ and, when necessary, we write $(X_s^{t,x})_{t \le s \le T}$ for the process defined by (1.1) which starts from x at time t, *i.e.* such that $X_t^{t,x} = x$.

We denote by $\mathcal{M}_d(\mathbb{R})$ the set of real $d \times d$ matrices, by "Id" the identity matrix of $\mathcal{M}_d(\mathbb{R})$ we denote by $\mathcal{B}$ the $2d \times d$ matrix: $\mathcal{B} = (\mathrm{Id}, 0_{\mathbb{R}^d \times \mathbb{R}^d})^*$. We write $\mathrm{GL}_d(\mathbb{R})$ the set of $d \times d$ invertible matrices with real coefficients. We recall that a denotes the square of the diffusion matrix σ, $a := \sigma\sigma^*$. The canonical Euclidean inner product on $\mathbb{R}^d$ is denoted by "$\cdot$".

Subsequently, we denote by c, C, c', C', c'' *etc.* a positive constant, depending only on known parameters in **(H)**, given just below, that may change from line to line and from an equation to another.

For any function from $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$, we use the notation D to denote the total space derivative, we denote by D_1 (resp. D_2) the derivative with respect to the first (resp. second) d-dimensional space component. In the same spirit, the notation D_z means the derivative w.r.t the variable z. Hence, for all integers n, D_z^n is the nth derivative w.r.t z and for all integers m the $n\times m$ cross differentiations w.r.t z, y are denoted by $D_z^n D_y^m$. Furthermore, the partial derivative ∂/∂_t is denoted by ∂_t.

Assumptions (H). We say that assumptions **(H)** hold if the following assumptions are satisfied.

(H1) Regularity of the coefficients: there exist $0<\beta_i^j\leq 1$, $1\leq i,j\leq 2$ and three positive constants C_1, C_2, C_σ such that for all t in $[0,T]$ and all (x_1,x_2) and (y_1,y_2) in $\mathbb{R}^d\times\mathbb{R}^d$

$$\begin{aligned}
&|F_1(t,x_1,x_2)-F_1(t,y_1,y_2)|\leq C_1(|x_1-y_1|^{\beta_1^1}+|x_2-y_2|^{\beta_1^2})\\
&|F_2(t,x_1,x_2)-F_2(t,y_1,y_2)|\leq C_2(|x_1-y_1|^{\beta_2^1}+|x_2-y_2|^{\beta_2^2})\\
&|\sigma(t,x_1,x_2)-\sigma(t,y_1,y_2)|\leq C_\sigma(|x_1-y_1|+|x_2-y_2|).
\end{aligned}$$

Moreover, the coefficients are supposed to be continuous w.r.t the time and the exponents β_i^2, $i=1,2$ are supposed to be strictly greater than $1/3$. Note that $\beta_2^1=1$ but we keep it for notational.

(H2) Uniform ellipticity of $\sigma\sigma^*$: The function $\sigma\sigma^*$ satisfies the uniform ellipticity hypothesis:

$$\exists\Lambda>1,\ \forall\zeta\in\mathbb{R}^d,\quad \Lambda^{-1}|\zeta|^2\leq[\sigma\sigma^*(t,x_1,x_2)\zeta]\cdot\zeta\leq\Lambda|\zeta|^2,$$

for all $(t,x_1,x_2)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$.

(H3-a) Differentiability and regularity of $x_1\mapsto F_2(.,x_1,.)$: For all $(t,x_2)\in[0,T]\times\mathbb{R}^d$, the function $F_2(t,.,x_2):\ x_1\mapsto F_2(t,x_1,x_2)$ is continuously differentiable and there exist $0<\eta<1$ and a positive constant $\bar{C}_2$ such that, for all (t,x_2) in $[0,T]\times\mathbb{R}^d$ and x_1,y_1 in $\mathbb{R}^d$

$$|D_1F_2(t,x_1,x_2)-D_1F_2(t,y_1,x_2)|\leq\bar{C}_2|x_1-y_1|^\eta.$$

(H3-b) Non-degeneracy of $(D_1F_2)(D_1F_2)^*$: There exists a closed convex subset $\mathcal{E}\subset\mathrm{GL}_d(\mathbb{R})$ (the set of $d\times d$ invertible matrices with real coefficients) such that for all t in $[0,T]$ and (x_1,x_2) in $\mathbb{R}^{2d}$ the matrix $D_1F_2(t,x_1,x_2)$ belongs to $\mathcal{E}$. We emphasize that this implies that

$$\exists\bar{\Lambda}>1,\ \forall\zeta\in\mathbb{R}^d,\quad \bar{\Lambda}^{-1}|\zeta|^2\leq[(D_1F_2)(D_1F_2)^*(t,x_1,x_2)\zeta]\cdot\zeta\leq\bar{\Lambda}|\zeta|^2,$$

for all $(t,x_1,x_2)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$.

Here are the main results of this paper:

Theorem 2.1. *Suppose that assumption **(H)** hold and in addition that*

$$\text{for i}=1,2,\quad \beta_{\mathrm{i}}^2>1/3. \tag{2.1}$$

Then, there exists a unique weak solution to (1.1).

Theorem 2.2. *There exists a drift (F_1,F_2) satisfying **(H)** and such that $\beta_2^2<1/3$ for which uniqueness fails.*

Remark. *Let us tell more about the regularity assumed on F_2 w.r.t. the non degenerate variable x_1: the uniform differentiability assumption together with the non-degeneracy of the derivative **(H3-b)** are crucial in that framework because they guarantee the hypoellipticity of the solution and so it allows the noise to propagate through the second component.*

We emphasize that one may object, as it is proved in the work [25] and [10], that the optimal condition on the drift of the non degenerate component F_1 that ensures weak uniqueness is an appropriate integrability condition. This is, in our opinion, true. But the main thing is that in those two works the system considered is linear in the second component: $F_2(t, x_1, x_2) = x_1$. This permits, in particular, to use Girssanov Theorem and to reduce the system to a degenerate Ornstein Uhlenbeck process. Here, the dependence of F_2 upon x_2 as well as the non linearity break down the arguments and as far as we can see the generalization in our case is non trivial.

The way we apply our parametrix method to investigate the smoothing properties of the semi-group of (1.1) is somehow global, in the sense that it does not allow to differentiate the different components (F_1, F_2) of the drift of (1.1). We then have to ask same kind of assumptions on the drift function F_1 and the linearized drift of F_2 (see system (5.2) below). This is the reason why both β_1^1 and η are strictly positive and β_1^2 and β_2^2 are supposed to be strictly greater than $1/3$ although our counter example only holds for the Hölder exponent of the drift of the second component in the degenerate variable. Finally, let us mention that in [20], the Author proved $\mathbb{L}_p$ estimate for the the semi-group of (1.1) by using parametrix approach when the drift functions are supposed to be Lipschitz in space. This could be a first step to extend our result in further work.

Finally, let us comment the almost sharpness of this result namely Theorem 2.2 which will be proved in Section 4. This exponent of $1/3$ can be immediately deduced from the heuristic rule previously given. Concerning the critical value for the Hölder exponent ($1/3$) our feeling in that weak uniqueness holds but we were not able to prove this result especially because our approach is not adapted to reach this critical case.

3 PDE result and proof of Theorem 2.1

Let us first begin by giving the smoothing properties of the PDE (1.6). Let $(F_1^n, F_2^n, a^n)_{n\geq 0}$ be a sequence of mollified coefficients (say infinitely differentiable with bounded derivatives of all orders greater than 1) satisfying **(H)** uniformly in n that converges to (F_1, F_2, a) uniformly on $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$ (such an example of coefficients can be found in [4]). Let us denote by $(\mathcal{L}^n)_{n\geq 0}$ the associated sequence of regularized versions of the operator $\mathcal{L}$ defined by (1.5). We have the following result.

Theorem 3.1. *Let $\mathcal{F}$ be the set of continuous in time and 1-Lipschitz in space functions $f : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^d$. For each n, the PDE (1.6) with $\mathcal{L}^n$ instead of $\mathcal{L}$ admits a unique classical solution u^n.*

*Moreover, there exist a positive $\mathcal{T}_{3.1}$, a positive $\delta_{3.1}$ and a positive ν, depending on known parameters in **(H)** only, such that for all T less than $\mathcal{T}_{3.1}$ the solution of the regularized PDE (1.6) with source term f satisfies:*

$$||D_2u^n||_\infty + ||D_1u^n||_\infty + ||D_1^2u^n||_\infty + ||D_1u^n||_{\infty,\infty,\nu} \leq CT^{\delta_{3.1}}, \tag{3.1}$$

where

$$||D_1u^n||_{\infty,\infty,\nu} = \sup_{t\in[0,T],x_1\in\mathbb{R}^d}\ \sup_{x_2\neq z_2} \frac{|D_1u^n(t,x_1,x_2) - D_1u^n(t,x_1,z_2)|}{|x_2-z_2| + |x_2-z_2|^{\beta_1^2} + |x_2-z_2|^{\beta_2^2} + |x_2-z_2|^{\nu}}. \tag{3.2}$$

Moreover, each classical solution u^n is uniformly bounded on every compact subset of $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$.

Let us just notice that the constant C above depends also on the class of functions $\mathcal{F}$ though the Lipschitz norm of functions belonging to this class.

Proof. The proof of this result is postponed to Section 5. □

We are now in position to prove uniqueness of the martingale problem associated to (1.1). Under our assumptions, it is clear from Theorem 6.1.7 of [21] that the system (1.1) has at least one weak solution (the linear growth assumption assumed here is not a problem to do so).

