Internship report

Stable driven McKean-Vlasov SDEs with singular kernels in the non-degenerate and kinetic case: Theory and approximation

Anna Bahrii

supervised by
Professor Stéphane Menozzi

September 9, 2024

Déclaration d'intégrité relative au plagiat

I, the undersigned, Anna Bahrii, certify that

– I am the author of the report;

– The results described in the report are the culmination of my work;

– I have not used third-party sources or results without clearly citing and referencing them according to the recommended bibliographic rules.

Done in Lviv, 09/09/2024

Signature:

Abstract

We study well-posedness of McKean-Vlasov stochastic differential equations driven by an α-stable process with a singular drift by exploiting regularization by noise techniques. We first prove weak and strong well-posedness of the *linear* SDE (i.e. with measure independent drift) where drift belongs to Hölder space. Therein, we also prove an optimal weak error between the Euler scheme to the solution of the SDE. We then discuss well-posedness results for *linear* and *non-linear* SDEs with distributional drift, in particular when it belongs to Lebesgue-Besov space.

Contents

1 Introduction

For a fixed $T > 0$, consider a *McKean-Vlasov stochastic differential equation* (SDE) given by

$$dX_t = \mathfrak{b}(t, X_t, \mu_{X_t})dt + dZ_t, \quad X_0 \sim \mu_0, \quad t \in [0, T], \tag{1}$$

where the driving noise $Z = (Z_t)_{t\geq 0}$ is a d-dimensional α-stable process with $\alpha \in (0, 2]$ (see Section 3 for definition), $\mu_{X_t} = \mathrm{Law}(X_t)$, μ_0 is a probability measure and the drift $\mathfrak{b}$ belongs to a suitable function space. We say that equation (1) is *non-degenerate* if the dimension of the noise Z is the same as the dimension of the process X, and *kinetic* when $X_t = (X_t^1, X_t^2)$ is given by

$$\begin{aligned} dX_t^1 &= \mathfrak{b}_1(t, X_t, \mu_{X_t})dt + dZ_t, \\ dX_t^2 &= \big(X_t^1 + \mathfrak{b}_2(t, X_t, \mu_{X_t^2})\big)dt, \quad X_0 \sim \mu_0, \quad t \in [0, T], \end{aligned}$$

where $\mu_{X_t^2} = \mathrm{Law}(X_t^2)$, and $\mathfrak{b}_1, \mathfrak{b}_2$ belong to suitable function spaces. Herein, we focus on the non-degenerate type.

It is well-known that when the driving noise in (1) is a Brownian motion ($\alpha = 2$) and $\mathfrak{b}$ is *Lipschitz continuous* in $\mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}_m(\mathbb{R}^d)$, where $m \in [1, \infty)$ and $\mathcal{P}_m(\mathbb{R}^d)$ is the set of probability measures with finite m-th moment equipped with the topology induced by the m-Wasserstein metric, then the McKean-Vlasov equation is well-defined in a *strong* sense. This observation goes back to McKean and can be proven using usual fixed point argument (see e.g. [Szn91], [CD22a]).

We are interested in well-posedness of the SDE (1) when the drift is not Lipschitz and is such that the ODE

$$dX_t = b(t, X_t)dt, \quad X_0 = x \in \mathbb{R}^d, \quad t \in [0, T],$$

is possibly ill-posed: the solution may exist but be not unique or not exist at all. We refer to such drifts as *singular*. The idea behind restoring the well-posedness in (1) with singular drifts is *regularization by noise* (see e.g. [Fla11]).

Precisely, random noise allows to compensate the singularity of the drift to obtain existence and uniqueness of solution. In this report, we prove well-posedness results when the drift is Hölder continuous (see Section 1), and we discuss results for $L_t^q - L_x^p$ and distributional drifts, in particular we focus on drifts that belong to *Besov spaces* $B_{p,q}^{\beta}$ ($\beta \in \mathbb{R}$, $p, q \geq 1$) in space, see Section 4 for definition and Section 5 for related results. Regularization effect can come from the driving noise (when the drift does not depend on measure) and from the driving noise together with the law dependence in the drift as in (1).

Let us first give an overview of the results for the measure independent drift in (1) ($\mathfrak{b}(t, X_t, \mu_{X_t}) = b(t, X_t)$) and for Brownian motion as the driving noise, i.e. we deal with *standard*, or *linear*, SDEs. An important result in this direction was proven by Zvonkin in [Zvo07]. He showed that 1-dimensional SDE is strongly well-posed as soon as the drift is bounded and Borel measurable. The fundamental idea behind the proof is to use what is called *Zvonkin transform* which consists in considering an auxiliary parabolic PDE with the drift as a source term. This trick allows us to express the drift using the regular solution of the PDE. This approach can be used for investigating even more singular drifts, and we explain it in more detail for Hölder drifts, see Section 1. See also Section 5 for more intricate setup. Zvonkin's result was extended to an arbitrary dimension by Veretennikov in [Ver81]. Krylov and Röcker proved in [KR05] that when $\mathfrak{b}(t, X_t, \mu_{X_t}) = b(t, X_t)$ belongs to $L^q_{\text{loc}}(\mathbb{R}_+, L^p(\mathbb{R}^d))$, where p, q satisfy *Krylov-Röckner condition*, or *Serrin type condition*,

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

then the SDE is strongly well-posed. In the strictly α-stable case ($\alpha < 2$), this condition extends to

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

which guarantees only weak well-posedness (see [XZ20]). See Section 5 for the context of usage of this classical result in a more intricate framework.

For the drift in a *standard* SDE being a distribution we mention several works. First, Flandoli, Issoglio and Russo in [FIR15] prove existence and uniqueness of the *virtual solution* (see Definition 5.1) using Zvonkin transform, when the driving noise is Brownian motion and drift belongs to a fractional Sobolev space. In [ABM20] authors prove strong well-posedness of a stable-driven SDE with symmetric noise when $d = 1$ and the drift is a distributional derivative of a β-Hölder function with $\beta > \frac{1-\alpha}{2}$, i.e. the drift belongs to a specific Besov space. The proof again relies on Zvonkin transform. Authors in [CM22] show well-posedness of the *martingale problem* and show existence and uniqueness of the weak solution when the underlying noise is *uniformly non-degenerate* and symmetric and $b \in L^r((0,T], B_{p,q}^{-1+\gamma}(\mathbb{R}^d, \mathbb{R}^d))$, $r, p, q \geq 1$, $\gamma \in (1/2, 1)$, where parameters satisfy

$$\alpha \in \left(\frac{1 + [d/p]}{1 - [1/r]}, 2 \right], \quad \gamma \in \left(\frac{3 - \alpha + [2d/p] + [2\alpha/r]}{2}, 1 \right).$$

See Section 5.1 for more details. It is possible to go below in the regularity index of the Besov space assuming additional structure on the drift. In [KP22] authors achieve the regularity $\beta > \frac{2-2\alpha}{3}$ assuming that the drift satisfies properties that allow to use *paracontrolled calculus*. In [DD14] authors also achieve the regularity $\beta > \frac{2-2\alpha}{3}$ for $\alpha = 2$ assuming that the drift has a rough paths structure.

For McKean-Vlasov type equations, one can refer to [Cha20] where the author extended Zvonkin transform for a drift that has general measure dependency and $\alpha = 2$, and proved strong well-posedness when $\mathfrak{b}$ belongs to Hölder space. Now, let us note that we are interested in a particular measure dependency in the drift. More precisely, we consider

$$\mathfrak{b}(t, X_t, \mu_{X_t}) = b * \mu_{X_t}(t, X_t) = \int_{\mathbb{R}^d} b(t, X_t - y)\mu_{X_t}(dy), \tag{2}$$

where b is called *interaction kernel* and $*$ denotes a spatial convolution. This special convolution dependency on the measure allows us to restore well-posedness of (1) for interaction kernel that is more singular than is allowed in the *linear* case. For example, as it will be explained in Section 5.3, having the drift of the form (2), one is able to show (take $\alpha = 2$ for simplicity) that strong well-posedness holds with the interaction kernel belonging to the Besov space $L^r((0,T], B_{\infty,\infty}^{-1+}(\mathbb{R}^d, \mathbb{R}^d))$, whereas in the *linear* case weak well-posedness holds for $b \in L^r((0,T], B_{\infty,\infty}^{-\frac{1}{2}+}(\mathbb{R}^d, \mathbb{R}^d))$. Thus, the additional regularization effect from the law allows us to gain $\frac{1}{2}$ of the regularity for the drift.

We start with analysis of the SDE with a time-homogeneous Hölder continuous drift that does not depend on the law of X_t, with deterministic initial condition and with a Brownian motion as driving noise. More precisely, we consider the drift $\mathfrak{b}(t, X_t, \mu_{X_t}) =: b(X_t) \in C^\beta(\mathbb{R}^d, \mathbb{R}^d)$, $\beta \in (0,1)$, and we are interested in the well-posedness of the SDE given by

$$dX_t = b(X_t)dt + dW_t, \quad X_0 = x \in \mathbb{R}^d, \tag{3}$$

where W is a d-dimensional Brownian motion. Equation (3) with a non-identity diffusion coefficient was thoroughly studied in [MP91], where authors also studied the convergence speed of the scheme. Section 1 of this report focuses on proving the results presented in this paper. See also [KM17] for an error for the densities of the scheme and the original equation.

A big part of the report is also focused on proving the optimal rate of convergence of the Euler scheme to the SDE (3). Although this is not directly related to the well-posedness questions, it is a great demonstration of the toolkit that is constantly used in the theory and it is an overview of possible problems and difficulties one may face in this field. We distinguish *strong* and *weak* error. Weak error involves the laws of the Euler scheme and the solution to (3). So first, one has to show that the density exists. Using Itô's formula, one can then derive that density satisfies a corresponding PDE (backward Kolmogorov equation, Fokker-Planck equation). Then, one can derive *Duhamel representation* for the density which allows us to investigate the weak error. Heat kernel estimates are of the essential use. Strong error involve trajectorial errors and is not covered here.

Finally, let us mention that McKean-Vlasov equations where first studied as a mean-field limit of interacting particle systems. Precisely, having $N \geq 1$ particles, one studies behaviour of a *typical* particle the law of which is given by empirical measure involving a full system of interacting particles and the behaviour of which is described by an SDE where the interaction kernel depends on the empirical measure. See [Szn91], [CD22a] for solid explanations. If the interaction kernel is regular enough, one can prove *propagation of chaos* for a given particle system provided that it is well-posed. Informally speaking, propagation of chaos property says that if at time 0 any $k > 0$ interacting particles converge in law to a product measure, then at any time > 0 the convergence persists. As before, in the case of Lipschitz kernel the system of SDEs is well-posed and propagation of chaos holds. For more singular kernels, we refer to [Tom23], [HRZ22], [Hao+24] and references therein. McKean-Vlasov equations give a probabilistic interpretation to some non-linear parabolic PDEs arising in physics, for example, Boltzmann type equations and Burgers equation. In particular, our interest to distributional drifts and Besov spaces comes from the fact that this type of drifts cover many related models. Nowadays, McKean-Vlasov equations are used in biology and social sciences to describe systems of indistinguishable particles, mean-field games and data science (see [CD22b] for concrete models).

Notion of solution Classical definitions of weak and strong solution is formulated as follows (see e.g. [SV97], [Bas97], [IW81] for more details).

Definition 1.1 (Martingale solution). *Let $\mathcal{L}^X$ be a generator associated to the SDE* (1)*. Let also $\Omega_2 = C([0,T], \mathbb{R}^d)$ (space of continuous functions) and $\Omega_\alpha = D([0,T], \mathbb{R}^d)$ (space of càdlàg functions) when $\alpha \in (0,2)$. We say that a probability measure $\mathbb{P}^\alpha$ on Ω_α is a* martingale solution *to the* martingale problem *associated with an operator $\mathcal{L}^X$ on the probability space $(\Omega_\alpha, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is a Borel σ-algebra, if by denoting $(X_t)_{t\geq 0}$ the canonical process on Ω_α, we have*

(i) $\mathbb{P}^\alpha(X_0 = 0) = 1$*;*

(ii) For all $f \in C_0^2(\mathbb{R}^d)$,

$$\Big(f(t, X_t) - \int_0^t (\partial_s + \mathcal{L}^X) f(s, X_s) ds\Big)_{0\leq t\leq T}$$

is a (square integrable if $\alpha = 2$) $\mathbb{P}^\alpha$-martingale.

Definition 1.2 (Weak solution). *A process $(X_t)_{t\geq 0}$ is called a weak solution to the SDE* (1) *if*

(i) It is continuous for $\alpha = 2$, and it is càdlàg for $\alpha \in (0,2)$;

(ii) There exists an α-stable noise $Z = (Z_t)_{t\geq 0}$ such that $Z_0 = 0$ a.s.;

(iii) $\forall t \in [0,T]$, (X_t, Z_t) satisfies (1) *a.s.*

Definition 1.3 (Strong solution). *Let $(Z_t)_{t\geq 0}$ be a given α-stable noise. A process $(X_t)_{t\geq 0}$ is called a strong solution to the SDE* (1) *if it is adapted to the filtration generated by $(Z_t)_{t\geq 0}$ and it satisfies* (1) *a.s.*

We note that often when working with singular drifts, in particular distributional, one has to understand *what* the solution to the equation is even in the "weak sense". This comes from the fact that having distribution in the equation, (1) is not well-defined as it is. Therefore, in Section 5 we sometimes introduce new definitions of solutions.

A weak solution is unique when for any solutions (X, Z), $(\tilde{X}, \tilde{Z})$ s.t. $\text{Law}(X_0) = \text{Law}(\tilde{X}_0)$, we have $\text{Law}(X, Z) = \text{Law}(\tilde{X}, \tilde{Z})$. A solution to (1) is *pathwise unique* if for any two weak solutions X, $\tilde{X}$ to (1) defined on the same probability space we have

$$\mathbb{P}(X_t = \tilde{X}_t,\ \forall t \geq 0) = 1.$$

Organization This report is organized as follows. In Section 1 we study the SDE and its Euler scheme with a time-homogeneous Hölder continuous drift and with Brownian motion as driving noise. This section is the main contribution of the report, therefore the statements are proved rigorously with a special attention to details and described methods. We also investigate the optimal convergence rate of a weak error between the associated Euler scheme and the weak solution of the Hölder SDE. The results of this section are to our best knowledge new. In Section 3 we introduce α-stable process. We discuss their main properties that are of an important use for studying SDEs with distributional drifts. In Section 4 we give background for Besov spaces from the viewpoint that is convenient in the framework of our problem. We do not demonstrate the proofs, although we refer an interested reader to the trustful sources. We start Section 5 with a short panorama on the theory that involves general distributional drifts. Then, in Section 5 we study well-posedness of the SDE (1) with Lebesgue-Besov drift distinguishing two cases: measure-dependent and -independent drift. We state main results of the related papers and try to give a general view on the proofs strategy referencing to the precise original statements when needed.

2 Brownian motion driven SDE with time-homogeneous Hölder continuous drift

In this chapter we are interested in two questions. First fundamental question consists in studying weak and strong well-posedness of the SDE (3). Second question would be to determine the rate of convergence of a corresponding Euler scheme of (3) to the original equation. As we will see, in the given setup existence of a weak solution can be easily proven using *Girsanov theorem*. For studying strong well-posedness in this specific case we introduce the parabolic PDE given by

$$\begin{cases} (\partial_t + b(x)\cdot\nabla + \frac{1}{2}\Delta)v(t,x) = b(x), \quad (t,x) \in [0,T) \times \mathbb{R}^d, \\ v(T,\cdot) = 0. \end{cases} \tag{4}$$

Remarking that it is important to have the "rough" drift b as a source term in (4), we use the so called *Zvonkin transform* (see [Zvo07] and [Ver81]), which allows us to express the integral of a not so smooth drift via a smooth solution to (4). Indeed, applying Itô's formula on $v(t, X_t)$, where v is a $C_b^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$ solution of (4), we get

$$\begin{aligned} v(t,X_t) &= v(0,X_0) + \int_0^t \partial_s v(s,X_s)ds + \int_0^t \nabla v(s,X_s)\cdot dX_s + \frac{1}{2}\int_0^t \Delta v(s,X_s)ds \\ &= v(0,x) + \int_0^t \partial_s v(s,X_s)ds + \int_0^t \nabla v(s,X_s)\cdot b(X_s)ds + \int_0^t \nabla v(s,X_s)\cdot dW_s + \frac{1}{2}\int_0^t \Delta v(s,X_s)ds \\ &= v(0,x) + \int_0^t b(X_s)ds + \int_0^t \nabla v(s,X_s)\cdot dW_s. \end{aligned}$$

And so

$$\int_0^t b(X_s)ds = v(t,X_t) - v(0,x) - \int_0^t \nabla v(s,X_s)\cdot dW_s.$$

Of course, important step in this analysis is to guarantee that the unique solution to (4) exists and holds the required properties. Here, the idea would be to rewrite (4) as

$$\begin{cases} (\partial_t + \frac{1}{2}\Delta)v(t,x) = b(x) - b(x)\cdot\nabla v(t,x), \quad (t,x) \in [0,T) \times \mathbb{R}^d, \\ v(T,\cdot) = 0, \end{cases} \tag{5}$$

considering $b(x)-b(x)\cdot\nabla v(t,x)$ as the source term, and to look for a fixed point in the space $C_b^{1+\frac{\beta-\gamma}{2},2+\beta-\gamma}([0,T]\times\mathbb{R}^d)$, $\gamma\in(0,\beta)$, for a map given by

$$\Phi(v)(t,x) := \int_t^T \int_{\mathbb{R}^d} g(s-t,x-y)(b(y)-b(y)\cdot\nabla v(s,y))dyds,$$

where $g(u,z)=(2\pi u)^{-\frac{d}{2}}\exp\big(-\frac{|z|^2}{2u}\big)$ is the Gaussian heat kernel (see Section 2.3.2 for more details).

Determining the rate of convergence of the Euler scheme of (3) is a big question of interest itself (see for example [FJM24] for stable-driven SDEs where drift is in a suitable Lebesgue space). However, in the context of studying well-posedness of (1), it serves as an introduction to the machinery that is used for investigating SDEs with Besov drifts. Here, we are primarily interested in the *weak error*, which involves estimating error between densities and thus heat kernel estimates. More precisely, we consider the PDE given by

$$\begin{cases}(\partial_t + b(x)\cdot\nabla + \frac{1}{2}\Delta)v(t,x) = 0, \quad (t,x)\in[0,T)\times\mathbb{R}^d,\\ v(T,x)=f(x), \quad x\in\mathbb{R}^d,\end{cases} \tag{6}$$

where $f\in C^2$ so that the unique solution v exists and is in $C_b^{1,2}([0,T]\times\mathbb{R}^d,\mathbb{R})$ (see Theorem IV.5.2 in [LS68] for proof). Next, define the Euler scheme associated with (3) in the following way: for $N\in\mathbb{N}$ let $h:=T/N$ be the time step, $t_i:=ih$, $i=0,...,N$, $\phi(s)=t_i$ for $s\in[t_i,t_{i+1})$. Then the *Euler scheme* for (3) is given by

$$X_t^h = x + \int_0^t b(X^h_{\phi(s)})ds + W_t, \quad x\in\mathbb{R}^d. \tag{7}$$

Finally, the *weak error* is defined as

$$\mathcal{E}(h,T,f,x) := \mathbb{E}_x[f(X_T^h)] - \mathbb{E}_x[f(X_T)].$$

Having justified that solution X_t^h to (7) admits density p^h, the key trick is to show that the density satisfies *Duhamel representation*

$$p^h(t,x,y) = g(t,y-x) - \int_0^t \mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_y g(t-s,y-X_s^h)]ds.$$

which can be done by relating a functional of a Brownian motion to a heat equation. Duhamel representation has an advantage of involving a Gaussian kernel. Indeed, it is a well-behaved function whose derivatives have a bound that involves a different Gaussian kernel. As we will see later, using the fact that $\int_0^t\int_{\mathbb{R}^d}\nabla_y g(t-s,y-z)dsdz=0$ allows us to add any slightly beneficial terms multiplied by this quantity for free and to profit of the regularity of these terms. This technique is called *cancellation argument*. This type of methods may also be adapted to α-stable processes for $\alpha\in(1,2]$ (see [FJM24]).

2.1 Well-posedness of the Hölder SDE

We first establish existence of a weak solution.

Proposition 2.1. *Equation* (3) *admits a weak solution for any initial condition $x\in\mathbb{R}^d$.*

Proof. A natural tool for proving this proposition is Girsanov theorem. Let X be a d-dimensional Brownian motion on a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_s)_{s\geq 0},\mathbb{Q})$ starting from $x\in\mathbb{R}^d$. Denote $Y_t:=\int_0^t b(X_s)\cdot dX_s$. We have that quadratic variation of Y_t is

$$\langle Y,Y\rangle_t = \int_0^t |b(X_s)|^2 ds,$$

Define the probability measure $\mathbb{P}$ in the following way:

$$\frac{d\mathbb{P}}{d\mathbb{Q}}\bigg|_{\mathcal{F}_t} = \exp\Big(Y_t - \frac{1}{2}\langle Y,Y\rangle_t\Big). \tag{8}$$

In order to apply Girsanov theorem, we have to ensure that (8) is a martingale. But $\mathbb{E}[e^{\frac{1}{2}\langle Y,Y\rangle_t}]<\infty$, $\forall t\geq 0$, since b is bounded, so by Novikov's criterion (8) is a $\mathbb{Q}$-martingale. Thus,

$$\tilde{W}_t = X_t - x - \langle X,Y\rangle_t \tag{9}$$

is a $\mathbb{P}$-Brownian motion. But

$$\begin{aligned}\langle X,Y\rangle_t &= \langle X,\int_0^{\cdot} b(X_s)dX_s\rangle_t\\ &= \langle \int_0^{\cdot} dX_s + x, \int_0^{\cdot} b(X_s)dX_s\rangle_t\\ &= \int_0^t b(X_s)ds.\end{aligned}$$

And so rearranging terms in (9) and using previous computations we get

$$\begin{aligned}X_t &= x + \langle X,Y\rangle_t + \tilde{W}_t\\ &= x + \int_0^t b(X_s)ds + \tilde{W}_t.\end{aligned}$$

Thus, $(X_t, \tilde{W}_t)$ is a weak solution to the SDE (3). □

Before studying strong well-posedness of (3), let us state a useful lemma.

Lemma 2.1. *For any weak solution of* (3) *there exists* $C_T > 0$ *such that*

$$\mathbb{E}[\sup_{s\in[0,T]} |X_s|^2] \leq C_T(1+|x|^2).$$

Proof. Assume that the underlying probability space is complete. Since $(X_s)_{s\in[0,T]}$ is a.s. continuous, $\sup_{s\in[0,T]}|X_s|^2$ is measurable. Then by direct computations and Doob's inequality we have,

$$\begin{aligned}\mathbb{E}[\sup_{s\in[0,T]} |X_s|^2] &= \mathbb{E}[\sup_{s\in[0,T]} |x + \int_0^s b(X_r)dr + W_s|^2]\\ &\leq C(|x|^2 + \mathbb{E}[\sup_{s\in[0,T]} (\int_0^s b(X_r)dr)^2] + \mathbb{E}[\sup_{s\in[0,T]} |W_s|^2])\\ &\leq C(|x|^2 + \|b\|_\infty^2 \sup_{s\in[0,T]} s^2 + 4\mathbb{E}[|W_T|^2])\\ &= C(|x|^2 + \|b\|_\infty^2 T^2 + 4T)\\ &\leq C_T(1+|x|^2),\end{aligned}$$

where $C_T := \|b\|_\infty^2 T^2 + 4T$. □

Proposition 2.2. *For* (3)*, pathwise uniqueness of solution holds.*

Proof. Let X, Y be two weak solutions of (3) for the Brownian motion W defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We first show that there exists a constant $C > 0$ such that

$$\mathbb{E}[\sup_{s\in[0,T]} |X_s - Y_s|^2] \leq C\Big((c_0 T^\theta)^2 \mathbb{E}[\sup_{s\in[0,T]} |X_s - Y_s|^2] + \Lambda^2 \int_0^T \mathbb{E}[\sup_{r\in[0,s]} |X_r - Y_r|^2] dr\Big),$$

denoting $\Lambda := \|D^2 v\|_\infty$ and for $\theta > 0$, $c_0 > 0$. Let $v \in C_b^{1,2}([0,T]\times\mathbb{R}^d, \mathbb{R})$ be the solution of the PDE (4).

Then,

$$
\begin{aligned}
\mathbb{E}[\sup_{s\in[0,T]}|X_s-Y_s|^2] &= \mathbb{E}[\sup_{s\in[0,T]}|\int_0^s(b(X_r)-b(Y_r))|^2dr] \\
&= \mathbb{E}[\sup_{s\in[0,T]}|v(s,X_s)-v(s,Y_s)-\int_0^s(\nabla v(r,X_r)-\nabla v(r,Y_r))\cdot dW_r)|^2] \\
&\le C\Big(\mathbb{E}[\sup_{s\in[0,T]}|v(s,X_s)-v(s,Y_s)|^2]+\mathbb{E}[\sup_{s\in[0,T]}|\int_0^s|\nabla v(r,X_r)-\nabla v(r,Y_r)|^2dr]\Big) \\
&\le C\Big(\|v\|_\infty^2\mathbb{E}[\sup_{s\in[0,T]}|X_s-Y_s|^2]+\mathbb{E}[\int_0^T|\nabla v(r,X_r)-\nabla v(r,Y_r)|^2dr]\Big) \\
&\le C\Big((c_0T^\theta)^2\mathbb{E}[\sup_{s\in[0,T]}|X_s-Y_s|^2]+\|D^2v\|_\infty^2\mathbb{E}[\int_0^T|X_r-Y_r|^2dr]\Big) \\
&= C\Big((c_0T^\theta)^2\mathbb{E}[\sup_{s\in[0,T]}|X_s-Y_s|^2]+\Lambda^2\int_0^T\mathbb{E}[\sup_{r\in[0,T]}|X_r-Y_r|^2dr]\Big),
\end{aligned}
$$

where for the second last inequality we use the fact $\|\nabla v\|_\infty \le c_0T^\theta$ (see later). Thanks to Grönwall's lemma, we have

$$
\mathbb{E}[\sup_{s\in[0,T]}|X_s-Y_s|^2]\le C(c_0T^\theta)^2\mathbb{E}[\sup_{s\in[0,T]}|X_s-Y_s|^2]e^{C\Lambda^2T},
$$

which implies that $\mathbb{E}[\sup_{s\in[0,T]}|X_s-Y_s|^2]=0$ for T such that $1-C(c_0T^\theta)^2e^{C\Lambda^2T}\ge 0$. Or, in other words, we have that a.s. for all $s\in[0,T]$, $X_s=Y_s$. ☐

Thanks to Proposition 2.1 and Proposition 2.2 and Yamada-Watanabe theorem (see [IW81]) we conclude that there exists a unique strong solution to (3) on the interval $[0,T]$.

2.2 Weak error for the Euler scheme

The final goal of this section is to prove the following theorem which is actually a new result in the literature.

Theorem 2.1. *For any $t\in(0,T]$ and $x,y\in\mathbb{R}^d$, the solution to (3) admits a density $p(t,x,y)$ and the solution to (7) admits a density $p^h(t,x,y)$. Moreover, for all $i=1,...,N$,*

$$
|(p^h-p)(t_i,x,y)|\le Cg_c(t_i,x-y)h^{\frac{1+\beta}{2}-\epsilon}(1+t_i^{-\frac{\beta}{2}}).
$$

We start with a simple computational approach that does not involve heat kernel estimates. Here, we only profit from the order of $|X^h_{\phi(t)}-X^h_t|$ and thus we obtain that $|\mathcal{E}(h,T,f,x)|$ is of order $h^{\beta/2}$. However, following a more tedious approach (Section 2.2.2-2.2.3), we are able to gain $1/2-\epsilon$, $\epsilon>0$, in the power of h which reflects the parabolic bootstrap of the underlying PDE (see e.g. [Fri83], [LS68]).

