A simple method for the existence of a density for stochastic evolutions with rough coefficients

Marco Romito*

Abstract

We extend the validity of a simple method for the existence of a density for stochastic differential equations, first introduced in [15], by proving local estimates for the density, existence for the density with summable drift, and by improving the regularity of the density.

Keywords: density of laws; stochastic differential equations; irregular coefficients; Besov spaces.
AMS MSC 2010: Primary 60H10, Secondary 60H07; 60H15; 34K50.
Submitted to EJP on October 13, 2017, final version accepted on November 3, 2018.
Supersedes arXiv:1707.05042.

Contents

*Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, I–56127 Pisa, Italia.
E-mail: marco.romito@unipi.it

1 Introduction

The purpose of this paper is to illustrate and improve a simple but effective method to prove existence and minimal regularity of a density for the solutions of stochastic differential equations with non–regular coefficients. The strong points of the method are its simplicity, flexibility and the dimension–free nature of the regularity estimates. Indeed, the simple method has been first introduced in [15] (see also [39, 41, 42, 43]), widely extending an idea in [20], to prove existence of densities for finite dimensional projections of an infinite dimensional stochastic equation. The method has been later used in several contexts, see for instance [14, 19, 44, 45, 1].

Indeed, the method we are discussing and extending fits into the general problem of studying existence and regularity of densities for stochastic equations with non–smooth coefficients. The problem has raised some interest recently. In addition to the aforementioned [20, 15], we would like to mention the approach in [5, 7, 6] based on some analytic criteria in spaces of Orlicz type, and interpolation. In [24, 25] the author get rid of the drift with a Girsanov transformation, and then use the Malliavin calculus. Malliavin calculus is also used in [31]. An atypical method based on optimization is instead introduced in [3, 4]. Different approaches, based on an explicit representation of the density (and so giving typically results as lower and upper bounds on the density, rather than regularity) have been given in [26, 29, 28]. Finally, [34] proves Malliavin differentiability of solutions of stochastic equations with non–differentiable coefficients. For a PDE approach see for instance [32].

In this paper, as well as in most of the aforementioned papers (possibly with non–Gaussian noise), we focus on the following "toy problem",

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t$$

and in Section 2 we illustrate our simple method for the existence of a density for the solutions of the equation above, under the assumptions $b \in L^\infty(\mathbb{R}^d)$, $\sigma \in C_b^\beta(\mathbb{R}^d; \mathbb{R}^{d \times d'})$,

for some $\beta \in (0,1)$, $d, d' \geq 1$, and $\sigma(y)\sigma(y)^\star \geq \delta I$, with $\delta > 0$ (weaker assumptions will be discussed in the subsequent sections). The method uses the idea of an auxiliary (much) simpler process introduced in [20], together with a smoothing lemma (see Appendix A) for duality in Besov spaces.

In this paper we use the simple method illustrated in Section 2 to extend the scope and the conclusions of the method itself. It is of foremost importance to notice that, even though we prove our results on the toy problem above, the results themselves are not difficult to extend to other contexts, such as equations driven by non–Gaussian noise, path–dependent equations, stochastic PDEs. We shall illustrate some of these examples in Section 7.

We first show in Section 2.2 that a higher order of approximation provides, whenever it is compatible with the regularity of the coefficients, a higher regularity for the density.

Even though the results of Section 2 assume uniform ellipticity of the diffusion, it is not difficult to adapt the method to the cases of a singular diffusion matrix as in Section 3.1, and of a hypo-elliptic diffusion. In the latter case we only discuss a simple example to illustrate the ideas. Unfortunately, but not unexpectedly, stronger regularity assumptions are necessary here, strong enough that in principle the problem may be amenable by a more standard approach such as the Malliavin calculus.

In Section 4 we prove a local version of the results of the simple method introduced in Section 2, to take into account the case when the coefficients are not globally regular or globally bounded.

In Section 5 we discuss the case of rougher coefficients (with respect to the assumptions of Section 2, namely bounded drift and Hölder covariance). In Section 5.1 we prove existence of a density for L^p drifts, with p larger than the dimension of the state spaces. Unfortunately we have not been able to lower the regularity requirements for the diffusion. We briefly discuss the issue in Section 5.2.

Finally, in Section 6 we slightly improve on the summability index of the Besov spaces for the regularity of the density. This is not yet a satisfactory result, since in the basic case of Section 2 one would expect Hölder regularity. We can achieve Hölder regularity only in dimension one (see Remark 6.4).

Our results aim to be explanatory, so we will not mix different developments of the method (for instance, we will not consider local results, as in Section 4, with rough drift, as in Section 5, or with singular diffusion, etc.).

In addition with the results explained so far for the simple toy model, in Section 7 we discuss a series of examples,

- path–dependent stochastic equations,

- improvements over [14] for a class of stochastic equations driven by α–stable noise,

- a singular stochastic PDE.

These examples should convince of the power and flexibility of our simple method. We notice in particular that the dimension–free nature of the regularity obtained is well suited for infinite dimensional problem (as well as in general for singular problems). It is sufficient indeed to apply the simple method to a series of approximating problems, to obtain uniform estimate in a Besov space with small but positive regularity. This ensures uniform integrability and thus convergence of the densities to the density of the limit problem.

Finally, in Appendix A we introduce our functional analytic framework, with the definition of all the function spaces we use throughout the paper, and we prove the crucial smoothing results for laws of random variables.

1.1 Basic notation

Throughout the paper we will use the following notations.

Given $x \in \mathbb{R}^d$ and $r > 0$, we shall denote by $B_r(x)$ the ball centred at x of radius r. Given a $m \times d$ matrix A, we shall denote by $A^\star$ its transpose, and by $|A|$ its "operator" norm, that is $|A| = \max_{x \in \mathbb{R}^d, |x|=1} |Ax|$. We shall use the symbol $\lesssim$ in our estimates to denote that the inequality holds up to a universal multiplicative numerical constant.

We shall denote by $C^\beta(\mathbb{R}^d; \mathbb{R}^m)$, respectively $C_b^\beta(\mathbb{R}^d; \mathbb{R}^m)$, the space of Hölder continuous functions, respectively bounded Hölder continuous, from $\mathbb{R}^d$ to $\mathbb{R}^m$, with Hölder exponent β. The space $C_b^\beta(\mathbb{R}^d; \mathbb{R}^m)$ is a normed space with norm $\| \cdot \|_{C_b^\beta} := \| \cdot \|_\infty + [\cdot]_{C_b^\beta}$, where $\| \cdot \|_\infty$ is the classical sup-norm, and

$$[f]_{C_b^\beta} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\beta},$$

is the Hölder seminorm. When the functions are real valued, we will adopt the shorter notation $C_b^\beta(\mathbb{R}^d)$. A separable version of these spaces can be obtained by taking the closure of bounded smooth function with respect to the norm $\| \cdot \|_{C_b^\beta}$.

More generally we shall use the *Zygmund spaces* $\mathscr{C}_b^\alpha(\mathbb{R}^d)$, that can be defined as the closure of bounded smooth functions with respect to the norm

$$\| \cdot \|_{\mathscr{C}_b^\alpha} := \| \cdot \|_\infty + [\cdot]_{\mathscr{C}^\alpha},$$

where

$$[f]_{\mathscr{C}_b^\alpha} := \left\| h \mapsto \frac{\|\Delta_h^m f\|_\infty}{|h|^s} \right\|_\infty.$$

Here the discrete increments are defined as,

$$(\Delta_h^1 f)(x) := f(x + h) - f(x),$$

$$(\Delta_h^n f)(x) := \Delta_h^1(\Delta_h^{n-1} f)(x) = \sum_{j=0}^n (-1)^{n-j} \binom{n}{j} f(x + jh). \tag{1.1}$$

In Appendix A we will define, in greater generality, Besov and Triebel–Lizorkin spaces. In particular we shall see that $\mathscr{C}_b^\alpha(\mathbb{R}^d) = B_{\infty,\infty}^\alpha(\mathbb{R}^d)$. We refer to Appendix A for a thorough introduction to these spaces.

2 The core idea

In this section we introduce the core idea around the simple method which is the main theme of the paper. The idea appears implicitly in [15] (and later, explicitly, in [14]). We repeat it here so that it will be the starting point of our improvements. We will moreover make explicit the dependence on the time when the density is evaluated and on the initial condition.

To get the gist of the idea, we focus on the simple toy model,

$$dX_t = b(X_t) \, dt + \sigma(X_t) \, dB_t, \tag{2.1}$$

where $b \in L^\infty(\mathbb{R}^d)$, $\sigma \in C_b^\beta(\mathbb{R}^d; \mathbb{R}^{d \times d'})$, with $\beta \in (0, 1)$, and $(B_t)_{t \geq 0}$ is a d'-dimensional Brownian motion. Assume moreover that

there is $\delta > 0$ such that $\quad det(\sigma(y)\sigma^\star(y)) \geq \delta > 0 \quad$ for all $y \in \mathbb{R}^d$. $\tag{2.2}$

In other words the diffusion coefficient is non–degenerate.

Our basic tool here is the smoothing Lemma A.1. Fix an integer $m \geq 1$ large and a function ϕ in $\mathscr{C}_b^\alpha(\mathbb{R}^d)$, with $\alpha \in (0, 1)$ to be chosen later. Our aim is to estimate

$\mathbb{E}[\Delta_h^m \phi(X_t)]$ and capture the regularizing effect of the density. To this end, consider a number $0 < \epsilon < t \wedge 1$ and the auxiliary process

$$Y_s^\epsilon = \begin{cases} X_s, & s \leq t - \epsilon, \\ X_{t-\epsilon} + \sigma(X_{t-\epsilon})(B_s - B_{t-\epsilon}), & s \geq t - \epsilon. \end{cases} \tag{2.3}$$

We decompose the quantity $\mathbb{E}[\Delta_h^m \phi(X_t)]$ in two terms, the *approximation error*,

$$\textbf{\textcolor{red}{Ae}} := \big| \mathbb{E}[\Delta_h^m \phi(X_t)] - \mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon)] \big|, \tag{2.4}$$

and the *probabilistic estimate*,

$$\textbf{\textcolor{red}{Pe}} := \big| \mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon)] \big|. \tag{2.5}$$

We will address (and denote) the two fundamental quantities we have defined as **Ae** and **Pe** in the rest of the paper.

For the first term we use the regularity of the test function ϕ,

$$\textbf{\textcolor{red}{Ae}} = \big| \mathbb{E}[\Delta_h^m \phi(X_t) - \Delta_h^m \phi(Y_t^\epsilon)] \big| \lesssim [\phi]_{\mathscr{C}_b^\alpha} \mathbb{E}[|X_t - Y_t^\epsilon|^\alpha] \lesssim [\phi]_{\mathscr{C}_b^\alpha} \mathbb{E}[|X_t - Y_t^\epsilon|^2]^{\frac{\alpha}{2}},$$

where

$$X_t - Y_t^\epsilon = \int_{t-\epsilon}^t b(X_r)\, dr + \int_{t-\epsilon}^t (\sigma(X_r) - \sigma(X_{t-\epsilon}))\, dB_r$$

and, by standard estimates on stochastic equations, we have,

$$\begin{aligned} \big| \mathbb{E}[|X_t - Y_t^\epsilon|^2] \big| &\lesssim \|b\|_{L^\infty}^2 \epsilon^2 + [\sigma]_{C^\beta}^2 \int_{t-\epsilon}^t \mathbb{E}[|X_r - X_{t-\epsilon}|^{2\beta}]\, dr \\ &\lesssim \|b\|_{L^\infty}^2 \epsilon^2 + [\sigma]_{C^\beta}^2 (\|b\|_{L^\infty}^2 + \|\sigma\|_{L^\infty}^2)^\beta) \epsilon^{1+\beta}, \end{aligned} \tag{2.6}$$

or, in other words, $\textbf{\textcolor{red}{Ae}} \lesssim \epsilon^{\frac{1}{2}\alpha(1+\beta)}$. For the second term we condition over the history $\mathscr{F}_{t-\epsilon}$ up to time $t - \epsilon$, as in [20],

$$\textbf{\textcolor{red}{Pe}} = \big| \mathbb{E}[\mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon)|\mathscr{F}_{t-\epsilon}]] \big| = \big| \mathbb{E}[\mathbb{E}[\Delta_h^m \phi(y + \sigma(y)\tilde{B}_\epsilon)]_{y=X_{t-\epsilon}}] \big|,$$

since, given $\mathscr{F}_{t-\epsilon}$, Y_t^ϵ is a Brownian motion, independent of $\mathscr{F}_{t-\epsilon}$, with starting point $X_{t-\epsilon}$ and covariance matrix $\sigma(X_{t-\epsilon})\sigma^\star(X_{t-\epsilon})$. Using the second formula in (1.1) and a change of variables,

$$\begin{aligned} \big| \mathbb{E}[\Delta_h^m \phi(y + \sigma(y)\tilde{B}_\epsilon)] \big| &= \Big| \int_{\mathbb{R}^d} \phi(y + x)\Delta_{-h}^m g_{\sigma(y)}(\epsilon, x)\, dx \Big| \\ &\leq \|\phi\|_{L^\infty} \|\Delta_{-h}^m g_{\sigma(y)}(\epsilon)\|_{L^1}, \end{aligned} \tag{2.7}$$

where $g_{\sigma(y)}(\epsilon)$ is the density of a Brownian motion with covariance matrix $\sigma(y)\sigma^\star(y)$ at time ϵ. It is elementary now to show that $\|\Delta_{-h}^m g_{\sigma(y)}(\epsilon)\|_{L^1} \leq c(1 \wedge (|h|/\sqrt{\epsilon}))^m$, with a number c that depends on $det(\sigma(y)\sigma^\star(y))^{-1}$ (and $\|\sigma\|_{L^\infty}$), and thus is uniformly bounded with respect to y by our non–degeneracy assumption. In conclusion, $\textbf{\textcolor{red}{Pe}} \lesssim \|\phi\|_{L^\infty}(|h|/\sqrt{\epsilon})^m$, and

$$\big| \mathbb{E}[\Delta_h^m \phi(X_t)] \big| \lesssim \|\phi\|_{\mathscr{C}_b^\alpha}\big(\epsilon^{\frac{1}{2}\alpha(1+\beta)} + \big(1 \wedge \tfrac{|h|}{\sqrt{\epsilon}}\big)^m\big).$$

An optimization in ϵ suggests the choice $\epsilon \approx |h|^{\frac{2m}{m+\alpha(1+\beta)}}$ (the requirement $\epsilon \leq t$ will be discussed below), thus

$$\big| \mathbb{E}[\Delta_h^m \phi(X_t)] \big| \lesssim \|\phi\|_{\mathscr{C}_b^\alpha} |h|^{\frac{\alpha m(1+\beta)}{m+\alpha(1+\beta)}}. \tag{2.8}$$

For m large, the exponent of $|h|$ is about $\alpha(1 + \beta)$. The smoothing Lemma A.1 yields that X_t has a density $p_t \in B_{1,\infty}^{\alpha\beta}(\mathbb{R}^d)$. Since α can be chosen arbitrarily close to 1, we conclude that $p_t \in B_{1,\infty}^{\beta-\eta}(\mathbb{R}^d)$ for all $\eta > 0$.

Remark 2.1. The regularity conditions on the coefficients can be replaced by growth conditions and finiteness of moments. For instance, assume that

$$\mathbb{E}\big[\sup_{s\in[0,t]}|X_s|^2\big] < \infty,$$

and that $|b(x)| \lesssim |x|^2$ (this is indeed the case analyzed in [15]). Then

$$\mathbb{E}\Big|\int_{t-\epsilon}^{t} b(X_s)\,ds\Big| \lesssim \epsilon\,\mathbb{E}\big[\sup_{s\in[0,t]}|X_s|^2\big],$$

and the estimate of the *approximation error* is as above.

We wish to give a version of the above computations that takes into account the size of the pre–factor in (2.8) in terms of the time t and the initial value X_0. We state the result in a way that will be convenient for the next sections. In terms of the proposition below, the above computations correspond to $a_0 = \beta$, $\theta = 2$.

Proposition 2.2. *Consider a solution X of (2.1) with initial condition $X_0 = x \in \mathbb{R}^d$. If there are numbers $a_0 > 0$, $\theta > 0$, $K_0 \geq 1$ such that*

$$\big|\,\mathbb{E}[\Delta_h^m \phi(X_t)]\,\big| \leq K_0^{\frac{\alpha}{\theta}}\big(\epsilon^{\frac{\alpha}{\theta}(1+a_0)} + \big(\tfrac{|h|}{\epsilon^{1/\theta}}\big)^m\big)\|\phi\|_{\mathscr{C}_b^\alpha}, \tag{2.9}$$

for every $\epsilon < 1$, with $\epsilon \leq \frac{t}{2}$, every $\alpha \in (0,1)$, and every $\phi \in \mathscr{C}_b^\alpha$, then

$$\|p_x(t)\|_{B_{1,\infty}^a} \leq cK_0^{\frac{a}{\theta a_0}+\delta}(1\wedge t)^{-\frac{1+a_0}{\theta a_0}a-\delta}, \tag{2.10}$$

for every $a \in (0, a_0)$ and small $\delta > 0$, where $c = c(a, a_0, \theta, \delta)$ does not depend on x. Here $p_x(t)$ is the density of X_t.

Proof. Indeed, if $a < a_0$, choose $\alpha = \frac{a+\delta_1}{a_0}$, with $\delta_1 < a_0 - a$, so that $\alpha < 1$. Consider also $\delta_2 < 1$, to be chosen later, and choose $m \geq 1$ as the smallest integer such that

$$\frac{m}{m + \alpha(1+a_0)} \geq 1 - \delta_2.$$

Consider two cases. If $|h|^{\theta(1-\delta_2)} < t$, then we choose $\epsilon = \frac{1}{2}|h|^{\theta(1-\delta_2)}$ so that

$$\big|\,\mathbb{E}[\Delta_h^m \phi(X_t)]\,\big| \lesssim K_0^{\frac{\alpha}{\theta}}\big(|h|^{\alpha(1+a_0)(1-\delta_2)} + |h|^{m\delta_2}\big)\|\phi\|_{\mathscr{C}_b^\alpha} \lesssim K_0^{\frac{\alpha}{\theta}}|h|^{\alpha(1+a_0)(1-\delta_2)}\|\phi\|_{\mathscr{C}_b^\alpha},$$

since $\delta_2 m \geq \alpha(1+a_0)(1-\delta_2)$ by the choice of m.

If on the other hand $t \leq |h|^{\theta(1-\delta_2)}$, then we choose $\epsilon = \frac{t}{2}$, so that

$$\begin{aligned}
\big|\,\mathbb{E}[\Delta_h^m \phi(X_t)]\,\big| &\lesssim K_0^{\frac{\alpha}{\theta}}\big(t^{\frac{\alpha}{\theta}(1+a_0)} + \big(\tfrac{|h|}{t^{1/\theta}}\big)^{\alpha(1+a_0)(1-\delta_2)}\big)\|\phi\|_{\mathscr{C}_b^\alpha} \\
&\lesssim K_0^{\frac{\alpha}{\theta}}\big((1\wedge t)^{-\frac{\alpha}{\theta}(1+a_0)(1-\delta_2)}|h|^{\alpha(1+a_0)(1-\delta_2)}\big)\|\phi\|_{\mathscr{C}_b^\alpha}.
\end{aligned}$$

In either case,

$$\big|\,\mathbb{E}[\Delta_h^m \phi(X_t)]\,\big| \lesssim K_0^{\frac{\alpha}{\theta}}(1\wedge t)^{-\frac{\alpha}{\theta}(1+a_0)(1-\delta_2)}|h|^{\alpha(1+a_0)(1-\delta_2)}\|\phi\|_{\mathscr{C}_b^\alpha}.$$

If we choose $\delta_2 = \frac{\delta_1}{2\alpha(1+a_0)}$, then $\alpha(1+a_0)(1-\delta_2) - \alpha = a + \frac{\delta_1}{2}$, thus $a < \alpha(1+a_0)(1-\delta_2) - \alpha$. Since by our choice of m we also have $m > \alpha(1+a_0)(1-\delta_2)$, by the smoothing Lemma A.1,

$$\|p_x(t)\|_{B_{1,\infty}^a} \lesssim K_0^{\frac{\alpha}{\theta}}(1\wedge t)^{-\frac{\alpha}{\theta}(1+a_0)(1-\delta_2)}.$$

By the choice of δ_2,

$$\frac{\alpha}{\theta}(1+a_0)(1-\delta_2) = \frac{1+a_0}{\theta a_0}a + \frac{2+a_0}{2\theta a_0}\delta_1 = \frac{1+a_0}{\theta a_0}a + \delta,$$

if we set $\delta_1 = \frac{2\theta a_0}{2+a_0}\delta$. With these positions,

$$\frac{\alpha}{\theta} = \frac{a}{\theta a_0} + \frac{\delta_1}{\theta a_0} = \frac{a}{\theta a_0} + \frac{2}{2+a_0}\delta \leq \frac{a}{\theta a_0} + \delta.$$

This yields the exponents in (2.10).

Finally, $\delta_1 < a_0 - a$ if $\delta < \frac{1}{2\theta a_0}(a_0 - a)(2 + a_0)$, while with our positions,

$$\delta_2 = \frac{\theta a_0 \delta}{(1+a_0)(2+a_0)\frac{a}{a_0} + 2\theta(1+a_0)\delta} \leq \frac{1}{2}. \qquad \square$$

Remark 2.3 (Time dependent coefficients)**.** It is not difficult to add time regularity to the coefficients. We will do so for the drift in Section 5.1. As for the diffusion coefficient, if for instance $\sigma \in L^\infty(0,T; C_b^\beta)$, and if we define the auxiliary process as

$$Y_s^\epsilon = X_{t-\epsilon} + \sigma(s, X_{t-\epsilon})(B_s - B_{t-\epsilon}),$$

if $s \geq t - \epsilon$, and as in (2.3) otherwise, then both the *approximation error* and the *probabilistic estimate* can be estimated with the same power of ϵ as above, and thus the result we get is the same.

Remark 2.4 (Time regularity)**.** It is possible to obtain regularity in time of the density. This has been done in [43] in the framework of finite dimensional projections of Navier–Stokes equations. Here one can prove that

$$\|p_x(t) - p_x(s)\|_{B_{1,\infty}^\alpha} \lesssim |t-s|^{\beta/2},$$

when $\alpha + \beta < 1$, where p_x is the density of the solution started at x. The estimate of the semi–norm $[\cdot]_{B_{1,\infty}^\alpha}$ (see (A.1) for its definition) is easy, since it is elementary to see that

$$\|\Delta_h^n(p_x(t) - p_x(s))\|_{L^1} \lesssim \begin{cases} \|\Delta_h^n p_x(t)\|_{L^1} + \|\Delta_h^n p_x(s)\|_{L^1}, & |h| \ll |t-s|, \\ \|p_x(t) - p_x(s)\|_{L^1}, & |t-s| \ll |h|. \end{cases}$$

From this, using the methods we have introduced in this section, it is easy to derive that $[p_x(t) - p_x(s)]_{B_{1,\infty}^\alpha} \lesssim |t-s|^{\beta/2}$. The L^1–estimate $\|p_x(t) - p_x(s)\|_{L^1} \lesssim |t-s|^{\frac{1}{2}-}$ is the most challenging, and in [43] it has been obtained using the Girsanov transformation.