Let $f:[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ be some continuous in time and 1-Lipschitz in space function, let u^n be the classical solution of the regularized version of the PDE (1.6) with source term f and let (X^1,X^2) be a weak solution of (1.1) starting from x at time 0. Let now suppose that T is less than $\mathcal{T}_{3.1}$ given in Theorem 3.1. Applying Itô's Formula on $u^n(t,X^1_t,X^2_t)$ we obtain that

$$\begin{aligned}
u^n(t,X^1_t,X^2_t) &= u^n(0,x_1,x_2)+\int_0^t(\partial_t+\mathcal{L})u^n(s,X^1_s,X^2_s)\mathrm{d}s\\
&\quad+\int_0^t D_xu^n(s,X^1_s,X^2_s)\mathcal{B}\sigma(s,X^1_s,X^2_s)\mathrm{d}B_s\\
&= u^n(0,x_1,x_2)+\int_0^t(\partial_t+\mathcal{L}^n)u^n(s,X^1_s,X^2_s)\mathrm{d}s\\
&\quad+\int_0^t(\mathcal{L}-\mathcal{L}^n)u^n(s,X^1_s,X^2_s)\mathrm{d}s\\
&\qquad+\int_0^t D_xu^n(s,X^1_s,X^2_s)\mathcal{B}\sigma(s,X^1_s,X^2_s)\mathrm{d}B_s\\
&= u^n(0,x_1,x_2)+\int_0^t f(s,X^1_s,X^2_s)\mathrm{d}s+\int_0^t(\mathcal{L}-\mathcal{L}^n)u^n(s,X^1_s,X^2_s)\mathrm{d}s\\
&\qquad+\int_0^t D_xu^n(s,X^1_s,X^2_s)\mathcal{B}\sigma(s,X^1_s,X^2_s)\mathrm{d}B_s,
\end{aligned}$$

since u^n is the solution of the regularized version of (1.6) and where we recall that $\mathcal{B}$ is the $2d\times d$ matrix: $\mathcal{B}=(\mathrm{Id},0_{\mathbb{R}^d\times\mathbb{R}^d})^*$.

Thanks to Theorem 3.1 and Arzelà -Ascoli Theorem, we know that we can extract a subsequence of $(u^n)_{n\geq0}$ and $(D_1u^n)_{n\geq0}$ that converge respectively to a function u and D_1u uniformly on compact subset of $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$. Thus, together with the uniform convergence of the regularized coefficients, we can deduce that

$$\left(u(t,X^1_t,X^2_t)-\int_0^t f(s,X^1_s,X^2_s)\mathrm{d}s-u(0,x_1,x_2)\right)_{0\leq t\leq T}, \tag{3.3}$$

is a $\mathbb{P}$-martingale by letting the regularization procedure tend to the infinity.

Let us now come back to the canonical space, and let $\mathbb{P}$ and $\tilde{\mathbb{P}}$ be two solutions of the martingale problem associated to (1.1) with initial condition (x_1,x_2) in $\mathbb{R}^d\times\mathbb{R}^d$. Thus, for all continuous in time

and Lipschitz in space functions $f : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ we have from (3.3) (recall that by definition we have that $u(T,\cdot,\cdot) = 0$),

$$u(0,x_1,x_2) = \mathbb{E}_{\mathbb{P}}\left[\int_0^T f(s,X_s^1,X_s^2)\mathrm{d}s\right] = \mathbb{E}_{\tilde{\mathbb{P}}}\left[\int_0^T f(s,X_s^1,X_s^2)\mathrm{d}s\right],$$

so that the marginal law of the canonical process are the same under $\mathbb{P}$ and $\tilde{\mathbb{P}}$. We extend the result on $\mathbb{R}^+$ thanks to regular conditional probabilities, see [21] Chapter 6.2. Uniqueness then follows from Corollary 6.2.4 of [21].

4 Counter example

We here prove Theorem 2.2. As we said, we feel that this counter example does not reduce to our current setting. Hence, we wrote it in a general form in order to adapt it to different cases. Let $\mathcal{W}$ be a random process with continuous path satisfying, in law,

$$(\mathcal{W}_t,\ t \geq 0) = (-\mathcal{W}_t,\ t \geq 0)\ ,\ \forall t \geq 0\, t^\gamma \mathcal{W}_1 = \mathcal{W}_t,$$

and $\mathbb{E}|\mathcal{W}_1| < +\infty$ for some given $\gamma > 0$. Let $\alpha < 1$ and $c_\alpha := (1-\alpha)^{1/(1-\alpha)}$. We suppose that $\mathcal{W}$ and α are such that there exists a weak solution of

$$X_t = x + \int_0^t \mathrm{sign}(\mathrm{X_s})|\mathrm{X_s}|^\alpha \mathrm{ds} + \mathcal{W}_\mathrm{t}, \tag{4.1}$$

for any $x \geq 0$ that satisfies Kolmogorov's criterion. Given $0 < \beta < 1$ we define for any continuous path Y from $\mathbb{R}^+$ to $\mathbb{R}$ the variable $\tau(Y)$ as

$$\tau(Y) = \inf\{t \geq 0\ :\ Y_t \leq (1-\beta)c_\alpha t^{1/(1-\alpha)}\}.$$

We now have the following Lemma whose proof is postponed to the end of this section:

Lemma 4.1. *Let X be a weak solution of (4.1) starting from some $x > 0$ and suppose that $\alpha < 1 - 1/\gamma$. Then, there exists a positive ρ, depending on α, β, γ and $\mathbb{E}|\mathcal{W}_1|$ only such that*

$$\mathbb{P}_x(\tau(X) \geq \rho) \geq 3/4. \tag{4.2}$$

We are now in position to give our counter-example. Note that if $(X,\mathcal{W})$ is a weak solution of (4.1) with the initial condition $x = 0$, then, $(-X,-\mathcal{W})$ is also a weak solution of (4.1). So that, if uniqueness in law holds X and $-X$ have the same law.

Let us consider a weak solution X^n of (4.1) starting from $1/n$, n being a positive integer. Since each X^n satisfies Kolmogorov's criterion, the sequence of law $(\mathbb{P}_{1/n})_{n\geq 0}$ of X^n is tight, so that we can extract a converging subsequence $(\mathbb{P}_{1/n_k})_{k\geq 0}$ to $\mathbb{P}_0$, the law of the weak solution X of (4.1) starting from 0. Since the bound in (4.2) does not depend on the initial condition we get that

$$\mathbb{P}_0(\tau(X) \geq \rho) \geq 3/4,$$

and, thanks to uniqueness in law

$$\mathbb{P}_0(\tau(-X) \geq \rho) \geq 3/4,$$

which is a contradiction. Choosing $\mathcal{W} = \int_0^\cdot W_s \mathrm{d}s$, so that $\gamma = 3/2$, we get that weak uniqueness fails as soon as

$$\alpha < 1 - 1/\gamma = 1/3.$$

We now prove Lemma 4.1 which allows to understand how the threshold above, exhibited in the introduction, also appears in our counter-example.

Proof of Lemma 4.1. Let X be a weak solution of (4.1) starting from $x > 0$. Since it has continuous path, we have almost surely that $\tau(X) > 0$. Then, note that on $[0, \tau(X)]$ we have:

$$\begin{aligned} X_t &= x + \int_0^t \mathrm{sign(X_s)|X_s|^\alpha ds} + \mathcal{W}_\mathrm{t} \\ &\geq (1-\beta)^\alpha c_\alpha t^{1/(1-\alpha)} + \mathcal{W}_t. \end{aligned}$$

Hence, choosing η such that $(1-\eta) = [(1-\beta)^\alpha + (1-\beta)]/2$ we get that:

$$X_t \geq (1-\eta)c_\alpha t^{1/(1-\alpha)} + (\beta-\eta)c_\alpha t^{1/(1-\alpha)} + \mathcal{W}_t,$$

for all t in $[0, \tau(X)]$.

Now let ρ be a positive number, set $\tilde{c}_\alpha = (\beta-\eta)c_\alpha$ and

$$A = \left\{ \tilde{c}_\alpha t^{1/(1-\alpha)} + \mathcal{W}_t > 0 \text{ for all } t \text{ in } (0, \rho] \right\}.$$

Note that on A we have

$$X_t \geq (1-\eta)c_\alpha t^{1/(1-\alpha)} \geq (1-\beta)c_\alpha t^{1/(1-\alpha)}$$

for all t in $[0, \rho]$. But this is compatible only with the event $\{\tau(X) \geq \rho\}$ so that $A \subset \{\tau(X) \geq \rho\}$. Hence

$$\mathbb{P}(\tau(X) \geq \rho) \geq \mathbb{P}(A). \tag{4.3}$$

We are now going to bound from below the probability of the event A. We have

$$\begin{aligned} \mathbb{P}(A^c) &= \mathbb{P}\left(\exists t \in (0,\rho] \ : \ \tilde{c}_\alpha t^{1/(1-\alpha)} + \mathcal{W}_t \leq 0\right) \\ &\leq \mathbb{P}\left(\exists t \in (0,\rho] \ : \ |\mathcal{W}_t| \geq \tilde{c}_\alpha t^{1/(1-\alpha)}\right) \\ &= \mathbb{P}\left(\exists t \in (0,1] \ : \ (\rho t)^\gamma |\mathcal{W}_1| \geq \tilde{c}_\alpha (\rho t)^{1/(1-\alpha)}\right) \\ &= \mathbb{P}\left(\exists t \in (0,1] \ : \ |\mathcal{W}_1| \geq \tilde{c}_\alpha (\rho t)^{-\delta}\right), \end{aligned}$$

where $\delta = \gamma - 1/(1-\alpha)$. Since $\alpha < 1 - 1/\gamma$, we get that $\delta > 0$ and we obtain from the previous computations that

$$\mathbb{P}(A^c) \leq \mathbb{P}\left(|\mathcal{W}_1| \geq \tilde{c}_\alpha \rho^{-\delta}\right) \leq \mathbb{E}|\mathcal{W}_1|\tilde{c}_\alpha^{-1}\rho^\delta,$$

from Markov inequality. Thus

$$\mathbb{P}(\tau(X) \geq \rho) \geq \mathbb{P}(A) \geq 1 - \mathbb{E}|\mathcal{W}_1|\tilde{c}_\alpha^{-1}\rho^\delta,$$

so that there exists a positive ρ such that

$$\mathbb{P}(\tau(X) \geq \rho) \geq 3/4.$$

□

5 Smoothing properties of the PDE

This section is dedicated to the proof of Theorem 3.1. This proof is in the same spirit and uses the same kind of tools as the one introduced in the previous work [4]. We nevertheless emphasize that our analysis is here quite different: although the tools are the same, the objective differs from the previously mentioned work. Firstly, our PDE does not have the same source term; secondly the controls we want to obtain on the solution are weaker; thirdly our regularity assumptions on the coefficients of the operator $\mathcal{L}$ in (1.6) are weaker so that we have to be careful.