2.2.1 First simple approach

Proposition 2.3. *For h small enough, there exists a constant C such that*

$$
|\mathcal{E}(h,T,f,x)|\le C\|b\|_{C^\beta}\|\nabla v\|_\infty h^{\beta/2}T.
$$

Proof. We first show that

$$
\mathcal{E}(h,T,f,x)=\int_0^T\mathbb{E}[(b(X^h_{\phi(t)})-b(X^h_t))\cdot\nabla v(t,X^h_t)]dt.
$$

Let $v \in C_b^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$ be the solution of the PDE (6). Applying Itô's formula on $v(T, X_T^h)$ we get

$$
\begin{aligned}
v(T, X_T^h) &= v(0, X_0^h) + \int_0^T \partial_t v(t, X_t^h)dt + \int_0^T \nabla v(t, X_t^h) \cdot dX_t^h + \frac{1}{2}\int_0^T \Delta v(t, X_t^h)dt \\
&= v(0, X_0^h) + \int_0^T (\partial_t + \frac{1}{2}\Delta) v(t, X_t^h)dt + \int_0^T b(t, X_{\phi(t)}^h) \cdot \nabla v(t, X_t^h)dt + \int_0^T \nabla v(t, X_t^h) \cdot dW_t \\
&= v(0, X_0^h) + \int_0^T (\partial_t + b(X_t^h) \cdot \nabla + \frac{1}{2}\Delta) v(t, X_t^h)dt + \int_0^T (b(t, X_{\phi(t)}^h) - b(X_t^h)) \cdot \nabla v(t, X_t^h)dt + \int_0^T \nabla v(t, X_t^h) \cdot dW_t \\
&= v(0, X_0^h) + \int_0^T (b(t, X_{\phi(t)}^h) - b(X_t^h)) \cdot \nabla v(t, X_t^h)dt + \int_0^T \nabla v(t, X_t^h) \cdot dW_t,
\end{aligned}
$$

and so $\mathbb{E}_x[f(X_T^h)] = v(0,x) + \int_0^T \mathbb{E}[(b(X_{\phi(t)}^h) - b(X_t^h)) \cdot \nabla v(t, X_t^h)]dt$, using Fubini's theorem for the second term. Similarly, we have

$$
\begin{aligned}
v(T, X_T) &= v(0, X_0) + \int_0^T \partial_t v(t, X_t)dt + \int_0^T \nabla v(t, X_t) \cdot dX_t + \frac{1}{2}\int_0^T \Delta v(t, X_t)dt \\
&= v(0, X_0) + \int_0^T \nabla v(t, X_t) \cdot dW_t,
\end{aligned}
$$

and $\mathbb{E}_x[f(X_T)] = v(0,x)$. Finally, we obtain

$$
\begin{aligned}
\mathcal{E}(h, T, f, x) &= v(0,x) + \int_0^T \mathbb{E}[(b(X_{\phi(t)}^h) - b(X_t^h)) \cdot \nabla v(t, X_t^h)]dt - v(0,x) \\
&= \int_0^T \mathbb{E}[(b(X_{\phi(t)}^h) - b(X_t^h)) \cdot \nabla v(t, X_t^h)]dt.
\end{aligned}
$$

Let us now analyse the error. We write

$$
\begin{aligned}
|\mathcal{E}(h, T, f, x)| &\le \int_0^T \left|\mathbb{E}[(b(X_{\phi(t)}^h) - b(X_t^h)) \cdot \nabla v(t, X_t^h)]\right| dt \\
&\le \|\nabla v\|_\infty \int_0^T \mathbb{E}[|b(X_{\phi(t)}^h) - b(X_t^h)|]dt \\
&= \|\nabla v\|_\infty \int_0^T \mathbb{E}\left[\frac{|b(X_{\phi(t)}^h) - b(X_t^h)|}{|X_{\phi(t)}^h - X_t^h|^\beta} |X_{\phi(t)}^h - X_t^h|^\beta\right] dt \\
&\le \|\nabla v\|_\infty \|b\|_{C^\beta} \int_0^T \mathbb{E}[|X_{\phi(t)}^h - X_t^h|^\beta]dt.
\end{aligned}
$$

Note that

$$
\begin{aligned}
|X_{\phi(t)}^h - X_t^h| &= |\int_0^{\phi(t)} b(X_{\phi(s)}^h)ds - \int_0^t b(X_{\phi(s)}^h)ds + W_{\phi(t)} - W_t| \\
&\le |\int_t^{\phi(t)} b(X_{\phi(s)}^h)ds| + |W_{\phi(t)} - W_t| \\
&\le \|b\|_\infty |\phi(t) - t| + |W_{\phi(t)} - W_t| \\
&\le \|b\|_\infty h + |W_{\phi(t)} - W_t|.
\end{aligned}
$$

Taking expectation we get,

$$
\begin{aligned}
\mathbb{E}[|X_{\phi(t)}^h - X_t|^\beta] &\le \|b\|_\infty^\beta h^\beta + \mathbb{E}[|W_{\phi(t)} - W_t|^\beta] \\
&= \|b\|_\infty^\beta h^\beta + \mathbb{E}[|W_{\phi(t)-t}|^\beta] \\
&\le \|b\|_\infty^\beta h^\beta + h^{\beta/2} \\
&\le C h^{\beta/2},
\end{aligned}
$$

for $h \le 1$. Combining the previous results, we obtain

$$\begin{aligned}|\mathcal{E}(h,T,f,x)| &\le \|\nabla v\|_\infty \|b\|_{C^\beta} \int_0^T Ch^{\beta/2}dt \\ &= Cv\|_\infty \|b\|_{C^\beta} h^{\beta/2}T.\end{aligned}$$

□

The demonstrated approach was the original approach of [MP91]. The obtained rate is somehow optimal when the diffusion coefficient is not identity (see [KM17])

2.2.2 Density estimates for the Euler scheme and the SDE

To prove Theorem 2.1, we have to first establish the existence of density for the SDE (3) and its Euler scheme (7) and the related properties. This section is devoted to proving the following propositions.

Proposition 2.4 (Properties and estimates for density of the SDE). *For any $t \in (0,T]$ and $x, y \in \mathbb{R}^d$, the solution to* (3) *admits a density $p(t,x,y)$ which admits Duhamel representation*

$$p(t,x,y) = g(t, y-x) - \int_0^t \mathbb{E}[b(X_s)\cdot \nabla_y g(t-s, y-X_s)]ds. \tag{10}$$

The density satisfies Aronson type bound, i.e. for all $t \in (0,T]$, all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, there exists $C > 0$ such that

$$p(t,x,y) \le Cg_c(t, y-x), \tag{11}$$

Moreover, for fixed $(t,x) \in (0,T] \times \mathbb{R}^d$, $y, y' \in \mathbb{R}^d$ s.t. $|y-y'| \le t^{1/2}$, there exists a constant $C_\beta > 0$ such that

$$|p(t,x,y) - p(t,x,y')| \le C_\beta \frac{|y-y'|^\beta}{t^{\beta/2}} g_c(t, y-x). \tag{12}$$

Proposition 2.5 (Properties and estimates for density of the Euler scheme). *For any $t \in (0,T]$ and $x, y \in \mathbb{R}^d$, the solution to* (7) *admits a density $p^h(t,x,y)$ which admits Duhamel representation*

$$p^h(t,x,y) = g(t, y-x) - \int_0^t \mathbb{E}[b(X^h_{\phi(s)})\cdot \nabla_y g(t-s, y-X^h_s)]ds. \tag{13}$$

The density satisfies Aronson type bound, i.e. for all $t \in (0,T]$, all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, there exists $C > 0$ such that

$$p^h(t_j,x,y) \le Cg_c(t_j, y-x),$$

Moreover, for fixed $(t,x) \in (0,T] \times \mathbb{R}^d$, $y, y' \in \mathbb{R}^d$ s.t. $|y-y'| \le t^{1/2}$, there exists a constant $C_\beta > 0$ such that

$$|p^h(t,x,y) - p^h(t,x,y')| \le C_\beta \frac{|y-y'|^\beta}{t^{\beta/2}} g_c(t, y-x).$$

Let us first prove Proposition 2.5. We will need the following auxiliary results concerning the Gaussian density.

Lemma 2.2 (Usual Gaussian controls). *For all $z \in \mathbb{R}^d$, $u \in (0,T]$, there exist $C, c > 0$ such that*

$$|\nabla g(u,z)| \le C\frac{g_c(u,z)}{u^{1/2}}, \quad |\partial_u g(u,z)| \le C\frac{g_c(u,z)}{u}.$$

Proof. Let $z \in \mathbb{R}^d$, $u \in (0,T]$ be fixed. Since it is always possible to bound $g(u,z)\frac{|z|}{u^{1/2}}$ by $Cg_c(u,z)$ for some $C > 0$ and $c > 0$, we easily obtain

$$\begin{aligned}|\nabla g(u,z)| &= g(u,z)\frac{|z|}{u} \\ &\le Cg_c(u,z)\frac{1}{u^{1/2}}.\end{aligned}$$

For the time derivative we have we similarly have

$$
\begin{aligned}
|\partial_u g(u,z)| &\le \frac{1}{2u}\frac{|z|^2}{u} g(u,z) + \frac{d}{2u} g(u,z) \\
&\le C\frac{1}{2u} g_c(u,z) + \frac{d}{2u} g(u,z) \\
&\le C g_c(u,z)\frac{1}{u}.
\end{aligned}
$$

□

Proposition 2.6 (Negligibility of drift in one time step)**.** *Let $j = 1, ..., N$ be fixed. Then, for any $z \in \mathbb{R}^d$, $s \in [0, t_j]$,*

$$\mathbb{E}[g_c(t_j - s, z - X_s^h)] \le C\mathbb{E}[g_c(t_j - \phi(s), z - X_{\phi(s)}^h)].$$

Proof. Using Markov property of the Euler scheme, the convolution property of two Gaussian densities and the fact that b is bounded, we obtain:

$$
\begin{aligned}
\mathbb{E}[g_c(t_j - s, z - X_s^h)] &= \mathbb{E}[\mathbb{E}_{\mathcal{F}_{\phi(s)}}[g_c(t_j - s, z - X_s^h)]] \\
&= \mathbb{E}[\mathbb{E}_{X^h_{\phi(s)}}[g_c(t_j - s, z - X_s^h)]] \\
&= \mathbb{E}[\int_{\mathbb{R}^d} p^h(s - \phi(s), y, X_{\phi(s)}^h + b(X_{\phi(s)}^h)(s - \phi(s))) g_c(t_j - s, z - y) dy] \\
&\le C\mathbb{E}[\int_{\mathbb{R}^d} g_c(s - \phi(s), y - X_{\phi(s)}^h - b(X_{\phi(s)}^h)(s - \phi(s))) g_c(t_j - s, z - y) dy] \\
&\le C\mathbb{E}[\int_{\mathbb{R}^d} g_c(s - \phi(s), y - X_{\phi(s)}^h) g_c(t_j - s, z - y) dy] \\
&\le C\mathbb{E}[g_c(t_j - \phi(s), z - X_{\phi(s)}^h)].
\end{aligned}
$$

□

Existence of density of the Euler scheme Let us first show that the solution to (7) admits density. Suppose $t \in (0, h)$. Then

$$
\begin{aligned}
X_t^h &= x + \int_0^t b(x) ds + W_t \\
&= x + b(x)t + W_t \\
&\sim \mathcal{N}(x + b(x)t, tI_d).
\end{aligned}
$$

It follows that for $t \in (0, h)$,

$$p^h(t, x, y) = (2\pi t)^{-d/2} \exp\Big(- \frac{|y - x - b(x)t|^2}{2t} \Big).$$

Suppose now $t \in (0, T]$. Then, conditionally on $X_{\phi(t)}^h = z \in \mathbb{R}^d$ and using Markov structure of the Euler scheme,

$$
\begin{aligned}
\mathbb{P}_x(X_t^h \in A) &= \mathbb{E}_x[\mathbb{E}[1_{X_t^h \in A} | \mathcal{F}_{\phi(t)}]] \\
&= \mathbb{E}_x[\mathbb{E}[1_{X_t^h \in A} | X_{\phi(t)}^h]] \\
&= \int_A p^h(t - \phi(t), z, dy) \\
&= \int_A p^h(t - \phi(t), z, y) dy,
\end{aligned}
$$

since $t - \phi(t) < h$. Thus, conditionally on $X_{\phi(t)}^h = z \in \mathbb{R}^d$ the density of X_t^h is given by $p^h(t - \phi(t), z, y)$.

Now, recall that $t_n = nh$ for $n = 1, ..., N$. Since $t \in (0, T]$, there exists $n = 1, ..., N$ s.t. $t \in (\phi(t), t_n]$.

$$
\begin{aligned}
\mathbb{P}_x(X_t^h \in A) &= \mathbb{E}[1_{X_t^h \in A} | X_0^h = x] \\
&= \int_{(\mathbb{R}^d)^{n-1}} \int_A p^h(t_1, x, z_1) p^h(t_2 - t_1, z_1, z_2) ... p^h(t - \phi(t), z_n, y) dy dz_1 ... dz_n.
\end{aligned}
$$

And so conditionally on $X_0^h = x \in \mathbb{R}^d$, the density of X_t^h is given by

$$p^h(t,x,y) = \int_{(\mathbb{R}^d)^{n-1}} p^h(t_1,x,z_1) \prod_{i=1}^{n-2} p^h(t_{i+1}-t_i, z_i, z_{i+1}) p^h(t_n - \phi(t), z_n, y) dz_1 ... dz_n.$$

Brownian motion and PDE The key tool for proving the Duhamel representation (13) is a PDE such that Brownian motion provides its solution. Consider the heat equation

$$\begin{cases} (\partial_t + \frac{1}{2}\Delta)u(t,x) = 0, & (t,x) \in [0,t) \times \mathbb{R}^d, \\ v(t,x) = f(x), & x \in \mathbb{R}^d, \end{cases} \tag{14}$$

where $f \in C_0^\infty(\mathbb{R}^d, \mathbb{R})$. Let W be a standard Brownian motion on $(\Omega, (\mathcal{F})_{s\geq 0}, \mathbb{P})$ and for fixed $t \in (0,T]$, $s \in [0,t]$, let

$$v(s,x) := \mathbb{E}[f(x + W_{t-s})] = \int_{\mathbb{R}^d} f(y) g(t-w, x-y) dy.$$

We show that v solves this heat equation $\mathbb{P}$-a.s and that it is unique solution belonging to $C_b^{1,2}([0,t] \times \mathbb{R}^d)$. Apply Itô's formula for $v(r, W_r)$ on $[0,s) \times \mathbb{R}^d$:

$$\begin{aligned} v(s, W_s) &= v(0, W_0) + \int_0^s \partial_r v(r, W_r) dr + \int_0^s \nabla v(r, W_r) \cdot dW_r + \frac{1}{2} \int_0^s \Delta v(r, W_r) dr \\ &= v(0,0) + \int_0^s \nabla v(r, W_r) \cdot dW_r + \int_0^s (\partial_r + \frac{1}{2}\Delta) v(r, W_r) dr. \end{aligned} \tag{15}$$

Note that

$$\begin{aligned} \mathbb{E}[v(s, W_s)] &= \mathbb{E}[f(W_s + W_{t-s})] \\ &= \mathbb{E}[f(W_s + W_t - W_s)] \\ &= \mathbb{E}[f(W_t)], \end{aligned}$$

and

$$v(0,0) = \mathbb{E}[f(W_t)].$$

And so taking expectation in (15), we obtain

$$\mathbb{E}[\int_0^s (\partial_r + \frac{1}{2}\Delta) v(r, W_r) dr] = 0.$$

But this means that for almost all $r \in [0,s]$ and for almost all $x \in \mathbb{R}^d$,

$$(\partial_r + \frac{1}{2}\Delta) v(r, W_r) = 0, \quad \mathbb{P}\text{-a.s.}$$

For the terminal condition we readily get for any $x \in \mathbb{R}^d$,

$$\begin{aligned} v(t,x) &= \mathbb{E}[f(x + W_{t-t})] \\ &= f(x). \end{aligned}$$

Therefore, we have shown that for a.a $(s,x) \in [0,t] \times \mathbb{R}^d$, $v(s,x)$ satisfies (14) $\mathbb{P}$-a.s. Moreover, by Theorem 3.1 in [Bas97], we we have that any solution to (14) is of the form $\mathbb{E}[f(x + W_{t-s})]$ and it is in $C_b^{1,2}([0,t] \times \mathbb{R}^d)$ (see [Fri83]).

Proposition 2.7 (Duhamel representation for the density of the Euler scheme)**.** *Denote the law of X_t^h by $\mu_{X_t^h}(dy) = p^h(t,x,y)dy$. Then for almost all $y \in \mathbb{R}^d$ it holds*

$$p^h(t,x,y) = g(t, y-x) - \int_0^t \mathbb{E}[b(X_{\phi(s)}^h) \cdot \nabla_y g(t-s, y - X_s^h)] ds. \tag{16}$$

Proof. Applying Itô's formula on $v(t, X_t^h)$, taking expectation and using (14), we obtain

$$\begin{aligned}\mathbb{E}[v(t, X_t^h)] &= v(0,x) + \mathbb{E}[\int_0^t (\partial_s + \frac{1}{2}\Delta)v(s, X_s^h)ds + \int_0^t \nabla v(s, X_s^h) \cdot b(X_{\phi(s)}^h)ds] \\ &= v(0,x) + \int_0^t \mathbb{E}[\nabla v(s, X_s^h) \cdot b(X_{\phi(s)}^h)]ds,\end{aligned}$$

using Fubini's theorem for the last equality. Using terminal condition in (14), we obtain

$$\mathbb{E}[f(X_t^h)] = v(0,x) + \int_0^t \mathbb{E}[\nabla v(s, X_s^h) \cdot b(X_{\phi(s)}^h)]ds.$$

Note since $v(s,y) = \mathbb{E}[f(y + W_{t-s})] = \int_{\mathbb{R}^d} f(y)g(t-s, x-y)dy$, we have

$$v(0,x) = \int_{\mathbb{R}^d} f(y)g(t, x-y)dy$$

and

$$\begin{aligned}\int_0^t \mathbb{E}[\nabla v(s, X_s^h) \cdot b(X_{\phi(s)}^h)]ds &= \int_0^t \mathbb{E}[b(X_{\phi(s)}^h) \cdot \nabla_z \int_{\mathbb{R}^d} f(y)g(t-s, X_s^h - y)dy]\big|_{z=X_s^h} ds \\ &= \int_0^t \mathbb{E}[b(X_{\phi(s)}^h) \int_{\mathbb{R}^d} f(y)\nabla_z g(t-s, X_s^h - y)dy]\big|_{z=X_s^h} ds \\ &= -\int_0^t \mathbb{E}[b(X_{\phi(s)}^h) \int_{\mathbb{R}^d} f(y)\nabla_y g(t-s, y - X_s^h)dy]ds \\ &= -\int_{\mathbb{R}^d} \int_0^t \mathbb{E}[b(X_{\phi(s)}^h) \cdot \nabla_y g(t-s, y - X_s^h)]f(y)dyds.\end{aligned}$$

Finally,

$$\begin{aligned}\mathbb{E}[f(X_t^h)] &= \int_{\mathbb{R}^d} f(y)p^h(t,x,y)dy \\ &= \int_{\mathbb{R}^d} f(y)g(t, x-y)dy - \int_{\mathbb{R}^d} \int_0^t \mathbb{E}[b(X_{\phi(s)}^h)\nabla_y g(t-s, y - X_s^h)]f(y)dyds,\end{aligned}$$

from which, since function f is arbitrary, we deduce that for almost all $y \in \mathbb{R}^d$,

$$p^h(t,x,y) = g(t, y-x) - \int_0^t \mathbb{E}[b(X_{\phi(s)}^h)\nabla_y g(t-s, y - X_s^h)]ds.$$

□

Lemma 2.3 (A priory (explosive) bounds on the density)**.** *For any $t \in (0,T]$,*

$$p^h(t,x,y) \le C^{\lfloor t/h \rfloor} g_c(t, x-y) \le C^N g_c(t, x-y). \tag{17}$$

Proof. We prove (17) by induction. Let $t \in (0,h)$. We show that there exists $C \ge 1$, $c \ge 1$ such that $p^h(t,x,y) \le C g_c(t, y-x)$. For such a t, we have

$$\begin{aligned}X_t^h &= x + \int_0^t b(X_{\phi(s)}^h)ds + W_t \\ &= x + b(x)t + W_t.\end{aligned}$$

Using inequality

$$|y - x - b(x)t|^2 \ge (1-\tilde{c})|y-x|^2 + (1-\tilde{c}^{-1})|b(x)t|^2$$

for $\tilde{c} \in (0,1)$, we obtain

$$\begin{aligned}
p^h(t,x,y) &= (2\pi t)^{-d/2} \exp\big(-\frac{|y-x-b(x)t|^2}{2t}\big) \\
&\le (2\pi t)^{-d/2} \exp\big(-\frac{(1-\tilde{c})|y-x|^2}{2t}\big) \exp\big(-\frac{(1-\tilde{c}^{-1})|b(x)t|^2}{2t}\big) \\
&= c^{d/2} \exp\big(-\frac{(1-\tilde{c}^{-1})|b(x)t|^2}{2t}\big)(2\pi ct)^{-d/2} \exp\big(-\frac{|y-x|^2}{2ct}\big) \\
&= C g_c(t, x-y),
\end{aligned}$$

for $c := \frac{1}{1-c} > 1$, $C := c^{d/2} \exp\big(-\frac{(1-\tilde{c}^{-1})|b(x)t|^2}{2t}\big)$. Now assume that (17) holds for $t \in [(n-1)h, nh)$ for $n \le N$. Then, since X_t^h is a Markov process, we can use Chapman-Kolmogorov representation for $t \in [nh, (n+1)h)$. Denote $t_n := nh$. Then,

$$\begin{aligned}
p^h(t_n,x,y) &= \int_{\mathbb{R}^d} p^h(t_{n-1},x,z) p^h(t_n - t_{n-1}, z, y) dz \\
&\le \int_{\mathbb{R}^d} C^{n-1} g_c(t_n, z-x) \cdot C g_c(t_n - t_{n-1}, y-z) dz \\
&\le C^n g_c(t_n, y-x).
\end{aligned}$$

Now, using this result for $t \in [nh, (n+1)h)$, we obtain,

$$\begin{aligned}
p^h(t,x,y) &= \int_{\mathbb{R}^d} p^h(t_n,x,z) p^h(t-t_n, z, y) dz \\
&\le C^n \int_{\mathbb{R}^d} g_c(t_n, x-z) g_c(t-t_n, z-y) dz \\
&\le C^n g_c(t, x-y) \\
&= C^{\lfloor t/h \rfloor} g_c(t, x-y),
\end{aligned}$$

since $n \le t/h \le n+1$. $\square$

Although the bound in Lemma 2.3 is explosive as N goes to infinity, it allows us to obtain a uniform bound on the density of the Euler scheme which is proven in the Proposition 2.8 by applying a Grönwall type lemma.

Proposition 2.8. *For T small enough, there exists $\tilde{C} > 0$ such that for all $j = 1, ..., N$,*

$$p^h(t_j, x, y) \le \tilde{C} g_c(t_j, y-x),$$

where $c > 0$, $g_c(u,z) = (2\pi uc)^{-\frac{d}{2}} \exp\big(-\frac{|z|^2}{2cu}\big)$ is the density of a centered Gaussian vector with covariance matrix cI_d.

Proof. Define for all $j = 1, ..., N$, $m_{t_j}^h := \sup_{z \in \mathbb{R}^d} \frac{p(t_j,x,z)}{g(t_j,z-x)}$. We first derive a bound on $m_{t_j}^h$. Let $j = 1, ..., N$ be fixed.

We can rewrite X_s^h as

$$\begin{aligned}
X_s^h &= X_{\phi(s)}^h + (X_s^h - X_{\phi(s)}^h) \\
&= X_{\phi(s)}^h + b(X_{\phi(s)}^h)(s - \phi(s)) + W_s - W_{\phi(s)} \\
&\overset{\text{law}}{=} X_{\phi(s)}^h + b(X_{\phi(s)}^h)(s - \phi(s)) + \mathcal{N}(s - \phi(s))^{1/2} \\
&=: X_{\phi(s)}^h + \Phi(X_{\phi(s)}^h),
\end{aligned}$$

where $\mathcal{N} \sim \text{Gaussian}(0, I_d)$. Then using the Duhamel representation (13) and Proposition 2.6, we write

$$
\begin{aligned}
\frac{p(t_j,x,z)}{g_c(t_j,z-x)} &= \frac{g_c(t_j,z-x)}{g_c(t_j,z-x)} - \frac{1}{g_c(t_j,z-x)}\int_0^{t_j}\mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_z g(t_j-s,z-X^h_s)]ds\\
&\le 1+\frac{1}{g_c(t_j,z-x)}\int_0^{t_j}\int_{\mathbb{R}^d}p^h(\phi(s),x,w)b(w)\mathbb{E}[\nabla_z g(t_j-s,z-w-\Phi(w))]dwds\\
&\le 1+\|b\|_\infty\frac{1}{g_c(t_j,z-x)}\int_0^{t_j}\int_{\mathbb{R}^d}p^h(\phi(s),x,w)\mathbb{E}[|\nabla_z g(t_j-s,z-w-\Phi(w))|]dwds\\
&\le 1+C\|b\|_\infty\frac{1}{g_c(t_j,z-x)}\int_0^{t_j}\int_{\mathbb{R}^d}p^h(\phi(s),x,w)\mathbb{E}[g_c(t_j-s,z-w-\Phi(w))](t_j-s)^{-1/2}dwds\\
&\le 1+C\|b\|_\infty\frac{1}{g_c(t_j,z-x)}\int_0^{t_j}\int_{\mathbb{R}^d}\frac{p^h(\phi(s),x,w)}{g_c(\phi(s),x-w)}g_c(\phi(s),x-w)g_c(t_j-s,z-w)(t_j-s)^{-1/2}dwds\\
&\le 1+C\|b\|_\infty\frac{1}{g_c(t_j,z-x)}\int_0^{t_j}\frac{m^h_{\phi(s)}}{(t_j-s)^{1/2}}\int_{\mathbb{R}^d}g_c(\phi(s),x-w)g_c(t_j-s,z-w)dwds\\
&\le 1+C\|b\|_\infty\frac{1}{g_c(t_j,z-x)}\int_0^{t_j}\frac{m^h_{\phi(s)}}{(t_j-s)^{1/2}}\int_{\mathbb{R}^d}g_c(s,x-w)g_c(t_j-s,z-w)dwds\\
&= 1+C\|b\|_\infty\int_0^{t_j}m^h_{\phi(s)}\frac{1}{(t_j-s)^{1/2}}ds.
\end{aligned}
$$

Taking supremum over all $z\in\mathbb{R}^d$, we obtain

$$
\begin{aligned}
m^h_{t_j} &\le C\Big(1+\|b\|_\infty\int_0^{t_j}m^h_{\phi(s)}\frac{1}{(t_j-s)^{1/2}}ds\Big)\\
&\le C\Big(1+\|b\|_\infty\sup_{s\in[0,h]}m^h_{\phi(s)}\int_0^{t_j}\frac{1}{(t_j-s)^{1/2}}ds\Big)\\
&\le C\Big(1+\|b\|_\infty\sup_{s\in[0,h]}m^h_{\phi(s)}t_j^{1/2}\Big).
\end{aligned}
$$

Since by Proposition 2.3 $\sup_{s\in[0,h]}m^h_{\phi(s)}<\bar{C}$, for some $\bar{C}>0$, from the previous computations we have that

$$
\begin{aligned}
p^h(t_j,x,y) &\le C\Big(1+\|b\|_\infty\bar{C}t_j^{1/2}\Big)g_c(t_j,y-x)\\
&=\tilde{C}g_c(t_j,y-x),
\end{aligned}
$$

where $\tilde{C}<+\infty$ for T small enough. □

From now on for establishing the desired bounds, we will be distinguishing two main cases for the position of space sensitivity with respect to the current time. The first regime called *diagonal regime*, and it includes the points $y,y'\in\mathbb{R}^d$ such that $|y-y'|\le t^{1/2}$, where $t\in[0,T]$ is a fixed time, i.e. the points are close to each other w.r.t. to the time. This setup allows us to exploit sensitivities of the underlying heat kernel. On the contrary, *off-diagonal regime* includes the points $y,y'\in\mathbb{R}^d$ such that $|y-y'|\ge t^{1/2}$ for $t\in[0,T]$, and in this case we can exploit the smoothness properties of the heat kernel. In the upcoming statements and their proofs, the separation of these regimes will be normally used.