2.1 Local solutions

In principle the solution of (2.1) may not be defined over all times. In this section we want to manage the case of solutions with explosion. We will see that, under suitable assumptions on the coefficients, the solution of (2.1) has a density on the event of non-explosion.

To this end we need to give a suitable definition of the solution. We define X_t as the local solution before explosion, and a cemetery site ∞ afterwards. Under the assumption of local existence and uniqueness, this defines a Markov process (this was done already for instance in [40] in a general framework). Denote by τ_∞ the explosion time, then $\{\tau_\infty > t\} = \{|X_t| < \infty\}$. We will use these events to localize the solution at a fixed time.

Let $\eta_R(x)$ be a smooth function equal to 1 when $|x| \leq R$, and to 0 when $|x| \geq R+1$. Our computations will prove that, if μ_t is the law of X_t (possibly with an atom at ∞), then $\eta_R(x)\mu_t(dx)$ has a density with respect to the Lebesgue measure. As long as the estimate on the density of $\eta_R\mu_t$ does not depend on R, by the uniform integrability ensured by the Besov bound, also the measure $\mathbb{1}_{\{|X_t|<\infty\}}\mu_t = \mathbb{1}_{\{\tau_\infty>t\}}\mu_t$ has a density, with similar Besov bounds by semi–continuity.

In more details, fix $t > 0$, $\alpha \in (0,1)$, $\phi \in \mathscr{C}_b^\alpha$, $m \geq 1$, and h such that $|h| \leq 1$. To prove that $\eta_R(x)\mu_t(dx)$ has a density, in view of the smoothing Lemma A.1 it is sufficient to prove that

$$\left| \mathbb{E}[\Delta_h^m \phi(X_t)\eta_R(X_t)] \right| \lesssim |h|^s \|\phi\|_{\mathscr{C}_b^\alpha},$$

for some $s > \alpha$. We decompose the discrete derivative as

$$\mathbb{E}[\Delta_h^m \phi(X_t)\eta_R(X_t)] = \mathbb{E}\big[\big(\Delta_h^m \phi(X_t) - \Delta_h^m \phi(Y_t^\epsilon)\big)\eta_R(X_t)\big] \\ + \mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon)(\eta_R(X_t) - \eta_R(X_{t-\epsilon}))] + \mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon)\eta_R(X_{t-\epsilon})]$$

The first term on the right-hand-side is the *approximation error*,

$$\left| \mathbb{E}\big[\big(\Delta_h^m \phi(X_t) - \Delta_h^m \phi(Y_t^\epsilon)\big)\eta_R(X_t)\big] \right| \lesssim [\phi]_{\mathscr{C}_b^\alpha} \mathbb{E}[|X_t - Y_t^\epsilon|^\alpha \eta_R(X_t)],$$

and the final estimate depends only on the regularity of the coefficients. The third term is the *probabilistic estimate*,

$$\left| \mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon)\eta_R(X_{t-\epsilon})] \right| = \left| \mathbb{E}\big[\eta_R(X_{t-\epsilon})\mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon) \,|\, \mathscr{F}_{t-\epsilon}]\big] \right| \\ \leq \mathbb{E}\big[|\mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon) \,|\, \mathscr{F}_{t-\epsilon}]|\big]$$

that can be estimated as in the previous section. Here $\mathscr{F}_{t-\epsilon}$ is the σ–field of events before time $t - \epsilon$.

Finally, the new term that accounts for explosion can be controlled as follows,

$$\left| \mathbb{E}[\Delta_h^m \phi(Y_t^\epsilon)(\eta_R(X_t) - \eta_R(X_{t-\epsilon}))] \right| \lesssim [\phi]_{\mathscr{C}_b^\alpha} |h|^\alpha \mathbb{E}[|\eta_R(X_t) - \eta_R(X_{t-\epsilon})|].$$

Moreover, we have that

$$|\eta_R(X_t) - \eta_R(X_{t-\epsilon})| \leq |X_t - X_{t-\epsilon}|\eta_{R+1}(X_t),$$

if both X_t and $X_{t-\epsilon}$ are smaller than $R + 1$ or larger than $R + 1$. In case $X_{t-\epsilon} < R + 1$ and $X_t \geq R + 1$,

$$|\eta_R(X_t) - \eta_R(X_{t-\epsilon})| = |\eta_R(X_{t-\epsilon})\eta_{R+1}(X_{t-\epsilon})| \\ = |\eta_R(X_{\tau_{R+1}}) - \eta_R(X_{t-\epsilon})|\eta_{R+1}(X_{t-\epsilon}) \leq |X_{\tau_{R+1}} - X_{t-\epsilon}|\eta_{R+1}(X_{t-\epsilon}),$$

and in both cases one can use the equation to obtain an estimate of the above quantities.

2.2 More regularity - I

By Proposition 2.2 it is immediately clear that in order to obtain more regularity for the density we need to improve the *approximation error*. If we think of the auxiliary process (2.3) as a (very) basic numerical approximation of the original process X, then we need to use a more refined numerical method. To do this it is necessary to have smoother coefficients. Moreover, due to the non–anticipative nature of the estimate of the *probabilistic estimate*, the numerical method needs to be explicit.

Let us focus, for the sake of clarity, on the drift term, namely consider

$$dX_t = b(X_t)\,dt + dB_t,$$

with $b \in C_b^\beta(\mathbb{R}^d)$. Similar considerations and conclusions are likewise possible for the diffusion coefficient.

To exploit the additional regularity of the drift b, it is meaningful to define the auxiliary process differently. Define the auxiliary process Y^ϵ in a general way as

$$Y_s^\epsilon = X_{t-\epsilon} + \int_{t-\epsilon}^s A_r\,dr + B_s - B_{t-\epsilon}, \qquad s \geq t - \epsilon, \tag{2.11}$$

and this time take $A_r = b(X_{t-\epsilon})$, rather than $A_r = 0$ as we have done before. The *probabilistic estimate* does not change, since we have only changed the mean, but in a way that is measurable with respect to the information at time $t - \epsilon$. To evaluate the *approximation error*, consider

$$\mathbb{E}[|X_t - Y_t^\epsilon|] = \mathbb{E}\left|\int_{t-\epsilon}^t (b(X_s) - b(X_{t-\epsilon}))\, ds\right| \le [b]_{C^\beta}\mathbb{E}\int_{t-\epsilon}^t |X_s - X_{t-\epsilon}|^\beta\, ds \lesssim \epsilon^{1+\frac{\beta}{2}}, \quad (2.12)$$

since, as in Section 2, $\mathbb{E}[|X_s - X_{t-\epsilon}|] \lesssim \sqrt{\epsilon}$. Thus, **Ae** $\lesssim [\phi]_{\mathscr{C}_b^\alpha}\epsilon^{\frac{\alpha}{2}(2+\beta)}$ and Proposition 2.2 ensures that the density is in $B_{1,\infty}^a$ for $a < 1 + \beta$.

It is now quantitatively clear that the more regular is b, the more regularity we obtain for the density. The (obvious) key idea is to find a good approximation of X. A natural candidate for Y are the Picard iterations for our equation. The next step to improve the regularity of the density is to choose an auxiliary process that ensures a smaller estimate of $|X_t - Y_t^\epsilon|$. Since in (2.12) the estimates depends on $|X_s - X_{t-\epsilon}|$, and in turn the size of this term corresponds to the size of the Brownian increments, a way to improve the difference might be to define the auxiliary process (2.11) with $A_s = b(X_{t-\epsilon} + B_s - B_{t-\epsilon})$ when $s \ge t - \epsilon$. Unfortunately this makes the term in the *probabilistic estimate* "anticipative" (with respect to the time $t - \epsilon$). If we are ready to consider an "anticipative" term, we are faced with the difficulty that it becomes now difficult to evaluate the law of the terms that contribute to the *probabilistic estimate*. The simplicity of the *probabilistic estimate* is indeed an implicit requirement of the simple method we are illustrating.

A workaround that, albeit "anticipative", keeps the term that will contribute to the *probabilistic estimate* simple, can be considered if b is more regular, namely if $b \in C_b^{1+\beta}(\mathbb{R}^d)$, by looking at iterated integrals of Brownian motion. Additional regularity of b cannot be avoided in general, and a way to see this is to look at the discussion on the Fokker–Planck equation at the beginning of Section 6. With this in mind, take

$$A_r = b(X_{t-\epsilon}) + b'(X_{t-\epsilon})(B_r - B_{t-\epsilon}) \quad (2.13)$$

in (2.11), then

$$\begin{aligned}
b(X_s) - A_s &= b(X_s) - b(X_{t-\epsilon}) - b'(X_{t-\epsilon})(B_r - B_{t-\epsilon}) \\
&= b'(X_{t-\epsilon})(X_s - X_{t-\epsilon}) + O(|X_s - X_{t-\epsilon}|^{1+\beta}) - b'(X_{t-\epsilon})(B_r - B_{t-\epsilon}) \\
&= b'(X_{t-\epsilon})\int_{t-\epsilon}^t b(X_r)\, dr + O(|X_s - X_{t-\epsilon}|^{1+\beta}),
\end{aligned}$$

thus $\mathbb{E}[|b(X_s) - A_s|] \lesssim \epsilon^{\frac{1}{2}(1+\beta)}$, therefore **Ae** $\lesssim [\phi]_{\mathscr{C}_b^\alpha}\epsilon^{\frac{\alpha}{2}(3+\beta)}$. With this choice of the auxiliary process, we see that the term providing the *probabilistic estimate* is again amenable to our analysis: given $y = X_{t-\epsilon}$, we have that

$$Y_t^\epsilon = y + \epsilon b(y) + \tilde{B}_\epsilon + b'(y)\int_0^\epsilon \tilde{B}_s\, ds,$$

where $\tilde{B}_r = B_{t-\epsilon+r} - B_{t-\epsilon}$ is independent of the history until time $t - \epsilon$. The random variable Y_t^ϵ is conditionally Gaussian with variance

$$\mathrm{Var}(Y_t^\epsilon|X_{t-\epsilon}) = \epsilon + b'(X_{t-\epsilon})\epsilon^2 + \frac{1}{3}b'(X_{t-\epsilon})^2\epsilon^3 \gtrsim \epsilon,$$

so that the *probabilistic estimate* has the same order in ϵ. In conclusion the density is in $B_{1,\infty}^a$ for all $a < \beta + 2$.

With additional regularity of b one can consider further stochastic Taylor approximations to improve the estimates. We refer to the classical [27, Chapter 10] for some possibilities.

As a concluding remark we notice that, as long as the coefficients have derivatives, one can more easily resort to Malliavin calculus. These consideration will become useful though in Section 3.2.

3 Singular diffusion coefficient

In this section we wish to consider the case when the diffusion matrix is singular. We will briefly discuss two cases. The first is when the diffusion matrix is non-invertible. The second case is the so–called hypo-elliptic case, when the diffusion matrix is singular, but the noise is transmitted to noise-less components by the drift.

3.1 Singular diffusion coefficient

First of all, we notice that in general we do not need a uniform estimate such as (2.2), since under the assumption of non–singularity one can use the results of Section 4. In this section we wish to investigate conditions that can ensure the existence of a density around points where the diffusion coefficient is zero or non–invertible.

An observation from [14] indeed says that the method from Section 2 can be slightly modified to take into account the case when the diffusion coefficient is not invertible. The idea is as in Section 2.1. Consider a solution X_t of (2.1) and let μ_t be its law. Fix an integer $m \geq 1$, then the smoothing Lemma A.1 is applied to the measure $|\sigma^{-1}(x)|^{-m}\mu_t(dx)$, where $|\cdot|$ is the operator norm of a matrix, see Section 1.1. If this measure has a density, then μ_t has a density on $\{y : \sigma(y) \text{ invertible}\}$.

To this end, it is sufficient to prove that

$$\Big| \mathbb{E}\Big[\frac{\Delta_h^m \phi(X_t)}{|\sigma^{-1}(X_t)|^m}\Big] \Big| \lesssim |h|^s \|\phi\|_{\mathscr{C}_b^\alpha}.$$

The decomposition here is

$$\mathbb{E}[|\sigma^{-1}(X_t)|^{-m}\Delta_h^m \phi(X_t)] = \mathbb{E}[\Delta_h^m \phi(X_t)(|\sigma^{-1}(X_t)|^{-m} - |\sigma^{-1}(X_{t-\epsilon})|^{-m})] +$$
$$+ \mathbb{E}[|\sigma^{-1}(X_{t-\epsilon})|^{-m}(\Delta_h^m \phi(X_t) - \Delta_h^m \phi(Y_t^\epsilon))] +$$
$$+ \mathbb{E}[|\sigma^{-1}(X_{t-\epsilon})|^{-m}\Delta_h^m \phi(Y_t^\epsilon)]$$

The first term can be controlled as

$$\big| \mathbb{E}[\Delta_h^m \phi(X_t)(|\sigma^{-1}(X_t)|^{-m} - |\sigma^{-1}(X_{t-\epsilon})|^{-m})] \big| \lesssim [\phi]_{\mathscr{C}_b^\alpha}|h|^\alpha[\sigma]_{C^\beta}\mathbb{E}[|X_t - X_{t-\epsilon}|^\beta]$$
$$\lesssim [\phi]_{\mathscr{C}_b^\alpha}|h|^\alpha\epsilon^{\frac{\beta}{2}},$$

while the second term plays the role of the *approximation error*, thus

$$\mathsf{Ae} \lesssim \|\sigma\|_{L^\infty}^m [\phi]_{\mathscr{C}_b^\alpha}\mathbb{E}[|X_t - Y_t^\epsilon|^\alpha] \lesssim [\phi]_{\mathscr{C}_b^\alpha}\epsilon^{\frac{\alpha}{2}(1+\beta)}.$$

Finally, the third term is the *probabilistic estimate*, and by conditioning,

$$\mathsf{Pe} = \big| \mathbb{E}\big[\big(|\sigma^{-1}(y)|^{-m}\mathbb{E}[\Delta_h^m \phi(y + \sigma(y)\tilde{B}_\epsilon)]\big)_{y=X_{t-\epsilon}}\big] \big| \lesssim \|\phi\|_{L^\infty}\epsilon^{-\frac{m}{2}}|h|^m.$$

The contribution of the first term is negligible. Indeed, if the method is successful, then $|h| \lesssim \sqrt{\epsilon}$, thus $|h|^\alpha\epsilon^{\beta/2} \lesssim \epsilon^{\frac{1}{2}(\alpha+\beta)} \leq \epsilon^{\frac{\alpha}{2}(1+\beta)}$, since $\alpha < 1$. In conclusion, we end up with the same estimate as in Section 2,

$$\Big| \mathbb{E}\Big[\frac{\Delta_h^m \phi(X_t)}{|\sigma^{-1}(X_t)|^m}\Big] \Big| \lesssim \|\phi\|_{\mathscr{C}_b^\alpha}(\epsilon^{\frac{\alpha}{2}(1+\beta)} + \epsilon^{-\frac{m}{2}}|h|^m),$$

so that we obtain the same conclusion, but on the measure $|\sigma^{-1}(x)|^{-m}\mu_t(dx)$. This proves the following result.

Proposition 3.1. *Let $b \in L^\infty(\mathbb{R}^d)$ and $\sigma \in C_b^\beta(\mathbb{R}^d : \mathbb{R}^{d \times d'})$. Let X be a solution of* (2.1) *and $t > 0$. Then X_t has a density with respect to the Lebesgue measure on the set $\{y : \sigma(y) \text{ invertible}\}$.*

Clearly, the regularity assumptions on the coefficients can be relaxed (for instance along the lines of Section 5).

We now wish to find some simple condition for the existence of a density on $\mathbb{R}^d$ even in the case of non–invertible diffusion coefficient. For the sake of simplicity, we consider our problem (2.1) without drift, namely,

$$dX_t = \sigma(X_t)\,dB_t.$$

Proposition 3.2. *Assume that $\sigma \in C_b^\beta(\mathbb{R}^d; \mathbb{R}^{d \times d})$, let $t > 0$ and assume that there is $\gamma > 0$ such that*

$$\mathbb{E}\Big[\int_{\frac{t}{2}}^t |\sigma(X_s)^{-1}|^\gamma\,ds\Big] < \infty.$$

Then X_t has a density p in $\mathbb{R}^d$ and there is $a_0 = a_0(\beta, \gamma) > 0$ such that $p \in B_{1,\infty}^a$ for all $a < a_0$.

Proof. Consider the auxiliary process as in (2.3). The *approximation error* is as in Section 2, thus **Ae** $\lesssim \epsilon^{\frac{\alpha}{2}(1+\beta)}[\phi]_{\mathscr{C}_b^\alpha}$.

For the *probabilistic estimate*, if $\tilde{B}_s = B_{t-\epsilon+s} - B_{t-\epsilon}$, then by conditioning,

$$\textbf{Pe} = \big|\,\mathbb{E}\big[\mathbb{E}[\Delta_h^m \phi(y + \sigma(y)\tilde{B}_\epsilon)]_{y=X_{t-\epsilon}}\big]\,\big|$$

Now, if $\sigma(y)$ is invertible,

$$\big|\,\mathbb{E}[\Delta_h^m \phi(y + \sigma(y)\tilde{B}_\epsilon)]\,\big| \lesssim \Big(1 \wedge \frac{|\sigma(y)^{-1}|\,|h|}{\sqrt{\epsilon}}\Big)^m \|\phi\|_{L^\infty},$$

and the inequality extends to $\sigma(y)$ non invertible, since by definition $|\sigma(y)^{-1}|^{-1} = \inf_{|z|=1}|\sigma(y)z|$ (and the understanding $\frac{1}{0} = \infty$). By the moment assumption,

$$\textbf{Pe} \lesssim \mathbb{E}\Big[\Big(1 \wedge \frac{|\sigma(X_{t-\epsilon})^{-1}|\,|h|}{\sqrt{\epsilon}}\Big)^m\Big]\|\phi\|_{L^\infty} \lesssim \mathbb{E}[|\sigma(X_{t-\epsilon})^{-1}|^\gamma]\Big(\frac{|h|}{\sqrt{\epsilon}}\Big)^\gamma \|\phi\|_{L^\infty}.$$

Choose now $0 < \lambda_1 < \lambda_2 < 1$ and consider the two terms **Ae** and **Pe** (here we have replaced ϵ with δ for convenience),

$$\big|\,\mathbb{E}[\Delta_h^m \phi(X_t)]\,\big| \lesssim \Big(\delta^{\frac{\alpha}{2}(1+\beta)} + \frac{|h|^\gamma}{\delta^{\gamma/2}}\mathbb{E}[|\sigma(X_{t-\delta})^{-1}|^\gamma]\Big)\|\phi\|_{\mathscr{C}_b^\alpha}.$$

Integrate the above inequality over $\delta \in (\lambda_1\epsilon, \lambda_2\epsilon)$ to obtain,

$$\big|\,\mathbb{E}[\Delta_h^m \phi(X_t)]\,\big| \lesssim \Big(\epsilon^{\frac{\alpha}{2}(1+\beta)} + \frac{|h|^\gamma}{\epsilon^{\gamma/2}}\mathbb{E}\int_{\frac{t}{2}}^t |\sigma(X_r)^{-1}|^\gamma\,dr\Big)\|\phi\|_{\mathscr{C}_b^\alpha}.$$

The proof can now be concluded with computations similar to those in Proposition 2.2. $\square$

Example 3.3. The result we have obtained is clearly non–optimal. Let us consider a simple example to figure out the issue.

Consider $X_t = B_t^2$, where $(B_t)_{t \geq 0}$ is a one–dimensional Brownian motion. It is known that $(X_t)_{t \geq 0}$ solves

$$dX_t = dt + 2\sqrt{X_t}\,dW_t,$$

with initial condition $X_0 = 0$, where $(W_t)_{t \geq 0}$ is another Brownian motion. The function $\sigma(x) = \sqrt{x}$ is clearly $\frac{1}{2}$–Hölder, and

$$\mathbb{E}[\sigma(X_t)^\gamma] = \mathbb{E}[|B_t|^{-\gamma}] < \infty$$

if and only if $\gamma < 1$. The previous considerations ensure that there is a density p_t for X_t on $\mathbb{R}$ and that has regularity $B_{1,\infty}^\alpha$, with $\alpha < \frac{2}{3}(\frac{5}{2} - 2\sqrt{\frac{3}{2}}) \ll \frac{1}{2}$. On the other hand here the density is explicitly known, with $p_t(x) = (2\pi t x)^{-\frac{1}{2}} e^{-x/(2t)}$. It is not difficult to prove that actually $p_t \in B_{1,\infty}^\alpha$, $\alpha \leq \frac{1}{2}$. The correct regularity can be recovered, at least away from the zero of the diffusion coefficient, using the methods of Section 4.

3.2 A hypo-elliptic example

Here we aim to consider a hypo-elliptic problem, that is a problem where the diffusion coefficient is not invertible, but the effect of the noise is propagated by the drift. We will only show a (very) elementary example, to convince that the simple method we are illustrating is effective also in this framework. It would be difficult, though, to give a general result.

Consider the following problem, with $X = (X^1, X^2)$,

$$\begin{cases} dX^1 = b_1(X^1, X^2) \, dt + dB, \\ dX^2 = b_2(X^1, X^2) \, dt, \end{cases} \tag{3.1}$$

where B is a one–dimensional standard Brownian motion. Here the diffusion coefficient is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

which is nowhere invertible, thus none of the results of the previous section is available. The key point is clearly the definition of the auxiliary process for the second component. The basic definition in (2.3), as well as the one in (2.11) with $A_r^2 = b_2(X_{t-\epsilon})$, are not suitable, because there is no term that propagates the noise. This is fundamental in view of the *probabilistic estimate*. Thus, we consider the auxiliary process in the second component as in (2.11), but with the process A defined similarly to (2.13). Moreover, only variations in the first component (the one that contains the noise) are taken into account. In other words,

$$Y_s^{\epsilon,2} = X_{t-s}^2 + \int_{t-\epsilon}^s \big(b_2(X_{t-\epsilon}) + \partial_{x_1} b_2(X_{t-\epsilon})(Y_r^1 - X_{t-\epsilon}^1)\big) \, dr,$$

for $s \geq t - \epsilon$. Here we need to assume that b_2 is differentiable. To ensure that the first component will "transfer" the regularizing effect of the noise to the second component, we need to add the "hypo-elliptic" assumption

$$\partial_{x_1} b_2(x) \neq 0.$$

We turn to the definition of the first component of the auxiliary process. There is a hidden requirement for the *approximation error*. Indeed, we will see in Proposition 3.4 below that **Pe** $\lesssim \epsilon^{-\frac{3}{2}m} |h|^m$. This is due to the fact that the random variable in the second component is smoother, thus gives a stronger singularity in time. To guess the right size of the *approximation error*, we see that, with the *probabilistic estimate* as above, if **Ae** $\lesssim \epsilon^{\alpha q}$, then by the simple optimization we have seen in Proposition 2.2, we need $q > \frac{3}{2}$. In other words, we need to choose an auxiliary process that approximates the

original process to the order $\epsilon^{3/2}$. Thus we assume $b_1 \in C_b^{1+\beta}(\mathbb{R}^d)$, with $\beta > 0$, and set for $s \geq t - \epsilon$,

$$Y_s^{\epsilon,1} = X_{t-\epsilon}^1 + \int_{t-\epsilon}^s \left(b_1(X_{t-\epsilon}) + \partial_{x_1} b_1(X_{t-\epsilon})(B_s - B_{t-\epsilon}) \right) ds + B_s - B_{1-\epsilon}.$$

Proposition 3.4. *Assume that $b_1 \in C_b^{1+\beta}(\mathbb{R}^d)$, with $\beta \in (0,1)$, and $b_2 \in C_b^1(\mathbb{R}^d)$, with $|\partial_{x_1} b_2(x)| \geq c_0 > 0$. Then for every $t > 0$ each solution X_t of problem* (3.1) *has a density in $B_{1,\infty}^a$ for every $a < \frac{1}{3}\beta$.*

We notice that on the one hand the result is not very satisfactory, since coefficients are assumed differentiable (although one can be more careful on which directional derivatives are really necessary), thus one can derive the existence of a density by the existence of the Malliavin derivative and the non–degeneracy of the Malliavin matrix. The last statement would follow from our hypo-elliptic assumption. On the other hand, at this level the simple method of this paper still provides the additional minimal regularity that ensures that smooth approximations of the solution have uniformly integrable (thus, weakly compact) densities.

