Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 163, 1–52.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP864

On multidimensional stable-driven stochastic differential equations with Besov drift

Paul-Éric Chaudru de Raynal* Stéphane Menozzi†‡

Abstract

We establish well-posedness results for multidimensional non degenerate α-stable driven SDEs with time inhomogeneous singular drifts in $\mathbb{L}^r - \mathbb{B}^{-1+\gamma}_{p,q}$ with $\gamma < 1$ and α in $(1,2]$, where $\mathbb{L}^r$ and $\mathbb{B}^{-1+\gamma}_{p,q}$ 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. This allows us in turn to define an *ad hoc* notion of weak solution, for which well-posedness holds as well. 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.

Keywords: SDEs with singular drifts; Besov spaces; stable processes; dynamics.
MSC2020 subject classifications: 60H10; 35R11; 60H50; 35B65.
Submitted to EJP on October 6, 2020, final version accepted on October 14, 2022.

1 Introduction

1.1 Statement of the problem and related literature

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 $(1,2]$ (thus including Brownian noise).

The main point here comes from the fact that the drift F is only supposed to belong to the space $\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d,\mathbb{R}^d))$, where $\mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d,\mathbb{R}^d)$ denotes a Besov space. In a

*Nantes Université, Laboratoire de Mathématiques Jean Leray, LMJL, UMR CNRS 6629, F-44000 Nantes, France. E-mail: pe.deraynal@univ-nantes.fr

†Université d'Evry Val d'Essonne-Paris Saclay, Laboratoire de Mathématiques et Modélisation d'Evry (LaMME), UMR CNRS 8071, 23 Boulevard de France 91037 Evry France.

‡Laboratory of Stochastic Analysis, HSE, Prokrovsky Boulevard 11, Moscow, Russian Federation.
E-mail: stephane.menozzi@univ-evry.fr

nutshell, when $p = q = \infty$, for any non integer $\beta > 0$, Besov spaces coincide with Hölder spaces $\mathbb{B}^{\beta}_{\infty,\infty}(\mathbb{R}^d, \mathbb{R}^d) = \mathcal{C}^{\beta}(\mathbb{R}^d, \mathbb{R}^d)$; when $\beta < 0$, this somehow indicates that the Hölder modulus blows up at rate β. The parameters p and q are related to the integrability of such a modulus. We refer to Section 2.6.4 of [Tri83] and Section 3.1 below for rigorous definition.

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" or "*formal* SDE (1.1)". 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 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 time inhomogeneous 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, they investigated the smoothness properties of the underlying PDE (with the drift as source term and null terminal condition) which *formally* writes[1],

$$\begin{cases} \partial_t u + F \cdot Du + \frac{1}{2}\Delta u = -F, \text{ on } [0, T) \times \mathbb{R}^d, \\ u(T, \cdot) = 0. \end{cases} \tag{1.2}$$

We said *formal* because above, it is not clear that the product $F \cdot Du$ is meaningful, since

[1] The PDE investigated in [FIR17] slightly differs from the one introduced here (additional potential term therein). However, the resulting Zvonkin transforms are *somehow* equivalent. We have thus chosen to present their result according to the approach we adopted here for the sake of clarity.

product between two distributions can only be defined under suitable constraints.

In their work, they managed to prove that this product indeed makes sense and that this PDE admits a unique *mild* solution (see Definition 8 therein), i.e.

$$u(t,x) = \int_t^T ds P_{s-t}[\{F(s,\cdot) + F(s,\cdot)\cdot Du(s,\cdot)\}](x),\ x \in \mathbb{R}^d, \tag{1.3}$$

where $(P_t)_t$ denotes the usual heat semi-group, which belongs to a suitable function space, allowing them to define the Zvonkin transform $\Phi(t,x) = x + u(t,x)$. The authors then introduce the notion of *virtual solution*. Namely, this is a process X defined on a stochastic basis such that:

$$X_t = x + u(0,x) - u(t,X_t) + \int_0^1 (\nabla u(s,X_s) + I) d\mathcal{W}_s.$$

The main advantage of such a definition is that the *bad* drift does not explicitly appears. It actually would from a formal expansion of u solving (1.2) through Itô's formula. This correspondence also gives that, for reasonable smooth drifts, classical solutions also are virtual solutions. Existence and uniqueness for virtual solutions are then established from the fact that Φ can be shown, from the well-posedness of the mild solution (1.3), to be a C^1-diffeomorphism for T small enough.

The SDE

$$Y_t = \Phi(0,x) + \int_0^t D\Phi(s, \Phi^{-1}(s,Y_s)) d\mathcal{W}_s, \tag{1.4}$$

has indeed itself a unique weak solution from the smoothness of u solving (1.2) and $X_t = \Phi^{-1}(t,Y_t)$ is a virtual solution for which uniqueness in law holds. As a consequence of this approach only few things can be said about the original dynamics.

In this last perspective, we can also refer to the work of Zhang and Zhao [ZZ17], who established in the time homogeneous case the well-posedness of the martingale problem for the generator associated with (1.1), which can in their framework additionally contain a non trivial smooth enough diffusion coefficient (see also Remarks 6 and 15 below). Therein, they obtained as well as some Krylov type density estimates in Bessel potential spaces for the solution. Also, they managed to derive a more precise description of the limit drift in the *formal* dynamics in (1.1), which is interpreted as a suitable limit of a sequence of mollified drifts, i.e. $\lim_n \int_0^t F_m(s,X_s)ds$ for a sequence of smooth functions $(F_m)_{m\geq 1}$ converging to F in a suitable sense.

The key point in these works, who heavily rely on PDE arguments, is to establish: (i) that the product $F \cdot Du$ in (1.2) is meaningful as a distribution (thus with the same regularity as F); (ii) that the semi-group $(P_t)_t$ associated with the noise maps the quantity in the bracket in the right hand side of (1.3) onto a suitable function space, with time integrable singularity, say at least $\mathcal{C}^{1^+}(\mathbb{R}^d,\mathbb{R}^d)$ to define Du properly. Let us illustrate how such constraints translate, assuming for a while that F is time homogeneous and belongs to the Besov-Hölder space $\mathbb{B}^{-1+\gamma}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d) = \mathcal{C}^{-1+\gamma}(\mathbb{R}^d,\mathbb{R}^d)$. From the smoothing effect of the noise (parabolic bootstrap), one expects that the semi-group maps $\mathcal{C}^{-1+\gamma}(\mathbb{R}^d,\mathbb{R}^d)$ onto $\mathcal{C}^{-1+\gamma+2}(\mathbb{R}^d,\mathbb{R}^d) = \mathcal{C}^{1+\gamma}(\mathbb{R}^d,\mathbb{R}^d)$ (Schauder estimates). The second constraint (ii) then gives $-1+\gamma+2 > 1 \Leftrightarrow \gamma > 0$. On the other hand, as Du belongs to $\mathcal{C}^{\gamma}(\mathbb{R}^d,\mathbb{R}^d)$, the first constraint (i) gives, from Bony's paraproduct rule, that the sum of the regularity indexes of F and Du must be strictly positive: $-1+\gamma+\gamma > 0 \Leftrightarrow \gamma > 1/2$.

This is indeed the threshold appearing in [FIR17] and [ZZ17] as well as the one previously obtained in [BC01]. This is precisely the threshold that will guarantee well-posedness of the corresponding Martingale problem and weak well-posedness of the associated dynamics, with the Definitions of Section 1.3 below (see Definitions 2 and 4), for a drift $F \in \mathbb{L}^\infty([0,T], \mathbb{B}^{-1+\gamma}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d))$ in the present work.

To bypass such a limit (*i.e.* $\gamma > 1/2$), one therefore has to use a suitable theory in order to first 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 $\mathbb{L}^\infty([0,T], \mathbb{B}_{\infty,\infty}^{(-2/3)^+}(\mathbb{R},\mathbb{R}))$. 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 non-linear Young integral involving the 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 most 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. We also refer to Kremp and Perkowski [KP22] for extensions of the approach of [DD16] to the stable case.

The pure jump case: $\alpha < 2$. In the pure jump case, there are a few works concerning the well-posedness of (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 any cases, the threshold obtained does not allow the Authors to consider singular/distributional drift, so that these results are not comparable with the present work.

On the other hand, 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 [ABM20] for $\alpha > 1$ and a time homogeneous $F \in \mathbb{B}_{\infty,\infty}^{-1+\gamma}(\mathbb{R},\mathbb{R})$ under the condition: $-1+\gamma > (1-\alpha)/2 \Leftrightarrow \gamma > (3-\alpha)/2$. 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 drifts belonging to a certain Kato class in the multidimensional setting.

Again, the result obtained by Athreya, Butkovsky 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 *formally* writes

$$\begin{cases} \partial_t u + F \cdot Du + L^\alpha u = -F, \text{ on } [0,T) \times \mathbb{R}, \\ u(T,\cdot) = 0, \end{cases} \tag{1.5}$$

where L^α is the generator of a non-degenerate α-stable process.

Considering again a time homogeneous F in $\mathcal{C}^{-1+\gamma}(\mathbb{R},\mathbb{R})$, one may reproduce the analysis done before in the Brownian setting to deduce that the solution u is expected to be in $\mathcal{C}^{-1+\gamma+\alpha}(\mathbb{R},\mathbb{R})$ (as the smoothing effect of the semi-group associated to the generator L^α is now of order α). The two constraints exposed in the Brownian setting ((i) to define the product $F \cdot Du$; (ii) to define the gradient Du) translate into $-1+\gamma+\alpha > 1 \Leftrightarrow \gamma > 2-\alpha$ for (ii) and $-1+\gamma-1+\gamma+\alpha-1 > 0 \Leftrightarrow \gamma > (3-\alpha)/2$ for (i). This is precisely the threshold that will guarantee weak well-posedness and pathwise uniqueness in the scalar case for a drift $F \in \mathbb{L}^\infty([0,T], \mathbb{B}^{-1+\gamma}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d))$ in the present work.

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 $\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d,\mathbb{R}^d)))$ when the noise driving the SDE is a 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 (1.5). Namely, we will provide a Schauder like 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 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 [CdRHM21], [CdRHM22]) or Schauder estimates for super-critical fractional operators [CdRMP20].

More precisely, our first main result consists in deriving the well-posedness of the martingale problem introduced in Definition 2 under suitable conditions on the parameters p, q, r and γ, see Theorem 3.

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 4, 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.

Inspired by the dynamics exhibited for the Martingale solution, we define next an ad hoc notion of weak solution in Definition 4 and prove the associated well-posedness result, see Theorem 6. Therein, we also manage to derive pathwise uniqueness in the scalar case, extending partially the previous results of [ABM20].

Eventually, we manage in Proposition 7 to collect all the results we have at hand to specify the dynamics of both the Martingale and weak solution. These results could be useful to investigate the numerical approximations of those singular SDEs (see equations (1.16) and (1.17)) and the recent work by De Angelis *et al.* [DGI19]

Let us conclude by mentioning that, while finishing the preparation of the present manuscript, we discovered a related work by Ling and Zhao [LZ22] 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 (1.5) with (homogeneous) drift belonging to Hölder-Besov spaces with negative regularity index (i.e. in $\mathbb{B}^{-1+\gamma}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d)$) 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 $F \in \mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d, \mathbb{R}^d))$). 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 6 and 15 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 [LZ22] 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.