Proposition 2.9 (Spatial sensitivity of the Euler scheme's density in the forward variable)**.** *For fixed $(t,x)\in(0,T]\times\mathbb{R}^d$ and $y,y'\in\mathbb{R}^d$ such that $|y-y'|\le t^{1/2}$, there exists $C_\beta>0$ such that*

$$
|p^h(t,x,y)-p^h(t,x,y')|\le C_\beta\frac{|y-y'|^\beta}{t^{\beta/2}}g_c(t,y-x).
$$

Proof. Let $(t,x)\in(0,T]\times\mathbb{R}^d$ be fixed and $y,y'\in\mathbb{R}^d$ such that $|y-y'|\le t^{1/2}$. From (13) we have

$$
\begin{aligned}
|p^h(t,x,y)-p^h(t,x,y')| &\le |g(t,y-x)-g(t,y'-x)| \qquad (18)\\
&\quad +|\int_0^t\mathbb{E}_x[b(X^h_{\phi(s)})\cdot(\nabla_y g(t-s,y-X^h_s)-\nabla_y g(t-s,y'-X^h_s))]ds| \qquad (19)
\end{aligned}
$$

We start with the first term. We rewrite it using fundamental theorem of calculus as

$$\begin{aligned}
|g(t, y-x) - g(t, y'-x)| &= \Big| \int_0^1 \nabla_\lambda g(t, z)\big|_{z=x-(y'+\lambda(y-y'))} \cdot (y-y') d\lambda \Big| \\
&\leq \int_0^1 |\nabla_\lambda g(t, z)\big|_{z=x-(y'+\lambda(y-y'))}||y-y'| d\lambda \\
&\leq C \int_0^1 g_c(t, x-(y'+\lambda(y-y'))) \frac{|y-y'|}{\sqrt{t}} d\lambda,
\end{aligned}$$

using Lemma 2.2 for the last inequality. Noting that

$$\begin{aligned}
|x-(y'+\lambda(y-y'))|^2 &= |x-y+(1-\lambda)(y-y')|^2 \\
&\geq |x-y|^2 - 2(1-\lambda)|x-y||y-y'| + (1-\lambda)^2|y-y'|^2 \\
&\geq (1-\epsilon)|x-y|^2 + (1-\epsilon^{-1})(1-\lambda)^2|y-y'|^2,
\end{aligned}$$

for $\epsilon > 0$, and using the fact that $|y-y'|^2 \leq t$, we obtain

$$\begin{aligned}
g_c(t, z) &\leq \exp\Big(-\frac{(1-\epsilon)|x-y|^2}{2ct}\Big) \exp\Big(\frac{(\epsilon^{-1}-1)(1-\lambda)^2|y-y'|^2}{2ct}\Big) \\
&\leq C \exp\Big(-\frac{|x-y|^2}{2ct}\Big),
\end{aligned} \tag{20}$$

where $c := \frac{c}{1-\epsilon}$, $C := \exp\Big(\frac{\epsilon^{-1}-1}{2c}\Big)$. And so we get

$$\begin{aligned}
|g(t, y-x) - g(t, y'-x)| &\leq C \int_0^1 g_c(t, x-y) \frac{|y-y'|}{\sqrt{t}} d\lambda \\
&= C g_c(t, x-y) \frac{|y-y'|^\beta}{t^{\beta/2}} \Big(\frac{|y-y'|}{t^{1/2}}\Big)^{1-\beta} \\
&\leq C g_c(t, x-y) \frac{|y-y'|^\beta}{t^{\beta/2}}.
\end{aligned} \tag{21}$$

For bounding the second term in (18) we consider so called *local diagonal in time regime* and *local off-diagonal in time regime*. More precisely, we write,

$$\begin{aligned}
&\Big| \int_0^t \mathbb{E}_x[b(X^h_{\phi(s)}) \cdot (\nabla_y g(t-s, y-X^h_s) - \nabla_y g(t-s, y'-X^h_s))] ds \Big| \\
&\qquad = \Big| \int_0^t \mathbb{E}_x[b(X^h_{\phi(s)}) \cdot (\nabla_y g(t-s, y-X^h_s) - \nabla_y g(t-s, y'-X^h_s))](\mathbf{1}_{D_{loc}}(s) + \mathbf{1}_{D^c_{loc}}(s)) ds \Big| \\
&\qquad =: |T_1 + T_2|(t, y, y') \\
&\qquad \leq |T_1(t, y, y')| + |T_2(t, y, y')|,
\end{aligned}$$

where $D_{loc} := \{s \geq t : \ |y-y'| \leq (t-s)^{1/2}\}$ is the diagonal regime w.r.t. the current running time. We start bounding T_1 using fundamental theorem of calculus and using the property of the set D_{loc}.

$$
\begin{aligned}
|T_1(t,y,y')| &= |\int_0^t \mathbb{E}_x[b(X^h_{\phi(s)})\cdot(\nabla_y g(t-s,y-X^h_s)-\nabla_y g(t-s,y'-X^h_s))]\mathbf{1}_{D_{loc}}(s)ds| \\
&\le \|b\|_\infty \int_0^t \mathbb{E}_x[|\nabla_y g(t-s,y-X^h_s)-\nabla_y g(t-s,y'-X^h_s)|]\mathbf{1}_{D_{loc}}(s)ds \\
&= \|b\|_\infty \int_0^t \mathbb{E}_x[|\int_0^1 H_g(t-s,y-X^h_s+\lambda(y-y'))(y-y')d\lambda|]\mathbf{1}_{D_{loc}}(s)ds \\
&\le \|b\|_\infty \int_0^t \mathbb{E}_x[\int_0^1 |H_g(t-s,y-X^h_s+\lambda(y-y'))||y-y'|d\lambda]\mathbf{1}_{D_{loc}}(s)ds \\
&\le C\|b\|_\infty \int_0^t\int_0^1 \mathbb{E}_x\Big[g_c(t-s,y-X^h_s+\lambda(y-y'))\Big]\frac{|y-y'|}{t-s}d\lambda\mathbf{1}_{D_{loc}}(s)ds \\
&\le C\|b\|_\infty \int_0^t \mathbb{E}_x\Big[g_c(t-s,y-X^h_s)\Big]\frac{|y-y'|}{t-s}\mathbf{1}_{D_{loc}}(s)ds \\
&= C\|b\|_\infty \int_0^t \mathbb{E}_x\Big[g_c(t-s,y-X^h_s)\Big]\Big(\frac{|y-y'|}{(t-s)^{1/2}}\Big)^{1-\beta}\frac{|y-y'|^\beta}{(t-s)^{(1+\beta)/2}}\mathbf{1}_{D_{loc}}(s)ds \\
&\le C\|b\|_\infty|y-y'|^\beta \int_0^t \mathbb{E}_x\Big[g_c(t-s,y-X^h_s)\Big]\frac{1}{(t-s)^{(1+\beta)/2}}ds.
\end{aligned}
$$

Note that since $\beta\in(0,1)$, $(t-s)^{-(1+\beta)/2}$ is integrable on the time interval $[0,t]$. Since by Proposition 2.8 we can bound the density of X^h_s, i.e.

$$
\begin{aligned}
\mathbb{E}_x\Big[g_c(t-s,y-X^h_s)\Big] &= \int_{\mathbb{R}^d} g_c(t-s,y-z)p^h(s,x,z)dz \\
&\le C\int_{\mathbb{R}^d} g_c(t-s,y-z)g_c(s,x-z)dz \\
&= Cg_c(t,y-x),
\end{aligned}
$$

we have

$$
\begin{aligned}
|T_1(t,y,y')| &\le C\|b\|_\infty g_c(t,y-x)|y-y'|^\beta\int_0^t \frac{1}{(t-s)^{(1+\beta)/2}}ds \\
&= C_\beta\|b\|_\infty g_c(t,y-x)\frac{|y-y'|^\beta}{t^{(\beta-1)/2}}.
\end{aligned}
\tag{22}
$$

For the second term we use triangular inequality, the property of the set D^c_{loc} and Proposition 2.8 again:

$$
\begin{aligned}
|T_2(t,y,y')| &= |\int_0^t \mathbb{E}_x[b(X^h_{\phi(s)})\cdot(\nabla_y g(t-s,y-X^h_s)-\nabla_y g(t-s,y'-X^h_s))]\mathbf{1}_{D^c_{loc}}(s)ds| \\
&\le \|b\|_\infty\int_0^t \mathbb{E}_x[|\nabla_y g(t-s,y-X^h_s)|+|\nabla_y g(t-s,y'-X^h_s)|]\mathbf{1}_{D^c_{loc}}(s)ds \\
&\le C\|b\|_\infty\int_0^t \mathbb{E}_x[g_c(t-s,y-X^h_s)\frac{1}{(t-s)^{1/2}}+g_c(t-s,y'-X^h_s)\frac{1}{(t-s)^{1/2}}]\mathbf{1}_{D^c_{loc}}(s)ds \\
&\le C\|b\|_\infty\int_0^t \mathbb{E}_x[g_c(t-s,y-X^h_s)+g_c(t-s,y'-X^h_s)]\frac{1}{(t-s)^{1/2}}\frac{|y-y'|^\beta}{(t-s)^{\beta/2}}ds \\
&= C\|b\|_\infty|y-y'|^\beta\int_0^t\Big(\int_{\mathbb{R}^d}(g_c(t-s,y-z)+g_c(t-s,y'-z))p^h(s,x,z)dz\Big)\frac{1}{(t-s)^{(1+\beta)/2}}ds \\
&\le C\|b\|_\infty|y-y'|^\beta\int_0^t\Big(\int_{\mathbb{R}^d}(g_c(t-s,y-z)+g_c(t-s,y'-z))g_c(s,x-z)dz\Big)\frac{1}{(t-s)^{(1+\beta)/2}}ds \\
&= C\|b\|_\infty|y-y'|^\beta\int_0^t\big(g_c(t,y-x)+g_c(t,y'-x)\big)\frac{1}{(t-s)^{(1+\beta)/2}}ds.
\end{aligned}
$$

Recall that we are in the global diagonal regime, i.e. $|y-y'| \le t^{1/2}$, thus similarly to (20), we have

$$
\begin{aligned}
g_c(t, y'-x) &= (2\pi c t)^{-\frac{d}{2}} \exp\Big(-\frac{|y'-y+y-x|^2}{2ct}\Big) \\
&\le (2\pi c t)^{-\frac{d}{2}} \exp\Big(\frac{C|y'-y|^2}{t}\Big) \exp\Big(-\frac{|y-x|^2}{2ct}\Big) \\
&= C g_c(t, y-x).
\end{aligned}
$$

Then,

$$
\begin{aligned}
|T_2(t,y,y')| &\le C\|b\|_\infty g_c(t,y-x)|y-y'|^\beta \int_0^t \frac{1}{(t-s)^{(1+\beta)/2}} ds \\
&= C_\beta \|b\|_\infty g_c(t,y-x) \frac{|y-y'|^\beta}{t^{(\beta-1)/2}}.
\end{aligned} \tag{23}
$$

Putting together bounds (21), (22) and (23), we get

$$
\begin{aligned}
|p^h(t,x,y) - p^h(t,x,y')| &\le |g(t,y-x) - g(t,y'-x)| + |T_1(t,y,y')| + |T_2(t,y,y')| \\
&\le C_\beta g_c(t,y-x) \frac{|y-y'|^\beta}{t^{\beta/2}}.
\end{aligned}
$$

□

Proof of Proposition 2.4 The proof of the existence of the density $p(t,\cdot,\cdot)$ for all $t \in (0,T]$ can be found in [JM24]. Knowing that the density exists, the Duhamel repsentation (10) is shown by the same argument as for the Duhamel representation for the Euler density (13) using the appropriate PDE. For the proof of Aronson estimate for the density p, we refer to [Aro67]. The spatial sensitivity bound (12) can be shown similarly to Proposition 2.9.

2.2.3 Weak error for densities

The idea for estimating $|(p^h - p)(t,x,\cdot)|$ for fixed $(t,x) \in (0,T] \times \mathbb{R}^d$ is to rewrite this difference as a sum of terms in such a way that each term exploits the regularity of underlying functions in the best possible way on different time intervals. Using Duhamel representation (10) and (13), we write

$$
\begin{aligned}
(p^h - p)(t,x,y) &= \int_0^t (\mathbb{E}[b(X_s)\cdot\nabla_y g(t-s,y-X_s)] - \mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_y g(t-s,y-X^h_s)])ds \\
&= \int_0^h (\mathbb{E}[b(X_s)\cdot\nabla_y g(t-s,y-X_s)] - \mathbb{E}[b(x)\cdot\nabla_y g(t-s,y-X^h_s)])ds \\
&\quad + \int_h^t (\mathbb{E}[b(X_s)\cdot\nabla_y g(t-s,y-X_s)] - \mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_y g(t-s,y-X^h_s)])ds,
\end{aligned}
$$

where for the first integral in the second equality we use the fact that $\phi(s) = 0$ for $s \in [0, h)$. Now, adding and subtracting two additional terms and regrouping them, we obtain,

$$\begin{aligned}
&\int_h^t (\mathbb{E}[b(X_s)\cdot\nabla_y g(t-s, y-X_s)] - \mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_y g(t-s, y-X^h_s)])ds \\
&\qquad = \int_h^t (\mathbb{E}[b(X_s)\cdot\nabla_y g(t-s, y-X_s)] - \mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_y g(t-s, y-X^h_s)] \\
&\qquad\quad \pm \mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_y g(t-s, y-X^h_{\phi(s)})] \pm \mathbb{E}[b(X_{\phi(s)})\cdot\nabla_y g(t-s, y-X_{\phi(s)})])ds \\
&\qquad = \int_h^t \mathbb{E}[b(X_{\phi(s)})\cdot\nabla_y g(t-s, y-X_{\phi(s)}) - b(X^h_{\phi(s)})\cdot\nabla_y g(t-s, y-X^h_{\phi(s)})]ds \\
&\qquad\quad + \int_h^t \mathbb{E}[b(X_s)\cdot\nabla_y g(t-s, y-X_s) - b(X_{\phi(s)})\cdot\nabla_y g(t-s, y-X_{\phi(s)}]ds \\
&\qquad\quad + \int_h^t \mathbb{E}[b(X^h_{\phi(s)})\cdot\nabla_y g(t-s, y-X^h_{\phi(s)}) - b(X^h_{\phi(s)})\cdot\nabla_y g(t-s, y-X^h_s)]ds \\
&\qquad = \int_h^t \int_{\mathbb{R}^d} (p(\phi(s), x, z) - p^h(\phi(s), x, z))b(z)\nabla g(t-s, y-z)dzds \\
&\qquad\quad + \int_h^t \int_{\mathbb{R}^d} (p(s, x, z) - p(\phi(s), x, z))b(z)\nabla g(t-s, y-z)dzds \\
&\qquad\quad + \int_h^t \mathbb{E}[b(X^h_{\phi(s)})\cdot(\nabla_y g(t-s, y-X^h_{\phi(s)}) - \nabla_y g(t-s, y-X^h_s))]ds.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
(p^h - p)(t, x, y) &= \int_h^t \int_{\mathbb{R}^d} (p(\phi(s), x, z) - p^h(\phi(s), x, z))b(z)\nabla g(t-s, y-z)dzds \\
&\quad + \int_h^t \int_{\mathbb{R}^d} (p(s, x, z) - p(\phi(s), x, z))b(z)\nabla g(t-s, y-z)dzds \\
&\quad + \int_h^t \mathbb{E}[b(X^h_{\phi(s)})\cdot(\nabla_y g(t-s, y-X^h_{\phi(s)}) - \nabla_y g(t-s, y-X^h_s))]ds \\
&\quad + \int_0^h \big(b(x)\mathbb{E}[\nabla_y g(t-s, y-X_s)] - \mathbb{E}[b(X_s)\nabla_y g(t-s, y-X^h_s)]\big)ds \\
&=: \sum_{l=1}^4 \Delta^h_l(t, x, y).
\end{aligned}$$

This representation allows us to get the desired estimates. First term Δ^h_1 is a "Grönwall like" term, meaning that the difference between the density of the solution to the original SDE and the density of the Euler scheme appears on both sides of inequality and the corresponding lemma can be applied. In the term Δ^h_2 we exploit smoothness in time of the density on the interval $[h, t]$. In bounding the third term Δ^h_3 we, roughly speaking, profit from the difference of gradients of the Gaussian densities where X^h_s and $X^h_{\phi(s)}$ appears respectively on the interval $[h, t]$. In the last term Δ^h_4 we profit from the smallness of the integration interval and the classical bounds on the Gaussian density.

We now state estimates for all terms Δ^h_l separately, and their proofs, which are quite technical, are postponed to Section 2.3.1.

Lemma 2.4 (Circular term)**.** *Let $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $i = 1, ..., N$ be fixed. Define $e^h_j := \sup_{y\in\mathbb{R}^d} \frac{|(p^h-p)(t_j,x,y)|}{g_c(t_j,y-x)}$ for $j = 1, ..., N$. Then for all $j = 1, ..., N$ there exists $\tilde{C} > 0$ such that $e^h_j < \tilde{C}$ and*

$$|\Delta^h_1(t_i, x, y)| \le C\|b\|_\infty t_i^{\frac{1}{2}} g_c(t_i, y-x) \max_{j=1,...,N} e^h_j.$$

Lemma 2.5 (Time sensitivity of the density)**.** *Let $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $i = 1, ..., N$ be fixed. Then for any $\epsilon > 0$ there exists a constant $C_{\beta,\epsilon}$ such that*

$$|\Delta^h_2(t_i, x, y)| \le C_{\beta,\epsilon} h^{\frac{1+\beta}{2}-\epsilon} t_i^{-\frac{\beta}{2}} g_c(t_i, y-x).$$

Lemma 2.6 (Main sensitivity of the Euler scheme)**.** *Let $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $i = 1, ..., N$ be fixed. Then there exists a constant $C_{\beta,\epsilon}$ such that*

$$|\Delta_3^h(t_i,x,y)| \le C_{\beta,\epsilon} h^{\frac{1+\beta}{2}-\epsilon} t_i^{-\frac{\beta}{2}} g_c(t_i, y-x).$$

Lemma 2.7 (First time step term)**.** *Let $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $i = 1, ..., N$ be fixed. Then there exists $C > 0$ and $c > 0$ such that*

$$|\Delta_4^h(t_i,x,y)| \le 2Ch\|b\|_\infty t_i^{-\frac{1}{2}} g_c(t_i, y-x).$$

Remark 2.1. *The desired bound on Δ_1^h in Lemma 2.4 will be obtained for $t_i \le T$ small enough up to some normalization constant. However, by applying Kolmogorov-Chapman representation and applying the same procedure at each small time step, we will obtain the result for any $t_i \le T$.*

To conclude the proof of Theorem 2.1, we combine Lemmas 2.4, 2.7, 2.5, 2.6 and Remark 2.1.

2.3 Technical results

In this section we provide proofs of the results stated in the previous subsection.

2.3.1 Smoothness properties for the scheme

Proof of Lemma 2.4. First, let us show that e_j^h is bounded for any $j = 1, ..., N$. For fixed $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ note that

$$\begin{aligned}|(p^h - p)(t_j,x,y)| &\le p^h(t_j,x,y) + p(t_j,x,y)\\ &\le C_1 g_{c_1}(t_j,x,y) + C_2 g_{c_2}(t_j,x,y)\\ &\le C g_c(t_j,x,y),\end{aligned}$$

where second inequality follows from Proposition 2.8 and (11), and $c = \max\{c_1, c_2\}$, $C \ge C_1, C_2$. So for any $j = 1, ..., N$

$$\frac{|(p^h - p)(t_j,x,y)|}{g_c(t_j,x,y)} \le C,$$

for some $C > 0$. Now let $i = 1, ..., N$ be fixed. Using Lemma 2.2 and boundness of e_j^h, we get

$$\begin{aligned}|\Delta_1^h(t_i,x,y)| &= |\int_h^{t_i}\int_{\mathbb{R}^d}(p(\phi(s),x,z) - p^h(\phi(s),x,z))b(z)\nabla g(t-s,y-z)dzds|\\ &\le \|b\|_\infty \int_h^{t_i}\int_{\mathbb{R}^d}\frac{|p(\phi(s),x,z) - p^h(\phi(s),x,z)|}{g_c(\phi(s),z-x)} g_c(\phi(s),z-x)|\nabla g(t-s,z-y)|dzds\\ &\le \|b\|_\infty \max_j e_j^h \int_h^{t_i}\int_{\mathbb{R}^d} g_c(\phi(s),z-x)|\nabla g(t-s,z-y)|dzds\\ &\le C\|b\|_\infty \max_j e_j^h \int_h^{t_i}\int_{\mathbb{R}^d} g_c(\phi(s),z-x) g_c(t_i-\phi(s),z-y)(t_i-s)^{-1/2}dzds\\ &= C\|b\|_\infty \max_j e_j^h g_c(t_i,y-x)\int_h^{t_i}(t_i-s)^{-1/2}ds\\ &= C\|b\|_\infty g_c(t_i,y-x)\max_j e_j^h (t_i-h)^{1/2}\\ &\le C\|b\|_\infty t_i^{1/2} g_c(t_i,y-x)\max_j e_j^h.\end{aligned}$$

∎

Proof of Lemma 2.7.

$$
\begin{aligned}
|\Delta_4^h(t_i,x,y)| &= |\int_0^h \big(b(x)\mathbb{E}[\nabla_y g(t_i-s,y-X_s)] - \mathbb{E}[b(X_s)\nabla_y g(t_i-s,y-X_s^h)]\big)ds| \\
&\le \|b\|_\infty \int_0^h \mathbb{E}[|\nabla_y g(t_i-s,y-X_s) - \nabla_y g(t_i-s,y-X_s^h)|]ds \\
&\le \|b\|_\infty \int_0^h \int |p(s,x,z)-p^h(s,x,z)||\nabla_y g(t_i-s,y-z)|dzds \\
&\le 2C\|b\|_\infty \int_0^h \int g_c(s,x-z)g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds \\
&= 2C\|b\|_\infty g_c(t_i,x-y)\int_0^h (t_i-s)^{-1/2}ds \\
&\le 2Ch\|b\|_\infty t_i^{-\frac{1}{2}} g_c(t_i,x-y).
\end{aligned}
$$

□

Terms Δ_2^h and Δ_3^h are a bit trickier. We first deal with Δ_2^h. For proving a sharp estimate we rewrite $p^h(s,x,z)-p(\phi(s),x,z)$ as a sum of terms where each of them allow to maximize the advantage of regularity of underlying functions on different time intervals. For fixed $s\in[h,T]$, we separate integrals on the intervals $(0,s/2]$, $(s/2,\phi(s)]$ and $(\phi(s),s]$:

$$
\begin{aligned}
p^h(s,x,z)-p(\phi(s),x,z) =& g(s,z-x)-g(\phi(s),z-x) - \int_0^s\int p(r,x,w)b(w)\cdot\nabla_z g(s-r,z-w)drdw \\
&+\int_0^{\phi(s)}\int p(r,x,w)b(w)\cdot\nabla_z g(\phi(s)-r,z-w)drdw \\
=& g(s,z-x)-g(\phi(s),z-x) - \int_0^{s/2}\int p(r,x,w)b(w)\cdot\nabla_z g(s-r,z-w)drdw \\
&-\int_{s/2}^{\phi(s)}\int p(r,x,w)b(w)\cdot\nabla_z g(s-r,z-w)drdw \\
&-\int_{\phi(s)}^{s}\int p(r,x,w)b(w)\cdot\nabla_z g(s-r,z-w)drdw \\
&+\int_0^{s/2}\int p(r,x,w)b(w)\cdot\nabla_z g(\phi(s)-r,z-w)drdw \\
&+\int_{s/2}^{\phi(s)}\int p(r,x,w)b(w)\cdot\nabla_z g(\phi(s)-r,z-w)drdw.
\end{aligned}
$$

On the interval $(0,s/2]$ we simply have integral

$$
\int_0^{s/2}\int p(r,x,w)b(w)(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))dwdr.
$$

On the intervals $(s/2,\phi(s)]$ and $(\phi(s),s]$ we use cancellation argument. More precisely, using the fact that g_c is a density and $\int\nabla_z g(t-s,z-w)dw = \nabla_z\int g(t-s,z-w)dw = \nabla_z 1 = 0$, we add or subtract for free terms of type

$$
\int_a^b\int p(r,x,z)b(z)\nabla_z g(t-s,z-w)drdw,
$$

for a suitable time interval $(a,b]$. We obtain,

$$
\begin{aligned}
&-\int_{\phi(s)}^{s}\int p(r,x,w)b(w)\cdot\nabla_z g(s-r,z-w)drdw \\
&\quad= -\int_{\phi(s)}^{s}\int p(r,x,w)b(w)\cdot\nabla_z g(s-r,z-w)drdw + \int_{\phi(s)}^{s}\int p(r,x,z)b(z)\cdot\nabla_z g(s-r,z-w)drdw \\
&\quad= \int_{\phi(s)}^{s}\int (p(r,x,z)b(z)-p(r,x,w)b(w))\cdot\nabla_z g(s-r,z-w)drdw,
\end{aligned}
$$

and

$$\int_{s/2}^{\phi(s)}\int p(r,x,w)b(w)\cdot\nabla_z g(\phi(s)-r,z-w)drdw-\int_{s/2}^{\phi(s)}\int p(r,x,w)b(w)\cdot\nabla_z g(s-r,z-w)drdw$$
$$=\int_{s/2}^{\phi(s)}\int p(r,x,w)b(w)\cdot(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))drdw$$
$$-\int_{s/2}^{\phi(s)}\int p(r,x,z)b(z)\cdot(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))drdw$$
$$=\int_{s/2}^{\phi(s)}\int (p(r,x,w)b(w)-p(r,x,z)b(z))\cdot(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))drdw$$

Combining all integrals together, we finally obtain

$$\begin{aligned}
p(s,x,z)-p(\phi(s),x,z)&=g(s,z-x)-g(\phi(s),z-x)\\
&+\int_{\phi(s)}^{s}\int (p(r,x,z)b(z)-p(r,x,w)b(w))\nabla_z g(s-r,z-w)dwdr\\
&+\int_0^{s/2}\int p(r,x,w)b(w)(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))dwdr\\
&+\int_{s/2}^{\phi(s)}\int (p(r,x,w)b(w)-p(r,x,z)b(z))(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))dwdr\\
&=:\sum_{l=1}^{4}\mathcal{T}_l(s,x,z).
\end{aligned}\tag{24}$$

While analyzing each of the terms separately, we will make some $|x-w|^\beta$ appear in the estimates. This will be possible by exploiting the smoothness of b and the smoothness in the forward variable of the density through the cancellation argument on the given time intervals. The factor $|x-w|^\beta$ will allow us to absorb the time singularity thanks to the properties of the heat kernel.

We now estimate each term separately.