Proof of Proposition 3.4. We first estimate the *approximation error*. We easily have

$$\mathbb{E}[|X_s^1 - X_{t-\epsilon}^1|] \leq \mathbb{E}[|B_s - B_{t-\epsilon}|] + \mathbb{E}\int_{t-\epsilon}^s |b_1(X_r)|\, dr \lesssim \sqrt{\epsilon},$$

$$\mathbb{E}[|X_s^2 - X_{t-\epsilon}^2|] \leq \mathbb{E}\int_{t-\epsilon}^s |b_2(X_r)|\, dr \leq \|b\|_{L^\infty}\epsilon,$$

thus,

$$\mathbb{E}[|X_t^1 - Y_t^1|] = \mathbb{E}\Big| \int_{t-\epsilon}^t \big(b_1(X_s) - b_1(X_{t-\epsilon}) - \partial x_1 b_1(X_{t-\epsilon})(X_s - X_{t-\epsilon}) + $$
$$+ \partial x_1 b_1(X_{t-\epsilon})(X_s - X_{t-\epsilon} - (B_s - B_{t-\epsilon})) \big)\, ds$$
$$\lesssim \epsilon^{\frac{1}{2}(\beta+3)}.$$

Likewise,

$$\mathbb{E}[|X_t^2 - Y_t^{\epsilon,2}|] \leq \mathbb{E}\int_{t-\epsilon}^t |b_2(X_s) - b_2(X_{t-\epsilon}) - \partial_{x_1} b_2(X_{t-\epsilon})(Y_s^1 - X_{t-\epsilon}^1)|\, ds$$

$$\leq \mathbb{E}\int_{t-\epsilon}^t \big(|\partial_{x_1} b_2(X_{t-\epsilon})(X_s^1 - Y_s^1) + \partial_{x_2} b_2(X_{t-\epsilon})(X_s^2 - X_{t-\epsilon}^2) \big)\, ds$$

$$\leq \|\nabla b\|_{L^\infty} \int_{t-\epsilon}^t \mathbb{E}[|X_s^1 - Y_s^1| + |X_s^2 - X_{t-\epsilon}^2|]\, ds$$

$$\lesssim \epsilon^2.$$

In conclusion the *approximation error* is $\mathbf{Ae} \lesssim [\phi]_{\mathscr{C}_b^\alpha} \epsilon^{\frac{\alpha}{2}(\beta+3)}$.

We turn to the *probabilistic estimate*. Conditional to the history up to time $t - \epsilon$, we can write

$$Y_t^1 = A_1 + \tilde{B}_\epsilon + A_2\tilde{C}_\epsilon, \qquad Y_t^2 = A_3 + A_4\tilde{C}_\epsilon \qquad \text{with} \qquad \tilde{C}_\epsilon := \int_0^\epsilon \tilde{B}_s\, ds,$$

where A_1, A_2, A_3, A_4 are measurable with respect to the past (of $t - \epsilon$), and $\tilde{B}_s = B_{s+t-\epsilon} - B_{t-\epsilon}$. In particular, $A_4 = \partial_{x_1} b_2(X_{t-\epsilon})$, thus $|A_4| \geq c_0 > 0$. The random variable $(\tilde{B}_\epsilon, \tilde{C}_\epsilon)$ is Gaussian, and its covariance matrix has eigenvalues of order ϵ and ϵ^3. Thus

$$\big| \mathbb{E}[\Delta_h \phi(Y_t)] \big| \lesssim \frac{|h|^m}{\epsilon^{\frac{3}{2}m}}[\phi]_{\mathscr{C}_b^\alpha},$$

and in conclusion

$$\left| \mathbb{E}[\Delta_h^m \phi(X_1)] \right| \lesssim [\phi]_{\mathscr{C}_b^\alpha} \epsilon^{\frac{1}{2}\alpha(3+\beta)} + \|\phi\|_{L^\infty} \epsilon^{-\frac{3}{2}m} |h|^m.$$

Our standard computations (those in Proposition 2.2), ensure that the density of X_t is in $B_{1,\infty}^a$ for all $a < \frac{1}{3}\beta$. $\qquad\square$

Remark 3.5. The regularity obtained in the previous proposition is the regularity of the joint density of X_t^1 and X_t^2. It is not difficult to check that the density of the random variable X_t^1 is more regular.

4 Local estimates

In this section we wish to obtain local estimates on the density. This might be useful if for instance

- we only have local regularity of the coefficients,
- or if the diffusion coefficient is non–zero or non–singular only in some part of the space,

and so on.

To this end, let us consider our toy model (2.1). We will localize the problem as in [13, Theorem 2.4], and then apply the method to the localized problem.

Theorem 4.1. *Assume that there is a ball $D \subset \mathbb{R}^d$ such that $b \in L^\infty(D)$, $\sigma \in C^\beta(D)$, for some $\beta > 0$ and that $det(\sigma(y)\sigma(y)^\star) > 0$ on D.*

If X^x is solution of (2.1), with initial condition $x \in \mathbb{R}^d$, then for every $t > 0$ the random variable X_t^x has a density $p_x(t)$ in a smaller ball $D' \subseteq D$, and $p_x(t) \in B_{1,\infty}^a(D')$ for every $a < \beta$.

We will devote the rest of the section to the proof of the theorem. Prior to this, we give a global version of the above result.

Corollary 4.2. *Assume that $b \in L_{loc}^\infty(\mathbb{R}^d)$, that σ is Hölder continuous on bounded sets, and that $det(\sigma(y)\sigma(y)^\star) > 0$ for all $y \in \mathbb{R}^d$. If X^x is solution of (2.1), with initial condition $x \in \mathbb{R}^d$, then for every $t > 0$ the random variable X_t^x has a density $p_x(t)$ on $\mathbb{R}^d$. Moreover, if $D \subset \mathbb{R}^d$ is a ball and $\sigma \in C^\beta(D)$, then $p_x(t) \in B_{1,\infty}^a(D')$ for every $a < \beta$ and every smaller ball $D' \subset D$.*

Proof. Notice that a probability measure that is locally absolutely continuous, is absolutely continuous, thus it is sufficient to prove the existence of a density in balls. This is immediate from the previous theorem. $\qquad\square$

4.1 Proof of Theorem 4.1

Let $x_0 \in \mathbb{R}^d$ and $r > 0$ be such that $B_{6r}(x_0) \subset D$. We shall study existence and regularity of the density around x_0 of the solution X of (2.1) at some fixed time t.

4.1.1 Localization

Let $\varphi \in C^\infty(\mathbb{R}^d)$ be such that $\varphi = 1$ on $B_1(0)$ and $\mathbb{1}_{B_1(0)} \leq \varphi \leq \mathbb{1}_{B_2(0)}$, and let $\varphi_r(x) := \varphi((x - x_0)/r)$. Set

$$m(t,r) = \mathbb{E}[\varphi_r(X_t)],$$

then we see that $r \mapsto m(t,r)$ is non–decreasing with limit 1 at $r = \infty$. Assume that $m(t,r) > 0$ for all $r > 0$ (otherwise $|X_t - x_0| > 0$ a.s. and the density of X_t is equal to 0 in a neighbourhood of x_0). By the definition of φ, if $\mathrm{Supp}(f) \subset B_r(x_0)$, then

$$\mathbb{E}[f(X_t)] = \mathbb{E}[f(X_t)\varphi_r(X_t)],$$

A simple method for the existence of a density

thus, in view of the smoothing Lemma A.1, it is sufficient to prove that

$$\left| \mathbb{E}[(\Delta_h^m \phi)(X_t)\varphi_r(X_t)] \right| \lesssim |h|^\gamma \|\phi\|_{\mathscr{C}_b^\alpha}, \tag{4.1}$$

for suitable α, γ.

To localize the dynamics, we proceed again as in [13], and consider a localizing function $\eta \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$, defined as $\eta(x) = x$ in $B_{4r}(0)$, and $\eta(x) = 5x/|x|$ outside $B_{5r}(0)$. Set $\bar{b}(x) = b(\eta(x))$ and $\bar{\sigma}(x) = \sigma(\eta(x))$. It is immediate to see that $\bar{b}$ and $\bar{\sigma}$ have (globally) the same (local) regularity properties of b, σ. In particular, $\bar{\sigma} \in C_b^\beta(\mathbb{R}^d)$. Let $\bar{X}^x$ be the solution of

$$d\bar{X} = \bar{b}(\bar{X}_s)\,ds + \bar{\sigma}(\bar{X}_s)\,dB_s,$$

with initial condition $\bar{X}_0^x = x$.

4.1.2 Decomposition

Fix a small parameter $\delta \in (0,1]$, with $\delta \ll t$, and set

$$\tau_i := \{s \geq t - \delta : X_s \in B_{3r}(x_0)\},$$
$$\tau_o := \{s \geq \tau_i : X_s \notin B_{4r}(x_0)\}.$$

Set moreover

$$I_t := \{\varphi_r(X_t) > 0, \tau_i = t - \delta, \tau_o > t\},$$
$$O_t := \{\varphi_r(X_t) > 0, (\tau_i > t - \delta \text{ or } (\tau_i = t - \delta \text{ and } \tau_o \leq t))\},$$

so that if $f \in C_b(\mathbb{R}^d)$,

$$\mathbb{E}[f(X_t)\varphi_r(X_t)] = \mathbb{E}[f(X_t)\varphi_r(X_t)\mathbb{1}_{I_t}] + \mathbb{E}[f(X_t)\varphi_r(X_t)\mathbb{1}_{O_t}] \tag{4.2}$$

Consider first the second term. In [13] it is proved that

$$O_t \subset \left\{\varphi_r(X_t) > 0, \sup_{[0,\delta]} |\bar{X}_s^{X_{\tau_i}} - X_{\tau_i}| \geq r|\right\},$$

and it is easy to see (recall that $\bar{X}$ satisfies an equation with globally bounded coefficients) that for every $q \geq 1$,

$$\mathbb{P}\left[\varphi_r(X_t) > 0, \sup_{[0,\delta]} |\bar{X}_s^{X_{\tau_i}} - X_{\tau_i}| \geq r|\right] \leq \mathbb{P}\left[\tau_i \leq t, \sup_{[0,\delta]} |\bar{X}_s^{X_{\tau_i}} - X_{\tau_i}| \geq r|\right]$$

$$\leq \frac{1}{r^q}\mathbb{E}\left[\mathbb{1}_{\tau_i \leq t}\mathbb{E}\left[\sup_{[0,\delta]} |\bar{X}_s^{X_{\tau_i}} - X_{\tau_i}|^q \,|\tau_i\right]\right]$$

$$\leq \frac{1}{r^q}\sup_{x \in B_{3r}(x_0)} \mathbb{E}\left[\sup_{[0,\delta]} |\bar{X}_s^x - x|^q\right]$$

$$\lesssim \frac{1}{r^q}(\|\bar{b}\|_{L^\infty}^q + \|\bar{\sigma}\|_{L^\infty}^q)\delta^{\frac{q}{2}}.$$

Therefore, for every $q \geq 1$,

$$\left|\mathbb{E}[f(X_t)\varphi_r(X_t)\mathbb{1}_{O_t}]\right| \leq \|f\|_{L^\infty}\|\varphi\|_{L^\infty}\mathbb{P}[O_t]$$

$$\lesssim \frac{\|\varphi\|_{L^\infty}}{r^q}(\|\bar{b}\|_{L^\infty}^q + \|\bar{\sigma}\|_{L^\infty}^q)\|f\|_{L^\infty}\delta^{\frac{q}{2}}. \tag{4.3}$$

Consider next the first term in (4.2). Conditional to $X_{t-\delta}$, the random variable $\mathbb{1}_{\{\tau_i = t - \delta\}}$ is $X_{t-\delta}$–measurable, (since $\{\tau_i = t - \delta\} = \{X_{t-\delta} \in B_{3r}(x_0)\}$), thus

$$\mathbb{E}\left[f(X_t)\varphi_r(X_t)\mathbb{1}_{\{\varphi_r(X_t)>0\}}\mathbb{1}_{\{\tau_i=t-\delta\}}\mathbb{1}_{\{\tau_o>t\}}\right] =$$
$$= \mathbb{E}\left[\mathbb{1}_{\{\tau_i=t-\delta\}}\mathbb{E}[f(X_t)\varphi_r(X_t)\mathbb{1}_{\{\varphi_r(X_t)>0\}}\mathbb{1}_{\{\tau_o>t\}}\,|X_{t-\delta}]\right]. \tag{4.4}$$

On the event I_t we have that $X_{t-\delta} \in B_{3r}(x_0)$ and $X_s \in B_{4r}(x_0)$ for all $s \in [t-\delta, t]$. Therefore $X_s = X_{s+\delta-t}^{X_{t-\delta}}$ for all $s \in [t-\delta, t]$. Thus, by the Markov property, on $\{\tau_i = t - \delta\}$,

$$
\begin{aligned}
\big| \mathbb{E}\big[f(X_t)\varphi_r(X_t)\mathbb{1}_{\{\varphi_r(X_t)>0\}}\mathbb{1}_{\{\tau_o>t\}}\big|X_{t-\delta}\big]\big| &= \\
&= \big| \mathbb{E}\big[f(\bar{X}_t)\varphi_r(\bar{X}_t)\mathbb{1}_{\{\varphi_r(\bar{X}_t)>0\}}\mathbb{1}_{\{\tau_o>t\}}\big|X_{t-\delta}\big]\big| \\
&= \Big| \mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\mathbb{1}_{\{\varphi_r(\bar{X}_\delta^x)>0\}}\mathbb{1}_{\{\bar{\tau}^x>\delta\}}\big]\Big|_{x=X_{t-\delta}}\Big| \\
&\leq \sup_{x\in B_{3r}(x_0)}\Big|\mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\mathbb{1}_{\{\bar{\tau}^x>\delta\}}\big]\Big|
\end{aligned}
\tag{4.5}
$$

where $\bar{\tau}^x$ is the first exit time of $\bar{X}^x$ from $B_{4r}(x_0)$. We consider now the above expectation. If $x \in B_{3r}(x_0)$,

$$
\begin{aligned}
\mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\mathbb{1}_{\{\varphi_r(\bar{X}_\delta^x)>0\}}\mathbb{1}_{\{\bar{\tau}^x>\delta\}}\big] &= \\
&= \mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\big] - \mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\mathbb{1}_{\{\bar{\tau}^x\leq\delta\}}\big].
\end{aligned}
\tag{4.6}
$$

The analysis of the first term on the right-hand side is postponed to the next section. We focus here on the second term. Indeed,

$$
\big| \mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\mathbb{1}_{\{\bar{\tau}^x\leq\delta\}}\big]\big| \leq \|f\|_{L^\infty}\|\varphi\|_{L^\infty}\mathbb{P}[\bar{\tau}^x\leq\delta].
$$

We have

$$
\bar{X}_{\bar{\tau}^x}^x - x = \int_0^{\bar{\tau}^x} \bar{b}(\bar{X}_s^x)\,ds + \int_0^{\bar{\tau}^x} \bar{\sigma}(\bar{X}_s^x)\,dB_s =: J_1 + J_2.
$$

On $\{\bar{\tau}^x \leq \delta\}$ and for $x \in B_{3r}(x_0)$, we have $|\bar{X}_{\bar{\tau}^x}^x - x| \geq r$, thus $|J_1 + J_2| \geq r$, hence $|J_1| \geq \frac{1}{2}r$ or $|J_2| \geq \frac{1}{2}r$. Therefore, for every $q \geq 1$,

$$
\begin{aligned}
\mathbb{P}[\bar{\tau}^x \leq \delta] &\leq \mathbb{P}[\bar{\tau}^x \leq \delta, |J_1| \geq \tfrac{1}{2}r] + \mathbb{P}[\bar{\tau}^x \leq \delta, |J_2| \geq \tfrac{1}{2}r] \\
&\lesssim \frac{1}{r^q}\|\bar{b}\|_{L^\infty}^q \mathbb{E}[(\bar{\tau}^x\wedge\delta)^q] + \frac{1}{r^q}\mathbb{E}\Big[\Big|\int_0^{\bar{\tau}^x\wedge\delta}\bar{\sigma}(\bar{X}_s^x)\,dB_s\Big|^q\Big] \\
&\lesssim \frac{1}{r^q}(\|\bar{b}\|_{L^\infty}^q + \|\bar{\sigma}\|_{L^\infty}^q)\delta^{q/2},
\end{aligned}
$$

and in conclusion

$$
\big| \mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\mathbb{1}_{\{\bar{\tau}^x\leq\delta\}}\big]\big| \lesssim \frac{\|\varphi\|_{L^\infty}}{r^q}(\|\bar{b}\|_{L^\infty}^q + \|\bar{\sigma}\|_{L^\infty}^q)\|f\|_{L^\infty}\delta^{\frac{q}{2}}.
\tag{4.7}
$$

Before turning to the next step, we summarise what we have proved so far. By (4.4), (4.5) and (4.6), we see that

$$
\big| \mathbb{E}[f(X_t)\varphi_r(X_t)\mathbb{1}_{I_t}]\big| \lesssim \sup_{x\in B_{3r}(x_0)}\Big|\mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\big]\Big| + \|f\|_{L^\infty}\delta^{\frac{q}{2}}.
$$

We use this estimate and (4.3) in (4.2) to finally obtain

$$
\big| \mathbb{E}[f(X_t)\varphi_r(X_t)]\big| \lesssim \sup_{x\in B_{3r}(x_0)}\Big|\mathbb{E}\big[f(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\big]\Big| + \|f\|_{L^\infty}\delta^{\frac{q}{2}},
\tag{4.8}
$$

where the constant in the inequality above depends on r, $\|\bar{b}\|_{L^\infty}$ and $\|\bar{\sigma}\|_{L^\infty}$.

4.1.3 The method

Let $\alpha \in (0,1)$ and $m \geq 1$ (to be suitably chosen later), and $h \in \mathbb{R}^d$ with $|h| \leq 1$. Consider $\phi \in \mathscr{C}_b^\alpha(\mathbb{R}^d)$, and apply (4.8) to get

$$
\big| \mathbb{E}[(\Delta_h^m\phi)(X_t)\varphi_r(X_t)]\big| \lesssim \sup_{x\in B_{3r}(x_0)}\Big|\mathbb{E}\big[(\Delta_h^m\phi)(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\big]\Big| + \|\phi\|_{L^\infty}\delta^{\frac{q}{2}},
\tag{4.9}
$$

A simple method for the existence of a density

We will apply our method to the new process $\bar{X}$. Under our standing assumptions, $\bar{X}$ has globally bounded drift and globally non–singular Hölder diffusion coefficient, so in principle the estimate should not be any different than what we did in Section 2. Indeed, we consider the extra parameter $\epsilon \in [0, \delta]$ and an extra process $\bar{Y}^x$, defined as in Section 2, and we do the decomposition

$$\mathbb{E}\big[(\Delta_h^m \phi)(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\big] = \mathbb{E}\big[(\Delta_h^m \phi)(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\big] - \mathbb{E}\big[(\Delta_h^m \phi)(\bar{Y}_\delta^x)\varphi_r(\bar{Y}_\delta^x)\big] + \\ + \mathbb{E}\big[(\Delta_h^m \phi)(\bar{Y}_\delta^x)\varphi_r(\bar{Y}_\delta^x)\big]$$

The *approximation error* is essentially the same,

$$\big| \mathbb{E}\big[(\Delta_h^m \phi)(\bar{X}_\delta^x)\varphi_r(\bar{X}_\delta^x)\big] - \mathbb{E}\big[(\Delta_h^m \phi)(\bar{Y}_\delta^x)\varphi_r(\bar{Y}_\delta^x)\big] \big| = \\ = \big| \mathbb{E}\big[(\Delta_h^m \phi)(\bar{X}_\delta^x)\big(\varphi_r(\bar{X}_\delta^x) - \varphi_r(\bar{Y}_\delta^x)\big)\big] + \mathbb{E}\big[\big((\Delta_h^m \phi)(\bar{X}_\delta^x) - (\Delta_h^m \phi)(\bar{Y}_\delta^x)\big)\varphi_r(\bar{Y}_\delta^x)\big] \big| \\ \lesssim \frac{\|D\varphi\|_{L^\infty}}{r}\|\phi\|_{L^\infty}\mathbb{E}|\bar{X}_\delta^x - \bar{Y}_\delta^x| + \|\varphi\|_{L^\infty}[\phi]_{\mathscr{C}_b^\alpha}\mathbb{E}|\bar{X}_\delta^x - \bar{Y}_\delta^x|^\alpha.$$

As in Section 2, we have

$$\mathbb{E}|\bar{X}_\delta^x - \bar{Y}_\delta^x| \lesssim \big(\|\bar{b}\|_{L^\infty} + [\bar{\sigma}]_{C^\beta}(\|\bar{b}\|_{L^\infty} + \|\bar{\sigma}\|_{L^\infty})^\beta\big)\epsilon^{\frac{1}{2}(1+\beta)},$$

thus

$$\textcolor{red}{\mathbf{Ae}} \lesssim (1 + \tfrac{1}{r})\|\varphi\|_{C^1}\|\phi\|_{\mathscr{C}_b^\alpha}\epsilon^{\frac{\alpha}{2}(1+\beta)}.$$