1.3 Overview of the paper

We state in this part our main assumptions and results, together with rigorous definition about the meaning of solution of the *formal* SDE (1.1).

Framework. As already said, we consider the problem of the solvability, in a sense to be specified later on, of the *formal* SDE (1.1) with drift $F \in \mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d, \mathbb{R}^d))$, with $p, q, r \geq 1$ and $\gamma \in (1/2, 1)$. Concerning the driving noise in (1.1), we will denote by L^α its generator. 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 $\varphi \in C_0^\infty(\mathbb{R}^d, \mathbb{R})$:

$$L^\alpha \varphi(x) = \text{p.v.} \int_{\mathbb{R}^d} \big[\varphi(x+z) - \varphi(x)\big] \nu(dz), \tag{1.6}$$

where, writing in polar coordinates $z = \rho\xi$, $(\rho, \xi) \in \mathbb{R}_+ \times \mathbb{S}^{d-1}$, the Lévy measure decomposes as $\nu(dz) = \mu(d\xi) d\rho/\rho^{1+\alpha}$ 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 $\lambda \in \mathbb{R}^d$:

$$\kappa^{-1}|\lambda|^\alpha \leq \int_{\mathbb{S}^{d-1}} |\lambda \cdot \xi|^\alpha \mu(d\xi) \leq \kappa |\lambda|^\alpha. \tag{1.7}$$

Observe in particular that a rather large class of spherical measures μ satisfy (1.7). 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 $t > 0$:

$$P^\alpha_t[\varphi](x) := \int_{\mathbb{R}^d} dy p_\alpha(t, y-x) \varphi(y), \tag{1.8}$$

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.

Notion of solution of the *formal* SDE (1.1). In order to foster an appropriate notion of solution for the *formal* SDE (1.1), let us start with an informal discussion which can be found as well in the work of Cannizzaro and Choukh [CC18]. We reproduce it for the sake of clarity. Let $\Omega_2 = \mathcal{C}([0,T],\mathbb{R}^d)$ and $\Omega_\alpha = \mathcal{D}([0,T],\mathbb{R}^d)$ (space of càdlàg functions) when $0<\alpha<2$. The Stroock and Varadhan formulation of the Martingale Problem associated with an operator $F\cdot D + L^\alpha$ reads as: find a probability measure $\mathbb{P}^\alpha$ on the space Ω_α equipped with the canonical filtration so that

(a) $\mathbb{P}^\alpha(X_0 = x) = 1$,
(b) For all ϕ in $\mathcal{E}$ (where $\mathcal{E}$ here stands for a rich enough function space):

$$\left(\phi(t,X_t) - \int_0^t (\partial_s + F\cdot D + L^\alpha)\phi(s,X_s)ds\right)_{0\le t\le T}, \tag{1.9}$$

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

Above, the class $\mathcal{E}$ has to be chosen so that it is sufficiently rich to characterize a Markov process through the martingale formulation, see e.g. [EK86]. In our current setting, the main issue comes from the fact that the operator involves a distributional part (the drift term F) so that even if the products in $(F\cdot D + L^\alpha)\phi$ are well defined, this term could only be a distribution with same regularity as F. To avoid such a consideration the idea consists in taking $\mathcal{E}$ as the set of maps for which $(\partial_t + F\cdot D + L^\alpha)\phi$ is indeed a function f, *i.e.* $\mathcal{E}$ should be the set of solutions of the Cauchy Problem $\mathscr{C}(F,L^\alpha,f,g,T)$:

$$(\partial_t + F\cdot D + L^\alpha)u = f, \quad u_T = g, \tag{1.10}$$

for f and g in large enough classes $\mathcal{F}$ and $\mathcal{G}$. Having this in mind, one may thus rewrite the associated Martingale Problem as: find a probability measure $\mathbb{P}^\alpha$ on the space Ω_α equipped with its canonical filtration so that

(a) $\mathbb{P}^\alpha(X_0 = x) = 1$,
(b) For all f,g in $\mathcal{F},\mathcal{G}$,

$$\left(u(t,X_t) - \int_0^t f(s,X_s)ds - u(0,x)\right)_{0\le t\le T}, \tag{1.11}$$

with u the solution (in a sense to be specified) of the Cauchy Problem $\mathscr{C}(F,L^\alpha,f,g,T)$, is a (square integrable if $\alpha=2$) martingale under $\mathbb{P}^\alpha$.

Remark 1 (On the link between both Martingale formulation (1.9) and (1.11) in a favorable setting)**.** Let us try to briefly illustrate the link between both formulations in a favorable setting. Assume first that F is bounded, continuous in time and is γ-Hölder continuous in space. A natural choice for $\mathcal{E}$ is $C_b^{1,\alpha+\gamma}([0,T]\times\mathbb{R}^d,\mathbb{R}^d)$ so that $(\partial_t + F\cdot D + L^\alpha)\phi$ belongs to $C_b^{0,\gamma}([0,T]\times\mathbb{R}^d,\mathbb{R}^d)$. This thus means that $\mathcal{F}$ should be this latter function space while $\mathcal{G}$ should be the space $C_b^{\alpha+\gamma}(\mathbb{R}^d)$. As a consequence, both formulations are equivalents. Indeed, under these conditions, for any $f,g\in\mathcal{F},\mathcal{G}$, Schauder estimates imply that there exists a unique element $u\in C_b^{1,\alpha+\gamma}([0,T]\times\mathbb{R}^d,\mathbb{R}^d)$ that solves $\mathscr{C}(F,L^\alpha,f,g,T)$ in a classical sense. Therefore (1.9) implies (1.11) and, conversely, (1.11) implies (1.9) choosing for any $\phi\in C_b^{1,\alpha+\gamma}([0,T]\times\mathbb{R}^d,\mathbb{R}^d)$: $f = (\partial_t + F\cdot D + L^\alpha)\phi$ and $g = \phi(T,\cdot)$.

Definition of the Martingale solutions and associated well-posedness results. Let us come back to our general case, $F\in\mathbb{L}^r([0,T],\mathbb{B}_{p,q}^{-1+\gamma}(\mathbb{R}^d,\mathbb{R}^d))$. The main point is now to notice that, in order to obtain the martingale formulation (1.11), we do not need to work with classical solutions. Up to a regularization argument and thanks to Itô's formula, it is indeed enough to work with *mild* solutions u in $C^{0,1}([0,T]\times\mathbb{R}^d,\mathbb{R})$ with bounded gradient. In our setting, the natural candidate for the *mild* solution *formally* writes

$$\forall (t,x)\in[0,T]\times\mathbb{R}^d,\quad u(t,x)=P^\alpha_{T-t}[g](x)-\int_t^T ds P^\alpha_{s-t}[\{f-F\cdot Du\}](s,x),$$

for e.g. continuously differentiable function g and continuous in time and bounded in space function f. This form requires the last term on the right hand side to be defined as a function, whereas the drift F is only assumed to be a distribution. We are thus again led to find appropriate conditions that guarantee that: (i) $F\cdot Du$ is well defined; (ii) the product is mapped onto the space of continuous in time and continuously differentiable in space functions with bounded gradient by the semi-group $(P^\alpha_t)_t$.

This suggests that the gradient Du should belong to a suitable Lebesgue-Besov space that depends both on the driving noise and the drift. The following lemma, whose proof can be found in Subsection 3.4, is useful to investigate wether the above integral makes sense as a function and also gives some hints concerning the function space in which the solution u should be sought:

Lemma 1. Let $p,q,r\ge 1$, $\alpha\in(1,2]$ and $\gamma\in(1/2,1)$ be such that they satisfy a *good relation, i.e.*

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

Define then

$$\theta:=\gamma-1+\alpha-\frac{d}{p}-\frac{\alpha}{r}. \tag{1.12}$$

Let G in $\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d,\mathbb{R}^d))$ and v belongs to $\mathbb{L}^\infty([0,T],\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d))$ for any $0<\varepsilon\ll 1$. Then, the map

$$\mathfrak{r}^v:[0,T]\times\mathbb{R}^d\ni(t,x)\mapsto\int_t^T ds P^\alpha_{s-t}[G\cdot v](s,x),$$

with P^α the semi-group generated by L^α, belongs to $C^{0,1}_b([0,T]\times\mathbb{R}^d,\mathbb{R})$. Moreover, the product $G\cdot v$ makes sense as an element of $\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d,\mathbb{R}))$.

We point out that the parameter θ would correspond to the parabolic bootstrap, which thus also reflects the impact of the integrability indexes p,r (but not q).

We are now in position to define the *mild* solution of the *formal* Cauchy problem and then state the associated well-posedness theorem.

Definition 1. Let $\alpha\in(1,2]$, $f:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}$ and $g:\mathbb{R}^d\to\mathbb{R}$. For any given fixed $T>0$, we say that $u:[0,T]\times\mathbb{R}^d\to\mathbb{R}$ is a mild solution of the *formal* Cauchy problem $\mathscr{C}(F,L^\alpha,f,g,T)$

$$(\partial_t+F\cdot D+L^\alpha)u(t,x)=f(t,x)\text{ on }[0,T)\times\mathbb{R}^d,\quad u(T,\cdot)=g(\cdot)\text{ on }\mathbb{R}^d, \tag{1.13}$$

if it belongs to $\mathcal{C}^{0,1}([0,T]\times\mathbb{R}^d,\mathbb{R})$ with Du in $\mathcal{C}^0_b([0,T],\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d))$ for any $0<\varepsilon\ll 1$ where θ is given by (1.12) and satisfies

$$\forall (t,x)\in[0,T]\times\mathbb{R}^d,\quad u(t,x)=P^\alpha_{T-t}[g](x)-\int_t^T ds P^\alpha_{s-t}[\{f-F\cdot Du\}](s,x), \tag{1.14}$$

with $(P^\alpha_t)_t$ the semi-group generated by $(L^\alpha_t)_t$.

Notice that, due to definition of the space in which the solution should be sought, Lemma 1 implies that the above mild formulation is meaningful. We now have the following well-posedness result proved in Section 3 (see in particular Subsection 3.4).

Theorem 2. Let $p,q,r \geq 1$, $\alpha \in (1,2]$ and $\gamma \in (1/2,1)$ satisfy a *good relation* (GR). For all f in $\mathcal{C}([0,T], \mathbb{B}^{\theta-\alpha}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}))$ and $g \in \mathcal{C}^1(\mathbb{R}^d,\mathbb{R})$ with $Dg \in \mathbb{B}^{\theta-1}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d)$, where θ is given by (1.12), the *formal* Cauchy problem $\mathscr{C}(F, L^\alpha, f, g, T)$ admits a unique solution in the sense of Definition 1. Moreover it satisfies that for all $(t \leq s)$ in $[0,T]^2$, x in $\mathbb{R}^d$:

$$\begin{aligned} |u(t,x) - u(s,x)| &\leq C|t-s|^{\frac{\theta}{\alpha}}, \\ |Du(t,x) - Du(s,x)| &\leq C|t-s|^{\frac{\theta-1}{\alpha}}. \end{aligned}$$

We can therefore define the associated Martingale Problem and the corresponding well-posedness result proved in Section 2 (see subsection 2.2, points (i) to (iii)).