Lemma 2.8. *Let $s\in(2h,T]$, $(x,z)\in\mathbb{R}^d\times\mathbb{R}^d$. Then there exists $C>0$ such that*

$$\sum_{l=1}^{3}|\mathcal{T}_l(s,x,z)|\le Cg_c(s,z-x)h^{\frac{1+\beta}{2}}(s^{-\frac{1+\beta}{2}}+1).$$

Proof. Using fundamental theorem of calculus and Lemma 2.2, the first term of (24) can be bounded as follows:

$$\begin{aligned}
|\mathcal{T}_1(s,x,z)|&=|g(s,z-x)-g(\phi(s),z-x)|\\
&=|\int_0^1\partial_t g(t,z-x)|_{t=s-\lambda(s-\phi(s))}(s-\phi(s))d\lambda|\\
&\le C\int_0^1 g_c(s-\lambda(s-\phi(s)),z-x)\frac{1}{s-\lambda(s-\phi(s))}hd\lambda\\
&=C\int_0^1 g_c(s-\lambda(s-\phi(s)),z-x)\frac{1}{s-\lambda(s-\phi(s))}h^{\frac{1+\beta}{2}}h^{\frac{1-\beta}{2}}d\lambda\\
&\le C\int_0^1 g_c(s-\lambda(s-\phi(s)),z-x)\frac{1}{s-\lambda(s-\phi(s))}h^{\frac{1+\beta}{2}}(s-\lambda(s-\phi(s)))^{\frac{1-\beta}{2}}d\lambda\\
&\le Cg_c(s,z-x)h^{\frac{1+\beta}{2}}s^{-\frac{1+\beta}{2}}.
\end{aligned}$$

For bounding the second term in (24), we first use Lemma 2.2

$$
\begin{aligned}
|\mathcal{T}_2(s,x,z)| &= |\int_{\phi(s)}^{s}\int (p(r,x,z)b(z)-p(r,x,w)b(w))\nabla_z g(s-r,z-w)dwdr| \\
&\leq \|b\|_\infty \int_{\phi(s)}^{s}\int |p(r,x,z)-p(r,x,w)||\nabla_z g(s-r,z-w)|dwdr \\
&\quad + \int_{\phi(s)}^{s}\int p(r,x,w)|b(z)-b(w)||\nabla_z g(s-r,z-w)|dwdr \\
&\leq C\|b\|_\infty \int_{\phi(s)}^{s}\int |p(r,x,z)-p(r,x,w)|g_c(s-r,z-w)(s-r)^{-1/2}dwdr \\
&\quad + C\|b\|_\infty \int_{\phi(s)}^{s}\int p(r,x,w)|z-w|^\beta g_c(s-r,z-w)(s-r)^{-1/2}dwdr \\
&=: I_1(s,x,z)+I_2(s,x,z).
\end{aligned}
$$

The second term I_2 is bounded using (11),

$$
\begin{aligned}
I_2(s,x,z) &\leq C\|b\|_\infty \int_{\phi(s)}^{s}(s-r)^{(\beta-1)/2}\int p(r,x,w)g_c(s-r,z-w)(s-r)^{-1/2}dwdr \\
&\leq C\|b\|_\infty \int_{\phi(s)}^{s}(s-r)^{(\beta-1)/2}\int g_c(r,x-w)g_c(s-r,z-w)dwdr \\
&= C\|b\|_\infty g_c(s,x-z)\int_{\phi(s)}^{s}(s-r)^{(\beta-1)/2}dr \\
&= C\|b\|_\infty g_c(s,x-z)(s-\phi(s))^{(\beta+1)/2} \\
&\leq C\|b\|_\infty g_c(s,x-z)h^{(\beta+1)/2}.
\end{aligned}
$$

For the term $I_1(s,x,z)$ we distinguish local diagonal regime term

$$
I_{11}(s,x,z) := C\|b\|_\infty \int_{\phi(s)}^{s}\int |p(r,x,z)-p(r,x,w)|1_{|w-z|\leq r^{1/2}}g_c(s-r,z-w)(s-r)^{-1/2}dwdr,
$$

and local off-diagonal regime term

$$
I_{12}(s,x,z) := C\|b\|_\infty \int_{\phi(s)}^{s}\int |p(r,x,z)-p(r,x,w)|1_{|w-z|\geq r^{1/2}}g_c(s-r,z-w)(s-r)^{-1/2}dwdr.
$$

$I_{11}(s,x,z)$ can be bounded using (12) and Lemma 2.2:

$$
\begin{aligned}
I_{11}(s,x,z) &\leq C\|b\|_\infty \int_{\phi(s)}^{s}\int \frac{1}{r^{\beta/2}}g_c(r,x-w)|z-w|^\beta g_c(s-r,z-w)\frac{1}{(s-r)^{1/2}}dwdr \\
&= C\|b\|_\infty \int_{\phi(s)}^{s}\int \frac{1}{r^{\beta/2}}g_c(r,x-w)\frac{|z-w|^\beta}{(s-r)^{\beta/2}}g_c(s-r,z-w)\frac{1}{(s-r)^{\frac{1-\beta}{2}}}dwdr \\
&\leq C\|b\|_\infty \int_{\phi(s)}^{s}\frac{1}{r^{\beta/2}}\frac{1}{(s-r)^{\frac{1-\beta}{2}}}\int g_c(r,x-w)g_c(s-r,z-w)dwdr \\
&= C\|b\|_\infty \int_{\phi(s)}^{s}\frac{1}{r^{\beta/2}}\frac{1}{(s-r)^{\frac{1-\beta}{2}}}g_c(s,x-z)dr \\
&\leq Cg_c(s,x-z)h^{\frac{1+\beta}{2}}s^{-\frac{\beta}{2}}.
\end{aligned}
$$

For $I_{12}(s,x,z)$ we use triangular inequality and Lemma 2.2 again to arrive at:

$$
\begin{aligned}
I_{12}(s,x,z) &= C\|b\|_\infty \int_{\phi(s)}^s \int |p(r,x,z)-p(r,x,w)| 1_{|w-z|\geq r^{1/2}} g_c(s-r,z-w)(s-r)^{-1/2} dwdr \\
&\leq C \int_{\phi(s)}^s \int (p(r,x,z)+p(r,x,w)) 1_{|w-z|\geq r^{1/2}} g_c(s-r,z-w)(s-r)^{-1/2} dwdr \\
&\leq C \int_{\phi(s)}^s \int (g_c(r,z-x)+g_c(r,w-x)) \frac{|w-z|^\beta}{r^{\beta/2}} g_c(s-r,z-w)(s-r)^{-1/2} dwdr \\
&= C \int_{\phi(s)}^s \int (g_c(r,z-x)+g_c(r,w-x)) \frac{1}{r^{\beta/2}} \frac{|w-z|^\beta}{(s-r)^{\beta/2}} g_c(s-r,z-w)(s-r)^{-\frac{1-\beta}{2}} dwdr \\
&\leq C \int_{\phi(s)}^s \frac{1}{r^{\beta/2}} \frac{1}{(s-r)^{\frac{1-\beta}{2}}} \int (g_c(r,z-x)+g_c(r,w-x)) g_c(s-r,z-w) dwdr \\
&= C \int_{\phi(s)}^s \frac{1}{r^{\beta/2}} \frac{1}{(s-r)^{\frac{1-\beta}{2}}} \int g_c(r,z-x) g_c(s-r,z-w) dwdr \\
&\quad + C \int_{\phi(s)}^s \frac{1}{r^{\beta/2}} \frac{1}{(s-r)^{\frac{1-\beta}{2}}} \int g_c(r,w-x) g_c(s-r,z-w) dwdr \\
&=: I_{21}(s,x,z) + I_{22}(s,x,z).
\end{aligned}
$$

Here, the terms can be bounded as

$$
\begin{aligned}
I_{22}(s,x,z) &= C \int_{\phi(s)}^s \frac{1}{r^{\beta/2}} \frac{1}{(s-r)^{\frac{1-\beta}{2}}} g_c(s,z-x) dr \\
&\leq C g_c(s,z-x) s^{-\beta/2} \int_{\phi(s)}^s \frac{1}{(s-r)^{\frac{1-\beta}{2}}} dr \\
&\leq C g_c(s,z-x) s^{-\frac{\beta}{2}} h^{\frac{1+\beta}{2}},
\end{aligned}
$$

and

$$
\begin{aligned}
I_{21}(s,x,z) &= C \int_{\phi(s)}^s \frac{1}{r^{\beta/2}} \frac{1}{(s-r)^{\frac{1-\beta}{2}}} g_c(r,z-x) \int g_c(s-r,z-w) dwdr \\
&= C \int_{\phi(s)}^s \frac{1}{r^{\beta/2}} \frac{1}{(s-r)^{\frac{1-\beta}{2}}} g_c(r,z-x) dr \\
&\leq C g_c(s,z-x) s^{-\beta/2} \int_{\phi(s)}^s \frac{1}{(s-r)^{\frac{1-\beta}{2}}} dr \\
&\leq C g_c(s,z-x) s^{-\frac{\beta}{2}} h^{\frac{1+\beta}{2}}.
\end{aligned}
$$

Thus,

$$
I_{12}(s,x,z) \leq C g_c(s,z-x) s^{-\frac{\beta}{2}} h^{\frac{1+\beta}{2}},
$$

and

$$
|\mathcal{T}_2(s,x,z)| \leq C g_c(s,z-x) s^{-\frac{\beta}{2}} h^{\frac{1+\beta}{2}}.
$$

For bounding the third term in (24), we use similarly fundamental theorem of calculus and Lemma 2.2:

$$
\begin{aligned}
|\mathcal{T}_3(s,x,z)| &= \Big|\int_0^{s/2}\int p(r,x,w)b(w)(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))dwdr\Big| \\
&\le \|b\|_\infty \int_0^{s/2}\int p(r,x,w)|\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w)|dwdr \\
&\le C\|b\|_\infty \int_0^{s/2}\int g_c(r,x-w)|\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w)|dwdr \\
&= C\|b\|_\infty \int_0^{s/2}\int g_c(r,x-w)\Big|\int_0^1 \partial_t\nabla_z g(t,z-w)|_{t=s-\lambda(s-\phi(s))-r}(s-\phi(s))d\lambda\Big|dwdr \\
&\le C\|b\|_\infty \int_0^{s/2}\int g_c(r,x-w)\int_0^1 g_c(t,z-w)t^{-3/2}|_{t=s-\lambda(s-\phi(s))-r}(s-\phi(s))d\lambda dwdr \\
&\le C\|b\|_\infty \int_0^{s/2}(s-r)^{-3/2}\int g_c(r,x-w)g_c(s-r,z-w)hdwdr \\
&= C\|b\|_\infty h g_c(s,x-z)\int_0^{s/2}(s-r)^{-3/2}dr \\
&\le C\|b\|_\infty g_c(s,x-z)s^{-1/2}h \\
&\le C\|b\|_\infty g_c(s,x-z)h^{\frac{1+\beta}{2}}s^{-\frac{\beta}{2}}.
\end{aligned}
$$

Note that using Young's inequality for products, we can write

$$
s^{-\frac{\beta}{2}} = s^{-\frac{\beta}{2}}\cdot 1 \le C(s^{-\frac{1+\beta}{2}}+1).
$$

Thus, combining the obtained bounds we get the desired bound on the sum,

$$
\begin{aligned}
\sum_{l=1}^{3}|\mathcal{T}_l(s,x,z)| &\le C g_c(s,z-x)h^{\frac{1+\beta}{2}}(s^{-\frac{1+\beta}{2}}+s^{-\frac{\beta}{2}}) \\
&\le C g_c(s,z-x)h^{\frac{1+\beta}{2}}(s^{-\frac{1+\beta}{2}}+1).
\end{aligned}
$$

□

Lemma 2.9. *Let $s\in(0,T]$, $(x,z)\in\mathbb{R}^d\times\mathbb{R}^d$. Then for any $\epsilon>0$ there exists $C>0$ such that*

$$
|\mathcal{T}_4(s,x,z)| \le C_{\beta,\epsilon} g_c(s,z-x)s^{-\frac{\beta}{2}}h^{\frac{1+\beta}{2}-\epsilon}.
$$

Proof. The techniques to be used are similar to the ones in the previous Proposition. First, note that

$$
\begin{aligned}
|\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w)| &= \Big|\int_0^1 \partial_\lambda\nabla_z g(s-\lambda(s-\phi(s))-r,z-w)(s-\phi(s))d\lambda\Big| \\
&\le \int_0^1 |\partial_\lambda\nabla_z g(s-\lambda(s-\phi(s))-r,z-w)|(s-\phi(s))d\lambda \\
&\le C\int_0^1 g_c(s-\lambda(s-\phi(s))-r,z-w)(s-\lambda(s-\phi(s))-r)^{-3/2}(s-\phi(s))d\lambda \\
&\le C g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s)).
\end{aligned}
$$

Using this computation, we rewrite $|\mathcal{T}_4(s,x,z)|$:

$$
\begin{aligned}
&|\mathcal{T}_4(s,x,z)| = \Big|\int_{s/2}^{\phi(s)}\int (p(r,x,w)b(w)-p(r,x,z)b(z))(\nabla_z g(\phi(s)-r,z-w)-\nabla_z g(s-r,z-w))dwdr\Big| \\
&\le \|b\|_\infty\int_{s/2}^{\phi(s)}\int |p(r,x,w)-p(r,x,z)|g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s))dwdr \\
&\quad + \int_{s/2}^{\phi(s)}\int p(r,x,w)|b(w)-b(z)|g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s))dwdr \\
&=: I_1(s,x,z)+I_2(s,x,z).
\end{aligned}
$$

For I_2 we have

$$\begin{aligned}
I_2(s,x,z) &\le \|b\|_\infty \int_{s/2}^{\phi(s)} \int p(r,x,w)|w-z|^\beta g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s))dwdr \\
&\le C\|b\|_\infty h \int_{s/2}^{\phi(s)} \int p(r,x,w) g_c(s-r,z-w)(s-r)^{(\beta-3)/2} dwdr \\
&\le C\|b\|_\infty h \int_{s/2}^{\phi(s)} (s-r)^{(\beta-3)/2} \int g_c(r,x-w) g_c(s-r,z-w) dwdr \\
&\le C\|b\|_\infty g_c(s,x-z) h \int_{s/2}^{\phi(s)} (s-r)^{(\beta-3)/2} dr \\
&= C\|b\|_\infty g_c(s,x-z) h((s-\phi(s))^{(\beta-1)/2} - (s/2)^{(\beta-1)/2}) \\
&\le C\|b\|_\infty g_c(s,x-z) h^{\frac{\beta+1}{2}-\epsilon}.
\end{aligned}$$

For I_1 we distinguish local diagonal and off-diagonal regimes:

$$\begin{aligned}
I_1(s,x,z) \le{}& \|b\|_\infty \int_{s/2}^{\phi(s)} \int \frac{|z-w|^\beta}{r^{\beta/2}} g_c(r,x-w) 1_{|z-w|\le r^{1/2}} g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s))dwdr \\
&+ \|b\|_\infty \int_{s/2}^{\phi(s)} \int (g_c(r,x-w)+g_c(r,x-z)) \frac{|z-w|^\beta}{r^{\beta/2}} 1_{|z-w|\ge r^{1/2}} g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s))dwdr \\
=:{}& I_{11} + I_{12}.
\end{aligned}$$

$$\begin{aligned}
I_{11}(s,x,z) &\le C\|b\|_\infty (s-\phi(s)) \int_{s/2}^{\phi(s)} \int \frac{(s-r)^{(\beta-3)/2}}{r^{\beta/2}} g_c(r,x-w) g_c(s-r,z-w) dwdr \\
&\le C\|b\|_\infty g_c(s,x-z)(s-\phi(s)) \int_{s/2}^{\phi(s)} \frac{(s-r)^{(\beta-3)/2}}{r^{\beta/2}} dr \\
&\le C\|b\|_\infty g_c(s,x-z) s^{-\beta/2} h((s-\phi(s))^{(\beta-1)/2} - (s/2)^{(\beta-3)/2}) \\
&\le C\|b\|_\infty g_c(s,x-z) s^{-\frac{\beta}{2}} h^{\frac{\beta-1}{2}-\epsilon}.
\end{aligned}$$

$$\begin{aligned}
I_{12}(s,x,z) &\le \|b\|_\infty \int_{s/2}^{\phi(s)} \int g_c(r,x-w) \frac{|z-w|^\beta}{r^{\beta/2}} g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s))dwdr \\
&+ \|b\|_\infty \int_{s/2}^{\phi(s)} \int g_c(r,x-z) \frac{|z-w|^\beta}{r^{\beta/2}} g_c(s-r,z-w)(s-r)^{-3/2}(s-\phi(s))dwdr \\
&=: I_{121} + I_{122}.
\end{aligned}$$

I_{121} is handled as I_{11}. For I_{122} we have

$$I_{122}(s,x,z) \le \|b\|_\infty g_c(s,x-z) \int_{s/2}^{\phi(s)} \frac{|z-w|^\beta}{r^{\beta/2}} (s-r)^{-3/2}(s-\phi(s))dr,$$

and it is again handled as I_{11}.

As a result, we obtain the desired bound. □

Proof of Lemma 2.5. Using Lemmas 2.8 and 2.9 and Lemma 2.2 we get,

$$\begin{aligned}
|\Delta_2^h(t_i,x,y)| &= |\int_h^{t_i}\int_{\mathbb{R}^d}(p(s,x,z)-p(\phi(s),x,z))b(z)\nabla_z g(t_i-s,y-z)dzds| \\
&\le \|b\|_\infty \int_h^{t_i}\int_{\mathbb{R}^d}|p(s,x,z)-p(\phi(s),x,z)||\nabla_z g(t_i-s,y-z)|dzds \\
&\le C\|b\|_\infty \int_h^{t_i}\int_{\mathbb{R}^d}\sum_{l=1}^4|\mathcal{T}_l(s,x,z)|g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds \\
&\le C\|b\|_\infty \int_h^{t_i} h^{\frac{1+\beta}{2}-\epsilon}(1+h^\epsilon(s^{-\frac{1+\beta}{2}}+1))(t_i-s)^{-1/2}\int_{\mathbb{R}^d}g_c(s,z-x)g_c(t_i-s,y-z)dzds \\
&= C\|b\|_\infty g_c(t_i,y-x)h^{\frac{1+\beta}{2}-\epsilon}\int_h^{t_i}(1+h^\epsilon(s^{-\frac{1+\beta}{2}}+1))(t_i-s)^{-1/2}ds
\end{aligned}$$

$$\begin{aligned}
\int_h^{t_i}(1+h^\epsilon(s^{-\frac{1+\beta}{2}}+1))(t_i-s)^{-1/2}ds &\le \int_0^{t_i}(1+h^\epsilon(s^{-\frac{1+\beta}{2}}+1))(t_i-s)^{-1/2}ds \\
&\le Ct_i^{1/2}+\int_0^{t_i}s^{-\frac{1+\beta}{2}}(t_i-s)^{-1/2}ds \\
&\le C(t_i^{1/2}+t_i^{-\beta/2}) \\
&\le Ct_i^{-\beta/2}.
\end{aligned}$$

Thus,

$$|\Delta_2^h(t_i,x,y)| \le C\|b\|_\infty g_c(t_i,y-x)h^{\frac{1+\beta}{2}-\epsilon}t_i^{-\frac{\beta}{2}}.$$

□

Proof of Lemma 2.6. Noting that

$$\begin{aligned}
X_s^h - X_{\phi(s)}^h &= \int_{\phi(s)}^s b(X_{\phi(r)}^h)dr + (W_s - W_{\phi(s)}) \\
&\overset{\text{law}}{=} b(X_{\phi(s)}^h)(s-\phi(s)) + (s-\phi(s))^{1/2}\mathcal{N},
\end{aligned}$$

where $\mathcal{N}$ is a normally distributed random variable with mean zero and covariance matrix I_d, we can rewrite $\Delta_3^h(t_i,x,y)$ as

$$\begin{aligned}
&|\Delta_3^h(t_i,x,y)| = |\int_h^{t_i-h}\mathbb{E}[b(X_{\phi(s)}^h)\cdot(\nabla_y g(t_i-s,y-X_s^h)-\nabla_y g(t_i-s,y-X_{\phi(s)}^h))]ds \\
&\quad + \int_{t_i-h}^{t_i}\mathbb{E}[b(X_{\phi(s)}^h)\cdot(\nabla_y g(t_i-s,y-X_s^h)-\nabla_y g(t_i-s,y-X_{\phi(s)}^h))]ds| \\
&\le |\int_h^{t_i-h}\int_{\mathbb{R}^d}p^h(\phi(s),x,z)b(z)\cdot\mathbb{E}[\nabla_y g(t_i-s,y-(z+b(z)(s-\phi(s))+(s-\phi(s))^{1/2}\mathcal{N}))-\nabla_y g(t_i-s,y-z)]dzds| \\
&\quad + |\int_{t_i-h}^{t_i}\mathbb{E}[b(X_{\phi(s)}^h)\cdot(\nabla_y g(t_i-s,y-X_s^h)-\nabla_y g(t_i-s,y-X_{\phi(s)}^h))]ds| \\
&=: S_1(s,x,z)+S_2(s,x,z).
\end{aligned}$$

Let us first deal with the term S_1, which is the integral on the time interval $[h,t_i-h]$. For a fixed $s\in[h,t_i-h]$, we write the expectation in the integral as

$$\begin{aligned}
&\mathbb{E}[\nabla_y g(t_i-s,y-(z+b(z)(s-\phi(s))+(s-\phi(s))^{1/2}\mathcal{N}))-\nabla_y g(t_i-s,y-z)] \\
&\qquad = \mathbb{E}[\nabla_y g(t_i-s,y-z+(s-\phi(s))^{1/2}\mathcal{N})-\nabla_y g(t_i-s,y-z)]+R(z,s),
\end{aligned}$$

where for any $\epsilon > 0$,

$$
\begin{aligned}
R(z,s) &= \mathbb{E}[\nabla_y g(t_i - s, y - (z + b(z)(s-\phi(s)) + (s-\phi(s))^{1/2}\mathcal{N})) - \nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N})] \\
&\le C\|b\|_\infty (s-\phi(s))^{1-\epsilon}(t_i - s)^{\frac{\epsilon}{2}-1} g_c(t_i - s, y - z) \\
&\le C\|b\|_\infty h^{1-\epsilon}(t_i - s)^{\frac{\epsilon}{2}-1} g_c(t_i - s, y - z) \\
&= O(h).
\end{aligned}
$$

Moreover, notice that

$$
\begin{aligned}
&|\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)]| \\
&\qquad = |\mathbb{E}[\int_0^1 \nabla_y^2 g(t_i - s, y - z + \lambda(s-\phi(s))^{1/2}\mathcal{N})(s-\phi(s))^{1/2}\mathcal{N} d\lambda]| \\
&\qquad = |\mathbb{E}[\int_0^1 (\nabla_y^2 g(t_i - s, y - z + \lambda(s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y^2 g_c(t_i - s, y - z))(s-\phi(s))^{1/2}\mathcal{N} d\lambda]| \\
&\qquad \le \int_0^1 \mathbb{E}[|\nabla_y^2 g(t_i - s, y - z + \lambda(s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y^2 g_c(t_i - s, y - z)||s-\phi(s)|^{1/2}\mathcal{N}|] d\lambda \\
&\qquad \le \int_0^1 (t_i - s)^{-1-\frac{\beta-\epsilon}{2}}(s-\phi(s))^{\frac{1+\beta-\epsilon}{2}}\mathbb{E}[|\mathcal{N}|^{1+\beta-\epsilon}] g_c(t_i - s, y - z) d\lambda \\
&\qquad \le (t_i - s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i - s, y - z).
\end{aligned}
$$

Then, S_1 can be bounded as

$$
\begin{aligned}
S_1(s,x,z) &= O(h) + |\int_h^{t_i - h}\int_{\mathbb{R}^d} p^h(\phi(s),x,z)b(z)\cdot\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)] dz ds| \\
&\le O(h) + |\int_h^{t_i/2}\int_{\mathbb{R}^d} p^h(\phi(s),x,z)b(z)\cdot\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)] dz ds| \\
&\quad + |\int_{t_i/2}^{t_i - h}\int_{\mathbb{R}^d} p^h(\phi(s),x,z)b(z)\cdot\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)] dz ds| \\
&= O(h) + \int_h^{t_i/2}\int_{\mathbb{R}^d} |p^h(\phi(s),x,z)b(z)|\cdot|\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)]| dz ds \\
&\quad + \int_{t_i/2}^{t_i - h}\int_{\mathbb{R}^d} |p^h(\phi(s),x,z)b(z) - p^h(\phi(s),x,y)b(y)| \\
&\quad \cdot |\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)]| dz ds \\
&\le O(h) + \|b\|_\infty \int_h^{t_i/2}\int_{\mathbb{R}^d} p^h(\phi(s),x,z)|\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)]| dz ds \\
&\quad + \|b\|_\infty \int_{t_i/2}^{t_i - h}\int_{\mathbb{R}^d} |p^h(\phi(s),x,z) - p^h(\phi(s),x,y)|(t_i - s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i - s, y - z) dz ds \\
&\quad + \|b\|_\infty \int_{t_i/2}^{t_i - h}\int_{\mathbb{R}^d} |y - z|^\beta p^h(\phi(s),x,z)(t_i - s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i - s, y - z) dz ds \\
&=: O(h) + I_1(s,x,z) + I_2(s,x,z) + I_3(s,x,z).
\end{aligned}
$$

For the first term I_1 using the bound

$$
\begin{aligned}
|\mathbb{E}[\nabla_y g(t_i - s, y - z + (s-\phi(s))^{1/2}\mathcal{N}) - \nabla_y g(t_i - s, y - z)]| &\le (s-\phi(s))^{1-\epsilon}|\nabla_y^3 g(t_i - s, y - z)| \\
&\le C(s-\phi(s))^{1-\epsilon}(t_i - s)^{-3/2} g_s(t_i - s, y - z),
\end{aligned}
$$

we have

$$\begin{aligned}
I_1(s,x,z) &\le \|b\|_\infty \int_h^{t_i/2} \int_{\mathbb{R}^d} p^h(\phi(s),x,z)(s-\phi(s))^{1-\epsilon}(t_i-s)^{-3/2} g_c(t_i-s,y-z)dzds \\
&\le \|b\|_\infty \int_h^{t_i/2} (s-\phi(s))^{1-\epsilon}(t_i-s)^{-3/2} \int_{\mathbb{R}^d} g_c(\phi(s),x-z) g_c(t_i-s,y-z)dzds \\
&\le \|b\|_\infty \int_h^{t_i/2} (s-\phi(s))^{1-\epsilon}(t_i-s)^{-3/2} \int_{\mathbb{R}^d} g_c(\phi(s),x-z) g_c(t_i-\phi(s),y-z)dzds \\
&\le \|b\|_\infty g_c(t_i,x-y)(s-\phi(s))^{1-\epsilon} \int_h^{t_i/2} (t_i-s)^{-3/2} ds \\
&\le \|b\|_\infty g_c(t_i,x-y)(s-\phi(s))^{1-\epsilon} \big((t_i/2)^{-1/2} - (t_i-h)^{-1/2}\big) \\
&\le \|b\|_\infty g_c(t_i,x-y) h^{1-\epsilon} t_i^{-1/2} \\
&\le \|b\|_\infty g_c(t_i,x-y) h^{\frac{\beta+1}{2}-\epsilon} t_i^{-\frac{\beta}{2}}.
\end{aligned}$$

For the second term I_2 as usual, we distinguish diagonal and off-diagonal regimes. The diagonal regime is handled as follows:

$$\begin{aligned}
I_{21}(s,x,z) &= \|b\|_\infty \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} |p^h(\phi(s),x,z) - p^h(\phi(s),x,y)| 1_{|z-y|\le \phi(s)^{1/2}} (t_i-s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&\le C\|b\|_\infty \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} |z-y|^\beta \phi(s)^{-\beta/2} g_c(\phi(s),x-z)(t_i-s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&\le C\|b\|_\infty \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} \phi(s)^{-\beta/2} g_c(\phi(s),x-z)(t_i-s)^{-1+\frac{\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&\le C\|b\|_\infty \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} s^{-\beta/2} g_c(s,x-z)(t_i-s)^{-1+\frac{\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&= C\|b\|_\infty g_c(t_i,x-y) h^{\frac{1+\beta-\epsilon}{2}} \int_{t_i/2}^{t_i-h} s^{-\beta/2}(t_i-s)^{-1+\frac{\epsilon}{2}} ds \\
&\le C\|b\|_\infty g_c(t_i,x-y) h^{\frac{1+\beta-\epsilon}{2}} t_i^{-\beta/2}.
\end{aligned}$$