The *probabilistic estimate* is slightly more delicate. By conditioning on the σ–field $\mathscr{F}_{\delta-\epsilon}$ of events known at time $\delta - \epsilon$,

$$\mathbb{E}\big[(\Delta_h^m \phi)(\bar{Y}_\delta^x)\varphi_r(\bar{Y}_\delta^x)\big] = \mathbb{E}\big[\mathbb{E}[(\Delta_h^m \phi)(\bar{Y}_\delta^x)\varphi_r(\bar{Y}_\delta^x)|\mathscr{F}_{\delta-\epsilon}]\big] \\ = \mathbb{E}\big[\mathbb{E}[(\Delta_h^m \phi)(u + \bar{\sigma}(u)\tilde{B}_\epsilon))\varphi_r(u + \bar{\sigma}(u)\tilde{B}_\epsilon)]\big|_{u=\bar{Y}_{\delta-\epsilon}^x}\big],$$

since

$$\bar{Y}_\delta^x = Y_{\delta-\epsilon}^x + \bar{\sigma}(Y_{\delta-\epsilon}^x)(B_\delta - B_{\delta-\epsilon}),$$

and $\tilde{B}_s = B_{\delta-\epsilon+s} - B_{\delta-\epsilon}$, $s \in [0, \epsilon]$, is a Brownian motion. It remains to analyze the term

$$\mathbb{E}[(\Delta_h^m \phi)(u + \bar{\sigma}(u)\tilde{B}_\epsilon))\varphi_r(u + \bar{\sigma}(u)\tilde{B}_\epsilon)] = \int_{\mathbb{R}^d}(\Delta_h^m \phi)(u + y)\varphi_r(u + y)g_\epsilon(y)\,dy,$$

where g_ϵ is the density of a centred Gaussian random vector with covariance $\bar{\sigma}(u)\bar{\sigma}(u)^\star\epsilon$. By a simple change of variables (see the first formula in(4.10)),

$$\big| \mathbb{E}[(\Delta_h^m \phi)(u + \bar{\sigma}(u)\tilde{B}_\epsilon))\varphi_r(u + \bar{\sigma}(u)\tilde{B}_\epsilon)] \big| = \Big| \int_{\mathbb{R}^d} \phi(u + y)\Delta_{-h}^m(\varphi_r(u + \cdot)g_\epsilon)(y)\,dy \Big| \\ \leq \|\phi\|_{L^\infty}\|\Delta_{-h}^m(\varphi_r(u + \cdot)g_\epsilon)\|_{L^1}.$$

The Leibniz formula for discrete derivatives (4.10) yields

$$\|\Delta_{-h}^m(\varphi_r(u + \cdot)g_\epsilon)\|_{L^1} \leq \sum_{k=0}^m \binom{m}{k}\|\Delta_{-h}^k g_\epsilon\|_{L^1}\|(\Delta_{-h}^{m-k}\varphi_r(u + \cdot))(\cdot + kh)\|_{L^\infty}.$$

On the one hand, since $\epsilon \leq 1$, and since by the assumptions of the theorem there is $\bar{\sigma}_0 > 0$ such that $\bar{\sigma}_0^2 \leq \det(\bar{\sigma}(u)\bar{\sigma}(u)^\star)$ for all $u \in \mathbb{R}^d$,

$$\|\Delta_{-h}^k g_\epsilon\|_{L^1} \lesssim \Big(\frac{|h|}{\sqrt{\epsilon \det(\bar{\sigma}(u)\bar{\sigma}(u)^\star)}}\Big)^k \lesssim \frac{1}{\bar{\sigma}_0^k}\frac{|h|^k}{\epsilon^{\frac{k}{2}}}.$$

On the other hand,

$$\|(\Delta_{-h}^{m-k}\varphi_r(u+\cdot))(\cdot+kh)\|_{L^\infty} \lesssim \frac{\|D^{m-k}\varphi\|_{L^\infty}}{r^{m-k}}|h|^{m-k},$$

thus,

$$\|\Delta_{-h}^m(\varphi_r(u+\cdot)g_\epsilon)\|_{L^1} \lesssim \frac{|h|^m}{\epsilon^{\frac{m}{2}}}\|\varphi\|_{C^m}\sum_{k=0}^m \binom{m}{k}\frac{1}{\bar\sigma_0^k}\frac{1}{r^{m-k}} = \|\varphi\|_{C^m}\left(\frac{1}{r}+\frac{1}{\bar\sigma_0}\right)^m\frac{|h|^m}{\epsilon^{\frac{m}{2}}},$$

and in conclusion,

$$\left|\mathbb{E}\left[(\Delta_h^m\phi)(\bar Y_\delta^x)\varphi_r(\bar Y_\delta^x)\right]\right| \lesssim \|\phi\|_{L^\infty}\|\varphi\|_{C^m}\left(\frac{1}{r}+\frac{1}{\bar\sigma_0}\right)^m\frac{|h|^m}{\epsilon^{\frac{m}{2}}}.$$

In conclusion, by (4.9) and the above computations we have that

$$\left|\mathbb{E}\left[(\Delta_h^m\phi)(X_t)\varphi_r(X_t)\right]\right| \lesssim \|\phi\|_{\mathscr{C}_b^\alpha}\left(\delta^{\frac{q}{2}}+\epsilon^{\frac{1}{2}\alpha(1+\beta)}+\frac{|h|^m}{\epsilon^{\frac{m}{2}}}\right),$$

with q, δ, ϵ yet to be chosen, with $0 < \epsilon \le \delta \le 1 \wedge t$. Choose δ to be a multiple of ϵ, and $q = \alpha(1+\beta)$, so that

$$\left|\mathbb{E}\left[(\Delta_h^m\phi)(X_t)\varphi_r(X_t)\right]\right| \lesssim \|\phi\|_{\mathscr{C}_b^\alpha}\left(\epsilon^{\frac{\alpha}{2}(1+\beta)}+\frac{|h|^m}{\epsilon^{\frac{m}{2}}}\right),$$

The same computations of Proposition 2.2 yield the result.

We finally mention two formulae on discrete derivatives we have used in the computations above,

$$\int_{\mathbb{R}^d}(\Delta_h^m\phi)(x)\psi(x)\,dx = \int_{\mathbb{R}^d}\phi(x)(\Delta_{-h}^m\psi)(x)\,dx,$$
$$\Delta_h^m(\phi\psi)(x) = \sum_{k=0}^m\binom{m}{k}(\Delta_h^m\phi)(x)(\Delta_h^{m-k}\psi)(x+kh).$$

(4.10)

5 Rougher coefficients

In this section we wish to discuss an extension of the basic result on existence given in Section 2. We will only be able to lower the required regularity of the drift coefficient.

We wish first to give a few remarks about the possible limitations in the case of rough coefficients. With Proposition 2.2 at hand, it is reasonable to expect to find a threshold of regularity for the coefficients under which the method fails. Indeed, the key requirement is $a_0 > 0$ in formula (2.9). In the simple case of equation (5.1) below, this would correspond to the estimate **Ae** $\lesssim \epsilon^{\alpha/2}$, that is

$$\mathbb{E}\left|\int_{t-\epsilon}^t b(X_s)\,ds\right|^\alpha \lesssim \epsilon^{\frac{\alpha}{2}}.$$

Notice that this estimate is readily available if one observes that

$$\int_{t-\epsilon}^t b(X_s)\,ds = X_t - X_{t-\epsilon} - (B_t - B_{t-\epsilon}).$$

In other words, in the case $a_0 = 0$, at short times, the size of the drift is the same as the size of the noise and in principle one can believe that the effect of the noise fails to be effective for densities.

It would be interesting to understand if there is a counterexample to existence of densities in the case $a_0 \le 0$, or if simply the method fails.

5.1 Rougher drift coefficient

In this section we shall study the existence of a density for solutions of the equation

$$dX_t = b(t, X_t)\, dt + dB_t, \tag{5.1}$$

with non–regular coefficient b. We have taken a simple diffusion coefficient to focus on the difficulties originated by the drift.

Theorem 5.1 below shows existence and regularity of a density under the assumption $b \in L^q(0, T; L^p(\mathbb{R}^d))$, with $\frac{2}{q} + \frac{d}{p} < 1$. It is interesting to notice that this same condition ensures existence and uniqueness of a strong solution [30] of (5.1). See Remark 5.2 for a wider discussion along these lines.

This setting could provide a testbed for the problem of existence of densities in the case $a_0 \leq 0$ (addressed at the beginning of Section 5).

Theorem 5.1. *Assume* $b \in L^q(0, T; L^p(\mathbb{R}^d))$, *with* $\frac{2}{q} + \frac{d}{p} < 1$. *Then for every initial condition* $x \in \mathbb{R}^d$ *the solution of (5.1) with initial condition* x *has a density* $p_x(t)$ *for every* $t > 0$. *Moreover, for every* $\gamma < 1 - \frac{2}{q}$,

$$\sup_{t \in (0,T]} \sup_{x \in \mathbb{R}^d} (1 \wedge t)^{e_\gamma} \|p_x(t)\|_{B_{1,\infty}^\gamma} < \infty,$$

where e_γ *is any number such that* $e_\gamma > \frac{1 - \frac{1}{q}}{1 - \frac{2}{q} - \frac{d}{p}}\gamma$.

Proof. Set $p' = \frac{p}{p-1}$ and $a = \frac{d}{p}$, so that $p' = \frac{d}{d-a}$ and thus $B_{1,\infty}^a \subset L^{p'}$. Set for every $\gamma > 0$,

$$\|p.\|_{\star,\gamma} := \sup_{t \in (0,T]} \sup_{x \in \mathbb{R}^d} (1 \wedge t)^{e_\gamma} \|p_x(t)\|_{B_{1,\infty}^\gamma},$$

where $\gamma \mapsto e_\gamma$ will be identified below in the proof. Now, by the Hölder inequality, and Sobolev's embeddings,

$$
\begin{aligned}
\mathbb{E} \int_{t-\epsilon}^t |b(s, X_s)|\, ds &= \int_{t-\epsilon}^t \int_{\mathbb{R}^d} |b(s, y)| p_x(s, y)\, dy \\
&\leq \int_{t-\epsilon}^t \|b(s)\|_{L^p} \|p_x(s)\|_{L^{p'}}\, ds \\
&\lesssim \int_{t-\epsilon}^t \|b(s)\|_{L^p} \|p_x(s)\|_{B_{1,\infty}^a}\, ds \\
&\leq \|p.\|_{\star,a} \int_{t-\epsilon}^t \|b(s)\|_{L^p} (1 \wedge s)^{-e_a}\, ds \\
&\leq \|p.\|_{\star,a} \|b\|_{L^q(0,T;L^p)} \Big(\int_{t-\epsilon}^t (1 \wedge s)^{-e_a q'}\, ds \Big)^{\frac{1}{q'}} \\
&\lesssim (1 \wedge t)^{-e_a} \epsilon^{\frac{1}{q'}} \|p.\|_{\star,a} \|b\|_{L^q(0,T;L^p)},
\end{aligned}
$$

where $q' = \frac{q}{q-1}$. We can thus estimate the *approximation error* with,

$$
\begin{aligned}
\mathbf{\color{red}{Ae}} \leq [\phi]_{\mathscr{C}_b^\alpha} \mathbb{E}\Big[\Big| \int_{t-\epsilon}^t b(s, X_s)\, ds \Big|^\alpha \Big] &\leq \\
&\leq [\phi]_{\mathscr{C}_b^\alpha} \mathbb{E}\Big[\int_{t-\epsilon}^t |b(s, X_s)|\, ds \Big]^\alpha \lesssim (1 \wedge t)^{-\alpha e_a} \epsilon^{\alpha/q'} \|p.\|_{\star,a}^\alpha [\phi]_{\mathscr{C}_b^\alpha}.
\end{aligned}
$$

The *probabilistic estimate* is as in the standard case (Section 2), so, if we set $K_0 = (1 \wedge t)^{-2\alpha e_a}(1 \vee \|p.\|_{\star,a}^{2\alpha})$ and $a_0 = \frac{2}{q'} - 1$, then we have (2.9), therefore (2.10) holds, that is

$$
\begin{aligned}
\|p_x(t)\|_{B_{1,\infty}^\gamma} &\lesssim K_0^{\frac{\gamma}{2a_0}+\delta} (1 \wedge t)^{-\frac{1+a_0}{2a_0}\gamma - \delta} \\
&= (1 \wedge t)^{-(\frac{\gamma}{a_0}+2\delta)e_a - \frac{1+a_0}{2a_0}\gamma - \delta}(1 \vee \|p.\|_{\star,a}^{\frac{\gamma}{a_0}+2\delta}).
\end{aligned} \tag{5.2}
$$

This formula shows an a–priori estimate for $\|p.\|_{\star,\gamma}$, once we have defined, through this same formula, the value of e_γ as

$$e_\gamma > \frac{\gamma}{a_0} e_a + \frac{1 + a_0}{2a_0} \gamma.$$

Unfortunately the estimate, as well as the value e_γ, depend on $\|p.\|_{\star,a}$ and e_a.

Notice that our assumption ensures that $a < a_0$, therefore if we choose $\gamma = a$, the restriction on e_a reads

$$e_a > \frac{1 + a_0}{2(a_0 - a)} a.$$

The constant in the inequality (5.2) does not depend on the initial condition, thus

$$\|p.\|_{\star,a} \lesssim (1 \vee \|p.\|_{\star,a}^{\frac{a}{a_0} + 2\delta}).$$

If we choose δ small enough that $\frac{a}{a_0} + 2\delta < 1$, this provides an a–priori estimate for $\|p.\|_{\star,a}$, with e_a chosen as above.

With the estimate of $\|p.\|_{\star,a}$ at hand, we look back at (5.2) and see that if

$$e_\gamma > \frac{1 + a_0}{2(a_0 - a)} \gamma,$$

then (5.2) provides an estimate for $\|p.\|_{\star,\gamma}$. $\qquad\square$

Remark 5.2. An alternative proof via a backward Kolmogorov equation is available to prove Theorem 5.1. The idea we wish to use has been extensively used in results on regularization by noise, see for instance [18], or [30], and [16] for an extension to random drifts. The idea, here, is to re–write the equation (5.1) as an equation with more regular coefficients. Notice though that in this way, to "solve" a Fokker–Planck equation, we end up using its dual Kolmogorov equation. Our original proof has the advantage to be based on elementary arguments that do not require to solve PDEs, and therefore can be readily used in more general cases, see for instance Section 7.2.2.

Indeed, for $b \in L^q(0, T; L^p(\mathbb{R}^d))$, $p \geq 2$ and $\frac{d}{p} + \frac{2}{q} < 1$, consider the following backward parabolic equation,

$$\begin{cases} \partial_s U + \frac{1}{2}\Delta U - \lambda U + b \cdot \nabla U = b, & s \leq t, \\ U(t, x) = 0, & x \in \mathbb{R}^d. \end{cases} \tag{5.3}$$

with $U \in L^q(0, t; W^{2,p}(\mathbb{R}^d)) \cap W^{1,q}(0, t; L^p(\mathbb{R}^d))$ (see [30, Theorem 10.3]). It is possible to give a quantitative estimate of the dependence on λ (with a minor modification from [17, Lemma 3.4], to make the dependence on λ more explicit),

$$\|U\|_{L^\infty([0,t]\times\mathbb{R}^d)} \leq c\lambda^{-a-\frac{1}{2}}. \qquad \|\nabla U\|_{L^\infty([0,t]\times\mathbb{R}^d)} \leq c\lambda^{-a}.$$

for every positive $a < \frac{1}{2}(1 - \frac{2}{q} - \frac{d}{p})$, and for λ large enough. By Itō's formula,

$$\int_{t-\epsilon}^t b(s, X_s)\, ds = U(t, X_t) - U(t - \epsilon, X_{t-\epsilon})$$
$$- \lambda \int_{t-\epsilon}^t U(s, X_s)\, ds - \int_{t-\epsilon}^t \nabla U(s, X_s) \cdot dB_s,$$

and this immediately allows to estimate the *approximation error* (with the optimal choice $\lambda = \epsilon^{-1}$).

5.1.1 An almost sure regularity result

We discuss a simple application of the previous result. The regularity of the density we have found in Theorem 5.1 can be used to obtain regularity of functionals of solutions of (5.1). Indeed, under the assumptions of Theorem 5.1, if $f \in L^q(0, T; L^p(\mathbb{R}^d))$, then the function $t \mapsto \mathbb{E}[\int_0^t f(r, X_r^x) \, dr]$ is Hölder continuous in $(0, T)$. To see this, we simply use the Hölder inequality, embeddings of Besov spaces, and the estimate of the previous theorem to obtain,

$$
\begin{aligned}
\left| \mathbb{E}\Big[\int_s^t f(r, X_r^x) \, dr \Big] \right| &\leq \int_s^t \|f(r)\|_{L^p} \|p_x(r)\|_{L^{p'}} \, ds \\
&\lesssim \int_s^t |f(r)|_{L^p} \|p_x(r)\|_{B_{1,\infty}^a} \, ds \lesssim \int_s^t \|f(r)\|_{L^p} (1 \wedge r)^{-e_a} \, dr \\
&\lesssim \|f\|_{L^q(L^p)} \Big(\int_s^t (1 \wedge r)^{-q' e_a} \, dr \Big)^{\frac{1}{q'}} \lesssim \frac{(t-s)^{\frac{1}{q'}}}{(1 \wedge s)^{e_a}} \|f\|_{L^q(L^p)},
\end{aligned}
$$

where $a = \frac{d}{p}$, and p' and q' are the conjugate Hölder exponents of p, q. Similar computations show that the same function is Hölder continuous on $[0, T]$ if $q' e_a < 1$, that is if $\frac{2d}{p} + \frac{2}{q} < 1$. More generally, we have the following result.

Proposition 5.3. *Assume $b \in L^q(0, T; L^p(\mathbb{R}^d))$, with $\frac{2}{q} + \frac{2d}{p} < 1$. Let X^x be the solution of (5.1) with initial condition $x \in \mathbb{R}^d$, and let $p_x(t)$ be its density for every $t > 0$. If $f \in L^q(0, T; L^p(\mathbb{R}^d))$, then the map*

$$
t \mapsto \int_0^t f(r, X_r^x) \, dr
$$

is a. s. Hölder continuous on $[0, T]$ with exponent smaller than $\frac{1}{q'} - e_a$, where q is the conjugate Hölder exponent of q and $a = \frac{d}{p}$.

Proof. The proof is elementary and uses the Markov property and Kolmogorov's continuity theorem. We only give a sketch under the assumption that f does not depend on t. The general case is entirely similar.

Let $a = \frac{d}{p}$, and set for brevity $g(r) = (1 \wedge r)^{-e_a}$. By the Markov property, if $r_1 < r_2 < \cdots < r_k$,

$$
\begin{aligned}
\left| \mathbb{E}[f(X_{r_1}^x) f(X_{r_2}^x) \ldots f(X_{r_k}^x)] \right| &= \left| \mathbb{E}\big[f(X_{r_1}^x) f(X_{r_2}^x) \ldots f(X_{r_{k-1}}^x) \mathbb{E}[f(X_{r_k}^x) | \mathscr{F}_{r_{k-1}}] \big] \right| \\
&= \left| \mathbb{E}\big[f(X_{r_1}^x) f(X_{r_2}^x) \ldots f(X_{r_{k-1}}^x) \mathbb{E}[f(X_{r_k - r_{k-1}}^y)]_{y = X_{r_{k-1}}^x} \big] \right| \\
&\lesssim \|f\|_{L^p} g(r_k - r_{k-1}) \mathbb{E}\big[f(X_{r_1}^x) f(X_{r_2}^x) \ldots f(X_{r_{k-1}}^x) \big] \\
&\lesssim \cdots \lesssim \|f\|_{L^p}^k g(r_1) g(r_2 - r_1) \ldots g(r_k - r_{k-1}),
\end{aligned}
$$

since, as above,

$$
\left| \mathbb{E}[f(X_r^x)] \right| \leq \|f\|_{L^p} \|p_x(r)\|_{L^{p'}} \leq \|f\|_{L^p} \|p_x(r)\|_{B_{1,\infty}^a} \lesssim \|f\|_{L^p} g(r).
$$

Thus, by k changes of variables,

$$\left| \mathbb{E}\left[\left(\int_s^t f(X_r^x) \, dr \right)^k \right] \right| =$$

$$= \left| \int_s^t \ldots \int_s^t \mathbb{E}[f(X_{r_1}^x) f(X_{r_2}^x) \ldots f(X_{r_k}^x)] \, dr_k \ldots dr_2 \, dr_1 \right|$$

$$= k! \left| \int_s^t dr_1 \int_{r_1}^t dr_2 \cdots \int_{r_{k-1}}^t dr_k \mathbb{E}[f(X_{r_1}^x) f(X_{r_2}^x) \ldots f(X_{r_k}^x)] \right|$$

$$\lesssim k! \|f\|_{L^p}^k \int_s^t dr_1 \int_{r_1}^t dr_2 \ldots \int_{r_{k-1}}^t dr_k \, g(r_1) g(r_2 - r_1) \ldots g(r_k - r_{k-1})$$

$$= k! \|f\|_{L^p}^k \int_0^{t-s} g(t - r_1) \int_0^{r_1} g(r_1 - r_2) \int_0^{r_2} g(r_2 - r_3) \ldots \int_0^{r_{k-1}} g(r_{k-1} - r_k) \, dr_k \ldots dr_2 \, dr_1.$$

If $t - s \leq 1$, then $r_k \leq r_{k-1} \leq \cdots \leq r_1 \leq t - s \leq 1$ and

$$\int_0^{r_{k-1}} g(r_{k-1} - r_k) \, dr_k = \frac{r_{k-1}^{1-e_a}}{1 - e_a} \lesssim (t - s)^{1-e_a},$$

therefore

$$\left| \mathbb{E}\left[\left(\int_s^t f(X_r^x) \, dr \right)^k \right] \right| \lesssim (t - s)^{(k-1)(1-e_a)} \int_0^{t-s} g(t - r_1) \, dr_1 \lesssim (t - s)^{k(1-e_a)},$$

since

$$\int_0^{t-s} g(t - r_1) \, dr_1 \lesssim \frac{t - s}{(1 \wedge t)^{e_a}} \leq (t - s)^{1-e_a}. \qquad \square$$

Remark 5.4. We remark that the above proposition ***does not*** show that the stochastic flow generated by (5.1) is Hölder continuous.

5.2 Rougher diffusion coefficient

We wish to briefly discuss some difficulties related to the extension of the "core" method with rougher diffusion coefficients.

First of all, regularity of the diffusion coefficient is a more delicate issue, and regularity itself might not be as significant as for the drift coefficient in view of densities. For instance by the Levy characterization theorem we know if $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d'}$ satisfies $|\sigma(x)| = 1$ for all x, then the solution of

$$dX_t = \sigma(X_t) \, dB_t, \tag{5.4}$$

is a Brownian motion and thus has a smooth density.

On the other hand the method we have introduced seems to strongly depend on an evaluation of the increments of σ. It would be thus reasonable to expect results if we lower the summability of the "derivatives" of σ, namely by requiring that $\sigma \in B_{p,q}^\beta$ for some $p, q \geq 1$ but finite (recall that $\mathscr{C}_b^\beta = B_{\infty,\infty}^\beta$).