Our main strategy for proving Theorem 3.1 rests upon *parametrix* approach (see [18],[13]). This perturbation approach consists in expanding the operator $\mathcal{L}$ around a well chosen proxy, denoted by $\tilde{\mathcal{L}}$, which enjoys suitable properties and can be handled. Hence we rewrite our PDE (1.6) as:

$$(\partial_t + \mathcal{L})u = f \Leftrightarrow (\partial_t + \tilde{\mathcal{L}})u = f + (\tilde{\mathcal{L}} - \mathcal{L})u,$$

and we call parametrix kernel the second term on the right hand side of the equation above (*i.e.* the approximation error). By doing so, we obtain a representation of our PDE solution in terms of a time space convolution of the source term and the parametrix kernel against the fundamental solution of our proxy. This is the reason why the choice of this proxy as well as its property are crucial in our analysis.

The idea is to obtain a Gaussian approximation which means that we aim at obtaining a proxy $\tilde{\mathcal{L}}$ which is the generator of a Gaussian process. This Gaussian process has to be as closed as possible of our original process. Having in mind the Kolmogorov Example, this means that the dependence w.r.t. the noise of each component has to be of linear form (see Section 2 of [4] for more details or [7] for a more general account on this topic).

We thus end with a proxy $\tilde{\mathcal{L}}$ which is an operator whose coefficients are the linearized version of the one of the operator $\mathcal{L}$. As already emphasized, it is the generator of a degenerate Gaussian process relying on Kolmogorov Example. Namely this process has a transition density whose covariance matrix is homogeneous to the covariance matrix of a Brownian motion and its time integral. This proxy is not new, it has been introduced by Delarue and Menozzi in [7] and then successfully used by Menozzi in [19, 20] to prove weak well posedness of a generalization of (1.1) when the drifts coefficients are Lispchitz continuous and the diffusion matrix σ is respectively Hölder continuous and solely continuous.

Since our Gaussian proxy (process, semi-group, generator, transition density) plays a central role in our analysis, we dedicated the first part of this section (Subsection 5.1) to the study and the properties of such a system. These tools have already been introduced in [4], we then collect from this work all the necessary ingredients and let the reader check the proof in the aforementioned paper. Then, we prove Theorem 3.1 in Subsection 5.2.

Let us notice that Theorem 3.1 concerns the solution of the regularized version of (1.6). Thus, for the sake of clarity, we forget the superscript n that follows from the regularization procedure and we suppose throughout this section that the coefficients F_1, F_2 and $a := \sigma\sigma^*$ are smooth (say infinitely differentiable with bounded derivatives of all orders greater than one). We then specify the dependence of the constants when necessary.

5.1 The frozen system

We here defined our Gaussian proxy. As already underlined, this process is a linear (w.r.t. the noise) version of (1.1). Because of our degenerate setting we have to be careful when doing this linearization: it has to be done around the forward flow associated to the deterministic version of (1.1) (*i.e.* when $\sigma = 0$ therein). Namely given any frozen point (τ, ξ) in $[0,T] \times \mathbb{R}^{2d}$, we consider the following system on $[\tau, T]$

$$\left\{ \begin{array}{ll} \dfrac{\mathrm{d}}{\mathrm{d}s}\theta^1_{\tau,s}(\xi) = F_1(s, \theta_{\tau,s}(\xi)), & \theta^1_{\tau,\tau}(\xi) = \xi_1, \\ \dfrac{\mathrm{d}}{\mathrm{d}s}\theta^2_{\tau,s}(\xi) = F_2(s, \theta_{\tau,s}(\xi)), & \theta^2_{\tau,\tau}(\xi) = \xi_2, \end{array} \right. \tag{5.1}$$

which is well posed under our regularized framework and we extend the definition of its solution on $[0,\tau)$ by assuming that for all $(v > r)$ in $[0,T]^2$, for all ξ in $\mathbb{R}^{2d}$, $\theta_{v,r}(\xi) = 0$. Given the solution $(\theta_{\tau,s}(\xi))_{s \le T}$ of this system, we define the linearized and frozen version of (1.1):

$$\left\{ \begin{array}{l} \mathrm{d}\tilde{X}^{1,t,x}_s = F_1(s, \theta_{\tau,s}(\xi))\mathrm{d}s + \sigma(s, \theta_{\tau,s}(\xi))\mathrm{d}W_s \\ \mathrm{d}\tilde{X}^{2,t,x}_s = \Big[F_2(s, \theta_{\tau,s}(\xi)) + D_1F_2(s, \theta_{\tau,s}(\xi))(\tilde{X}^{1,t,x}_s - \theta^1_{\tau,s}(\xi))\Big]\,\mathrm{d}s \end{array} \right. \tag{5.2}$$

for all s in $(t,T]$, any t in $[0,T]$, and for any initial condition x in $\mathbb{R}^{2d}$ at time t. We then have the following proposition:

Proposition 5.1. [4, Proposition 3.1] *Under our assumptions:*

(i) There exists a unique (strong) solution of (5.2) with mean

$$(m^{\tau,\xi}_{t,s})_{t \le s \le T} = (m^{1,\tau,\xi}_{t,s}, m^{2,\tau,\xi}_{t,s})_{t \le s \le T},$$

where

$$\begin{aligned} m^{1,\tau,\xi}_{t,s}(x) &= x_1 + \int_t^s F_1(r, \theta_{\tau,r}(\xi))\mathrm{d}r, \\ m^{2,\tau,\xi}_{t,s}(x) &= x_2 + \int_t^s \Big[F_2(r, \theta_{\tau,r}(\xi)) + D_1F_2(r, \theta_{\tau,r}(\xi))(x_1 - \theta^1_{\tau,r}(\xi)) \\ &\qquad + D_1F_2(r, \theta_{\tau,r}(\xi)) \int_t^r F_1(v, \theta_{\tau,v}(\xi))\mathrm{d}v\Big]\mathrm{d}r, \end{aligned} \tag{5.3}$$

and uniformly non-degenerate covariance matrix $(\tilde{\Sigma}_{t,s})_{t \le s \le T}$:

$$\tilde{\Sigma}_{t,s} = \begin{pmatrix} \int_t^s \sigma\sigma^*(r, \theta_{\tau,r}(\xi))\mathrm{d}r & \int_t^s R_{r,s}(\tau,\xi)\sigma\sigma^*(r, \theta_{\tau,r}(\xi))\mathrm{d}r \\ \int_t^s \sigma\sigma^*(r, \theta_{\tau,r}(\xi))R^*_{r,s}(\tau,\xi)\mathrm{d}r & \int_t^s R_{t,r}(\tau,\xi)\sigma\sigma^*(r, \theta_{\tau,r}(\xi))R^*_{t,r}(\tau,\xi)\mathrm{d}r \end{pmatrix}, \tag{5.4}$$

where:

$$R_{t,r}(\tau,\xi) = \left[\int_t^r D_1F_2(v, \theta_{\tau,v}(\xi))\mathrm{d}v\right], \quad t \le r \le s \le T.$$

(ii) This solution is a Gaussian process with transition density:

$$\begin{aligned} &\tilde{q}(t, x_1, x_2; s, y_1, y_2) \\ &= \frac{3^{d/2}}{(2\pi)^{d/2}}(\det[\tilde{\Sigma}_{t,s}])^{-1/2} \exp\left(-\frac{1}{2}|\tilde{\Sigma}^{-1/2}_{t,s}(y_1 - m^{1,\tau,\xi}_{t,s}(x), y_2 - m^{2,\tau,\xi}_{t,s}(x))^*|^2\right), \end{aligned} \tag{5.5}$$

for all s in $(t,T]$.

(iii) This transition density $\tilde{q}$ is the fundamental solution of the PDE driven by $\tilde{\mathcal{L}}^{\tau,\xi}$ and given by:

$$\begin{aligned}\tilde{\mathcal{L}}^{\tau,\xi} &:= \frac{1}{2}Tr\left[a(t,\theta_{\tau,t}(\xi))D^2_{x_1}\right] + \left[F_1(t,\theta_{\tau,t}(\xi))\right]\cdot D_{x_1} \\ &\quad + \left[F_2(t,\theta_{\tau,t}(\xi)) + D_1F_2(t,\theta_{\tau,t}(\xi))\left(x_1 - \theta^1_{\tau,t}(\xi)\right)\right]\cdot D_{x_2}. \end{aligned} \tag{5.6}$$

*(iv) There exist two positive constants c and C, depending only on known parameters in **(H)**, such that*

$$\tilde{q}(t,x_1,x_2;s,y_1,y_2) \le C\hat{q}_c(t,x_1,x_2;s,y_1,y_2), \tag{5.7}$$

where

$$\hat{q}_c(t,x_1,x_2;s,y_1,y_2) = \frac{c}{(s-t)^{2d}}\exp\left(-c\left(\frac{\left|y_1 - m^{1,\tau,\xi}_{t,s}(x)\right|^2}{s-t} + \frac{\left|y_2 - m^{2,\tau,\xi}_{t,s}(x)\right|^2}{(s-t)^3}\right)\right),$$

and

$$\begin{aligned}&\left|D^{N^{x_1}}_{x_1}D^{N^{x_2}}_{x_2}D^{N^{y_1}}_{y_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right| \\ \le\ & C(s-t)^{-[3N^{x_2}+N^{x_1}+N^{y_1}]/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2),\end{aligned} \tag{5.8}$$

for all s in $(t,T]$ and any integers $N^{x_1}, N^{x_2}, N^{y_1}$ less than 2.