For the off-diagonal regime we have,

$$\begin{aligned}
I_{22}(s,x,z) &= \|b\|_\infty \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} |p^h(\phi(s),x,z) - p^h(\phi(s),x,y)| 1_{|z-y|\ge \phi(s)^{1/2}} (t_i-s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&\le C\|b\|_\infty \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} (g_c(\phi(s),x-z) + g_c(\phi(s),x-y)) \frac{|z-y|^\beta}{\phi(s)^{\beta/2}} (t_i-s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&=: I_{221}(s,x,z) + I_{222}(s,x,z) \\
&= I_{21}(s,x,z) + I_{222}(s,x,z).
\end{aligned}$$

$$\begin{aligned}
I_{222}(s,x,z) &\le C\|b\|_\infty g_c(t_i,x-y) \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} \frac{|z-y|^\beta}{\phi(s)^{\beta/2}} (t_i-s)^{-1-\frac{\beta-\epsilon}{2}} h^{\frac{1+\beta-\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&\le C\|b\|_\infty g_c(t_i,x-y) h^{\frac{1+\beta-\epsilon}{2}} \int_{t_i/2}^{t_i-h} \int_{\mathbb{R}^d} \frac{1}{\phi(s)^{\beta/2}} (t_i-s)^{-1+\frac{\epsilon}{2}} g_c(t_i-s,y-z)dzds \\
&= C\|b\|_\infty g_c(t_i,x-y) h^{\frac{1+\beta-\epsilon}{2}} \int_{t_i/2}^{t_i-h} \frac{1}{\phi(s)^{\beta/2}} (t_i-s)^{-1+\frac{\epsilon}{2}} ds \\
&\le C\|b\|_\infty g_c(t_i,x-y) h^{\frac{1+\beta-\epsilon}{2}} \int_{t_i/2}^{t_i-h} s^{-\beta/2} (t_i-s)^{-1+\frac{\epsilon}{2}} ds \\
&\le C\|b\|_\infty g_c(t_i,x-y) h^{\frac{1+\beta-\epsilon}{2}} t_i^{-\beta/2}.
\end{aligned}$$

Finaly, for the third term I_3 we have,

$$\begin{aligned}I_3(s,x,z) &\le C\|b\|_\infty h^{\frac{1+\beta-\epsilon}{2}}\int_{t_i/2}^{t_i-h}(t_i-s)^{-1+\frac{\epsilon}{2}}\int_{\mathbb{R}^d} g_c(t_i-s,y-z)dzds\\ &= C\|b\|_\infty h^{\frac{1+\beta-\epsilon}{2}}\int_{t_i/2}^{t_i-h}(t_i-s)^{-1+\frac{\epsilon}{2}}ds\\ &\le C\|b\|_\infty h^{\frac{1+\beta-\epsilon}{2}} t_i^{\frac{\epsilon}{2}}\end{aligned}$$

Let us now deal with the term S_2 which is the last time step integral.

$$\begin{aligned}S_2(s,x,z) &= |\int_{t_i-h}^{t_i}\mathbb{E}[(b(X^h_{\phi(s)})-p^h(\phi(s),x,y)b(y))\cdot(\nabla_y g(t_i-s,y-X^h_s)-\nabla_y g(t_i-s,y-X^h_{\phi(s)}))]ds|\\ &\le \int_{t_i-h}^{t_i}\mathbb{E}[|b(X^h_{\phi(s)})-p^h(\phi(s),x,y)b(y)||\nabla_y g(t_i-s,y-X^h_s)|]ds\\ &\quad+\int_{t_i-h}^{t_i}\mathbb{E}[|b(X^h_{\phi(s)})-p^h(\phi(s),x,y)b(y)||\nabla_y g(t_i-s,y-X^h_{\phi(s)})|]ds\\ &\le \int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)b(z)-p^h(\phi(s),x,y)b(y)|\mathbb{E}[|\nabla_y g(t_i-s,y-X^h_s)|]dzds\\ &\quad+\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)b(z)-p^h(\phi(s),x,y)b(y)||\nabla_y g(t_i-s,y-z)|dzds\\ &\le \int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)b(z)-p^h(\phi(s),x,y)b(y)|\\ &\qquad\cdot\mathbb{E}[|\nabla_y g(t_i-s,y-(z+b(z)(s-\phi(s))+N(s-\phi(s))^{\frac{1}{2}}))|]dzds\\ &\quad+\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)b(z)-p^h(\phi(s),x,y)b(y)||\nabla_y g(t_i-s,y-z)|dzds\\ &\le C\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)b(z)-p^h(\phi(s),x,y)b(y)|\\ &\qquad\cdot\mathbb{E}[g_c(t_i-s,y-(z+b(z)(s-\phi(s))+N(s-\phi(s))^{\frac{1}{2}}))](t_i-s)^{-1/2}dzds\\ &\quad+C\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)b(z)-p^h(\phi(s),x,y)b(y)|g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds.\end{aligned}$$

From Proposition 2.6 we have,

$$\mathbb{E}[|g(t_i-s,y-(z+b(z)(s-\phi(s))+N(s-\phi(s))^{\frac{1}{2}}))|]\le Cg_c(t_i-s,y-z),$$

and so

$$S_2(s,x,z)\le C\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)b(z)-p^h(\phi(s),x,y)b(y)|g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds.$$

Then, proceeding as before we get,

$$\begin{aligned}S_2(s,x,z)&\le C\|b\|_\infty\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}|p^h(\phi(s),x,z)-p^h(\phi(s),x,y)|g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds\\ &\quad+C\|b\|_\infty\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d}p^h(\phi(s),x,z)|z-y|^\beta g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds.\end{aligned}$$

Considering diagonal and off-diagonal regimes in the first integral as before, we obtain

$$
\begin{aligned}
S_2(s,x,z) \leq{}& C\|b\|_\infty \int_{t_i-h}^{t_i}\int_{\mathbb{R}^d} \phi(s)^{-\beta/2}|z-y|^\beta g_c(\phi(s),x-z)g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds \\
&+ C\|b\|_\infty \int_{t_i-h}^{t_i}\int_{\mathbb{R}^d} \phi(s)^{-\beta/2}|z-y|^\beta g_c(\phi(s),x-y)g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds \\
&+ C\|b\|_\infty \int_{t_i-h}^{t_i}\int_{\mathbb{R}^d} g_c(s,x-z)|z-y|^\beta g_c(t_i-s,y-z)(t_i-s)^{-1/2}dzds \\
\leq{}& C\|b\|_\infty \int_{t_i-h}^{t_i}\int_{\mathbb{R}^d} \phi(s)^{-\beta/2} g_c(s,x-z)g_c(t_i-s,y-z)(t_i-s)^{(\beta-1)/2}dzds \\
&+ C\|b\|_\infty g_c(t_i,x-y)\int_{t_i-h}^{t_i}\int_{\mathbb{R}^d} \phi(s)^{-\beta/2} g_c(t_i-s,y-z)(t_i-s)^{(\beta-1)/2}dzds \\
&+ C\|b\|_\infty \int_{t_i-h}^{t_i}\int_{\mathbb{R}^d} g_c(s,x-z)g_c(t_i-s,y-z)(t_i-s)^{(\beta-1)/2}dzds \\
\leq{}& C\|b\|_\infty g_c(t_i,x-y)\Big(\int_{t_i-h}^{t_i}\phi(s)^{-\beta/2}(t_i-s)^{(\beta-1)/2}ds\Big) \\
&+ C\|b\|_\infty g_c(t_i,x-y)\int_{t_i-h}^{t_i}(t_i-s)^{(\beta-1)/2}ds \\
\leq{}& C\|b\|_\infty g_c(t_i,x-y)h^{\frac{\beta+1}{2}}(t_{i-1}^{-\frac{\beta}{2}}+1).
\end{aligned}
$$

□

2.3.2 Well-posedness of the PDE

Recall that we are interested in establishing well-posedness of the PDE given by

$$
\begin{cases}
(\partial_t + \frac{1}{2}\Delta)v(t,x) = b(x) - b(x)\cdot\nabla v(t,x), & (t,x)\in[0,T)\times\mathbb{R}^d, \\
v(T,\cdot) = 0,
\end{cases} \tag{25}
$$

by looking for a fixed point in the space $\mathcal{A} := C_b^{1+\frac{\beta-\gamma}{2},2+\beta-\gamma}([0,T]\times\mathbb{R}^d)$, for $\gamma\in(0,\beta)$ small enough, for a map given by

$$
\Phi(v)(t,x) := \int_t^T\int_{\mathbb{R}^d} g(s-t,x-y)(b(y)-b(y)\cdot\nabla v(s,y))dyds,
$$

$\mathcal{A}$ is a space of bounded functions such that their time derivative is bounded and $\frac{\beta-\gamma}{2}$-Hölder, their first order space derivative is bounded and second order space derivative is bounded and $\beta-\gamma$-Hölder. Precisely, this space is equipped with the norm

$$
\begin{aligned}
\|f\|_{\mathcal{A}} ={}& \|f\|_{L^\infty([0,T]\times\mathbb{R}^d)} + \sup_{t\in[0,T]}\|D_x f(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} \\
&+ \sup_{x\in\mathbb{R}^d}\|\partial_t f(\cdot,x)\|_{C^{\frac{\beta-\gamma}{2}}([0,T])} + \sup_{t\in[0,T]}\|D_x^2 f(t,\cdot)\|_{C^{\beta-\gamma}(\mathbb{R}^d)},
\end{aligned} \tag{26}
$$

where $\|f\|_{C^\alpha(\mathcal{R})} = \|f\|_{L^\infty(\mathcal{R})} + [f]_\alpha := \|f\|_{L^\infty(\mathcal{R})} + \sup_{\substack{x,x'\in\mathcal{R}\\ x\neq x'}}\frac{|f(x)-f(x')|}{|x-x'|^\alpha}$ for $\alpha\in(0,1)$ and appropriate space $\mathcal{R}$.

We have to show that the map Φ maps $\mathcal{A}$ into itself and that it is a contraction. Let us first show that $\|\Phi(v)\|_{\mathcal{A}} \leq K\|v\|_{\mathcal{A}}$ for some $K>0$ and any $v\in\mathcal{A}$. For this we estimate each term of norm (26) of $\Phi(v)$ separately. For simplifying the writing in what comes next, for $(s,x)\in[0,T]\times\mathbb{R}^d$ denote

$$
\phi(s,x) := b(x) - b(x)\cdot\nabla v(s,x).
$$

Note that the following bound holds:

$$
\begin{aligned}
|\phi(s,x)| &\leq \|b\|_\infty(1+\|\nabla v\|_\infty) \\
&\leq \|b\|_\infty(1+\|v\|_{\mathcal{A}}),
\end{aligned}
$$

and for any $y \in \mathbb{R}^d$,

$$\begin{aligned}|\phi(s,y)-\phi(s,x)| &= |b(y)(1-\nabla v(s,y)) - b(x)(1-\nabla v(s,x))| \\ &\le \|1-\nabla v(s,\cdot)\|_\infty |b(y)-b(x)| \\ &\le \|b\|_\infty \|1-\nabla v(s,\cdot)\|_\infty |y-x|^\beta.\end{aligned}$$

For the term $\|\Phi(v)\|_{L^\infty([0,T]\times\mathbb{R}^d)}$ we easily obtain the following bound:

$$\begin{aligned}\|\Phi(v)\|_{L^\infty([0,T]\times\mathbb{R}^d)} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\int_t^T \int_{\mathbb{R}^d} g(s-t,x-y)\phi(s,y)dyds| \\ &\le \|b\|_\infty(1+\|v\|_{\mathcal{A}}) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T \int_{\mathbb{R}^d} g(s-t,x-y)dyds \\ &\le \|b\|_\infty(1+\|v\|_{\mathcal{A}}) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T ds \\ &= T\|b\|_\infty(1+\|v\|_{\mathcal{A}}).\end{aligned}$$

For bounding the other terms, we rely on the cancellation technique, i.e. using the fact that the terms

$$\int_{\mathbb{R}^d} \partial_t g(s-t,x-y)\phi(s,x)dy = 0$$

and

$$\int_{\mathbb{R}^d} \nabla_x^2 g(s-t,x-y)\phi(s,x)dy = 0$$

can be introduced for free.

We write $\sup_{x\in\mathbb{R}^d} \|\partial_t \Phi(v)(\cdot,x)\|_{L^\infty([0,T])}$ as

$$\sup_{x\in\mathbb{R}^d} \|\partial_t \Phi(v)(\cdot,x)\|_{L^\infty([0,T])} = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\partial_t \int_t^T \int_{\mathbb{R}^d} g(s-t,x-y)\phi(s,y)dyds|,$$

and

$$\begin{aligned}\partial_t \int_t^T \int_{\mathbb{R}^d} g(s-t,x-y)\phi(s,y)dyds &= \lim_{\epsilon\to 0} \partial_t \int_{t+\epsilon}^T \int_{\mathbb{R}^d} g(s-t,x-y)\phi(s,y)dyds \\ &= \lim_{\epsilon\to 0} \int_{\mathbb{R}^d} g(\epsilon,x-y)\phi(t+\epsilon,y)dy + \int_t^T \int_{\mathbb{R}^d} \partial_t g(s-t,x-y)\phi(s,y)dyds \\ &= \phi(t,x) + \int_t^T \int_{\mathbb{R}^d} \partial_t g(s-t,x-y)\phi(s,y)dyds \\ &=: \phi(t,x)+\psi(t,x).\end{aligned} \tag{27}$$

For the term $|\psi(t)|$ using Lemma 2.2 we have,

$$\begin{aligned}|\int_t^T \int_{\mathbb{R}^d} \partial_t g(s-t,x-y)\phi(s,y)dyds| &= |\int_t^T \int_{\mathbb{R}^d} \partial_t g(s-t,x-y)(\phi(s,y)-\phi(s,x))dyds| \\ &\le \int_t^T \int_{\mathbb{R}^d} |\partial_t g(s-t,x-y)||\phi(s,y)-\phi(s,x)|dyds \\ &\le C\|b\|_\infty(1+\|v\|_{\mathcal{A}}) \int_t^T \int_{\mathbb{R}^d} g_c(s-t,x-y)(s-t)^{-1}|x-y|^\beta dyds \\ &\le C\|b\|_\infty(1+\|v\|_{\mathcal{A}}) \int_t^T \int_{\mathbb{R}^d} g_c(s-t,x-y)(s-t)^{\beta/2-1} dyds \\ &= C\|b\|_\infty(1+\|v\|_{\mathcal{A}})(T-t)^{\beta/2},\end{aligned}$$

thus

$$\sup_{x\in\mathbb{R}^d} \|\partial_t \Phi(v)(\cdot,x)\|_{L^\infty([0,T])} \le \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} (|\phi(t,x)| + |\psi(t,x)|)$$
$$\le C\|b\|_\infty(1+\|v\|_{\mathcal{A}})(1+T^{\beta/2}).$$

For bounding the term $\sup_{t\in[0,T]} \|D_x\Phi(v)(t,\cdot)\|_{L^\infty(\mathbb{R}^d)}$ we use the same approach as before together with Lemma 2.2:

$$\begin{aligned}
\sup_{t\in[0,T]} \|D_x\Phi(v)(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\int_t^T \int_{\mathbb{R}^d} \nabla_x g(s-t,x-y)\phi(s,y)dyds| \\
&\le \|b\|_\infty(1+\|v\|_{\mathcal{A}}) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T \int_{\mathbb{R}^d} |\nabla_x g(s-t,x-y)|dyds \\
&\le C\|b\|_\infty(1+\|v\|_{\mathcal{A}}) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T \int_{\mathbb{R}^d} g_c(s-t,x-y)(s-t)^{-1/2}dyds \\
&\le C\|b\|_\infty(1+\|v\|_{\mathcal{A}}) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T (s-t)^{-1/2}ds \\
&= CT^{1/2}\|b\|_\infty(1+\|v\|_{\mathcal{A}}).
\end{aligned}$$

Following cancellation technique approach together again with Lemma 2.2, we obtain:

$$\begin{aligned}
\sup_{t\in[0,T]} \|D_x^2\Phi(v)(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\int_t^T \int_{\mathbb{R}^d} \nabla_x^2 g(s-t,x-y)\phi(s,y)dyds| \\
&= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\int_t^T \int_{\mathbb{R}^d} \nabla_x^2 g(s-t,x-y)(\phi(s,y)-\phi(s,x))dyds| \\
&\le C \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T \int_{\mathbb{R}^d} \|1-\nabla v(s,\cdot)\|_{\mathcal{A}} \frac{g_c(s-t,x-y)}{s-t}|y-x|^\beta dyds \\
&\le C\|b\|_\infty(1+\|v\|_\infty) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T \int_{\mathbb{R}^d} g_c(s-t,x-y)\frac{|y-x|^\beta}{(s-t)^{\beta/2}}(s-t)^{\beta/2-1}dyds \\
&\le C\|b\|_\infty(1+\|v\|_{\mathcal{A}}) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T \int_{\mathbb{R}^d} g_c(s-t,x-y)(s-t)^{\beta/2-1}dyds \\
&= C\|b\|_\infty(1+\|v\|_{\mathcal{A}}) \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_t^T (s-t)^{\beta/2-1}ds \\
&\le CT^{\frac{\beta}{2}}\|b\|_\infty(1+\|v\|_{\mathcal{A}}).
\end{aligned}$$

Next, we proceed to the Hölder modulus estimates which are the most delicate. The supremum in time of the Hölder modulus of $D_x^2 v(t,x)$ writes as

$$\begin{aligned}
\sup_{t\in[0,T]} [D_x^2\Phi(v)(t,\cdot)]_{\beta-\gamma} &= \sup_{\substack{(t,x,x')\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d \\ x\neq x'}} \frac{|D_x^2\Phi(v)(t,x)-D_x^2\Phi(v)(t,x')|}{|x-x'|^{\beta-\gamma}} \\
&= \sup_{\substack{(t,x,x')\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d \\ x\neq x'}} \frac{|\int_t^T \int_{\mathbb{R}^d} (\nabla_x^2 g(s-t,x-y)-\nabla_x^2 g(s-t,x'-y))\phi(s,y)dyds|}{|x-x'|^{\beta-\gamma}}
\end{aligned}$$

In this case, spatial regularity will be derived from global time smallness. To this end, we consider diagonal ($|x-x'| \le (T-t)^{1/2}$) and off-diagonal ($|x-x'| \ge (T-t)^{1/2}$) regimes. In the off-diagonal regime, using cancellation technique and triangle inequality, we write

$$
\begin{aligned}
&|\int_t^T \int_{\mathbb{R}^d} (\nabla_x^2 g(s-t,x-y) - \nabla_x^2 g(s-t,x'-y))\phi(s,y) 1_{|x-x'|\geq (T-t)^{1/2}} dyds| \\
&\quad = |\int_t^T \int_{\mathbb{R}^d} \nabla_x^2 g(s-t,x-y)(\phi(s,y)-\phi(s,x)) 1_{|x-x'|\geq (T-t)^{1/2}} dyds \\
&\quad\quad - \int_t^T \int_{\mathbb{R}^d} \nabla_x^2 g(s-t,x'-y)(\phi(s,y)-\phi(s,x')) 1_{|x-x'|\geq (T-t)^{1/2}} dyds| \\
&\quad \leq \sum_{z\in\{x,x'\}} \int_t^T \int_{\mathbb{R}^d} |\nabla_x^2 g(s-t,z-y)||\phi(s,y)-\phi(s,z)| 1_{|x-x'|\geq (T-t)^{1/2}} dyds.
\end{aligned}
$$

For $z \in \{x,x'\}$, using Lemma 2.2,

$$
\begin{aligned}
&\frac{1}{|x-x'|^{\beta-\gamma}} \int_t^T \int_{\mathbb{R}^d} |\nabla_x^2 g(s-t,z-y)||\phi(s,y)-\phi(s,z)| 1_{|x-x'|\geq (T-t)^{1/2}} dyds \\
&\quad \leq C\|b\|_\infty (1+\|v\|_{\mathcal{A}}) \int_t^T \int_{\mathbb{R}^d} \frac{g_c(s-t,z-y)}{(s-t)|x-x'|^{\beta-\gamma}} |y-z|^\beta 1_{|x-x'|\geq (T-t)^{1/2}} dyds \\
&\quad \leq C\|b\|_\infty (1+\|v\|_{\mathcal{A}}) \int_t^T \int_{\mathbb{R}^d} \frac{g_c(s-t,z-y)}{(s-t)(s-t)^{\frac{\beta-\gamma}{2}}} |y-z|^\beta dyds \\
&\quad \leq C\|b\|_\infty (1+\|v\|_{\mathcal{A}}) \int_t^T \int_{\mathbb{R}^d} g_c(s-t,z-y)(s-t)^{\gamma/2-1} dyds \\
&\quad \leq C\|b\|_\infty (1+\|v\|_{\mathcal{A}})(T-t)^{\gamma/2}.
\end{aligned}
$$

In the diagonal regime, we consider again two cases: local diagonal ($|x-x'| \leq (s-t)^{1/2}$) and local off-diagonal ($|x-x'| \geq (s-t)^{1/2}$) regimes (for $s \in [t,T]$). In the local off-diagonal we proceed similarly to the global diagonal regime

$$
\begin{aligned}
&|\int_t^{t+|x-x'|^2} \int_{\mathbb{R}^d} (\nabla_x^2 g(s-t,x-y) - \nabla_x^2 g(s-t,x'-y))\phi(s,y) dyds| \\
&\quad \leq \sum_{z\in\{x,x'\}} \int_t^{t+|x-x'|^2} \int_{\mathbb{R}^d} |\nabla_x^2 g(s-t,z-y)||\phi(s,y)-\phi(s,z)| dyds.
\end{aligned}
$$

Again for $z \in \{x,x'\}$ we have in the local off-diagonal regime,

$$
\begin{aligned}
&\int_t^{t+|x-x'|^2} \int_{\mathbb{R}^d} |\nabla_x^2 g(s-t,z-y)||\phi(s,y)-\phi(s,z)| dyds \\
&\quad \leq C\|b\|_\infty (1+\|v\|_{\mathcal{A}}) \int_t^{t+|x-x'|^2} \int_{\mathbb{R}^d} g_c(s-t,z-y) \frac{|z-y|^\beta}{s-t} dyds \\
&\quad \leq C\|b\|_\infty (1+\|v\|_{\mathcal{A}}) \int_t^{t+|x-x'|^2} \int_{\mathbb{R}^d} g_c(s-t,z-y)(s-t)^{\beta/2-1} dyds \\
&\quad = C\|b\|_\infty (1+\|v\|_{\mathcal{A}})(s-t)^{\beta/2}\Big|_t^{t+|x-x'|^2} \\
&\quad = C\|b\|_\infty (1+\|v\|_{\mathcal{A}})|x-x'|^\beta.
\end{aligned}
$$

In the local diagonal regime, we use fundamental theorem of calculus, cancellation technique and Lemma 2.2

to obtain the following bound:

$$
\begin{aligned}
&|\int_{t+|x-x'|^2}^{T}\int_{\mathbb{R}^d}(\nabla_x^2 g(s-t,x-y)-\nabla_x^2 g(s-t,x'-y))\phi(s,y)dyds|\\
&=|\int_{t+|x-x'|^2}^{T}\int_{\mathbb{R}^d}\int_0^1 \nabla_x^3 g(s-t,x'+\lambda(x-x')-y)(x-x')(\phi(s,y)-\phi(s,x'+\lambda(x-x')))d\lambda dyds|\\
&\leq \int_{t+|x-x'|^2}^{T}\int_{\mathbb{R}^d}\int_0^1 |\nabla_x^3 g(s-t,x'+\lambda(x-x')-y)||x-x'||\phi(s,y)-\phi(s,x'+\lambda(x-x'))|d\lambda dyds\\
&\leq C\|b\|_\infty(1+\|v\|_{\mathcal{A}})\int_{t+|x-x'|^2}^{T}\int_{\mathbb{R}^d}\int_0^1 g_c(s-t,x'+\lambda(x-x')-y)\frac{|x-x'|}{(s-t)^{3/2}}|y-(x'+\lambda(x-x'))|^\beta d\lambda dyds\\
&\leq C\|b\|_\infty(1+\|v\|_{\mathcal{A}})\int_{t+|x-x'|^2}^{T}\int_{\mathbb{R}^d}\int_0^1 g_c(s-t,x'+\lambda(x-x')-y)\frac{|x-x'|}{(s-t)^{\frac{3-\beta}{2}}}d\lambda dyds\\
&= C\|b\|_\infty(1+\|v\|_{\mathcal{A}})|x-x'|\int_{t+|x-x'|^2}^{T}(s-t)^{\frac{\beta-3}{2}}ds\\
&= -C\|b\|_\infty(1+\|v\|_{\mathcal{A}})|x-x'|(s-t)^{\frac{\beta-1}{2}}\Big|_{t+|x-x'|^2}^{T}\\
&= C\|b\|_\infty(1+\|v\|_{\mathcal{A}})(|x-x'|^\beta-|x-x'|(T-t)^{\frac{\beta-1}{2}})\\
&\leq C\|b\|_\infty(1+\|v\|_{\mathcal{A}})|x-x'|^\beta.
\end{aligned}
$$

Combining the obtained bounds, we get

$$
\begin{aligned}
\sup_{t\in[0,T]}[D_x^2\Phi(v)(t,\cdot)]_{\beta-\gamma} &\leq C\|b\|_\infty(1+\|v\|_{\mathcal{A}})(T^{\gamma/2}+\sup_{\substack{(t,x,x')\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\\ x\neq x'}}\frac{|x-x'|^\beta}{|x-x'|^{\beta-\gamma}}1_{|x-x'|\leq(T-t)^{1/2}})\\
&\leq C\|b\|_\infty(1+\|v\|_{\mathcal{A}})(T^{\gamma/2}+\sup_{t\in[0,T]}(T-t)^{\gamma/2})\\
&= C\|b\|_\infty(1+\|v\|_{\mathcal{A}})T^{\gamma/2}.
\end{aligned}
$$

The supremum in space of the Hölder modulus of $D_x^2 v(t,x)$ writes as

$$
\sup_{x\in\mathbb{R}^d}[\partial_t\Phi(v)(\cdot,x)]_{\frac{\beta-\gamma}{2}}=\sup_{\substack{(t,t',x)\in[0,T]\times[0,T]\times\mathbb{R}^d\\ t\neq t'}}\frac{|\partial_t\Phi(v)(t,x)-\partial_{t'}\Phi(v)(t',x)|}{|t-t'|^{\frac{\beta-\gamma}{2}}}
$$

For $r\in\{t,t'\}$, proceeding as in (27) we have

$$
\partial_r\int_r^T\int_{\mathbb{R}^d}g(s-r,x-y)(b(y)-b(y)\cdot\nabla v(s,y))dyds=\phi(r,x)+\psi(r,x),
$$

and

$$
\begin{aligned}
|\phi(t,x)-\phi(t',x)| &\leq \|b\|_\infty|\nabla v(t,x)-\nabla v(t',x)|\\
&\leq \|b\|_\infty\|v\|_{\mathcal{A}}|t-t'|^{1/2+(\beta-\gamma)/2}.
\end{aligned}
$$

Without loss of generality, assume that $t < t'$. We divide the integral ψ into the intervals $[t, t']$ and $[t', T]$:

$$\begin{aligned}
|\psi(t,x) - \psi(t',x)| &= |\int_t^{t'} \int_{\mathbb{R}^d} \partial_t g(s-t, x-y)\phi(s,y)dyds \\
&\quad - \int_{t'}^{T} \int_{\mathbb{R}^d} (\partial_{t'} g(s-t', x-y) - \partial_t g(s-t, x-y))\phi(s,y)dyds)| \\
&\leq |\int_t^{t'} \int_{\mathbb{R}^d} \partial_t g(s-t, x-y)(\phi(s,y) - \phi(s,x))dyds| \\
&\quad + |\int_{t'}^{2t'-t} \int_{\mathbb{R}^d} (\partial_{t'} g(s-t', x-y) - \partial_t g(s-t, x-y))(\phi(s,y) - \phi(s,x))dyds| \\
&\quad + |\int_{2t'-t}^{T} \int_{\mathbb{R}^d} (\partial_{t'} g(s-t', x-y) - \partial_t g(s-t, x-y))(\phi(s,y) - \phi(s,x))dyds| \\
&=: T_1(x) + T_2(x) + T_3(x).
\end{aligned}$$

For the term T_1 we simply use the Gaussian bound from Lemma 2.2:

$$\begin{aligned}
T_1(x) &\leq C\|b\|_\infty(1 + \|v\|_{\mathcal{A}}) \int_t^{t'} \int_{\mathbb{R}^d} g_c(s-t, x-y)(s-t)^{-1}|x-y|^\beta dyds \\
&\leq C\|b\|_\infty(1 + \|v\|_{\mathcal{A}}) \int_t^{t'} \int_{\mathbb{R}^d} g_c(s-t, x-y)(s-t)^{\beta/2-1} dyds \\
&\leq C\|b\|_\infty(1 + \|v\|_{\mathcal{A}})|t'-t|^{\beta/2}.
\end{aligned}$$