A strategy to estimate the *approximation error* could be as in Theorem 5.1. Assume for instance σ is non–singular and $\sigma \in L^\infty \cap B_{p,q}^\beta$, and let $p_x(s)$ be the density of the solution X_s^x of (5.4) with initial condition $X_0 = x$, fix $t > 0$ and let Y be the auxiliary process introduced in Section 2. The *approximation error* would be

$$\mathsf{Ae} \lesssim [\phi]_{\mathscr{C}_b^\alpha} \left(\int_{t-\epsilon}^t \mathbb{E}[|\sigma(X_r^x) - \sigma(X_{t-\epsilon}^x)|^2] \, dr \right)^{\frac{\alpha}{2}},$$

and the Markov property yields

$$\mathbb{E}[|\sigma(X_r^x) - \sigma(X_{t-\epsilon}^x)|^2] = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{E}[|\sigma(z) - \sigma(y)|^2] p_y(t-s, z) p_x(s, y) \, dz \, dy.$$

On the one hand the above formula contains increments of σ that could be estimate using the regularity of σ. On the other hand the term $p_y(t - s, \cdot)$ gives a too singular contribution, if we expect a singularity in time as in Theorem 5.1. Indeed, one can see that the estimate of the term above is successful only when $\beta > \frac{d}{p}$, that is whenever σ, by Sobolev's embeddings, is a Hölder function.

6 More regularity - II

In this section we shall see that the method introduced in Section 2 is not optimal, and suggest a partial probabilistic proof that goes in the direction of the optimal result.

Consider for simplicity the problem with constant diffusion (although our considerations equally hold with a non singular–diffusion) and bounded drift,

$$dX_t = b(X_t) \, dt + dB_t,$$

with initial condition $X_0 = x$. The computations in Section 2 show that X_t has a density $p_x(t) \in B_{1,\infty}^{1-}$. It is easy to be convinced though that the expected regularity is $B_{\infty,\infty}^{1-}$ (that is Hölder). Indeed, ideally the density should solve the associated Fokker–Planck equation

$$\partial_t p_x = \frac{1}{2} \Delta p_x - \nabla \cdot (b p_x),$$

with initial condition $p_x(0) = \delta_x$, or

$$p_x(t, y) = g(t, y - x) + \int_0^t \nabla g(t - s) \star (b p_x)(y) \, dt,$$

where g is the heat kernel. It is not difficult to prove, using this mild formulation, that

$$\|\Delta_h^2 p_x(t)\|_{L^1} \le \|\Delta_h^2 g(t)\|_{L^1} + \left\| \int_0^t (\Delta_h^2 \nabla g(t - s)) \star (b p_x) \, dt \right\|_{L^1}$$

$$\le \|\Delta_h^2 g(t)\|_{L^1} + \|b\|_{L^\infty} \int_0^t \|\Delta_h^2 \nabla g(t - s)\|_{L^1} \, ds,$$

and conclude with estimates for the heat kernel. Similar computations (with a caveat) also show Hölder bounds. The trick is to bound the quantity

$$\mathcal{F}_T = \sup_{[0,T]} t^{\frac{d}{2}} \|p_\cdot(t)\|_{L^\infty},$$

although we need to know a–priori that this quantity is finite. A similar quantity has been considered in Section 5.

We turn to a result that improves slightly (but without getting the optimal result) the summability of the density. We will work under the same assumptions of Section 2.

Theorem 6.1. Let $b \in L^\infty(\mathbb{R}^d)$ and $\sigma \in C_b^\beta(\mathbb{R}^d; \mathbb{R}^{d \times d'})$, with $\beta \in (0, 1)$. Assume moreover (2.2). Let X^x be a solution of (2.1) with initial condition $x \in \mathbb{R}^d$. Then the density $p_x(t)$ of X_t^x is in $B_{p,\infty}^a(\mathbb{R}^d)$ for every $a < \beta$ and every $p < \frac{d}{d-\beta}$.

Proof. By the computations in Section 2 we know that under the standing assumptions the density $p_x(t) \in B_{1,\infty}^a$ for every $a < \beta$. By Sobolev's embeddings, $p_x(t) \in L^p$ for every $p < \frac{d}{d-\beta}$.

A simple method for the existence of a density

Fix $p \in (1, \frac{d}{d-\beta})$ and denote by q the conjugate Hölder exponent of p. Notice that, by our choice of p, we have that $q > \frac{d}{\beta}$.

We will use the smoothing Lemma A.3. To this end let $\phi \in F^\alpha_{q,\infty}(\mathbb{R}^d)$, with $\alpha \in (\frac{d}{q}, 1)$. With these values of α and q, we can use the characterization of $F^\alpha_{q,\infty}$ in terms of differences given in formula (A.2). In particular, if we set

$$\Phi(x) := \sup_{h \neq 0} \frac{|\Delta_h \phi(x)|}{|h|^\alpha}, \tag{6.1}$$

then $\Phi \in L^q(\mathbb{R}^d)$, $[\phi]_{B^\alpha_{q,\infty}} = \|\Phi\|_{L^q}$, and $|\Delta_h \phi(x)| \leq |h|^\alpha \Phi(x)$. As usual, we consider $\mathbb{E}[\Delta^m_h \varphi(X_t)]$ and split it into *probabilistic estimate* and *approximation error*.

The *probabilistic estimate* is not the source of issues. We only have to make a Hölder inequality in formula (2.7) in $L^p - L^q$ rather than $L^1 - L^\infty$, to get

$$\mathbf{Pe} = \left| \mathbb{E}[\Delta^m_h \phi(Y^\epsilon_t)] \right| \lesssim \epsilon^{-\frac{d}{2q}} \frac{|h|^m}{\epsilon^{\frac{m}{2}}} \|\phi\|_{L^q},$$

where Y^ϵ is the process defined in formula (2.3).

We turn to the *approximation error*,

$$\mathbf{Ae} = \left| \mathbb{E}[\phi(X^x_t) - \phi(Y^\epsilon_t)] \right| = \left| \mathbb{E}\left[\mathbb{E}[\phi(X^x_t) - \phi(Y^\epsilon_t) | X_{t-\epsilon}] \right] \right|.$$

Notice that Y^ϵ_t is a function of $X^x_{t-\epsilon}$ and of a Brownian motion $\tilde{B}_s = B_{t-\epsilon+s} - B_{t-\epsilon}$, $s \geq 0$ that is independent of the history until time $t - \epsilon$, and X is a Markov process. Thus

$$\mathbb{E}[\phi(X^x_t) - \phi(Y^\epsilon_t) | X^x_{t-\epsilon}] = \mathbb{E}[\phi(X^y_\epsilon) - \phi(y + \sigma(y)\tilde{B}_\epsilon)]|_{y = X^x_{t-\epsilon}},$$

and

$$X^y_\epsilon = y + \sigma(y)\tilde{B}_\epsilon + \int_0^\epsilon b(X^y_s)\, ds + \int_0^\epsilon (\sigma(X^y_s) - \sigma(y))\, d\tilde{B}_s.$$

Therefore, by the Hölder inequality, and with the same computations used to obtain the estimate (2.6)

$$\left| \mathbb{E}[|\phi(X^y_\epsilon) - \phi(y + \sigma(y)\tilde{B}_\epsilon)] \right| \leq$$
$$\leq \mathbb{E}\left[\left| \int_0^\epsilon b(X^y_s)\, ds + \int_0^\epsilon (\sigma(X^y_s) - \sigma(y))\, d\tilde{B}_s \right|^\alpha \Phi(y + \sigma(y)\tilde{B}_\epsilon) \right]$$
$$\leq \mathbb{E}\left[\left| \int_0^\epsilon b(X^y_s)\, ds + \int_0^\epsilon (\sigma(X^y_s) - \sigma(y))\, d\tilde{B}_s \right|^{\alpha p} \right]^{\frac{1}{p}} \mathbb{E}[|\Phi(y + \sigma(y)\tilde{B}_\epsilon)|^q]^{\frac{1}{q}}$$
$$\lesssim \epsilon^{\frac{\alpha}{2}(1+\beta)} \mathbb{E}[|\Phi(y + \sigma(y)\tilde{B}_\epsilon)|^q]^{\frac{1}{q}}.$$

We thus have, using the computations above and again the Hölder inequality,

$$\mathbf{Ae} = \left| \mathbb{E}\left[\mathbb{E}[\phi(X^y_\epsilon) - \phi(y + \sigma(y)\tilde{B}_\epsilon)]|_{y = X^x_{t-\epsilon}} \right] \right|$$
$$= \left| \int_{\mathbb{R}^d} \mathbb{E}[\phi(X^y_\epsilon) - \phi(y + \sigma(y)\tilde{B}_\epsilon)] p_x(t-\epsilon, y)\, dy \right|$$
$$\lesssim \epsilon^{\frac{\alpha}{2}(1+\beta)} \int_{\mathbb{R}^d} \mathbb{E}[|\Phi(y + \sigma(y)\tilde{B}_\epsilon)|^q]^{\frac{1}{q}} p_x(t-\epsilon, y)\, dy$$
$$\lesssim \epsilon^{\frac{\alpha}{2}(1+\beta)} \|p_x(t-\epsilon)\|_{L^p} \left(\mathbb{E}\int_{\mathbb{R}^d} |\Phi(y + \sigma(y)\tilde{B}_\epsilon)|^q\, dy \right)^{\frac{1}{q}}.$$

Let $g_{\sigma(y),\epsilon}$ be the density of $y + \sigma(y)\tilde{B}_\epsilon$, then by the non–degeneracy assumption on σ it is easy to see that

$$g_{\sigma(y),\epsilon}(z) \lesssim \frac{1}{(2\pi\epsilon)^{d/2}} e^{-\frac{c}{\epsilon}|z-y|^2},$$

for a constant c that depends on σ, thus

$$\mathbb{E} \int_{\mathbb{R}^d} |\Phi(y + \sigma(y)\tilde{B}_\epsilon)|^q \, dy = \int_{\mathbb{R}^d} |\Phi(z)|^q \int_{\mathbb{R}^d} g_{\sigma(y),\epsilon}(z) \, dy \, dz \lesssim \|\Phi\|_{L^q}^q \sim [\phi]_{F_{q,\infty}^\alpha}.$$

In conclusion,

$$\mathsf{Ae} \lesssim \epsilon^{\frac{\alpha}{2}(1+\beta)} \|p_x(t-\epsilon)\|_{L^p} [\phi]_{F_{q,\infty}^\alpha}.$$

Since we work with $\epsilon < 1 \wedge \frac{t}{2}$, by (2.10) we have that

$$\|p_x(t-\epsilon)\|_{L^p} \lesssim \|p_x(t-\epsilon)\|_{B_{1,\infty}^{d/q}} \lesssim c_t$$

Using the same computations of Proposition 2.2 and the smoothing Lemma A.3, we finally obtain that $p_x(t) \in B_{p,\infty}^a$ for every $a < \beta$. $\qquad\square$

Remark 6.2. With a more careful analysis of the proof of Proposition 2.2, it is not difficult to obtain an estimate of the norm of the density in $B_{p,\infty}^a$ in terms of t, as in (2.10).

Remark 6.3. Notice that, since by the theorem above we now know that the density is in $B_{\frac{d}{d-\beta}-,\infty}^{\beta-}$, we can deduce a stronger summability (this was the starting point of the proof above). Unfortunately the limitation in the characterization (A.2) of Triebel–Lizorkin spaces prevents us to iterate the above proof and deduce Hölder bounds.

Remark 6.4. In the case $d = 1, \beta > \frac{1}{2}$ we have the embedding of $B_{\frac{d}{d-\beta}-,\infty}^{\beta-}$ in spaces of Hölder functions. We can thus conclude that the density is Hölder continuous, as expected.

7 Examples and applications

In this section we show a few applications of the simple method and its improvements illustrated in the first part of the paper.

Before presenting the examples, we wish to notice that, in the estimate of the *probabilistic estimate*, there are two main key points. The first is the estimate of the small time asymptotics of the "noise" part of the equation. The second is that the method we have illustrated depends very much on the fact that for times close to the final time the noise is independent of the past. In principle this might rule out, for instance, processes such as the fractional Brownian motion as the source of noise. In the particular case of the fractional Brownian motion though, one can use a suitable integral representation, and in that case is easy to be convinced that most of the results (and in particular the basic results of Section 2) hold true with minimal modifications (essentially we have the Hurst index as the value of the parameter θ in Proposition 2.2).

Moreover we notice that, in view of the analysis of infinite dimensional problems, our method seems to be tailored to a white in time noise, and it seems so far unlikely that it can be applied to problems driven by a spatial noise, with no temporal components (see for instance [9] for results in this direction based on Malliavin's calculus).

7.1 A path-dependent SDE

Here we use only the basic method (from Section 2), because as such we do not have a Markov evolution. This result can be essentially found already in [7], we provide it for the purpose of illustrating the method. Notice that the example includes, with suitable adjustments, also the case of equations with delay.

Given $T > 0$ and two integers $d, d' \geq 1$, let $b : [0,T] \times C([0,T];\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T] \times C([0,T];\mathbb{R}^d) \to \mathbb{R}^{d \times d'}$ be such that for every $t \geq 0$, if $\omega|_{[0,t]} = \omega'|_{[0,t]}$, then

$b(t, \omega) = b(t, \omega')$ and $\sigma(t, \omega) = \sigma(t, \omega')$. In other words, b and σ, when evaluated at time t, depend only on the part of the path up to time t. This is to ensure adaptness in the equation.

Given $0 \le s \le t \le T$, set for every $\omega \in C([0, T]; \mathbb{R}^d)$,

$$\delta_{s,t}(\omega) := \sup_{r \in [s,t]} |\omega_r - \omega_s|.$$

Assume that

- b, σ are bounded,
- there are $c > 0$ and $\beta \in (0, 1)$ such that

$$|\sigma(t, \omega) - \sigma(s, \omega)| \le c d_{s,t}(\omega)^\beta,$$

  for every $\omega \in C([0, T]; \mathbb{R}^d)$ and every $0 \le s \le t \le T$,
- there is $\lambda_0 > 0$ such that $\sigma(t, \omega)\sigma(t, \omega)^\star \ge \lambda_0 I$, for every $t \in [0, T]$ and every $\omega \in C([0, T]; \mathbb{R}^d)$.

Consider the following path-dependent stochastic equation,

$$dX_t = b(t, \mathbf{X}) \, dt + \sigma(t, \mathbf{X}) \, dB_t, \tag{7.1}$$

where $(B_t)_{t \ge 0}$ is a d'-dimensional Brownian motion, and $\mathbf{X} = (X_t)_{t \in [0,T]}$.

Proposition 7.1. *Under the above assumption, if $(X_t)_{t \ge 0}$ is a solution of (7.1), then for every $t \in (0, T]$ the random variable X_t has a density with respect to the Lebesgue measure. Moreover, the density is in $B_{1,\infty}^a(\mathbb{R}^d)$ for every $a < \beta$.*

Proof. We use the auxiliary process (2.3). Notice that by our first assumption the term $\sigma(t - \epsilon, \mathbf{X})$ is measurable with respect to the history up to $t - \epsilon$.

We have,

$$X_t - Y_t^\epsilon = \int_{t-\epsilon}^t b(s, \mathbf{X}) \, ds + \int_{t-\epsilon}^t \big(\sigma(s, \mathbf{X}) - \sigma(t - \epsilon, \mathbf{X}) \big) \, dB_s,$$

thus the *approximation error* is,

$$\mathbf{Ae} \lesssim [\varphi]_{\mathscr{C}^\alpha} \left(\|b\|_{L^\infty} \epsilon^\alpha + \mathbb{E}\Big[\Big| \int_{t-\epsilon}^t \big(\sigma(s, \mathbf{X}) - \sigma(t - \epsilon, \mathbf{X}) \big) \, dB_s \Big|^\alpha \Big] \right)$$

$$\lesssim [\varphi]_{\mathscr{C}^\alpha} \left(\|b\|_{L^\infty} \epsilon^\alpha + \mathbb{E}\Big[\int_{t-\epsilon}^t |\sigma(s, \mathbf{X}) - \sigma(t - \epsilon, \mathbf{X})|^2 \, ds \Big]^{\frac{\alpha}{2}} \right)$$

$$\lesssim [\varphi]_{\mathscr{C}^\alpha} \left(\|b\|_{L^\infty} \epsilon^\alpha + \mathbb{E}\Big[\int_{t-\epsilon}^t \delta_{t-\epsilon,s}(\mathbf{X})^{2\beta} \, ds \Big]^{\frac{\alpha}{2}} \right).$$

To estimate $\mathbb{E}[\delta_{t-\epsilon,s}(\mathbf{X})^{2\beta}]$, we notice that

$$\delta_{t-\epsilon,s}(\mathbf{X}) \le \epsilon \|b\|_{L^\infty} + \sup_{[t-\epsilon,t]} \Big| \int_{t-\epsilon}^s \sigma(r, \mathbf{X}) \, dB_r \Big|,$$

so that $\mathbb{E}[\delta_{t-\epsilon,s}(\mathbf{X})^{2\beta}] \le \epsilon^\beta$ and $\mathbf{Ae} \lesssim [\varphi]_{\mathscr{C}^\alpha} \epsilon^{\frac{\alpha}{2}(1+\beta)}$.

By our second assumption, the *probabilistic estimate* can be estimated as in Section 2, to get $\mathbf{Pe} \lesssim \epsilon^{-m/2} |h|^m \|\varphi\|_{L^\infty}$. Proposition 2.2 concludes the proof. $\square$

Remark 7.2. Notice that in the derivation of the results of Sections 4, 5, and 6, the Markov property was used at some stage, thus such results cannot be adapted to this setting.

Remark 7.3. The result we have obtained here is slightly less general than in [7, Section 3.1]. Indeed, in [7] they have the same assumptions, but they have a weaker form of continuity of the diffusion coefficient, namely,

$$|\sigma(t, \omega) - \sigma(s, \omega)| \lesssim (-\log \delta_{s,t}(\omega))^{-2-\epsilon},$$

for some $\epsilon > 0$. The density in this case is in $L^{e_{\log}}$, the Orlicz space with Young function $e_{\log}(u) = (1 + |u|) \log(1 + |u|)$. We believe that in principle, if one would consider Besov spaces with norm as in [7], one might extend the result we have given above with this weaker assumption. We do not consider this, due to the analytical difficulties involved.

7.2 Lévy noise driven SDEs

In this section we wish to slightly extend the results of [14] in the direction of Sections 5 and 6. We briefly recall the setting from [14].

Assumption 7.4. *Let $(Z_t)_{t \geq 0}$ be a Lévy process with characteristics function $k \mapsto e^{-t\Psi(k)}$, with*

$$\Psi(u) = \int_{\mathbb{R}^d} (1 - e^{i\langle u, y\rangle} + i\langle u, y\rangle \mathbb{1}_{\{|y| \leq 1\}}) \, \mu(dy),$$

and there is $\alpha \in (0, 2)$ such that the Lévy measure μ satisfies,

- *$\int_{|y| \geq 1} |y|^\gamma \mu(dy) < \infty$ for all $\gamma \in [0, \alpha)$,*
- *there is $c > 0$ such that $\int_{|y| \leq \lambda} |y|^2 \mu(dy) \leq c\lambda^{2-\alpha}$ for all $\lambda \in (0, 1]$,*
- *there are $c > 0$, $r > 0$ such that $\int (1 - \cos\langle u, y\rangle)\mu(dy) \geq c|u|^\alpha$, for all $u \in \mathbb{R}^d$ with $|u| \geq r$.*

We consider the following equation, driven by the α-stable-like process $(Z_t)_{t \geq 0}$ defined above,

$$dX_t = b(t, X_t) \, dt + \sigma(X_t) \, dZ_t, \tag{7.2}$$

where $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ is bounded measurable, and $\sigma \in C_b^\beta(\mathbb{R}^d; \mathbb{R}^{d \times d'})$, with

$$\sigma(y)\sigma(y)^\star > 0, \qquad \text{uniformly in } y \in \mathbb{R}^d. \tag{7.3}$$

Notice that here the non–degeneracy assumption could be weakened following the lines of Section 3.

In the rest of the section we will restrict to the (simpler) case $\alpha \in (1, 2)$. While this greatly simplifies the computations of Section 7.2.1, it is a necessary assumption in Section 7.2.2 to obtain sufficiently smallness in the *approximation error*.

7.2.1 More regularity

In this section we will adapt the ideas of Section 6 to problem (7.2). From [14] we know that, under the above assumptions on the coefficients and on the driving process, any solution of (7.2) has a density in $B_{1,\infty}^s$, for $s < \beta \wedge (\alpha - 1)$. We wish to improve the summability index. As far as we know, under the standing assumptions on the coefficients there is not yet a general result of existence or uniqueness for (7.2). So in the rest of this section we assume that either there is a solution of (7.2) that is a Markov process, or that there is a solution that can be obtained by approximation of (7.2) with smooth coefficients. In the latter case the Besov bound ensures the uniform integrability of the approximating densities and thus that the limit problem has a density with the same Besov regularity.

Theorem 7.5. *Let $(Z_t)_{t\geq 0}$ be a Lévy process as in Assumption 7.4, with $\alpha > 1$, let b be bounded measurable and $\sigma \in C_b^\beta(\mathbb{R}^d; \mathbb{R}^{d\times d'})$ such that (7.3) holds, and set $\kappa = \min(\alpha - 1, \beta)$. For every $x \in \mathbb{R}^d$ let $(X_t^x)_{t\geq 0}$ be a solution (as specified above) of (7.2). Then for every $t > 0$ and $x \in \mathbb{R}^d$, the density $p_x(t)$ of the random variable X_t^x is in $B_{p,\infty}^a(\mathbb{R}^d)$ for every $p \in (1, \frac{d}{d-\kappa})$ and every $a < \kappa(1 \wedge \frac{\alpha}{p})$.*

We proceed with the proof of the additional regularity of the density. We will need first a slight modification of [14, Lemma 3.3] (which in turn is a quantitative version of [46, Theorem 1.2]).

Lemma 7.6. *Let $(Z_t)_{t\geq 0}$ be a Lévy process as in Assumption 7.4, and let g_t be the density of Z_t, for each $t > 0$. Then for all integers $m \geq 1$, all $h \in \mathbb{R}^d$, $|h| \leq 1$, and all $p \in [1, \infty]$,*

$$\|\Delta_h^m g_t\|_{L^p} \lesssim (1 \wedge t)^{-\frac{m}{\alpha} - \frac{d}{\alpha q}} |h|^m,$$

where q is the conjugate Hölder exponent of p.

Proof of Theorem 7.5. Define for $s \geq t - \epsilon$,

$$Y_s^{y,\epsilon} = y + \sigma(y)(Z_t - Z_{t-\epsilon}) \tag{7.4}$$

and set $Y_s^\epsilon = Y_s^{X_{t-\epsilon}, \epsilon}$. Up to the change of the driving process, this is the same as the auxiliary process (2.3).

We proceed as in Section 6. Let p_x be the density of the solution of (7.2) with initial condition x. We know that $p_x(t) \in B_{1,\infty}^a$ for all $a < \kappa$, then by Sobolev's embeddings $p_x \in L^p$ for all $p < \frac{d}{d-\kappa}$. Fix $p \in (1, \frac{d}{d-\kappa})$, and let $\phi \in F_{q,\infty}^\theta(\mathbb{R}^d)$, with $\theta \in (\frac{d}{q}, 1)$, where q is the conjugate Hölder exponent of p. Define Φ as in (6.1), so that $\Delta_h\phi(x) \leq |h|^\theta\Phi(x)$, $\Phi \in L^q(\mathbb{R}^d)$ and $\|\Phi\|_{L^q} = [\phi]_{F_{q,\infty}^\theta}$.