We now introduce:

Definition 5.2. *In order to prove the estimate (3.1) of Theorem 3.1, let us define, for all $x \in \mathbb{R}^{2d}$, $y \in \mathbb{R}^{2d}$: and any $\gamma \in [0,1]$:*

$$\Delta^2_\gamma(x-y) = |x_2 - y_2| + |x_2 - y_2|^{\beta^2_1} + |x_2 - y_2|^{\beta^2_2} + |x_2 - y_2|^\gamma \tag{5.9}$$

We set for all measurable function $\varphi : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$, for all $t < s$ in $[0,T]^2$ and ξ and x in $\mathbb{R}^{2d}$:

$$\left[\tilde{P}^\xi_{t,s}\varphi\right](s,x) = \int_{\mathbb{R}^{2d}}\varphi(s,y)\tilde{q}(t,x;s,y)\mathrm{d}y, \tag{5.10}$$

and

$$\left[\hat{P}^\xi_{t,s}\varphi\right](s,x) = \int_{\mathbb{R}^{2d}}\varphi(s,y)\hat{q}_c(t,x;s,y)\mathrm{d}y. \tag{5.11}$$

Finally, we have the following Proposition regarding the smoothing properties of $\tilde{P}$ defined above:

Proposition 5.3. *Under our assumptions there exist three positive constants C, C' and C'', depending on known parameters in **(H)** only such that for all $t < s$ in $[0,T]^2$, ξ and x in $\mathbb{R}^{2d}$, ζ in $\mathbb{R}^{2d}$ and all measurable function $\varphi : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$:*

(i) Gaussian decay of the derivatives:

$$\left|D_{x_i}\left[\tilde{P}^\xi_{t,s}\varphi\right](s,x)\right| \le C'(s-t)^{-i+1/2}\left[\hat{P}^\xi_{t,s}|\varphi|\right](s,x), \quad i = 1,2,$$

(ii) Centering argument:

$$\left|D_{x_1}\left[\tilde{P}^\xi_{t,s}\varphi\right](s,x)\right| \le C''(s-t)^{-1+1/2}\left[\hat{P}^\xi_{t,s}|\varphi - \varphi(\cdot,\zeta)|\right](s,x),$$

$$\left|D_{x_2}\left[\tilde{P}^\xi_{t,s}\varphi\right](s,x)\right| \le C''(s-t)^{-2+1/2}\left[\hat{P}^\xi_{t,s}|\varphi - \varphi(\cdot,\cdot,\zeta_2)|\right](s,x),$$

Moreover, for all $i = 1, 2$ and for all γ in $(0, 1]$ we have

(iii) Gaussian smoothing:

$$\begin{aligned} &\left[\hat{P}^{\xi}_{t,s}|(\cdot - \theta_{t,s}(\xi))_i|^{\gamma}\right](s,x)\Big|_{(\tau,\xi)=(t,x)} \\ = &\left[\hat{P}^{x}_{t,s}|(\cdot - m^{t,x}_{t,s}(x))_i|^{\gamma}\right](s,x) \leq C'''(s-t)^{(i-1/2)\gamma}. \end{aligned} \tag{5.12}$$

Note that this last property implies in particular that if $\gamma \leq [\beta_1^2 \wedge \beta_2^2]$ and $T < 1$ we have

$$\left[\hat{P}^{\xi}_{t,s}\Delta^2_{\gamma}(\cdot - \theta_{t,s}(\xi))\right](s,x)\Big|_{(\tau,\xi)=(t,x)} = \left[\hat{P}^{x}_{t,s}\Delta^2_{\gamma}(\cdot - m^{t,x}_{t,s}(x))\right](s,x) \leq C'''(s-t)^{3\gamma/2}.$$

Finally, given x_2 and z_2 in $\mathbb{R}^{2d}$ and defining $\mathcal{S} := \left\{s \in (t,T] : \ |x_2 - z_2| \leq (s-t)^{3/2}\right\}$ we have the following Hölder estimate:

(iv) Hölder estimate in diagonal regime: for all s in $\mathcal{S}$ we have that for all measurable function $\varphi : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$:

$$\begin{aligned} &\left|D^n_{x_1}\left[\tilde{P}^{\xi}_{t,s}\varphi\right](s,x_1,x_2) - D^n_{x_1}\left[\tilde{P}^{\xi}_{t,s}\varphi\right](s,x_1,z_2)\right| \\ &\quad \leq C(s-t)^{-n/2-3\bar{\nu}/2}\left[\hat{P}^{\xi}_{t,s}|\varphi|\right](s,x_1,x_2)|x_2 - z_2|^{\bar{\nu}}, \end{aligned}$$

for every $0 < \bar{\nu} < 1$ and n in $\mathbb{N}$.

Remark. *Concerning (ii), in the following we say that we center the function φ around $\varphi(\cdot,\zeta)$ and $\varphi(\cdot,\cdot,\zeta_2)$ respectively.*

Proof. Assertion (i) follows from Definition 5.2, the inversion of the integral and the derivative operators and then estimate (5.8) in Proposition 5.1. For the first assertion in (ii), we have by definition of $\tilde{P}$ that the quantity $\left[\tilde{P}^{\xi}_{t,s}\varphi(\cdot,\zeta)\right](s,x)$ does not depend on x so that $D_{x_i}\left[\tilde{P}^{\xi}_{t,s}\varphi(\cdot,\zeta)\right](s,x) = 0$ for $i = 1, 2$. Hence, the assertion follows from the following splitting:

$$\forall \zeta \in \mathbb{R}^{2d}, \ \varphi = \varphi - \varphi(\cdot,\zeta) + \varphi(\cdot,\zeta).$$

The second assertion in (iii) follows from the same arguments. The last assertion (iii) of the Proposition follows from the Gaussian decay of $\hat{q}$. Indeed, by definition we have

$$\begin{aligned} &\left[\hat{P}^{\xi}_{t,s}|(\cdot - \theta_{t,s}(\xi))_i|^{\gamma}\right](s,x) \\ &= \int_{\mathbb{R}^{2d}} |y_i - \theta^i_{t,s}(\xi)|^{\gamma}\hat{q}_c(t,x;s,y)\mathrm{d}y \\ &= \int_{\mathbb{R}^{2d}} \left\{(s-t)^{(i-1/2)\gamma}\left|\frac{y_i - \theta^i_{t,s}(\xi)}{(s-t)}\right|^{\gamma}\frac{c}{(s-t)^{2d}}\right. \\ &\quad \left.\times \exp\left(-c\left(\frac{\left|y_1 - m^{1,\tau,\xi}_{t,s}(x)\right|^2}{s-t} + \frac{\left|y_2 - m^{2,\tau,\xi}_{t,s}(x)\right|^2}{(s-t)^3}\right)\right)\right\}\mathrm{d}y \end{aligned} \tag{5.13}$$

Note that for all s in $[t,T]$, the mean $(m^{1,t,x}_{t,s}(x), m^{2,t,x}_{t,s}(x))$ satisfies the ODE (5.1) with initial data (t,x). Hence, the forward transport function defined by (5.1) with the initial data $(\tau,\xi) = (t,x)$ is

equal to the mean: $\theta_{t,s}(x) = m^{t,x}_{t,s}(x)$ (recall that under our regularized setting (5.1) has a unique solution). We deduce the result by letting $(\tau,\xi) = (t,x)$ and by using the following inequality:

$$\forall \eta > 0,\ \forall q > 0,\ \exists \bar{C} > 0 \text{ s.t. } \forall \sigma > 0,\ \sigma^q e^{-\eta\sigma} \leq \bar{C}.$$

For the proof of the last assertion, note that on $\mathcal{S}$ we have for any measurable function $\varphi : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$:

$$\begin{aligned}
&\left| D_{x_1}\left[\tilde{P}^{\xi}_{t,s}\varphi\right](s,x_1,x_2) - D_{x_1}\left[\tilde{P}^{\xi}_{t,s}\varphi\right](s,x_1,z_2)\right| \\
&\quad \leq \sup_{\lambda\in(0,1)} \left| D_{x_2}D_{x_1}\left[\tilde{P}^{\xi}_{t,s}\varphi\right](s,x_1,\lambda x_2 + (1-\lambda)z_2)\right| |x_2 - z_2| \\
&\quad \leq C(s-t)^{-2}\left[\hat{P}^{\xi}_{t,s}|\varphi|\right](s,x_1,x_2)|x_2 - z_2| \\
&\quad \leq C(s-t)^{-1/2-3\bar{\nu}/2}\left[\hat{P}^{\xi}_{t,s}|\varphi|\right](s,x_1,x_2)|x_2 - z_2|^{\bar{\nu}},
\end{aligned}$$

for every $0 < \bar{\nu} < 1$. The last inequality essentially follows from the definitions (5.10) and (5.11), the form of $\tilde{q}$ and convexity inequality (see also the computations done in the proof of Claim 4.4 p29 of [4]). □

5.2 Estimation of the solution

Let us now expand the regularized solution of (1.6) to a a first order parametrix: we rewrite this PDE as

$$(\partial_t + \tilde{\mathcal{L}}^{\tau,\xi})u(t,x) = -\left(\mathcal{L} - \tilde{\mathcal{L}}^{\tau,\xi}\right)u(t,x) + f(t,x),$$

on $[0,T) \times \mathbb{R}^{2d}$ with terminal condition 0 at time T. Thus, using the definitions given in the previous subsection, we obtain that for every (t,x) in $[0,T] \times \mathbb{R}^{2d}$, the solution u writes

$$\begin{aligned}
u(t,x) \quad = \quad & \int_t^T \Bigg\{ \left[\tilde{P}^{\xi}_{t,s} f\right](s,x) - \left[\tilde{P}^{\xi}_{t,s}(F_1 - F_1(s,\theta_{t,s}(\xi)))\cdot D_1 u\right](s,x) \\
& - \left[\tilde{P}^{\xi}_{t,s}(F_2 - F_2(s,\theta_{t,s}(\xi)) - D_1F_2(s,\theta_{t,s}(\xi))(\cdot - \theta_{t,s}(\xi))_1)\cdot D_2 u\right](s,x) \\
& - \left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\mathrm{Tr}\left[(a - a(s,\theta_{t,s}(\xi)))D_1^2 u\right]\right](s,x)\Bigg\}\mathrm{d}s \qquad (5.14) \\
=: \quad & \int_t^T \mathcal{I}^{1,\xi}_{t,s}(x) + \mathcal{I}^{2,\xi}_{t,s}(x) + \mathcal{I}^{3,\xi}_{t,s}(x) + \mathcal{I}^{4,\xi}_{t,s}(x)\mathrm{d}s, \qquad (5.15)
\end{aligned}$$

by choosing $\tau = t$. We made this choice for the freezing time τ in the following.