For the term T_2, we first use triangular inequality,

$$T_2(x) \leq \int_{t'}^{2t'-t} \int_{\mathbb{R}^d} \big(|\partial_{t'} g(s-t', x-y)| + |\partial_t g(s-t, x-y)|\big)|\phi(s,y) - \phi(s,x)|dyds$$

Using Lemma 2.2 and the fact that $\frac{t'-t}{s-t} < \frac{t'-t}{s-t'}$, for $\theta \in (0, \beta/2)$ we have that:

$$\begin{aligned}
|\partial_{t'} g(s-t', x-y)| &\leq C\frac{1}{s-t'} g_c(s-t', x-y) \\
&\leq C\frac{|t'-t|^\theta}{(s-t')^{1+\theta}} g_c(s-t', x-y),
\end{aligned}$$

and

$$|\partial_t g(s-t, x-y)| \leq C\frac{|t'-t|^\theta}{(s-t')^{1+\theta}} g_c(s-t, x-y).$$

Thus,

$$\begin{aligned}
T_2(x) &\leq C\|b\|_\infty(1 + \|v\|_{\mathcal{A}}) \int_{t'}^{2t'-t} \int_{\mathbb{R}^d} (g_c(s-t', x-y) + g_c(s-t, x-y))\frac{|t'-t|^\theta}{(s-t')^{1+\theta}}|x-y|^\beta dyds \\
&\leq C\|b\|_\infty(1 + \|v\|_{\mathcal{A}}) \int_{t'}^{2t'-t} \int_{\mathbb{R}^d} (g_c(s-t', x-y) + g_c(s-t, x-y))\frac{|t'-t|^\theta}{(s-t')^{1-\beta/2+\theta}} dyds \\
&\leq C\|b\|_\infty(1 + \|v\|_{\mathcal{A}}) \int_{t'}^{2t'-t} \frac{|t'-t|^\theta}{(s-t')^{1-\beta/2+\theta}} ds \\
&\leq C\|b\|_\infty(1 + \|v\|_{\mathcal{A}})|t'-t|^\theta |t'-t|^{\beta/2+\theta} \\
&= C\|b\|_\infty(1 + \|v\|_{\mathcal{A}})|t'-t|^{\beta/2}.
\end{aligned}$$

For the third term T_3 we use Taylor expansion and obtain the following bounds:

$$\begin{aligned}
T_3(x) &= \int_{2t'-t}^{T}\int_{\mathbb{R}^d}|\int_0^1 \partial_r^2 g(s-\lambda(t'-t)-t,x-y)||t'-t||\phi(s,y)-\phi(s,x)|d\lambda dyds\\
&\le \int_{2t'-t}^{T}\int_{\mathbb{R}^d}\int_0^1 g_c(s-\lambda(t'-t)-t,x-y)\frac{|t'-t|^\theta}{(s-\lambda(t'-t)-t)^{1+\theta}}|x-y|^\beta d\lambda dyds\\
&\le C\int_{2t'-t}^{T}\int_{\mathbb{R}^d}\int_0^1 g_c(s-\lambda(t'-t)-t,x-y)\frac{|t'-t|^\theta}{(s-\lambda(t'-t)-t)^{1+\theta-\beta/2}}d\lambda dyds\\
&= \int_{2t'-t}^{T}\int_0^1 \frac{|t'-t|^\theta}{(s-\lambda(t'-t)-t)^{1+\theta-\beta/2}}d\lambda ds\\
&\le \int_{2t'-t}^{T}\frac{|t'-t|^\theta}{(s-t')^{1+\theta-\beta/2}}ds\\
&\le |t'-t|^\theta((T-t')^{\beta/2-\theta}-|t'-t|^{\beta/2-\theta})\\
&\le |t'-t|^\theta T^{\beta/2-\theta}.
\end{aligned}$$

Finally, taking $\theta=\frac{\beta-\gamma}{2}$,

$$\begin{aligned}
\sup_{x\in[0,T]}[\partial_t\Phi(v)(\cdot,x)]_{\frac{\beta-\gamma}{2}} &\le \sup_{\substack{(t,t',x)\in[0,T]\times[0,T]\times\mathbb{R}^d\\ t\neq t'}}\frac{|\phi(t,x)-\phi(t',x)|+|\psi(t,x)-\psi(t',x)|}{|t'-t|^{\frac{\beta-\gamma}{2}}}\\
&\le C\|b\|_\infty \sup_{\substack{(t,t',x)\in[0,T]\times[0,T]\times\mathbb{R}^d\\ t\neq t'}}\frac{\|v\|_{\mathcal{A}}|t-t'|^{1/2+(\beta-\gamma)/2}+(1+\|v\|_{\mathcal{A}})|t'-t|^{(\beta-\gamma)/2}T^{\gamma/2}}{|t'-t|^{\frac{\beta-\gamma}{2}}}\\
&\le C\|b\|_\infty \sup_{\substack{(t,t',x)\in[0,T]\times[0,T]\times\mathbb{R}^d\\ t\neq t'}}(\|v\|_{\mathcal{A}}|t-t'|^{\frac12}+(1+\|v\|_{\mathcal{A}})T^{\frac{\gamma}{2}})\\
&\le C\|b\|_\infty(\|v\|_{\mathcal{A}}T^{\frac12}+2(1+\|v\|_{\mathcal{A}})T^{\frac{\gamma}{2}}).
\end{aligned}$$

Combining all the previous norms bounds, we obtain that when T is small enough, Φ is an endomorphism.

Now, let $v_0=0$ and $v_{n+1}:=\Phi(v_n)$ for $n\ge 0$. We aim at showing that Φ is a contraction, i.e. that there exists $k<1$ such that $\|\Phi(v_{n+1})-\Phi(v_n)\|_{\mathcal{A}}\le k\|v_{n+1}-v_n\|_{\mathcal{A}}$. As before, for $(s,x)\in[0,T]\times\mathbb{R}^d$ we introduce

$$\phi_n(s,x):=b(x)-b(x)\cdot\nabla v_n(s,x),$$

and note that the following bounds hold:

$$\begin{aligned}
|\phi_{n+1}(s,x)-\phi_n(s,x)| &= |b(x)\cdot(\nabla v_{n+1}(s,x)-\nabla v_n(s,x))|\\
&\le \|b\|_\infty\|\nabla(v_{n+1}-v_n)\|_\infty\\
&\le \|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}},
\end{aligned}\tag{28}$$

and

$$\begin{aligned}
|\phi_{n+1}(s,y)-\phi_{n+1}(s,x)-\phi_n(s,y)+\phi_n(s,x)| &\le |\phi_{n+1}(s,y)-\phi_n(s,y)|+|\phi_{n+1}(s,x)-\phi_n(s,x)|\\
&\le 2\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}.
\end{aligned}\tag{29}$$

Using same strategy as for $\|\Phi\|_{L^\infty([0,T]\times\mathbb{R}^d)}$ and using (28), we get the bound for $\|\Phi(v_{n+1})-\Phi(v_n)\|_{L^\infty([0,T]\times\mathbb{R}^d)}$ given by

$$\begin{aligned}
\|\Phi(v_{n+1})-\Phi(v_n)\|_{L^\infty([0,T]\times\mathbb{R}^d)} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|\int_t^T\int_{\mathbb{R}^d} g(s-t,x-y)(\phi_{n+1}(s,y)-\phi_n(s,y))dyds|\\
&\le T\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}.
\end{aligned}$$

For bounding $\|\partial_t(\Phi(v_{n+1})-\Phi(v_n))(\cdot,x)\|_{L^\infty([0,T])}$, we introduce, similarly to $\phi(t)$,

$$\psi_n(t,x):=\int_t^T\int_{\mathbb{R}^d}\partial_t g(s-t,x-y)\phi_n(s,y)dyds,$$

increments of which satisfies the bound

$$|\psi_{n+1}(t,x)-\psi_n(t,x)| \le C\|b\|_\infty \|v_{n+1}-v_n\|_{\mathcal{A}}(T-t)^{\beta/2}.$$

Thus,

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d}\|\partial_t(\Phi(v_{n+1})-\Phi(v_n))(\cdot,x)\|_{L^\infty([0,T])} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\partial_t \int_t^T \int_{\mathbb{R}^d} g(s-t,x-y)(\phi_{n+1}(s,y)-\phi_n(s,y))dyds| \\
&\le \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} (|\phi_{n+1}(t,x)-\phi_n(t,x)|+|\psi_{n+1}(t,x)-\psi_n(t,x)|) \\
&\le C\|b\|_\infty\|\nabla v_{n+1}-\nabla v_n\|_\infty + C\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}T^{\beta/2} \\
&\le C\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}T^{1/2} + C\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}T^{\beta/2}.
\end{aligned}$$

The quantity $\|D_x(\Phi(v_{n+1})-\Phi(v_n))(t,\cdot)\|_{L^\infty(\mathbb{R}^d)}$ readily satisfies the following bound:

$$\begin{aligned}
\sup_{t\in[0,T]}\|D_x(\Phi(v_{n+1})-\Phi(v_n))(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\int_t^T \int_{\mathbb{R}^d} \nabla_x g(s-t,x-y)(\phi_{n+1}(s,y)-\phi_n(s,y))dyds| \\
&\le CT^{1/2}\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}.
\end{aligned}$$

And similarly for $\|D_x^2(\Phi(v_{n+1})-\Phi(v_n))(t,\cdot)\|_{L^\infty(\mathbb{R}^d)}$:

$$\begin{aligned}
\sup_{t\in[0,T]}\|D_x^2(\Phi(v_{n+1})-\Phi(v_n))(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\int_t^T \int_{\mathbb{R}^d} \nabla_x^2 g(s-t,x-y)(\phi_{n+1}(s,y)-\phi_n(s,y))dyds| \\
&\le C\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}T^{\beta/2}.
\end{aligned}$$

Finally,

$$\begin{aligned}
&\sup_{t\in[0,T]}[D_x^2(\Phi(v_{n+1})-\Phi(v_n))(t,\cdot)]_{\beta-\gamma} \\
&\quad= \sup_{\substack{(t,x,x')\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\\ x\neq x'}} \frac{|\int_t^T\int_{\mathbb{R}^d}(\nabla_x^2 g(s-t,x-y)-\nabla_x^2 g(s-t,x'-y))(\phi_{n+1}(s,y)-\phi_n(s,y))dyds|}{|x-x'|^{\beta-\gamma}} \\
&\quad\le C\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}T^{\gamma/2},
\end{aligned}$$

and

$$\begin{aligned}
\sup_{x\in[0,T]}[\partial_t(\Phi(v_{n+1})-\Phi(v_n))(\cdot,x)]_{\frac{\beta-\gamma}{2}} &= \sup_{\substack{(t,t',x)\in[0,T]\times[0,T]\times\mathbb{R}^d\\ t\neq t'}} \frac{|\partial_t\Phi(v)(t,x)-\partial_{t'}\Phi(v)(t',x)|}{|t-t'|^{\frac{\beta-\gamma}{2}}} \\
&= \sup_{\substack{(t,t',x)\in[0,T]\times[0,T]\times\mathbb{R}^d\\ t\neq t'}} \Big(|\partial_t\int_t^T\int_{\mathbb{R}^d} g(s-t,x-y)(\phi_{n+1}(s,y)-\phi_n(s,y))dyds \\
&\qquad - \partial_{t'}\int_{t'}^T\int_{\mathbb{R}^d} g(s-t',x-y)(\phi_{n+1}(s,y)-\phi_n(s,y))dyds||t-t'|^{-\frac{\beta-\gamma}{2}}\Big) \\
&\le C\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}(T^{\gamma/2}+T^{1/2}).
\end{aligned}$$

Putting all parts of the norm $\|\cdot\|_{\mathcal{A}}$ together, we arrive at

$$\begin{aligned}
\|\Phi(v_{n+1})-\Phi(v_n)\|_{\mathcal{A}} &\le C\|b\|_\infty\|v_{n+1}-v_n\|_{\mathcal{A}}(T+T^{\beta/2}+T^{1/2}+T^{\gamma/2}) \\
&:= k\|v_{n+1}-v_n\|_{\mathcal{A}},
\end{aligned}$$

for $T>0$ such that $C\|b\|_\infty(T+T^{\beta/2}+T^{1/2}+T^{\gamma/2})<1$. Thus Φ is a contraction map and by fixed point theorem, there exists a unique $v^*\in\mathcal{A}$ such that $\Phi(v^*)=v^*$, and $v^*=\lim_{n\to\infty}v_n$.

3 Stable processes: definition and properties

This section aims at defining α-stable processes and describing their useful properties. First, recall that *Lévy process* $(X_t)_{t\geq 0}$ on $\mathbb{R}^d$ is an *additive process* that has stationary increments and is continuous in probability, i.e. it satisfies:

(1) $X_0 = 0$ a.s.;

(2) For any $t_0 < t_1 < ... < t_n$, $X_{t_0}, X_{t_1} - X_{t_0}, ..., X_{t_n} - X_{t_{n-1}}$ are i.i.d.;

(3) For any $s, t > 0$, $X_{s+t} - X_s \overset{\text{law}}{=} X_t$;

(4) For any $\epsilon > 0$ and $t \geq 0$, $\lim_{s\to t} \mathbb{P}(|X_s - X_t| > \epsilon) = 0$.

For any $t \geq 0$, the law of a Lévy process $\mu^t = \text{Law}(X_t)$ is *infinitely divisible*, i.e. for any $n > 0$, there exists $\mu_n^t \in \mathcal{P}(\mathbb{R}^d)$ s.t. $\mu^t = \mu_n^t * ... * \mu_n^t$. Thanks to this property, Lévy processes enjoy *Lévy-Khintchine representation*, which uniquely characterize the process by its characteristic function $\bar{\mu}(z) := \int_{\mathbb{R}^d} e^{i\langle z,x\rangle}\mu(dx)$.

Theorem 3.1 (Theorem 8.1, [Sat99])**.** *If μ is an infinitely divisible distribution on $\mathbb{R}^d$, then there exist unique symmetric nonnegative-definite $d \times d$ matrix A, vector $\gamma \in \mathbb{R}^d$ and measure ν on $\mathbb{R}^d$ satisfying*

$$\nu(\{0\}) = 0, \qquad \int_{\mathbb{R}^d} (|x|^2 \wedge 1)\nu(dx) < \infty, \tag{30}$$

such that

$$\hat{\mu}(z) = \exp\Big(-\frac{1}{2}\langle z, Az\rangle + i\langle \gamma, z\rangle + \int_{\mathbb{R}^d} (e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle 1_{|x|\leq 1}(x))\nu(dx)\Big). \tag{31}$$

Conversely, if A is a symmetric nonnegative-definite $d \times d$ matrix, $\gamma \in \mathbb{R}^d$, and ν is a measure ν on $\mathbb{R}^d$ satisfying (30)*, then there exists an infinitely divisible distribution μ whose characteristic function is given by* (31)*.*

We call the triple (A, γ, ν) the *generating triplet* of μ. In particular, if $(X_t)_{t\geq 0}$ is a Lévy process, then for any $t \geq 0$, $(A_t, \gamma_t, \nu_t) := (tA, t\gamma, t\nu)$ is the generating triplet of μ^t. The matrix A is called the *Gaussian covariance matrix*, the measure ν is the *Lévy measure* of μ.

Example 3.1.
- *If $\nu = 0$, then μ is a Gaussian law;*
- *If $A = 0$, then μ is called purely non-Gaussian;*
- *If $A = 0$, $\nu = 0$, $\gamma = a \in \mathbb{R}^d$, then $\mu = \delta_a$.*

Generally speaking, stable processes are Lévy processes that enjoy *self-similarity property*: change of time scale has the same effect as change of spatial scale and addition of a linear motion. First, they can be defined through the characteristic function.

Definition 3.1. *An infinitely divisible distribution $\mu \in \mathcal{P}(\mathbb{R}^d)$ is* stable *if for any $a > 0$, there exists $b > 0$ and $c \in \mathbb{R}^d$ s.t.*

$$\hat{\mu}^a(z) = \hat{\mu}(bz)e^{i\langle c,x\rangle}.$$

We say that a Lévy process is stable *if its distribution at $t = 1$ is stable.*

Definition 3.2. *A stochastic process $(X_t)_{t\geq 0}$ is called* broad-sense self-similar *if for any $a > 0$, there exists $b > 0$ and $c(t) : [0, \infty) \to \mathbb{R}^d$ s.t.*

$$(X_{at})_{t\geq 0} \overset{\text{law}}{=} (bX_t + c(t))_{t\geq 0}. \tag{32}$$

By Proposition 13.5 in [Sat99], we have that if $(X_t)_{t\geq 0}$ is a Lévy process on $\mathbb{R}^d$ then it is stable if and only if it is broad-sense selfsimilar. Moreover, by Theorem 13.11 in [Sat99], there exists $H > 0$ s.t. for any $a > 0$ for which there exists $b, c(t)$ satisfying (32), such that $b = a^H$. The index H is called the *self-similarity index* of the process, and $\alpha = \frac{1}{H}$ is called the *stability index* of the process. By Theorem 13.15 in [Sat99], it holds that $\alpha \in (0, 2]$. Thus, we say that a stable process with the index α is an *α-stable process*. Note that $\alpha = 2$ if and only if μ is Gaussian (Theorem 14.1, [Sat99]), and $(X_t)_{t\geq 0}$ is called a *pure jump* process if $\alpha \in (1, 2)$.

Theorem 3.2 (Theorem 14.3, [Sat99])**.** *An infinitely divisible distribution $\mu \neq \delta_0$ with the generating triplet (A, γ, ν) is α-stable with $\alpha \in (0,2)$ if and only if $A = 0$ and there exists a finite measure λ on the unit sphere $S := \{x \in \mathbb{R}^d : |x| = 1\}$ s.t. for any $B \in \mathcal{B}(\mathbb{R}^d)$,*

$$\nu(B) = \int_S \lambda(d\xi) \int_0^\infty 1_B(r\xi) \frac{dr}{r^{1+\alpha}}.$$

The measure λ on S is uniquely determined by μ and is called a spherical part *of the Lévy measure ν, and $r^{-1-\alpha}dr$ is called a* radial part *of ν.*

Observe that, as α decreases, $r^{-1-\alpha}$ get smaller for $r \in (0,1)$ and bigger for $r \in (1,\infty)$. This can be understood as an α-stable process moves mostly by big jumps if α is close to 0, and mostly by small jumps if α is close to 2.

Another powerful result about Lévy processes is *Lévy-Itô decomposition* which allows us to write the process as a sum of a purely jump process and continuous process.

Theorem 3.3 (Lévy-Itô decomposition, Theorem 19.2, [Sat99])**.** *Let $(X_t)_{t\geq 0}$ be an additive process on $\mathbb{R}^d$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the generating triplet $(A_t, \gamma_t, \nu_t)_{t\geq 0}$. For $B \in \mathcal{B}((0,\infty) \times (\mathbb{R}^d \backslash \{0\}))$, define $\Omega_0 := \{\omega \in \Omega : \lim_{t\downarrow 0} \mathbb{P}(|X_t| > \epsilon) = 0,\ \forall \epsilon > 0\}$ and*

$$N(B, \omega) := \begin{cases} \#\{s : (s, X_s(\omega) - X_{s-}(\omega)) \in B\}, & \omega \in \Omega_0, \\ 0, & \omega \notin \Omega_0. \end{cases}$$

Denote $D(a,b] := \{x \in \mathbb{R}^d : a < |x| \leq b\}$, $D(a,b) := \{x \in \mathbb{R}^d : a < |x| < b\}$ for $a, b \in [0,\infty)$. Then the following hold:

(i) $\{N(B) : B \in \mathcal{B}((0,\infty) \times (\mathbb{R}^d \backslash \{0\}))\}$ is a Poisson random measure on $(0,\infty) \times (\mathbb{R}^d \backslash \{0\})$ with intensity measure ν;

(ii) $\exists \Omega_1 \in \mathcal{F}$, $\mathbb{P}(\Omega_1) = 1$, s.t. for any $\omega \in \Omega_1$,

$$\begin{aligned} N_t(\omega) &= \lim_{\epsilon\downarrow 0} \int_{(0,t]\times D(\epsilon,1]} \big(xN(d(s,x),\omega) - x\nu_t(d(s,x))\big) + \int_{(0,t]\times D(1,\infty)} xN(d(s,x),\omega) \\ &=: \lim_{\epsilon\downarrow 0} \int_{(0,t]\times D(\epsilon,1]} x\tilde{N}(d(s,x),\omega) + \int_{(0,t]\times D(1,\infty)} xN(d(s,x),\omega) \end{aligned}$$

is defined for all $t \in [0,\infty)$ and the convergence is uniform in $t \in [0,T]$, $\forall T > 0$. The process $(N_t)_{t\geq 0}$ is an additive process on $\mathbb{R}^d$ with the generating triplet $(0, 0, \nu_t)_{t\geq 0}$;

(iii) Define $M_t(\omega) = X_t(\omega) - N_t(\omega)$, $\omega \in \Omega_1$. There exists $\Omega_2 \in \mathcal{F}$, $\mathbb{P}(\Omega_2) = 1$, s.t. for any $\omega \in \Omega_2$, $M_t(\omega)$ is continuous. The process $(M_t)_{t\geq 0}$ is additive with the generating triplet $(A_t, \gamma_t, 0)_{t\geq 0}$;

(iv) $(N_t)_{t\geq 0}$ and $(M_t)_{t\geq 0}$ are independent.

We call $\tilde{N}$ the *compensated Poisson measure*, and the term $\lim_{\epsilon\downarrow 0} \int_{(0,t]\times D(\epsilon,1]} x\tilde{N}(d(s,x),\omega)$ the *compensated sum of jumps.*

Finally, we refresh some definitions and state related results that are important for proving main results of the next sections. We say that $\rho \in \mathcal{P}(\mathbb{R}^d)$ is *symmetric* if for any $B \in \mathcal{B}(\mathbb{R}^d)$, $\rho(B) = \rho(-B)$. We say that $\rho \in \mathcal{P}(\mathbb{R}^d)$ is *rotation invariant* if for any $B \in \mathcal{B}(\mathbb{R}^d)$ and any orthogonal matrix U, $\rho(B) = \rho(U^{-1}B)$, where $U^{-1}B = \{U^{-1}x, x \in B\}$.

Theorem 3.4 (Theorem 14.13, [Sat99])**.** *If μ is symmetric and α-stable for $\alpha \in (0,2)$ on $\mathbb{R}^d$, then its characteristic function is given by*

$$\hat{\mu}(z) = \exp\Big(- \int_S |\langle z, \xi\rangle|^\alpha \lambda(d\xi)\Big),$$

where λ is a symmetric finite non-zero measure on S, uniquely determined by μ.

Theorem 3.5 ([Kol00]). *Let $(X_t)_{t\geq 0}$ be a symmetric α-stable process, $\alpha \in (1,2)$ with the generating triplet (A,γ,ν). If the spherical measure λ of the Lévy measure ν satisfies the condition of* uniform non-degeneracy*, i.e. if there exists $\kappa \geq 1$ s.t.*

$$\kappa^{-1}|z|^\alpha \leq \int_S |\langle z,\xi\rangle|^\alpha \lambda(d\xi) \leq \kappa|z|^\alpha, \quad \forall z \in \mathbb{R}^d,$$

then for any $t > 0$, the process $(X_t)_{t\geq 0}$ admits a continuous density $p_\alpha(t,\cdot)$.

The generator of a Lévy process $(X_t)_{t\geq 0}$ with a semigroup $P_t^\alpha f(x) := \mathbb{E}[f(X_t)|X_0 = x]$ is given by

$$\begin{aligned}\mathcal{L}^\alpha f(x) =& \frac{1}{2}\sum_{j,k=1}^d A_{jk}\frac{\partial^2 f}{\partial x_j \partial x_k}(x) + \sum_{j=1}^d \gamma_j \frac{\partial f}{\partial x_j}(x)\\ &+ \int_{\mathbb{R}^d}\big(f(x+y) - f(x) - \sum_{j=1}^d y_j \frac{\partial f}{\partial x_j}(x)1_{|x|\leq 1}(y)\big)\nu(dy),\end{aligned}$$

for $f \in C_0^2(\mathbb{R}^d)$, where (A,γ,ν) is the generating triplet of $(X_t)_{t\geq 0}$.

Example 3.2.
- *If $(X_t)_{t\geq 0}$ is a Brownian motion, then $\mathcal{L}^2 f(x) = \frac{1}{2}\Delta f(x)$;*
- *If $(X_t)_{t\geq 0}$ is a compound Poisson process, then $\mathcal{L}^\alpha f(x) = \int_{\mathbb{R}^d}\big(f(x+y) - f(x)\big)\nu(dx)$;*
- *If $(X_t)_{t\geq 0}$ is a pure jump process and its spherical measure is symmetric non-degenerate (see Theorem 3.5), then $\mathcal{L}^\alpha f(x) = p.v. \int_{\mathbb{R}^d}\big(f(x+y) - f(x)\big)\nu(dx)$.*

Proposition 3.1 (Bounds and Sensitivities for the stable density). *Let $(X_t)_{t\geq 0}$ be a symmetric α-stable process such that the conditions of Theorem 3.5 are satisfied. Then for any $t > 0$, X_t admits a density $p_\alpha(t,\cdot)$ such that there exists $C > 0$ s.t. for all $l \in \{1,2\}$, $t > 0$, $y \in \mathbb{R}^d$,*

$$|D_y^l p_\alpha(t,y)| \leq \frac{C}{t^{l/\alpha}} q_\alpha(t,y), \quad |\partial_t^l p_\alpha(t,y)| \leq \frac{C}{t^l} q_\alpha(t,y),$$

where $(q_\alpha(t,\cdot))_{t>0}$ is a family of probability measures on $\mathbb{R}^d$ s.t. $q_\alpha(t,y) = t^{-d/\alpha}q_\alpha(1,t^{-1/\alpha}y)$, $t > 0$, and for all $\gamma \in [0,\alpha)$, there exists a constant $c = c(\alpha,\eta,\gamma) > 0$ s.t.

$$\int_{\mathbb{R}^d} q_\alpha(t,y)|y|^\gamma dy \leq ct^{\gamma/\alpha}.$$

Proof. For $\alpha = 2$, see Lemma 2.2. For $\alpha \in (0,2)$, see Lemma 11, [CM22]. $\square$

4 Besov spaces: definition and properties

The classical definition of Besov spaces is based on Littlewood-Paley decomposition. For this, we consider functions $\phi \in C_0^\infty(\mathbb{R}^d,\mathbb{R})$ (i.e. real-valued smooth functions with compact support) such that $\phi(0) \neq 0$, and $\phi_j(x) = 2^j\phi(2^j x)$. By $\mathcal{F}$, $\mathcal{F}^{-1}$ we denote Fourier transform and inverse Fourier transform respectively. Finally, by $\mathcal{S}'(\mathbb{R}^d)$ we denote the space of tempered distributions. Then Besov space $B_{p,q}^\beta(\mathbb{R}^d,\mathbb{R}^k)$ with $p,q,\beta \in \mathbb{R}$, $k \geq 1$, is defined as follows:

$$B_{p,q}^\beta(\mathbb{R}^d,\mathbb{R}^k) := \{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{B_{p,q}^\beta} < +\infty\},$$

where

$$\|f\|_{B_{p,q}^\beta} := \|\mathcal{F}^{-1}(\phi\mathcal{F}(f))\|_{L^p} + \begin{cases}(\sum_{j\in\mathbb{N}} 2^{\beta jq}\|\mathcal{F}^{-1}(\phi_j\mathcal{F}(f))\|_{L^p}^q)^{1/q}, & 1 \leq q < \infty,\\ \sup_{j\in\mathbb{N}}(2^{\beta j}\|\mathcal{F}^{-1}(\phi_j\mathcal{F}(f))\|_{L^p}), & q = \infty.\end{cases}$$

However, by Section 2.5.10 in [Tri83] (see also Section 5.3 in [Lem02]), $f \in B_{p,q}^\beta$ if and only if $\|f\|_{\mathcal{H}_{p,q,\tilde{\alpha}}^\beta} < \infty$, where this quantity is defined as

$$\|f\|_{\mathcal{H}^{\beta}_{p,q,\tilde\alpha}} := \|\mathcal{F}^{-1}(\phi\mathcal{F}(f))\|_{L^p} + \mathcal{T}^{\beta}_{p,q}(f)$$
$$:= \|\mathcal{F}^{-1}(\phi\mathcal{F}(f))\|_{L^p} + \begin{cases} (\int_0^1 \frac{dv}{v} v^{(n-\beta/\tilde\alpha)q} \|\partial_v^n p^{\tilde\alpha}(v,\cdot) * f\|_{L^p}^q)^{1/q}, & 1 \le q < \infty, \\ \sup_{v\in(0,1]}(v^{(n-\beta/\tilde\alpha)}\|\partial_v^n p^{\tilde\alpha}(v,\cdot) * f\|_{L^p}), & q = \infty, \end{cases}$$

where $\tilde\alpha \in [1,2]$ and $0 \le n < \beta/\tilde\alpha$. This equivalent definition of Besov spaces is called *thermic characterization* as it involves a heat kernel. The term $\mathcal{T}^{\beta,\tilde\alpha}_{p,q}(f)$ is referred to as the thermic part, $\tilde\alpha$ is called the thermic characterization index, β corresponds to the regularity index and p is an integrability index. From now on, we take $\|\cdot\|_{B^{\beta}_{p,q}} := \|\cdot\|_{\mathcal{H}^{\beta}_{p,q,\tilde\alpha}}$.