The *probabilistic estimate* is obtained through the previous lemma and the non–degeneracy assumption on σ,

$$\mathbf{\color{red}{Pe}} = \big| \mathbb{E}[\mathbb{E}[\Delta_h^m \phi(Y_\epsilon^{y,\epsilon})]_{y=X_{t-\epsilon}}] \big| \lesssim \epsilon^{-\frac{m}{\alpha} - \frac{d}{\alpha q}} |h|^m \|\phi\|_{L^q}.$$

We turn to the analysis of the *approximation error*. We use the Markov property,

$$\mathbf{\color{red}{Ae}} = \big| \mathbb{E}[\mathbb{E}[\phi(X_\epsilon^y) - \phi(Y_\epsilon^{y,\epsilon})]_{y=X_{t-\epsilon}}] \big|,$$

and by the Hölder inequality, for each $y \in \mathbb{R}^d$,

$$\begin{aligned}
\big| \mathbb{E}[\phi(X_\epsilon^y) - \phi(Y_\epsilon^{y,\epsilon})] \big| &\leq \mathbb{E}[|X_\epsilon^y - Y_\epsilon^{y,\epsilon}|^\theta \Phi(Y_\epsilon^{y,\epsilon})] \\
&\leq \mathbb{E}[|X_\epsilon^y - Y_\epsilon^{y,\epsilon}|^{\theta p}]^{\frac{1}{p}} \mathbb{E}[|\Phi(Y_\epsilon^{y,\epsilon})|^q]^{\frac{1}{q}} \\
&\lesssim \epsilon^{\frac{\theta}{\alpha}(\kappa+1)} \mathbb{E}[|\Phi(Y_\epsilon^{y,\epsilon})|^q]^{\frac{1}{q}}
\end{aligned}$$

where we have used [14, Lemma 3.1-(ii)], and we need $\theta p < \alpha$ (this will limit the final regularity when $\alpha < p$). We thus have,

$$\begin{aligned}
\mathbf{\color{red}{Ae}} &\lesssim \epsilon^{\frac{\theta}{\alpha}(\kappa+1)} \int_{\mathbb{R}^d} \mathbb{E}[|\Phi(Y_\epsilon^{y,\epsilon})|^{q_1}]^{\frac{1}{q_1}} p_x(t - \epsilon, y) \, dy \\
&\leq \epsilon^{\frac{\theta}{\alpha}(\kappa+1)} \|p_x(t - \epsilon)\|_{L^p} \left(\mathbb{E} \int_{\mathbb{R}^d} |\Phi(Y_\epsilon^{y,\epsilon})|^q \, dy \right)^{\frac{1}{q}}.
\end{aligned}$$

First, we need a bound of $\|p_x(t - \epsilon)\|_{L^p}$. By our choice of p, $B_{1,\infty}^a \subset L^p$ for a suitable $a < \kappa$. Proposition 2.2, together with the estimates in [14], provides an estimate in time of $\|p_x(t - \epsilon)\|_{B_{1,\infty}^a}$. With our standard choice $\epsilon \leq \frac{t}{2}$, we readily have that $\|p_x(t - \epsilon)\|_{L^p}$ is bounded by a number that depends on t, but that is uniform in x and ϵ.

It remains to estimate the last term. Denote by $g_{y,\epsilon}$ the density of $\sigma(y)(Z_t - Z_{t-\epsilon})$, and by g_ϵ the density of $Z_t - Z_{t-\epsilon}$. It is easy to see that $g_{y,\epsilon}(z) = \det(\sigma(y))^{-1} g_\epsilon(\sigma(y)^{-1}z)$, thus

$$\mathbb{E}\int_{\mathbb{R}^d} |\Phi(Y_\epsilon^{y,\epsilon})|^q \, dy = \int_{\mathbb{R}^d} |\Phi(z)|^q \Big(\int_{\mathbb{R}^d} \det(\sigma(y))^{-1} g_\epsilon(\sigma(y)^{-1}(z-y)) \, dy \Big) dz,$$

and it is sufficient to prove that the inner integral is uniformly bounded in z. This follows from computations similar to those in Section 6, using the estimate in [14, Lemma 3.3].

In conclusion, **Ae** $\lesssim \epsilon^{\frac{\theta}{\alpha}(1+\kappa)}$, therefore, using the same computations of Proposition 2.2 and the smoothing Lemma A.3, we finally obtain that $p_x(t) \in B^a_{p,\infty}(\mathbb{R}^d)$, for every $a < \kappa\big(1 \wedge \frac{\alpha}{p}\big)$. $\qquad\square$

7.2.2 Rougher drift

In this section we extend the results of Section 5.1 to α-stable-like drivers. We consider problem (7.2) with $b \in L^q(0,T; L^p(\mathbb{R}^d))$ (and same assumptions as before on σ).

As in Section 6, our result is an a–priori estimate on the regularity of the density. In Section 6 the equation we consider has a unique strong solution [30], thus the result is rigorous. Here, under the assumptions on the coefficients we consider, we do not know if there is a solution, or if it is a Markov process. So, as in the above Section 7.2.1, we assume that either there is a solution of (7.2) that is a Markov process, or that there is a solution that can be obtained by approximation of (7.2) with smooth coefficients. In the latter case the Besov bound ensures the uniform integrability of the approximating densities and thus that the limit problem has a density with the same Besov regularity.

Theorem 7.7. *Let $(Z_t)_{t \geq 0}$ be a Lévy process as in Assumption 7.4, with $\alpha > 1$, let $b \in L^q(0,T; L^p(\mathbb{R}^d))$, $\sigma \in C_b^\beta(\mathbb{R}^d; \mathbb{R}^{d \times d'})$, such that (7.3) holds. For every $x \in \mathbb{R}^d$ let $(X_t^x)_{t \geq 0}$ be a solution (as specified above) of (7.2). Assume there is $e \geq 0$ such that*

$$eq' < 1, \qquad \alpha\kappa > 1, \qquad \frac{d}{p} < \alpha\kappa - 1, \qquad e > \frac{\kappa d}{p(\alpha\kappa - 1) - d}, \tag{7.5}$$

where $\kappa = \min(\frac{1}{q'}, \frac{1+\beta}{\alpha}, \frac{1}{\alpha} + \frac{\beta}{q'} - \frac{1}{2}\beta e)$, and q' is the conjugate Hölder exponent of q.

Then for every $t > 0$ and $x \in \mathbb{R}^d$, the random variable X_t^x has a density $p_x(t)$ in $B^b_{1,\infty}(\mathbb{R}^d)$ for every $b < \alpha\kappa - 1$.

We recall that, if $p_x(t)$ is the density of a solution X_t^x of (7.2) with initial condition x, we have set

$$\|p_\cdot\|_{\star,\gamma} = \sup_{t \in (0,T]} \sup_{x \in \mathbb{R}^d} (1 \wedge t)^{e_\gamma} \|p_x(t)\|_{B^\gamma_{1,\infty}},$$

with e_γ to be suitably chosen. Set $a = \frac{d}{p}$, so that $p' = \frac{d}{d-a}$ and $B^a_{1,\infty} \subset L^{p'}$, where p' is the Hölder conjugate exponent of p. We start with some estimates of the contribution of the drift.

Lemma 7.8. *If $s \leq t$ and $\epsilon < 1$, $\epsilon \leq \frac{t}{2}$,*

$$\Big| \mathbb{E}\Big[\int_{t-\epsilon}^s b(r, X_r) \, dr \Big] \Big| \lesssim (1 \wedge t)^{-e_a} \epsilon^{\frac{1}{q'}} \|b\|_{L^q(L^p)} \|p_\cdot\|_{\star,a},$$

where q' is the Hölder conjugate exponent of q. If moreover $q' e_a < 1$, then

$$\mathbb{E}\Big[\Big(\int_{t-\epsilon}^s b(r, X_r) \, dr \Big)^2 \Big] \lesssim (1 \wedge t)^{-e_a} \epsilon^{\frac{2}{q'} - e_a} \|b\|_{L^q(L^p)}^2 \|p_\cdot\|_{\star,a}^2.$$

Proof. The first estimate follows as in Theorem 5.1, since we have,

$$\big|\,\mathbb{E}[b(r, X_s)]\,\big| \lesssim (1 \wedge s)^{-e_a} \|b(r)\|_{L^p} \|p.\|_{\star,a}.$$

Likewise,

$$\mathbb{E}\Big[\Big(\int_{t-\epsilon}^{s} b(r, X_r)\, dr\Big)^2\Big] = 2\int_{t-\epsilon}^{s} dr_1 \int_{r_1}^{t} \mathbb{E}[b(r_1, X_{r_1}) b(r_2, X_{r_2})],$$

and, by the Markov property and the estimate above (twice),

$$\big|\,\mathbb{E}[b(r_1, X_{r_1}) b(r_2, X_{r_2})]\,\big| = \big|\,\mathbb{E}[b(r_1, X_{r_1})\mathbb{E}[b(r_2, X_{r_2-r_1}^y)]_{y=X_{r_1}}]\,\big|$$
$$\lesssim (1 \wedge (r_2 - r_1))^{-e_a} \|b(r_1)\|_{L^p} \|p.\|_{\star,a} \mathbb{E}[|b(r_1, X_{r_1})|]$$
$$\lesssim \|b(r_1)\|_{L^p} \|b(r_2)\|_{L^p} \|p.\|_{\star,a}^2 (1 \wedge (r_2 - r_1))^{-e_a} (1 \wedge r_1)^{-e_a}.$$

Thus, by the Hölder inequality (twice),

$$\mathbb{E}\Big[\Big(\int_{t-\epsilon}^{s} b(r, X_r)\, dr\Big)^2\Big] \lesssim$$

$$\lesssim \|b\|_{L^q(L^p)}^2 \|p.\|_{\star,a}^2 \Big(\int_{t-\epsilon}^{t} (1 \wedge r_1)^{-e_a q'} \int_{r_1}^{t} (1 \wedge (r_2 - r_1))^{-e_a q'}\, dr_2\, dr_1\Big)^{\frac{1}{q'}}$$

$$\lesssim \|b\|_{L^q(L^p)}^2 \|p.\|_{\star,a}^2 (1 \wedge t)^{-e_a} \epsilon^{\frac{2}{q'} - e_a},$$

by elementary computations, since $q' e_a < 1$. $\square$

From the above estimates we immediately deduce the following result.

Lemma 7.9. *If $\gamma < \alpha$ and $s \in [t - \epsilon, t]$,*

$$\mathbb{E}[|X_s - X_{t-\epsilon}|^\gamma]^{\frac{1}{\gamma}} \lesssim (1 + \|\sigma\|_{L^\infty} + \|b\|_{L^q(L^p)})(1 \vee \|p.\|_{\star,a})(1 \wedge t)^{-\frac{e_a}{2}} \epsilon^{\kappa_1},$$

where $\kappa_1 = \min(\frac{1}{\alpha}, \frac{1}{q'} - \frac{e_a}{2})$

Proof. We have that

$$X_s - X_{t-\epsilon} = \int_{t-\epsilon}^{s} b(r, X_r)\, dr + \int_{t-\epsilon}^{s} \sigma(X_{r-})\, dZ_r,$$

thus

$$\mathbb{E}[|X_s - X_{t-\epsilon}|^\gamma] \lesssim \mathbb{E}\Big[\Big|\int_{t-\epsilon}^{s} b(r, X_r)\, dr\Big|^\gamma\Big] + \mathbb{E}\Big[\Big|\int_{t-\epsilon}^{s} \sigma(X_{r-})\, dZ_r\Big|^\gamma\Big] = \boxed{\text{D}} + \boxed{\text{s}}.$$

By [14, Lemma A.2-(i)],

$$\boxed{\text{s}} \lesssim (s - (t - \epsilon))^{\gamma/\alpha} \|\sigma\|_{L^\infty}^\gamma \leq \epsilon^{\gamma/\alpha} \|\sigma\|_{L^\infty}^\gamma$$

For $\boxed{\text{D}}$ we use the previous lemma, since

$$\boxed{\text{D}} \leq \mathbb{E}\Big[\Big|\int_{t-\epsilon}^{s} b(r, X_r)\, dr\Big|^2\Big]^{\frac{\gamma}{2}} \lesssim \|b\|_{L^q(L^p)}^\gamma \|p.\|_{\star,a}^\gamma (1 \wedge t)^{-e_a \gamma/2} \epsilon^{\frac{\gamma}{2}(\frac{2}{q'} - e_a)}.$$ $\square$

Proof of Theorem 7.7. We are ready to estimate the *approximation error* and the *probabilistic estimate*. We use the auxiliary process as in formula (7.4) of previous section. The *probabilistic estimate* can be immediately deduced from Lemma 7.6 (with $p = 1$), thus **Pe** $\lesssim \epsilon^{-m/\alpha} |h|^m \|\phi\|_{L^\infty}$. We turn to the *approximation error*. We have,

$$\textbf{Ae} \lesssim [\phi]_{\mathscr{C}^\theta} \mathbb{E}[|X_t - Y_t^\epsilon|^\theta]$$

$$\lesssim [\phi]_{\mathscr{C}^\theta} \Big(\mathbb{E}\Big|\int_{t-\epsilon}^{t} b(s, X_s)\, ds\Big|^\theta + \mathbb{E}\Big|\int_{t-\epsilon}^{t} (\sigma(X_{s-}) - \sigma(X_{t-\epsilon})\, dZ_s\Big|^\theta\Big).$$

For the first term we use the first statement of Lemma 7.8 (recall that $\theta < 1$),

$$\mathbb{E}\left|\int_{t-\epsilon}^{t} b(s, X_s)\, ds\right|^{\theta} \leq \mathbb{E}\left[\left|\int_{t-\epsilon}^{t} b(s, X_s)\, ds\right|\right]^{\theta} \lesssim (1 \wedge t)^{-e_a\theta} \epsilon^{\frac{\theta}{q'}} \|p.\|_{\star,a}^{\theta}.$$

For the second term we use [14, Lemma A.2-(i)] and Lemma 7.9: let γ be such that $\alpha < \gamma \leq 2$ and $\gamma\beta < \alpha$, then

$$\mathbb{E}\left|\int_{t-\epsilon}^{t} (\sigma(X_{s-}) - \sigma(X_{t-\epsilon})\, dZ_s\right|^{\theta} \lesssim \epsilon^{\theta/\alpha} \sup_{[t-\epsilon,t]} \mathbb{E}[|\sigma(X_{s-}) - \sigma(X_{t-\epsilon})|^{\gamma}]^{\frac{\theta}{\gamma}}$$

$$\lesssim \epsilon^{\theta/\alpha} [\sigma]_{\mathscr{C}^\beta}^{\theta} \sup_{[t-\epsilon,t]} \mathbb{E}[|X_s - X_{t-\epsilon}|^{\gamma\beta}]^{\frac{\theta}{\gamma}}$$

$$\lesssim (1 \wedge t)^{-\frac{1}{2}\beta\theta e_a} (1 \vee \|p.\|_{\star,a})^{\beta\theta} \epsilon^{\theta(\kappa_1\beta + \frac{1}{\alpha})}$$

If we put the two estimates together,

$$\mathbf{Ae} \lesssim (1 \wedge t)^{-\theta e_a} (1 \vee \|p.\|_{\star,a})^{\theta} \epsilon^{\theta\kappa} [\phi]_{\mathscr{C}_b^\theta},$$

where $\kappa = \min(\frac{1}{q'}, \frac{1+\beta}{\alpha}, \frac{1}{\alpha} + \frac{\beta}{q'} - \frac{1}{2}\beta e_a)$. With the positions $K_0 = (1 \wedge t)^{-\alpha e_a}(1 \wedge \|p.\|_{\star,a})^{\alpha}$ and $a_0 = \alpha\kappa - 1$, Proposition 2.2 yields that

$$\|p_x(t)\|_{B_{1,\infty}^a} \lesssim (1 \wedge \|p.\|_{\star,a})^{\frac{a}{a_0}+\alpha\delta}(1 \wedge t)^{-\frac{a}{a_0}e_a - \alpha\delta e_a - \frac{1+a_0}{\alpha a_0}a - \delta},$$

All the computations above show that under the following conditions,

$$q'e_a < 1, \qquad \alpha\kappa > 1, \qquad a = \frac{d}{p} < a_0,$$

$$e_a > \frac{a}{a_0}e_a + \frac{1+a_0}{\alpha a_0}a,$$

the norm $\|p.\|_{\star,a}$ is finite, and thus $\|p.\|_{\star,b} < \infty$ for every $b < a_0$, for a suitable value of e_b that can be easily computed as above by Proposition 2.2. □

It is not immediately apparent that the conditions of Theorem 7.7 may be verified, so we provide a couple of particular cases. The first matches Theorem 5.1. Notice that under slightly different assumptions on the drift b that are essentially the same as those in the corollary below, existence and uniqueness hold, see [11].

Corollary 7.10. *Under the assumptions of Theorem 7.7 on the solution and on the coefficients, assume additionally that σ is constant. Then the parameters κ, e can be chosen as $\kappa = \frac{1}{q'}$ and $e = 0$. In particular, the conclusions of the theorem hold if*

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

The second particular case applies for p large. In that case we expect the number e of Theorem 7.7 to be small. As a matter of facts, $e \to 0$ as $p \to \infty$. To further simplify, we consider the case $\kappa = \frac{1}{q'}$.

Corollary 7.11. *Under the same assumptions of Theorem 7.7, if*

- $\frac{2d}{p} + \frac{\alpha}{q} < \alpha - 1$,
- $q \geq \frac{\alpha}{2-\alpha}$,
- *either $\beta \geq \alpha - 1$, or $\beta < \alpha - 1$ and $q \leq \frac{\alpha}{\alpha-\beta-1}$,*

then any number e such that

$$\frac{\kappa d}{p(\alpha\kappa - 1) - d} < e < \frac{1}{q'}$$

meets the conditions (7.5)*, and thus the conclusions of Theorem 7.7 hold.*

Proof. Let us first prove that $\kappa = \frac{1}{q'}$ if $eq' < 1$. Indeed, by the third of the assumptions above, $\frac{1}{q'} \leq \frac{\beta+1}{\alpha}$. Moreover, by the second of the assumptions above, $\frac{1}{q'} \leq 2(\frac{1}{q'} - \frac{1}{\alpha})$, thus if $eq' < 1$, then $e < 2(\frac{1}{q'} - \frac{1}{\alpha})$, therefore $\frac{1+\beta}{\alpha} \leq \frac{1}{\alpha} + \frac{\beta}{q'} - \frac{1}{2}\beta e$, and in conclusion $\kappa = \frac{1}{q'}$.

The condition $\alpha\kappa > 1$ holds since $\beta > 0$ and $q' < \alpha$. The condition $\frac{d}{p} < \alpha\kappa - 1$ holds by the first of the assumptions of the corollary. To ensure that e can be chosen in a non–empty interval, we need to check that

$$\frac{\kappa d}{p(\alpha\kappa - 1) - d} < \frac{1}{q'},$$

and this follows from the first assumption of the corollary. □

7.3 The 3D Navier–Stokes equations with noise

The problem of existence of densities for finite dimensional projections of the solutions of the Navier-Stokes equations driven by noise has been the motivating example, discussed in [15], that has led to the development of the dimension-free method illustrated in this paper. The results have been further improved in [43] proving Hölder regularity in time with values in Besov spaces in time, and in [41] proving Hölder regularity in space of the densities. Unlike Section 6, the result of Hölder regularity is optimal but has been proved by analytical methods.

7.4 A singular equation: Φ_d^4

In this section we consider the following singular stochastic PDE,

$$\partial_t X = \Delta X + X - X^3 + \xi, \tag{7.6}$$

on the torus $\mathbb{T}_d$, with periodic boundary conditions, in two dimensions (but see Remarks 7.14 and 7.19 for the three dimensional case), where ξ is space-time white noise.

The equation is generally understood as the limit of a family of regularized problems. To this end, let η be a smooth compactly supported function, set $\eta_\delta(x) = \delta^{-d}\eta(\delta^{-1}x)$, and $\xi_\delta = \eta_\delta \star \xi$. Consider

$$\partial_t X_\delta - \Delta X_\delta = -X_\delta^3 + (1 + 3c_\delta)X_\delta + \xi_\delta, \tag{7.7}$$

where c_δ is a suitable number that depends on δ and η. Define I_δ, V_δ, and W_δ as

$$(\partial_t - \Delta)\mathsf{I}_\delta = -\mathsf{I}_\delta + \xi_\delta, \qquad \mathsf{V}_\delta = \mathsf{I}_\delta^2 - c_\delta, \qquad \mathsf{W}_\delta = \mathsf{I}_\delta^3 - 3c_\delta\mathsf{I}_\delta.$$

where I_δ is chosen as the stationary process that solves the above equation. With this choice, V_δ and W_δ are also stationary processes. The constants c_δ are chosen so that V_δ and W_δ have a limit as $\delta \downarrow 0$, see [12, 35] in dimension two, and [21, 22, 23, 10, 37] for the three dimensional case. We will denote by $\mathsf{I}, \mathsf{V}, \mathsf{W}$ the limit quantities as $\delta \downarrow 0$.

7.4.1 Densities for the solution

Set $X_\delta = \mathsf{I}_\delta + R_\delta$, then R_δ solves

$$(\partial_t - \Delta)R_\delta = -R_\delta^3 - 3\mathsf{I}_\delta R_\delta^2 + (1 - 3\mathsf{V}_\delta)R_\delta + 2\mathsf{I}_\delta - \mathsf{W}_\delta.$$

We prove that the solution has a density at each time. Since the solution is distribution valued, we will prove the existence of a joint density of the solution tested over a finite but arbitrary number of test functions. Here we denote by $\langle \cdot, \cdot \rangle$ the evaluation of distributions over test functions.

Theorem 7.12. *Let X be a solution of (7.6) on the two dimensional torus, fix $t > 0$ and an integer $n \geq 1$, and let $\varphi_1, \varphi_2, \ldots, \varphi_n$ be smooth periodic functions such that the matrix $(\langle \varphi_i, \varphi_j \rangle)_{i,j=1,\ldots,n}$ is non–singular.*

Then the random vector $(\langle X(t), \varphi_1 \rangle, \langle X(t), \varphi_1 \rangle, \ldots, \langle X(t), \varphi_n \rangle)$ has a density with respect to the Lebesgue measure in $B_{1,\infty}^a(\mathbb{R}^n)$, for every $a < 1$.

Proof. We will perform our estimates on the solution of (7.7). The Besov bound of the densities will ensure the existence of a limit density.

Equation (7.7) in mild form and evaluated on each test function φ_i, $i = 1, 2, \ldots, n$, yields the evolution

$$dX_{\delta,i} = \langle \Delta X_\delta + (1 + 3c_\delta)X_\delta - X_\delta^3, \varphi_i \rangle + \langle \varphi_i, \xi_\delta \rangle,$$

of the random vector $(X_{\delta,1}, X_{\delta,2}, \ldots, X_{\delta,n})$, defined as $X_{\delta,i}(t) = \langle X_\delta(t), \varphi_i \rangle$.