We next assume without loss of generality that $T < 1$. We are now in position to prove the main estimates of Theorem 3.1. This is done by proving the following results and then using a circular argument: letting the time interval being sufficiently small to obtain each estimate separately.

Proposition 5.4. *There exists four positive constants C_1, C_2, C' and C'', and three positive numbers δ, δ' and δ'', depending on known parameters in **(H)** only, such that:*

$$\begin{aligned}
||D_1^n u||_\infty &\leq C_n T^\delta \left(1 + ||D_1 u||_\infty + ||D_2 u||_\infty + ||D_1^2 u||_\infty\right), \quad n = 1, 2, && (5.16)\\
||D_2 u||_\infty &\leq C' T^{\delta'} \Big(1 + ||D_1 u||_{\infty,\infty,\nu} + ||D_1 u||_{\infty,\infty,\nu} \\
&\qquad + ||D_2 u||_\infty + ||D_1^2 u||_\infty\Big), && (5.17)\\
||D_1 u||_{\infty,\infty,\nu} &\leq C'' T^{\delta''} \left(1 + ||D_1 u||_\infty + ||D_2 u||_\infty + ||D_1^2 u||_\infty\right), && (5.18)
\end{aligned}$$

where $||\cdot||_{\infty,\infty,\nu}$ is defined by (3.2) and such that:

$$\nu = \sup\{I \in \mathbb{R} : \ I < (1 + \beta_1^1)/3\}. \tag{5.19}$$

Before entering in the (long) proof of this result, let us mention that one may guess that the bound on the regularized solution depends on the regularity of the source term of the PDE f. This is indeed true, and this is why there is a 1 in the bounds above.

The main strategy consists in estimating the derivatives of the time integrands $\mathcal{I}^j$, $j = 1, \ldots, 4$, of the representation (5.14) and then to invert the differentiation and integration operators. Hence, int the following, we estimate the derivatives of these time integrands and investigate their Hölder regularity. Without loss of generality, we suppose in the following that $T < 1$.

Proof of estimate (5.16). Let $n \in \{1, 2\}$ and s in $(t, T]$. We have from (5.14) and Proposition 5.3 (assertions (i) and (ii)):

$$\begin{aligned}
&\left| D_{x_1}^n \left\{ \mathcal{I}_{t,s}^{1,\xi}(x) + \mathcal{I}_{t,s}^{2,\xi}(x) + \mathcal{I}_{t,s}^{3,\xi}(x) + \mathcal{I}_{t,s}^{4,\xi}(x) \right\} \right| \\
= &\Bigg| D_{x_1}^n \Bigg\{ \left[\tilde{P}_{t,s}^\xi f\right](s,x) - \left[\tilde{P}_{t,s}^\xi (F_1 - F_1(s, \theta_{t,s}(\xi))) \cdot D_1 u\right](s,x) \\
&\quad - \left[\tilde{P}_{t,s}^\xi (F_2 - F_2(s, \theta_{t,s}(\xi)) - D_1 F_2(s, \theta_{t,s}(\xi))(\cdot - \theta_{t,s}(\xi))_1) \cdot D_2 u\right](s,x) \\
&\quad - \left[\tilde{P}_{t,s}^\xi \frac{1}{2} \mathrm{Tr}\left[(a - a(s, \theta_{t,s}(\xi)))D_1^2 u\right]\right](s,x) \Bigg\} \Bigg| \\
\leq &\ C(s-t)^{-n/2} \Bigg\{ \left[\hat{P}_{t,s}^\xi \big| f - f(s, \theta_{t,s}(\xi)) \big|\right](s,x) \\
&\quad + \left[\hat{P}_{t,s}^\xi \big| (F_1 - F_1(s, \theta_{t,s}(\xi))) \cdot D_1 u \big|\right](s,x) \\
&\quad + \left[\hat{P}_{t,s}^\xi \big| (F_2 - F_2(s, \theta_{t,s}(\xi)) - D_1 F_2(s, \theta_{t,s}(\xi))(\cdot - \theta_{t,s}(\xi))_1) \cdot D_2 u \big|\right](s,x) \\
&\quad + \left[\hat{P}_{t,s}^\xi \big| \frac{1}{2} \mathrm{Tr}\left[(a - a(s, \theta_{t,s}(\xi)))D_1^2 u\right] \big|\right](s,x) \Bigg\}.
\end{aligned}$$

By using the regularity of the coefficients assumed in **(H)**, Lipschitz regularity on f and expanding F_2 around the forward transport θ:

$$F_2(t,\cdot) = F_2(t,\theta_{t,s}(\xi)) + \int_{\theta_{t,s}(\xi)}^{\cdot} D_1 F_2(t,z)\mathrm{d}z$$

we get that

$$\begin{aligned}
&\left| D^n_{x_1}\left\{\mathcal{I}^{1,\xi}_{t,s}(x)+\mathcal{I}^{2,\xi}_{t,s}(x)+\mathcal{I}^{3,\xi}_{t,s}(x)+\mathcal{I}^{4,\xi}_{t,s}(x)\right\}\right| \\
\leq\ & C(s-t)^{-n/2}\Bigg\{ \left[\hat{P}^{\xi}_{t,s}\left|(\cdot-\theta_{t,s}(\xi))_1\right|\right](s,x) + \left[\hat{P}^{\xi}_{t,s}\left|(\cdot-\theta_{t,s}(\xi))_2\right|\right](s,x) \\
&+||D_1u||_\infty\left[\hat{P}^{\xi}_{t,s}\Big(\left|(\cdot-\theta_{t,s}(\xi))_1\right|^{\beta^1_1}+\left|(\cdot-\theta_{t,s}(\xi))_2\right|^{\beta^2_1}\Big)\right](s,x) \\
&+||D^2_1u||_\infty\left[\hat{P}^{\xi}_{t,s}\Big(\left|(\cdot-\theta_{t,s}(\xi))_1\right|+\left|(\cdot-\theta_{t,s}(\xi))_2\right|\Big)\right](s,x) \\
&+||D_2u||_\infty\left[\hat{P}^{\xi}_{t,s}\Big(\left|(\cdot-\theta_{t,s}(\xi))_1\right|^{1+\eta}+\left|(\cdot-\theta_{t,s}(\xi))_2\right|^{\beta^2_2}\Big)\right](s,x)\Bigg\}.
\end{aligned}$$

By letting $\xi = x$ we obtain from estimate (5.12) in Proposition 5.3 that

$$\begin{aligned}
&\left| D^n_{x_1}\left\{\mathcal{I}^{1,\xi}_{t,s}(x)+\mathcal{I}^{2,\xi}_{t,s}(x)+\mathcal{I}^{3,\xi}_{t,s}(x)+\mathcal{I}^{4,\xi}_{t,s}(x)\right\}\right| \\
\leq\ & C(s-t)^{-n/2}\Big((s-t)^{1/2}+(s-t)^{3/2}+||D_1u||_\infty\Big((s-t)^{\beta^1_1/2}+(s-t)^{3\beta^2_1/2}\Big) \\
&+||D^2_1u||_\infty\Big((s-t)^{1/2}+(s-t)^{3/2}\Big)+||D_2u||_\infty\Big((s-t)^{(1+\eta)/2}+(s-t)^{3\beta^2_2/2}\Big)\Big),
\end{aligned}$$

where all the time-singularities in the right hand side are integrables. Therefore

$$\begin{aligned}
|D^n_{x_1}u(t,x)| \leq\ & CT^{1-n/2}\Big(T^{1/2}+T^{3/2}+||D_1u||_\infty\Big(T^{\beta^1_1/2}+T^{3\beta^2_1/2}\Big) \\
&+||D^2_1u||_\infty\Big(T^{1/2}+T^{3/2}\Big)+||D_2u||_\infty\Big(T^{(1+\eta)/2}+T^{3\beta^2_2/2}\Big)\Big).
\end{aligned}$$

□

Proof of estimate (5.17). We now estimate the derivative of the solution in the degenerate direction. We are unfortunately not able to repeat the previous strategy since the Gaussian smoothing (see eq. (5.13)) of part of the coefficients are not strong enough to smooth the time singularity coming from the derivative (of order $-3/2$) in the degenerate direction. To overcome this problem we are lead to re-center some of the terms around the solution of the PDE (*i.e.* to use the second estimate in (*ii*) of

Proposition 5.3 on the function D_1u): namely, $\mathcal{I}^2$ and $\mathcal{I}^4$ in (5.14). In order to clarify our exposition, we hence estimate each term $\mathcal{I}^{j,\xi}_{t,s},\ j=1,\dots,4$ defined in (5.14) separately.