Thermic characterization of Besov spaces is beneficial in the scope of proving well-posedness of SDEs with Lebesgue-Besov drifts thanks to the properties of convolution with stable densities. For what comes next, we choose $\tilde\alpha = \alpha$ and $n = 1$ which is enough to use cancellation argument.

We now state fundamental properties of Besov spaces.

Lemma 4.1. *(1) Continuous embedding with Schwartz spaces. For any $p, q \in [1,\infty]$, $\beta \in \mathbb{R}$,*

$$\mathcal{S}(\mathbb{R}^d) \hookrightarrow B^{\beta}_{p,q} \hookrightarrow \mathcal{S}'(\mathbb{R}^d).$$

(2) Continuous embedding with Lebesgue spaces. For any $1 \le p \le \infty$,

$$\begin{aligned} B^0_{p,1} \hookrightarrow L^p \hookrightarrow B^0_{p,\infty}, &\quad 1 \le p \le \infty, \\ B^{\beta}_{p,q} \hookrightarrow L^\infty, &\quad \beta > d/p \text{ or } \beta = d/p,\ 0 < q \le 1, \\ B^0_{\infty,q} \hookrightarrow L^\infty \hookrightarrow B^0_{\infty,\infty}, &\quad 0 < q \le 1. \end{aligned}$$

(3) Continuous embedding with space of complex-valued bounded and uniformly continuous functions on $\mathbb{R}^d$. For any $1 \le p \le \infty$,

$$\begin{aligned} B^0_{\infty,q} \hookrightarrow C \hookrightarrow B^0_{\infty,\infty}, &\quad 0 < q \le 1, \\ B^{\beta}_{p,q} \hookrightarrow C, &\quad \beta > d/p \text{ or } \beta = d/p,\ 0 < q \le 1. \end{aligned}$$

(4) Continuous embedding between Besov spaces. For any $p_0, p_1, q_0, q_1 \in [1,\infty]$ s.t. $p_0 \le p_1$, $q_0 \le q_1$ and $\beta_1 - d/p_0 \le \beta_0 - d/p_0$,

$$B^{\beta_0}_{p_0,q_0} \hookrightarrow B^{\beta_1}_{p_1,q_1}.$$

(5) Continuous embedding with probability measures space. For any $\epsilon > 0$, $p \in [1,\infty]$, $q \in [1,\infty)$,

$$\mathcal{P}(\mathbb{R}^d) \hookrightarrow \cap_{p'\ge 1} B^{-d/p'}_{p,\infty}, \quad \mathcal{P}(\mathbb{R}^d) \hookrightarrow \cap_{p'\ge 1} B^{-d/p'-\epsilon}_{p,q}.$$

(6) Convolution Young inequality. Let $p, q \in [1,\infty]$, $\beta \in \mathbb{R}$. For any $\delta \in \mathbb{R}$, $p_1, p_2 \in [1,\infty]$ s.t. $1 + p^{-1} = p_1^{-1} + p_2^{-1}$ and $q_1, q_2 \in (0,\infty]$ s.t. $q_1^{-1} \ge (q^{-1} - q_2^{-1}) \vee 0$, there exists a constant $c = c(d) > 0$ s.t.

$$\|f * g\|_{B^{\beta}_{p,q}} \le c\|f\|_{B^{\beta-\delta}_{p_1,q_1}} \|g\|_{B^{\delta}_{p_2,q_2}}$$

(7) Bound on the Besov norm of heat kernel. For any $p, q \in [1,\infty]$, $\beta \in \mathbb{R}$, there exists a constant $c = c(\alpha, p, q, \beta, d) > 0$ s.t. for any multi-index $\mathbf{a} \in \mathbb{N}^d$ with $|\mathbf{a}| \le 1$ and $0 < v < s < \infty$,

$$\|\partial^{\mathbf{a}} p^{\alpha}_{s-v}\|_{B^{\beta}_{p,q}} \le \frac{c}{(s-v)^{\frac{\beta}{\alpha} + \frac{d}{\alpha}(1-\frac{1}{l}) + \frac{|\mathbf{a}|}{\alpha}}},$$

where $\partial^{\mathbf{a}} = \frac{\partial^{|\mathbf{a}|}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}$, $|\mathbf{a}| = a_1 + \dots + a_d$.

(8) Duality inequality. For $p, q \in [1, m]$, $\beta \in \mathbb{R}$ and $f \in B^{\beta}_{p,q}$, $g \in B^{-\beta}_{p',q'}$, where p', q' are conjugates of p, q respectfully (i.e. $1/p + 1/p' = 1$, $1/q + 1/q' = 1$),

$$|\int_{\mathbb{R}^d} f(y)g(y)dy| \le \|f\|_{B^{\beta}_{p,q}} \|g\|_{B^{-\beta}_{p',q'}}.$$

Proof. (1) See Theorem 2.4 in [Saw18].

(2) See Proposition 2.1 in [Saw18], Theorem 3.3.1 in [ST95] and Theorem 3.1.1 in [ST95]

(3) See Theorem 3.1.1 in [ST95] and Section 2.8.3 in [Tri83].

(4) See Proposition 2(i), (ii) of Section 2.3.2 in [Tri83] for monotonicity in parameters q, β and Theorem 2.5 in [Saw18] for monotonicity in parameter p.

(5) See Lemma 5 in [CJM22].

(6) See Theorem 3 in [Bur90].

(7) Proof follows from Proposition 3.1.

(8) See Proposition 3.6 in [Lem02]. □

Remark 4.1. *Note that for $\beta \in \mathbb{R}$ in ((8)), one of the maps is a distribution so $\int_{\mathbb{R}^d} f(y)g(y)dy$ must be given a proper meaning. Since $C_c^\infty(\mathbb{R}^d)$ is dense in $B_{p,q}^\beta$ for any $p, q \in [1, \infty)$, $\beta \in \mathbb{R}$ (see Theorem 2.4 in [Saw18]), fg can be thought as a limit of product of some functions in $C_c^\infty(\mathbb{R}^d)$ for which the integral is well-defined.*

Remark 4.2. *For $p, q = \infty$, $\beta > 0$, Besov space $B_{\infty,\infty}^\beta$ is understood as Hölder space. For more examples of particular spaces that cover Besov space, see Section 2.3.5, [Tri83] and Section 2.1.2, [RS96].*

5 Non-degenerate stable-driven SDEs with Besov drift

In this section we make an excursion into the theory of stochastic differential equations with distributional drifts. We focus more on the Besov drifts which includes discussing the results for usual and McKean-Vlasov equations (Sections 5.2, 5.3).

When dealing with distributional drifts, one part of the job is to give a meaning to such a drift. In the modern literature, there are three main approaches. First approach consists in using *Zvonkin transform* introduced in Section 1. In this case, the integral of a drift can be understood and expressed via a smooth solution to the Zvonkin PDE:

$$\int_0^t b(X_s)ds = u(t, X_t) - u(0, x) - \int_0^t (\partial_s + \mathcal{L}^X)u(s, X_s) \cdot dZ_t.$$

For the distributional drifts, this approach was first developed in [FIR15].

Second approach, developed in [Zha10], [ABM20], consists in usual approximation technique. We consider a mollified drift b_n which is a C_b^∞ function and $b_n \xrightarrow{n\to\infty} b$ for which the corresponding SDE is strongly well-defined. Then the *formal* integral of the distributional drift is then a Dirichlet process and is understood as

$$\text{“}\int_0^t b(X_s)ds\text{”} = \lim_{n\to\infty} \int_0^t b_n(X_s)ds.$$

Third approach, the most precise one developed in [DD14], [CM22], reconstructs the distributional drift as a Young integral

$$\text{“}\int_0^t b(X_s)ds\text{”} = \int_0^t \mathcal{B}(s, X_s, ds),$$

where $\mathcal{B}(v, x, s - v) := \int_v^s \int_{\mathbb{R}^d} b(r, y)p_\alpha(r - v, y - x)dydr$ (see Section 5.2 for details).

5.1 Background for the SDEs with distributional drift

In [FIR15], authors consider a d-dimensional SDE given for $t \geq 0$ by

$$dX_t = b(t, X_t)dt + dW_t, \quad X_0 = x \in \mathbb{R}, \tag{33}$$

where $(W_t)_{t\geq 0}$ is a d-dimensional Brownian motion, and $b \in L^\infty([0, T], H_q^{-\beta} \cap H_{\tilde{q}}^{-\beta})$, where

$$H_p^s(\mathbb{R}^d) := (I - \frac{1}{2}\Delta)^{s/2}(L^p(\mathbb{R}^d))$$

is a *fractional Sobolev space* of order $s \in \mathbb{R}$ (see [Tri83]) and $\beta \in (0, \frac{1}{2})$, $q \in (\frac{d}{1-\beta}, \frac{d}{p})$, $\tilde{q} := \frac{d}{1-\beta}$. Authors prove well-posedness of (33) introducing notion of *virtual solution* which involves using Zvonkin transform. More precisely, well-posedness of the following formal parabolic PDE is investigated:

$$\begin{cases} \partial_t + \frac{1}{2}\Delta u + b \cdot \nabla u - (\lambda+1)u = -b, & \text{on } [0,T] \times \mathbb{R}^d, \\ u(T, \cdot) = 0, & \text{on } \mathbb{R}^d, \end{cases} \tag{34}$$

where $\lambda > 0$. Additional work has to be done to understand when $b \cdot \nabla u$ is well-defined due to the distributional nature of the drift, thus we call it formal. The above PDE allows us to get rid of the bad drift in (33) and write it using the solution of the PDE which is regular enough. Thus, the notion of virtual solution involves rather the solution of the PDE than the drift.

Definition 5.1. *A process $(X_t)_{t\geq 0}$ is called* virtual solution *to the SDE* (33) *if it satisfies the integral equation*

$$X_t = x + u(0,x) - u(t, X_t) + (\lambda+1)\int_0^t u(s, X_s)ds + \int_0^t \nabla u(s, X_s)dW_s + W_t,$$

for all $t \in [0,T]$, where u is the unique mild *solution to* (34).

The existence and uniqueness of the virtual solution to (33) is established by proving that $\Phi(t,x) = x + u(t,x)$ is a C^1-diffeomorphism for T small enough. The SDE

$$Y_t = \Phi(0,x) + \int_0^t D\Phi(s, \Phi^{-1}(s, Y_s))dW_s$$

has a unique weak solution thanks to the smoothness of the solution u to (34) and $X_t = \Phi^{-1}(t, Y_t)$ is then a virtual solution for which uniqueness in law holds.

In [ABM20], authors consider a time-homogeneous 1-dimensional SDE given for $t \geq 0$ by

$$dX_t = b(X_t)dt + dZ_t, \quad X_0 = x \in \mathbb{R}, \tag{35}$$

where $(Z_t)_{t\geq 0}$ is a symmetric 1-dimensional α-stable noise, and $b \in C^\beta$ with

$$\beta > \frac{1-\alpha}{2}. \tag{36}$$

This work is an extension of the Brownian motion driven SDE studied in [BC01]. The state-of-art in [ABM20] is a strong well-posedness result for the SDE (35) under condition (36). The main idea behind the proof is to use Zvonkin transform by considering a *resolvent equation* which includes too studying its well-posedness. However, unlike in the Brownian motion case, the resulting solution of the SDE may not be a semimartingale but a Dirichlet process, which is defined as follows.

Definition 5.2. *We say that an adapted process $(X_t)_{t\in[0,T]}$, $T > 0$, is a* Dirichlet process *if*

$$X_t = M_t + A_t, \quad t \in [0,T],$$

where M is a square integrable martingale and A is an adapted process of zero energy*, i.e. $A_0 = 0$ and*

$$\lim_{\delta\to 0} \sup_{\pi_T : |\pi_T| < delta} \mathbb{E}[\sum_{t_i \in \pi_T} |A_{t_{i+1}-t_i}|^2] = 0.$$

Authors prove existence and uniqueness of the virtual solution to (35) that involves the regular solution to the resolvent equation. Having this, they prove how to derive strong well-posedness from the virtual well-posedness, which is not an immediate step.

5.2 "Linear" SDEs with Besov drift driven by stable noise

Consider the *formal* SDE for a fixed $T > 0$, $t \in [0,T]$,

$$dX_t = b(t, X_t)dt + Z_t, \quad X_0 \sim \mu_0, \tag{37}$$

where $(Z_t)_{t\geq 0}$ is a d-dimensional symmetric α-stable process, $\alpha \in (1, 2]$, $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$. We say that the SDE (37) is *formal* because we are interested in the drift b that belongs to a suitable Lebesgue-Besov space, and thus, written as it is, (37) may not make sense.

In this section, we state the main results of [CM22] about well-posedness of (37) when the drift belongs to the appropriate Lebesgue-Besov. The well-posedness is studied in the sense of martingale problem and weak solutions (see below for proper definitions). Although these problems are usually equivalent when the SDE coefficients are regular enough (see e.g. [SV97]), it is not the case with the distributional drift. Having a canonical process from the martingale solution, it is not obvious how to reconstruct the dynamics and the drift in particular. The authors in [CM22] (see also [DD14]) overcome this problem and give a precise meaning to the distributional drift. In this section, we give a general idea of how this type of questions are addressed. The main well-posedness result reads as follows.

Theorem 5.1 (Theorem 3,6, [CM22])**.** *Let $p, q, r \geq 1$, $\alpha \in (1, 2]$, $\gamma \in (1/2, 1)$ and $b \in L^r([0, T], B^{-1+\gamma}_{p,q}(\mathbb{R}^d, \mathbb{R}^d))$.*

1. *Assume that γ satisfies*
$$\alpha \in \left(\frac{1 + [d/p]}{1 - [1/r]}, 2\right], \quad \gamma \in \left(\frac{3 - \alpha + [d/p] + [\alpha/r]}{2}, 1\right). \tag{38}$$
Then there exists a unique martingale solution to the martingale problem associated with the SDE (37). Moreover, the canonical process under the martingale solution is strong Markov.

2. *Assume that γ satisfies*
$$\alpha \in \left(\frac{1 + [d/p]}{1 - [1/r]}, 2\right], \quad \gamma \in \left(\frac{3 - \alpha + [2d/p] + [2\alpha/r]}{2}, 1\right). \tag{39}$$
Then there exists a unique weak solution to the formal *SDE (37). Moreover, if $d = 1$, then the solution is pathwise unique.*

Martingale solution Let us first give the intuition behind studying well-posedness of the martingale problem. Let $\mathcal{L}^X$ be a generator associated to the SDE (37), and $\mathcal{E}$ be a rich enough class of functions to be specified later on. Let also $\Omega_2 = C([0, T], \mathbb{R}^d)$ (space of continuous functions) and $\Omega_\alpha = D([0, T], \mathbb{R}^d)$ (space of càdlàg functions) when $\alpha \in (0, 2)$. We say that a probability measure $\mathbb{P}^\alpha$ on Ω_α is a *martingale solution* to the *martingale problem* associated with an operator $\mathcal{L}^X$ on the probability space $(\Omega_\alpha, \mathcal{F}, \mathbb{P})$ if

(i) $\mathbb{P}^\alpha(X_0 = 0) = 1$;

(ii) For all $\phi \in \mathcal{E}$,
$$\left(\phi(t, X_t) - \int_0^t (\partial_s + \mathcal{L}^X)\phi(s, X_s)ds\right)_{0\leq t\leq T} \tag{40}$$
is a (square integrable if $\alpha = 2$) $\mathbb{P}^\alpha$-martingale.

In our problem, we have
$$\mathcal{L}^X = b \cdot D + \mathcal{L}^\alpha,$$
where D is a generalized derivative and $\mathcal{L}^\alpha$ is a generator of the α-stable process $(Z_t)_{t\geq 0}$. In the classical setup when b is a bounded measurable function, then it is enough to take $\mathcal{E} = C_0^\infty(\mathbb{R}^d)$ (see [SV97]) to characterize a Markov process through the martingale formulation. However, recall that we consider singular drifts that are not functions but distributions. Therefore, we are facing two problems:

(1) Ensure that $\forall \phi \in \mathcal{E}$, $b \cdot D\phi$ is well-defined and is at least a distribution with some regularity properties;

(2) Ensure that the integrand in (40) is well-defined and is a function with some regularity properties.

To address these problems, we introduce the auxiliary *formal* Cauchy problem with the corresponding notion of solution. As before, here, we say that the problem is *formal* due to distributional nature of the drift b. The *formal* Cauchy problem reads as follows:

$$(\partial_t + b \cdot D + \mathcal{L}^\alpha)u(t,x) = f(t,x),\ (t,x) \in [0,T) \times \mathbb{R}^d, \quad u(T,\cdot) = g(\cdot), \tag{41}$$

where $f \in \mathcal{F}$, $g \in \mathcal{G}$, and $\mathcal{F}, \mathcal{G}$ are rich enough classes of functions. Then (ii) can be reformulated as

(ii)$'$ For all $f \in \mathcal{F}$, $g \in \mathcal{G}$,

$$\Big(u(t,X_t) - u(0,x) - \int_0^t f(s,X_s)ds\Big)_{0\le t\le T}$$

is a (square integrable if $\alpha = 2$) $\mathbb{P}^\alpha$-martingale, where u is the solution (in the sense to be specified) to the Cauchy problem (41).

We note that for obtaining the well-posed martingale problem, we do not need to work with classical solutions. Instead, we introduce *mild* solutions, the precise definition of which will get the problems (1)-(2) resolved (see Lemma 1, [CM22]).

Definition 5.3. *Let $\alpha \in (1,2]$, $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$, $g : \mathbb{R}^d \to \mathbb{R}$. For any fixed $T > 0$, $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is a* mild solution *to the formal Cauchy problem* (41) *if*

(i) $u \in C^{0,1}([0,T] \times \mathbb{R}^d, \mathbb{R})$ and $Du \in C_b^0([0,T], B^{\theta-1-\epsilon}_{\infty,\infty}(\mathbb{R}^d, \mathbb{R}^d))$, where $\epsilon > 0$ and

$$\theta := \gamma - 1 + \alpha - \frac{d}{p} - \frac{\alpha}{r}; \tag{42}$$

(ii) u satisfies the Duhamel representation

$$u(t,x) = P^\alpha_{T-}(g)(x) - \int_t^T P^\alpha_{s-t}(f - b \cdot Du)(s,x)ds.$$

Then, the following theorem about well-posedness of the Cauchy problem can be proven.

Theorem 5.2 (Theorem 2, [CM22])**.** *Let $p, q, r \geq 1$, $\alpha \in (1,2]$ and $\gamma \in (1/2,1)$ satisfy* (38)*. Then for all $f \in C([0,T], B^{\theta-\alpha}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}))$, $g \in C^1(\mathbb{R}^d,\mathbb{R})$ with $Dg \in B^{\theta-1}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d)$, where θ is given by* (42)*, the formal Cauchy problem admits a unique mild solution.*

Finally, (ii) in the martingale problem can be reformulated as

(ii)$''$ For all $f \in C([0,T], \mathcal{S}(\mathbb{R}^d,\mathbb{R}))$, $g \in C^1(\mathbb{R}^d,\mathbb{R})$ with $Dg \in B^{\theta-1}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d)$,

$$\Big(u(t,X_t) - u(0,x) - \int_0^t f(s,X_s)ds\Big)_{0\le t\le T}$$

is a (square integrable if $\alpha = 2$) $\mathbb{P}^\alpha$-martingale, where u is the mild solution to the Cauchy problem (41).

Note that thanks to the embedding property (1) of Schwartz space into Besov space and its density in Besov space (see [Tri83]), we are able to formulate (ii)$''$ for $f(s,\cdot) \in \mathcal{S}(\mathbb{R}^d,\mathbb{R})$, $\forall s \in [0,T]$. This reads as $\mathcal{E} = \mathcal{S}(\mathbb{R}^d,\mathbb{R})$ in the sense of definition (ii).

With all this in hand, one can prove first part of Theorem 5.1 with the martingale problem formulation (i)-(ii)$''$. The strategy for proving it is the following:

- Introduce the *mollified* Cauchy problem, where precisely the drift b is mollified to obtain b_m, a sequence of smooth functions, s.t. $\|b - b_m\|_{L^r([0,T],B^{-1+\gamma}_{p,q})} \xrightarrow[m\to\infty]{} 0$ (see Remark 3, [CM22]);
- Prove the Schauder type estimates on the solution u_m of the mollified Cauchy problem (see Proposition 8, 9 and Corollary 10, [CM22]). The key element for proving this step is Duhamel representation of the mollified PDE solution and thermic characterization of Besov space. Precisely this combination allows us to use heat kernel bounds (7) and other Besov space properties stated in Lemma 4.1. Note that this step does not distinguish pure jump driving noise and Brownian noise as the heat kernel bounds hold for any $\alpha \in (1,2]$;
- Noting that the martingale problem of the mollified SDE is well-posed, prove tightness of the sequence of probability measures induced by its martingale problem, identify its limit probability measure and prove its uniqueness (see Section 2.2, [CM22]).
- The identified limit is exactly the martingale solution of (37) in the sense of definition (i)-(ii)$''$.

Weak solution and pathwise uniqueness When the drift of the SDE is smooth enough, equivalence between a martingale problem and weak well-posedness is direct. Specifically, having a martingale solution, one can recover the underlying noise from the associated canonical process and build the dynamics. However, in the case of Besov drift, the difficulty lies in specifying the drift, thus the strategy with recovering the noise does not work. However, one can consider an *enlarged* martingale problem that allows us to keep track of the noise. And so the strategy is to build simultaneously the martingale solution associated to (37) and the noise (X, Z) as the solution of an enlarged martingale problem. This results into constructing the drift as the difference between the dynamics and the noise.

First, we have to introduce a *Young integral* which will allow us to give a meaning to the drift.

Definition 5.4. *Let $\tau > 0$, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le\tau}, \mathbb{P})$ be a filtered probability space and let $(\psi_t)_{0\le t\le\tau}$ be a progressively measurable process. Let $(A(s,t))_{0\le s\le t\le\tau}$ be a continuous and progressively measurable map, i.e. for any $0 \le s \le t$,*

$$\Omega \times \{s' \in [0,s],\ t' \in [0,t],\ s' \le t'\} \ni (\omega, s', t') \mapsto A(s', t')$$

is $\mathcal{F} \otimes \mathcal{B}(\{s' \in [0,s],\ t' \in [0,t],\ s' \le t'\})$-measurable and

$$\{s' \in [0,\tau],\ t' \in [0,\tau],\ s' \le t'\} \ni (s,t) \mapsto A(s,t)$$

is continuous. For $l \ge 1$, we say that

$$\int_0^\tau \psi_t A(t, t+dt) := \lim_{\substack{\Delta \text{ partition of } [0,\tau],\\ |\Delta|\to 0}} \sum_{t_i\in\Delta} \psi_{t_i} A(t_i, t_{i+1}) \quad \text{in } L^l(\Omega, \mathbb{P})$$

is a L^l-stochastic Young integral *of ψ w.r.t. the pseudo increment A when it exists.*

We choose the following definition of weak solution of the SDE (37).

Definition 5.5. *We say that a pair $(Y, \mathcal{Z})$ of adapted processes on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{s\ge 0}, \mathbb{P})$ is a* weak solution *of the formal SDE* (37) *if*

(i) $\mathcal{Z}$ is an $(\mathcal{F}_t)_{t\ge 0}$ α-stable process;

(ii) For any $t \in [0,T]$, $x \in \mathbb{R}^d$, $(Y, \mathcal{Z})$ satisfies

$$Y_t = x + \int_0^t \mathcal{B}(s, Y_s, ds) + \mathcal{Z}_t,\ \mathbb{P}-a.s., \quad \mathbb{E}[|\int_0^t \mathcal{B}(s, Y_s, ds)|] < \infty,$$

where for any $0 \le v \le s \le T$,

$$\mathcal{B}(s, Y_s, s-v) := \int_v^s dr \int_{\mathbb{R}^d} b(r,y) p_\alpha(r-v, y-x) \tag{43}$$

is understood as an L^1-stochastic Young integral*.*

By considering an enlarged martingale problem, we can prove the following theorem, which allows us to interpret the distributional drift.

Theorem 5.3 (Theorem 4, [CM22])**.** *Let $p, q, r \ge 1$, $\alpha \in (1,2]$ and $\gamma \in (1/2, 1)$ satisfy* (39)*. Then there exists a probability measure $\mathbf{P}^\alpha$ on $C([0,T], \mathbb{R}^{2d})$ if $\alpha = 2$ and $D([0,T], \mathbb{R}^{2d})$ if $\alpha \in (1,2)$ s.t. the canonical process, denoted by $(X, \mathcal{Z})$, satisfies*

1. *The law of X under $\mathbf{P}^\alpha$ is a solution of the martingale problem associated with* (37)*, and the law of $\mathcal{Z}$ under $\mathbf{P}^\alpha$ is a Brownian motion if $\alpha = 2$ and an α-stable process if $\alpha \in (1,2)$;*
2. *$(X, \mathcal{Z})$ satisfies (ii).*

Note that the above theorem gives existence of a weak solution for (37). For the proof of uniqueness of the weak solution and pathwise uniqueness, see Section 5, [CM22]. Moreover, the representation of the drift as 43 is precise in the sense that it allows us to perform numerical approximations and related Monte-Carlo methods. It is also precise in the sense that when b is a Hölder function, then (43) coincides with b, which, in particular, is stated in the following proposition.

Proposition 5.1 (Proposition 7, [CM22])**.** *Either the martingale solution or the weak solution of the formal SDE* (37) *is:*

1. *A virtual solution of the formal SDE* (37)*;*
2. *A Dirichlet process;*
3. *It holds that for any smooth approximating sequence of drift* $(b_m)_{m\geq 1}$ *s.t.* $\|b - b_m\|_{L^r(B_{p,q}^{-1+\gamma})} \xrightarrow[m\to\infty]{} 0$,

$$\lim_{m\to\infty} \| \int_0^t \mathcal{B}(s, X_s, ds) - \int_0^t b_m(s, X_s)ds\|_{L^l} = 0, \quad l \in [1, \alpha),$$

with $l = 1$ *for the weak solution;*

4. *If* $b \in C^\beta(\mathbb{R}_+, \mathbb{R}^d)$*, we have*

$$\int_0^t \mathcal{B}(s, X_s, ds) = \int_0^t b(s, X_s)ds, \quad a.s.$$

5.3 McKean-Vlasov SDEs with Besov kernel driven by stable noise

We now consider the formal McKean-Vlasov SDE

$$dX_t = b * \mu_{X_t}(t, X_t)dt + dZ_t, \quad X_0 \sim \mu_0, \tag{44}$$

where $(Z_t)_{t\geq 0}$ is a d-dimensional symmetric α-stable process, $\alpha \in (1, 2]$, $\mu_{X_t} = \text{Law}(X_t)$, and $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$.

The special convolution dependency of the drift on the measure allows us to prove well-posedness of the SDE (44) for b being less regular than in (37). By proving first that the law of the solution is itself smooth enough, which is achieved using the nice properties of the stable noise, we can regularize the singular interaction kernel and define properly the drift. We are also able to obtain well-posedness in a strong sense unlike in Theorem 5.1.