Fix $t > 0$, then the auxiliary process for the simple method (the term analogous to (2.3)) here is given as $dY_{\delta,i}^\epsilon = \langle \varphi_i, \xi_\delta \rangle$. By the non–degeneracy assumption on the test functions it follows that the random vector $(\langle \varphi_1, \xi_\delta \rangle, \langle \varphi_2, \xi_\delta \rangle, \ldots, \langle \varphi_n, \xi_\delta \rangle)_{i,j=1,2,\ldots,n}$ is a non–degenerate Gaussian random vector, for δ small enough. It easily follows then that the *probabilistic estimate* is given by **Pe** $\lesssim \epsilon^{-m/2}|h|^m\|\phi\|_{L^\infty}$.

For the *approximation error* we need to compute

$$\mathbb{E} \int_{t-\epsilon}^t \langle \Delta X_\delta + (1 + 3c_\delta)X_\delta - X_\delta^3, \varphi_i \rangle\, ds.$$

This is immediate, since X_δ has moments in (negative) Besov spaces [35]. This is read in terms of uniform bounds on moments of X_δ in negative Besov spaces. In conclusion **Ae** $\lesssim \epsilon^\alpha[\phi]_{\mathscr{C}_b^\alpha}$. Proposition 2.2 concludes the proof. □

Remark 7.13. We actually expect that the densities in the previous proposition should be smooth. Following the lines of Section 2.2, one could define a infinite dimensional auxiliary process Y^ϵ with the drift frozen at $t - \epsilon$, as in formula (2.13). This would allow to compute an *approximation error* at the level of the infinite dimensional processes in the appropriate Besov spaces of distributions. When evaluated over test functions, this error would provide the *approximation error* for the method.

Remark 7.14. The above result holds also in dimension three, with the same proof. The difference is that, as we shall see below, the correct interpretation of the equation is more involved.

In dimension three we additionally introduce the diagram $\mathbf{\Psi}_\delta$, the stationary solution of $(\partial_t - \Delta)\mathbf{\Psi}_\delta = \mathbf{\Psi}_\delta$. Again, the constant c_δ is chosen so that $\mathbf{V}_\delta$ and $\mathbf{\Psi}_\delta$ have a limit as $\delta \downarrow 0$, see [37]. Set $X_\delta = \mathbf{I}_\delta - \mathbf{\Psi}_\delta + R_\delta$, then R_δ solves

$$(\partial_t - \Delta)R_\delta = -R_\delta^3 - 3(R_\delta - \mathbf{\Psi}_\delta)\mathbf{V}_\delta + P_\delta(R_\delta),$$

where

$$P_\delta(R_\delta) = 3(\mathbf{\Psi}_\delta - \mathbf{I}_\delta)R_\delta^2 + (1 + 6\mathbf{I}_\delta\mathbf{\Psi}_\delta - 3\mathbf{\Psi}_\delta^2)R_\delta + (\mathbf{I}_\delta - \mathbf{\Psi}_\delta) + \mathbf{\Psi}_\delta^3 - 3\mathbf{I}_\delta\mathbf{\Psi}_\delta^2.$$

This is not enough yet to give a meaning to the limit equation, because the term $R_\delta\mathbf{V}_\delta$ is not well defined in the limit, given the expected regularity of the limit R. Without

giving too many details (that would be beyond the scope of this paper), we know that the equation can be suitably reformulated following for instance [37] (see also [10]) as,

$$(\partial_t - \Delta)R_\delta = -R_\delta^3 - 3R_\delta^H \ominus \mathsf{V}_\delta - 3R_\delta \oslash \mathsf{V}_\delta + Q_\delta(R_\delta) + 3\mathsf{V}_\delta \oslash \mathsf{V}_\delta$$

where $R_\delta = R_\delta^L + R_\delta^H$, R_δ^L is solution of

$$(\partial_t - \Delta)R_\delta^L = -3(R_\delta - \mathsf{V}_\delta) \oslash \mathsf{V}_\delta,$$

Q_δ gathers more regular terms, and $\ominus$, $\oslash$ are defined in terms of the Bony paraproduct (see for instance [8]).

At this stage one can proceed as in the two-dimensional case, since the stochastic diagrams as well as the remainder have uniform bounds in time on moments in negative Besov spaces. Thus, Theorem 7.12 holds for (7.6) also in dimension $d = 3$.

Corollary 7.15. *Let X be a solution of (7.6) on the three dimensional torus, fix $t > 0$ and an integer $n \geq 1$, and let $\varphi_1, \varphi_2, \ldots, \varphi_n$ be smooth periodic functions such that the matrix $(\langle \varphi_i, \varphi_j \rangle)_{i,j=1,\ldots,n}$ is non–singular.*

Then the random vector $(\langle X(t), \varphi_1 \rangle, \langle X(t), \varphi_1 \rangle, \ldots, \langle X(t), \varphi_n \rangle)$ has a density with respect to the Lebesgue measure in $B_{1,\infty}^a(\mathbb{R}^n)$, for every $a < 1$.

7.4.2 Densities for the remainder

In this section we wish to delve into another direction. As a second application of the simple method in this framework, we wish to investigate the existence of a density for the remainder R. A possible approach could be based on the previous considerations and some arguments from Section 3.2. Indeed, it would be sufficient to prove (for instance in dimension two) that (X, I) has a joint density (when evaluated over test functions). Since both X and I are driven by the same noise, this can be only possible if the drift is hypoelliptic. This idea has two drawbacks: the first is that it would give densities for R against test functions, while R is a *bona fide* function. The second is that, as we have seen in Section 3.2, the method requires good estimates on the drift, while giving back low regularity.

We wish to consider here a different idea, that uses the non–linearity directly. The equation for R is (in dimension two),

$$(\partial_t - \Delta)R = -R^3 - 3\mathsf{I}R^2 + (1 - 3\mathsf{V})R + 2\mathsf{I} - \mathsf{V}.$$

Our idea is to understand the terms not depending on R as the "noise" of the evolution, since our simple method presented in this paper is essentially based on small time estimates for the density of the "noise" object.

Unfortunately the idea does not work here in dimension two (and apparently in dimension three as well, see Remark 7.19), since I and V are "too good", that is with regularity comparable with the "non–noise" terms $R^3 + 3\mathsf{I}R^2 - (1 - 3\mathsf{V})R$. As such, we would end up with a formula similar to (2.9), but with $a_0 = 0$. For this reason in the rest of the section we will discuss an example very close to problem (7.6) and that is amenable to our analysis.

We consider the following problem,

$$(\partial_t - \Delta)X = -X^3 + (-\Delta)^{\gamma/2}\xi, \tag{7.8}$$

with periodic boundary conditions on the two dimensional torus, where $\gamma \in (0, \frac{2}{5})$ and ξ is space-time white noise. As above, the problem makes sense when suitably renormalized

or as a limit of approximated problems. If ξ_δ is a spatial smooth approximation of the noise, we study the problem

$$(\partial_t - \Delta)X_\delta = -X_\delta^3 + 3c_\delta(t)X_\delta + (-\Delta)^{\gamma/2}\xi_\delta,$$

where $c_\delta(t) = \mathbb{E}[\mathord{\uparrow}_\delta(t)^2]$, and $\mathord{\uparrow}_\delta$ solves $(\partial_t - \Delta)\mathord{\uparrow}_\delta = (-\Delta)^{\gamma/2}\xi_\delta$. Set $\mathbf{V}_\delta(t) = \mathord{\uparrow}_\delta(t)^2 - c_\delta(t)$ and $\mathbf{W}_\delta(t) = \mathord{\uparrow}_\delta(t)^3 - 3c_\delta(t)\mathord{\uparrow}_\delta(t)$. Denote respectively by $\mathord{\uparrow}$, $\mathbf{V}$ and $\mathbf{W}$ their limits as $\delta \to \infty$. From now on we will drop for simplicity the subscript δ, even though all computations are rigorous only at the level of approximations. The Besov bound from our method will provide the uniform integrability necessary to bring the argument rigorously in the limit as $\delta \to 0$ to the solution of (7.8).

The choice $\gamma < \frac{2}{5}$ is for convenience. In this regime it is sufficient to decompose the solution as $X = \mathord{\uparrow} + R$, and the remainder solves

$$(\partial_t - \Delta)R = -R^3 - 3\mathord{\uparrow}R^2 - 3\mathbf{V}R - \mathbf{W}.$$

One can prove, for instance following the lines of [36, Theorem 1.1], that $\mathord{\uparrow}$ has regularity $\mathscr{C}^{-\gamma-}$, $\mathbf{V}$ has regularity $\mathscr{C}^{-2\gamma-}$ and $\mathbf{W}$ has regularity $\mathscr{C}^{-3\gamma-}$. Thus we expect that R has regularity $\mathscr{C}^{(2-3\gamma)-}$, so that $\mathord{\uparrow}R^2$ and $\mathbf{V}R$ are well defined when $\gamma < \frac{2}{5}$[1]. In the rest of the section we will *assume* that we have a solution for the auxiliary equation with the above mentioned regularity. In principle the solution might be defined only up to a random time, but on the one hand we may guess that the results of [37] extend to this case, and on the other hand our method can take local solutions into account as well (as in Section 2.1).

The main result for the densities of the remainder is as follows.

Theorem 7.16. *Let X be a solution of (7.8), and set $R_t = X_t - \mathord{\uparrow}_t$. Then for every $t > 0$ and every $x \in \mathbb{T}_2$ the random variable $R_t(x)$ has a density with respect to the Lebesgue measure on $\mathbb{R}$. Moreover the density is in $B_{1,\infty}^a$ for every $a < \frac{\gamma}{2-3\gamma}$.*

To prove the theorem we first identify the approximation problem, as in formula (2.3) for the toy problem. Fix $0 < s < t$ and write the equation for R in mild form,

$$R_t = e^{\Delta(t-s)}R_s - \int_s^t e^{\Delta(t-r)}(R_r^3 + 3\mathord{\uparrow}_r R_r^2 + 3\mathbf{V}_r R_r)\,dr - \int_s^t e^{\Delta(t-r)}\mathbf{W}_r\,dr.$$

The terms in the first integral in the formula above are more regular than the second integral, thus should provide a smaller *approximation error*. We will treat the last term as "noise", although clearly this new "noise" has no independent increments. We define our auxiliary process for $s \leq r \leq t$, as

$$S_r = e^{\Delta(r-s)}R_s - \int_s^r e^{\Delta(r-u)}\mathbf{W}_u\,du. \tag{7.9}$$

Lemma 7.17. *Under the standing assumptions, given $\phi \in \mathscr{C}^\alpha$ for some $\alpha \in (0,1)$,*

$$\mathbf{\textit{Ae}} \approx \big| \mathbb{E}[\phi(R_t(x))] - \mathbb{E}[\phi[S_t(x)]] \big| \lesssim [\phi]_{\mathscr{C}^\alpha}(t-s)^{1-\gamma-\delta},$$

for every $0 \leq s < t$, $x \in \mathbb{T}_2$, and every $\delta > 0$.

Proof. We have that

$$S_t - R_t = \int_s^t e^{\Delta(t-r)}(R_r^3 + 3\mathord{\uparrow}_r R_r^2 + 3\mathbf{V}_r R_r)\,dr \tag{7.10}$$

---

[1] If γ is above $\frac{2}{5}$ but below $\frac{1}{2}$, we need to decompose X with the additional term $\mathbf{Y}$. The case $\gamma = \frac{1}{2}$ is, analytically, equivalent to Φ_3^4.

Among the terms in the right-hand-side of formula above, the least regular is the one containing $\mathsf{V}_r R_r$. By [36, Proposition 2.3],

$$\|\mathsf{V}_r R_r\|_{\mathscr{C}^{-2\gamma-\delta}} \lesssim \|\mathsf{V}_r\|_{\mathscr{C}^{-2\gamma-\delta}} \|R_r\|_{\mathscr{C}^{2-3\gamma-\delta}},$$

with δ such that $0 < \delta < 1 - \frac{5}{2}\gamma$. Therefore,

$$\left|\left(\int_s^t e^{\Delta(t-r)}\,\mathsf{V}_r R_r\,dr\right)(x)\right| \lesssim \left(\int_s^t (t-r)^{-\gamma-\delta}\,dr\right)\left(\sup_{[0,t]}\|R\|_{\mathscr{C}^{2-3\gamma-\delta}}\right)\left(\sup_{[0,t]}\|\mathsf{V}\|_{\mathscr{C}^{-2\gamma-\delta}}\right)$$
$$\lesssim (t-s)^{1-\gamma-\delta}.$$

The terms containing R_r^3 and $\mathsf{I}_r R_r^2$ are more regular and give a smaller contribution in terms of powers of $t-s$. $\qquad\square$

We wish now to "extract" the component of our "noise" that is independent of the past (that is before time s). To this end notice that if $s < r < t$, then

$$\mathsf{I}_r = e^{\Delta(r-s)}\,\mathsf{I}_s + \int_s^r e^{\Delta(r-u)}\,dW_u =: \mathsf{I}_{s\to r} + \mathsf{I}_{s,r},$$

with $\mathsf{I}_{s\to r}$ measurable with respect to the history up to time s (and smooth for $r > s$), and $\mathsf{I}_{s,r}$ independent of the history up to time s. By squaring,

$$\mathsf{V}_r = \mathsf{I}_r^2 - \mathbb{E}[\mathsf{I}_r^2] = \left(\mathsf{I}_{s\to r}^2 - \mathbb{E}[\mathsf{I}_{s\to r}^2]\right) + 2\mathsf{I}_{s\to r}\mathsf{I}_{s,r} + \left(\mathsf{I}_{s,r}^2 - \mathbb{E}[\mathsf{I}_{s,r}^2]\right),$$

and we set

$$\mathsf{V}_{s,r} = \mathsf{I}_{s,r}^2 - \mathbb{E}[\mathsf{I}_{s,r}^2], \qquad \mathsf{V}_{s\to r} = \mathsf{I}_{s\to r}^2 - \mathbb{E}[\mathsf{I}_{s\to r}^2].$$

Likewise,

$$\mathsf{W}_r = \mathsf{I}_r^3 - 3\mathbb{E}[\mathsf{I}_r^2]\mathsf{I}_r$$
$$= \left(\mathsf{I}_{s\to r}^3 - 3\mathbb{E}[\mathsf{I}_{s\to r}^2]\mathsf{I}_{s\to r}\right) + 3\mathsf{V}_{s\to r}\mathsf{I}_{s,r} + 3\mathsf{I}_{s\to r}\mathsf{V}_{s,r} + \left(\mathsf{I}_{s,r}^3 - 3\mathsf{I}_{s,r}\mathbb{E}[\mathsf{I}_{s,r}^2]\right),$$

and we set

$$\mathsf{W}_{s,r} = \mathsf{I}_{s,r}^3 - 3\mathsf{I}_{s,r}\mathbb{E}[\mathsf{I}_{s,r}^2], \qquad \mathsf{W}_{s\to r} = \mathsf{I}_{s\to r}^3 - 3\mathbb{E}[\mathsf{I}_{s\to r}^2]\mathsf{I}_{s\to r}.$$

In conclusion,

$$S_t = e^{\Delta(t-s)}\,R_s - \int_s^t e^{\Delta(t-r)}\left(\mathsf{W}_{s\to r} + 3\mathsf{V}_{s\to r}\mathsf{I}_{s,r} + 3\mathsf{I}_{s\to r}\mathsf{V}_{s,r} + \mathsf{W}_{s,r}\right)dr.$$

Lemma 7.18. *Under the standing assumptions, given $\phi \in L^\infty$,*

$$\textbf{\textit{Pe}} = \left|\mathbb{E}[\Delta_h^m \phi(S_t(x))]\right| \lesssim \|\phi\|_{L^\infty}\frac{|h|^m}{(t-s)^{m(1-\frac{3}{2}\gamma+\delta)}},$$

for every $0 \le s < t$, $x \in \mathbb{T}_2$, $m \ge 1$, $h \in \mathbb{T}_2$ with $|h| \le 1$, and every $\delta > 0$.

Proof. Since

$$\mathbb{E}[\Delta_h^m \phi(S_t(x))] = \mathbb{E}\big[\mathbb{E}[\Delta_h^m \phi(S_t(x))\,|\,\mathscr{F}_s]\big], \tag{7.11}$$

where $\mathscr{F}_s$ is the σ–field of events until time s, the result follows if we can estimate the small time asymptotic of the L^1 norm of the density (and of its derivatives) of $S_t(x)$ given $\mathscr{F}_s$.

To this end we recall that for a real random variable X with density g, we have by [38, Theorem 2.1.4], the representation

$$D^m g(x) = \mathbb{E}[\mathbb{1}_{\{X>x\}} H_{(1,1,\ldots,1)}],$$

(there are $(m+1)$ ones in the subscript above) for the density and its derivatives ($m \geq 0$), as well as

$$D^m g(x) = \mathbb{E}[\mathbb{1}_{\{X<x\}} H_{(1,1,\ldots,1)}],$$

where the terms $H.$ come from integration by parts and are defined in [38, Proposition 2.1.4]. Moreover, by [38, formula (2.28)],

$$\mathbb{E}[|H_{(1,1,\ldots,1)}|^p]^{\frac{1}{p}} \lesssim \|\mathcal{M}_X^{-1} \mathcal{D} X\|_{m+1,2^m r}^{m+1},$$

for every $r > p$, where $\mathcal{D}$ is the Malliavin derivative, and $\mathcal{M}_X$ is the Malliavin matrix of X. By integrating over x, we see that

$$\|D^m g\|_{L^1} \lesssim \mathbb{E}[|X| \, |H_{(1,1,\ldots,1)}|], \tag{7.12}$$

so that our task is to understand how this expectation scales in powers of $(t-s)$ when we take

$$X = -\int_0^{t-s} e^{\Delta(t-s-r)} \big(U_3(r) + 3U_2(r)\mathsf{I}_r + 3U_1(r)\mathsf{V}_r + \mathsf{V}_r\big)\, dr,$$

and U_i, $i = 0,\ldots,3$ are given elements with U_i uniformly bounded in time with values in $\mathscr{C}^{-i\cdot\gamma-\delta}$ for $i = 1,2,3$, for every $\delta > 0$. Notice that the choice of X above corresponds, up to an additive constant, to S_t when conditioned over $\mathscr{F}_s$, and with a translation in time of the stochastic diagrams (that does not change the law). By the chain rule for the Malliavin derivative,

$$\mathcal{D} X = -\int_0^{t-s} e^{\Delta(t-s-r)} \big(3U_2(r)\mathcal{D}\mathsf{I}_r + 6U_1(r)\mathsf{I}_r \mathcal{D}\mathsf{I}_r + 3\mathsf{V}_r \mathcal{D}\mathsf{I}_r\big)\, dr,$$

and similarly for the second and third derivative (the fourth derivative is zero since the random variable X above is in the third Wiener chaos). In particular,

$$\mathcal{D}_u \mathsf{I}_r = \mathbb{1}_{\{u \leq r\}} e^{\Delta(r-u)} \sum |k|^\gamma e_k,$$

where $(e_k)_{k \in \mathbb{Z}^2}$ is the Fourier basis of complex exponentials.

We see that for every small $\delta > 0$,

$$\mathbb{E}\Big[\Big|\Big(\int_0^{t-s} e^{\Delta(t-s-r)} \mathsf{V}_r\Big)(x)\Big|^2\Big] \approx (t-s)^{2-3\gamma-2\delta},$$

and likewise,

$$\mathbb{E}\Big[\Big|\Big(\int_0^{t-s} e^{\Delta(t-s-r)} U_1(r) \mathsf{V}_r\Big)(x)\Big|^2\Big] \approx (t-s)^{2-2\gamma-2\delta} \sup_{[0,t-s]} \|U_1\|_{\mathscr{C}^{2\gamma+\delta}}.$$

Notice that when we will evaluate the external expectation in (7.11), we will see that $\sup_{[0,t-s]} \|U_1\|_{\mathscr{C}^{2\gamma+\delta}} \approx (t-s)^{-\frac{\gamma}{2}}$ (since we will replace $U_1(r)$ by $\mathsf{I}_{s\to r}$). In conclusion X scales as $(t-s)^{1-\frac{3}{2}\gamma}$, up to small corrections. With similar computations we see that $\mathcal{D} X$ scales also as $(t-s)^{1-\frac{3}{2}\gamma}$ (due to the additional contribution of $\mathcal{D}\mathsf{I}$), as well as the second and third Malliavin derivatives. Therefore, using formula (7.12), the conclusion of the lemma follows. $\qquad\square$

We are ready to complete the proof of the main result of this section.

Proof of Theorem 7.16. In view of Lemma 7.17 and Lemma 7.18, with the choice $s = t - \epsilon$, this is a simple application of Proposition 2.2, with $\theta = 1/(1 - \frac{3}{2}\gamma + \delta)$ and $a_0 = (1 - \gamma)\theta - 1$. $\qquad\square$

Remark 7.19. One can expect that Theorem 7.16 might hold also for (7.6) in dimension three. In that case the role of the "noise" should be played by

$$\int_s^t e^{\Delta(t-r)} \mathbf{V}_r \oslash \mathbf{V}_r \, dr.$$

Since the remainder has regularity $\mathscr{C}^{1-}$, the *approximation error* should be of the order $\mathbf{Ae} \approx \epsilon^{\frac{1}{2}-}$, while the *probabilistic estimate* has a density with increments of order $|h|/\sqrt{\epsilon}$. This prevents the application of Proposition 2.2 (in short, it would correspond to the case $a_0 = 0$).

To circumvent the problem, one possibility could be to use Hölder continuity in time of the stochastic diagrams. Unfortunately this would deteriorate the space regularity (see for instance [36, Theorem 1.1]), yielding a final dependence of the *approximation error* of order $\mathbf{Ae} \approx \epsilon^{1/2-}$ (or worse). Another possibility could be to improve the *approximation error* with a smarter definition of the auxiliary process, as in Section 2.2.

A   Weaker versions of the smoothing lemma

In the first pages of Malliavin's seminal paper [33] on a probabilistic proof of the Hörmander theorem there is a classical smoothing lemma. This is the link between the existence of a density and probabilistic integration by parts and the Malliavin calculus. The lemma says roughly that if for a $\mathbb{R}^d$-valued random variable X,

$$\big| \mathbb{E}[D^\alpha \phi(X)] \big| \leq c_n \|\phi\|_{L^\infty},$$

for all $|\alpha| \leq n$ and all test functions ϕ, then X has a density, with respect to the Lebesgue measure, in C^{n-d-1}. In this section we give a generalization of this lemma in Besov spaces.

A.1   Besov and Triebel–Lizorkin spaces

Besov spaces, together with the Triebel–Lizorkin spaces, are a scale of function spaces introduced to capture the fine properties of regularity of functions, beyond on the one hand the Sobolev spaces, and on the other hand the spaces of continuous functions. Indeed, Besov spaces contain both. The main references we shall use on this subject are [47, 48].