From Proposition 5.3, (ii) and by using the Lipschitz regularity of f and the regularity of the coefficients assumed in **(H)**

$$\begin{aligned}\left|D_{x_2}\mathcal{I}^{1,\xi}_{t,s}(x)\right| &= \left|D_{x_2}\left[\tilde{P}^{\xi}_{t,s}(f-f(s,\cdot,\theta^2_{t,s}(\xi)))\right](s,x)\right| \\ &\leq C(s-t)^{-3/2}\left[\hat{P}^{\xi}_{t,s}\left|(\cdot-\theta_{t,s}(\xi))_2\right|\right](s,x),\end{aligned} \tag{5.20}$$

$$\begin{aligned}&\left|D_{x_2}\mathcal{I}^{2,\xi}_{t,s}(x)\right| \\ =\ &\Big|D_{x_2}\left[\tilde{P}^{\xi}_{t,s}(F_1-F_1(s,\cdot,\theta^2_{t,s}(\xi)))D_1u\right](s,x) \\ &+D_{x_2}\left[\tilde{P}^{\xi}_{t,s}(F_1(s,\cdot,\theta^2_{t,s}(\xi)))-F_1(s,\theta_{t,s}(\xi)))(D_1u-D_1u(s,\cdot,\theta^2_{t,s}(\xi))\right](s,x)\Big| \\ \leq\ & C(s-t)^{-3/2}||D_1u||_\infty\left[\hat{P}^{\xi}_{t,s}\Big(\left|(\cdot-\theta_{t,s}(\xi))_2(\cdot)\right|^{\beta^2_1}\Big)\right](s,x) \\ &+C(s-t)^{-3/2}||D_1u||_{\infty,\infty,\nu}\left[\hat{P}^{\xi}_{t,s}\Big(\left|(\cdot-\theta_{t,s}(\xi))_1\right|^{\beta^1_1}\Delta^2_\nu(\cdot-\theta_{t,s}(\xi))\Big)\right](s,x),\end{aligned} \tag{5.21}$$

where $||D_1u||_{\infty,\infty,\nu}$ is defined by (3.2) and Δ^2_ν is defined by (5.9) and where we used the second estimate in (ii) of Proposition 5.3 on D_1u. We also have

$$\begin{aligned}&\left|D_{x_2}\mathcal{I}^{3,\xi}_{t,s}(x)\right| \\ =\ &\left|D_{x_2}\left[\tilde{P}^{\xi}_{t,s}(F_2-F_2(s,\theta_{t,s}(\xi))-D_1F_2(s,\theta_{t,s}(\xi)))\cdot D_2u\right](s,x)\right| \\ \leq\ & C(s-t)^{-3/2}||D_2u||_\infty\left[\hat{P}^{\xi}_{t,s}\Big(\left|(\cdot-\theta_{t,s}(\xi))_1\right|^{1+\eta}+\left|(\cdot-\theta_{t,s}(\xi))_2\right|^{\beta^2_2}\Big)\right](s,x).\end{aligned} \tag{5.22}$$

We now deal with the term $\mathcal{I}^4$ which is the delicate part. We first define for all measurable function $\varphi:[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$, for all $t<s$ in $[0,T]^2$, ξ, x in $\mathbb{R}^{2d}$ and $l=1,\dots,d$:

$$\left[\tilde{P}'^{,l,\xi}_{t,s}\varphi\right](s,x)=\int_{\mathbb{R}^{2d}}\varphi(s,y)D_{y_{1l}}\tilde{q}(t,x;s,y)\mathrm{d}y, \tag{5.23}$$

where y_{1l} denotes the lth component of the d-dimensional variable y_1. Note that the derivative in the integral above is w.r.t. the integration variable.

Now we can split $a-a(s,\theta_{t,s}(\xi))$ as $a-a(s,\cdot,\theta^2_{t,s}(\xi))+a(s,\cdot,\theta^2_{t,s}(\xi))-a(s,\theta_{t,s}(\xi))$ so that

$$\begin{aligned}D_{x_2}\mathcal{I}^{4,\xi}_{t,s}(x) &= -D_{x_2}\left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\mathrm{Tr}\left[(a-a(s,\cdot,\theta_{t,s}(\xi)))D^2_1u\right]\right](s,x) \\ &\quad -D_{x_2}\left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\mathrm{Tr}\left[(a(s,\cdot,\theta^2_{t,s}(\xi))-a(s,\theta_{t,s}(\xi)))D^2_1u\right]\right](s,x).\end{aligned}$$

Applying now an integration by parts on the second term in the right hand side leads to

$$\begin{aligned} D_{x_2}\mathcal{I}^{4,\xi}_{t,s}(x) &= -D_{x_2}\left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\mathrm{Tr}\left[(a-a(s,\cdot,\theta_{t,s}(\xi)))D^2_1u\right]\right](s,x) \\ &\quad +\sum_{l=1}^{d} D_{x_2}\left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\left[(D_{1l}a_{\cdot l}(s,\cdot,\theta^2_{t,s}(\xi)))\cdot D_1u\right]\right](s,x) \\ &\quad +\sum_{l=1}^{d} D_{x_2}\left[\tilde{P}^{\prime,l,\xi}_{t,s}\frac{1}{2}\left[(a-a(s,\cdot,\theta_{t,s}(\xi)))_{\cdot l}\right]\cdot D_1u\right](s,x), \end{aligned}$$

where $a_{\cdot l}$ denotes the lth column of the matrix a. Hence, by using Proposition 5.3 (ii) with D_1u we obtain

$$\begin{aligned} & D_{x_2}\mathcal{I}^{4,\xi}_{t,s}(x) \\ = \; & -D_{x_2}\left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\mathrm{Tr}\left[(a-a(s,\cdot,\theta_{t,s}(\xi)))D^2_1u\right]\right](s,x) \\ & +\sum_{l=1}^{d} D_{x_2}\left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\left[(D_{1l}a_{\cdot l}(s,\cdot,\theta^2_{t,s}(\xi)))\cdot\left[D_1u-D_1u(s,\cdot,\theta^2_{t,s}(\xi))\right]\right]\right](s,x) \\ & +\sum_{l=1}^{d} D_{x_2}\left[\tilde{P}^{\prime,l,\xi}_{t,s}\frac{1}{2}\left[(a-a(s,\cdot,\theta_{t,s}(\xi)))_{\cdot l}\right]\cdot\left[D_1u-D_1u(s,\cdot,\theta^2_{t,s}(\xi))\right]\right](s,x). \end{aligned}$$

We can finally deduce that:

$$\begin{aligned} \left|D_{x_2}\mathcal{I}^{4,\xi}_{t,s}(x)\right| \leq\; & C(s-t)^{-3/2}\Bigg\{||D^2_1u||_\infty\left[\hat{P}^{\xi}_{t,s}\Big(\big|(\cdot-\theta_{t,s}(\xi))_2\big|\Big)\right](s,x) \\ & +||D_1u||_{\infty,\infty,\nu}\Bigg(\left[\hat{P}^{\xi}_{t,s}\Big(\Delta^2_\nu(\cdot-\theta_{t,s}(\xi))\Big)\right](s,x) \\ & +(s-t)^{-1/2}\left[\hat{P}^{\xi}_{t,s}\Big(\big|(\cdot-\theta_{t,s}(\xi))_1\big|\Delta^2_\nu(\cdot-\theta_{t,s}(\xi))\Big)\right](s,x)\Bigg)\Bigg\}, \end{aligned} \tag{5.24}$$

Where $||D_1u||_{\infty,\infty,\nu}$ is defined by (3.2) and Δ^2_ν is defined by (5.9).

Hence, collecting all the previous estimates ((5.20), (5.21), (5.22), (5.24)) we eventually deduce

$$\begin{aligned}
&\left|D_{x_2}\left\{\mathcal{I}^{1,\xi}_{t,s}(x)+\mathcal{I}^{2,\xi}_{t,s}(x)+\mathcal{I}^{3,\xi}_{t,s}(x)+\mathcal{I}^{4,\xi}_{t,s}(x)\right\}\right| \\
\leq\ & C(s-t)^{-3/2}\Bigg\{\left[\hat{P}^{\xi}_{t,s}\left|(\cdot-\theta_{t,s}(\xi))_2\right|\right](s,x) \\
&+||D_1u||_\infty\left[\hat{P}^{\xi}_{t,s}\left(\left|(\cdot-\theta_{t,s}(\xi))_2(\cdot)\right|^{\beta^2_1}\right)\right](s,x) \\
&+||D_1u||_{\infty,\infty,\nu}\left[\hat{P}^{\xi}_{t,s}\left(\left|(\cdot-\theta_{t,s}(\xi))_1\right|^{\beta^1_1}\Delta^2_\nu(\cdot-\theta_{t,s}(\xi))\right)\right](s,x) \\
&+||D_2u||_\infty\left[\hat{P}^{\xi}_{t,s}\left(\left|(\cdot-\theta_{t,s}(\xi))_1\right|^{1+\eta}+\left|(\cdot-\theta_{t,s}(\xi))_2\right|^{\beta^2_2}\right)\right](s,x) \\
&+||D^2_1u||_\infty\left[\hat{P}^{\xi}_{t,s}\left(\left|(\cdot-\theta_{t,s}(\xi))_2\right|\right)\right](s,x) \\
&+||D_1u||_{\infty,\infty,\nu}\Bigg(\left[\hat{P}^{\xi}_{t,s}\left(\Delta^2_\nu(\cdot-\theta_{t,s}(\xi))\right)\right](s,x) \\
&\quad+(s-t)^{-1/2}\left[\hat{P}^{\xi}_{t,s}\left(\left|(\cdot-\theta_{t,s}(\xi))_1\right|\Delta^2_\nu(\cdot-\theta_{t,s}(\xi))\right)\right](s,x)\Bigg)\Bigg\}.
\end{aligned}$$