Theorem 5.4 (Theorem 1, [CJM22])**.** *Let* $T > 0$ *be fixed,* $t \in [0, T)$*, and* $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$*. Let* $b \in L^r((t, T], B^\beta_{p,q}(\mathbb{R}^d, \mathbb{R}^d))$*.*

1. *Let*

$$\beta > 1 - \alpha + \frac{d}{p} + \frac{\alpha}{r}. \tag{45}$$

Then the McKean-Vlasov SDE (44) *admits a unique weak solution s.t. its marginal laws* $(\mu_s^{t,\mu})_{s\in(0,T]}$ *admit a density* $\rho_{t,\mu}(s, \cdot)$ *for a.a.* $s \in (t, T]$ *that satisfies*

$$\int_t^T \|\rho_{t,\mu}(s, \cdot)\|^{\bar{r}}_{B^{-\beta}_{p',q'}} ds < +\infty,$$

for any $\bar{r} \in [r', (-\beta/\alpha + d/(\alpha p))^{-1})$*, where* $1/p + 1/p' = 1/q + 1/q' = 1/r + 1/r' = 1$.

2. *Let*

$$\beta > 2 - \frac{3}{2}\alpha + \frac{d}{p} + \frac{\alpha}{r}. \tag{46}$$

Assume that $(Z_t)_{t\geq 0}$ *is rotationally invariant. Then the McKean-Vlasov SDE* (44) *admits a unique strong solution.*

Globally, for proving weak well-posedness we aim to show that $\mathfrak{b} = b * \mu_{X_t}$ satisfies a particular case of Theorem 5.1, i.e. that the non-linear drift belongs to the appropriate Lebesgue-Besov space. More precisely, we are able to show the well-posedness of the corresponding martingale problem if $\mathfrak{b} \in L^r([0, T], L^\infty(\mathbb{R}^d))$, where

$$\frac{\alpha}{r} < \alpha - 1, \tag{47}$$

see Lemma 14, [CJM22]. Thanks to the embedding property (2) of Besov spaces with Lebesgue spaces, we are indeed in the framework of Theorem 5.1. Actually, we are able to show that for any $r_0, \bar{r}$ s.t. $r_0^{-1} = \bar{r}^{-1} + r^{-1}$ and $r' \leq \bar{r} < (-\beta/\alpha + d/\alpha p)^{-1}$, there exists $C > 0$ s.t.

$$\|b * \rho_{t,\mu}\|_{L^{r_0}((t,T),B^0_{\infty,1})} \leq C\|b\|_{L^r(B^\beta_{p,q})}\|\rho_{t,\mu}\|_{L^{\bar{r}}(B^{-\beta}_{p',q'})} < \infty,$$

where $\rho_{t,\mu}$ is the density of the law of the solution to the equation (44) for $t \in [0, T]$. Thus, the big part of this problem is to be devoted to proving that the density $\rho_{t,\mu}$ indeed exists and satisfies the above regularity properties. The idea behind the proof is to consider a *mollified* interaction kernel in (44) so that it becomes smooth and bounded, and thus the mollified convolution drift is well-defined for any measure. The modified SDE is well-posed and admits a density which can be shown to have the desired regularity properties using the auxiliary Fokker-Planck equation and Duhamel representation. Then, by stability argument we are able to conclude the result for the law of the solution to (44).

Strong well-posedness will follow from the weak well-posedness by verifying that the drift satisfies *Krylov-Röckner type conditions.*

Weak well-posedness Let $b_\epsilon(s, \cdot) := (b^\epsilon_{sp}(\cdot, \cdot) * \eta_\epsilon)(s)$, where $b^\epsilon_{sp}(s, \cdot) := (p^\alpha(\epsilon, \cdot) * b(s, \cdot))(\cdot)$ is a spatial approximation of the drift by the density of the α-stable process, $\epsilon > 0$, $s \in [0, T]$. Then $b_\epsilon(s, \cdot)$ is a mollification in time variable of this drift spatial approximation. By Proposition 2, [CJM22], if $b \in L^r((t, T], B^\beta_{p,q}(\mathbb{R}^d, \mathbb{R}^d))$ for $p, q \geq 1$, $\beta \in (-1, 0]$, then

$$\|b - b^\epsilon\|_{L^{\bar{r}}((t,T],B^{\tilde{\beta}}_{p,q}(\mathbb{R}^d,\mathbb{R}^d))} \xrightarrow[\epsilon \to 0]{} 0 \quad \forall \tilde{\beta} < \beta, \tag{48}$$

where $\bar{r} = r$ if $r < \infty$, and $\bar{r} < \infty$ if $r = \infty$. Moreover, there exists $\kappa \geq 1$ s.t. $\sup_{\epsilon>0} \|b^\epsilon\|_{L^{\bar{r}}((t,T],B^\beta_{p,q})} \leq \kappa\|b^\epsilon\|_{L^r((t,T],B^\beta_{p,q})}$. Thus, by [FKM21] ($\alpha \in (1, 2)$) and [Szn91] ($\alpha = 2$), the SDE (44) where the drift is replaced by its mollification $\mathfrak{b}^\epsilon = b^\epsilon * \mu$, is well-posed in a strong sense for any $\epsilon > 0$. In fact, the time marginals of the law of the solution of the mollified SDE is absolutely continuous w.r.t. the Lebesgue measure and so admits a smooth density (see [CG92]) which we denote by $\rho^\epsilon_{t,\mu}$. For any $s \in (t, T]$, $y \in \mathbb{R}^d$ and $\epsilon > 0$, the density $\rho^\epsilon_{t,\mu}$ solves in a weak sense the Fokker-Planck equation given by

$$\begin{cases} \partial_s \rho^\epsilon_{t,\mu}(s, y) + \operatorname{div}\big(\mathfrak{b}^\epsilon(s, y)\rho^\epsilon_{t,\mu}(s, y)\big) - \mathcal{L}^\alpha \rho^\epsilon_{t,\mu}(s, y) = 0, \\ \rho^\epsilon_{t,\mu}(t, \cdot) = \mu, \end{cases} \tag{49}$$

and satisfies the Duhamel representation

$$\rho^\epsilon_{t,\mu}(s, y) = p^\alpha_{s-t} * \mu(y) - \int_s^t \Big(\big(\mathfrak{b}^\epsilon(v, \cdot)\rho^\epsilon_{t,\mu}(v, \cdot)\big) * \nabla p^\alpha_{s-v}\Big)(y)dv, \tag{50}$$

see Lemma 3, [CJM22]. This can be proven by applying Itô's formula to $\phi(v, X^{\epsilon,t,\mu}_v) := p^\alpha_{s-v} * f(X^{\epsilon,t,\mu}_v)$, where $X^{\epsilon,t,\mu}_v$ is the solution of the SDE with the mollified drift and f is a smooth and bounded function on $\mathbb{R}^d$. For proving finiteness of the appeared integrals, one can use the nice properties of Besov spaces, in particular Properties (2), (6) and (7) of Lemma 4.1.

Introduce the following parameter, which we call a *gap to singularity*:

$$\Gamma := \beta - \Big(1 - \alpha + \frac{d}{p} + \frac{\alpha}{r}\Big) > 0.$$

Proposition 5.2. *Let $(t, \mu) \in [0, T] \times \mathcal{P}(\mathbb{R}^d)$ and let $(\epsilon_k)_{k\geq1}$ be a decreasing sequence going to zero. Let $(\rho^{\epsilon_k}_{t,\mu})_{k\geq1}$ be a sequence with its elements solving (49). Then there exists $\rho_{t,\mu} \in L^{\bar{r}}(B^{-\beta+v\Gamma}_{p',q'})$ with $\bar{r} \in \Big[r', \frac{1}{\alpha}\Big(-\beta+v\Gamma+\frac{d}{p}\Big)^{-1}\Big)$ with r' as in Theorem 5.1, $v \in [0, 1)$ s.t.*

(i) For any $s \in [t, T]$, $\rho_{t,\mu}(s, \cdot) \in \mathcal{P}(\mathbb{R}^d)$ is the limit point of $(\rho^{\epsilon_k}_{t,\mu})_{k\geq1}$;

(ii) For almost any $s \in (t, T]$, $\rho_{t,\mu}(s, \cdot)$ is absolutely continuous w.r.t. the Lebesgue measure;

(iii) $\rho_{t,\mu}$ *solves weakly the non-linear Fokker-Planck equation*

$$\begin{cases} \partial_s \rho_{t,\mu}(s,y) + div\big(\mathfrak{b}(s,y)\rho_{t,\mu}(s,y)\big) - \mathcal{L}^\alpha \rho_{t,\mu}(s,y) = 0, \\ \rho_{t,\mu}(t,\cdot) = \mu, \end{cases} \tag{51}$$

and satisfies the Duhamel representation

$$\rho_{t,\mu}(s,y) = p^\alpha_{s-t} * \mu(y) - \int_s^t \Big(\big(\mathfrak{b}(v,\cdot)\rho_{t,\mu}(v,\cdot)\big) * \nabla p^\alpha_{s-v}\Big)(y)dv;$$

(iv) For any $s \in [t,T]$, $(\rho_{t,\mu}(s,\cdot))_{s\in[t,T]}$ *is a unique weak solution of* (51) *in* $L^{\bar r}(B^{-\beta+v\Gamma}_{p',q'})$.

Idea of proof. (i) First, we establish a priori estimates on $\rho^{\epsilon_k}_{t,\mu}$ for fixed $k \geq 1$. More precisely, Lemma 7, [CJM22], gives us that there exists a constant $C = C(d,\alpha,r,\beta,p,v) > 0$ and $\theta = \theta(d,\alpha,\beta,p,v) > 0$ s.t.

$$\int_t^T \|\rho^\epsilon_{t,\mu}\|^{\bar r}_{B^{-\beta+v\Gamma}_{p',q'}} ds \leq C(T-t)^\theta, \quad \forall \epsilon > 0,$$

and for any $t < S < T$ and $\bar r' \in \Big[r', \frac{1}{\alpha}\Big(-\beta + v\Gamma + \frac{d}{p}\Big)^{-1}\Big)$

$$\|\rho^\epsilon_{t,\mu}\|_{L^{\bar r'}_w((t,S],B^{-\beta+v\Gamma}_{p',q'})} \leq C(S-t)^{\bar{\bar\delta}/\bar r'},$$

$$\Big(\int_t^S (S-s)^{-\bar r'/\alpha}\|\rho^\epsilon_{t,\mu}(s,\cdot)\|^{\bar r'}_{B^{-\beta+v\Gamma}_{p',q'}} ds\Big)^{1/\bar r'} \leq C(S-t)^{\bar{\bar\delta}/\bar r'}, \tag{52}$$

for some $\bar{\bar\delta} = \bar{\bar\delta}(\bar r') > 0$. These estimates can be proven using Duhamel representation (50) and Besov space properties (2), (4), (6) and (7) in Lemma 4.1. Having these bounds in hand, using (48) and using again Lemma 4.1, we can show that $(\rho^{\epsilon_k}_{t,\mu})_{k\geq 1}$ is a Cauchy sequence in $L^{\bar r}(B^{-\beta+v\Gamma}_{p',q'}) \cap L^\infty((t,T],L^1)$ and

$$\|\rho^{\epsilon_k}_{t,\mu} - \rho_{t,\mu}\|_{L^{\bar r}(B^{-\beta+v\Gamma}_{p',q'})} + \sup_{s\in(t,T]} \|(\rho^{\epsilon_k}_{t,\mu} - \rho_{t,\mu})(s,\cdot)\|_{L^1} \xrightarrow[k\to\infty]{} 0,$$

see Lemma 8, [CJM22].

(ii) Follows from the proof of (i), see Lemma 8, [CJM22].

(iii) Using the fact that $\rho^\epsilon_{t,\mu}$ satisfies Duhamel representation and solves Fokker-Planck equation, we are able to obtain the claimed results the stability property mentioned in the proof of (i), convergence (48) and Lemma 4.1, see Lemma 9, [CJM22].

(iv) We prove it using Duhamel representation (52) and stability property mentioned in the proof of (i) that holds for non-mollified density, see Lemma 10, [CJM22]. □

To derive weak well-posedness of the McKean-Vlasov equation (44), we manipulate the equivalent (in this setup) martingale problem which has to be reformulated comparing to the linear case in Section 5.2.

Definition 5.6. *We say that a measure* $\mathbb{P}^\alpha \in \mathcal{P}(\mathbb{R}^d)$ *is a* martingale solution *to the* non-linear martingale problem *associated with an operator* $\mathcal{L}^X$ *on the probability space* $(\Omega_\alpha, \mathcal{F}, \mathbb{P})$ *on* $[t,T]$ *if*

(i) $\mathbb{P}^\alpha \circ x_t^{-1} = \mu$*;*

(ii) For a.a. $s \in (t,T]$, $\mathbb{P}^\alpha \circ x_s^{-1}$ *is absolutely continuous w.r.t. the Lebesgue measure and its density belongs to* $L^{r'}((t,T],B^{-\beta}_{p',q'})$*;*

(iii) For all $\phi \in C^1([t,T],C^2_0(\mathbb{R}^d))$,

$$\Big(\phi(s,x_s) - \phi(t,x_t) - \int_t^s (\partial_r + \mathcal{L}^x)\phi(r,x_r)dr\Big)_{t\leq s\leq T}$$

is a (square integrable if $\alpha = 2$*)* $\mathbb{P}^\alpha$*-martingale.*

In our case, $\mathcal{L}^x = \mathfrak{b}_{\mathbb{P}^\alpha \circ x^{-1}} \cdot \nabla + \mathcal{L}^\alpha$. Having this definition, we now have to relate the Fokker-Planck equation (51) and the martingale problem associated to (44). Actually, using Proposition 5.2, one can prove that the solution to the non-linear martingale problem related to the mollified McKean-Vlasov SDE in $\mathcal{P}(\Omega_\alpha)$ equipped with weak topology to a solution to the non-linear martingale problem related to (44) as soon as (45) holds (see Proposition 12, [CJM22]).

Proposition 5.3. *Under assumption* (45)*, for any* $(t, \mu) \in [0, T] \times \mathcal{P}(\mathbb{R}^d)$ *the SDE* (44) *admits a unique weak solution s.t. its marginal laws* $(\mu_s^{t,\mu})_{s\in[t,T]}$ *have a density for a.a.* $s \in (t, T]$*, i.e.* $\mu_s^{t,\mu}(dx) = \rho_{t,\mu}(s, x)dx$ *and* $\rho_{t,\mu} \in L^{\bar{r}}(B^{-\beta}_{p',q'})$.

While existence of a weak solution can be derived from a non-linear martingale problem as described before, uniqueness of the solution is derived from the Krylov-Röckner type condition (47), see Proposition 13 and Lemma 14, [CJM22].

Strong well-posedness Having the weak well-posedness result in hand, we can prove strong well-posedness by considering the convolution drift as a *linear* drift $\mathfrak{b}$ and showing that it satisfies the Krylov-Röckner type condition. More precisely, let

$$\mathfrak{b}(\cdot, \cdot) = \int_{\mathbb{R}^d} b(\cdot, \cdot)\rho_{t,\mu}(\cdot, y)dy.$$

Then

1. If $\alpha = 2$, then $\mathfrak{b} \in L^s(L^l)$, where

$$\frac{2}{s} + \frac{d}{l} < 1$$

   with s, l to be specified explicitly;

2. If $\alpha \in (1, 2)$, then $(I - \Delta)^{\gamma/2}\mathfrak{b} \in L^s(L^l)$, where

$$l \in \Big(\frac{2d}{\alpha} \vee 2, \infty\Big), \quad s \in \Big(\frac{\alpha}{\alpha - 1}, \infty\Big), \quad \frac{\alpha}{s} + \frac{d}{l} < \alpha - 1,$$

   again with s, l to be specified explicitly.

Thus, we have the following strong well-posedness result.

Proposition 5.4. *Under assumption* (46)*, for any* $(t, \mu) \in [0, T] \times \mathcal{P}(\mathbb{R}^d)$ *the SDE* (44) *admits a unique strong solution s.t. its law* $\mu^{t,\mu}$ *belongs to* $L^{\bar{r}}(B^{-\beta}_{p',q'})$ *and for a.a.* $s \in (t, T]$*,* $\mu^{t,\mu}(dx) = \rho_{t,\mu}(s, x)dx$.

For detailed proof, see Proposition 15, [CJM22].

Perspectives

One of the questions of interest for us is to investigate how much more irregular the interaction kernel can be in (44) so that solution exists in some sense. Recalling that for a Lebesgue-Besov interaction kernel strong well-posedness was shown by exploiting the particular Krylov-Röckner type condition, the open question is what can we say about the interaction kernel b if

$$\mathfrak{b} := b * \mu_{X_t} \in L^r(B^\beta_{p,q}),$$

for which weak well-posedness holds by Theorem 5.1. Thanks to the regularizing effect of the law, one hopes to go below index proven in Theorem 5.4. Similarly to the *linear* equation, additional structure on the interaction kernel can be assumed as it was mentioned in the introduction (e.g. rough paths) and potentially it would allow us to obtain well-posedness even for more irregular Besov drifts. Similar questions arise in the study of *kinetic* equations which are not covered in this report and deserve additional attention. We refer to [Cha+23], [HRZ22] and references therein for known results in this direction.

Acknowledgments I want to thank my advisor Stéphane Menozzi for a thorough supervision and guidance during my internship which allowed me to go into depth of the topic. I also want to thank the lab LaMME for a warm welcome and for the support of people whom I met there.

References

[ABM20] Siva Athreya, Oleg Butkovsky, and Leonid Mytnik. "Strong existence and uniqueness for stable stochastic differential equations with distributional drift". In: *The Annals of Probability* 48.1 (2020), pp. 178–210. DOI: 10.1214/19-AOP1358. URL: https://doi.org/10.1214/19-AOP1358.

[Aro67] Donald Aronson. "Bounds for the Fundamental Solution of a Parabolic Equation". In: *Bulletin of The American Mathematical Society - BULL AMER MATH SOC* 73 (Nov. 1967). DOI: 10.1090/S0002-9904-1967-11830-5.

[Bas97] Richard Bass. *Diffusions and Elliptic Operators*. Springer New York, NY, 1997. URL: https://doi.org/10.1007/b97611.

[BC01] Richard Bass and Zhen-Qing Chen. "Stochastic differential equations for Dirichlet processes". In: *Probability Theory and Related Fields* 121 (Nov. 2001), pp. 422–446. DOI: 10.1007/s004400100151.

[Bur90] Victor Burenkov. "On estimates of Fourier transforms and convolutions in Nikol'skij-Besov spaces". In: *Proc. Steklov Inst. Math.* 187 (1990), pp. 35–44. ISSN: 0081-5438.

[CD22a] Louis-Pierre Chaintron and Antoine Diez. "Propagation of chaos: A review of models, methods and applications. I. Models and methods". In: *Kinetic and Related Models* 15.6 (2022), p. 895. ISSN: 1937-5077. DOI: 10.3934/krm.2022017. URL: http://dx.doi.org/10.3934/krm.2022017.

[CD22b] Louis-Pierre Chaintron and Antoine Diez. "Propagation of chaos: A review of models, methods and applications. II. Applications". In: *Kinetic and Related Models* 15.6 (2022), pp. 1017–1173. ISSN: 1937-5093. DOI: 10.3934/krm.2022018. URL: https://www.aimsciences.org/article/id/62948a7a2d80b70dfd2582aa.

[CG92] Agnès Coquio and J. B Gravereaux. "Calcul de Malliavin et régularité de la densité d'une probabilité invariante d'une châine de Markov". fr. In: *Annales de l'I.H.P. Probabilités et statistiques* 28 (1992), pp. 431–478. URL: http://dml.mathdoc.fr/item/AIHPB_1992__28_4_431_0.

[Cha+23] Paul-Eric Chaudru de Raynal et al. "Heat kernel and gradient estimates for kinetic SDEs with low regularity coefficients". In: *Bulletin des Sciences Mathématiques* 183 (2023), p. 103229. ISSN: 0007-4497. DOI: https://doi.org/10.1016/j.bulsci.2023.103229. URL: https://www.sciencedirect.com/science/article/pii/S0007449723000039.

[Cha20] Paul-Eric Chaudru de Raynal. "Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift". In: *Stochastic Processes and their Applications* 130.1 (2020), pp. 79–107. ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2019.01.006. URL: https://www.sciencedirect.com/science/article/pii/S0304414919300511.

[CJM22] Paul-Eric Chaudru de Raynal, Jean-François Jabir, and Stéphane Menozzi. *Multidimensional Stable driven McKean-Vlasov SDEs with distributional interaction kernel – a regularization by noise perspective.* May 2022. DOI: 10.48550/arXiv.2205.11866.

[CM22] Paul-Eric Chaudru de Raynal and Stéphane Menozzi. "On multidimensional stable-driven stochastic differential equations with Besov drift". In: *Electronic Journal of Probability* 27 (2022), pp. 1–52. DOI: 10.1214/22-EJP864. URL: https://doi.org/10.1214/22-EJP864.

[DD14] François Delarue and Roland Diel. "Rough paths and 1d SDE with a time dependent distributional drift: application to polymers". In: *Probability Theory and Related Fields* 165 (Feb. 2014). DOI: 10.1007/s00440-015-0626-8.

[FIR15] Franco Flandoli, Elena Issoglio, and Francesco Russo. "Multidimensional stochastic differential equations with distributional drift". In: *to appear in Transactions of the American Mathematical Society* (Apr. 2015). DOI: 10.1090/tran/6729.

[FJM24] Mathis Fitoussi, Benjamin Jourdain, and Stéphane Menozzi. *Weak well-posedness and weak discretization error for stable-driven SDEs with Lebesgue drift.* 2024. arXiv: 2405.08378.

[FKM21] Noufel Frikha, Valentin Konakov, and Stéphane Menozzi. "Well-posedness of some non-linear stable driven SDEs". In: *Discrete & Continuous Dynamical Systems - A* 41.2 (2021), pp. 849–898. ISSN: 1078-0947. DOI: 10.3934/dcds.2020302. URL: https://www.aimsciences.org/article/id/53aeae71-076b-4d93-aff8-dceaf9cb8895.

[Fla11] Franco Flandoli. *Random Perturbation of PDEs and Fluid Dynamic Models: École d'Été de Probabilités de Saint-Flour XL – 2010*. Vol. 2015. Springer Berlin, Heidelberg, Jan. 2011. ISBN: 978-3-642-18230-3. DOI: 10.1007/978-3-642-18231-0.

[Fri83] Avner Friedman. *Partial Differential Equations of Parabolic Type*. 1983. URL: https://api.semanticscholar.org/CorpusID:117905500.

[Hao+24] Zimo Hao et al. *Propagation of chaos for moderately interacting particle systems related to singular kinetic Mckean-Vlasov SDEs*. 2024. arXiv: 2405.09195 [math.AP]. URL: https://arxiv.org/abs/2405.09195.

[HRZ22] Zimo Hao, Michael Röckner, and Xicheng Zhang. *Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions*. 2022. arXiv: 2204.07952 [math.PR]. URL: https://arxiv.org/abs/2204.07952.

[IW81] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*. Vol. 24. North-Holland Mathematical Library. Amsterdam: North-Holland Publishing Co., 1981, pp. xiv+464. ISBN: 0-444-86172-6.

[JM24] Benjamin Jourdain and Stéphane Menozzi. "Convergence rate of the Euler–Maruyama scheme applied to diffusion processes with drift coefficient and additive noise". In: *The Annals of Applied Probability* 34.1B (2024), pp. 1663–1697. DOI: 10.1214/23-AAP2006. URL: https://doi.org/10.1214/23-AAP2006.

[KM17] Valentin Konakov and Stéphane Menozzi. "Weak error for the Euler scheme approximation of diffusions with non-smooth coefficients". In: *Electronic Journal of Probability* 22.none (2017), pp. 1–47. DOI: 10.1214/17-EJP53. URL: https://doi.org/10.1214/17-EJP53.

[Kol00] Vassili Kolokoltsov. "Symmetric Stable Laws and Stable-Like Jump-Diffusions". In: *Proceedings of the London Mathematical Society* 80 (2000). URL: https://doi.org/10.1112/S0024611500012314.

[KP22] Helena Kremp and Nicolas Perkowski. "Multidimensional SDE with distributional drift and Lévy noise". In: *Bernoulli* 28 (Aug. 2022). DOI: 10.3150/21-BEJ1394.

[KR05] Nicolai Krylov and Michael Röckner. "Strong solutions of stochastic equations with singular time dependent drift". In: *Probab. Theory Related Fields* 131 (Feb. 2005), pp. 154–196. DOI: 10.1007/s00440-004-0361-z.

[Lem02] Pierre Gilles Lemarié–Rieusset. *Recent Developments in the Navier-Stokes Problem*. CRC Press, 2002. URL: https://doi.org/10.1201/9780367801656.

[LS68] Olga Ladyzhenskaja and Vsevolod Solonnikov. *Linear and Quasilinear Equations of Parabolic Type*. 1968. URL: https://doi.org/10.1090/mmono/023.

[MP91] Remigijus Mikulevicius and Eckhard Platen. *Rate of Convergence of the Euler Approximation for Diffusion Processes*. Published Paper Series. 1991. URL: https://EconPapers.repec.org/RePEc:uts:ppaper:1991-3.

[RS96] Thomas Runst and Winfried Sickel. *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations*. Berlin, New York: De Gruyter, 1996. ISBN: 9783110812411. DOI: doi:10.1515/9783110812411. URL: https://doi.org/10.1515/9783110812411.

[Sat99] Ken-iti Sato. *Lévy Processes and Infinitely Divisible Distributions*. New York: Cambridge University Press, 1999. URL: https://api.semanticscholar.org/CorpusID:209888537.

[Saw18] Yoshihiro Sawano. *Theory of Besov Spaces*. Springer Singapore, 2018. URL: https://doi.org/10.1007/978-981-13-0836-9.

[ST95] Winfried Sickel and Hans Triebel. "Hölder Inequalities and Sharp Embeddings in Function Spaces of B^s_{pq} and F^s_{pq} Type". In: *Zeitschrift für Analysis und ihre Anwendungen* 14 (Mar. 1995), pp. 105–140. DOI: 10.4171/ZAA/666.

[SV97] Daniel Stroock and Srinivasa Varadhan. *Multidimensional Diffusion Processes*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1997. ISBN: 9783540903536. URL: https://books.google.de/books?id=DuDsmoyqCy4C.

[Szn91] Alain-Sol Sznitman. "Topics in propagation of chaos". In: *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*. Ed. by Paul-Louis Hennequin. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, pp. 165–251. ISBN: 978-3-540-46319-1.

[Tom23] Milica Tomašević. "Propagation of chaos for stochastic particle systems with singular mean-field interaction of type". In: *Electronic Communications in Probability* 28.none (2023), pp. 1–13. DOI: 10.1214/23-ECP539. URL: https://doi.org/10.1214/23-ECP539.

[Tri83] Hans Triebel. *Theory Of Function Spaces*. Birkhäuser Basel, 1983. URL: https://doi.org/10.1007/978-3-0346-0416-1.

[Ver81] Alexander Veretennikov. "On strong solutions and explicit formulas for solutions of stochastic integral equations". In: *Mathematics of The Ussr-sbornik* 39 (1981), pp. 387–403. URL: 10.1070/SM1981v039n03ABEH001522.

[XZ20] Longjie Xie and Xicheng Zhang. "Ergodicity of stochastic differential equations with jumps and singular coefficients". In: *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 56.1 (2020), pp. 175–229. DOI: 10.1214/19-AIHP959. URL: https://doi.org/10.1214/19-AIHP959.

[Zha10] Xicheng Zhang. "Stochastic Volterra equations in Banach spaces and stochastic partial differential equation". In: *Journal of Functional Analysis* 258.4 (2010), pp. 1361–1425. ISSN: 0022-1236. DOI: https://doi.org/10.1016/j.jfa.2009.11.006. URL: https://www.sciencedirect.com/science/article/pii/S0022123609004650.

[Zvo07] Alexander Zvonkin. "A transformation of the phase space of a diffusion process that removes the drift". In: *Mathematics of the USSR-Sbornik* 22 (Oct. 2007), p. 129. DOI: 10.1070/SM1974v022n01ABEH001689.