A general definition with the Littlewood–Paley decomposition is (briefly) as follows. Let $(\varphi_n)_{n \geq 0}$ be a band–limited decomposition of the frequency space. For a distribution f each $f_n = \mathcal{F}^{-1}(\varphi_n \hat{f})$ is a Schwartz function and $f = \sum_n f_n$. Then

$$\|f\|_{B_{p,q}^s} := \big\| (2^{ns} \|f_n\|_{L^p})_{n \geq 0} \big\|_{\ell^q}, \qquad \text{and} \qquad \|f\|_{F_{p,q}^s} := \big\| \|(2^{ns} f_n)_{n \geq 0}\|_{\ell^q} \big\|_{L^p},$$

($p < \infty$ for the $F_{p,q}^s$ norm). Notice that to define the $B_{p,q}^s$ norm, the Littlewood–Paley decomposition is first averaged over position, and then over frequencies, while the opposite happens for the $F_{p,q}^s$ norm. We define the spaces $B_{p,q}^s(\mathbb{R}^d)$ and $F_{p,q}^s(\mathbb{R}^d)$, with $s \in \mathbb{R}$ and $1 \leq p, q \leq \infty$ (with the exception of $p = \infty$ for the F–space) as the *closure of the Schwartz space* with respect to the above norms. In this way the spaces we obtain are separable, regardless of the index (but this is an issue only if $p = \infty$ or $q = \infty$). Since $\|\cdot\|_{B_{p,p}^s} = \|\cdot\|_{F_{p,p}^s}$ for all p, we define $F_{\infty,\infty}^s := B_{\infty,\infty}^s$. The spaces obtained do not depend on the band–limited decomposition (although the norms do), and different choices of the decomposition give raise to equivalent norms.

A.1.1 Definition via the difference operator

Given $\alpha > 0$, we shall denote by $C^\alpha(\mathbb{R}^d)$ the standard Hölder space, namely the space of functions with $[\alpha]$ derivatives such that the derivatives of order $[\alpha]$ are Hölder continuous of exponent $(\alpha - [\alpha])$.

A special role in this paper is played by the *Zygmund spaces* $\mathscr{C}_b^\alpha(\mathbb{R}^d) = B_{\infty,\infty}^\alpha(\mathbb{R}^d)$ that, for non–integer values of α, coincide with the (separable version of the) Hölder spaces. With this in mind, we recall an alternative definition of Besov spaces that is better suited for our purposes (see [47, Theorem 2.5.12] or [48, Theorem 2.6.1] for further details). Define

$$(\Delta_h^1 f)(x) = f(x+h) - f(x),$$

$$(\Delta_h^n f)(x) = \Delta_h^1(\Delta_h^{n-1} f)(x) = \sum_{j=0}^n (-1)^{n-j} \binom{n}{j} f(x+jh),$$

then the following norms, for $s > 0$, $1 \le p \le \infty$, $1 \le q \le \infty$,

$$\|f\|_{L^p} + [f]_{B_{p,q}^s}$$

are equivalent norms of $B_{p,q}^s(\mathbb{R}^d)$ for the given range of parameters. Here we have set

$$[f]_{B_{p,q}^s} := \left\| h \mapsto \frac{\|\Delta_h^m f\|_{L^p}}{|h|^s} \right\|_{L^q\left(B_1(0); \frac{dh}{|h|^d}\right)}. \tag{A.1}$$

where m is any integer such that $s < m$, and $B_1(0)$ is the unit ball in $\mathbb{R}^d$.

A similar definition can be given for Triebel–Lizorkin spaces (see [47, Theorem 2.5.10]). Let $1 \le p < \infty$, $1 \le q \le \infty$, and $s > \frac{d}{p \wedge q}$, and set

$$[f]_{F_{p,q}^s} := \left\| x \mapsto \left\| h \mapsto |h|^{-s} \Delta_h^m f(x) \right\|_{L^q\left(\mathbb{R}^d; \frac{dh}{|h|^d}\right)} \right\|_{L^p} \tag{A.2}$$

where m is any integer such that $m > s$. Then $\|f\|_{L^p} + [f]_{F_{p,q}^s}$ is an equivalent norm in $F_{p,q}^s(\mathbb{R}^d)$. Unfortunately the representation in terms of differences only holds for s large.

A.1.2 Besov spaces on domains

Given a bounded domain D with smooth boundary, define for $s \in \mathbb{R}$ and $1 \le p, q \le \infty$,

$$B_{p,q}^s(D) := \{f : \text{there is } g \in B_{p,q}^s \text{ such that } g|_D = f\},$$

with norm

$$\|f\|_{B_{p,q}^s(D)} := \inf\{\|g\|_{B_{p,q}^s} : g \in B_{p,q}^s(\mathbb{R}^d) \text{ and } g|_D = f\}.$$

See [47, Chapter 3] for further details.

A.2 Smoothing results

This lemma, in a weaker form, is implicitly given in [15] (see also [43, Lemma 4.1], and see [14, Lemma 2.1] for a real analytic proof).

Lemma A.1 (fractional integration by parts). *Let X be a $\mathbb{R}^d$–valued random variable. If there are an integer $m \ge 1$, a real number $s > 0$, a real $\alpha > 0$, with $\alpha < s < m$, and a constant $K > 0$ such that for every $\phi \in \mathscr{C}_b^\alpha(\mathbb{R}^d)$ and $h \in \mathbb{R}^d$, with $|h| \le 1$,*

$$\left| \mathbb{E}[\Delta_h^m \phi(X)] \right| \le K |h|^s \|\phi\|_{\mathscr{C}_b^\alpha},$$

then X has a density f_X with respect to the Lebesgue measure on $\mathbb{R}^d$. Moreover $f_X \in B_{1,\infty}^{s-\alpha}(\mathbb{R}^d)$ and

$$\|f\|_{B_{1,\infty}^{s-\alpha}} \lesssim (1 + K). \tag{A.3}$$

Proof. Let μ be the law of X. Fix a smooth function ϕ. Let $(\varphi_\epsilon)_{\epsilon>0}$ be a smoothing kernel, namely $\varphi_\epsilon = \epsilon^{-d}\varphi(x/\epsilon)$, with $\varphi \in C_c^\infty(\mathbb{R}^d)$, $0 \leq \varphi \leq 1$, and $\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$. Let $f_\epsilon = \varphi_\epsilon \star \mu$, then easy computations show that $f_\epsilon \geq 0$, $\int_{\mathbb{R}^d} f_\epsilon(x)\,dx = 1$ and that

$$\left| \int_{\mathbb{R}^d} \Delta_h^m \phi(x) f_\epsilon(x)\,dx \right| = \left| \int \varphi_\epsilon(x) \mathbb{E}[\Delta_h^m \phi(x-X)]\,dx \right| \leq K|h|^s \|\phi\|_{\mathscr{C}_b^\alpha}.$$

On the other hand, by a simple change of variables,

$$\int_{\mathbb{R}^d} \Delta_h^m \phi(x) f_\epsilon(x)\,dx = \int_{\mathbb{R}^d} \Delta_{-h}^m f_\epsilon(x) \phi(x)\,dx. \tag{A.4}$$

Set $g_\epsilon = (I - \Delta_d)^{-\alpha/2} f_\epsilon$, and $\psi = (I - \Delta_d)^{\alpha/2}\phi$, where Δ_d is the d–dimensional Laplace operator. We have by [2, Theorem 10.1] that $\|g_\epsilon\|_{L^1} \leq c\|f_\epsilon\|_{L^1}$. Moreover, by [47, Theorem 2.5.7,Remark 2.2.2/3]), we know that for $\alpha > 0$, $\mathscr{C}_b^\alpha(\mathbb{R}^d) = B_{\infty,\infty}^\alpha(\mathbb{R}^d)$, and by [47, Theorem 2.3.8] we know that $(I - \Delta_d)^{-\alpha/2}$ is a continuous operator from $B_{\infty,\infty}^0(\mathbb{R}^d)$ to $B_{\infty,\infty}^\alpha(\mathbb{R}^d)$. Hence, by (A.4) it follows that

$$\left| \int_{\mathbb{R}^d} \Delta_h^m g_\epsilon(x) \psi(x)\,dx \right| = \left| \int_{\mathbb{R}^d} \Delta_h^m f_\epsilon(x) \phi(x)\,dx \right| \leq K|h|^s \|\phi\|_{\mathscr{C}_b^\alpha} \leq cK|h|^s \|\psi\|_{B_{\infty,\infty}^0}$$

Notice that by [47, Theorem 2.11.2 and Remark 2.11.2/2], $B_{\infty,\infty}^0(\mathbb{R}^d)$ is the dual of $B_{1,1}^0(\mathbb{R}^d)$, moreover $B_{1,1}^0(\mathbb{R}^d) \hookrightarrow L^1(\mathbb{R}^d)$ by [47, Proposition 2.5.7], therefore $\|\Delta_h^m g_\epsilon\|_{L^1} \leq \|\Delta_h^m g_\epsilon\|_{B_{1,1}^0} \leq cK|h|^s$, hence $\|g_\epsilon\|_{B_{1,\infty}^s} \leq c(1+K)$. Again, since $(I - \Delta_d)^{\alpha/2}$ maps continuously $B_{1,\infty}^s(\mathbb{R}^d)$ into $B_{1,\infty}^{s-\alpha}(\mathbb{R}^d)$, it finally follows that $\|f_\epsilon\|_{B_{1,\infty}^{s-\alpha}} \leq c\|g_\epsilon\|_{B_{1,\infty}^s}$.

By Sobolev's embeddings and [47, formula 2.2.2/(18)], we have for every $r < s - \alpha$ and $1 \leq p \leq d/(d-r)$ that $B_{1,\infty}^{s-\alpha}(\mathbb{R}^d) \hookrightarrow B_{1,1}^r(\mathbb{R}^d) = W^{r,1}(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$. In particular, $(f_\epsilon)_{\epsilon>0}$ is uniformly integrable in $L^1(\mathbb{R}^d)$, therefore there is f_X such that $\mu = f_X\,dx$ and $(f_\epsilon)_{\epsilon>0}$ converges weakly in $L^1(\mathbb{R}^d)$ to f_μ. Formula (A.3) follows by semi–continuity. $\square$

By the proof, it is clear that the Lemma applies in the case of a positive finite measure (so, not necessarily of mass one).

The properties of Besov spaces we have used hold regardless of the summability parameters of the spaces, thus one can show with arguments entirely similar with those above the following result.

Corollary A.2. *Let X be a $\mathbb{R}^d$–valued random variable. If there are an integer $m \geq 1$, numbers α, s, p, q, with $0 < \alpha < s < m$, $1 < p \leq \infty$ and $1 \leq q \leq \infty$, and a constant $K > 0$ such that for every $\phi \in B_{p,q}^\alpha(\mathbb{R}^d)$ and $h \in \mathbb{R}^d$, with $|h| \leq 1$,*

$$\left| \mathbb{E}[\Delta_h^m \phi(X)] \right| \leq K|h|^s \|\phi\|_{B_{p,q}^\alpha},$$

then X has a density f_X with respect to the Lebesgue measure on $\mathbb{R}^d$. Moreover, for every $r < s - \alpha$, $f_X \in B_{p',\infty}^r(\mathbb{R}^d)$ and

$$\|f\|_{B_{p',\infty}^r} \lesssim (1+K),$$

where p' is the conjugate Hölder exponent of p.

We will also use the following version of the smoothing lemma in the paper.

Lemma A.3. *Let $1 < p < \infty$ and $f \in L^p(\mathbb{R}^d)$. Assume there are an integer $m \geq 1$, two real numbers $s > 0$ and $\alpha > 0$, with $\alpha < s < m$, and a constant $K > 0$ such that for every $\phi \in C_c^\infty(\mathbb{R}^d)$ and $h \in \mathbb{R}^d$, with $|h| \leq 1$,*

$$\left| \int_{\mathbb{R}^d} \Delta_h^m \phi(x) f(x)\,dx \right| \leq K|h|^s \|\phi\|_{F_{q,\infty}^\alpha},$$

where q is the Hölder conjugate exponent of p. Then $f \in B_{p,\infty}^{s-\alpha}$ and

$$\|f\|_{B_{p,\infty}^{s-\alpha}} \lesssim (\|f\|_{L^p} + K). \tag{A.5}$$

Proof. Fix a smooth function ϕ and set $g = (I - \Delta_d)^{-\alpha/2}f$, $\psi = (I - \Delta_d)^{\alpha/2}\phi$, then by integration by parts,

$$\left|\int_{\mathbb{R}^d} (\Delta_h^m g)(x)\psi(x)\,dx\right| = \left|\int_{\mathbb{R}^d} g(x)(\Delta_{-h}^m \psi)(x)\,dx\right| =$$
$$= \left|\int_{\mathbb{R}^d} f(x)(\Delta_{-h}^m \phi)(x)\,dx\right| \le K|h|^s\|\phi\|_{F_{q,\infty}^\alpha} \le K|h|^s\|\psi\|_{F_{q,\infty}^0},$$

since by [47, Theorem 2.3.8], $(I - \Delta_d)^{-\alpha/2}$ maps $F_{q,\infty}^\alpha$ into $F_{q,\infty}^0$.

We know by [47, Theorem 2.11.2] that the space $F_{q,\infty}^0$ is the dual of $F_{q,1}^0$, therefore we deduce from the inequality above that $\|\Delta_h^m g\|_{F_{p,1}^0} \le K|h|^s$. On the other hand we know by [47, Proposition 2.5.7] that $F_{p,1}^0 \subset L^p$, hence $\|\Delta_h^m g\|_{L^p} \lesssim K|h|^s$.

Finally, by [2], $\|g\|_{L^p} \lesssim \|f\|_{L^p}$, and in conclusion $g \in B_{p,\infty}^s$ and $\|g\|_{B_{p,\infty}^s} \lesssim (K + \|f\|_{L^p})$, and then (A.5) follows, since $\|f\|_{B_{p,\infty}^{s-\alpha}} = \|(I - \Delta_d)^{\alpha/2}g\|_{B_{p,\infty}^{s-\alpha}} = \|g\|_{B_{p,\infty}^s}$. $\square$

Remark A.4. Since $F_{\infty,\infty}^\alpha = B_{\infty,\infty}^\alpha$ for all $\alpha \in \mathbb{R}$, the case $p = \infty$ in the lemma above is already covered by Lemma A.1.

References

[1] Randolf Altmayer, *Estimating occupation time functionals*, arXiv:1706.03418, 2017.

[2] Nachman Aronszajn and Kennan T. Smith, *Theory of Bessel potentials I*, Ann. Inst. Fourier (Grenoble) **11** (1961), 385–475. MR-0143935

[3] David Baños and Paul Krühner, *Optimal density bounds for marginals of Itô processes*, Commun. Stoch. Anal. **10** (2016), no. 2, 131–150. MR-3605421

[4] David Baños and Paul Krühner, *Hölder continuous densities of solutions of SDEs with measurable and path dependent drift coefficients*, Stochastic Process. Appl. **127** (2017), no. 6, 1785–1799. MR-3646431

[5] Vlad Bally and Lucia Caramellino, *Regularity of probability laws by using an interpolation method*, 2012, arXiv:1211.0052.

[6] Vlad Bally and Lucia Caramellino, *Integration by parts formulas, Malliavin calculus, and regularity of probability laws*, Stochastic integration by parts and functional Itô calculus, Adv. Courses Math. CRM Barcelona, Birkhäuser/Springer, [Cham], 2016, pp. 1–114. MR-3497714

[7] Vlad Bally and Lucia Caramellino, *Convergence and regularity of probability laws by using an interpolation method*, Ann. Probab. **45** (2017), no. 2, 1110–1159. MR-3630294

[8] Jean-Michel Bony, *Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires*, Ann. Sci. École Norm. Sup. (4) **14** (1981), no. 2, 209–246. MR-631751

[9] Giuseppe Cannizzaro, Peter K. Friz, and Paul Gassiat, *Malliavin calculus for regularity structures: the case of gPAM*, J. Funct. Anal. **272** (2017), no. 1, 363–419. MR-3567508

[10] Rémi Catellier and Khalil Chouk, *Paracontrolled distributions and the 3-dimensional stochastic quantization equation*, arXiv:1310.6869, to appear on Annals of Probability, 2018.

[11] Zhen-Qing Chen and Longmin Wang, *Uniqueness of stable processes with drift*, Proc. Amer. Math. Soc. **144** (2016), no. 6, 2661–2675. MR-3477084

[12] Giuseppe Da Prato and Arnaud Debussche, *Ergodicity for the 3D stochastic Navier-Stokes equations*, J. Math. Pures Appl. (9) **82** (2003), no. 8, 877–947. MR-2005200

[13] Stefano De Marco, *Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions*, Ann. Appl. Probab. **21** (2011), no. 4, 1282–1321. MR-2857449

[14] Arnaud Debussche and Nicolas Fournier, *Existence of densities for stable-like driven SDE's with Hölder continuous coefficients*, J. Funct. Anal. **264** (2013), no. 8, 1757–1778. MR-3022725

[15] Arnaud Debussche and Marco Romito, *Existence of densities for the 3D Navier–Stokes equations driven by Gaussian noise*, Probab. Theory Related Fields **158** (2014), no. 3-4, 575–596. MR-3176359

[16] Romain Duboscq and Anthony Réveillac, *Stochastic regularization effects of semi-martingales on random functions*, J. Math. Pures Appl. (9) **106** (2016), no. 6, 1141–1173. MR-3565418

[17] Ennio Fedrizzi and Franco Flandoli, *Hölder flow and differentiability for SDEs with nonregular drift*, Stoch. Anal. Appl. **31** (2013), no. 4, 708–736. MR-3175794

[18] Franco Flandoli, Massimiliano Gubinelli, and Enrico Priola, *Well-posedness of the transport equation by stochastic perturbation*, Invent. Math. **180** (2010), no. 1, 1–53. MR-2593276

[19] Nicolas Fournier, *Finiteness of entropy for the homogeneous Boltzmann equation with measure initial condition*, Ann. Appl. Probab. **25** (2015), no. 2, 860–897. MR-3313757

[20] Nicolas Fournier and Jacques Printems, *Absolute continuity for some one–dimensional processes*, Bernoulli **16** (2010), no. 2, 343–360. MR-2668905

[21] Martin Hairer, *A theory of regularity structures*, Invent. Math. **198** (2014), no. 2, 269–504. MR-3274562

[22] Martin Hairer, *Introduction to regularity structures*, Braz. J. Probab. Stat. **29** (2015), no. 2, 175–210. MR-3336866

[23] Martin Hairer, *Regularity structures and the dynamical Φ_3^4 model*, Current developments in mathematics 2014, Int. Press, Somerville, MA, 2016, pp. 1–49. MR-3468250

[24] Masafumi Hayashi, Arturo Kohatsu-Higa, and Gô Yûki, *Local Hölder continuity property of the densities of solutions of SDEs with singular coefficients*, J. Theoret. Probab. **26** (2013), no. 4, 1117–1134. MR-3119987

[25] Masafumi Hayashi, Arturo Kohatsu-Higa, and Gô Yûki, *Hölder continuity property of the densities of SDEs with singular drift coefficients*, Electron. J. Probab. **19** (2014), no. 77, 22. MR-3256877

[26] Lorick Huang, *Density estimates for SDEs driven by tempered stable processes*, arXiv:1504.04183, 2015.

[27] Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992. MR-1214374

[28] Victoria Knopova and Alexei Kulik, *Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 1, 100–140. MR-3765882

[29] Arturo Kohatsu-Higa and Libo Li, *Regularity of the density of a stable-like driven SDE with Hölder continuous coefficients*, Stoch. Anal. Appl. **34** (2016), no. 6, 979–1024. MR-3544166

[30] Nicolai V. Krylov and Michael Röckner, *Strong solutions of stochastic equations with singular time dependent drift*, Probab. Theory Related Fields **131** (2005), no. 2, 154–196. MR-2117951

[31] Seiichiro Kusuoka, *Existence of densities of solutions of stochastic differential equations by Malliavin calculus*, J. Funct. Anal. **258** (2010), no. 3, 758–784. MR-2558176

[32] Claude Le Bris and Pierre-Louis Lions, *Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients*, Comm. Partial Differential Equations **33** (2008), no. 7-9, 1272–1317. MR-2450159

[33] Paul Malliavin, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976) (New York-Chichester-Brisbane), Wiley, 1978, pp. 195–263. MR-536013

[34] Thilo Meyer-Brandis and Frank Proske, *Construction of strong solutions of SDE's via Malliavin calculus*, J. Funct. Anal. **258** (2010), no. 11, 3922–3953. MR-2606880

[35] Jean-Christophe Mourrat, Hendrik Weber, *Global well-posedness of the dynamic Φ^4 model in the plane*, Ann. Probab. **45** (2017), no. 4, 2398–2476. MR-3693966

[36] Jean-Christophe Mourrat, Hendrik Weber, and Weijun Xu, *Construction of Φ_3^4 diagrams for pedestrians*, From particle systems to partial differential equations, Springer Proc. Math. Stat., vol. 209, Springer, Cham, 2017, pp. 1–46. MR-3746744

[37] Jean-Christophe Mourrat and Hendrik Weber, *The dynamic Φ_3^4 model comes down from infinity*, Comm. Math. Phys. **356** (2017), no. 3, 673–753. MR-3719541

[38] David Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, Berlin, 2006. MR-2200233

[39] Marco Romito, *Unconditional existence of densities for the Navier-Stokes equations with noise*, Mathematical analysis of viscous incompressible fluid, RIMS Kôkyûroku, vol. 1905, Kyoto University, 2014, pp. 5–17.

[40] Marco Romito, *Uniqueness and blow-up for a stochastic viscous dyadic model*, Probab. Theory Related Fields **158** (2014), no. 3-4, 895–924. MR-3176368

[41] Marco Romito, *Hölder continuity of the densities for the Navier–Stokes equations with noise*, Stoch. Partial Differ. Equ. Anal. Comput. **4** (2016), no. 3, 691–711. MR-3538013

[42] Marco Romito, *Some probabilistic topics in the Navier–Stokes equations*, Recent Progress in the Theory of the Euler and Navier–Stokes Equations (James C. Robinson, José L. Rodrigo, Witold Sadowski, and Alejandro Vidal-López, eds.), London Math. Soc. Lecture Note Ser., vol. 430, Cambridge Univ. Press, Cambridge, 2016, pp. 175–232. MR-3497693

[43] Marco Romito, *Time regularity of the densities for the Navier–Stokes equations with noise*, J. Evol. Equations **16** (2016), no. 3, 503–518. MR-3551234

[44] Marta Sanz-Solé and André Süß, *Absolute continuity for SPDEs with irregular fundamental solution*, Electron. Commun. Probab. **20** (2015), no. 14, 11. MR-3314649

[45] Marta Sanz-Solé and André Süß, *Non-elliptic SPDEs and ambit fields: existence of densities*, Stochastics of environmental and financial economics—Centre of Advanced Study, Oslo, Norway, 2014–2015, Springer Proc. Math. Stat., vol. 138, Springer, Cham, 2016, pp. 121–144. MR-3451171

[46] René L. Schilling, Paweł Sztonyk, and Jian Wang, *Coupling property and gradient estimates of Lévy processes via the symbol*, Bernoulli **18** (2012), no. 4, 1128–1149. MR-2995789

[47] Hans Triebel, *Theory of function spaces*, Monographs in Mathematics, vol. 78, Birkhäuser Verlag, Basel, 1983. MR-781540

[48] Hans Triebel, *Theory of function spaces. II*, Monographs in Mathematics, vol. 84, Birkhäuser Verlag, Basel, 1992. MR-1163193

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm