On Multidimensional stable-driven Stochastic Differential Equations with Besov drift

Paul-Eric Chaudru de Raynal, Stéphane Menozzi

On Multidimensional stable-driven Stochastic Differential Equations with Besov drift

Paul-Éric Chaudru de Raynal* **and Stéphane Menozzi†**

July 28, 2019

Abstract

We establish well-posedness results for multidimensional non degenerate α-stable driven SDEs with time inhomogeneous singular drifts in $\mathbb{L}^r - \mathbb{B}_{p,q}^{-1+\gamma}$ with $\gamma < 1$ and α in $(1,2]$, where $\mathbb{L}^r$ and $\mathbb{B}_{p,q}^{-1+\gamma}$ stand for Lebesgue and Besov spaces respectively. Precisely, we first prove the well-posedness of the corresponding martingale problem and then give a precise meaning to the dynamics of the SDE. Our results rely on the smoothing properties of the underlying PDE, which is investigated by combining a perturbative approach with duality results between Besov spaces.

1 Introduction

1.1 Statement of the problem

We are here interested in providing a well-posedness theory for the following *formal* d-dimensional stable driven SDE. For a fixed $T > 0$, $t \in [0,T]$:

$$X_t = x + \int_0^t F(s, X_s)ds + \mathcal{W}_t, \tag{1.1}$$

where in the above equation $(\mathcal{W}_s)_{s\geq 0}$ is a d-dimensional symmetric α-stable process, for some α in $(1,2]$.

The main point here comes from the fact that the drift F is only supposed to belong to the space $\mathbb{L}^r([0,T], \mathbb{B}_{p,q}^{-1+\gamma}(\mathbb{R}^d, \mathbb{R}^d))$, where $\mathbb{B}_{p,q}^{-1+\gamma}(\mathbb{R}^d, \mathbb{R}^d)$ denotes a Besov space (see Section 2.6.4 of [Tri83] and Section 3.1.2 below). The parameters (p, q, γ, r) s.t. $1/2 < \gamma < 1$, $p, q, r \geq 1$ will have to satisfy some constraints to be specified later on in order to give a meaning to (1.1). Importantly, assuming the parameter γ to be strictly less than 1 implies that F can even not be a function, but just a distribution, so that it is not clear that the integral part in (1.1) has any meaning, at least as this. This is the reason why, at this stage, we talk about "*formal* d-dimensional stable SDE". There are many approaches to tackle such a problem which mainly depend on the choice of the parameters p, q, γ, r, α and the dimension d. Let us now try to review some of them.

The Brownian setting: $\alpha = 2$. There already exists a rather large literature about singular/distributional SDEs of type (1.1). Let us first mention the work by Bass and Chen [BC01] who derived in the Brownian scalar case the strong well-posedness of (1.1) when the drift writes (still formally) as $F(t,x) = F(x) = aa'(x)$, for a spatial function a being β-Hölder continuous with $\beta > 1/2$ and for a multiplicative noise associated with a^2, i.e. the additive noise $\mathcal{W}_t$ in (1.1) must be replaced by $\int_0^t a(X_s)d\mathcal{W}_s$. The key point in this setting is that the underlying generator associated with the SDE writes as $L = (1/2)\partial_x\big(a^2\partial_x\big)$. From this specific divergence form structure, the authors manage to use the theory of Dirichlet forms of Fukushima *et al.* (see [FOT10]) to give a proper meaning to (1.1). Importantly, the formal integral corresponding to the drift has to be understood as a Dirichlet process. Also, in the particular case where the distributional derivative of a is a signed Radon measure, the authors give an explicit expression of the drift of the SDE in terms of the local time (see Theorem 3.6 therein). In the multi-dimensional Brownian case, Bass and Chen have also established weak well-posedness of SDE of type (1.1) when the homogeneous drift belongs to the Kato class, see [BC03].

Many authors have also recently investigated SDEs of type (1.1) in both the scalar and multidimensional Brownian setting for time inhomogeneous drifts in connection with some physical applications. From these

*Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA, 73000 Chambéry, France. pe.deraynal@univ-smb.fr

†Laboratoire de Modélisation Mathématique d'Evry (LaMME), CNRS UNMR 8071, Université d'Evry Val d'Essonne, Paris Saclay, 23 Boulevard de France 91037 Evry, France and Laboratory of Stochastic Analysis, HSE, Shabolovka 31, Moscow, Russian Federation. stephane.menozzi@univ-evry.fr

works, it clearly appears that handling time inhomogeneous distributional drift can be a more challenging question. Indeed, in the time homogeneous case, denoting by $\mathbf{F}$ an antiderivative of F, one can observe that the generator of (1.1) can be written in the form $(1/2)\exp(-2\mathbf{F}(x))\partial_x\big(\exp(2\mathbf{F}(x))\partial_x\big)$ and the dynamics can again be investigated within the framework of Dirichlet forms (see e.g. the works by Flandoli, Russo and Wolf, [FRW03], [FRW04]). The crucial point is that in the time inhomogeneous case such connection breaks down. In this framework, we can mention the work by Flandoli, Issoglio and Russo [FIR17] for drifts in fractional Sobolev spaces. The authors establish therein the existence and uniqueness of what they call *virtual solutions* to (1.1): such solutions are defined through the diffeomorphism induced by the Zvonkin transform in [Zvo74] which is precisely designed to get rid of the *bad* drift through Itô's formula. Namely, after having established appropriate smoothness properties of the underlying PDE:

$$\begin{cases} \partial_t u + F \cdot Du + \frac{1}{2}\Delta u - (\lambda+1)u = -F, \\ u(T) = 0, \end{cases} \tag{1.2}$$

introducing $\Phi(t,x) = x + u(t,x)$ (Zvonkin transform), it is indeed seen from Itô's formula that, at a formal level, $Y_t = \Phi(t, X_t)$ satisfies:

$$Y_t = \Phi(0,x) + \int_0^t (\lambda+1)(s, \Phi^{-1}(s,Y_s))ds + \int_0^t D\Phi(s, \Phi^{-1}(s,Y_s))d\mathcal{W}_s, \tag{1.3}$$

which itself has a unique weak solution from the smoothness of u solving (1.2). Since Φ can be shown to be a C^1-diffeomorphism for λ large enough, a *virtual* solution to (1.1) is then rigorously defined setting $X_t = \Phi^{-1}(t, Y_t)$.

We can also refer to the work of Zhang and Zhao [ZZ17], who derived in the time homogeneous case the well-posedness of the martingale problem for the generator associated with (1.1), which also contain non trivial smooth enough diffusion coefficients (see also Remaks 7 and 14 below). Therein, they obtained as well as some Krylov type density estimates in Bessel potential spaces for the solution. Also, they manage to obtain more precise description of the limit drift in the dynamics in (1.1), which is interpreted as a suitable limit of a sequence of mollified drifts, i.e. $\lim_n \int_0^t F_n(s,X_s)ds$ for a sequence of smooth functions $(F_n)_{n\geq 1}$ converging to F.

The key point in these works, who heavily rely on PDE arguments, is to establish that the product $F \cdot Du$ in (1.2) is in some sense meaningful, which is a real issue since F is meant to be a distribution. This in particular implies to derive some sufficient smoothness properties for the gradient Du. Such estimates are usually obtained, and this will also be the case in the current work, through a Duhamel type perturbative argument (or mild representation of the solution) that leads to some *natural* constraints on the parameter of the space in which the drift is assumed to belong. To make things simple, if F is the derivative in space of a Hölder function $\mathbf{F}$ with Hölder exponent γ, it follows from the usual parabolic bootstrap that the gradient of the solution u to (1.2) can only be expected to live in a Hölder space of regularity index $-1+\gamma+\alpha-1$, $\alpha = 2$. Thus, in order to give a meaning to the product $F \cdot Du$ as a distribution (more specifically as an element of a suitable Besov-Hölder space) in (1.2), one has to assume that γ is such that $-1+\gamma+\alpha-1+\gamma > 1 \Leftrightarrow \gamma > 1/2$, recalling that $\alpha = 2$. This is indeed the threshold appearing in [FIR17] and [ZZ17] as well as the one previously obtained in [BC01]. Note that in such a case, the product also makes sense as a Young integral, i.e. $\int_{\mathbb{R}^d} Du(s,y) \cdot F(s,y)dy = \int_{\mathbb{R}^d} Du(s,y)d\mathbf{F}(s,y)$, which is again coherent with the thresholds.

To bypass such a limit, one therefore has to use a suitable theory in order to give a meaning to the product $F \cdot Du$. This is, for instance, precisely the aim of either rough paths, regularity structures or paracontrolled calculus. However, as a price to pay to enter this framework, one has to add some structure to the drift assuming that this latter can be enhanced into a rough path structure. In the scalar Brownian setting, and in connection with the KPZ equation, Delarue and Diel [DD16] used such specific structure to extend the previous results for an inhomogeneous drift which can be viewed as the generalized derivative of $\mathbf{F}$ with Hölder regularity index greater than $1/3$ (i.e. assuming that F belongs to $\mathbb{L}^\infty([0,T], \mathbb{B}^{(-1/3)^+}_{\infty,\infty})$). Importantly, in [DD16] the authors derived a very precise description of the meaning of the *formal* dynamics (1.1): they show that the drift of the solution may be understood as stochastic-Young integral against a mollification of the distribution by the transition density of the underlying noise. As far as we know, it appears to us that such a description is the more accurate that can be found in the literature on stochastic processes (see [CG16] for a pathwise version and Remark 18 in [DD16] for some comparisons between the two approaches). With regard to the martingale problem, the result of [DD16] has then been extended to the multidimensional setting by Cannizzaro and Choukh [CC18], but nothing is said therein about the dynamics.

The pure jump case: $\alpha < 2$. In the pure jump case, there are a few works concerning the well-posedness (1.1) in the singular/distributional case. Even for drifts that are functions, strong uniqueness was shown rather

recently. Let us distinguish two cases: the *sub-critical case* $\alpha \geq 1$, in this case the noise dominates the drift (in term of self-similarity index α) and the *super-critical* case $\alpha < 1$ where the noise does not dominate. In the first case, we can refer for bounded Hölder drifts to Priola [Pri12] who proved that strong uniqueness holds (for time homogeneous) functions F in (1.1) which are β Hölder continuous provided $\beta > 1 - \alpha/2$. In the second case, the strong well-posedness has been established under the same previous condition by Chen *et al.* [CZZ17]. Those results are multi-dimensional.

In the current distributional framework, and in the scalar case, the martingale problem associated with the formal generator of (1.1) has been recently investigated by Athreya, Butkovski and Mytnik [ABM18] for $\alpha > 1$ and a time homogeneous $F \in \mathbb{B}^{-1+\gamma}_{\infty,\infty}$ under the condition: $-1 + \gamma > (1 - \alpha)/2$. After specifying how the associated dynamics can be understood, viewing namely the drift as a Dirichlet process (similarly to what was already done in the Brownian case in [BC01]), they eventually manage to derive strong uniqueness under the previous condition. Note that results in that direction have also been derived by Bogachev and Pilipenko in [BP15] for drift belonging to a certain Kato class in the multidimensional setting.

Again, the result obtained by Athreya, Butkovski and Mytnik relies on the Zvonkin transform and hence requires to have a suitable theory for the associated PDE. In our pure-jump time inhomogeneous framework, it writes

$$\begin{cases} \partial_t u + F \cdot Du + L^\alpha u - (\lambda + 1)u = -F, \\ u(T) = 0, \end{cases} \tag{1.4}$$

where L^α is the generator of a non-degenerate α-stable process. Reproducing the previous reasoning concerning the expected parabolic bootstrap properties induced by the stable process, we can now expect that, when $F(t, \cdot)$ is the generalized derivative of a Hölder function $\mathbf{F}$ with regularity index γ (or putting in the Besov space terminology $F \in \mathbb{L}^\infty([0,T], \mathbb{B}^{-1+\gamma}_{\infty,\infty})$), the gradient of the solution of the above PDE has Hölder regularity index $-1 + \gamma + \alpha - 1$: we gain the stability index as regularity order. Again, in order to give a meaning to the product $F \cdot Du$ as a distribution (more specifically as an element of a suitable Besov-Hölder space) in (1.4), one has to assume that γ is such that $-1 + \gamma + \alpha - 1 + \gamma > 1 \Leftrightarrow \gamma > (3 - \alpha)/2$. This is precisely the threshold that will guarantee weak well-posedness holds for a drift $F \in \mathbb{L}^\infty([0,T], \mathbb{B}^{-1+\gamma}_{\infty,\infty})$.

1.2 Aim of the paper.

In the current work, we aim at investigating a rather large framework by considering the d-dimensional case $d \geq 1$, with a distributional, potentially singular in time, inhomogeneous drift (in $\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q})$) when the noise driving the SDE is symmetric α-stable process, α in $(1, 2]$. This setting thus includes both the Brownian and pure-jump case. In the latter case, we will also be able to consider driving noises with singular spectral measures. As previously done for the aforementioned results, our strategy relies on the idea by Zvonkin. The core of the analysis therefore consists in obtaining suitable *a priori* estimates on an associated underlying PDE of type (1.2) or (1.4). Namely, we will provide a Schauder type theory for the mild solution of such PDE for a large class of data. This result is also part of the novelty of our approach since these estimates are obtained thanks to a rather robust methodology based on heat-kernel estimates on the transition density of the driving noise together with duality results between Besov spaces viewed through their thermic characterization (see Section 3.1.2 below and Triebel [Tri83] for additional properties on Besov spaces and their characterizations). This approach does not distinguish the pure-jump and Brownian setting provided the heat-kernel estimates hold. It has for instance also been successfully applied in various frameworks, to derive Schauder estimates and strong uniqueness for a degenerate Brownian chain of SDEs (see [CdRHM18a], [CdRHM18b]) or Schauder estimates for super-critical fractional operators [CdRMP19].

Our first main result consists in deriving the well-posedness of the martingale problem introduced in Definition 1 under suitable conditions on the parameters p, q, r and γ, see Theorem 1. As a by-product of our proof, we also manage to obtain through Krylov type estimates that the canonical process associated with the solution of the martingale problem also possesses a density belonging to an appropriate Lebesgue-Besov space (see Corollary 2).

Then, under slightly reinforced conditions on p, q, r and γ, we are able to reconstruct the dynamics for the canonical process associated with the solution of the martingale problem, see Theorem 3, specifying how the Dirichlet process associated with the drift writes. In the spirit of [DD16], we in particular exhibit a main contribution in this drift that could be useful to investigate the numerical approximations of those singular SDEs (see equations (1.13) and (1.14)) and the recent work by De Angelis *et al.* [DGI19].

Let us conclude by mentionning that, while finishing the preparation of the present manuscript, we discovered a brand new preprint of Ling and Zhao [LZ19] which somehow presents some overlaps with our results. Therein, the Authors investigate *a priori* estimates for the elliptic version of the PDE of type (1.2) or (1.4) with (homogeneous) drift belonging to Hölder-Besov spaces with negative regularity index (i.e. in $\mathbb{B}^{-1+\gamma}_{\infty,\infty}$) and including a non-trivial diffusion coefficient provided the spectral measure of the driving noise is absolutely continuous. As an application, they derive the well-posedness of the associated martingale problem and prove that the drift can be understood as a Dirichlet process. They also obtained quite sharp regularity estimates on the density of the solution and succeeded in including the limit case $\alpha = 1$.

In comparison with their results, we here manage to handle the case of an inhomogeneous and singular in time drift which can also have additional space singularities, since the integrability indexes of the parameter p, q for the Besov space are not supposed to be $p = q = \infty$ (recall that we assume $F \in \mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q})$). Although we did not include it, we could also handle in our framework an additional non-trivial diffusion coefficient under their standing assumptions, we refer to Remarks 7 and 14 below concerning this point. It also turns out that we obtain more accurate version of the dynamics of the solution which is here, as mentioned above, tractable enough for practical purposes. We eventually mention that, as a main difference with our approach, the controls in [LZ19] are mainly obtained through Littlewood-Paley decompositions whereas we rather exploit the thermic characterization and the parabolic framework for the PDE. In this regard, we truly think that the methodology to derive the a priori estimates in both works can be seen as complementary. Eventually, we mention that we also manage to derive pathwise uniqueness in the scalar case.

The paper is organized as follows. We introduce our main assumptions and state our results in the next paragraph. Section 2 is dedicated to the proof of the main results concerning the SDE: we state in Subsection 2.1 the key *a priori* controls for the underlying PDE (with both the mollified and initial rough coefficients) and then describe in Subsection 2.2 how to pass from the PDE results to the SDE itself, following somehow the procedure considered by Delarue and Diel [DD16]. In Section 3, we prove the *a priori* control for the PDE introducing to this end the auxiliary mathematical tools needed (heat kernel estimates, thermic characterization of Besov spaces). Section 4 is then devoted to the reconstruction of the dynamics from the solution to the martingale problem and Section 5 to the pathwise uniqueness in dimension one. Eventually, we postpone to Appendix A the proof of some technical results.

1.3 Assumptions and main results

Framework. We will denote by L^α the generator associated with the driving stable process $(\mathcal{W}_s)_{s\geq 0}$. When $\alpha = 2$, $L^2 = (1/2)\Delta$ where Δ stands for the usual Laplace operator on $\mathbb{R}^d$. In the pure-jump stable case $\alpha \in (1,2)$, for all $\varphi \in C_0^\infty(\mathbb{R}^d, \mathbb{R})$:

$$L^\alpha \varphi(x) = \text{p.v.} \int_{\mathbb{R}^d} \big[\varphi(x+z) - \varphi(x)\big] \nu(dz), \tag{1.5}$$

where, writing in polar coordinates $z = \rho\xi$, $\rho \in \mathbb{R}_+ \times \mathbb{S}^{d-1}$, the Lévy measure decomposes as $\nu(dz) = \frac{\mu(d\xi)}{\rho^{1+\alpha}}$ with μ a symmetric non degenerate measure on the sphere $\mathbb{S}^{d-1}$. Precisely, we assume:

(UE) There exists $\kappa \geq 1$ s.t. for all $\lambda \in \mathbb{R}^d$:

$$\kappa^{-1}|\lambda|^\alpha \leq \int_{\mathbb{S}^{d-1}} |\langle \lambda, \xi\rangle|^\alpha \mu(d\xi) \leq \kappa |\lambda|^\alpha. \tag{1.6}$$

Observe in particular that a rather large class of spherical measures μ satisfy (1.6). From the Lebesgue measure, which actually leads, up to a normalizing constant, to $L^\alpha = -(-\Delta)^{\alpha/2}$ (usual fractional Laplacian of order α corresponding to the generator of the isotropic stable process), to sums of Dirac masses in each direction, i.e. $\mu_{\text{Cyl}} = \sum_{j=1}^d c_j(\delta_{e_j} + \delta_{-e_j})$, with $(e_j)_{j\in[\![1,d]\!]}$ standing for the canonical basis vectors, which for $c_j = 1/2$ then yields $L^\alpha = -\sum_{j=1}^d (-\partial^2_{x_j})^{\alpha/2}$ corresponding to the cylindrical fractional Laplacian of order α associated with the sum of scalar symmetric α-stable processes in each direction. In particular, it is clear that under **(UE)**, the process $\mathcal{W}$ admits a smooth density in positive time (see e.g. [Kol00]). Correspondingly, L^α generates a semi-group that will be denoted from now on by $P^\alpha_t = \exp(tL^\alpha)$. Precisely, for all $\varphi \in B_b(\mathbb{R}^d, \mathbb{R})$ (space of bounded Borel functions), and all $t > 0$:

$$P^\alpha_t[\varphi](x) := \int_{\mathbb{R}^d} dy p_\alpha(t, y-x)\varphi(y), \tag{1.7}$$

where $p_\alpha(t,\cdot)$ stands for the density of $\mathcal{W}_t$. Further properties associated with the density p_α, in particular concerning the integrability properties of its derivatives, are stated in Section 3.1.

Main results. As already mentioned, the SDE (1.1) is stated at a *formal* level. Indeed, the drift being only a distribution, the dynamics (1.1) cannot have a clear meaning as this stage. Our first main result concerns the weak well-posedness for (1.1) in terms of the Stroock and Varadhan formulation of martingale problem (see [SV79]). However, still in our particular setting and in order to avoid discussion on pointwise products involving distributions, we slightly modify the definition of such a formulation.

Definition 1. *Let $\alpha \in (1,2]$. For any given fixed $T>0$, we say that the martingale problem with data $\big(L^\alpha, F, x\big)$, $x \in \mathbb{R}^d$ is well posed if there exists a unique probability measure $\mathbb{P}^\alpha$ on $\mathcal{C}([0,T],\mathbb{R}^d)$ if $\alpha = 2$ and on the Skorokhod space $\mathcal{D}([0,T],\mathbb{R}^d)$ of $\mathbb{R}^d$-valued càdlàg functions if $\alpha \in (1,2)$, s.t. the canonical process $(X_t)_{0\le t\le T}$ satisfies the following conditions:*

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

(ii) For any $f \in \mathcal{C}([0,T],\mathbb{L}^\infty(\mathbb{R}^d))$, the process

$$\Big(u(t,X_t) - \int_0^t f(s,X_s)ds - u(0,x_0)\Big)_{0\le t\le T} \tag{1.8}$$

is a $\mathbb{P}^\alpha$-martingale where $u \in \mathcal{C}^{0,1}([0,T],\mathbb{R}^d)$ is the mild solution of

$$\begin{aligned} \partial_t u(t,x) + L^\alpha u(t,x) + F(t,x)\cdot Du(t,x) &= f(t,x), \quad \text{on } [0,T)\times\mathbb{R}^d, \\ u(T,x) &= 0, \quad \text{on } \mathbb{R}^d. \end{aligned} \tag{1.9}$$

Having such a definition at hand, we may state our first existence and uniqueness result related to (1.1).

Theorem 1. *Let $p,q,r \ge 1$, $\alpha \in \Big(\frac{1+[d/p]}{1-[1/r]}, 2\Big]$. Then, for all $\gamma \in \Big(\frac{3-\alpha+[d/p]+[\alpha/r]}{2}, 1\Big)$, for all $x \in \mathbb{R}^d$ the martingale problem with data $\big(L^\alpha, F, x\big)$ is well posed in the sense of Definition 1.*

Remark 1 (On space-time integrability.)**.** *Observe that the constraint $\alpha \in (\frac{1+[d/p]}{1-[1/r]}, 2]$ imposes that $\frac{1+[d/p]}{1-[1/r]} < 2 \iff \frac{d}{p} + \frac{2}{r} < 1$, which is precisely as well the constraint appearing for the relation between time and space integrability in the work by Krylov and Röckner [KR05] which addresses strong uniqueness in the Brownian case for drifts locally in $\mathbb{L}^r(\mathbb{R}_+,\mathbb{L}^p) = \mathbb{L}^r(\mathbb{R}_+,\mathbb{B}^0_{p,p})$.*

As a consequence of the proof of Theorem 1 we also derive the following corollary.

Corollary 2 (Krylov type estimates and associated density bounds for the canonical process)**.** *Under the previous assumptions, the following Krylov type estimate holds for the canonical process $(X_t)_{t\ge0}$. Define:*

$$\theta = \gamma - 1 + \alpha - \frac{d}{p} - \frac{\alpha}{r}. \tag{1.10}$$

For all $f \in C^\infty$,

$$|\mathbb{E}^{\mathbb{P}^\alpha}[\int_0^T f(s,X_s)ds]| \le C\|f\|_{\mathbb{L}^r([0,T],\mathbb{B}^{\theta-\alpha}_{p,q})}. \tag{1.11}$$

with $r > \alpha/(\theta - d/p) > 1$ and $T > 0$. This in particular implies that X_t admits for almost all $t > 0$ a density $\mathfrak{p}_\alpha(\cdot,x,\cdot) : (t,y) \mapsto \mathfrak{p}_\alpha(t,x,y)$ in $\mathbb{L}^{r'}([0,T],\mathbb{B}^{-\theta+\alpha}_{p',q'})$ with $1/m + 1/m' = 1$, $m \in \{p,q,r\}$.

Remark 2. *Note that there is no constraint on the parameter q. This comes from the fact that such a parameter does not play any role in the estimate. The density $\mathfrak{p}_\alpha$ thus belongs to $\mathbb{L}^{r'}([0,T],\mathbb{B}^{-\theta+\alpha}_{p',\infty})$.*

*We emphasize that this estimate seems to be not optimal for us. Roughly speaking, the expected regularity should be the one needed to define pointwise the gradient of the solution of the associated PDE (1.9). As suggested by the analysis done in point **(i)** of Section 3.3, one may be able to prove that the density belongs to $\mathbb{L}^{r'}([0,T],\mathbb{B}^{\alpha+\gamma-2-\alpha/r-d/p}_{p',\infty})$. Note that when $p = r = \infty$, this threshold is, at least formally, the one that could be obtained through the result of Debussche and Fournier [DF13] where density estimates for (time homogeneous) stable driven SDEs with Hölder diffusion coefficients and bounded measurable drifts are obtained. We refrain to go further in that direction as such estimate is not the main concern of our work.*

The following theorem connects the solution of the martingale problem with the dynamics of the *formal* SDE (1.1). Namely, it specifies, in our current singular framework, how the dynamics of (1.1) has to be understood. We decompose it into two terms: the first one is the driving α-stable process and the other one is a drift obtained as the stochastic-Young limit of a regularized version of the initial drift by the density of the driving process.

Theorem 3. *If we now reinforce the assumptions of Theorem 1, assuming*

$$\gamma \in \left(\frac{3 - \alpha + [2d/p] + [2\alpha/r]}{2}, 1 \right), \tag{1.12}$$

it then holds that:

$$X_t = x + \int_0^t \mathscr{F}(s, X_s, ds) + \mathcal{W}_t, \tag{1.13}$$

where for any $0 \leq v \leq s \leq T$, $x \in \mathbb{R}^d$,

$$\mathscr{F}(v, x, s - v) \quad = \quad \int_v^s dr \int_{\mathbb{R}^d} dy F(r, y) p_\alpha(r - v, y - x), \tag{1.14}$$

with p_α the (smooth) density of $\mathcal{W}$ and where the integral in (1.13) *is understood as a $\mathbb{L}^\ell$ limit of the associated Riemann sum (called $\mathbb{L}^\ell$ stochatic-Young integral), $1 \leq \ell < \alpha$.*

Corollary 4 (Associated $\mathbb{L}^\ell$ stochastic-Young integral, $1 \leq \ell < \alpha$)**.** *Under the above assumptions, for any $1 \leq \ell < \alpha$ one can define a stochastic-Young integral w.r.t. the quantities in* (1.13)*. Namely, for any $1 \leq \ell < \alpha$, there exist $1 \leq q < \ell$ and $q' \geq 1$ satisfying $1/q' + 1/q = 1/\ell$ such that for any predictable process $(\psi_s)_{s \in [0,t]}$, $(1 - 1/\alpha - \varepsilon_2)$-Hölder continuous in $\mathbb{L}^{q'}$ with $0 < \varepsilon_2 < (\theta - 1)/\alpha$, one has*

$$\int_0^t \psi_s dX_s = \int_0^t \psi_s \mathscr{F}(s, X_s, ds) + \int_0^t \psi_s d\mathcal{W}_s. \tag{1.15}$$

Eventually, in the particular case $d = 1$, we are able to derive pathwise uniqueness for the solution of (1.1) under suitable conditions. We hence recover and generalize part of the previous existing results of Bass and Chen [BC01] and Athreya *et al.* [ABM18].

Theorem 5. *Under the assumption of Theorem 3, when $d = 1$, pathwise uniqueness holds for the formal equation* (1.1)*, i.e. two weak solutions $(X, \mathcal{W})$ and $(X', \mathcal{W})$ satisfying* (1.13) *are a.s. equal.*

Remark 3. *Pay attention that, in the above result, we do not claim that strong uniqueness holds. This mainly comes from a measurability argument. In [ABM18], the Authors built the drift as a Dirichlet process and then recover the noise part of the dynamics as the difference between the solution and the drift allowing them in turn to work under a more standard framework (in term of measurability), and thus to use the Yamada-Watanabe Theorem. Here, we mainly recover the noise in a canonical way, through the martingale problem, and then build the drift as the difference between the solution and the noise. Such a construction allows us to give a precise meaning to the drift and the loss of measurability can be seen as the price to pay for it. Nevertheless, at this stage, one may restart with the approach of Athreya* et al. *[ABM18] to define an ad hoc noise as the difference between the process and the drift (which reads as a Dirichlet process), identify the objects obtained with the two approaches and then obtain suitable measurability conditions to apply the Yamada-Watanabe Theorem.*

Notations. Throughout the document, we denote by $c, c' ...$ some positive constants depending on the non-degeneracy constant κ in **(UE)** and on the set of parameters $\{\alpha, p, q, r, \gamma\}$. The notation $C, C' ...$ is used when the constants also depend in a non-decreasing way on time T. Other possible dependencies are also explicitly indicated.

2 Proof of the main results

2.1 The underlying PDE

As underlined by Definition 1, it turns out that the well-posedness of the martingale problem associated with (1.1) heavily relies on the construction of a suitable theory for the Cauchy problem (1.9). Hence, we start this part by introducing, for data $T > 0$, $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$, the following *formal* Cauchy problem:

$$\begin{aligned} \partial_t u(t, x) + L^\alpha u(t, x) + F(t, x) \cdot Du(t, x) \quad &= \quad f(t, x), \quad \text{on } [0, T] \times \mathbb{R}^d, \\ u(T, x) \quad &= \quad g(x), \quad \text{on } \mathbb{R}^d, \end{aligned} \tag{2.1}$$

with L^α as in (1.5). Obviously, as is, it is not clear that the scalar product $F(t,x)\cdot Du(t,x)$ makes sense, and this is why the above PDE is, for the time being, only stated formally. Here, the data f and g are functions belonging to some spaces to be specified later on.

The aim of this section is to provide a "$(p,q,r,\gamma)-$ well posedness theory" for the PDE (2.1) which will in turn allow us to establish our main results for the *formal* SDE (1.1). As a key intermediate tool we need to introduce what we will later on call the *mollified* PDE associated with (2.1). Namely, denoting by $(F_m)_{m\in\mathbb{N}}$ as sequence of smooth functions such that $\|F-F_m\|_{\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q})}\to 0$ when $m\to\infty$, we introduce the *mollified* PDE:

$$\begin{array}{rcl}\partial_t u_m(t,x)+L^\alpha u_m(t,x)+F_m(t,x)\cdot Du_m(t,x) &=& f(t,x),\quad \text{on } [0,T]\times\mathbb{R}^d,\\ u_m(T,x) &=& g(x),\quad \text{on } \mathbb{R}^d,\end{array}\tag{2.2}$$

for which we are able to obtain the following controls.

Proposition 6. *Let f,g be smooth functions where g has as well at most linear growth. Let $(u_m)_{m\geq 0}$ denote the sequence of classical solutions of the* mollified *PDE* (2.2)*. It satisfies that*

$$\forall\, p,q,r\geq 1,\ \forall\,\alpha\in\left(\frac{1+\frac{d}{p}}{1-\frac{1}{r}},2\right],\ \forall\gamma\in\left(\frac{3-\alpha+\frac{d}{p}+\frac{\alpha}{r}}{2},1\right),\tag{2.3}$$

recalling from (1.10) *that $\theta-1:=\gamma-2+\alpha-d/p-\alpha/r>0$, there exist positive constants $C:=C(\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})})$, $C_T:=C(T,\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})})$, $\varepsilon>0$ depending on the known parameters γ,p,q,r and κ in* **(UE)**, *s.t. for all $m\geq 0$*

$$\begin{array}{rcl}|u_m(t,x)| &\leq& C(1+|x|),\\ \|Du_m\|_{\mathbb{L}^\infty\left(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}\right)} &\leq& C_T(\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}}+\|f\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-\alpha}_{\infty,\infty})}),\\ \forall 0\leq t\leq s\leq T,\ x\in\mathbb{R}^d,\ |u_m(t,x)-u_m(s,x)| &\leq& C|t-s|^{\frac{\theta}{\alpha}},\ |Du_m(t,x)-Du_m(s,x)|\leq C|t-s|^{\frac{\theta-1}{\alpha}},\end{array}\tag{2.4}$$

where $\varepsilon<<1$ can be chosen as small as desired and $T\mapsto C_T$ is a non-decreasing function.

Remark 4 (About the Schauder type estimates)**.** *Let us first mention that, when the terminal condition g is bounded, then the solution u_m is itself bounded, i.e. $|u_m(t,x)|\leq C$. Note as well that, w.r.t. to the expected parabolic bootstrap, for $s\in[0,T)$, to give some sense to the product $\int_{\mathbb{R}^d}[Du_m\cdot F_m](s,y)dy$ (as a spatial Young integral) uniformly in m, we must have $-1+\gamma+(-1+\gamma+\alpha)>1\iff\gamma>(3-\alpha)/2$. The additional constraint $\gamma>(3-\alpha)/2+(d/p+\alpha/r)/2$ stands, in some sense, as a compensation for the lack of boundedness in time and space of the drift and disappears when $p=q=r=\infty$.*

Corollary 7 (Zvonkin type theory for the *mollified* PDE)**.** *Let k in $\{1,\dots,d\}$ and consider the* mollified *PDE* (2.2) *with terminal condition $g\equiv 0$ and source $f=F^k_m$ (the k^{th} component of F_m). Under the above assumptions, there exists a positive constant $C_T:=C(T,\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})})$ s.t. for each k and all $m\geq 0$, the sequence of classical solutions $(u^k_m)_{m\geq 0}$ of the* mollified *PDE* (2.2) *satisfies:*

$$\|u^k_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}+\|Du^k_m\|_{\mathbb{L}^\infty\left(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}\right)}\leq C_T,\tag{2.5}$$

where $C_T\downarrow 0$ when $T\downarrow 0$. Moreover, there exists $C:=C(T,\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma)}_{p,q})})>0$ such that (2.4) *holds.*

Remark 5 (On the spatial smoothness of the mollified PDE)**.** *From the conditions on γ,α and the definition of θ in* (1.10)*, we carefully point out that:*

$$\theta=\gamma-1+\alpha-\frac{d}{p}-\frac{\alpha}{r}>1.$$

This reflects the spatial smoothness of the underlying PDE. In particular, the condition $\theta>1$ provides a pointwise gradient estimate for the solution of the mollified PDE. This key condition rewrites: $\theta>1\iff\gamma-2+\alpha-[d/p]-[\alpha/r]>0$. It will be implied assuming that $\gamma>[3-\alpha+d/p+\alpha/r]/2$, since in this case $[3-\alpha+d/p+\alpha/r]/2-2+\alpha-[d/p]-[\alpha/r]>0\iff\alpha>[1+d/p]/[1-1/r]$.

Remark 6. *Of course, to derive strong well-posedness in the multidimensional setting some controls of the second order derivatives are needed. This is what Krylov and Röckner do in [KR05] in the Sobolev setting. Let us also specify that, in connection with Theorem 5 and Remark 3, in the scalar setting weak and strong*

uniqueness are somehow closer since, from the PDE viewpoint, they do not require to go up to second order derivatives. Indeed, the strategy is then to develop for two weak solutions X^1X^2 of (1.15), a regularized version of $|X^1_t - X^2_t|$, which somehow makes appear a kind of "local-time" term which is handled through the Hölder controls on the gradients (see the proof of Theorem 5 and e.g. Proposition 2.9 in [ABP18]), whereas in the multidimensional setting, for strong uniqueness, the second derivatives get in.

This, in turn, allows us to derive a well-posedness theory, in the mild sense, for the *formal* PDE (2.1) summarized in the following theorem.

Theorem 8 (Mild theory for the associated PDE and Schauder type regularization result)**.** *Let the assumptions of Theorem 1 hold. For θ defined in* (1.10)*, so that in particular $\theta - 1 > 0$, we assume that g has linear growth and $Dg \in \mathbb{B}^{\theta-1}_{\infty,\infty}$ and $f \in \mathbb{L}^\infty([0,T], \mathbb{B}^{\theta-\alpha}_{\infty,\infty})$. Then, the PDE* (2.1) *admits a unique mild solution which admits the following Duhamel type representation*

$$u(t,x) = P^\alpha_{T-t}[g](x) + \int_t^T ds P^\alpha_{s-t}[\{f + F \cdot Du\}](s,x),$$

with P^α the semi-group generated by L^α. Furthermore, the unique mild solution satisfies the bounds (2.4) *of Proposition 6 (replacing u_m by u).*

Remark 7 (On the corresponding parabolic bootstrap)**.** *Observe that, when $p = r = +\infty$, we almost have a Schauder type result, namely $\theta = \gamma - 1 + \alpha$ in* (2.4) *and we end up with the corresponding parabolic bootstrap effect for both the solution of the mollified PDE* (2.2) *and the mild solution of* (2.1)*, up to the small exponent ε which can be chosen arbitrarily small.*

Remark 8 (About additional diffusion coefficients)**.** *It should be noted at this point that we are confident about the extension of the results to differential operator L^α involving non-trivial diffusion coefficient, provided this last is Hölder-continuous in space. Sketches of proofs in this direction are given in the Remark 14 following the proof of Proposition 6, Theorem 8 and Corollary 7. However, we avoid investing this direction for sake of clarity and in order to focus on the more (unusual) drift component.*

2.2 From PDE to SDE results

We here state the procedure to go from the "(p,q,r,γ) − well posedness theory" for the PDE (2.2) deriving from Proposition 6, Corollary 7 and Theorem 8 to the corresponding one for the SDE.

It is quite standard to derive well-posedness results for a probabilistic problem through PDE estimates. When the drift is a function, such a strategy goes back to e.g. Zvonkin [Zvo74] or Stroock and Varadhan [SV79]. Such strategy has been made quite systematic in the distributional setting by Delarue and Diel in [DD16] who provide a very robust framework. To investigate the meaning and well-posedness of (1.1), we adapt their procedure to the current setting.

Points **(i)** to **(iii)** allow to derive the rigorous proof of Theorem 1 provided Proposition 6, Corollary 7 and Theorem 8 hold. Point **(iv)** concerns the meaning of the *formal* dynamics (1.1) and gives some highlights to the (more involved) proof of Theorem 3. Eventually, we explain in point **(v)** how the PDE results obtained in Proposition 6, Corollary 7 and Theorem 8 can be used to derive the pathwise uniqueness for the *formal* SDE (1.1) (or more precisely for the stochastic dynamical system obtained in point **(iv)**). This gives a flavor of the proof of Theorem 5.

(i) Tightness of the sequence of probability measure induced by the solution of the mollified SDE (1.1)**.** Here, we consider the regular framework induced by the *mollified* PDE (2.2). Note that in this regularized framework, for any m, the martingale problem associated with L^α_m is well posed. We denote by $\mathbb{P}^\alpha_m$ the associated solution. Let us generically denote by $(X^m_t)_{t\geq 0}$ the associated canonical process. Note that the underlying space where such a process is defined differs according to the values of α: when $\alpha = 2$ the underlying space is $\mathcal{C}([0,T], \mathbb{R}^d)$ while it is $\mathcal{D}([0,T], \mathbb{R}^d)$ when $\alpha < 2$.

Assume w.l.o.g. $s > v$, let $u_m = (u^1_m, \dots, u^d_m)$ where each u^i_m is the solution of (2.2) with terminal condition $g \equiv 0$ and source term $f = F^k_m$ (i.e. the k^{th} component of F_m). Let us define for any $s \geq v$ in $[0,T]^2$ and for

any $\alpha \in (1,2]$ the process

$$M_{v,s}(\alpha, u_m, X^m) = \begin{cases} \int_v^s Du_m(r, X_r^m) \cdot dW_r, \\ \text{where } W \text{ is a Brownian motion, if } \alpha = 2; \\ \int_v^s \int_{\mathbb{R}^d \setminus \{0\}} \{u_m(r, X_{r^-}^m + x) - u_m(r, X_{r^-}^m)\} \tilde{N}(dr, dx), \\ \text{where } \tilde{N} \text{ is the compensated Poisson measure, if } \alpha < 2. \end{cases} \tag{2.6}$$

Note that this process makes sense since the solution u_m of the *mollified* PDE (2.2) is bounded. Next, applying Itô's formula we obtain

$$X_s^m - X_v^m \quad = \quad M_{v,s}(\alpha, u_m, X^m) + \mathcal{W}_s - \mathcal{W}_v - [u_m(v, X_v^m) - u_m(s, X_s^m)]. \tag{2.7}$$

In order to prove that $(\mathbb{P}_m^\alpha)_{m\in\mathbb{N}}$ actually forms a tight sequence of probability measures on $\mathcal{C}([0,T], \mathbb{R}^d)$ (resp. on $\mathcal{D}([0,T], \mathbb{R}^d)$), it is sufficient to prove that there exists $c, \tilde{p}$ and $\eta > 0$ such that $\mathbb{E}^{\mathbb{P}_m^\alpha}[|X_s^m - X_v^m|^{\tilde{p}}] \leq c|v-s|^{1+\eta}$ (resp. $\mathbb{E}^{\mathbb{P}_m^\alpha}[|X_s^m - X_0^m|^{\tilde{p}}] \leq cs^\eta$) thanks to the Kolmogorov (resp. Aldous) Criterion. We refer e.g. for the latter to Proposition 34.9 in Bass [Bas11]. Writing

$$[u_m(v, X_v^m) - u_m(s, X_s^m)] = u_m(v, X_v^m) - u_m(v, X_s^m) + u_m(v, X_s^m) - u_m(s, X_s^m), \tag{2.8}$$

the result follows in small time thanks to Corollary 7 (choosing $1 < \tilde{p} < \alpha$ in the pure jump setting) and (2.4) for the regularity in time.

(ii) Identification of the limit probability measure. Let us now prove that the limit is indeed a solution of the martingale problem associated with L^α. Let $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be some measurable, continuous in time and bounded in space function, let u_m be the classical solution of the *mollified* PDE (2.2) with source term f and terminal condition $g \equiv 0$. Applying Itô's Formula for each $u_m(t, X_t^m)$ we obtain that

$$u_m(t, X_t^m) - u_m(0, x_0) - \int_0^t f(s, X_s^m) ds = M_{0,t}(\alpha, u_m, X^m),$$

where $M(\alpha, u_m, X^m)$ is defined by (2.6). From this definition, if we are able to control uniformly in m the modulus of continuity of u_m and of Du_m, then from Arzelà -Ascoli Theorem, we know that we can extract a subsequence $(m_k)_{k\geq 0}$ s.t. $(u_{m_k})_{k\geq 0}$ and $(Du_{m_k})_{m_k \geq 0}$ converge uniformly on compact subsets of $[0,T] \times \mathbb{R}^d$ to functions u and Du respectively.

In particular, equation (3.25) holds for the limit functions u, Du. Hence, this implies that u is the unique mild solution of PDE (2.1). Thus, together with a uniform control of the moment of X^m (which also follows from (2.7) and above conditions on u_m), we deduce that

$$\left(u(t, X_t) - \int_0^t f(s, X_s) ds - u(0, x_0) \right)_{0 \leq t \leq T}, \tag{2.9}$$

is a $\mathbb{P}^\alpha$-martingale (square integrable when $\alpha = 2$) by letting the regularization procedure tend to the infinity.

(iii) Uniqueness of the limit probability measure. We now come back to the canonical space (which again depends on the current value of α), and let $\mathbb{P}^\alpha$ and $\tilde{\mathbb{P}}^\alpha$ be two solutions of the martingale problem associated with data (L^α, F, x_0), x_0 in $\mathbb{R}^d$. Thus, for all continuous in time and measurable and bounded in space functions $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ we have, setting again $g \equiv 0$, from Theorem 8

$$u(0, x_0) = \mathbb{E}^{\mathbb{P}^\alpha} \left[\int_0^T f(s, X_s) ds \right] = \mathbb{E}^{\tilde{\mathbb{P}}^\alpha} \left[\int_0^T f(s, X_s) ds \right],$$

so that the marginal laws of the canonical process are the same under $\mathbb{P}^\alpha$ and $\tilde{\mathbb{P}}^\alpha$. We extend the result on $\mathbb{R}_+$ thanks to regular conditional probabilities, see Chapter 6.2 in [SV79] . Uniqueness then follows from Corollary 6.2.4 of [SV79].

(iv) Reconstructing the dynamics associated with the *formal* SDE (1.1)**.** This part requires to introduce an enhanced martingale problem (considering $(X, \mathcal{W})$ as canonical process). Working within this

enlarged setting allows to recover the drift part of the dynamics by studying the difference between the increment of the process and the associated stable noise on small time intervals which are further meant to be infinitesimal. It turns out that, for any time interval $[v,s]$ considering $\mathfrak{f}(v,X_v,s-v) := u^s(v,X_v) - X_v$ (where we denote by u^s the solution of (2.1) on the time-interval $[0,s]$, with terminal condition $u^s(s,x) = x$ and $f = 0$), which can be expanded as in (1.14), we establish sufficient quantitive controls to be able to give a meaning through stochastic-Young type integration to $\int_0^t \mathfrak{f}(v,X_v,dv)$ which in turns is the limit drift of the dynamics. Observe as well from (1.14) that, on a small time interval, the Euler approximation of the drift writes as $\mathscr{F}(v,X_v,s-v) = \int_v^s dr \int_{\mathbb{R}^d} dy F(r,y) p_\alpha(r-v,y-X_v)$ which is nothing but the convolution of the initial distributional drift with the density of the driving noise. Importantly, this limit drift is also a Dirichlet process. We eventually mention that, the previous explicit representation of the drift could also be useful in order to derive numerical approximations for the SDE (1.13). We can to this end mention the recent work by De Angelis *et al.* [DGI19] who considered in the Brownian scalar case some related issues.

(v) About the strong well-posedness for (1.13)**.** Having at hand a representation for the dynamics, it is tempting to wonder if pathwise uniqueness holds for the SDE (1.13) or even if it admits a strong solution. As mentioned in the previous point, we know in particular that the drift part in (1.13) reads as a Dirichlet process. The point is then to apply the Itô formula for Dirichlet processes to expand any weak solution of (1.13) along the solution of the *mollified* PDE (2.2) with source term F_m and terminal condition 0. This yields to

$$X_t^{Z,m} := X_t - u_m(t,X_t) = x - u_m(0,x) + \mathcal{W}_t - M_{0,t}(\alpha,u_m,X) + R_{0,t}(\alpha,F_m,\mathscr{F},X). \tag{2.10}$$

where $M_{0,t}(\alpha,u_m,X)$ is as in (2.6) with X instead of X^m therein and $R_{0,t}(\alpha,F_m,\mathscr{F},X) := \int_0^t \mathscr{F}(s,X_s,ds) - F_m(s,X_s)ds$. This strategy is slightly different from the one implemented in [ABM18], where the authors consider the limit PDE itself getting rid of the remainder term $R_{0,t}(\alpha,F_m,\mathscr{F},X)$. We feel our approach is more adapted with the way we reconstruct the dynamics of the drift. The remainder will indeed be handled through some *a priori* controls for the drift leandig to pathwise uniqueness. Again, the by-product is that we are faced with some measurability issues (see also Remark 3). We think that reproducing the strategy of [ABM18] would lead to strong well-posedness as soon as the parameters satisfy the previous condition (1.12) in Theorem 3.

3 PDE analysis

This part is dedicated to the proofs of Proposition 6, Corollary 7 and Theorem 8. It is thus the core of this paper as these results allow to recover, specify and extend, most of the previous results on SDEs with distributional drifts discussed in the introduction. Especially, as they are handled, the proofs are essentially the same in the diffusive ($\alpha = 2$) and pure jump ($\alpha < 2$) setting as they only require heat kernel type estimates on the density of the associated underlying noise. We first start by introducing the mathematical tools in Section 3.1. Then, we provide a primer on the PDE (2.1) by investigating the smoothing properties of the Green kernel associated with the stable noise in Section 3.2. Eventually, we derive in Section 3.3 the proofs of Proposition 6, Corollary 7 and Theorem 8. We importantly point out that, from now on and in all the current section, we assume without loss of generality that $T \leq 1$.

3.1 Mathematical tools

In this part, we give the main mathematical tools needed to prove Proposition 6 and Theorem 8.

3.1.1 Heat kernel estimates for the density of the driving process.

Under **(UE)**, it is rather well known that the following properties hold for the density p_α of $\mathcal{W}$. For the sake of completeness we provide a complete proof.

Lemma 9 (Bounds and Sensitivities for the stable density)**.** *There exists $C := C((\mathbf{A}))$ s.t. for all $\ell \in \{1,2\}$, $t > 0$, and $y \in \mathbb{R}^d$:*

$$|D_y^\ell p_\alpha(t,y)| \leq \frac{C}{t^{\ell/\alpha}} q_\alpha(t,y), \ |\partial_t^\ell p_\alpha(t,y)| \leq \frac{C}{t^\ell} q_\alpha(t,y), \tag{3.1}$$

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

$$\int_{\mathbb{R}^N} q_\alpha(t,y)|y|^\gamma dy \leq C_\gamma t^{\frac{\gamma}{\alpha}}, \quad t > 0. \tag{3.2}$$

Remark 9. *From now on, for the family of stable densities* $\big(q(t,\cdot)\big)_{t>0}$*, we also use the notation* $q(\cdot) := q(1,\cdot)$*, i.e. without any specified argument* $q(\cdot)$ *stands for the density* $q(t,\cdot)$ *at time* $t = 1$*.*

Proof. We focus here on the pure jump case $\alpha \in (1,2)$. Indeed, for $\alpha = 2$ the density of the driving Brownian motion readily satisfies the controls of (3.1) with q_α replaced by a suitable Gaussian density.

Let us recall that, for a given fixed $t > 0$, we can use an Itô-Lévy decomposition at the associated characteristic stable time scale for $\mathcal{W}$ (i.e. the truncation is performed at the threshold $t^{\frac{1}{\alpha}}$) to write $\mathcal{W}_t := M_t + N_t$ where M_t and N_t are independent random variables. More precisely,

$$N_s = \int_0^s \int_{|x|>t^{\frac{1}{\alpha}}} xN(du,dx), \quad M_s = \mathcal{W}_s - N_s, \;\; s \geq 0, \tag{3.3}$$

where N is the Poisson random measure associated with the process $\mathcal{W}$; for the considered fixed $t > 0$, M_t and N_t correspond to the *small jumps part* and *large jumps part* respectively. A similar decomposition has been already used in [Wat07], [Szt10] and [HM16], [HMP19] (see in particular Lemma 4.3 therein). It is useful to note that the cutting threshold in (3.3) precisely yields for the considered $t > 0$ that:

$$N_t \overset{(\text{law})}{=} t^{\frac{1}{\alpha}} N_1 \;\text{ and }\; M_t \overset{(\text{law})}{=} t^{\frac{1}{\alpha}} M_1. \tag{3.4}$$

To check the assertion about N we start with

$$\mathbb{E}[e^{i\langle \lambda, N_t\rangle}] = \exp\Big(t \int_{\mathbb{S}^{d-1}} \int_{t^{\frac{1}{\alpha}}}^{\infty} \Big(\cos(\langle \lambda, r\xi\rangle) - 1\Big) \frac{dr}{r^{1+\alpha}} \mu_S(d\xi)\Big), \;\; \lambda \in \mathbb{R}^d$$

(see [Sat99]). Changing variable $r/t^{1/\alpha} = s$ we get that $\mathbb{E}[e^{i\langle \lambda, N_t\rangle}] = \mathbb{E}[e^{i\langle \lambda, t^{1/\alpha} N_1\rangle}]$ for any $\lambda \in \mathbb{R}^d$ and this shows the assertion (similarly we get the statement for M). The density of $\mathcal{W}_t$ then writes

$$p_\alpha(t,x) = \int_{\mathbb{R}^d} p_M(t, x-\xi) P_{N_t}(d\xi), \tag{3.5}$$

where $p_M(t,\cdot)$ corresponds to the density of M_t and P_{N_t} stands for the law of N_t. From Lemma A.2 in [HMP19] (see as well Lemma B.1 in [HM16]), $p_M(t,\cdot)$ belongs to the Schwartz class $\mathscr{S}(\mathbb{R}^N)$ and satisfies that for all $m \geq 1$ and all $\ell \in \{0,1,2\}$, there exist constants $\bar{C}_m,\; C_m$ s.t. for all $t > 0,\; x \in \mathbb{R}^d$:

$$|D_x^\ell p_M(t,x)| \leq \frac{\bar{C}_m}{t^{\frac{\ell}{\alpha}}} p_{\bar{M}}(t,x), \;\text{ where }\; p_{\bar{M}}(t,x) := \frac{C_m}{t^{\frac{d}{\alpha}}} \left(1 + \frac{|x|}{t^{\frac{1}{\alpha}}}\right)^{-m} \tag{3.6}$$

where C_m is chosen in order that $p_{\bar{M}}(t,\cdot)$ *be a probability density.*

We carefully point out that, to establish the indicated results, since we are led to consider potentially singular spherical measures, we only focus on integrability properties similarly to [HMP19] and not on pointwise density estimates as for instance in [HM16]. The main idea thus consists in exploiting (3.3), (3.5) and (3.6). The derivatives on which we want to obtain quantitative bounds will be expressed through derivatives of $p_M(t,\cdot)$, which also give the corresponding time singularities. However, as for general stable processes, the integrability restrictions come from the large jumps (here N_t) and only depend on its index α. A crucial point then consists in observing that the convolution $\int_{\mathbb{R}^d} p_{\bar{M}}(t, x-\xi) P_{N_t}(d\xi)$ actually corresponds to the density of the random variable

$$\bar{\mathcal{W}}_t := \bar{M}_t + N_t, \;\; t > 0 \tag{3.7}$$

(where $\bar{M}_t$ has density $p_{\bar{M}}(t,.)$ and is independent of N_t; to have such decomposition one can define each $\bar{\mathcal{W}}_t$ on a product probability space). Then, the integrability properties of $\bar{M}_t + N_t$, and more generally of all random variables appearing below, come from those of $\bar{M}_t$ and N_t.

One can easily check that $p_{\bar{M}}(t,x) = t^{-\frac{d}{\alpha}} p_{\bar{M}}(1, t^{-\frac{1}{\alpha}} x)$, $t > 0,\; x \in \mathbb{R}^d$. Hence

$$\bar{M}_t \overset{(\text{law})}{=} t^{\frac{1}{\alpha}} \bar{M}_1, \quad N_t \overset{(\text{law})}{=} t^{\frac{1}{\alpha}} N_1.$$

By independence of $\bar{M}_t$ and N_t, using the Fourier transform, one can easily prove that

$$\bar{\mathcal{W}}_t \overset{(\text{law})}{=} t^{\frac{1}{\alpha}} \bar{\mathcal{W}}_1. \tag{3.8}$$

Moreover, $\mathbb{E}[|\bar{\mathcal{W}}_t|^\gamma] = \mathbb{E}[|\bar{M}_t + N_t|^\gamma] \leq C_\gamma t^{\frac{\gamma}{\alpha}} (\mathbb{E}[|\bar{M}_1|^\gamma] + \mathbb{E}[|N_1|^\gamma]) \leq C_\gamma t^{\frac{\gamma}{\alpha}},\; \gamma \in (0,\alpha)$. This shows that the density of $\bar{\mathcal{W}}_t$ verifies (3.2). The controls on the spatial derivatives are derived similarly using (3.6) for $\ell \in \{1,2\}$ and the same previous argument. The bound for the time derivatives follow from the Kolmogorov equation $\partial_t p_\alpha(t,z) = L^\alpha p_\alpha(t,z)$ and (3.5) using the fact that for all $x \in \mathbb{R}^d,\; |L^\alpha p_M(t,x)| \leq C_m t^{-1} \bar{p}_M(t,x)$ (see again Lemma 4.3 in [HMP19] for details). ∎

3.1.2 Thermic characterization of Besov norm.

In the sequel, we will intensively use the thermic characterisation of Besov spaces, see e.g. Section 2.6.4 of Triebel [Tri83]. Precisely, for $\vartheta \in \mathbb{R}, q \in (0,+\infty], p \in (0,\infty]$, $\mathbb{B}^{\vartheta}_{p,q}(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{\mathcal{H}^{\vartheta}_{p,q},\alpha} < +\infty\}$ where $\mathcal{S}(\mathbb{R}^d)$ stands for the Schwartz class and

$$\|f\|_{\mathcal{H}^{\vartheta}_{p,q},\alpha} := \|\varphi(D)f\|_{\mathbb{L}^p(\mathbb{R}^d)} + \Big(\int_0^1 \frac{dv}{v} v^{(n-\frac{\vartheta}{\alpha})q} \|\partial_v^n \tilde{p}_\alpha(v,\cdot) \star f\|^q_{\mathbb{L}^p(\mathbb{R}^d)}\Big)^{\frac{1}{q}}, \tag{3.9}$$

with $\varphi \in C_0^\infty(\mathbb{R}^d)$ (smooth function with compact support) s.t. $\varphi(0) \neq 0$, $\varphi(D)f := (\varphi \hat{f})^\vee$ where $\hat{f}$ and $(\varphi\hat{f})^\vee$ respectively denote the Fourier transform of f and the inverse Fourier transform of $\varphi\hat{f}$. The parameter n is an integer s.t. $n > \vartheta/\alpha$ and for $v > 0$, $z \in \mathbb{R}^d$, $\tilde{p}_\alpha(v,\cdot)$ denotes the density of the d-dimensional isotropic stable process at time v. In particular $\tilde{p}_v$ satisfies the bounds of Lemma 9 and in that case the upper-bounding density can be specified. Namely, in that case (3.1) holds with $q_\alpha(t,x) = C_\alpha t^{-d/\alpha}(1+|x|/t^{1/\alpha})^{-(d+\alpha)}$.

Importantly, it is well known that $\mathbb{B}^{\vartheta}_{p,q}(\mathbb{R}^d,\mathbb{R})$ and $\mathbb{B}^{-\vartheta}_{p',q'}(\mathbb{R}^d,\mathbb{R})$ where p',q' are the conjugates of p,q respectively are in duality. Namely, for $(p,q) \in (1,\infty]^2$, $\mathbb{B}^{\vartheta}_{p,q} = (\mathbb{B}^{-\vartheta}_{p',q'})^*$, see e.g. Theorem 4.1.3 in [AH96] or Proposition 3.6 in [LR02]. In particular, for all $(f,g) \in \mathbb{B}^{\vartheta}_{p,q}(\mathbb{R}^d,\mathbb{R}) \times \mathbb{B}^{-\vartheta}_{p',q'}(\mathbb{R}^d,\mathbb{R})$ which are also functions:

$$|\int_{\mathbb{R}^d} f(y)g(y)dy| \leq \|f\|_{\mathbb{B}^{\vartheta}_{p,q}} \|g\|_{\mathbb{B}^{-\vartheta}_{p',q'}}. \tag{3.10}$$

In the following we call *thermic part* the second term in the right hand side of (3.9). This contribution will be denoted by $\mathcal{T}^{\vartheta}_{p,q}[f]$.

Remark 10. *As it will be clear in the following, the first part of the r.h.s. in* (3.9) *will be the easiest part to handle (in our case) and will give negligible contributions. For that reason, we will only focus on the estimation of the* thermic part *of the Besov norm below. See Remark 15 in the proof of Lemma 10 in Appendix A for details.*

3.1.3 Auxiliary estimates

We here provide some useful estimates whose proofs are postponed to Appendix A. We refer to the next Section 3.2 for a flavor of those proofs as well as for applications of such results.

Lemma 10. *Let $\Psi : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$. Assume that for all s in $[0,T]$ the map $y \mapsto \Psi(s,y)$ is in $\mathbb{B}^{\beta}_{\infty,\infty}(\mathbb{R}^d)$ for some $\beta \in (0,1]$. Define for any α in $(1,2]$, for all $\eta \in \{0,1,\alpha\}$, the differential operator $\mathscr{D}^\eta$ by*

$$\hat{\mathscr{D}}^\eta := \begin{cases} \mathrm{Id} \text{ if } \eta = 0, \\ -i\xi \text{ if } \eta = 1, \\ |\xi|^\alpha \text{ if } \eta = \alpha, \end{cases} \tag{3.11}$$

and let $p_\alpha(t,\cdot)$ be the density of $\mathcal{W}_t$ defined in (3.5)*. Then, there exists a constant $C := C((\mathbf{UE}),T) > 0$ such that for any γ in $(1-\beta,1)$, any $p',q' \geq 1$, all $t < s$ in $[0,T]^2$, for all x in $\mathbb{R}^d$*

$$\|\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot - x)\|_{\mathbb{B}^{1-\gamma}_{p',q'}} \leq \|\Psi(s,\cdot)\|_{\mathbb{B}^{\beta}_{\infty,\infty}} \frac{C}{(s-t)^{\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{\eta}{\alpha}\right]}}, \tag{3.12}$$

where p is the conjugate of p'. Also, for any γ in $(1-\beta,1]$ all $t < s$ in $[0,T]^2$, for all x,x' in $\mathbb{R}^d$ it holds that for all $\beta' \in (0,1)$,

$$\|\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\|_{\mathbb{B}^{1-\gamma}_{p',q'}} \leq \|\Psi(s,\cdot)\|_{\mathbb{B}^{\beta}_{\infty,\infty}} \frac{C}{(s-t)^{\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{\eta+\beta'}{\alpha}\right]}} |x-x'|^{\beta'}, \tag{3.13}$$

up to a modification of C:= $C((\mathbf{UE}),T,\beta')$.

3.2 A primer on PDE (2.1): reading almost optimal regularity through Green kernel estimates

Equation (2.1) can be rewritten as

$$\begin{aligned} \partial_t u(t,x) + L^\alpha u(t,x) &= f(t,x) - F(t,x)\cdot Du(t,x), \quad \text{on } [0,T]\times\mathbb{R}^d, \\ u(T,x) &= g(x), \quad \text{on } \mathbb{R}^d, \end{aligned} \tag{3.14}$$

viewing the first order term as a source (depending here on the solution itself). In order to understand what type of smoothing effects can be expected for rough source we first begin by investigating the smoothness of the following equation:

$$\begin{aligned} \partial_t w(t,x) + L^\alpha w(t,x) &= \Phi(t,x), \quad \text{on } [0,T]\times\mathbb{R}^d, \\ w(T,x) &= 0, \quad \text{on } \mathbb{R}^d, \end{aligned} \tag{3.15}$$

The parallel with the initial problem (2.1), rewritten in (3.14), is rather clear. We will aim at applying the results obtained below for the solution of (3.15) to $\Phi = f - F\cdot Du$ (where the roughest part of the source will obviously be $F\cdot Du$).

Given a map Φ in $\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})$ we now specifically concentrate on the gain of regularity which can be obtained through the fractional operator L^α for the solution w of (3.15) w.r.t. the data Φ. Having a lot of parameters at hand, this will provide a primer to understand what could be, at best, attainable for the *target* PDE (3.14)-(2.1).

The solution of (3.15) corresponds to the Green kernel associated with Φ defined as:

$$G^\alpha\Phi(t,x) = \int_t^T ds \int_{\mathbb{R}^d} dy \Phi(s,y) p_\alpha(s-t, y-x). \tag{3.16}$$

Since to address the well-posedness of the martingale problem we are led to contol, in some sense, gradients, we will here try to do so for the Green kernel introduced in (3.16) solving the linear problem (3.15) with *rough* source. Namely for a multi-index $\eta\in\mathbb{N}^d$, $|\eta| := \sum_{i=1}^d \eta_i \le 1$, we want to control $D^\eta_x G^\alpha\Phi(t,x)$

Avoiding harmonic analysis techniques, which could in some sense allow to average non-integrable singularities, our approach allows to obtain *almost optimal regularity* thresholds that could be attainable on u. Thanks to the Hölder inequality (in time) and the duality on Besov spaces (see equation (3.10)) we have that:

$$\begin{aligned} |D^\eta_x G^\alpha\Phi(t,x)| &= \left| \int_t^T ds \int_{\mathbb{R}^d} dy \Phi(s,y) D^\eta_x p(s-t,y-x)\right| \\ &\le \|\Phi\|_{\mathbb{L}^r((t,T],\mathbb{B}^{-1+\gamma}_{p,q})} \|D^\eta_x p_\alpha(\cdot - t, \cdot - x)\|_{\mathbb{L}^{r'}((t,T],\mathbb{B}^{1-\gamma}_{p',q'})}, \end{aligned}$$

where p', q' and r' are the conjugate exponents of p, q and r. Let us first focus, for $s\in(t,T]$ on the thermic part of $\|D^\eta_x p_\alpha(s-t,\cdot-x)\|_{\mathbb{B}^{1-\gamma}_{p',q'}}$. We have with the notations of Section 3.1.2:

$$\begin{aligned} \left(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot-x)]\right)^{q'} &= \int_0^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot)\star D^\eta_x p_\alpha(s-t,\cdot-x)\|^{q'}_{\mathbb{L}^{p'}} \\ &= \int_0^{(s-t)} \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot)\star D^\eta_x p_\alpha(s-t,\cdot-x)\|^{q'}_{\mathbb{L}^{p'}} \\ &\quad + \int_{(s-t)}^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot)\star D^\eta_x p_\alpha(s-t,\cdot-x)\|^{q'}_{\mathbb{L}^{p'}} \\ &=: \left(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(\cdot-t,\cdot-x)]|_{[0,(s-t)]}\right)^{q'} + \left(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot-x)]|_{[(s-t),1]}\right)^{q'}. \end{aligned}$$

In the above equation, we split the time interval into two parts. On the upper interval, for which there are no time singularities, we use directly convolution inequalities and the available controls for the derivatives of the heat kernel (see Lemma 9). On the lower interval we have to equilibrate the singularities in v and use cancellation techniques involving the sensitivities of $D^\eta_x p_\alpha$ (which again follow from Lemma 9).

Let us begin with the upper part. Using the $\mathbb{L}^1 - \mathbb{L}^{p'}$ convolution inequality, we have from Lemma 9:

$$\begin{aligned} \left(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot-x)]|_{[(s-t),1]}\right)^{q'} &\le \int_{(s-t)}^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot)\|^{q'}_{\mathbb{L}^1} \|D^\eta_x p_\alpha(s-t,\cdot-x)\|^{q'}_{\mathbb{L}^{p'}} \\ &\le \frac{C}{(s-t)^{(\frac{d}{p\alpha}+\frac{|\eta|}{\alpha})q'}} \int_{(s-t)}^1 \frac{dv}{v} \frac{1}{v^{\frac{1-\gamma}{\alpha}q'}} \\ &\le \frac{C}{(s-t)^{[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}]q'}}. \end{aligned} \tag{3.17}$$

Indeed, we used for the second inequality that equation (3.1) and the self similarity of q_α give:

$$\begin{aligned}\|D_x^\eta p_\alpha(s-t,\cdot-x)\|_{\mathbb{L}^{p'}} &= \Big(\int_{\mathbb{R}^d}\big(\partial_x^\eta p_\alpha(s-t,x)\big)^{p'}dx\Big)^{1/p'}\\ &\le \frac{C_{p'}}{(s-t)^{\frac{|\eta|}{\alpha}}}\Big((s-t)^{-\frac{d}{\alpha}(p'-1)}\int_{\mathbb{R}^d}\frac{dx}{(s-t)^{\frac{d}{\alpha}}}\big(q_\alpha(1,\frac{x}{(s-t)^{\frac{1}{\alpha}}})\big)^{p'}\Big)^{1/p'}\\ &\le C_{p'}(s-t)^{-[\frac{d}{\alpha p}+\frac{|\eta|}{\alpha}]}\Big(\int_{\mathbb{R}^d}d\tilde{x}\big(q(1,\tilde{x})\big)^{p'}\Big)^{1/p'} \le \bar{C}_{p'}(s-t)^{-[\frac{d}{\alpha p}+\frac{|\eta|}{\alpha}]}, \quad (3.18)\end{aligned}$$

recalling that $p^{-1}+(p')^{-1}=1$ and $p\in(1,+\infty]$, $p'\in[1,+\infty)$ for the last inequality.

Hence, the map $s\mapsto \mathcal{T}^{1-\gamma}_{p',q'}[D_x^\eta p_\alpha(s-t,\cdot-x)]|_{[(s-t),1]}$ belongs to $\mathbb{L}^{r'}((t,T],\mathbb{R}^+)$ as soon as

$$-r'\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}\right]>-1 \Longleftrightarrow |\eta|<\alpha(1-\frac{1}{r})+\gamma-1-\frac{d}{p}. \quad (3.19)$$

On the other hand, still from (3.1) (see again the proof of Lemma 4.3 in [HMP19] for details), one derives that there exists C s.t. for all $\beta\in(0,1]$ and all $(x,y,z)\in(\mathbb{R}^d)^2$,

$$|D_x^\eta p_\alpha(s-t,z-x)-D_x^\eta p_\alpha(s-t,y-x)|\le\frac{C}{(s-t)^{\frac{\beta+|\eta|}{\alpha}}}|z-y|^\beta\Big(q_\alpha(s-t,z-x)+q_\alpha(s-t,y-x)\Big). \quad (3.20)$$

Indeed, (3.20) is direct if $|z-y|\ge(1/2)(s-t)^{1/\alpha}$ (off-diagonal regime). It suffices to exploit the bound (3.1) for $D_x^\eta p_\alpha(s-t,y-x)$ and $D_x^\eta p_\alpha(s-t,z-x)$ and to observe that $\big(|z-y|/(s-t)^{1/\alpha}\big)^\beta\ge 1$. If now $|z-y|\le(1/2)(s-t)^{1/\alpha}$ (diagonal regime), it suffices to observe from (3.6) that, with the notations of the proof of Lemma 9 (see in particular (3.5)), for all $\lambda\in[0,1]$:

$$\begin{aligned}|D_x^\eta Dp_M(s-t,y-x+\lambda(y-z))| &\le \frac{C_m}{(s-t)^{\frac{|\eta|+1}{\alpha}}}p_{\bar{M}}(s-t,y-x-\lambda(y-z))\\ &\le \frac{C_m}{(s-t)^{\frac{|\eta|+1+d}{\alpha}}}\frac{1}{\Big(1+\frac{|y-x-\lambda(z-y)|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m}\\ &\le \frac{C_m}{(s-t)^{\frac{|\eta|+1+d}{\alpha}}}\frac{1}{\Big(\frac{1}{2}+\frac{|y-x|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m}\le 2\frac{C_m}{(s-t)^{\frac{|\eta|+1}{\alpha}}}p_{\bar{M}}(s-t,y-x).\end{aligned} \quad (3.21)$$

Therefore, in the diagonal case (3.20) follows from (3.21) and (3.5) writing $|D_x^\eta p_\alpha(s-t,z-x)-D_x^\eta p_\alpha(s-t,y-x)|\le\int_0^1 d\lambda|D_x^\eta Dp_\alpha(s-t,y-x+\lambda(y-z))\cdot(y-z)|\le 2C_m(s-t)^{-(|\eta|+1)/\alpha}q_\alpha(s-t,y-x)|z-y|\le\tilde{C}_m(s-t)^{-(|\eta|+\beta)/\alpha}q_\alpha(s-t,y-x)|z-y|^\beta$ for all $\beta\in[0,1]$ (exploiting again that $|z-y|\le(1/2)(s-t)^{1/\alpha}$ for the last inequality). From (3.20) we now derive:

$$\begin{aligned}&\|\partial_v\tilde{p}_\alpha(v,\cdot)\star D_x^\eta p_\alpha(s-t,\cdot-x)\|_{L^{p'}}\\ =\;&\Big(\int_{\mathbb{R}^d}dz\Big|\int_{\mathbb{R}^d}dy\partial_v\tilde{p}_\alpha(v,z-y)D_x^\eta p_\alpha(s-t,y-x)\Big|^{p'}\Big)^{1/p'}\\ =\;&\Big(\int_{\mathbb{R}^d}dz\Big|\int_{\mathbb{R}^d}dy\partial_v\tilde{p}_\alpha(v,z-y)\Big[D_x^\eta p_\alpha(s-t,y-x)-D_x^\eta p_\alpha(s-t,z-x)\Big]\Big|^{p'}\Big)^{1/p'}\\ \le\;&\frac{1}{(s-t)^{\frac{|\eta|+\beta}{\alpha}}}\Big(\int_{\mathbb{R}^d}dz\Big|\int_{\mathbb{R}^d}dy|\partial_v\tilde{p}_\alpha(v,z-y)|\;|z-y|^\beta\big[q_\alpha(s-t,y-x)+q_\alpha(s-t,z-x)\big]\Big|^{p'}\Big)^{1/p'}\\ \le\;&\frac{C_{p'}}{(s-t)^{\frac{|\eta|+\beta}{\alpha}}}\Bigg[\Big(\int_{\mathbb{R}^d}dz\Big|\int_{\mathbb{R}^d}dy|\partial_v\tilde{p}_\alpha(v,z-y)|\;|z-y|^\beta q_\alpha(s-t,y-x)\Big|^{p'}\Big)^{1/p'}\\ &+\Big(\int_{\mathbb{R}^d}dz\big(q_\alpha(s-t,z-x)\big)^{p'}\Big(\int_{\mathbb{R}^d}dy|\partial_v\tilde{p}_\alpha(v,y-z)|\;|y-z|^\beta\Big)^{p'}\Big)^{1/p'}\Bigg].\end{aligned} \quad (3.22)$$

From the $\mathbb{L}^1-\mathbb{L}^{p'}$ convolution inequality and Lemma 9 (see also (3.18)) we thus obtain:

$$\|\tilde{p}_\alpha(v,\cdot)\star D_x^\eta p_\alpha(s-t,\cdot-x)\|_{\mathbb{L}^{p'}}\le\frac{C_{p'}}{(s-t)^{\frac{|\eta|+\beta+\frac{d}{p}}{\alpha}}}v^{-1+\frac{\beta}{\alpha}}.$$

Hence,

$$\Big(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot-x)]|_{[0,(s-t)]}\Big)^{q'} \leq \frac{C}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}+\frac{\beta}{\alpha}\right]q'}}\int_0^{(s-t)}\frac{dv}{v}v^{(1-\frac{1-\gamma}{\alpha}-1+\frac{\beta}{\alpha})q'}$$
$$\leq \frac{C}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}+\frac{\beta}{\alpha}+\frac{1-\gamma-\beta}{\alpha}\right]q'}} = \frac{C}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}+\frac{1-\gamma}{\alpha}\right]q'}}, \quad (3.23)$$

provided $\beta+\gamma>1$ for the second inequality (which can be assumed since we can choose β arbitrarily in $(0,1)$). The map $s\mapsto \mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot-x)]|_{[0,(s-t)]}$ hence belongs to $\mathbb{L}^{r'}((t,T],\mathbb{R}^+)$ under the same previous condition on η than in (3.19). Let us eventually mention that the above arguments somehow provide the lines of the proof of Lemma 10 for $\Psi=1$. The proof in its whole generality is provided in Appendix A.

Remark 11 (Pointwise gradient estimate on G^α)**.** *The condition in* (3.19) *then precisely gives that the gradient of the Green kernel will exist pointwise (with uniform bound depending on the Besov norm of* Φ*) as soon as:*

$$1<\alpha(1-\frac{1}{r})+\gamma-1-\frac{d}{p} \iff \gamma>2-\alpha(1-\frac{1}{r})+\frac{d}{p}. \quad (3.24)$$

In particular, provided (3.24) *holds, the same type of arguments would also lead to a Hölder control of the gradient in space of index* $\zeta<\alpha(1-1/r)+\gamma-1-d/p-1$*. The previous computations somehow provide the almost optimal regularity that could be attainable for* u *(through what can be derived from* w *solving* (3.15)*). The purpose of the next section will precisely be to prove that these arguments can be adapted to that framework. The price to pay will be some additional constraint on the* γ *because we will precisely have to handle the product* $F\cdot Du$*.*

Remark 12 (On the second integrability parameter "q" in the Besov norm)**.** *Eventually, we emphasize that the parameter* q *does not play a key role in the previous analysis. Indeed, all the thresholds appearing do not depend on this parameter. Since for all* γ,p *we have that for all* $q<q'$ *that* $B^\gamma_{p,q}\subset B^\gamma_{p,q'}$ *the above analysis suggests that it could be enough to consider the case* $q=\infty$*. Nevertheless, as it does not provide any additional difficulties, we let the parameter* q *vary in the following.*

3.3 Uniform estimates of the solution of the mollified version of PDE (2.1) and associated (uniform) Hölder controls.

This part is dedicated to the proof of Proposition 6 and Corollary 7. It is known that, under **(UE)** and for $\vartheta>\alpha$, if $g\in\mathbb{B}^\vartheta_{\infty,\infty}$ is also bounded and $f\in\mathbb{B}^{\vartheta-\alpha}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R})$, there exists a unique classical solution $u:=u_m\in\mathbb{L}^\infty([0,T],\mathbb{B}^\vartheta_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}))$ to the *mollified* PDE (2.2). This is indeed the usual Schauder estimates for sub-critical stable operators (see e.g. Priola [Pri12] or Mikulevicius and Pragarauskas who also address the case of a multiplicative noise [MP14]). It is clear that the following Duhamel representation formula holds for u_m. With the notations of (1.7):

$$u_m(t,x)=P^\alpha_{T-t}[g](x)+G^\alpha f(t,x)+\mathfrak{r}_m(t,x), \quad (3.25)$$

where the Green kernel G^α is defined by (3.16) and where the remainder term $\mathfrak{r}_m$ is defined as follows:

$$\mathfrak{r}_m(t,x):=\int_t^T ds P^\alpha_{T-s}[\langle F_m(s,\cdot),Du_m(s,\cdot)\rangle](x). \quad (3.26)$$

It is plain to check that, if we now relax the boundedness assumption on g, supposing it can have linear growth, there exists $C:=C(d)>0$ such that

$$\big\|DP^\alpha_{T-t}[g]\big\|_{\mathbb{L}^\infty([0,T],\mathbb{B}^{\vartheta-1}_{\infty,\infty})}+\|G^\alpha f\|_{\mathbb{L}^\infty([0,T],\mathbb{B}^\vartheta_{\infty,\infty})}\leq C\big(\|f\|_{\mathbb{L}^\infty([0,T],\mathbb{B}^{\vartheta-\alpha}_{\infty,\infty})}+\|Dg\|_{\mathbb{B}^{\vartheta-1}_{\infty,\infty}}\big).$$

We also refer to the section concerning the smoothness in time below for specific arguments related to a terminal condition with linear growth.

In the following, we will extend the previous bounds in order to consider *singular* sources as well. In order to keep the notations as clear as possible, we drop the superscript m associated with the mollifying procedure for the rest of the section.

(i) Gradient bound. Let us first control the terminal condition. We have, integrating by parts and using usual cancelation arguments,

$$|DP^{\alpha}_{T-t}[g](x)| \leq \sum_{j=1}^{d} |\partial_{x_j} P^{\alpha}_{T-t}[g](x)| \quad \leq \quad \sum_{j=1}^{d} \left| \int_{\mathbb{R}^d} dy \partial_j g(y) p_\alpha(T-t, y-x) \right| \leq \sum_{j=1}^{d} C \|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}}. \quad (3.27)$$

We now turn to control the Green kernel part. Write

$$\begin{aligned} |DG^{\alpha} f(t,x)| \leq \sum_{j=1}^{d} |\partial_{x_j} G^{\alpha} f(t,x)| &= \sum_{j=1}^{d} \left| \int_t^T ds \int_{\mathbb{R}^d} dy f(s,y) \partial_{x_j} p_\alpha(s-t, y-x) \right| \\ &\leq \sum_{j=1}^{d} \|f\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-\alpha}_{\infty,\infty})} \|\partial_{x_j} p_\alpha(\cdot - t, \cdot - x)\|_{\mathbb{L}^1(\mathbb{B}^{\alpha-\theta}_{1,1})}. \end{aligned}$$

From the very definition (1.10) of θ we have $\theta - \alpha + 1 < 1$ and $(\theta - \alpha + 1) + 1 > 1$. We can thus apply Lemma 10 (see eq. (3.12) with $\gamma = \theta - \alpha + 1$, $\beta = 1$, $\eta = 1$ and $\Psi = 1$ therein) to obtain

$$\|\partial_{x_j} p_\alpha(s-t, \cdot - x)\|_{\mathbb{B}^{\alpha-\theta}_{1,1}(\mathbb{R}^d)} \leq \frac{C}{(s-t)^{\left[\frac{\alpha-\theta}{\alpha} + \frac{1}{\alpha}\right]}}.$$

Recalling $\theta > 1$, we thus obtain

$$\|DG^{\alpha} f\|_{\mathbb{L}^\infty} \leq C(T-t)^{\frac{\theta-1}{\alpha}} \|f\|_{\mathbb{L}^\infty([0,T], \mathbb{B}^{\theta-\alpha}_{\infty,\infty})}. \quad (3.28)$$

Let us now focus on first gradient estimate of $\mathfrak{r}$. Using the Hölder inequality and then Besov duality we have,

$$\begin{aligned} |D\mathfrak{r}(t,x)| \leq \sum_{j=1}^{d} |\partial_{x_j} \mathfrak{r}(t,x)| &\leq \sum_{j=1}^{d} \sum_{k=1}^{d} \left| \int_t^T ds \int_{\mathbb{R}^d} dy F_k(s,y) \partial_{y_k} u(s,y) \partial_{x_j} p_\alpha(s-t, y-x) \right| \\ &\leq \sum_{j=1}^{d} \sum_{k=1}^{d} \|F_k\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})} \|\partial_k u \partial_{x_j} p_\alpha(\cdot - t, \cdot - x)\|_{\mathbb{L}^{r'}(\mathbb{B}^{1-\gamma}_{p',q'})}, \quad (3.29) \end{aligned}$$

so that the main issue consists in establishing the required control on the map $(t,T] \ni s \mapsto \|\partial_k u(s,\cdot) \partial_{x_j} p_\alpha(\cdot - t, \cdot - x)\|_{\mathbb{B}^{1-\gamma}_{p',q'}}$ for any j,k in $[\![1,d]\!]$. Note that since for all s in $[0,T]$ the map $y \mapsto u(s,y)$ is in $\mathbb{B}^{\vartheta}_{\infty,\infty}$ for any $\vartheta \in (\alpha, \alpha+1]$, we have in particular from the very definition of θ (see eq. (1.10)) and assumptions on γ that there exists $\varepsilon > 0$ such that $\theta - 1 - \varepsilon > 0$, $\theta - 1 - \varepsilon + \gamma > 1$ and for all s in $[0,T]$ the map $y \mapsto \partial_k u(s,y)$ is in $\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}$. One can hence apply Lemma 10 so that (see eq. (3.12) with $\beta = \theta - 1 - \varepsilon$, $\eta = 1$ and $\Psi(s,\cdot) = \partial_k u(s,\cdot)$ therein)

$$\|\partial_k u(s,\cdot) \partial_{x_j} p_\alpha(s-t, \cdot - x)\|_{\mathbb{B}^{1-\gamma}_{p',q'}} \leq \|\partial_k u(s,\cdot)\|_{\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}} \frac{C}{(s-t)^{\left[\frac{1-\gamma}{\alpha} + \frac{d}{p\alpha} + \frac{1}{\alpha}\right]}}.$$

This map hence belongs to $\mathbb{L}^{r'}((t,T], \mathbb{R}_+)$ as soon as

$$-r' \left[\frac{d}{p\alpha} + \frac{1}{\alpha} + \frac{1-\gamma}{\alpha} \right] > -1 \Leftrightarrow \gamma > 2 - \alpha + \frac{\alpha}{r} + \frac{d}{p}, \quad (3.30)$$

which follows from the assumptions on γ. We then obtain, after taking the $\mathbb{L}^{r'}((t,T], \mathbb{R}_+)$ norm of the above estimate, that

$$|D\mathfrak{r}(t,x)| \leq C T^{\frac{\theta-1}{\alpha}} \|Du\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})}. \quad (3.31)$$

(ii) Hölder norm of the gradient. As in the above proof we obtain gradient bounds depending on the spatial Hölder norm of Du, we now have to precisely estimate this quantity. The main difficulty is induced by

the remainder term:

$$\begin{aligned}|D\mathfrak{r}(t,x)-D\mathfrak{r}(t,x')| &\leq \sum_{j=1}^{d}|\partial_j\mathfrak{r}(t,x)-\partial_j\mathfrak{r}(t,x')|\\ &\leq \sum_{j=1}^{d}\sum_{k=1}^{d}\left|\int_t^T ds\int_{\mathbb{R}^d}dyF_k(s,y)\left(\partial_{y_k}u(s,y)\left(\partial_{x_j}p_\alpha(s-t,y-x)-\partial_{x_j}p_\alpha(s-t,y-x')\right)\right)\right|\\ &\leq \sum_{j=1}^{d}\sum_{k=1}^{d}\|F_k\|_{\mathbb{L}^r(\mathbb{B}_{p,q}^{-1+\gamma})}\|\partial_k u\left(\partial_{x_j}p_\alpha(\cdot-t,\cdot-x)-\partial_{x_j}p_\alpha(\cdot-t,\cdot-x')\right)\|_{\mathbb{L}^{r'}(\mathbb{B}^{1-\gamma}_{p',q'})},\end{aligned}$$

using again the Hölder inequality and duality between the considered Besov spaces (see Section 3.1.2). Hence, the main issue consists in establishing the required control on the map

$$(t,T]\ni s\mapsto \|\partial_k u(s,\cdot)\left(\partial_{x_j}p_\alpha(s-t,\cdot-x)-\partial_{x_j}p_\alpha(s-t,\cdot-x')\right)\|_{\mathbb{L}^{r'}(\mathbb{B}^{1-\gamma}_{p',q'})},$$

for any j,k in $[\![1,d]\!]$. Since $\theta-1-\varepsilon<1$, one can again apply Lemma 10 so that (see eq. (3.13) with $\beta=\theta-1-\varepsilon$, $\beta'=\theta-1-\varepsilon$, $\eta=1$ and $\Psi(s,\cdot)=\partial_k u(s,\cdot)$ therein):

$$\begin{aligned}&\|\partial_k u(s,\cdot)\big(\partial_{x_j}p_\alpha(s-t,\cdot-x)-\partial_{x_j}p_\alpha(s-t,\cdot-x')\big)\|_{\mathbb{B}^{1-\gamma}_{p',q'}}\\ \leq\ &\|\partial_k u(s,\cdot)\|_{\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}}\frac{C}{(s-t)^{\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{1+(\theta-1-\varepsilon)}{\alpha}\right]}}|x-x'|^{\theta-1-\varepsilon}\leq\frac{C\|\partial_k u\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})}}{(s-t)^{\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{1+(\theta-1-\varepsilon)}{\alpha}\right]}}|x-x'|^{\theta-1-\varepsilon}.\end{aligned}$$

The above map hence belongs to $\mathbb{L}^{r'}((t,T],\mathbb{R}^+)$ as soon as

$$-r'\left[\frac{d}{p\alpha}+\frac{1+(\theta-1-\varepsilon)}{\alpha}+\frac{1-\gamma}{\alpha}\right]>-1\Leftrightarrow\theta-1-\varepsilon<\gamma-\left(2-\alpha+\frac{\alpha}{r}+\frac{d}{p}\right),\tag{3.32}$$

which readily follows from the very definition of θ (see eq. (1.10)) and the fact that $\varepsilon>0$. We then obtain

$$|D\mathfrak{r}(t,x)-D\mathfrak{r}(t,x')|\leq CT^{\frac{\varepsilon}{\alpha}}\|Du\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})}|x-x'|^{\theta-1-\varepsilon}.\tag{3.33}$$

Remark 13. *Note that assuming that θ is fixed, we readily obtain from* (3.32) *together with the constraint $\theta-1-\varepsilon+\gamma>1$ the initial constraint*

$$\gamma>\frac{3-\alpha+\frac{d}{p}+\frac{\alpha}{r}}{2}.\tag{3.34}$$

In comparison with the threshold obtained when investigating the smoothing effect of the Green kernel (see eq. (3.24) *and the related discussion) this additional regularity allows to define the product $F\cdot Du$. Indeed, if one wants to define it* e.g. *as a Young integral, one has to require the sum of the local regularity indexes of the two maps to be greater than one: $\theta-1-\varepsilon+\gamma>1$. Extensions are possible and there already exist robust theories to bypass such a constraint (rough path in dimension 1, paracontrolled distribution or regularity structures) but, to the best of our knowledges, it requires the map F to be enhanced to a rough distribution $\tilde{F}$, which significantly restraints the possible choices of the drift.*

Let us eventually estimate the Hölder moduli of the gradients of the first and second terms in the Duhamel representation (3.25). We first note that, for the Green kernel, the proof follows from the above lines. When doing so, we obtain that

$$|DG^\alpha f(t,x)-DG^\alpha f(t,x')|\leq CT^\varepsilon\|f\|_{\mathbb{B}^{\theta-\alpha}_{\infty,\infty}}|x-x'|^{\theta-1-\varepsilon}.\tag{3.35}$$

Concerning the terminal condition, we have on the one hand, when $(T-t)^{\frac{1}{\alpha}}\leq|x-x'|$ (off-diagonal regime), that:

$$\begin{aligned}|DP^\alpha_{T-t}[g](x)-DP^\alpha_{T-t}[g](x')| &= \left|\int_{\mathbb{R}^d}dyDg(y)\big(p_\alpha(T-t,y-x)-p_\alpha(T-t,y-x')\big)\right|\\ &\leq \Bigg|\int_{\mathbb{R}^d}dy\big(Dg(y)-Dg(x)\big)p_\alpha(T-t,y-x)+Dg(x)-Dg(x')\\ &\quad-\int_{\mathbb{R}^d}dy\big(Dg(y)-Dg(x')\big)p_\alpha(T-t,y-x')\Bigg|\\ &\leq C\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}}|x-x'|^{\theta-1-\varepsilon}.\end{aligned}\tag{3.36}$$

On the other hand, when $(T-t)^{\frac{1}{\alpha}} > |x-x'|$ (diagonal regime), we have using cancellations arguments

$$\begin{aligned}
&|DP^\alpha_{T-t}[g](x) - DP^\alpha_{T-t}[g](x')| \\
\leq\ & \Big| \int_{\mathbb{R}^d} [p_\alpha(T-t, y-x) - p_\alpha(T-t, y-x')] Dg(y) dy \Big| \\
\leq\ & \Big| \int_0^1 d\lambda \int_{\mathbb{R}^d} [D_x p_\alpha\big(T-t, y-(x'+\mu(x-x'))\big)\cdot(x-x')][Dg(y) - Dg(x'+\mu(x-x'))] dy \Big| \\
\leq\ & \|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}} (T-t)^{-\frac{1}{\alpha}+\frac{\theta-1}{\alpha}} |x-x'| \leq C(T-t)^\varepsilon \|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}} |x-x'|^{\theta-1-\varepsilon}.
\end{aligned}$$

Hence

$$|DP^\alpha_{T-t}[g](x) - DP^\alpha_{T-t}[g](x')| \leq C(T^\varepsilon + 1)\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}} |x-x'|^{\theta-1-\varepsilon}. \tag{3.37}$$

Putting together estimates (3.27), (3.28), (3.31), (3.33), (3.35) and (3.37) we deduce that

$$\forall\, \alpha \in \left(\frac{1+\frac{d}{p}}{1-\frac{1}{r}}, 2\right], \forall \gamma \in \left(\frac{3-\alpha+\frac{d}{p}+\frac{\alpha}{r}}{2}, 1\right],\ \exists C(T) > 0 \text{ s.t. } \|Du\|_{\mathbb{L}^\infty\left(\mathbb{B}^{\gamma-2+\alpha-\frac{d}{p}-\frac{\alpha}{r}-\varepsilon}_{\infty,\infty}\right)} < C_T. \tag{3.38}$$

In particular, when $g = 0$, $\lim C_T = 0$ when T tends to 0.

(iii) Smoothness in time for u and Du. We restart here from the Duhamel representation (3.25). Namely,

$$u(t,x) = P^\alpha_{T-t}[g](x) + G^\alpha[f](t,x) + \mathfrak{r}(t,x),$$

where from (3.26), the remainder term writes:

$$\mathfrak{r}(t,x) = \int_t^T ds \int_{\mathbb{R}^d} dy \langle F(s,y), Du(s,y)\rangle p_\alpha(s-t, y-x).$$

We now want to control for a fixed $x \in \mathbb{R}^d$ and $0 \leq t < t' \leq T$ the difference:

$$u(t',x) - u(t,x) = \big(P^\alpha_{T-t'} - P^\alpha_{T-t}\big)[g](x) + \big(G^\alpha f(t',x) - G^\alpha f(t,x)\big) + \big(r(t',x) - r(t,x)\big). \tag{3.39}$$

For the first term in the r.h.s. of (3.39) we write:

$$\begin{aligned}
\big(P^\alpha_{T-t'} - P^\alpha_{T-t}\big)[g](x) &= \int_{\mathbb{R}^d} \big[p_\alpha(T-t', y-x) - p_\alpha(T-t, y-x)\big] g(y) dy \\
&= -\int_{\mathbb{R}^d} \int_0^1 d\lambda \big[\partial_s p_\alpha(s, y-x)\big]\Big|_{s=T-t-\lambda(t'-t)} g(y) dy (t'-t).
\end{aligned}$$

From the Fubini's theorem and usual cancellation arguments we get:

$$\big(P^\alpha_{T-t'} - P^\alpha_{T-t}\big)[g](x) = -(t'-t)\int_0^1 d\lambda \Big[\int_{\mathbb{R}^d} \partial_s p(s, y-x)\big(g(y) - g(x) - Dg(x)\cdot(y-x)\big) dy\Big]\Big|_{s=T-t-\lambda(t'-t)}.$$

We indeed recall that, because of the symmetry of the driving process $\mathcal{W}$, and since $\alpha > 1$, one has for all $s > 0$, $\int_{\mathbb{R}^d} p(s, y-x)(y-x) dy = 0$. Recalling as well that we assumed $Dg \in \mathbb{B}^{\theta-1}_{\infty,\infty}$, we therefore derive from Lemma 9:

$$\begin{aligned}
|\big(P^\alpha_{T-t'} - P^\alpha_{T-t}\big)[g](x)| &\leq (t'-t)\int_0^1 d\lambda \Big[\frac{C\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}}}{s} \int_{\mathbb{R}^d} q_\alpha(s, y-x)|y-x|^\theta dy\Big]\Big|_{s=T-t-\lambda(t'-t)} \\
&\leq C(t'-t)\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}} \int_0^1 d\lambda s^{-1+\frac{\theta}{\alpha}}\big|_{s=T-t-\lambda(t'-t)},
\end{aligned}$$

recalling from (1.10) that $\theta < \alpha$ for the last inequality. Observe now that since $0 \leq t < t' \leq T$, one has $s = T-t-\lambda(t'-t) \geq (1-\lambda)(t'-t)$ for all $\lambda \in [0,1]$. Hence,

$$\begin{aligned}
|\big(P^\alpha_{T-t'} - P^\alpha_{T-t}\big)[g](x)| &\leq C(t'-t)\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}} \int_0^1 \frac{d\lambda}{(1-\lambda)^{1-\frac{\theta}{\alpha}}} (t'-t)^{-1+\frac{\theta}{\alpha}} \\
&\leq C(t'-t)^{\frac{\theta}{\alpha}} \|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}},
\end{aligned} \tag{3.40}$$

which is the expected control. We now focus on the remainder term r since the control of the Green kernel is easier and can be derived following the same lines of reasoning. Write

$$\mathfrak{r}(t',x)-\mathfrak{r}(t,x)=\int_{t'}^{T}ds\big(P^{\alpha}_{s-t'}-P^{\alpha}_{s-t}\big)[\langle F(s,\cdot)Du(s,\cdot)\rangle](x)+\int_{t}^{t'}dsP^{\alpha}_{s-t}[\langle F(s,\cdot)Du(s,\cdot)\rangle](x). \tag{3.41}$$

From Lemma 10 (see eq. (3.12) with $\beta=\theta-1-\varepsilon$ and $\eta=0$) it can be deduced (see computations in point **(i)** of the current section) that

$$|\int_{t}^{t'}dsP^{\alpha}_{s-t}[\langle F(s,\cdot),Du(s,\cdot)\rangle](x)|\leq C|t-t'|^{\frac{\theta}{\alpha}}. \tag{3.42}$$

Let us now focus on

$$\begin{aligned}\int_{t'}^{T}ds\big(P^{\alpha}_{s-t'}-P^{\alpha}_{s-t}\big)[\langle F(s,\cdot)Du(s,\cdot)\rangle](x) &= \int_{t'}^{T}ds\int_{0}^{1}d\lambda\Big\{\partial_w P^{\alpha}_{s-w}[\langle F(s,\cdot)Du(s,\cdot)\rangle](x)\Big\}\Big|_{w=t+\lambda(t'-t)}(t'-t)\\ &= \int_{0}^{1}d\lambda\int_{t'}^{T}ds\Big\{L^{\alpha}P^{\alpha}_{s-w}[\langle F(s,\cdot)Du(s,\cdot)\rangle](x)\Big\}\Big|_{w=t+\lambda(t'-t)}(t'-t).\end{aligned} \tag{3.43}$$

We have

$$\begin{aligned}&\int_{t'}^{T}ds|L^{\alpha}P^{\alpha}_{s-w}[\langle F(s,\cdot)Du(s,\cdot)\rangle](x)|\\ \leq\ &\sum_{k=1}^{d}\int_{t'}^{T}ds\Big|\int_{\mathbb{R}^d}dyF_k(s,y)\partial_{y_k}u(s,y)L^{\alpha}p_{\alpha}(s-w,y-x)\Big|\\ \leq\ &\sum_{k=1}^{d}\|F_k\|_{\mathbb{L}^r([t',T],\mathbb{B}^{-1+\gamma}_{p,q})}\|\partial_k uL^{\alpha}p_{\alpha}(\cdot-w,\cdot-x)\|_{\mathbb{L}^{r'}([t',T],\mathbb{B}^{1-\gamma}_{p',q'})}.\end{aligned} \tag{3.44}$$

Applying Lemma 10 (see eq. (3.12) with $\beta=\theta-1-\varepsilon$ and $\eta=\alpha$ therein), we get:

$$\|\partial_k u(s,\cdot)L^{\alpha}p_{\alpha}(s-w,\cdot-x)\|_{\mathbb{B}^{1-\gamma}_{p',q'}}\leq\|\partial_k u(s,\cdot)\|_{\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}}\frac{C}{(s-w)^{\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+1\right]}}.$$

Thus, from (3.38) (recall from (1.10) that $\gamma-2+\alpha-\frac{d}{p}-\frac{\alpha}{r}-\varepsilon=\theta-1-\varepsilon$):

$$\|\partial_k uL^{\alpha}p_{\alpha}(\cdot-w,\cdot-x)\|_{\mathbb{L}^{r'}([t',T],\mathbb{B}^{1-\gamma}_{p',q'})} \leq C(t'-w)^{\frac{1}{r'}-\left(\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+1\right)}=C(t'-w)^{\frac{\theta}{\alpha}-1}. \tag{3.45}$$

Therefore, from (3.45) and (3.44), we derive:

$$\int_{t'}^{T}ds|L^{\alpha}P^{\alpha}_{s-w}[\langle F(s,\cdot)Du(s,\cdot)\rangle](x)| \leq C\sum_{k=1}^{d}\|F_k\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})}(t'-w)^{\frac{\theta}{\alpha}-1},$$

which in turn, plugged into (3.43), gives:

$$\begin{aligned}&|\int_{t'}^{T}ds\big(P^{\alpha}_{s-t'}-P^{\alpha}_{s-t}\big)[\langle F(s,\cdot)Du(s,\cdot)\rangle](x)|\leq\int_{0}^{1}d\lambda\int_{t'}^{T}ds\Big|L^{\alpha}P^{\alpha}_{s-w}[\langle F(s,\cdot)Du(s,\cdot)\rangle](x)\Big|\Big|_{w=t+\lambda(t'-t)}(t'-t)\\ \leq\ &C\sum_{k=1}^{d}\|F_k\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})}\int_{0}^{1}d\lambda(t'-(t+\lambda(t'-t)))^{\frac{\theta}{\alpha}-1}(t'-t)\\ \leq\ &C\sum_{k=1}^{d}\|F_k\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})}(t'-t)^{\frac{\theta}{\alpha}}.\end{aligned} \tag{3.46}$$

From (3.46), (3.42) and (3.41) we thus obtain:

$$\big|\mathfrak{r}(t',x)-\mathfrak{r}(t,x)\big|\leq C\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})}(t'-t)^{\frac{\theta}{\alpha}}. \tag{3.47}$$

The Hölder control of the Green kernel $G^\alpha f$ follows from similar arguments. Indeed, repeating the above proof it is plain to check that there exists $C \geq 1$ s.t. for all $0 \leq t < t' \leq T$, $x \in \mathbb{R}^d$:

$$\Big|\big(G^\alpha f(t',x) - G^\alpha f(t,x)\big)\Big| \leq C\|f\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-\alpha}_{\infty,\infty})}(t'-t)^{\frac{\theta}{\alpha}}. \tag{3.48}$$

The final control of (2.4) concerning the smoothness in time then follows plugging (3.40), (3.47) and (3.48) into (3.39). The control concerning the time sensitivity of the spatial gradient would be obtained following the same lines.

(iv) Conclusion: proof of Proposition 6, Corollary 7 and Theorem 8. Points **(i)** to **(iii)** conclude the proof of Proposition 6. Let us eventually notice that Corollary 7 is a direct consequence of the above computations. Indeed, replacing the source term f by the k^{th} coordinate of F in the Green kernel, the proof follows from the control obtained for the remainder term in the Duhamel representation (3.25). Eventually, the proof of Theorem 8 follows from compactness arguments together with the Schauder like control of Proposition 6.

Remark 14 (About additional diffusion coefficients)**.** *Let us first explain how, in the diffusive setting, $\alpha = 2$ the diffusion coefficient can be handled. Namely, this would lead to consider for the PDE with mollified coefficients an additional term in the Duhamel formulation that would write:*

$$u_m(t,x) = P^{\alpha,\xi,m}_{s-t}[g](x) + \int_t^T ds P^{\alpha,\xi,m}_{s-t}[\Big\{f(s,\cdot) + F_m \cdot Du_m(s,\cdot) + \frac{1}{2}\mathrm{Tr}\big((a_m(s,\cdot) - a_m(s,\xi))D^2u_m(s,\cdot)\big)\Big\}](x), \tag{3.49}$$

for an auxiliary parameter ξ which will be taken equal to x after potential differentiations in (3.49)*. Here, $P^{\alpha,\xi,m}_{s-t}$ denotes the two-parameter semi-group associated with $\big(\frac{1}{2}\mathrm{Tr}\big(a_m(v,\xi)D^2\big)\big)_{v\in[s,t]}$ (mollified diffusion coefficient frozen at point ξ). Let us focus on the second order term. Recall from the above proof of Proposition 6 that we aim at estimating the gradient pointwise, deriving as well some Hölder continuity for it. Hence, focusing on the additional term, we write for the gradient part:*

$$\begin{aligned}
& D_x \int_t^T ds P^{\alpha,\xi,m}_{s-t}[\frac{1}{2}\mathrm{Tr}\big((a_m(s,\cdot) - a_m(s,\xi))D^2u_m(s,\cdot)\big)](x) \\
= & \int_t^T ds \int_{\mathbb{R}^d} D_x p^{\xi,m}_\alpha(t,s,x,y)\frac{1}{2}\mathrm{Tr}\big((a_m(s,y) - a_m(s,\xi))D^2u_m(s,y)\big)dy \\
& = \frac{1}{2}\sum_{i,j=1}^d \int_t^T ds \int_{\mathbb{R}^d} \Big(D_x p^{\xi,m}_\alpha(t,s,x,y)\big((a_{m,i,j}(s,y) - a_{m,i,j}(s,\xi))\big)\Big)D_{y_iy_j}u_m(s,y)dy.
\end{aligned}$$

From the previous Proposition 6, we aim at establishing that Du_m has Hölder index $\theta - 1 - \varepsilon = \gamma - 2 + \alpha - d/p - \alpha/r - \varepsilon$ and therefore $D_{y_iy_j}u_m \in \mathbb{B}^{\theta-2-\varepsilon}_{\infty,\infty}$. Assume for a while that $p = q = r = +\infty$. The goal is now to bound the above term through Besov duality. Namely, taking $\xi = x$ after having taken the gradient w.r.t. x for the heat kernel, we get:

$$\begin{aligned}
& |D_x \int_t^T ds P^{\alpha,\xi,m}_{s-t}[\frac{1}{2}\mathrm{Tr}\big((a_m(s,\cdot) - a_m(s,\xi))D^2u_m(s,\cdot)\big)](x)|\Big|_{\xi=x} \\
\leq & \sum_{i,j=1}^d \int_t^T ds \|\Big(D_x p^{\xi,m}_\alpha(t,s,x,\cdot)\big((a_{m,i,j}(s,\cdot) - a_{m,i,j}(s,\xi))\big)\Big)\|_{\mathbb{B}^{2+\varepsilon-\theta}_{1,1}}\Big|_{\xi=x}\|\partial^2_{i,j}u_m(s,\cdot)\|_{\mathbb{B}^{\theta-2-\varepsilon}_{\infty,\infty}}.
\end{aligned}$$

Now, in the considered case $\theta - 2 - \varepsilon = \gamma - 1 - \varepsilon$. Recalling that $D_x p^{\xi,m}_\alpha(t,s,x,\cdot) \in \mathbb{B}^{1/2-\tilde{\varepsilon}}_{1,1}$ for any $\tilde{\varepsilon} > 0$ for $\gamma > 1/2 = (3-\alpha)/2$ and ε small enough, we will indeed have that $D_x p^{\xi,m}_\alpha(t,s,x,\cdot)\big((a_{m,i,j}(s,\cdot) - a_{m,i,j}(s,\xi)\big) \in \mathbb{B}^{2+\varepsilon-\theta}_{1,1}$ provided the bounded function a itself has the same regularity, i.e. $2+\varepsilon-\theta$, the integrability of the product deriving from the one of the heat kernel. Since $\|\partial^2_{i,j}u_m(s,\cdot)\|_{\mathbb{B}^{\theta-2-\varepsilon}_{\infty,\infty}} \leq C\|Du_m(s,\cdot)\|_{\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}}$, see e.g. Triebel [Tri83], this roughly means that, the same Schauder estimate should hold with a diffusion coefficient $a \in \mathbb{L}^\infty([0,T],\mathbb{B}^{2+\varepsilon-\theta}_{\infty,\infty})$. Similar thresholds also appear more generally in [ZZ17]. The general diffusive case for $p,q,r \geq 1$ and γ satisfying the conditions of Theorem 1 can be handled similarly through duality arguments.

For the pure jump case, we illustrate for simplicity what happens if the diffusion coefficient is scalar. Namely, when $L^{\alpha,\sigma}\varphi(x) = \text{p.v.}\int_{\mathbb{R}^d}\big(\varphi(x+\sigma(x)z) - \varphi(x)\big)\nu(dz) = -\sigma^\alpha(x)(-\Delta)^{\alpha/2}\varphi(x)$, where σ is a non-degenerate

diffusion coefficient. Introducing $L^{\alpha,\sigma,\xi}\varphi(x) = \text{p.v.}\int_{\mathbb{R}^d}\big(\varphi(x+\sigma(\xi)z)-\varphi(x)\big)\nu(dz) = -\sigma^\alpha(\xi)(-\Delta)^{\alpha/2}\varphi(x)$*, we rewrite for the Duhamel formula, similarly to* (3.49)*:*

$$u_m(t,x) = P^{\alpha,\xi,m}_{s-t}[g](x) + \int_t^T ds P^{\alpha,\xi,m}_{s-t}[\big\{f(s,\cdot) + F_m\cdot Du_m(s,\cdot) + (L^{\alpha,\sigma_m} - L^{\alpha,\sigma_m,\xi})u_m(s,\cdot)\big)\big\}](x). \quad (3.50)$$

Focusing again on the non-local term, we write for the gradient part:

$$\begin{aligned} & D_x\int_t^T ds P^{\alpha,\xi,m}_{s-t}[\big(\sigma^\alpha_m(s,\cdot)-\sigma^\alpha_m(s,\xi)\big)\Delta^{\frac{\alpha}{2}}u_m(s,\cdot)\big)](x) \\ = & -\int_t^T ds\int_{\mathbb{R}^d} D_x p^{\xi,m}_\alpha(t,s,x,y)\big(\sigma^\alpha_m(s,y)-\sigma^\alpha_m(s,\xi)\big)(-\Delta)^{\frac{\alpha}{2}}u_m(s,y)dy. \end{aligned}$$

Consider again the case $p=q=r=\infty$*. Since* $Du_m \in \mathbb{L}^\infty([0,T],\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})$*, we thus have that* $-(-\Delta)^{\alpha/2}u_m \in \mathbb{L}^\infty([0,T],\mathbb{B}^{\theta-\alpha-\varepsilon}_{\infty,\infty})$*, where* $\theta-\alpha-\varepsilon = -1+\gamma-\varepsilon$*. Still by duality one has to control the norm of the term* $D_x p^{\xi,m}_\alpha(t,s,x,y)\big(\sigma^\alpha_m(s,y)-\sigma^\alpha_m(s,\xi)\big)$ *in the Besov space* $\mathbb{B}^{1-\gamma+\varepsilon}_{1,1}$*. Since* $\gamma > (3-\alpha)/2$ *and* $D_x p^{\xi,m}_\alpha(t,s,x,y) \in \mathbb{B}^{1-1/\alpha}_{1,1}$*, this will be the case provided* $\sigma \in \mathbb{L}^\infty([0,T],\mathbb{B}^{1-\gamma+\varepsilon}_{\infty,\infty})$ *for* ε *small enough observing that* $1-\gamma+\varepsilon<(\alpha-1)/2$*.*

Note that, in comparison with the result obtained in [LZ19], the above threshold is precisely the one appearing in [LZ19] in this specific case. The general matrix case for σ *is more involved. It requires in [LZ19] the Bony decomposition. We believe it could also be treated through the duality approach considered here but postpone this discussion to further research. In the scalar case, the analysis for general* p,q,r,γ *as in Theorem 1 could be performed similarly.*

4 Building the dynamics

In this part, we aim at proving Theorem 3 and Corollary 4. We restrict here to the pure jump case $\alpha \in (1,2)$, since the diffusive one was already considered in [DD16]. We adapt here their procedure to the current framework: we first recover the noise through the martingale problem, then recover a drift as the difference between the weak solution and the noise obtained before and estimate its contribution. This is the purpose of Proposition 11 below. Having such tools at hand, we recover the dynamics of the weak solution of the *formal* SDE (1.1) by giving a meaning of each of the above quantities as $\mathbb{L}^\ell$ stochastic-Young integrals (for $\ell<\alpha$). More precisely, the $\mathbb{L}^\ell$ stochastic-Young integral are defined for a suitable class of integrand consisting in the predictable processes $(\psi_s)_{0\le s\le T}$ defined in Corollary 4, leading e.g. to the application of Itô's formula for the dynamics (1.13).

Proposition 11. *Let* $\alpha \in (1,2)$*. For any initial point* $x \in \mathbb{R}^d$*, one can find a probability measure on* $\mathcal{D}([0,T],\mathbb{R}^{2d})$ *(still denoted by* $\mathbb{P}^\alpha$*) s.t. the canonical process* $(X_t,\mathcal{W}_t)_{t\in[0,T]}$ *satisfies the following properties:*

(i) Under $\mathbb{P}^\alpha$*, the law of* $(X_t)_{t\ge0}$ *is a solution of the martingale problem associated with data* (L^α,F,x)*,* $x\in\mathbb{R}^d$ *and the law of* $(\mathcal{W}_t)_{t\ge0}$ *corresponds to the one of a* d*-dimensional stable process with generator* L^α*.*

(ii) For any $1\le q<\alpha$*, there exists a constant* $C := C(\alpha,p,q,r,\gamma)$ *s.t. for any* $0\le v<s\le T$*:*

$$\mathbb{E}^{\mathbb{P}^\alpha}[|X_s - X_v - (\mathcal{W}_s-\mathcal{W}_v)|^q]^{\frac{1}{q}} \le C(s-v)^{\frac{1}{\alpha}+\frac{\theta-1}{\alpha}}, \quad (4.1)$$

(iii) Let $(\mathcal{F}_v)_{v\ge0} := \big(\sigma((X_w,\mathcal{W}_w)_{0\le w\le v})\big)_{v\ge0}$ *denote the filtration generated by the couple* $(X,\mathcal{W})$*. For any* $0\le v<s\le T$*, it holds that:*

$$\mathbb{E}^{\mathbb{P}^\alpha}[X_s - X_v|\mathcal{F}_v] = \mathfrak{f}(v,X_v,s-v) = \mathbb{E}^{\mathbb{P}^\alpha}[u^s(v,X_v) - u^s(s,X_v)|\mathcal{F}_v],$$

with $\mathfrak{f}(v,X_v,s-v) := u^s(v,X_v) - X_v$*, recalling that* u^s *denotes here the solution of equation* (2.1) *on the time interval* $[0,s]$ *with final condition* $u^s(s,x)=x$ *and* $f=0$*.*

Furthermore, the following decomposition holds:

$$\begin{aligned} \mathfrak{f}(v,X_v,s-v) &= \mathscr{F}(v,X_v,s-v) + \mathscr{R}(v,X_v,s-v), \\ |\mathscr{F}(v,X_v,s-v)| &= \Big|\int_v^s dw\int_{\mathbb{R}^d} dy F(w,y)p_\alpha(w-s,y-X_v)\Big| \\ &\le C\|F\|_{\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q})}(s-v)^{\frac12+\chi},\ \chi\in(0,1/2], \\ |\mathscr{R}(v,X_v,s-v)| &\le C(s-v)^{1+\varepsilon'},\ \varepsilon'>0. \end{aligned} \quad (4.2)$$

Proof.(i) Coming back to point **(i)** in Section 2.2 we have that the couple $\big((X^m_t,\mathcal{W}^m_t)_{t\in[0,T]}\big)_{m\geq 0}$ is tight (pay attention that the stable noise $\mathcal{W}^m$ feels the mollifying procedure as it is obtained through solvability of the martingale problem) so that it converges, along a subsequence, to the couple $(X_t,\mathcal{W}_t)_{t\in[0,T]}$.

(ii) Let $0\leq v<s$. With the notations of (2.7), letting for notational convenience $u^s_m\equiv u_m=(u^1_m,\dots,u^d_m)$ where each u^i_m, i in $\{1,\dots,d\}$ is chosen as the solution of (2.2) with terminal condition x_i (i.e. the i^{th} coordinate of $x=(x_1,\dots,x_d)\in\mathbb{R}^d$) at time s and source term $f\equiv 0$ we obtain, from Itô's formula

$$\begin{aligned}
& X^m_s-X^m_v\\
= \;& M^{s,m}_{v,s}(\alpha,u_m,X^m)+[u_m(v,X^m_v)-u_m(s,X^m_v)] \qquad (4.3)\\
= \;& \int_v^s\int_{\mathbb{R}^d\setminus\{0\}}\{u_m(w,X^m_{w^-}+x)-u_m(w,X^m_{w^-})\}\tilde N^m(dw,dx)+[u_m(v,X_v)-u_m(s,X_v)]\\
= \;& \mathcal{W}^m_s-\mathcal{W}^m_v+[u_m(v,X^m_v)-u_m(s,X^m_v)]+\int_v^s\int_{|x|\leq 1}\{u_m(w,X^m_{w^-}+x)-u_m(w,X^m_{w^-})-x\}\tilde N^m(dw,dx)\\
& +\int_v^s\int_{|x|\geq 1}\{u_m(w,X^m_{w^-}+x)-u_m(w,X^m_{w^-})-x\}\tilde N^m(dw,dx).\\
:= \;& \mathcal{W}^m_s-\mathcal{W}^m_v++[u_m(v,X^m_v)-u_m(s,X^m_v)]+\mathcal{M}^m_S(v,s)+\mathcal{M}^m_L(v,s). \qquad (4.4)
\end{aligned}$$

From the smoothness properties of u_m established in Proposition 6 (in particular $|u^s_m(v,X^m_v)-u^s_m(s,X^m_v)]|\leq C(s-v)^{\theta/\alpha}$ and the gradient is uniformly bounded) we have

$$\begin{aligned}
|\mathcal{U}(w,X^m_{w^-},x)|:=\big|u_m(w,X^m_{w^-}+x)-u_m(w,X^m_{w^-})-x\big| &= \Big|\int_0^1 d\lambda(Du_m(w,X^m_{w^-}+\lambda x)-I)\cdot x\Big|\\
&\leq C(s-w)^{\frac{\theta-1}{\alpha}}|x|, \qquad (4.5)
\end{aligned}$$

recalling that for all z in $\mathbb{R}^d$, $u_m(s,z)=z$. Note that $\big(\mathcal{M}^m_S(v,s)\big)_{0\leq v<s\leq T}$ and $\big(\mathcal{M}^m_L(v,s)\big)_{0\leq v<s\leq T}$ are respectively $\mathbb{L}^2$ and $\mathbb{L}^q$ martingales associated respectively with the "small" and "large" jumps. Let us first handle the "large" jumps. We have by the Burkholder-Davies-Gundy (BDG) inequality that

$$\mathbb{E}\big[|\mathcal{M}^m_L(v,s)|^q\big]\leq C_\ell\mathbb{E}\big[[\mathcal{M}^m_L]^{\frac{q}{2}}_{(v,s)}\big],$$

where $[\mathcal{M}^m_L]_{(v,s)}$ denotes the corresponding bracket given by $\sum_{v\leq w\leq s}|\mathcal{U}(w,X^m_{w^-},\Delta\mathcal{W}^m_w)|^2\mathbf{1}_{|\Delta\mathcal{W}^m_w|\geq 1}$. Using the linear growth of $\mathcal{U}$ w.r.t. its third variable (uniformly w.r.t. the second one) from (4.5) together with the fact that $q/2\leq 1$ we obtain

$$\begin{aligned}
\Big(\sum_{v\leq w\leq s}|\mathcal{U}(w,X^m_{w^-},\Delta\mathcal{W}^m_w)|^2\mathbf{1}_{|\Delta\mathcal{W}^m_w|\geq 1}\Big)^{q/2} &\leq C(s-w)^{q\frac{\theta-1}{\alpha}}\Big(\sum_{v\leq w\leq s}|\Delta\mathcal{W}^m_w|^2\mathbf{1}_{|\Delta\mathcal{W}^m_w|\geq 1}\Big)^{q/2}\\
&\leq C(s-w)^{q\frac{\theta-1}{\alpha}}\sum_{v\leq w\leq s}|\Delta\mathcal{W}^m_w|^q\mathbf{1}_{|\Delta\mathcal{W}^m_w|\geq 1}.
\end{aligned}$$

We then readily get from the compensation formula that

$$\mathbb{E}\big[|\mathcal{M}^m_L(v,s)|^q\big]\leq C(s-w)^{1+q\frac{\theta-1}{\alpha}}\int|x|^q\mathbf{1}_{|x|\geq 1}\nu(dx)\leq C'(s-w)^{1+q\frac{\theta-1}{\alpha}}\leq C'(s-w)^{\frac{q}{\alpha}+q\frac{\theta-1}{\alpha}}.$$

We now deal with the "small" jumps and split them w.r.t. their characteristic scale writing

$$\begin{aligned}
\mathcal{M}^m_S(v,s) &= \mathcal{M}^m_{S,1}(v,s)+\mathcal{M}^m_{S,2}(v,s)\\
&=: \int_v\int_{|x|>(s-v)^{\frac{1}{\alpha}}}\mathcal{U}(w,X^m_{w^-},x)\tilde N^m(dw,dx)+\int_v\int_{|x|\leq(s-v)^{\frac{1}{\alpha}}}\mathcal{U}(w,X^m_{w^-},x)\tilde N^m(dw,dx).
\end{aligned}$$

In the off-diagonal regime (namely for $\mathcal{M}^m_{S,1}(v,s)$), we do not face any integrability problem w.r.t. the Lévy measure. The main idea consists then in using first the BDG inequality, then the compensation formula and

(4.5) and eventually usual convexity arguments together with the compensation formula again to obtain

$$\begin{aligned}
\mathbb{E}[|\mathcal{M}^m_{S,1}(v,s)|^q] &= \mathbb{E}\left[\left|\int_v^s \int_{|x|>|s-v|^{\frac{1}{\alpha}}} \mathcal{U}(w, X^m_{w^-}, x)\tilde{N}^m(dr,dx)\right|^q\right] \\
&\leq C_q \mathbb{E}\left[\left(\sum_{v\leq w\leq s} |\mathcal{U}(w, X^m_{w^-}, \Delta\mathcal{W}^m_w)|^2 \mathbf{1}_{|\Delta\mathcal{W}^m_w|>|v-s|^{\frac{1}{\alpha}}}\right)^{\frac{q}{2}}\right] \\
&\leq C_q (s-v)^{1+q\frac{\theta-1}{\alpha}} \int_{|x|>|v-s|^{\frac{1}{\alpha}}} |x|^q \nu(dx) \\
&\leq C_q |v-s|^{\frac{q}{\alpha}+\frac{\theta-1}{\alpha}}.
\end{aligned}$$

In the diagonal regime (i.e. for $\mathcal{M}^m_{S,2}(v,s)$) we use the BDG inequality and (4.5) to recover integrability w.r.t. the Lévy measure and then use the additional integrability to obtain better estimate. Namely:

$$\begin{aligned}
\mathbb{E}[|\mathcal{M}^m_{S,2}(v,s)|^q] &= C\mathbb{E}\left[\left|\int_v^s \int_{|x|\leq|v-s|^{\frac{1}{\alpha}}} \mathcal{U}(w, X^m_{w^-}, x)\tilde{N}^m(dw,dx)\right|^q\right] \\
&\leq C_q \left(\int_v^s \int_{|x|\leq|v-s|^{\frac{1}{\alpha}}} |\mathcal{U}(w, X^m_{w^-}, x)|^2 dw\nu(dx)\right)^{\frac{q}{2}} \\
&\leq C_q \left((s-v)^{1+2\frac{\theta-1}{\alpha}} \int_{|x|\leq|v-s|^{\frac{1}{\alpha}}} |x|^2 \nu(dx)\right)^{\frac{q}{2}} \\
&\leq C_q (s-v)^{\frac{q}{\alpha}+q\frac{\theta-1}{\alpha}}.
\end{aligned}$$

Using the above estimates on the q-moments of $\mathcal{M}^m_L(v,s)$, $\mathcal{M}^m_{S,1}(v,s)$ and $\mathcal{M}^m_{S,2}(v,s)$ the statement follows passing to the limit in m.

(iii) Letting $(\mathcal{F}^m_v)_{v\geq 0} := \big(\sigma((X^m_w, \mathcal{W}^m_w)_{0\leq w\leq v})\big)_{v\geq 0}$, restarting from (4.3) and taking the conditional expectation w.r.t. $\mathcal{F}^m$ yields

$$\mathbb{E}[X^m_s - X^m_v|\mathcal{F}^m_v] = \mathbb{E}[u_m(v, X^m_v) - u_m(s, X^m_v)|\mathcal{F}^m_v] = u^s_m(v, X^m_v) - X^m_v.$$

Passing to the limit in m, it can be deduced that

$$\mathbb{E}[X_s - X_v|\mathcal{F}_v] = u^s(v, X_v) - X_v =: \mathfrak{f}(v, X_v, s-v),$$

where u is the mild solution of (2.1) with terminal condition x at time s and source term $f \equiv 0$. From the mild definition of u in Theorem 8 we obtain that for all $(w,y) \in [s,v]\times\mathbb{R}^d$:

$$\begin{aligned}
Du(w,y) &= \int_{\mathbb{R}^d} dy'\{y'\otimes Dp_\alpha(s-w, y'-y)\} + \int_w^s dw' \int_{\mathbb{R}^d} dy[Du(w',y')F(w',y')]\otimes Dp_\alpha(w'-w, y'-y) \\
&= I + \int_w^s dw' \int_{\mathbb{R}^d} dy'[Du(w',y')F(w',y')]\otimes Dp_\alpha(w'-w, y'-y),
\end{aligned}$$

integrating by parts to derive the last inequality. We thus get:

$$\begin{aligned}
\mathbb{E}[X_s - X_v|\mathcal{F}_v] &= u(v,X_v) - u(s,X_v) = \int_v^s dw \int_{\mathbb{R}^d} dy Du(w,y)F(w,y)p_\alpha(w-v, y-X_v) \\
&= \int_v^s dw \int_{\mathbb{R}^d} dy F(w,y)p_\alpha(w-v, y-X_v) \\
&\quad + \int_v^s dw \int_{\mathbb{R}^d} dy \int_w^s dw' \int_{\mathbb{R}^d} dy'\big[[Du(w',y')F(w',y')]\otimes D_y p_\alpha(w'-w, y'-y)\big]F(w,y) \\
&\quad \times p_\alpha(w-v, y-X_v), \qquad (4.6)
\end{aligned}$$

where we have again plugged the mild formulation of Du. Let us first prove that the first term in the above

has the right order. Thanks to Lemma 10 (with $\eta = 0$ and $\Psi = \mathrm{Id}$ therein) that:

$$\begin{aligned} & \mathscr{F}(v, X_v, s-v) \\ := \; & \Big| \int_v^s dw \int_{\mathbb{R}^d} dy F(w,y) p_\alpha(w-v, y-X_v) \Big| \\ \leq \; & C \|F\|_{\mathbb{L}^r([0,T], \mathbb{B}_{p,q}^{-1+\gamma})} (s-v)^{1-(\frac{1}{r}+\frac{d}{p\alpha}+\frac{1-\gamma}{\alpha})} \\ \leq \; & C \|F\|_{\mathbb{L}^r([0,T], \mathbb{B}_{p,q}^{-1+\gamma})} (s-v)^{\frac{1}{2}+\left[\frac{1}{2}-(\frac{1}{r}+\frac{d}{p\alpha}+\frac{1-\gamma}{\alpha})\right]}. \end{aligned} \tag{4.7}$$

Let us now prove that $\chi := \frac{1}{2} - (\frac{1}{r} + \frac{d}{p\alpha} + \frac{1-\gamma}{\alpha}) > 0$. Recall that we have assumed in Theorem 1 that $\gamma > [3-\alpha(1-\frac{1}{r})+\frac{d}{p}]/2$. Note carefully that, for $\alpha > (1-\frac{d}{p})/(1-\frac{1}{r})$ it also holds that $\gamma > [3-\alpha(1-\frac{1}{r})+\frac{d}{p}]/2 > 2-\alpha+\alpha/r+d/p$ which was the natural condition appearing in the analysis of the Green kernel to give a pointwise meaning to the underlying gradient. This eventually gives that $\chi > 0$.

Let us now prove that the second in the r.h.s. of (4.6) is a negligible perturbation. Setting with the notations of Section 3.3:

$$\begin{aligned} \psi_{v,w,s}(y) \; &:= \; p_\alpha(w-v, y-X_v) \int_w^s dw' \int_{\mathbb{R}^d} dy' [Du(w',y') F(w',y')] \otimes D_y p_\alpha(w'-r, y'-y) \\ &= \; p_\alpha(w-v, y-X_v) D\mathfrak{r}(w,y) \end{aligned}$$

we write:

$$\mathscr{R}(v, X_v, s-v) := \int_v^s dw \int_{\mathbb{R}^d} dy \psi_{v,w,s}(y) F(w,y).$$

We thus have the following estimate:

$$|\mathscr{R}(v, X_v, s-v)| \leq \|F\|_{\mathbb{L}^r([0,T], \mathbb{B}_{p,q}^{-1+\gamma})} \|\psi_{v,\cdot,s}(\cdot)\|_{\mathbb{L}^{r'}([0,T], \mathbb{B}_{p',q'}^{1-\gamma})}. \tag{4.8}$$

Let us now consider the thermic part of $\|\psi_{v,\cdot,s}(\cdot)\|_{\mathbb{L}^{r'}([0,T], \mathbb{B}_{p',q'}^{1-\gamma})}$. With the same previous notations[1]:

$$\begin{aligned} \Big(\mathcal{T}_{p',q'}^{1-\gamma}(\psi_{v,w,s}(\cdot))\Big|_{[(w-v),1]} \Big)^{q'} \; &\leq \; C(w-v)^{-\frac{1-\gamma}{\alpha}q'} \|D\mathfrak{r}(w,\cdot)\|_\infty^{q'} \|p_\alpha(w-v,\cdot,-X_v)\|_{\mathbb{L}^{p'}}^{q'} \\ &\leq \; C(w-v)^{-\frac{1-\gamma}{\alpha}q'} (s-w)^{\frac{(\theta-1)}{\alpha}q'} (w-v)^{-\frac{d}{\alpha p}q'}, \end{aligned} \tag{4.9}$$

using (3.31) and (3.18) for the last inequality. Hence,

$$\Big(\int_v^s dw \Big(\mathcal{T}_{p',q'}^{1-\gamma}(\psi_{v,w,s}(\cdot))\Big|_{[(w-v),1]} \Big)^{r'} \Big)^{1/r'} \; \leq \; C(s-v)^{\frac{1}{r'}+\frac{\theta-1}{\alpha}-\frac{d}{\alpha p}-\frac{1-\gamma}{\alpha}}. \tag{4.10}$$

Observe that, for this term to be a remainder on small time intervals, we need:

$$\frac{1}{r'} + \frac{\theta-1}{\alpha} - \frac{d}{\alpha p} - \frac{1-\gamma}{\alpha} > 1 \iff \gamma - 1 + \theta - 1 - \frac{d}{p} - \frac{\alpha}{r} > 0.$$

Recalling the definition of θ in (1.10), we obtain the condition:

$$\gamma > \frac{3-\alpha+\frac{2d}{p}+\frac{2\alpha}{r}}{2}. \tag{4.11}$$

This stronger condition appears only in the case where one is interested in expliciting exactly the dynamics in terms of a drift which actually writes as the mollified version of the initial one along the density of the driving noise (regularizing kernel). Note that if one chooses to work in a bounded setting, i.e. for $p = r = \infty$, (4.11) again corresponds to the condition appearing in Theorem 1.

[1] Pay attention that, in order to absorb some singularities we cannot here directly appeal to Lemma 10 but simply exploit some $\mathbb{L}^\infty$ of $D\mathfrak{r}(t,\cdot)$ in terms of $(T-t)^{\frac{\theta}{\alpha}}$.

Let us now deal with the second term from the thermic characterization. Restarting from (3.33) and (3.1), exploiting as well (3.38), we get for $\beta = \theta - 1 - \varepsilon$:

$$\begin{aligned} & |D\mathfrak{r}(w,y)p_\alpha(w-v,y-x) - D\mathfrak{r}(w,z)p_\alpha(w-v,z-x)| \qquad (4.12)\\ \leq\ & C\left[\left(\|D\mathfrak{r}(w,\cdot)\|_{\dot{\mathbb{B}}^\beta_{\infty,\infty}} + \frac{\|D\mathfrak{r}(w,\cdot)\|_{\mathbb{L}^\infty}}{(r-v)^{\frac{\beta}{\alpha}}}\right)(q_\alpha(w-v,y-x)+q_\alpha(w-v,z-x))\right]|y-z|^\beta \\ \leq\ & C\Big((s-w)^{\frac{\varepsilon}{\alpha}} + \frac{(s-w)^{\frac{\theta-1}{\alpha}}}{(w-v)^{\frac{\beta}{\alpha}}}\Big)(q_\alpha(w-v,y-x)+q_\alpha(w-v,z-x))\,|y-z|^\beta, \end{aligned}$$

recalling also (3.31) for the last inequality and denoting by $\|\cdot\|_{\dot{\mathbb{B}}^\beta_{\infty,\infty}}$ the homogeneous Besov norm (Hölder modulus of order β). Hence:

$$\begin{aligned} \Big(\mathcal{T}^{1-\gamma}_{p',q'}(\psi_{v,w,s}(\cdot))\Big|_{[0,(w-v)]}\Big)^{q'} &\leq \frac{C}{(w-v)^{(\frac{d}{p\alpha})q'}}\int_0^{w-v}\frac{d\bar v}{\bar v}\bar v^{(\frac{\gamma-1+\beta}{\alpha})q'}\Big((s-w)^{\frac{\varepsilon}{\alpha}} + \frac{(s-w)^{\frac{\theta-1}{\alpha}}}{(w-v)^{\frac{\beta}{\alpha}}}\Big)^{q'},\\ \Big(\int_v^s dw\Big(\mathcal{T}^{1-\gamma}_{p',q'}(\psi_{v,w,s}(\cdot))\Big|_{[0,(w-v)]}\Big)^{r'}\Big)^{1/r'} &\leq \Big(\int_v^s dw(w-v)^{(\frac{\gamma-1+\beta}{\alpha}-\frac{d}{p\alpha})r'}\Big((s-w)^{\frac{\varepsilon}{\alpha}} + \frac{(s-w)^{\frac{\theta-1}{\alpha}}}{(w-v)^{\frac{\beta}{\alpha}}}\Big)^{r'}\Big)^{1/r'}\\ &\leq C(s-v)^{\frac{1}{r'}+(\frac{\gamma-1+\beta}{\alpha}-\frac{d}{p\alpha})+\frac{\varepsilon}{\alpha}} = C(s-v)^{\frac{1}{r'}+(\frac{\gamma-1+\theta-1}{\alpha}-\frac{d}{p\alpha})}, \qquad (4.13)\end{aligned}$$

which precisely gives a contribution homogeneous to the one of (4.10).We eventually derive that, under the condition (4.11), the remainder in (4.8) is s.t. there exists $\varepsilon' := -\frac{1}{r}+(\frac{\gamma-1+\theta-1}{\alpha}-\frac{d}{p\alpha}) > 0$ for which

$$|\mathscr{R}(v,X_v,s-v)| \leq C(s-v)^{1+\varepsilon'}, C := C(\|F\|_{\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q})}). \qquad (4.14)$$

□

Having this result at hand, one can now appeal to the construction implemented in Section 4.4 of [DD16] in order to conclude the proof of Theorem 3. Let us try to sum up how such a construction can be adapted in our setting. As in Section 4.4.1 of [DD16], we introduce in a generic way the process $(A(s,t))_{0\leq s\leq t\leq T}$ as $(i)A(t,t+h) = X_{t+h} - X_t$ or $(ii)A(t,t+h) = \mathcal{W}_{t+h} - \mathcal{W}_t$ or $(iii)A(t,t+h) = \mathfrak{f}(t,X_t,h)$. We then claim that the following estimates hold: for any $1 \leq q < \alpha$ there exists $\varepsilon_0 \in (0, 1-1/\alpha]$, $\varepsilon_1, \varepsilon_1' > 0$ and a constant $C := C(p,q,r,\gamma,q,T) > 0$ such that

$$\begin{aligned} \mathbb{E}[|\mathbb{E}[A(t,t+h)|\mathcal{F}_t]|^q]^{\frac{1}{q}} &\leq Ch^{\frac{1}{\alpha}+\varepsilon_0},\\ \mathbb{E}[|A(t,t+h)|^q]^{\frac{1}{q}} &\leq Ch^{\frac{1}{\alpha}},\\ \mathbb{E}[|\mathbb{E}[A(t,t+h)+A(t+h,t+h')-A(t,t+h')|\mathcal{F}_t|^q]^{\frac{1}{q}}] &\leq C(h')^{1+\varepsilon_1},\\ \mathbb{E}[|A(t,t+h)+A(t+h,t+h')-A(t,t+h')|^q]^{\frac{1}{q}} &\leq C(h')^{\frac{1}{\alpha}+\varepsilon_1'}. \qquad (4.15)\end{aligned}$$

Then, we aim at defining for any $T > 0$ the stochastic integral $\int_0^T \psi_s A(t,t+dt)$, for the class of predictable processes $(\psi_s)_{s\in[0,t]}$, $((1-1/\alpha)-\varepsilon_2)$-Hölder continuous in $\mathbb{L}^{q'}$ with $q' \geq 1$ such that $1/q' + 1/q = 1/\ell$, $\ell < \alpha$ and $0 < \varepsilon_2 < \varepsilon_0$, as an $\mathbb{L}^\ell$ limit of the associated Riemann sum: for $\Delta = \{0 = t_0 < t_1, \ldots, t_N = T\}$

$$S(\Delta) := \sum_{i=0}^{N-1}\psi_{t_i}A(t_i,t_{i+1}) \to \int_0^T \psi_s A(t,t+dt), \quad \text{in } \mathbb{L}^\ell, \qquad (4.16)$$

which justifies the fact that such an integral is called $\mathbb{L}^\ell$ stochastic-Young integral by the Authors. To do so, the main idea in [DD16] consists in splitting the process A as the sum of a drift and a martingale:

$$A(t,t+h) = A(t,t+h) - \mathbb{E}[A(t,t+h)|\mathcal{F}_t] + \mathbb{E}[A(t,t+h)|\mathcal{F}_t] := M(t,t+h) + R(t,t+h), \qquad (4.17)$$

and define $\mathbb{L}^\ell$-stochastic-Young integral w.r.t. each of these terms. We then have

Theorem 12 (Theorem 16 of [DD16])**.** *There exists $C = C(q,q',p,q,r,\gamma) > 0$ such that, given two subdivisions $\Delta \subset \Delta'$ of $[0,T]$, such that $\pi(\Delta) < 1$,*

$$\|S(\Delta) - S(\Delta')\|_{\mathbb{L}^\ell} \leq C\max\{T^{1/\alpha}, T\}(\pi(\Delta))^\eta, \qquad (4.18)$$

where $\pi(\Delta)$ denotes the step size of the subdivision Δ and with $\eta = \min\{\epsilon_0 - \varepsilon_2, \varepsilon_1, \varepsilon_1'\}$.

Proof. The main point consists in noticing that the proof in [DD16] remains valid in our setting (for parameter $\ell = p$ therein) and that the only difference is the possible presence of jumps. To handle that, the key idea is then to split the martingale part (which in our current framework may involve jumps) into two parts: an $\mathbb{L}^2$-martingale (which includes the compensated small jumps) and an $\mathbb{L}^\ell$-martingale (which includes the compensated large jumps). The first part can be handled using the BDG inequality (and this is what is done in [DD16]) and the other part by using the compensation formula (such a strategy is somehow classical in the pure-jump setting and has been implemented to prove point *(ii)* in Proposition 11 above). □

Thus, we obtain that for any fixed t in $[0,T]$ we are able to define an additive (on $[0,T]$) integral $\int_0^t \psi_s A(s,s+ds)$. The main point consists now in giving a meaning on this quantity as a process (i.e. that all the time integrals can be defined simultaneously). In the current pure-jump setting, we rely on the Aldous criterion, whereas in the diffusive framework of [DD16], the Kolmogorov continuity criterion was used. Thanks to Theorem 12, one has

$$\Big\| \int_t^{t+h} \psi_s A(s,s+ds) - \psi_t A(t,t+dt) \Big\|_{\mathbb{L}^\ell} \leq C h^{\frac{1}{\alpha}+\eta}, \tag{4.19}$$

so that one can apply Proposition 34.9 in Bass [Bas11] and Proposition 4.8.2 in Kolokoltsov [Kol11] to the sequence $\Big(\int_0^t \psi_s A(s,s+ds) \Big)_{s\leq t}$ and deduce that the limit is stochastically continuous.

Eventually, following Section 4.6 of [DD16] we can thus define the the processes $\Big(\int_0^t \psi_s dX_s \Big)_{0\leq t\leq T}$ and $\Big(\int_0^t \psi_s \mathfrak{f}(s,X_s,ds) \Big)_{0\leq t\leq T}$ for any $(\psi_s)_{0\leq s\leq T}$ with $\varepsilon_2 < (\theta-1)/\alpha$. Let us conclude by emphasizing the following fact underlined in [DD16]. When building the $\mathbb{L}^\ell$ stochastic-Young version of the drift, one has from (4.17) that

$$R(t,t+h) = \mathbb{E}[X_{t+h} - X_t | \mathcal{F}_t], \quad M(t,t+h) = X_{t+h} - X_t - \mathbb{E}[X_{t+h} - X_t | \mathcal{F}_t].$$

Thanks to Proposition 11 we have that $\Big(\int_0^t \psi_s R(t,t+dt) \Big)_{0\leq t\leq T} = \Big(\int_0^t \psi_s \mathfrak{f}(t,X_t,dt) \Big)_{0\leq t\leq T}$, so that the l.h.s. is well defined. Also, we have that $\Big(\int_0^t \psi_s (R(t,t+dt) - \mathscr{F}(t,X_t,dt)) \Big)_{0\leq t\leq T} = \Big(\int_0^t \psi_s \mathfrak{r}(t,X_t,dt) \Big)_{0\leq t\leq T}$ is well defined and is null since the bound appearing in the increment of the l.f.s. is greater than one. Hence,

$$\Big(\int_0^t \psi_s \mathfrak{f}(t,X_t,dt) \Big)_{0\leq t\leq T} = \Big(\int_0^t \psi_s \mathscr{F}(t,X_t,dt) \Big)_{0\leq t\leq T}.$$

On the other hand, we have that $\Big(\int_0^t \psi_s M(t,t+dt) \Big)_{0\leq t\leq T}$ is well defined as well and that $\Big(\int_0^t \psi_s M(t,t+dt) - d\mathcal{W}_t \Big)_{0\leq t\leq T} = \Big(\int_0^t \psi_s \hat{M}(t,t+dt) \Big)_{0\leq t\leq T}$ where

$$\hat{M}(t,t+h) = X_{t+h} - X_t - (\mathcal{W}_{t+h} - \mathcal{W}_t) - \mathbb{E}[X_{t+h} - X_t - (\mathcal{W}_{t+h} - \mathcal{W}_t) | \mathcal{F}_t],$$

is an $\mathbb{L}^q$ martingale with q moment bounded by $C_q h^{q[1+(\theta-1)/\alpha]}$ so that it is null as well, meaning that when reconstructing the drift as above, we indeed get that only the "original" noise part in the dynamics matter. □

Proof of (4.15). The proof follows from Proposition 11 and Theorem 8. Note that the two last estimates are equals to 0 in case $(i)-(ii)$ since the process A is additive. □

We eventually conclude this part with the following Lemma.

Lemma 13. *Under the previous assumptions we have that for any smooth functions $(F_m)_{m\in\mathbb{N}}$ satisfying*

$$\lim_{m\to\infty} \|F - F_m\|_{\mathbb{L}^r([0,T],\mathbb{B}_{p,q}^{-1+\gamma}(\mathbb{R}^d))} = 0,$$

that for all t in $[0,T]$,

$$\lim_{m\to\infty} \Big\| \int_0^t \psi_s \mathscr{F}(s,X_s,ds) - \int_0^t \psi_s F_m(s,X_s) ds \Big\|_{\mathbb{L}^\ell} = 0 \tag{4.20}$$

Proof. We want to investigate:

$$\lim_{m\to\infty} \mathbb{E} \left| \int_0^t \psi_s \mathscr{F}(s,X_s,ds) - \int_0^t \psi_s F_m(s,X_s) ds \right|^\ell. \tag{4.21}$$

Coming back to the definition of such integrals, this means that we want to control

$$\lim_{m\to\infty}\mathbb{E}\left|\lim_{N\to\infty}\sum_{i=0}^{N-1}\psi_{t_i}\int_{t_i}^{t_{i+1}}ds\left\{\int dyF(s,y)p_\alpha(s-t_i,y-X_{t_i})-F_m(t_i,X_{t_i})\right\}\right|^\ell.$$

We have the following decomposition:

$$\begin{aligned}
&\lim_{m\to\infty}\mathbb{E}\left|\lim_{N\to\infty}\sum_{i=0}^{N-1}\psi_{t_i}\int_{t_i}^{t_{i+1}}ds\left\{\int dyF(s,y)p_\alpha(s-t_i,y-X_{t_i})-F_m(t_i,X_{t_i})\right\}\right|^\ell\\
\leq\ &\lim_{m\to\infty}\mathbb{E}\left|\lim_{N\to\infty}\sum_{i=0}^{N-1}\psi_{t_i}\int_{t_i}^{t_{i+1}}ds\left\{\int dy[F(s,y)-F_m(s,y)]p_\alpha(s-t_i,y-X_{t_i})\right\}\right|^\ell\\
&+\lim_{m\to\infty}\mathbb{E}\left|\lim_{N\to\infty}\sum_{i=0}^{N-1}\psi_{t_i}\int_{t_i}^{t_{i+1}}ds\int dy[F_m(s,y)-F_m(t_i,X_{t_i})]p_\alpha(s-t_i,y-X_{t_i})\right\}\right|^\ell\\
&:=\lim_{m\to\infty}\|\lim_{\pi(\Delta)\to0}S_m^1(\Delta)\|_{\mathbb{L}^\ell}+\lim_{m\to\infty}\|\lim_{\pi(\Delta)\to0}S_m^2(\Delta)\|_{\mathbb{L}^\ell}
\end{aligned}$$

with the previous notations. Note that $\lim_{m\to\infty}\|S_m^1(\Delta)\|_{\mathbb{L}^\ell}=0$, uniformly w.r.t. Δ and that for each m, $\|S_m^1(\Delta)\|_{\mathbb{L}^\ell}$ tends to some $\|S_m^1\|_{\mathbb{L}^\ell}$ as $\pi(\Delta)\to0$. One can hence invert both limits and therefore deduce that

$$\lim_{m\to\infty}\lim_{\pi(\Delta)\to0}\|S_m^1(\Delta)\|_{\mathbb{L}^\ell}=\lim_{\pi(\Delta)\to0}\lim_{m\to\infty}\|S_m^1(\Delta)\|_{\mathbb{L}^\ell}=0.$$

For the second term, we note that due to the regularity of F_m (using e.g. its $\mathbb{L}^r(\mathbb{B}^1_{p,q})$ norm) that

$$\mathbb{E}\left|\int_{t_i}^{t_{i+1}}ds\int dy[F_m(s,y)-F_m(t_i,X_{t_i})]p_\alpha(s-t_i,y-X_{t_i})\right|^\ell\leq C_m(t_{i+1}-t_i)^{\ell(\frac{1}{2}+\frac{1}{\alpha}+\chi)},$$

so that $\lim_{m\to\infty}\lim_{\pi(\Delta)\to0}\|S_m^2(\Delta)\|_{\mathbb{L}^\ell}=0$. This concludes the proof. □

5 Pathwise uniqueness in dimension one

The aim of this part is to prove Theorem 5, adapting to this end the proof of Proposition 2.9 in [ABM18] to our current inhomogeneous and parabolic (for the auxiliary PDE concerned) framework. Let us consider $(X^1,\mathcal{W})$ and $(X^2,\mathcal{W})$ two weak solutions of (1.13). With the notations of (2.10), we consider the two corresponding Itô-Zvonkin transforms $X_t^{Z,m,i}:=X_t^i-u_m(t,X_t^i)=x-u_m(0,x)+\mathcal{W}_t-M_{0,t}(\alpha,u_m,X^i)+R_{0,t}(\alpha,F_m,\mathscr{F},X^i)$, $i\in\{1,2\}$. We point out that we here use the mollified PDE, keeping therefore the remainder term and dependence in m for the martingale part. This is mainly to avoid passing to the limit for the martingale term (as Athreya *et al.* [ABM18] do but which requires many additional technical lemmas therein). Of course, we will have to control the remainders, which is precisely possible from Lemma 13. From now on, we assume that $\alpha<2$. The case $\alpha=2$ is indeed easier and can be handled following the arguments below.

As a starting point, we now expand, for a smooth approximation of the absolute value

$$V_n(x)=\begin{cases}|x|,\ |x|\geq\frac{1}{n},\\ \frac{3}{8n}+\frac{3}{4}nx^2-\frac{1}{8}n^3x^4,\ |x|\leq\frac{1}{n},\end{cases}$$

the quantity $V_n(X_t^{Z,m,1}-X_t^{Z,m,2})$ approximating $|X_t^{Z,m,1}-X_t^{Z,m,2}|$. For fixed m,n we can apply Itô's formula

to obtain:

$$
\begin{aligned}
& V_n(X_t^{Z,m,1} - X_t^{Z,m,2}) \\
= \;& V_n(0) + \int_0^t V_n'(X_t^{Z,m,1} - X_t^{Z,m,2})\big[\mathscr{F}(s, X_s^1, ds) - F_m(s, X_s^1)ds - (\mathscr{F}(s, X_s^2, ds) - F_m(s, X_s^2)ds)\big] \\
& + \int_0^t [V_n(X_s^{Z,m,1} - X_s^{Z,m,2} + h_m(X_s^1, X_s^2, r)) - V_n(X_t^{Z,m,1} - X_t^{Z,m,2})]\tilde{N}(ds, dr) \\
& + \int_0^t \int_{|r|\geq 1} \psi_n(X_s^{Z,m,1} - X_s^{Z,m,2}, h_m(X_s^1, X_s^2, r))\nu(dr)ds \\
& + \int_0^t \int_{|r|\leq 1} \psi_n(X_s^{Z,m,1} - X_s^{Z,m,2}, h_m(X_s^1, X_s^2, r))\nu(dr)ds \\
=: \;& \frac{3}{8n} + \Delta R_{0,t}^{m,n} + \Delta M_{0,t}^{m,n} + \Delta C_{0,t,L}^{m,n} + \Delta C_{0,t,S}^{m,n}, \qquad (5.1)
\end{aligned}
$$

recalling that $X_0^{Z,m,1} = X_0^{Z,m,2}$, using the definition of V_n and denoting for all $(x_1, x_2, r) \in \mathbb{R}^3$:

$$
\begin{aligned}
h_m(x_1, x_2, r) &= u_m(x_1 + r) - u_m(x_1) - [u_m(x_2 + r) - u_m(x_2)], \qquad (5.2) \\
\psi_n(x_1, r) &= V_n(x_1 + r) - V_n(x_1) - V_n'(x_1)r.
\end{aligned}
$$

The point is now to take the expectations in (5.1). Since $\Delta M_{0,t}^{m,n}$ is a martingale, we then readily get $\mathbb{E}[\Delta M_{0,t}^{m,n}] = 0$. On the other hand, since $|V_n'(x)| \leq 2$, we also have from Lemma 13 that:

$$
\mathbb{E}[|\Delta R_{0,t}^{m,n}|] \underset{m}{\to} 0. \qquad (5.3)
$$

It now remains to handle the compensator terms. For the *large* jumps, we readily write:

$$
\mathbb{E}[|\Delta C_{0,t,L}^{m,n}|] \leq 2\|V_n'\|_\infty \|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)} \int_0^t \mathbb{E}[|X_s^1 - X_s^2|]ds \leq C \int_0^t \mathbb{E}[|X_s^1 - X_s^2|]ds, \qquad (5.4)
$$

observing that $|h_m(x_1, x_2, r)| \leq 2\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}|x_1 - x_2|$. Also, from Corollary 7, $\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)} \leq C_T \underset{T\to 0}{\longrightarrow} 0$ uniformly in m (as the terminal condition of the PDE is 0). In particular, for T small enough one has $\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)} \leq 1/4$ and

$$
|x_1 - u_m(t, x_1) - (x_2 - u_m(t, x_2))| \geq |x_1 - x_2| - |u_m(t, x_1) - u_m(t, x_2)| \geq |x_1 - x_2|(1 - \|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}) \geq \frac{3}{4}|x_1 - x_2|. \qquad (5.5)
$$

Hence,

$$
|h_m(X_s^1, X_s^2, r)| \leq 2\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}|X_s^1 - X_s^2| \leq \frac{2}{3}|X_s^{Z,m,1} - X_s^{Z,m,2}|. \qquad (5.6)
$$

Therefore, if $|X_s^{Z,m,1} - X_s^{Z,m,2}| \geq 3/n$, it is readily seen that $\psi_n(X_s^{Z,m,1} - X_s^{Z,m,2}, h_m(X_s^1, X_s^2, r)) = 0$. We thus have:

$$
\begin{aligned}
|\mathbb{E}[C_{0,t,S}^{m,n}]| &= \Big|\mathbb{E}[\int_0^t \int_{|r|\leq 1} \mathbb{I}_{|X_s^{Z,m,1} - X_s^{Z,m,2}| \leq \frac{3}{n}} \psi_n(X_s^{Z,m,1} - X_s^{Z,m,2}, h_m(X_s^1, X_s^2, r))\nu(dr)ds]\Big| \\
&\leq Cn\mathbb{E}[\int_0^t \int_{|r|\leq 1} \mathbb{I}_{|X_s^{Z,m,1} - X_s^{Z,m,2}| \leq \frac{3}{n}} |h_m(X_s^1, X_s^2, r)|^2 \nu(dr)ds], \qquad (5.7)
\end{aligned}
$$

using for the last inequality the definition of V_n which gives that there exists C s.t. for all $y \in \mathbb{R}$, $|V_n''(y)| \leq Cn|y|^2$. We now use the definition of h_m and the smoothness of u_m in order to balance the explosive contribution in n and to keep an exponent of r which allows to integrate the small jumps. From (5.2) and usual interpolation techniques (see e.g. Lemma 5.5 in [ABM18] or Lemma 4.1 in [Pri12]) we get:

$$
|h_m(X_s^1, X_s^2, r)| \leq \|u_m\|_{\mathbb{L}^\infty(\mathbb{B}^\gamma_{\infty,\infty})}|X_s^1 - X_s^2|^{\eta_1} r^{\eta_2}, \;\; (\eta_1, \eta_2) \in (0, 1)^2, \;\; \eta_1 + \eta_2 = \eta < \theta - \varepsilon.
$$

The point is now to apply the above identity with γ_1 large enough in order to get rid of the explosive term in (5.7) (i.e. $\eta_1 > 1/2$) and with γ_2 sufficiently large in order to guarantee the integrability of the Lévy measure (i.e. $\eta_2 > \alpha/2$). This suggests to choose $\eta_1 = 1/2 + \tilde{\varepsilon}/2$ and $\eta_2 = \alpha/2 + \tilde{\varepsilon}/2$, with $\tilde{\varepsilon} > 0$ meant to be small. In order to satisfy such constraints, we obtain that γ must satisfy $\gamma > [3 - \alpha + 2d/p + 2\alpha/r]/2$, which is

precisely the thresholds appearing when reconstructing the dynamics (see condition (1.12) in Theorem 3 and computations leading to (4.11) in the proof of Proposition 11). Hence,

$$\begin{aligned}|\mathbb{E}[C^{m,n}_{0,t,S}]| &\leq Cn\mathbb{E}\left[\int_0^t\int_{|r|\leq 1}\mathbb{I}_{|X^{Z,m,1}_s-X^{Z,m,2}_s|\leq\frac{3}{n}}|X^1_s-X^2_s|^{1+\tilde{\varepsilon}}r^{\alpha+\tilde{\varepsilon}}\frac{dr}{r^{1+\alpha}}ds\right]\\ &\leq Cn\mathbb{E}\left[\int_0^t\mathbb{I}_{|X^{Z,m,1}_s-X^{Z,m,2}_s|\leq\frac{3}{n}}|X^{Z,m,1}_s-X^{Z,m,2}_s|^{1+\tilde{\varepsilon}}ds\right]\leq Cn^{-\tilde{\varepsilon}},\end{aligned}\tag{5.8}$$

using (5.5) and the definition of $(X^{Z,m,i})_{i\in\{1,2\}}$ for the last but one inequality. Plugging (5.8), (5.4) into (5.1) (taking therein the expectations) and recalling that $\mathbb{E}[\Delta M^{m,n}_{0,t}]=0$, eventually yields:

$$\mathbb{E}[V_n(X^{Z,m,1}_t-X^{Z,m,2}_t)]\leq\frac{3}{8n}+\mathbb{E}[|\Delta R^{m,n}_{0,t}|]+C\int_0^t\mathbb{E}[|X^1_s-X^2_s|]ds+\frac{C}{n^{\tilde{\varepsilon}}}.$$

Passing to the limit, first in m recalling that $\mathbb{E}[|\Delta R^{m,n}_{0,t}|]\underset{m}{\to}0$ uniformly in n, gives (from the smoothness properties of $(u_m)_{m\geq1}$ in Proposition 6, see also point **(ii)** in Section 2.2):

$$\mathbb{E}[V_n(X^{Z,1}_t-X^{Z,2}_t)]\leq\frac{3}{8n}+C\int_0^t\mathbb{E}[|X^1_s-X^2_s|]ds+\frac{C}{n^{\tilde{\varepsilon}}},\ X^{Z,i}_t:=X^i_t-u(t,X^i_t),\ i\in\{1,2\}.$$

Take now the limit in n and write from (5.5) (which also holds replacing u_m by u):

$$\frac{3}{4}\mathbb{E}[|X^1_t-X^2_t|]\ \leq\ \mathbb{E}[|X^{Z,1}_t-X^{Z,2}_t|]\leq C\int_0^t\mathbb{E}[|X^1_s-X^2_s|]ds,$$

which readily gives from the Gronwall Lemma $\mathbb{E}[|X^1_t-X^2_t|]=0$. □

A Proof of Lemma 10

We start with the proof of estimate (3.12). Having in mind the thermic characterization of the Besov norm (3.9), the main point consists in establishing suitable controls on the thermic part of (3.9) (i.e. the second term in the r.h.s. therein) viewed as the map

$$s\mapsto\mathcal{T}^{1-\gamma}_{p',q'}[\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)].$$

Splitting the interval $[0,1]$ in function of the current time increment $s-t$ (meant to be small) considering $[0,1]=[0,s-t]\cup]s-t,1]$ (low and high cut-off), we write:

$$\begin{aligned}&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)]\Big)^{q'}\\ =\ &\int_0^1\frac{dv}{v}v^{(1-\frac{1-\gamma}{\alpha})q'}\|\partial_v\tilde{p}_\alpha(v,\cdot)\star\big(\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)\big)\|^{q'}_{\mathbb{L}^{p'}}\\ =\ &\int_0^{(s-t)}\frac{dv}{v}v^{(1-\frac{1-\gamma}{\alpha})q'}\|\partial_v\tilde{p}_\alpha(v,\cdot)\star\big(\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)\big)\|^{q'}_{\mathbb{L}^{p'}}\\ &+\int_{(s-t)}^1\frac{dv}{v}v^{(1-\frac{1-\gamma}{\alpha})q'}\|\partial_v\tilde{p}_\alpha(v,\cdot)\star\big(\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)\big)\|^{q'}_{\mathbb{L}^{p'}}\\ =:\ &\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)]|_{[0,(s-t)]}\Big)^{q'}+\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)]|_{[(s-t),1]}\Big)^{q'}.\end{aligned}\tag{A.1}$$

For the high cut-off, the singularity induced by the differentiation of the heat kernel in the thermic part is always integrable. Hence using $\mathbb{L}^1-\mathbb{L}^{p'}$ convolution inequalities we have

$$\begin{aligned}&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)]|_{[(s-t),1]}\Big)^{q'}\\ \leq\ &\int_{(s-t)}^1\frac{dv}{v}v^{(1-\frac{1-\gamma}{\alpha})q'}\|\partial_v\tilde{p}_\alpha(v,\cdot)\|^{q'}_{\mathbb{L}^1}\|\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)\|^{q'}_{\mathbb{L}^{p'}}.\end{aligned}$$

From (9) and similarly to (3.18), we have

$$\|\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)\|_{\mathbb{L}^{p'}} \quad \leq \quad \frac{\bar{C}_{p'}}{(s-t)^{\frac{d}{\alpha p}+\frac{|\eta|}{\alpha}}}.$$

We thus obtain

$$\begin{aligned}
&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)]|_{[(s-t),1]}\Big)^{q'}\\
\leq\ & \|\Psi(s,\cdot)\|^{q'}_{\mathbb{L}^\infty}\frac{C}{(s-t)^{(\frac{d}{p\alpha}+\frac{\eta}{\alpha})q'}}\int_{(s-t)}^1\frac{dv}{v}\frac{1}{v^{\frac{1-\gamma}{\alpha}q'}}\\
\leq\ & \frac{C\|\Psi\|^{q'}_{\mathbb{L}^\infty(\mathbb{L}^\infty)}}{(s-t)^{\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{\eta}{\alpha}\right]q'}}. \qquad (A.2)
\end{aligned}$$

To deal with the low cut-off of the thermic part, we need to smoothen the singularity induced by the differentiation of the heat kernel of the thermic characterization. Coming back to the very definition (A.1) of this term, we note that

$$\begin{aligned}
&\|\partial_v\tilde{p}_\alpha(v,\cdot)\star\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)\|_{\mathbb{L}^{p'}} \qquad (A.3)\\
=\ & \Big(\int_{\mathbb{R}^d}dz\Big|\int_{\mathbb{R}^d}dy\partial_v\tilde{p}_\alpha(v,z-y)\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,y-x)|^{p'}\Big)^{1/p'}\\
=\ & \Big(\int_{\mathbb{R}^d}dz\Big|\int_{\mathbb{R}^d}dy\partial_v\tilde{p}_\alpha(v,z-y)\Big[\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,y-x)-\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,z-x)\Big]\Big|^{p'}\Big)^{1/p'}.
\end{aligned}$$

To smoothen the singularity, one then needs to establish a suitable control on the Hölder moduli of the product $\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)$. We claim that for all $(t<s,x)$ in $[0,T]^2\times\mathbb{R}^d$, for all (y,z) in $(\mathbb{R}^d)^2$:

$$\begin{aligned}
&|\Psi(s,y)\mathscr{D}^\eta p_\alpha(s-t,y-x)-\Psi(s,z)\mathscr{D}^\eta p_\alpha(s-t,z-x)| \qquad (A.4)\\
\leq\ & C\left[\left(\frac{\|\Psi(s,\cdot)\|_{\dot{\mathbb{B}}^\beta_{\infty,\infty}}}{(s-t)^{\frac{\eta}{\alpha}}}+\frac{\|\Psi(s,\cdot)\|_{\mathbb{L}^\infty}}{(s-t)^{\frac{\eta+\beta}{\alpha}}}\right)(q_\alpha(s-t,y-x)+q_\alpha(s-t,z-x))\right]|y-z|^\beta\\
\leq\ & \frac{C}{(s-t)^{\frac{\eta+\beta}{\alpha}}}\|\Psi(s,\cdot)\|_{\mathbb{B}^\beta_{\infty,\infty}}(q_\alpha(s-t,y-x)+q_\alpha(s-t,z-x))\,|y-z|^\beta.
\end{aligned}$$

This readily gives, using $\mathbb{L}^1-\mathbb{L}^{p'}$ convolution estimates and (3.18), that

$$\begin{aligned}
\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Psi(s,\cdot)\mathscr{D}^\eta p(s-t,\cdot-x)]|_{[0,(s-t)]}\Big)^{q'} \quad &\leq \quad \frac{C\|\Psi(s,\cdot)\|^{q'}_{\mathbb{B}^\beta_{\infty,\infty}}}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{\eta}{\alpha}+\frac{\beta}{\alpha}\right]q'}}\int_0^{s-t}\frac{dv}{v}v^{(1-\frac{1-\gamma}{\alpha}-1+\frac{\beta}{\alpha})q'}\\
&\leq \quad \frac{C\|\Psi(s,\cdot)\|^{q'}_{\mathbb{B}^\beta_{\infty,\infty}}}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{\eta}{\alpha}+\frac{\beta}{\alpha}+\frac{1-\gamma-\beta}{\alpha}\right]q'}}. \qquad (A.5)
\end{aligned}$$

Putting together estimates (A.2) and (A.5) into (A.1) yields the estimate (3.12) in Lemma 10.

Remark 15 (On the control of the first term in the r.h.s. (3.9))**.** *This term is easily handled by the $\mathbb{L}^{p'}$ norm of the product $\Psi(s,\cdot)\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)$ and hence on $\mathbb{L}^{p'}$ norm of $\mathscr{D}^\eta p_\alpha$ times the $\mathbb{L}^\infty$ norm of Ψ. This, in view of* (3.18)*, clearly brings only a negligible contribution in comparison with the one of the thermic part.*

To conclude with (3.12), it remains to prove (A.4). From (3.1) (see again the proof of Lemma 4.3 in [HMP19] for details), we claim that there exists C s.t. for all $\beta'\in(0,1]$ and all $(x,y,z)\in(\mathbb{R}^d)^2$,

$$|\mathscr{D}^\eta p_\alpha(s-t,z-x)-\mathscr{D}^\eta p_\alpha(s-t,y-x)|\leq\frac{C}{(s-t)^{\frac{\beta'+\eta}{\alpha}}}|z-y|^{\beta'}\Big(q_\alpha(s-t,z-x)+q_\alpha(s-t,y-x)\Big). \qquad (A.6)$$

Indeed, (A.6) is direct if $|z-y|\geq[1/2](s-t)^{1/\alpha}$ (off-diagonal regime). It suffices to exploit the bound (3.1) for $\mathscr{D}^\eta p_\alpha(s-t,y-x)$ and $\mathscr{D}^\eta p_\alpha(s-t,z-x)$ and to observe that $\big(|z-y|/(s-t)^{1/\alpha}\big)^{\beta'}\geq 1$. If now

$|z-y| \leq [1/2](s-t)^{1/\alpha}$ (diagonal regime), it suffices to observe from (3.6) that, with the notations of the proof of Lemma 9 (see in particular (3.5)), for all $\lambda \in [0,1]$:

$$\begin{aligned}
|\mathscr{D}^\eta Dp_M(s-t, y-x+\lambda(y-z))| &\leq \frac{C_m}{(s-t)^{\frac{\eta+1}{\alpha}}} p_{\bar{M}}(s-t, y-x-\lambda(y-z)) \\
&\leq \frac{C_m}{(s-t)^{\frac{\eta+1+d}{\alpha}}} \frac{1}{\Big(1+\frac{|y-x-\lambda(z-y)|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m} \\
&\leq \frac{C_m}{(s-t)^{\frac{\eta+1+d}{\alpha}}} \frac{1}{\Big(\frac{1}{2}+\frac{|y-x|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m} \leq 2\frac{C_m}{(s-t)^{\frac{\eta+1}{\alpha}}} p_{\bar{M}}(s-t, y-x). \quad \text{(A.7)}
\end{aligned}$$

Therefore, in the diagonal case (A.6) follows from (A.7) and (3.5) writing $|\mathscr{D}^\eta p_\alpha(s-t, z-x) - \mathscr{D}^\eta p_\alpha(s-t, y-x)| \leq \int_0^1 d\lambda |\mathscr{D}^\eta Dp_\alpha(s-t, y-x+\lambda(y-z))\cdot(y-z)| \leq 2C_m(s-t)^{-[(\eta+1)/\alpha]} q_\alpha(s-t, y-x)|z-y| \leq \tilde{C}_m(s-t)^{-[(\eta+\beta')/\alpha]} q_\alpha(s-t, y-x)|z-y|^{\beta'}$ for all $\beta' \in [0,1]$ (exploiting again that $|z-y| \leq [1/2](s-t)^{1/\alpha}$ for the last inequality). We conclude the proof of (A.4) noticing that for all s in $(0,T]$ the map $\mathbb{R}^d \ni y \mapsto \Psi(s,y)$ is β-Hölder continuous and choosing $\beta' = \beta$ in the above estimate.

We now prove (3.13). Splitting again the thermic part of the Besov norm into two parts (high and low cut-off) we write

$$\begin{aligned}
&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big]\Big)^{q'} \\
= &\int_0^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot) \star \Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)\|^{q'}_{\mathbb{L}^{p'}} \\
= &\int_0^{(s-t)} \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot) \star \Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)\|^{q'}_{\mathbb{L}^{p'}} \\
&+\int_{(s-t)}^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot) \star \Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)\|^{q'}_{\mathbb{L}^{p'}} \\
=: &\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)]|_{[0,(s-t)]}\Big)^{q'} \\
&+\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)]|_{[(s-t),1]}\Big)^{q'}.
\end{aligned}$$

Proceeding as we did before for the high cut-off and using (A.6), we have for any β' in $[0,1]$:

$$\begin{aligned}
&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)]|_{[(s-t),1]}\Big)^{q'} \\
\leq &\int_{(s-t)}^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde{p}_\alpha(v,\cdot)\|^{q'}_{\mathbb{L}^1} \|\Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)\|^{q'}_{\mathbb{L}^{p'}} \\
\leq &\frac{C\|\Psi(s,\cdot)\|^{q'}_{\mathbb{L}^\infty}}{(s-t)^{(\frac{d}{p\alpha}+\frac{\eta+\beta'}{\alpha})q'}} \int_{(s-t)}^1 \frac{dv}{v} \frac{1}{v^{\frac{1-\gamma}{\alpha}q'}} |x-x'|^{\beta' q'} \\
\leq &\frac{C\|\Psi(s,\cdot)\|^{q'}_{\mathbb{L}^\infty}}{(s-t)^{\left[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{\eta+\beta'}{\alpha}\right]q'}} |x-x'|^{\beta' q'}.
\end{aligned}$$

To deal with the low cut-off, we proceed as we did for (A.3) in order to smoothen the singularity induced by the differentiation of the thermic kernel. We are hence led to control the Hölder moduli of $\Psi(s,\cdot)\Big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x) - \mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\Big)$. We claim that for any β' in $(0,1]$ and all $(t<s,x)$ in $[0,T]^2\times\mathbb{R}^d$, we have that for all (y,z) in $(\mathbb{R}^d)^2$:

$$\begin{aligned}
&\Big|\Psi u(s,y)\Big(\mathscr{D}^\eta p_\alpha(s-t,y-x) - \mathscr{D}^\eta p_\alpha(s-t,y-x')\Big) - \Psi(s,z)\Big(\mathscr{D}^\eta p_\alpha(s-t,z-x) - \mathscr{D}^\eta p_\alpha(s-t,z-x')\Big)\Big| \\
\leq &\frac{C}{(s-t)^{\frac{\eta+\beta+\beta'}{\alpha}}} \|\Psi(s,\cdot)\|_{\mathbb{B}^\beta_{\infty,\infty}} \Big(q_\alpha(s-t,y-x) + q_\alpha(s-t,z-x) + q_\alpha(s-t,y-x') + q_\alpha(s-t,z-x')\Big) \\
&\times|y-z|^\beta |x-x'|^{\beta'}. \qquad \text{(A.8)}
\end{aligned}$$

Repeating the computations in (A.3) and using the above estimate, we obtain that:

$$\begin{aligned}
&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[\Big(\Psi(s,\cdot)\big(\mathscr{D}^\eta p_\alpha(s-t,\cdot-x)-\mathscr{D}^\eta p_\alpha(s-t,\cdot-x')\big)\Big)]|_{[0,(s-t)]}\Big)^{q'}\\
\leq\ & \frac{C\|\Psi(s,\cdot)\|^{q'}_{\mathbb{B}^\beta_{\infty,\infty}}}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{\eta+\beta'}{\alpha}+\frac{\beta}{\alpha}\right]q'}}\int_0^{(s-t)}\frac{dv}{v}v^{(1-\frac{1-\gamma}{\alpha}-1+\frac{\beta}{\alpha})q'}|x-x'|^{\beta' q'}\leq\frac{C\|\Psi(s,\cdot)\|^{q'}_{\mathbb{B}^\beta_{\infty,\infty}}}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{\eta+\beta'}{\alpha}+\frac{1-\gamma}{\alpha}\right]q'}}|x-x'|^{\beta' q'},
\end{aligned}$$

provided

$$\beta+\gamma>1. \tag{A.9}$$

It thus remains to prove (A.8). It directly follows from (A.6) that:

$$\begin{aligned}
&\Big|\Psi(s,y)\Big(\mathscr{D}^\eta p_\alpha(s-t,y-x)-\mathscr{D}^\eta p_\alpha(s-t,y-x')\Big)-\Psi(s,z)\Big(\mathscr{D}^\eta p_\alpha(s-t,z-x)-\mathscr{D}^\eta p_\alpha(s-t,z-x')\Big)\Big|\\
\leq\ & \|\Psi(s,\cdot)\|_{\dot{\mathbb{B}}^\beta_{\infty,\infty}}|z-y|^\beta\frac{C}{(s-t)^{\frac{\eta+\beta'}{\alpha}}}|x-x|^{\beta'}\big(q_\alpha(s-t,y-x)+q_\alpha(s-t,y-x')\big) \qquad \text{(A.10)}\\
&+\|\Psi(s,\cdot)\|_{\mathbb{L}^\infty}\Big|\big(\mathscr{D}^\eta p_\alpha(s-t,y-x)-\mathscr{D}^\eta p_\alpha(s-t,y-x')\big)-\big(\mathscr{D}^\eta p_\alpha(s-t,z-x)-\mathscr{D}^\eta p_\alpha(s-t,z-x')\big)\Big|.
\end{aligned}$$

Setting:

$$\Delta(s-t,x,x',y,z):=\Big|\big(\mathscr{D}^\eta p_\alpha(s-t,y-x)-\mathscr{D}^\eta p_\alpha(s-t,y-x')\big)-\big(\mathscr{D}^\eta p_\alpha(s-t,z-x)-\mathscr{D}^\eta p_\alpha(s-t,z-x')\big)\Big|,$$

it now remains to control this term. Precisely,

- If $|x-x'|\geq(s-t)^{1/\alpha}/4$, we write:

$$\begin{aligned}
&\Delta(s-t,x,x',y,z) \qquad \text{(A.11)}\\
\leq\ & \big|\mathscr{D}^\eta p_\alpha(s-t,y-x)-\mathscr{D}^\eta p_\alpha(s-t,z-x)\big|+\big|\mathscr{D}^\eta p_\alpha(s-t,y-x')-\mathscr{D}^\eta p_\alpha(s-t,z-x')\big|\\
\underset{\text{(A.6)}}{\leq}\ & \frac{C}{(s-t)^{\frac{\eta+\beta}{\alpha}}}|y-z|^\beta\big(q_\alpha(s-t,y-x)+q_\alpha(s-t,y-x')+q_\alpha(s-t,z-x)+q_\alpha(s-t,z-x')\big)\\
\leq\ & \frac{4C}{(s-t)^{\frac{\eta+\beta+\beta'}{\alpha}}}|y-z|^\beta|x-x'|^{\beta'}\big(q_\alpha(s-t,y-x)+q_\alpha(s-t,y-x')+q_\alpha(s-t,z-x)+q_\alpha(s-t,z-x')\big).
\end{aligned}$$

- If $|z-y|\geq(s-t)^{1/\alpha}/4$, we write symmetrically:

$$\begin{aligned}
&\Delta(s-t,x,x',y,z) \qquad \text{(A.12)}\\
\leq\ & \big|\mathscr{D}^\eta p_\alpha(s-t,y-x)-\mathscr{D}^\eta p_\alpha(s-t,y-x')\big|+\big|\mathscr{D}^\eta p_\alpha(s-t,z-x)-\mathscr{D}^\eta p_\alpha(s-t,z-x')\big|\\
\underset{\text{(A.6)}}{\leq}\ & \frac{C}{(s-t)^{\frac{\eta+\beta'}{\alpha}}}|x-x'|^{\beta'}\big(q_\alpha(s-t,y-x)+q_\alpha(s-t,y-x')+q_\alpha(s-t,z-x)+q_\alpha(s-t,z-x')\big)\\
\leq\ & \frac{4C}{(s-t)^{\frac{\eta+\beta+\beta'}{\alpha}}}|y-z|^\beta|x-x'|^{\beta'}\big(q_\alpha(s-t,y-x)+q_\alpha(s-t,y-x')+q_\alpha(s-t,z-x)+q_\alpha(s-t,z-x')\big).
\end{aligned}$$

- If $|z-y|\leq(s-t)^{1/\alpha}/4$ and $|x-x'|\leq(s-t)^{1/\alpha}/4$, we get:

$$\begin{aligned}
&\Delta(s-t,x,x',y,z) \qquad \text{(A.13)}\\
\leq\ & \int_0^1 d\lambda\int_0^1 d\mu|D_x^2\mathscr{D}^\eta p_\alpha(s-t,z-x'+\mu(y-z)-\lambda(x-x'))||x-x'||z-y|\\
\leq\ & \frac{C}{(s-t)^{\frac{\eta+\beta+\beta'}{\alpha}}}|y-z|^\beta|x-x'|^{\beta'}\big(q_\alpha(s-t,y-x)+q_\alpha(s-t,y-x')+q_\alpha(s-t,z-x)+q_\alpha(s-t,z-x')\big)
\end{aligned}$$

proceeding as in (A.7) and exploiting (3.5) for the last identity. Plugging (A.13), (A.12) and (A.11) into (A.10) eventually yields the control (A.8).

□

Acknowledgments.

For the first Author, this work has been partially supported by the ANR project ANR-15-IDEX-02. For the second author, the article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

[ABM18] S. Athreya, O. Butkovsky, and L. Mytnik. Strong existence and uniqueness for stable stochastic differentialequations with distributional drift. *arXiv:1801.03473*, 2018.

[AH96] D. R. Adams and L. I. Hedberg. *Function spaces and Potential Theory*. Springer, 1996.

[Bas11] R. F. Bass. *Stochastic processes*, volume 33. Cambridge series in statistical and probabilistic mathematics, 2011.

[BC01] R. F. Bass and Z. Q. Chen. Stochastic differential equations for Dirichlet processes. *Probability Theory and Related Fields*, 121(3):422–446, November 2001.

[BC03] R. F. Bass and Z.-Q. Chen. Brownian motion with singular drift. *Ann. Probab.*, 31–2:791–817, 2003.

[BP15] V. I. Bogachev and A. Y. Pilipenko. Strong solutions to stochastic equations with lévy noise and a discontinuous drift coefficient. *Doklady Mathematics*, 92:471–475, 2015.

[CC18] G. Cannizzaro and K. Chouk. Multidimensional sdes with singular drift and universal construction of the polymer measure with white noise potential. *Ann. Probab.*, 46(3):1710–1763, 05 2018.

[CdRHM18a] P-E. Chaudru de Raynal, I. Honoré, and S. Menozzi. Sharp Schauder Estimates for some Degenerate Kolmogorov Equations. *arXiv:1810.12227*, 2018.

[CdRHM18b] P.-E. Chaudru de Raynal, I. Honoré, and S. Menozzi. Strong regularization by Brownian noise propagating through a weak Hörmander structure. *arXiv:1810.12225*, 2018.

[CdRMP19] P. E. Chaudru de Raynal, S. Menozzi, and E. Priola. Schauder estimates for drifted fractional operators in the supercritical case. *arXiv:1902.02616*, 2019.

[CG16] R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of ODEs. *Stochastic Processes and their Applications*, 126(8):2323–2366, August 2016.

[CZZ17] Z.-Q. Chen, X. Zhang, and G. Zhao. Well-posedness of supercritical SDE driven by L\'evy processes with irregular drifts. *arXiv:1709.04632*, Sep 2017.

[DD16] F. Delarue and R. Diel. Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probab. Theory Related Fields*, 165(1-2):1–63, 2016.

[DF13] A. Debussche and N. Fournier. Existence of densities for stable-like driven sde's with hölder continuous coefficients. *Journal of Functional Analysis*, 264(8):1757 – 1778, 2013.

[DGI19] T. De Angelis, M. Germain, and E. Issoglio. A numerical scheme for stochastic differential equations with distributional drift. *arXiv e-prints*, page arXiv:1906.11026, Jun 2019.

[FIR17] F. Flandoli, E. Issoglio, and F. Russo. Multidimensional stochastic differential equations with distributional drift. *Transactions of the American Mathematical Society*, 369(10.1090/tran/6729):1665–1688, 2017.

[FOT10] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes*. De Gruyter, 2010.

[FRW03] F. Flandoli, F. Russo, and J. Wolf. Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.*, 40:493D542, 2003.

[FRW04] F. Flandoli, F. Russo, and J. Wolf. Some SDEs with distributional drift. II. Lyons-Zheng structure, itô's formula and semimartingale characterization. *Random Oper. Stochastic Equations*, 12-2:145D184, 2004.

[HM16] L. Huang and S. Menozzi. A Parametrix Approach for some Degenerate Stable Driven SDEs. *Annales Instit. H. Poincaré*, 52(4):1925–1975, 2016.

[HMP19] L. Huang, S. Menozzi, and E. Priola. L^p Estimates For Degenerate Non-Local Kolmogorov Operators. *Journal de Mathématiques Pures et Appliquées*, 121:162–215, 2019.

[Kol00] V. N. Kolokoltsov. Symmetric Stable Laws and Stable-Like Jump-Diffusions. *Proceedings of the London Mathematical Society*, 80(3):725–768, May 2000.

[Kol11] V. N. Kolokoltsov. *Marko-v process, semigroups and genertaors*, volume 38. De Gruyter, Studies in Mathematics, 2011.

[KR05] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probability Theory and Related Fields*, 131:154–196, 2005.

[LR02] P.-G. Lemarie-Rieusset. *Recent developments in the Navier-Stokes problem*. CRC Press, 2002.

[LZ19] C. Ling and G. Zhao. Non-local elliptic equation in hölder space and the martingale problem. *Arxiv:1907.00588v1*, 2019.

[MP14] R. Mikulevicius and H. Pragarauskas. On the cauchy problem for integro-differential operators in hölder classes and the uniqueness of the martingale problem. *Potential Anal.*, 40(4):539–563, 2014.

[Pri12] E. Priola. Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.*, 49– 2:421–447, 2012.

[Sat99] K. Sato. *Lévy processes and Infinitely divisible Distributions*. Cambridge University Press, 1999.

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.

[Szt10] P. Sztonyk. Estimates of tempered stable densities. *J. Theoret. Probab.*, 23(1):127–147, 2010.

[Tri83] H. Triebel. *Theory of function spaces, II*. Birkhauser, 1983.

[Wat07] T. Watanabe. Asymptotic estimates of multi-dimensional stable densities and their applications. *Transactions of the American Mathematical Society*, 359(6):2851–2879, 2007.

[Zvo74] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

[ZZ17] X. Zhang and G. Zhao. Heat kernel and ergodicity of SDEs with distributional drifts. *arXiv:1710.10537*, 2017.