By letting $\xi=x$ we obtain from estimate (5.12) in Proposition 5.3 that

$$\begin{aligned}
&\left|D_{x_2}\left\{\mathcal{I}^{1,\xi}_{t,s}(x)-\mathcal{I}^{2,\xi}_{t,s}(x)-\mathcal{I}^{3,\xi}_{t,s}(x)-\mathcal{I}^{4,\xi}_{t,s}(x)\right\}\right| \\
\leq\ & C(s-t)^{-3/2}\Bigg((s-t)^{3/2}+||D_1u||_\infty(s-t)^{3\beta^2_1/2} \\
&+||D_1u||_{\infty,\infty,\nu}(s-t)^{\beta^1_1/2+3\nu/2}+||D_2u||_\infty\Big((s-t)^{(1+\eta)/2} \\
&+(s-t)^{3\beta^2_2/2}\Big)+||D^2_1u||_\infty(s-t)^{3/2}+2||D_1u||_{\infty,\infty,\nu}(s-t)^{3\nu/2}.
\end{aligned}$$

Since $\beta^2_j>1/3$ (condition (2.1)), and ν is given by (5.19), all the time-singularities of the right hand side above are integrable on $(t,T]$. Hence, we deduce from (5.14) and the estimate above that there exists a positive δ', depending on known parameters in **(H)** only, such that:

$$|D_{x_2}u(t,x)| \leq CT^{\delta'}\Big(1+||D_1u||_{\infty,\infty,\nu}+||D_1u||_\infty+||D_2u||_\infty+||D^2_1u||_\infty\Big).$$

□

Proof of estimate (5.18). Finally, we compute the Hölder semi norm of $D_{x_1}u$. Let $x_2 \neq z_2$ belong to $\mathbb{R}^d$. We have from (5.14):

$$\begin{aligned} & D_{x_1}u(t,x_1,x_2) - D_{x_1}u(t,x_1,z_2) \\ = \quad & D_{x_1}\int_t^T \left\{ \sum_{j=1}^4 \mathcal{I}_{t,s}^{j,\xi}(x_1,x_2) - \mathcal{I}_{t,s}^{j,\xi}(x_1,z_2) \right\} \mathrm{d}s. \end{aligned} \tag{5.25}$$

We first estimate for any s in $(t,T]$ the quantity:

$$\left| D_{x_1} \sum_{j=1}^4 \mathcal{I}_{t,s}^{j,\xi}(x_1,x_2) - \mathcal{I}_{t,s}^{j,\xi}(x_1,z_2) \right|. \tag{5.26}$$

To do this, we split the time interval w.r.t. the characteristic time-scale of the second space variable: let $\mathcal{S} := \left\{ s \in (t,T] : \ |x_2 - z_2| \leq (s-t)^{3/2} \right\}$. Note that on $\mathcal{S}$, the result is quite obvious: from the last assertion of Proposition 5.3 we have from **(H)** that (5.26) is bounded on $\mathcal{S}$ by (see the computations already done for the proof of (5.16)):

$$\begin{aligned} & C'(s-t)^{-1/2-3\bar{\nu}/2}\Big((s-t)^{1/2} + (s-t)^{3/2} + ||D_1 u||_\infty \big((s-t)^{\beta_1^1/2} + (s-t)^{3\beta_1^2/2} \big) \\ & \quad + ||D_2 u||_\infty \big((s-t)^{(1+\eta)/2} + (s-t)^{3\beta_2^2/2} \big) \Big) |x_2 - z_2|^{\bar{\nu}} \\ & \quad + + ||D_1^2 u||_\infty \big((s-t)^{1/2} + (s-t)^{3/2} \big), \end{aligned} \tag{5.27}$$

where we chose $\xi = x$. Note that this term is integrable on $(t,T]$ for all $\bar{\nu} < (1+\beta_1^1)/3$.

We now estimate (5.26) on $\mathcal{S}^c$. On a first hand, we have from the computations done when estimating $D_{x_1}u$ that:

$$\begin{aligned} & \left| D_{x_1} \sum_{j=1}^4 \mathcal{I}_{t,s}^{j,\xi}(x_1,x_2) \right| \\ \leq \quad & C(s-t)^{-1/2-3\bar{\nu}/2} \Big\{ (s-t)^{1/2} + (s-t)^{3/2} + ||D_1 u||_\infty \Big((s-t)^{\beta_1^1} \\ & \quad + (s-t)^{\beta_1^2} \Big) + ||D_2 u||_\infty \Big((s-t)^{(1+\eta)/2} + (s-t)^{3\beta_2^2/2} \Big) \\ & \quad + ||D_1^2 u||_\infty \Big((s-t)^{1/2} + (s-t)^{3/2} \Big) \Big\} |x_2 - z_2|^{\bar{\nu}}, \end{aligned} \tag{5.28}$$

since for any positive number $\bar{\nu}$ on $\mathcal{S}^c$ we have $1 \leq (s-t)^{-3\bar{\nu}/2}|x_2 - z_2|^{\bar{\nu}}$ and since we chose $\xi = x$ and then used Proposition 5.3. We emphasize that this is the same bound as (5.27) so that all the time singularity above are again integrables provided $\bar{\nu} < (1+\beta_1^1)/3$. It thus only remains to estimate the last part of (5.26) on $\mathcal{S}^c$, namely

$$
\begin{aligned}
&\left| D_{x_1} \sum_{j=1}^{4} \mathcal{I}_{t,s}^{j,\xi}(x_1,z_2) \right| \\
= &\Bigg| D_{x_1} \Bigg\{ \left[\tilde{P}_{t,s}^{\xi} f \right](s,x_1,z_2) - \left[\tilde{P}_{t,s}^{\xi}(F_1 - F_1(s,\theta_{t,s}(\xi))) \cdot D_1 u \right](s,x_1,z_2) \\
&- \left[\tilde{P}_{t,s}^{\xi}(F_2 - F_2(s,\theta_{t,s}(\xi)) - D_1 F_2(s,\theta_{t,s}(\xi))) \cdot D_2 u \right](s,x_1,z_2) \\
&- \left[\tilde{P}_{t,s}^{\xi} \frac{1}{2} \mathrm{Tr}\left[(a - a(s,\theta_{t,s}(\xi)))D_1^2 u \right] \right](s,x_1,z_2) \Bigg\} \Bigg|.
\end{aligned}
$$

The main issue here is that the Gaussian smoothing of Proposition 5.3 can not be applied immediately (last part of (iii)), the semi-group being evaluating at point (s,x_1,z_2) and the freezing point being previously chosen as $\xi = (x_1,x_2)$. To smooth the time singularity our control has to be of the form

$$
\left[\hat{P}_{t,s}^{x} |(\cdot - m_{t,s}^{t,x}(x_1,z_2))_2|^{\gamma} \right](s,x_1,z_2),
$$

which is (see the proof of assertion (iii) of Proposition 5.3) bounded by $C(s-t)^{3\gamma/2}$. To do so, the main idea consists in re-centering all the terms above around $m_{t,s}^{2,t,\xi}(x_1,z_2)$. When choosing $\xi = x$, we end with the remaining difference $|\theta_{t,s}^2(x) - m_{t,s}^{2,t,x}(x_1,z_2)|$ which is, fortunately, bounded by $|x_2 - z_2|$ (this follows from the definition (5.1) and (5.3) of θ and m respectively).

Let us first begin with the term $D_{x_1}\mathcal{I}_{t,s}^{1,\xi}(x_1,z_2) = D_{x_1}\left[\tilde{P}_{t,s}^{\xi} f \right](s,x_1,z_2)$. Splitting first f as

$$
\Big(f - f(s,\theta_{t,s}^1(\xi), m_{t,s}^{2,t,\xi}(\xi_1,z_2)) \Big) + f(s,\theta_{t,s}^1(\xi), m_{t,s}^{2,t,\xi}(\xi_1,z_2)),
$$

we get from Proposition 5.3

$$
\begin{aligned}
&\left| D_{x_1}\mathcal{I}_{t,s}^{1,\xi}(x_1,z_2) \right| \\
= &\left| D_{x_1} \left[\tilde{P}_{t,s}^{\xi}(f - f(s,\theta_{t,s}(\xi))) \right](s,x_1,z_2) \right| \\
\leq &\; C(s-t)^{-1/2} \left[\hat{P}_{t,s}^{\xi} \Big(|(\cdot - \theta_{t,s}(\xi))_1| + |(\cdot - m_{t,s}^{t,\xi}(\xi_1,z_2))_2| \Big) \right](s,x_1,z_2).
\end{aligned} \tag{5.29}
$$

Next, we consider the term

$$
D_{x_1}\mathcal{I}_{t,s}^{2,\xi}(x_1,z_2) = D_{x_1}\left[\tilde{P}_{t,s}^{\xi}(F_1 - F_1(s,\theta_{t,s}(\xi))) \cdot D_1 u \right](s,x_1,z_2).
$$

Splitting $(F_1 - F_1(s,\theta_{t,s}(\xi))) \cdot D_1 u$ as

$$
\begin{aligned}
&\Big(F_1 - F_1(s,\theta_{t,s}^1(\xi), m_{t,s}^{2,t,\xi}(x_1,z_2)) \Big) \cdot D_1 u + \Big(F_1(s,\theta_{t,s}^1(\xi), m_{t,s}^{2,t,\xi}(x_1,z_2)) \\
&- F_1(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi)) \Big) \cdot D_1 u,
\end{aligned}
$$

we get

$$\begin{aligned}
&\left|D_{x_1}\mathcal{I}^{2,\xi}_{t,s}(x_1,z_2)\right| \\
= &\left|D_{x_1}\left[\tilde{P}^{\xi}_{t,s}(F_1-F_1(s,\theta_{t,s}(\xi)))\cdot D_1u\right](s,x_1,z_2)\right| \\
\leq &\ C(s-t)^{-1/2}||D_1u||_\infty\Big[\hat{P}^{\xi}_{t,s}\Big(|(\cdot-\theta_{t,s}(\xi))_1|^{\beta^1_1}+|(\cdot-m^{t,\xi}_{t,s}(x_1,z_2))_2|^{\beta^2_1} \\
&+|\theta^2_{t,s}(\xi)-m^{2,\xi}_{t,s}(x_1,z_2)|^{\beta^2_1}\Big)\Big](s,x_1,z_2). \qquad (5.30)
\end{aligned}$$

We now deal with the integrand in $\mathcal{I}^{3,\xi}_{t,s}(x_1,z_2)$, we split $(F_2-F_2(s,\theta_{t,s}(\xi))-D_1F_2(s,\theta_{t,s}(\xi)))\cdot D_2u$ as

$$\begin{aligned}
&\Big(F_2-F_2(s,\cdot,m^{2,t,\xi}_{t,s}(x_1,z_2))\Big)\cdot D_2u \\
&+\Big(F_2(s,\cdot,m^{2,t,\xi}_{t,s}(x_1,z_2))-F_2(s,\theta_{t,s}(\xi))-D_1F_2(s,\theta_{t,s}(\xi))\Big)\cdot D_2u,
\end{aligned}$$

and we obtain

$$\begin{aligned}
&\left|D_{x_1}\mathcal{I}^{3,\xi}_{t,s}(x_1,z_2)\right| \qquad (5.31) \\
= &\left|D_{x_1}\left[\tilde{P}^{\xi}_{t,s}(F_2-F_2(s,\theta_{t,s}(\xi))-D_1F_2(s,\theta_{t,s}(\xi)))\cdot D_2u\right](s,x_1,z_2)\right| \\
\leq &\ C(s-t)^{-1/2}||D_2u||_\infty\Big[\hat{P}^{\xi}_{t,s}\Big(|(\cdot-m^{t,\xi}_{t,s}(x_1,z_2))_2|^{\beta^2_2} \\
&+|(\cdot-\theta_{t,s}(\xi))_1|^{1+\eta}+|\theta^2_{t,s}(\xi)-m^{2,t,\xi}_{t,s}(x_1,z_2)|^{\beta^2_2}\Big)\Big](s,x_1,z_2).
\end{aligned}$$

Finally, we deal with $D_{x_1}\mathcal{I}^{4,\xi}_{t,s}(x_1,z_2)$. We write $(1/2)\mathrm{Tr}\left[(a-a(s,\theta_{t,s}(\xi))D^2u_1\right]$ as

$$\begin{aligned}
&\frac{1}{2}\mathrm{Tr}\left[\Big(a-a(s,\theta^1_{t,s}(\xi),m^{2,t,\xi}_{t,s}(x_1,z_2))\Big)D^2_1u\right] \\
&+\frac{1}{2}\mathrm{Tr}\left[\Big(a(s,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi))-a(s,\theta^1_{t,s}(\xi),m^{2,t,\xi}_{t,s}(x_1,z_2))\Big)D^2_1u\right],
\end{aligned}$$

and we obtain

$$\begin{aligned}
\left|D_{x_1}\mathcal{I}^{4,\xi}_{t,s}(x_1,z_2)\right| &= \left|D_{x_1}\left[\tilde{P}^{\xi}_{t,s}\frac{1}{2}\mathrm{Tr}\left[(a-a(s,\theta_{t,s}(\xi))D^2u_1\right]\right](s,x_1,z_2)\right| \\
&\leq C(s-t)^{-1/2}||D^2_1u||_\infty\Big[\hat{P}^{\xi}_{t,s}\Big(|(\cdot-\theta_{t,s}(\xi))_1| \qquad (5.32) \\
&\qquad +|(\cdot-m^{2,t,\xi}_{t,s}(x_1,z_2))|+|\theta^2_{t,s}(\xi)-m^{2,t,\xi}_{t,s}(x_1,z_2)|\Big)\Big](s,x_1,z_2).
\end{aligned}$$

Hence, putting the previous estimates together ((5.29),(5.30),(5.31),(5.32)), letting $\xi = x$ we get that on $\mathcal{S}^c$

$$\begin{aligned}
&\left| D_{x_1} \sum_{j=1}^{4} \mathcal{I}_{t,s}^{j,\xi}(s, x_1, z_2) \right| \qquad (5.33)\\
\leq\ & C(s-t)^{-1/2} \Bigg\{ \Big\{ \big((s-t)^{(1-3\bar{\nu})/2} + (s-t)^{3(1-\bar{\nu})/2} \big) \\
&+ ||D_1 u||_\infty \Big((s-t)^{(\beta_1^1 - 3\bar{\nu})/2} + (s-t)^{3(\beta_1^2 - \bar{\nu})/2} \Big) \\
&+ ||D_2 u||_\infty \Big((s-t)^{3(\beta_2^2 - \bar{\nu})/2} + (s-t)^{(1+\eta-\bar{\nu})/2} \Big) \\
&+ ||D_1^2 u||_\infty \Big((s-t)^{(1-3\bar{\nu})/2} + (s-t)^{3(1-\bar{\nu})/2} \Big) \Big\} |x_2 - z_2|^{\bar{\nu}} \\
&+ ||D_1 u||_\infty |x_2 - z_2|^{\beta_1^2} + ||D_2 u||_\infty |x_2 - z_2|^{\beta_2^2} + ||D_1^2 u||_\infty |x_2 - z_2| \Bigg\}
\end{aligned}$$

since $1 \leq (s-t)^{-3\bar{\nu}/2} |x_2 - z_2|^{\bar{\nu}}$ and this holds for all $\bar{\nu}$ such that $\bar{\nu} < (1 + \beta_1^1)/3]$. Putting together estimates (5.27), (5.28) and (5.33), we can invert the differentiation and integration operators in (5.25) and we deduce that there exists a positive δ'' depending on known parameters in **(H)** only such that

$$||D_1 u||_{\infty,\infty,\nu} \leq C T^{\delta''} (||D_1 u||_\infty + ||D_2 u||_\infty + ||D_1^2 u||_\infty),$$

where $|| \cdot ||_{\infty,\infty,\nu}$ is defined in Theorem 3.1. □

Acknowledgments

I would like to thank François Delarue and Mario Maurelli for valuable contribution to this work and the anonymous referees for valuables remarks and suggestions.

References

[1] L. Beck, F. Flandoli, M. Gubinelli and M. Maurelli, Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness, *arXiv:1401.1530 [math]*,

[2] G. Cannizzaro and K. Chouk, Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential, *To appear in Annals of Probability*, *arXiv:1501.04751 [math]*,

[3] R. Catellier and M. Gubinelli, Averaging along irregular curves and regularisation of ODEs, *Stochastic Processes and their Applications*, **126** (2016), 2323–2366,

[4] P. E. Chaudru de Raynal, Strong existence and uniqueness for degenerate SDE with Hölder drift, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, **53** (2017), 259–286,

[5] F. Delarue and R. Diel, Rough paths and 1d SDE with a time dependent distributional drift: application to polymers, *Probability Theory and Related Fields*, 1–63,

[6] F. Delarue and F. Flandoli, The transition point in the zero noise limit for a 1d Peano example, *Discrete and Continuous Dynamical Systems*, **34** (2014), 4071–4083,

[7] F. Delarue and S. Menozzi, Density estimates for a random noise propagating through a chain of differential equations, *Journal of Functional Analysis*, **259** (2010), 1577–1630,

[8] M. Di Francesco and S. Polidoro, Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form, *Advances in Differential Equations*, **11** (2006), 1261–1320,

[9] R. J. DiPerna and P.-L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, *Inventiones Mathematicae*, **98** (1989), 511–547,
[10] E. Fedrizzi, F. Flandoli, E. Priola and J. Vovelle, Regularity of stochastic kinetic equations, *Electronic Journal of Probability*, **22**,
[11] F. Flandoli, E. Issoglio and F. Russo, Multidimensional stochastic differential equations with distributional drift, *Transactions of the American Mathematical Society*, **369** (2017), 1665–1688,
[12] F. Flandoli, *Random perturbation of PDEs and fluid dynamic models*, vol. 2015 of Lecture Notes in Mathematics, Springer, Heidelberg, 2011, Lectures from the 40th Probability Summer School held in Saint-Flour, 2010.
[13] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1964,
[14] M. Hairer, Introduction to regularity structures, *Brazilian Journal of Probability and Statistics*, **29** (2015), 175–210,
[15] L. Hörmander, Hypoelliptic second order differential equations, *Acta Mathematica*, **119** (1967), 147–171,
[16] A. Kolmogorov, {Zuf{ä}llige Bewegungen. (Zur Theorie der Brownschen Bewegung.).}, *Ann. of Math., II. Ser.*, **35** (1934), 116–117.
[17] N. V. Krylov and M. R\"ockner, Strong solutions of stochastic equations with singular time dependent drift, *Probability Theory and Related Fields*, **131** (2005), 154–196,
[18] H. P. McKean Jr. and I. M. Singer, Curvature and the eigenvalues of the Laplacian, *Journal of Differential Geometry*, **1** (1967), 43–69,
[19] S. Menozzi, Parametrix techniques and martingale problems for some degenerate Kolmogorov equations, *Electronic Communications in Probability*, **16** (2011), 234–250,
[20] S. Menozzi, Martingale problems for some degenerate Kolmogorov equations, *Stochastic Processes and their Applications*,
[21] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, vol. 233 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1979,
[22] A. J. Veretennikov, Strong solutions and explicit formulas for solutions of stochastic integral equations, *Matematicheski\u\ Sbornik. Novaya Seriya*, **111(153)** (1980), 434–452, 480,
[23] F. Wang and X. Zhang, Degenerate SDE with Hölder–Dini Drift and Non-Lipschitz Noise Coefficient, *SIAM Journal on Mathematical Analysis*, **48** (2016), 2189–2226,
[24] X. Zhang, Strong solutions of SDES with singular drift and Sobolev diffusion coefficients, *Stochastic Processes and their Applications*, **115** (2005), 1805–1818,
[25] X. Zhang, Stochastic Hamiltonian flows with singular coefficients, *arXiv:1606.04360 [math]*,
[26] A. K. Zvonkin, A transformation of the phase space of a diffusion process that will remove the drift, *Mat. Sb. (N.S.)*, **93(135)** (1974), 129–149, 152,

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: `pe.deraynal@univ-smb.fr`