Definition 2. Let $\Omega_2 = \mathcal{C}([0,T],\mathbb{R}^d)$ and $\Omega_\alpha = \mathcal{D}([0,T],\mathbb{R}^d)$ when $0 < \alpha < 2$. For any $\alpha \in (0,2]$, we say that a probability measure $\mathbb{P}^\alpha$ on Ω_α equipped with its canonical filtration is a solution of the Martingale Problem associated with (F, L^α, x) for $x \in \mathbb{R}^d$ if

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

(ii) $\forall f \in \mathcal{C}([0,T], \mathcal{S}(\mathbb{R}^d,\mathbb{R}))$, $g \in \mathcal{C}^1(\mathbb{R}^d,\mathbb{R})$ with $Dg \in \mathbb{B}^{\theta-1}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d)$,

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

is a (square integrable if $\alpha = 2$) martingale under $\mathbb{P}^\alpha$ where u is the mild solution of the Cauchy Problem $\mathscr{C}(F, L^\alpha, f, g, T)$ in the sense of Definition 1.

Theorem 3. Let $p,q,r \geq 1$, $\alpha \in (1,2]$ and $\gamma \in (1/2,1)$ satisfy a *good relation* (GR). Then, the Martingale Problem associated with (F, L^α, x) for $x \in \mathbb{R}^d$, is well-posed in the sense of Definition 2. Moreover, the canonical process under $\mathbb{P}^\alpha$ is strong Markov.

Note that, according to the notations of the previous paragraph, we chose above to work with $\mathcal{E} = \mathcal{S}(\mathbb{R}^d)$, where $\mathcal{S}$ stands for the class of Schwartz functions. This is mainly motivated by our approach based on Besov spaces as this class is continuously embedded into any Besov spaces $\mathbb{B}^s_{l,m}(\mathbb{R}^d,\mathbb{R})$, $s \in \mathbb{R}$, $1 \leq l,m \leq \infty$, see e.g. paragraph 2.3.3 in [Tri83]. Note that it is as well dense in $\mathbb{B}^s_{l,m}(\mathbb{R}^d,\mathbb{R})$ provided $l,m < \infty$.

Building the dynamics. One may then wonder if something can be said about the dynamics of the underlying process. In other words, the next step consists in linking the Martingale Problem and the *formal* SDE (1.1). For sufficiently smooth drifts, the starting point to build a dynamics consists in recovering the noise from the canonical process associated with the martingale solution. Since in the current framework the difficulty consists in specifying the drift, it therefore seems natural to have a noise at hand to precisely recover the drift. This is precisely what leads to consider an *enlarged* martingale problem which *includes* the noise. Similar issues were e.g. discussed in Kurtz [Kur10].

For singular drifts this procedure was performed in the specific setting of [DD16]. We here extend this connection to our general setting and explain how one can relate the Martingale solution with the dynamics of (1.1), and how this latter has to be understood. To do so, we will however be led to slightly reinforce the *good relation* (GR). Namely, we say that $p,q,r \geq 1$, $\alpha \in (1,2]$ and $\gamma \in (1/2,1)$ satisfy a *good relation for the dynamics* if the following relation holds:

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

Note that when $p = r = \infty$, the above condition and the previous one are equivalent.

Having such a condition at hand, our main strategy consists in following the ideas of [DD16]. Namely, we show that the canonical process can be decomposed into two parts: a stable driving noise, *i.e.* an α-stable process with generator L^α and a drift term, defined as a kind of stochastic non-linear Young integral of a regularized version of the initial drift by the density of the driving process. To the best of our knowledge, non-linear Young integrals have been introduced in [CG16] and further extended from a probabilistic perspective in [DD16]. The rigorous definition of this last object is the following.

Definition 3. Let $\tau > 0$, $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{0\le t\le\tau}, \tilde{\mathbb{P}})$ be a filtered probability space and let $(\psi_t)_{0\le t\le\tau}$ be a progressively measurable process on it. Let $(A(s,t))_{0\le s\le t\le\tau}$ be a continuous and progressively measurable map in the sense that for any $0 \le s \le t$,

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

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

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

is continuous. For $\ell \ge 1$, we call $\mathbb{L}^\ell$-*stochastic non-linear Young integral* of ψ with respect to the pseudo increment A the limit in $\mathbb{L}^\ell(\tilde{\Omega}, \tilde{\mathbb{P}})$

$$\lim_{\substack{\Delta \text{ partition of } [0,\tau] \\ |\Delta|\to 0}} \sum_{t_i\in\Delta} \psi_{t_i} A(t_i, t_{i+1}) =: \int_0^\tau \psi_t A(t, t+dt), \tag{1.15}$$

when it exists.

Having such tools at hand, the strategy consists in building simultaneously the Martingale solution and the noise $(X, \mathcal{W})$ as the solution of a kind of enlarged Martingale Problem whose entries for the underlying canonical process $(X, \mathcal{W})$ are, respectively, associated with the solution of the martingale problem $(L^\alpha + F \cdot D)$ and the corresponding driving noise. This will allow to build the drift as the difference between them. Indeed, having such a canonical process at hand, we decompose the increment of the process X as

$$X_{t+h} - X_t = \{\mathbb{E}[X_{t+h} - X_t|\mathcal{F}_t]\} + \{X_{t+h} - X_t - \mathbb{E}[X_{t+h} - X_t|\mathcal{F}_t]\},$$

where $\mathcal{F}_t := \sigma(X_s, \mathcal{W}_s,\ 0 \le s \le t)$. Clearly, the first difference in the above right hand side stands for a drift term, while the second stands for a martingale part. It thus suffices to relate both parts with (i) the original drift F and (ii) the α-stable noise $\mathcal{W}$ previously built. This is the purpose of the next result proved in Section 4 (see in particular Subsection 4.3 and Proposition 13 below).

Theorem 4. Let $p, q, r \ge 1$, $\alpha \in (1,2]$ and $\gamma \in (1/2, 1)$ satisfy a *good relation for the dynamics* (GR-D). It then holds that there exists a probability measure $\mathbf{P}^\alpha$ on $\mathcal{C}([0,T], \mathbb{R}^{2d})$ when $\alpha = 2$ and $\mathcal{D}([0,T], \mathbb{R}^{2d})$ when $0 < \alpha < 2$ such that the canonical process, denoted by $(X, \mathcal{W})$, satisfies

(i) The law of X under $\mathbf{P}^\alpha$ is a solution of the Martingale problem associated with (F, L^α, x), $x \in \mathbb{R}^d$ and the law of $\mathcal{W}$ under $\mathbf{P}^\alpha$ is a Brownian motion if $\alpha = 2$ and an α-stable process with generator L^α if $\alpha < 2$.

(ii) The dynamics of the canonical process reads

$$X_t = x + \int_0^t \mathscr{F}(s, X_s, ds) + \mathcal{W}_t, \quad \mathbf{P}^\alpha - a.s. \tag{1.16}$$

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

$$\mathscr{F}(v,x,s-v) := \int_v^s dr \int_{\mathbb{R}^d} dy F(r,y) p_\alpha(r-v,y-x), \tag{1.17}$$

with p_α the (smooth) density of $\mathcal{W}$ and where the integral in (1.16) is understood as an $\mathbb{L}^\ell$-*stochastic non-linear Young integral*, for any $1 \le \ell < \alpha$, in the sense of Definition 3.

In fact, we prove a stronger result concerning the dynamics. We show that it is possible to define a stochastic non-linear Young integral against the dynamics, leading in turn to use Itô calculus. This is done for a suitable class of progressively measurable processes ψ satisfying appropriate Hölder regularity conditions. For any $\mathfrak{q}' \ge 1$, any $\beta \in (0,1)$, we set

$$\mathcal{H}^\beta_{\mathfrak{q}'}(\Omega,\mathcal{F},(\mathcal{F}_t)_t,\mathbb{P}) := \Big\{ (\psi_t)_{t\in[0,T]} \text{ progressively measurable}, \tag{1.18}$$
$$\sup_{t\in[0,T]} \mathbb{E}^{\frac{1}{\mathfrak{q}'}}[|\psi_t|^{\mathfrak{q}'}] + \sup_{s\neq t\in[0,T]} \frac{\mathbb{E}^{\frac{1}{\mathfrak{q}'}}[|\psi_s-\psi_t|^{\mathfrak{q}'}]}{|t-s|^\beta} < +\infty \Big\}.$$

As a corollary of the proof of Theorem 4, we have the following result.

Corollary 5 (Associated $\mathbb{L}^\ell$-stochastic non-linear Young integral, $1 \le \ell < \alpha$)**.** Under the above assumptions, one can define a stochastic non-linear Young integral w.r.t. the quantities in (1.16). Namely, for any $1 \le \ell, \mathfrak{q} < \alpha$, for which there exists $\mathfrak{q}' \ge 1$ satisfying $1/\mathfrak{q}' + 1/\mathfrak{q} = 1/\ell$, one has

$$\int_0^t \psi_s dX_s = \int_0^t \psi_s \mathscr{F}(s,X_s,ds) + \int_0^t \psi_s d\mathcal{W}_s, \tag{1.19}$$

for any $\psi \in \mathcal{H}^{1-1/\alpha-\varepsilon_2}_{\mathfrak{q}'}$, for all $0 < \varepsilon_2 < (\theta-1)/\alpha$ and where the first term in the above right hand side is defined as an $\mathbb{L}^\ell$-stochastic non-linear Young integral.

Further properties and weak formulation. The previously described dynamics for the Martingale solution strongly suggests that a notion of weak solution associated with the *formal* SDE (1.1) can somehow be considered. This leads to the following definition.

Definition 4. We call weak solution of the *formal* SDE (1.1) a pair $(Y,\mathcal{Z})$ of adapted processes on a filtered probability space $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\ge0},\mathbb{P})$ such that $\mathcal{Z}$ is an $\{\mathcal{F}_t\}_{t\ge0}$ α-stable process and $(Y,\mathcal{Z})$ satisfies

$$Y_t = x + \int_0^t \mathscr{F}(s,Y_s,ds) + \mathcal{Z}_t, \quad \mathbb{P}-a.s., \qquad \mathbb{E}|\int_0^t \mathscr{F}(s,Y_s,ds)| < +\infty \tag{1.20}$$

for any t in $[0,T]$ and where for any $0 \le v \le s \le T$, $x \in \mathbb{R}^d$,

$$\mathscr{F}(v,x,s-v) = \int_v^s dr \int_{\mathbb{R}^d} dy F(r,y) p_\alpha(r-v,y-x) \tag{1.21}$$

with p_α the (smooth) density of $\mathcal{Z}$ and where the integral in (1.20) is understood as an $\mathbb{L}^1$-*stochastic non-linear Young integral*, in the sense of Definition 3.

We say that weak uniqueness holds for (1.1) if for any two weak solutions $(Y,\mathcal{Z})$, $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\ge0},\mathbb{P})$ and $(\tilde{Y},\tilde{\mathcal{Z}})$, $(\tilde{\Omega},\tilde{\mathcal{F}},\{\tilde{\mathcal{F}}_t\}_{t\ge0},\tilde{\mathbb{P}})$ with the same initial condition, then $(Y_t)_{t\ge0} \overset{(\text{law})}{=} (\tilde{Y}_t)_{t\ge0}$.

We then have the following well-posedness result whose proof is postponed to Section 5 (see Subsection 5.1) and thoroughly exploits that we can apply some stochastic calculus arguments to the dynamics and the smoothness of the underlying PDE (see Corollary 5 and Theorem 2 above).

Theorem 6. Let $p, q, r \geq 1$, $\alpha \in (1, 2]$ and $\gamma \in (1/2, 1)$ satisfy a *good relation for the dynamics* (GR-D). Then,

(i) the *formal* SDE (1.1) admits a unique weak solution in the sense of Definition 4;

(ii) if $d = 1$, pathwise uniqueness holds, *i.e.* the paths of two weak solutions defined on the same probability basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}, \mathcal{Z})$ coincide a.s. whenever they start from the same initial condition.

Moreover, one can define an associated $\mathbb{L}^1$-stochastic non-linear Young calculus *i.e.* Corollary 5 hold with $\ell = 1$ therein.

Remark 2 (About the connections between the Martingale and weak solutions)**.**

- Although we obtain well-posedness for both the Martingale problem and the weak formulation associated with the *formal* SDE (1.1), we do not claim that equivalence holds between them. In fact, what we are able to prove is that weak existence implies the existence of a Martingale solution and that uniqueness for the Martingale problem implies weak uniqueness, see Remark 17. It is nevertheless more involved to prove that existence of a Martingale solution gives the weak existence and that weak uniqueness implies uniqueness for the Martingale problem. The main issue relies on the fact that the very definition of the Martingale problem does not allow one to define an associated stochastic calculus, while the very definition of a weak solution does. Without the help of Itô's formula, we are not able to go from the Martingale solution to the weak one, which prevents us from obtaining the equivalence between both formulations. We can also refer to [IR22] for related issues.
- Pay attention that, in the above result (ii), we do not claim that strong uniqueness holds. This mainly comes from a measurability argument. In [ABM20], 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 terms 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.* [ABM20] 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.

We emphasize that the above results give, to the best of our knowledge, the most accurate description of the dynamics of the *formal* SDE (1.1) we found in the literature on SDE with distributional drift. We emphasize as well that such a description includes all those we found and discussed in the Introduction. A crucial point is that, moreover, the above dynamics (for whenever the Martingale solution or the weak solution) coincide with the classical one when the drift is time-space Hölder continuous. All those facts are collected in the following proposition whose proof is given in Section 5 (see Subsection 5.2).

Proposition 7. Assume (GR-D) holds. Then, either the Martingale solution or the weak solution of the formal SDE (1.1) is:

(i) A virtual solution of the *formal* SDE (1.1);

(ii) A Dirichlet process;

(iii) It holds that for any smooth approximating sequence $(F_m)_{m\geq 1}$ such that

$$\lim_{m\to+\infty} \|F - F_m\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})} = 0^2,$$

$$\lim_{m\to\infty} \left\| \int_0^t \mathscr{F}(s, X_s, ds) - \int_0^t F_m(s, X_s) ds \right\|_{\mathbb{L}^\ell} = 0, \quad 1 \leq \ell < \alpha, \tag{1.22}$$

with $\ell = 1$ for the weak solution;

(iv) If F is time-space β-Hölder continuous, the previous construction coincide with the *"usual"* drift:

$$\int_0^t \mathscr{F}(s, X_s, ds) = \int_0^t F(s, X_s) ds, \quad a.s..$$

We eventually mention that the previous explicit representation and properties of the drift could also be useful in order to derive numerical approximations for the SDE (1.16). 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.

Organization of this paper. The paper is organized as follows. Section 2 is dedicated to the proof of Theorem 3 (well-posedness of the Martingale problem). To do so, we collect some material (i.e. intermediate results) on the PDE, which is used all along this work. These intermediate results are proved in Section 3 and allow, in turn, to prove Theorem 2. As mentioned before, Section 3 is dedicated to the PDE analysis. In Section 4, we mainly reconstruct the dynamics associated with the Martingale solution. Proof of Theorem 4 can be found therein. Having this dynamics at hand, we then investigate the weak solution of the *formal* SDE (1.1) and further properties of the obtained drift in Section 5. This last section thus contains the proofs of Theorem 6 and Proposition 7.

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. We also sometimes shorten $\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^N, \mathbb{R}^M))$, for N, M in $\mathbb{N}$, with the notation $\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q})$ or $\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})$ when there are no ambiguities.

2 Well-posedness of the Martingale problem: proof of Theorem 3

In this Section, we mainly prove Theorem 3. To do so, we need some additional material on the PDE, which is collected in Subsection 2.1 below. This material is of crucial importance in our work as it will be used to prove Theorems 2, 4, 6 and Proposition 7 as well. All these PDE results are proved in Section 3 below (see Subsection 3.4). The proof of Theorem 3 is then derived in Subsection 2.2.

[2] See Remark 3 for the case when p and/or r are/is $+\infty$.

2.1 The underlying PDE

As underlined in Definitions 1 and 2, it turns out that the well-posedness of the Martingale Problem associated with (F, L^α, x), $x \in \mathbb{R}^d$, heavily relies on the construction of a suitable theory for the Cauchy problem $\mathscr{C}(F, L^\alpha, f, g, T)$ (see Definition 1) for some data f and g belonging to some appropriate function spaces to be specified later on. We recall that, because of the scalar product $F \cdot Du$ therein, the aforementioned PDE is only stated formally. Only the mild formulation, given by (1.14), is licit thanks to Lemma 1.

Hence, as a key intermediate tool we need to introduce what we will later on call the *mollified* Cauchy problem. Namely, denoting by $(F_m)_{m\in\mathbb{N}^*}$ a sequence of smooth functions such that $\|F - F_m\|_{\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q})} \to 0$ when $m \to \infty$, the *mollified* Cauchy problem $\mathscr{C}(F_m, L^\alpha, f, g, T)$ reads as

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

Remark 3 (Smooth approximating sequence of the drift)**.** When $p, q, r < \infty$, such a sequence can be obtained from $C_0^\infty([0,T]\times\mathbb{R}^d,\mathbb{R}^d)$ functions (or in the Schwartz class in space), see e.g. Theorem 4.1.3 in [AH96]. When $p = \infty$, one can approximate F by a sequence of $C_b^\infty([0,T]\times\mathbb{R}^d,\mathbb{R}^d)$ functions in $\mathbb{L}^{r'}([0,T],\mathbb{B}^{-1+\gamma'}_{p,q})$ for any $\gamma' < \gamma$, $r' = r$ if $r < +\infty$ and $r' < +\infty$ if $r = \infty$ (taking e.g. the time-space convolution with a gaussian kernel with variance $m^{-1}\mathrm{Id}$), see e.g. Definition 2.2 and Lemma 1.3 in [ABLM21] again for the spatial part. Up to an abuse of notation, one can denote by γ, r these indexes and still use the analysis done below.

We start with the following control which, in some sense, is the counterpart of Theorem 2 for the *mollified* Cauchy problem.

Proposition 8. Assume that the parameters p, q, r, α and γ satisfy a *good relation* (GR). Let f, g be smooth functions where g has as well at most linear growth. Let $(u_m)_{m\geq1}$ denote the sequence of classical solutions of the *mollified* PDE (2.1) i.e. of $\big(\mathscr{C}(F_m, L^\alpha, f, g, T)\big)_{m\geq1}$. Then, there exist positive constants $C := C(\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})})$, $C_T := C(T, \|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})})$, depending on the known parameters γ, p, q, r and κ in **(UE)**, s.t. for all $m \geq 1$:

$$
\begin{aligned}
\forall x\in\mathbb{R}^d,\quad |u_m(t,x)| &\leq C(1+|x|),\\
\|Du_m\|_{\mathbb{L}^\infty\left(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}\right)} &\leq C_T(\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}} + \|f\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-\alpha}_{\infty,\infty})}),
\end{aligned} \tag{2.2}
$$

$$
\begin{aligned}
\forall 0\leq t\leq s\leq T,\ x\in\mathbb{R}^d,\ |u_m(t,x)-u_m(s,x)| &\leq C|t-s|^{\frac{\theta}{\alpha}},\\
|Du_m(t,x)-Du_m(s,x)| &\leq C|t-s|^{\frac{\theta-1}{\alpha}},
\end{aligned}
$$

where $\varepsilon \ll 1$ can be chosen as small as desired, $T \mapsto C_T$ is a non-decreasing function and where, from the definition in (1.12), $\theta - 1 = \gamma - 2 + \alpha - d/p - \alpha/r > 0$.

Moreover, the sequence $(u_m, Du_m)_{m\geq1}$ converges toward the solution (u, D_u), in the sense of Definition 1, of the Cauchy problem $\mathscr{C}(F, L^\alpha, f, g, T)$ uniformly on compact subsets of $[0,T]\times\mathbb{R}^d$.

Remark 4. Let us mention that, when the terminal condition g is bounded, then the solution u_m is itself bounded, i.e. $\sup_{m\geq1}|u_m(t,x)| \leq C$.

Remark 5 (On the spatial smoothness of the mollified PDE)**.** From the conditions on γ, α and the definition of θ in (1.12), 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 (On the corresponding parabolic bootstrap)**.** Observe that, when $p = r = +\infty$, we almost have a Schauder type result, namely $\theta = \gamma - 1 + \alpha$ in (2.2) and we end up with the corresponding parabolic bootstrap effect for both the solution of $\mathscr{C}(F_m, L^\alpha, f, g, T)$ and $\mathscr{C}(F, L^\alpha, f, g, T)$, up to the small exponent ε which can be chosen arbitrarily small.

One may wonder why the control is given with $\varepsilon > 0$. This is one of the specificities of the distributional setting which makes makes it difficult to adapt the arguments of [CdRMP20] to control the remainder. In the quoted reference in the so-called diagonal regime we directly used the supremum of the gradient whereas the Hölder modulus is here needed to give a proper meaning to the distribution $F_m \cdot Du_m$ uniformly in m. We believe that, up to an additional scaling argument, similar to the one considered in [CdRHM21] the control should still hold with $\varepsilon = 0$. On the other hand the time regularity obtained corresponds to the expected “heuristic” parabolic bootstrap.

Remark 7 (About additional diffusion coefficients)**.** It should be noted at this point that we are confident about the extension of the results to differential operators L^α involving non-trivial diffusion coefficient, provided this latter is Hölder-continuous in space. Sketches of proofs in this direction are given in the Remark 15 in Subsection 3.4. However, we avoid investigating this point for the sake of clarity and in order to focus on the more (unusual) drift component.

The following Proposition and Corollary provide a Zvonkin type theory for the *mollified* and *formal* Cauchy problem, respectively $\mathscr{C}(F_m, L^\alpha, -F_m^k, 0, T)$ and $\mathscr{C}(F, L^\alpha, -F^k, 0, T)$, where for any k in $\{1, \ldots, d\}$, F^k denotes the $k^{\rm th}$ component of F and $(F_m^k)_{m\geq 1}$ denotes its mollification, see Remark 3.

Proposition 9 (Zvonkin type theory for the *mollified* PDE)**.** Let p, q, r, α and γ satisfy a *good relation* (GR) and let k in $\{1, \ldots, d\}$. 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 $m \geq 1$, the sequence of classical solutions $(u_m^k)_{m\geq 1}$ of $\big(\mathscr{C}(F_m, L^\alpha, -F_m^k, 0, T)\big)_{m\geq 1}$ satisfies:

$$\|u_m^k\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)} + \|Du_m^k\|_{\mathbb{L}^\infty\left(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}\right)} \leq C_T,$$

where $T \mapsto C_T$ is a non-decreasing function and for which the two last lines in (2.2) hold as well. Moreover, for any k in $\{1, \ldots, d\}$, the sequence $(u_m^k, Du_m^k)_{m\geq 1}$ converges towards the solution (u^k, Du^k),in the sense of Definition 1, of the Cauchy problem $\mathscr{C}(F, L^\alpha, -F^k, 0, T)$, uniformly on compact subsets of $[0, T] \times \mathbb{R}^d$.

Corollary 10 (Zvonkin type theory for the *formal* PDE)**.** Let $k \in \{1, \ldots, d\}$. Under the above assumptions, the *formal* Cauchy problem $\mathscr{C}(F, L^\alpha, -F^k, 0, T)$ admits a unique solution u^k in the sense of Definition 1 which moreover satisfies that there exists a positive constant $C_T := C(T, \|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})})$ s.t.

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

where $T \mapsto C_T$ is a non-decreasing function.

Remark 8. Of course, in order to use the Zvonkin type theory to derive strong well-posedness in the multidimensional setting some controls of the second order derivatives are needed. This is what Krylov and Röckner did in [KR05] in the Sobolev setting. Let us also specify that, in connection with Theorem 6, 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 $(X^1,\mathcal{W}),(X^2,\mathcal{W})$ of (1.16), a regularized version of $|X_t^1-X_t^2|$, 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 6-(ii) in Subsection 5.1-(ii) and e.g. Proposition 2.9 in [ABP18]), whereas in the multidimensional setting, for strong uniqueness, the second derivatives get in.

2.2 From PDE to SDE results: proof of Theorem 3

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]. This approach 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 3 provided Proposition 8, Proposition 9, Corollary 10 and Theorem 2 hold.

(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 (2.1). Note that in this regularized framework, for any m, the Martingale Problem associated with (F_m,L^α,x), $x\in\mathbb{R}^d$, is well posed. We denote by $\mathbb{P}_m^\alpha$ the associated solution. Let us generically denote by $(X_t^m)_{t\ge0}$ 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 $\alpha<2$. Recall that we denoted it by Ω_α.

Let $u_m=(u_m^1,\dots,u_m^d)$ where each u_m^k, $k\in\{1,\dots,d\}$ is the solution of the mollified Cauchy problem $\mathscr{C}(F_m,L^\alpha,-F_m^k,0,T)$. Let us define for any $s\ge v$ in $[0,T]^2$ and for any $\alpha\in(1,2]$ the process

$$M_{v,s}(\alpha,u_m,X^m)=\begin{cases}\displaystyle\int_v^s Du_m(r,X_r^m)\cdot dW_r,\\ \text{where } W \text{ is a Brownian motion, if } \alpha=2;\\ \displaystyle\int_v^s\int_{\mathbb{R}^d\setminus\{0\}}\{u_m(r,X_{r^-}^m+x)-u_m(r,X_{r^-}^m)\}\tilde N(dr,dx),\\ \text{where } \tilde N \text{ is the compensated Poisson measure,}\\ \text{if } \alpha<2.\end{cases}\tag{2.3}$$

Note that this process makes sense for any $m\ge1$, thanks to regularity estimates on u_m from Proposition 9. Next, applying Itô's formula to $(X_r^m+u_m(r,X_r^m))_{r\in[v,s]}$ we obtain

$$X_s^m-X_v^m=M_{v,s}(\alpha,u_m,X^m)+\mathcal{W}_s-\mathcal{W}_v+[u_m(v,X_v^m)-u_m(s,X_s^m)].\tag{2.4}$$

In order to prove that $(\mathbb{P}_m^\alpha)_{m\in\mathbb{N}^*}$ actually forms a tight sequence of probability measures on Ω_α, whose limit is denoted by $\mathbb{P}^\alpha$, it is sufficient to prove that there exists $c,\tilde p$ and $\eta>0$ such that $\mathbb{E}^{\mathbb{P}_m^2}[|X_s^m-X_v^m|^{\tilde p}]\le c|v-s|^{1+\eta}$ or $\mathbb{E}^{\mathbb{P}_m^\alpha}[|X_s^m-X_0^m|^{\tilde p}]\le cs^\eta$, for $\alpha\in(1,2)$ thanks to the Kolmogorov (resp. Aldous) Criterion. We refer e.g. for the latter to Proposition 34.9 in Bass [Bas11]. Writing

$$\begin{aligned}&[u_m(v,X_v^m)-u_m(s,X_s^m)]\\ =\ &u_m(v,X_v^m)-u_m(v,X_s^m)+u_m(v,X_s^m)-u_m(s,X_s^m),\end{aligned}$$

the result follows in small time, thanks to Proposition 9 (choosing $1<\tilde p<\alpha$ in the pure jump setting).

(ii) Identification of the limit probability measure. Let us now prove that the limit, denoted by $\mathbb{P}^\alpha$, is indeed a solution of the martingale problem associated with (F, L^α, x), $x \in \mathbb{R}^d$, in the sense of Definition 2. Let $f \in C^0([0,T], \mathcal{S}(\mathbb{R}^d))$ and g be a continuous function with gradient in $\mathbb{B}^{\theta-1}_{\infty,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ and let for any $m \geq 1$ u_m be the classical solution of the Cauchy problem $\mathscr{C}(F_m, L^\alpha, f, g, T)$. 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 (2.3). Thus, using the convergence of $(u_m, Du_m)_{m\geq 1}$ to the solution (u, Du) of $\mathscr{C}(F, L^\alpha, f, g, T)$ on every compact subsets of $[0,T] \times \mathbb{R}^d$, which follows from Proposition 8, together with a uniform control of the moment of X^m (which also derives from (2.4) and the above conditions on u_m), we deduce that

$$\left(u(t, X_t) - \int_0^t f(s, X_s) ds - u(0, x) \right)_{0 \leq t \leq T}, \tag{2.5}$$

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

(iii) Uniqueness of the limit probability measure. We now come back to the canonical space Ω_α, and let $\mathbb{P}^\alpha$ and $\tilde{\mathbb{P}}^\alpha$ be two solutions of the Martingale Problem associated with (F, L^α, x), $x \in \mathbb{R}^d$. Thus, for all $f \in C^0([0,T], \mathcal{S}(\mathbb{R}^d))$ we have, for the solution u of the Cauchy problem $\mathscr{C}(F, L^\alpha, f, 0, T)$

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

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]. The strong Markov property follows from Theorem 4.2 in [EK86].

3 PDE analysis

This part is dedicated to the proofs of Proposition 8, Proposition 9, Theorem 2 as well as Corollary 10 and Lemma 1. 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 Subsection 3.1. Then, we provide a primer on the *formal* Cauchy problem $\mathscr{C}(F, L^\alpha, f, g, T)$ by investigating the smoothing properties of the Green kernel associated with the stable noise in Subsection 3.2. Uniform (w.r.t. the mollification) estimates of the solution of the Cauchy problem $\mathscr{C}(F_m, L^\alpha, f, g, T)$ are investigated in Subsection 3.3. Eventually, we derive in Subsection 3.4 the proofs of Proposition 8, Proposition 9, Theorem 2 as well as Corollary 10 and Lemma 1. 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 8 and Theorem 2.

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

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

where $\big(q_\alpha(t,\cdot)\big)_{t>0}$ is a family of probability densities on $\mathbb{R}^d$ such that

$$q_\alpha(t,y) = t^{-d/\alpha}\, q_\alpha(1, t^{-1/\alpha} y),\ t > 0,\ y \in \mathbb{R}^d$$

and for all $\gamma \in [0,\alpha)$, there exists a constant $c := c(\alpha,\eta,\gamma)$ s.t.

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

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,

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

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

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

(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

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

(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). □

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]. The *thermic* terminology comes from the fact that such a norm involves convolution with a suitable heat kernel.

In the following, we denote by $\mathcal{S}(\mathbb{R}^d)$ the Schwartz class. For $f\in\mathcal{S}'(\mathbb{R}^d)$, $\varphi\in C_0^\infty(\mathbb{R}^d)$ (smooth function with compact support) s.t. $\varphi(0)\neq 0$ we set $\varphi(D)f := (\varphi\hat{f})^\vee$ 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 thermic characterisation of Besov spaces we will use in the current work reads as follows: for $\vartheta\in\mathbb{R}, m\in(0,+\infty], l\in(0,\infty]$, $\mathbb{B}^\vartheta_{l,m}(\mathbb{R}^d,\mathbb{R}) := \{f\in\mathcal{S}'(\mathbb{R}^d) : \|f\|_{\mathcal{H}^\vartheta_{l,m},\tilde{\alpha}} < +\infty\}$, for any $\tilde{\alpha}\in[1,2]$, with

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

$$=: \quad \|\varphi(D)f\|_{\mathbb{L}^l(\mathbb{R}^d)} + \mathcal{T}^{\vartheta}_{l,m}[f],$$

where in the above definition "$\star$" stands for the spatial convolution.

The parameter n is an integer s.t. $n > \vartheta/\tilde{\alpha}$ and $\tilde{p}_{\tilde{\alpha}}$ denotes the isotropic $\tilde{\alpha}$ stable heat kernel on $\mathbb{R}^d$ (or the gaussian heat kernel if $\tilde{\alpha} = 2$). In Section 2.6.4. of [Tri83], the thermic characterization is presented either with the Gaussian heat kernel ($\tilde{\alpha} = 2$) or with the Cauchy-Poisson kernel ($\tilde{\alpha} = 1$). It actually turns out from the main characterization of Besov spaces, see Section 2.5.1 in [Tri83], that many kernels can actually be used. We chose the isotropic stable one, with stability index associated with the one of the driving noise in (1.1), *i.e.* $\tilde{\alpha} = \alpha$, since it precisely allows to benefit from the stability by convolution when the underlying stable semi-group, around which we perform the Duhamel expansion (see (3.26), (3.16)), is involved. This is particularly well adapted to the computations in the proof of Lemma 12, see Appendix A. In particular $\tilde{p}_\alpha$ satisfies the bounds of Lemma 11 and in that case the upper-bounding density can be specified. Namely, in that case (3.1) holds with $q_\alpha(t,x) = C_\alpha t^{-d/\alpha}(1 + |x|/t^{1/\alpha})^{-(d+\alpha)}$.

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

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

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

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 a negligible contribution. For that reason, we will only focus on the estimation of the *thermic part* of the Besov norm below. See Remark 18 in the proof of Lemma 12 in Appendix A for details.

Remark 11. One may wonder why we chose to work with the Thermic characterization of Besov spaces instead of the one deriving from the dyadic Littlewood-Paley decomposition. Note first of all that such characterizations are equivalent, we can e.g. refer to Section 2.5.1 of [Tri83] (main Theorem) or Chapters 3 and 5 in [LR02]. The main point here is that we are led to handle a mild (or Duhamel) formulation of a parabolic PDE which itself involves convolutions of distributions by a heat kernel (associated with the driving noise). Such a framework hence naturally leads to consider the thermic characterization and to choose therein a heat kernel which is also compatible with the one of the driving noise. Let us mention that the Littlewood-Paley was in this same context of Schauder type estimates successfully used by Zhang and his co-authors, see e.g. [HWZ20] in the degenerate non-local kinetic case. The thermic approach seems more natural and direct to us for our goal.

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 12. 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 spatial operator $\mathscr{D}^\eta$ by

$$\mathscr{D}^\eta := \begin{cases} \mathrm{Id} \text{ if } \eta = 0, \\ \nabla \text{ if } \eta = 1, \\ L^\alpha \text{ if } \eta = \alpha, \end{cases} \tag{3.11}$$

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 $\beta' \in (0,1)$,

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

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

3.2 A primer on PDE. The *formal* Cauchy problem $\mathscr{C}(F,L^\alpha,f,g,T)$: reading almost optimal regularity through Green kernel estimates

Equation (1.13) can be, still formally, 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 sources, 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 (1.13), 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)-(1.13).

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 $\eta \in \mathbb{N}^d, |\eta| := \sum_{i=1}^d \eta_i \leq 1$, we want to control $D_x^\eta G^\alpha \Phi(t,x)$.

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:

$$|D_x^\eta G^\alpha \Phi(t,x)| = \left| \int_t^T ds \int_{\mathbb{R}^d} dy \Phi(s,y) D_x^\eta p(s-t,y-x) \right|$$

$$\leq \quad \|\Phi\|_{\mathbb{L}^r((t,T],\mathbb{B}^{-1+\gamma}_{p,q})}\|D^\eta_x p_\alpha(\cdot - t, \cdot - x)\|_{\mathbb{L}^{r'}((t,T],\mathbb{B}^{1-\gamma}_{p',q'})},$$

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 (see (3.9)):

$$\begin{aligned}
&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot - x)]\Big)^{q'}\\
= \quad & \int_0^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde p_\alpha(v,\cdot) \star D^\eta_x p_\alpha(s-t,\cdot - x)\|^{q'}_{\mathbb{L}^{p'}}\\
= \quad & \int_0^{(s-t)} \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde p_\alpha(v,\cdot) \star D^\eta_x p_\alpha(s-t,\cdot - x)\|^{q'}_{\mathbb{L}^{p'}}\\
& + \int_{(s-t)}^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde p_\alpha(v,\cdot) \star D^\eta_x p_\alpha(s-t,\cdot - x)\|^{q'}_{\mathbb{L}^{p'}}\\
=: \quad & \Big(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot - x)]|_{[0,(s-t)]}\Big)^{q'} \qquad (3.17)\\
& + \Big(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot - x)]|_{[(s-t),1]}\Big)^{q'}.
\end{aligned}$$

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 11). 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 11).

Let us begin with the upper part (i.e. the second term in (3.17)). Using the $\mathbb{L}^1 - \mathbb{L}^{p'}$ convolution inequality, we have from Lemma 11:

$$\begin{aligned}
&\Big(\mathcal{T}^{1-\gamma}_{p',q'}[D^\eta_x p_\alpha(s-t,\cdot - x)]|_{[(s-t),1]}\Big)^{q'}\\
\leq \quad & \int_{(s-t)}^1 \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha})q'} \|\partial_v \tilde p_\alpha(v,\cdot)\|^{q'}_{\mathbb{L}^1} \|D^\eta_x p_\alpha(s-t,\cdot - x)\|^{q'}_{\mathbb{L}^{p'}}\\
\leq \quad & \frac{C}{(s-t)^{(\frac{d}{p\alpha}+\frac{|\eta|}{\alpha})q'}} \int_{(s-t)}^1 \frac{dv}{v} \frac{1}{v^{\frac{1-\gamma}{\alpha}q'}} \leq \frac{C}{(s-t)^{[\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}]q'}}. \qquad (3.18)
\end{aligned}$$

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

$$\begin{aligned}
&\|D^\eta_x p_\alpha(s-t,\cdot - x)\|_{\mathbb{L}^{p'}} = \Big(\int_{\mathbb{R}^d} \big|D^\eta_x p_\alpha(s-t,x)\big|^{p'} dx\Big)^{1/p'}\\
\leq \quad & \frac{C_{p'}}{(s-t)^{\frac{|\eta|}{\alpha}}} \Big((s-t)^{-\frac{d}{\alpha}(p'-1)} \int_{\mathbb{R}^d} \frac{dx}{(s-t)^{\frac{d}{\alpha}}} \big(q_\alpha(1, \frac{x}{(s-t)^{\frac{1}{\alpha}}})\big)^{p'}\Big)^{1/p'}\\
\leq \quad & C_{p'}(s-t)^{-[\frac{d}{\alpha p}+\frac{|\eta|}{\alpha}]} \Big(\int_{\mathbb{R}^d} d\tilde x \big(q(1,\tilde x)\big)^{p'}\Big)^{1/p'} \leq \bar C_{p'}(s-t)^{-[\frac{d}{\alpha p}+\frac{|\eta|}{\alpha}]},
\end{aligned} \qquad (3.19)$$

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

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

$$-r'\left[\frac{1-\gamma}{\alpha} + \frac{d}{p\alpha} + \frac{|\eta|}{\alpha}\right] > -1 \iff |\eta| < \alpha(1-\frac{1}{r}) + \gamma - 1 - \frac{d}{p}. \qquad (3.20)$$

We now focus on the lower part (i.e. the first term in (3.17)). 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 $(x,y,z) \in (\mathbb{R}^d)^3$,

$$\begin{aligned} & |D_x^\eta p_\alpha(s-t,z-x) - D_x^\eta p_\alpha(s-t,y-x)| \\ \leq\ & \frac{C}{(s-t)^{\frac{\beta+|\eta|}{\alpha}}} |z-y|^\beta \Big(q_\alpha(s-t,z-x) + q_\alpha(s-t,y-x) \Big). \end{aligned} \tag{3.21}$$

Indeed, (3.21) is direct if $|z-y| \geq (1/2)(s-t)^{1/\alpha}$ (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 $\big(|z-y|/(s-t)^{1/\alpha}\big)^\beta \geq 1$. 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 11 (see in particular (3.5)), for all $\lambda \in [0,1]$:

$$\begin{aligned} & |D_x^\eta D p_M(s-t, y-x+\lambda(y-z))| \\ \leq\ & \frac{C_m}{(s-t)^{\frac{|\eta|+1}{\alpha}}} p_{\bar M}(s-t, y-x-\lambda(y-z)) \\ \leq\ & \frac{C_m}{(s-t)^{\frac{|\eta|+1+d}{\alpha}}} \frac{1}{\Big(1 + \frac{|y-x-\lambda(z-y)|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m} \\ \leq\ & \frac{C_m}{(s-t)^{\frac{|\eta|+1+d}{\alpha}}} \frac{1}{\Big(\frac{1}{2} + \frac{|y-x|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m} \leq 2 \frac{C_m}{(s-t)^{\frac{|\eta|+1}{\alpha}}} p_{\bar M}(s-t, y-x). \end{aligned} \tag{3.22}$$

Therefore, in the diagonal case, (3.21) follows from (3.22) and (3.5) writing

$$\begin{aligned} |D_x^\eta p_\alpha(s-t,z-x) - D_x^\eta p_\alpha(s-t,y-x)| \leq & \int_0^1 d\lambda |D_x^\eta D p_\alpha(s-t, y-x+\lambda(y-z)) \cdot (y-z)| \\ \leq & 2C_m (s-t)^{-(|\eta|+1)/\alpha} q_\alpha(s-t,y-x)|z-y| \\ \leq & \tilde C_m (s-t)^{-(|\eta|+\beta)/\alpha} q_\alpha(s-t,y-x)|z-y|^\beta, \end{aligned}$$

for all $\beta \in [0,1]$ (exploiting again that $|z-y| \leq (1/2)(s-t)^{1/\alpha}$ for the last inequality). From (3.21) we now derive:

$$\begin{aligned} & \|\partial_v \tilde p_\alpha(v,\cdot) \star D_x^\eta p_\alpha(s-t,\cdot - x)\|_{L^{p'}} \\ =\ & \Big(\int_{\mathbb{R}^d} dz \Big| \int_{\mathbb{R}^d} dy \partial_v \tilde p_\alpha(v,z-y) D_x^\eta p_\alpha(s-t,y-x) \Big|^{p'} \Big)^{1/p'} \\ =\ & \Big(\int_{\mathbb{R}^d} dz \Big| \int_{\mathbb{R}^d} dy \partial_v \tilde p_\alpha(v,z-y) \\ & \times \Big[D_x^\eta p_\alpha(s-t,y-x) - D_x^\eta p_\alpha(s-t,z-x) \Big] \Big|^{p'} \Big)^{1/p'} \\ \leq\ & \frac{1}{(s-t)^{\frac{|\eta|+\beta}{\alpha}}} \Big(\int_{\mathbb{R}^d} dz \Big| \int_{\mathbb{R}^d} dy |\partial_v \tilde p_\alpha(v,z-y)|\, |z-y|^\beta \\ & \times \big[q_\alpha(s-t,y-x) + q_\alpha(s-t,z-x) \big] \Big|^{p'} \Big)^{1/p'} \\ \leq\ & \frac{C_{p'}}{(s-t)^{\frac{|\eta|+\beta}{\alpha}}} \Bigg[\Big(\int_{\mathbb{R}^d} dz \Big| \int_{\mathbb{R}^d} dy |\partial_v \tilde p_\alpha(v,z-y)|\, |z-y|^\beta q_\alpha(s-t,y-x) \Big|^{p'} \Big)^{1/p'} \\ & + \Big(\int_{\mathbb{R}^d} dz \big(q_\alpha(s-t,z-x) \big)^{p'} \Big(\int_{\mathbb{R}^d} dy |\partial_v \tilde p_\alpha(v,z-y)|\, |z-y|^\beta \Big)^{p'} \Big)^{1/p'} \Bigg]. \end{aligned} \tag{3.23}$$

From the $\mathbb{L}^1 - \mathbb{L}^{p'}$ convolution inequality and Lemma 11 (see also (3.19)) we thus obtain:

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

Hence,

$$\begin{aligned} &\Big(\mathcal{T}^{1-\gamma}_{p',q'}[D_x^\eta p_\alpha(s-t,\cdot-x)]|_{[0,(s-t)]}\Big)^{q'} \\ \le\ & \frac{C}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}+\frac{\beta}{\alpha}\right]q'}} \int_0^{(s-t)} \frac{dv}{v} v^{(1-\frac{1-\gamma}{\alpha}-1+\frac{\beta}{\alpha})q'} \\ \le\ & \frac{C}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}+\frac{\beta}{\alpha}+\frac{1-\gamma-\beta}{\alpha}\right]q'}} = \frac{C}{(s-t)^{\left[\frac{d}{p\alpha}+\frac{|\eta|}{\alpha}+\frac{1-\gamma}{\alpha}\right]q'}}, \end{aligned} \tag{3.24}$$

provided $\beta+\gamma>1$ for the second inequality (which can be assumed since we can choose β arbitrarily in $(0,1)$). The map $s \mapsto \mathcal{T}^{1-\gamma}_{p',q'}[D_x^\eta p_\alpha(s-t,\cdot-x)]|_{[0,(s-t)]}$ hence belongs to $\mathbb{L}^{r'}((t,T],\mathbb{R}^+)$ under the same previous condition on η than in (3.20). Let us eventually mention that the above arguments somehow provide the lines of the proof of Lemma 12 for $\Psi = 1$. The proof in its whole generality is provided in Appendix A.

Remark 12 (Pointwise gradient estimate on G^α)**.** The condition in (3.20) 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:

$$1 < \alpha(1-\frac{1}{r}) + \gamma - 1 - \frac{d}{p} \iff \gamma > 2 - \alpha(1-\frac{1}{r}) + \frac{d}{p}. \tag{3.25}$$

In particular, provided (3.25) 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 $F \cdot Du$ on the way.

Remark 13 (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, none of the thresholds appearing depend on this parameter. Since for all γ, p we have that for all $q < q'$ that $B^\gamma_{p,q} \hookrightarrow B^\gamma_{p,q'}$ 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 Cauchy problem $\mathscr{C}(F_m, L^\alpha, f, g, T)$ and associated (uniform) Hölder controls

It is known that, under **(UE)** and for $\vartheta > \alpha$, if $g \in \mathbb{B}^\vartheta_{\infty,\infty}(\mathbb{R}^d,\mathbb{R})$ is also bounded and $f \in L^\infty([0,T],\mathbb{B}^{\vartheta-\alpha}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}))$, for any $m \ge 1$ there exists a unique classical solution $u_m \in \mathbb{L}^\infty([0,T],\mathbb{B}^\vartheta_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}))$ to the Cauchy problem $\mathscr{C}(F_m, L^\alpha, f, g, T)$. 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 (1.8):

$$u_m(t,x) = P^\alpha_{T-t}[g](x) - G^\alpha f(t,x) + \mathfrak{r}_m(t,x), \tag{3.26}$$

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

$$\mathfrak{r}_m(t,x) := \int_t^T ds P^\alpha_{T-s}[F_m(s,\cdot)\cdot Du_m(s,\cdot)](x). \tag{3.27}$$

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

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

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. Note also that the following analysis will allow us to drop the above condition $\vartheta > \alpha$, i.e. the above Duhamel representation holds in our setting with $\vartheta = \theta$.

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

$$\begin{aligned} |DP^\alpha_{T-t}[g](x)| \ & \leq \ \sum_{j=1}^d |\partial_{x_j} P^\alpha_{T-t}[g](x)| \leq \sum_{j=1}^d \left| \int_{\mathbb{R}^d} dy \partial_j g(y) p_\alpha(T-t, y-x) \right| \\ & \leq \ \sum_{j=1}^d C\|Dg\|_{\mathbb{B}^{\theta-1}_{\infty,\infty}}. \end{aligned} \tag{3.28}$$

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 (1.12) of θ we have $\theta - \alpha + 1 < 1$ and $(\theta - \alpha + 1) + 1 > 1$. We can thus apply Lemma 12 (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

$$\|DG^\alpha f\|_{\mathbb{L}^\infty} \leq C(T-t)^{\frac{\theta-1}{\alpha}} \|f\|_{\mathbb{L}^\infty([0,T],\mathbb{B}^{\theta-\alpha}_{\infty,\infty})}. \tag{3.29}$$

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

$$
\begin{aligned}
|D\mathfrak{r}(t,x)| &\le \sum_{j=1}^d |\partial_{x_j}\mathfrak{r}(t,x)|\\
&\le \sum_{j=1}^d\sum_{k=1}^d \left|\int_t^T ds\int_{\mathbb{R}^d} dy F_k(s,y)\partial_{y_k}u(s,y)\partial_{x_j}p_\alpha(s-t,y-x)\right|\\
&\le \sum_{j=1}^d\sum_{k=1}^d \|F_k\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})}\|\partial_k u\partial_{x_j}p_\alpha(\cdot-t,\cdot-x)\|_{\mathbb{L}^{r'}(\mathbb{B}^{1-\gamma}_{p',q'})},
\end{aligned}
\tag{3.30}
$$

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. (1.12)) 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 12 so that (see eq. (3.12) with $\beta=\theta-1-\varepsilon$, $\eta=1$ and $\Psi(s,\cdot)=\partial_k u(s,\cdot)$ therein)

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

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

$$
-r'\left[\frac{d}{p\alpha}+\frac{1}{\alpha}+\frac{1-\gamma}{\alpha}\right]>-1\Leftrightarrow \gamma>2-\alpha+\frac{\alpha}{r}+\frac{d}{p},
\tag{3.31}
$$

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

$$
|D\mathfrak{r}(t,x)|\le C(T-t)^{\frac{\theta-1}{\alpha}}\|Du\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})}\le CT^{\frac{\theta-1}{\alpha}}\|Du\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})},
\tag{3.32}
$$

recalling from (1.12) that $\theta=\gamma-1+\alpha-d/p-\alpha/r$.

(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:

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

using again the Hölder inequality and duality between the considered Besov spaces (see Section 3.1). 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 12 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):

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

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

$$-r'\left[\frac{d}{p\alpha}+\frac{1+(\theta-1-\varepsilon)}{\alpha}+\frac{1-\gamma}{\alpha}\right]>-1\Leftrightarrow\theta-1-\varepsilon<\gamma-\left(2-\alpha+\frac{\alpha}{r}+\frac{d}{p}\right),\qquad(3.33)$$

which readily follows from the very definition of θ (see eq. (1.12)) and the fact that $\varepsilon>0$. We then obtain

$$\begin{aligned}|D\mathfrak{r}(t,x)-D\mathfrak{r}(t,x')|\ &\le\ C(T-t)^{\frac{\varepsilon}{\alpha}}\|Du\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})}|x-x'|^{\theta-1-\varepsilon}\\ &\le\ CT^{\frac{\varepsilon}{\alpha}}\|Du\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty})}|x-x'|^{\theta-1-\varepsilon}.\end{aligned}\qquad(3.34)$$

Remark 14. Note that assuming that θ is fixed, we readily obtain from (3.33) together with the constraint $\theta-1-\varepsilon+\gamma>1$ the initial constraint

$$\gamma>\frac{3-\alpha+\frac{d}{p}+\frac{\alpha}{r}}{2}.\qquad(3.35)$$

In comparison with the threshold obtained when investigating the smoothing effect of the Green kernel (see eq. (3.25) and the related discussion) this additional regularity[3] allows to handle the product $F\cdot Du$, in the sense that it allows to give a meaning to this product as a distribution. Indeed, as suggested by Lemma 12 (replacing therein the heat kernel by a smooth test function), a sufficient condition to define the product $F\cdot Du$ is to obtain estimate on the Hölder modulus of the map Du of order β for some $\beta>1-\gamma=-(-1+\gamma)$ (see (3.12)). The threshold (3.35) precisely reflects that constraint and appears as the price to pay to define such a modulus.

Let us eventually estimate the Hölder moduli of the gradients of the first and second terms in the Duhamel representation (3.26). We first note that, for the Green kernel, the proof follows from the above lines. When doing so, we obtain that

$$|DG^\alpha f(t,x)-DG^\alpha f(t,x')|\le CT^{\frac{\varepsilon}{\alpha}}\|f\|_{\mathbb{B}^{\theta-\alpha}_{\infty,\infty}}|x-x'|^{\theta-1-\varepsilon}.\qquad(3.36)$$

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

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

[3] recall that since we assumed in (GR) $\alpha>\frac{1+\frac{d}{p}}{1-\frac{1}{r}}$, it indeed holds that $\frac{3-\alpha+\frac{d}{p}+\frac{\alpha}{r}}{2}>2-\alpha+\frac{d}{p}+\frac{\alpha}{r}$.

recalling that p_α is a density for the first inequality. On the other hand, when $(T-t)^{\frac{1}{\alpha}} > |x-x'|$ (diagonal regime), we have using cancellations arguments

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

Hence

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

Putting together estimates (3.28), (3.29), (3.32), (3.34), (3.36) and (3.37) we deduce that

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

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 (3.26). Namely,

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

where from (3.27), the remainder term writes:

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

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

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

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:

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

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 11:

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

recalling from (1.12) 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,

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

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

$$\begin{aligned}
\mathfrak{r}(t', x) - \mathfrak{r}(t, x) &= \int_{t'}^T ds\big(P^\alpha_{s-t'} - P^\alpha_{s-t}\big)[F(s,\cdot)\cdot Du(s,\cdot)](x) \\
&\quad + \int_t^{t'} ds P^\alpha_{s-t}[F(s,\cdot)\cdot Du(s,\cdot)](x). \qquad (3.41)
\end{aligned}$$

From Lemma 12 (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

$$|\int_t^{t'} ds P^\alpha_{s-t}[F(s,\cdot)\cdot Du(s,\cdot)](x)| \leq C\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})}|t'-t|^{\frac{\theta}{\alpha}}. \qquad (3.42)$$

Similarly,

$$\begin{aligned}
&|\int_{t'}^{[t'+\frac{1}{2}(t'-t)]\wedge T} ds(P^\alpha_{s-t'} - P^\alpha_{s-t})[F(s,\cdot)\cdot Du(s,\cdot)](x)| \\
\leq & |\int_{t'}^{[t'+\frac{1}{2}(t'-t)]\wedge T} ds(P^\alpha_{s-t'}[F(s,\cdot)\cdot Du(s,\cdot)](x)| \\
&+ |\int_{t'}^{[t'+\frac{1}{2}(t'-t)]\wedge T} ds P^\alpha_{s-t}[F(s,\cdot)\cdot Du(s,\cdot)](x)| \leq C|t'-t|^{\frac{\theta}{\alpha}}. \qquad (3.43)
\end{aligned}$$

Assuming now w.l.o.g. that $[t' + \frac{1}{2}(t'-t)] \wedge T = t' + \frac{1}{2}(t'-t)$, let us focus on

$$\begin{aligned}
&\int_{t'+\frac{1}{2}(t'-t)}^T ds\big(P^\alpha_{s-t'} - P^\alpha_{s-t}\big)[F(s,\cdot)\cdot Du(s,\cdot)](x) \\
=\ & \int_{t'+\frac{1}{2}(t'-t)}^T ds\int_0^1 d\lambda\Big\{\partial_w P^\alpha_{s-w}[F(s,\cdot)\cdot Du(s,\cdot)](x)\Big\}\Big|_{w=t+\lambda(t'-t)}(t'-t) \\
=\ & \int_0^1 d\lambda\int_{t'+\frac{1}{2}(t'-t)}^T ds\Big\{L^\alpha P^\alpha_{s-w}[F(s,\cdot)\cdot Du(s,\cdot)](x)\Big\}\Big|_{w=t+\lambda(t'-t)}(t'-t).
\end{aligned}$$

(3.44)

Keeping the notation $w = t + \lambda(t'-t)$, we have

$$\begin{aligned}
&\int_{t'+\frac12(t'-t)}^T ds |L^\alpha P^\alpha_{s-w}[F(s,\cdot)\cdot Du(s,\cdot)](x)| \\
\le\ & \sum_{k=1}^d \int_{t'+\frac12(t'-t)}^T ds \Big| \int_{\mathbb{R}^d} dy F_k(s,y)\partial_{y_k} u(s,y) L^\alpha p_\alpha(s-w,y-x)\Big| \\
\le\ & \sum_{k=1}^d \|F_k\|_{\mathbb{L}^r([t',T],\mathbb{B}^{-1+\gamma}_{p,q})} \|\partial_k u L^\alpha p_\alpha(\cdot - w, \cdot - x)\|_{\mathbb{L}^{r'}([t'+\frac12(t'-t),T],\mathbb{B}^{1-\gamma}_{p',q'})}. \qquad (3.45)
\end{aligned}$$

Applying Lemma 12 (see eq. (3.12) with $\beta = \theta - 1 - \varepsilon$ and $\eta = \alpha$ therein), we get:

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

Thus, from (3.38) (recall from (1.12) that $\gamma - 2 + \alpha - \frac{d}{p} - \frac{\alpha}{r} - \varepsilon = \theta - 1 - \varepsilon$) and since on the considered time interval $s \in [t' + \frac12(t'-t), T]$ there is no time-singularity (indeed $s - w = s - (t + \lambda(t'-t)) \ge t' + \frac12(t'-t) - w \ge \frac12(t'-t)$ for all $\lambda \in [0,1]$):

$$\begin{aligned}
\|\partial_k u L^\alpha p_\alpha(\cdot - w,\cdot - x)\|_{\mathbb{L}^{r'}([t'+\frac12(t'-t),T],\mathbb{B}^{1-\gamma}_{p',q'})} &\le C(t' + \frac{1}{2}(t'-t) - w)^{\frac{1}{r'}-\left(\frac{1-\gamma}{\alpha}+\frac{d}{p\alpha}+1\right)} \\
&= C(t'-t)^{\frac{\theta}{\alpha}-1}. \qquad (3.46)
\end{aligned}$$

Therefore, from (3.46) and (3.45), we derive:

$$\int_{t'+\frac12(t'-t)}^T ds |L^\alpha P^\alpha_{s-w}[F(s,\cdot)\cdot Du(s,\cdot)](x)| \le C \sum_{k=1}^d \|F_k\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})} (t'-t)^{\frac{\theta}{\alpha}-1},$$

which in turn, plugged into (3.44) and using (3.43) as well gives:

$$|\int_{t'}^T ds \big(P^\alpha_{s-t'} - P^\alpha_{s-t}\big)[F(s,\cdot)\cdot Du(s,\cdot)](x)| \le C\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})} (t'-t)^{\frac{\theta}{\alpha}}. \qquad (3.47)$$

From (3.47), (3.42) and (3.41) we thus obtain:

$$\big|\mathfrak{r}(t',x) - \mathfrak{r}(t,x)\big| \le C\|F\|_{\mathbb{L}^r(\mathbb{B}^{-1+\gamma}_{p,q})} (t'-t)^{\frac{\theta}{\alpha}}. \qquad (3.48)$$

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 $C \ge 1$ s.t. for all $0 \le t < t' \le T$, $x \in \mathbb{R}^d$:

$$\Big|\big(G^\alpha f(t',x) - G^\alpha f(t,x)\big)\Big| \le C\|f\|_{\mathbb{L}^\infty(\mathbb{B}^{\theta-\alpha}_{\infty,\infty})} (t'-t)^{\frac{\theta}{\alpha}}. \qquad (3.49)$$

The final control of (2.2) concerning the smoothness in time then follows plugging (3.40), (3.48) and (3.49) into (3.39). The control concerning the time sensitivity of the spatial gradient would be obtained following the same lines.

3.4 Proofs of Theorem 2, Proposition 8, Proposition 9, Corollary 10 and Lemma 1

Points *(i)* to *(iii)* conclude the proof of Proposition 8 up to the convergence assertion. Let us notice that the previous analysis allows to obtain Proposition 9 up to this assertion as well. Indeed, in such a case, the map f_m is the k^{th} coordinate of $-F_m$ and should thus be estimated in term of its $\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d,\mathbb{R}))$ norm. The associate control for this source can be obtained following exactly the same strategy as the one we used to handle the remainder term $\mathfrak{r}$ in the Duhamel representation (3.26) except that we do not need to deal with the additional gradient of the solution which previously led to some additional specific constraints to define $Du_m \cdot F_m$ uniformly in m. The fact that the constant therein are decreasing w.r.t. time follows from the fact that $g \equiv 0$, see (3.38).

Eventually, end of the proof of Proposition 8 and the proof of Theorem 2 follows from compactness arguments together with the Schauder like control of Propositions 8 and the previous analysis. Uniqueness follows from the Schauder like control of Propositions 8 as well, as the underlying PDE is linear. The representation (1.14) holds through similar computations. Corollary 10 is derived in the same way through Proposition 9, whose proof is concluded following the same lines as for Proposition 8.

Lemma 1 follows from the above calculations as well. The fact that the product $F \cdot Du$ makes sense is an easy consequence of Lemma 12 (see estimate (3.13) with the heat kernel therein replaced by a smooth test function) together with the regularity of Du.

Remark 15 (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:

$$\begin{aligned} u_m(t,x) &= P^{\alpha,\xi,m}_{s-t}[g](x) - \int_t^T ds P^{\alpha,\xi,m}_{s-t}\Big[\Big\{ f(s,\cdot) - F_m \cdot Du_m(s,\cdot) \\ &\quad + \frac{1}{2}\mathrm{Tr}\big((a_m(s,\cdot) - a_m(s,\xi)) D^2 u_m(s,\cdot)\big)\Big\}\Big](x), \end{aligned} \tag{3.50}$$

for an auxiliary parameter ξ which will be taken equal to x after potential differentiations in (3.50). 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 8 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:

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

From the previous Proposition 8, we aim at establishing that Du_m has Hölder index $\theta - 1 - \varepsilon = \gamma - 2 + \alpha - d/p - \alpha/r - \varepsilon$ and therefore $D_{y_i y_j} u_m \in \mathbb{B}^{\theta-2-\varepsilon}_{\infty,\infty}$. 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:

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

Now, in the considered case $\theta-2-\varepsilon=\gamma-1-\varepsilon$. Recalling that $D_x p_\alpha^{\xi,m}(t,s,x,\cdot)\in \mathbb{B}_{1,1}^{1/2-\tilde\varepsilon}$ for any $\tilde\varepsilon>0$ for $\gamma>1/2=(3-\alpha)/2$ and ε small enough, we will indeed have that $D_x p_\alpha^{\xi,m}(t,s,x,\cdot)\big((a_{m,i,j}(s,\cdot)-a_{m,i,j}(s,\xi)\big)\in \mathbb{B}_{1,1}^{2+\varepsilon-\theta}$ provided the bounded function a itself has the same regularity, i.e. $2+\varepsilon-\theta=1-\gamma+\varepsilon$, the integrability of the product deriving from the one of the heat kernel. Since $\|\partial_{i,j}^2 u_m(s,\cdot)\|_{\mathbb{B}^{\theta-2-\varepsilon}_{\infty,\infty}}\le C\|Du_m(s,\cdot)\|_{\mathbb{B}^{\theta-1-\varepsilon}_{\infty,\infty}}$, 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 $p,q,r\ge 1$ and γ satisfying the conditions of Theorem 3 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 $L^{\alpha,\sigma}\varphi(x)=\mathrm{p.v.}\int_{\mathbb{R}^d}\big(\varphi(x+\sigma(x)z)-\varphi(x)\big)\nu(dz)=-\sigma^\alpha(x)(-\Delta)^{\alpha/2}\varphi(x)$, where σ is a non-degenerate diffusion coefficient. Introducing $L^{\alpha,\sigma,\xi}\varphi(x)=\mathrm{p.v.}\int_{\mathbb{R}^d}\big(\varphi(x+\sigma(\xi)z)-\varphi(x)\big)\,\nu(dz)=-\sigma^\alpha(\xi)(-\Delta)^{\alpha/2}\varphi(x)$, we rewrite for the Duhamel formula, similarly to (3.50):

$$\begin{aligned}u_m(t,x)\;&=\;P^{\alpha,\xi,m}_{s-t}[g](x)-\int_t^T ds P^{\alpha,\xi,m}_{s-t}\Big[\Big\{f(s,\cdot)-F_m\cdot Du_m(s,\cdot)\\&\qquad+(L^{\alpha,\sigma_m}-L^{\alpha,\sigma_m,\xi})u_m(s,\cdot)\big)\Big\}\Big](x).\end{aligned}\tag{3.51}$$

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

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

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 $D_xp_\alpha^{\xi,m}(t,s,x,y)\,\big(\sigma_m^\alpha(s,y)-\sigma_m^\alpha(s,\xi)\big)$ in the Besov space $\mathbb{B}^{1-\gamma+\varepsilon}_{1,1}$. Since $\gamma>(3-\alpha)/2$ and $D_xp_\alpha^{\xi,m}(t,s,x,y)\in\mathbb{B}^{1-1/\alpha-\tilde\varepsilon}_{1,1},\tilde\varepsilon>0$ meant to be small, this will be the case provided the coefficient $\sigma\in\mathbb{L}^\infty([0,T],\mathbb{B}^{1-\gamma+\varepsilon}_{\infty,\infty})$ for ε small enough observing that $1-\gamma+\varepsilon<(\alpha-1)/2<1-\frac{1}{\alpha}-\tilde\varepsilon$.

Note that, in comparison with the result obtained in [LZ22], the above threshold is precisely the one appearing in [LZ22] in this specific case. The general matrix case for σ is more involved. It requires in [LZ22] 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 3 could be performed similarly.

4 Dynamics of the *formal* SDE (1.1)

In this part, we aim at proving Theorem 4 and Corollary 5. We restrict here to the pure jump case $\alpha\in(1,2)$, since the diffusive one was already considered in [DD16]. We adapt below their procedure to the current framework.

In subsection 4.1, we first recover the noise through an enlarged martingale problem (point (i) of Proposition 13 below), then recover a drift as the difference between the Martingale solution and the noise obtained before and estimate its contribution (point (ii) of Proposition 13 below). With this contribution at hand, we show that the drift decomposes as a principal part plus a remainder which has a *negligible* contribution (point (iii) of Proposition 13 below). Then, we recall in Subsection 4.2 how the general construction of the stochastic non-linear Young integral from [DD16] translates in our setting. Eventually, we derive in subsection 4.3 the dynamics associated with the solution of the Martingale Problem and define the class of processes to which an associated Itô's formula holds. This last part thus conclude the proof of Theorem 4 and Corollary 5. Importantly, we suppose throughout this section that (GR-D) holds.

4.1 Shape of the drift

Proposition 13. Let $\alpha \in (1,2)$. For any initial point $x \in \mathbb{R}^d$, one can find a probability measure $\mathbf{P}^\alpha$ on $\mathcal{D}([0,T],\mathbb{R}^{2d})$ s.t. the canonical process $(X_t,\mathcal{W}_t)_{t\in[0,T]}$ satisfies the following properties:

(i) Under $\mathbf{P}^\alpha$, the law of $(X_t)_{t\ge 0}$ is a solution of the Martingale Problem associated with data (L^α, F, x), $x \in \mathbb{R}^d$ and the law of $(\mathcal{W}_t)_{t\ge 0}$ corresponds to the one of a d-dimensional stable process with generator L^α.

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

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

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

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

with $\mathfrak{f}(v,X_v,s-v) := u(v,X_v) - X_v$, where u is the mild solution of the Cauchy problem $\mathscr{C}(F,L^\alpha,0,x,s)$ (note that the dependence of $\mathfrak{f}(v,X_v,s-v)$ on s is precisely through the Cauchy problem $\mathscr{C}(F,L^\alpha,0,x,s)$).

Furthermore, the following decomposition holds:

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

Remark 16. The above proposition gives a first information on the shape of the drift. Indeed, using the decomposition

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

one can see that the infinitesimal increment of the canonical process X involves a drift part (first term in the above r.h.s.) and a martingale part (second and third terms in the above r.h.s.). The main point being now that

$$\mathbb{E}[X_{t+h} - X_t|\mathcal{F}_t] = \mathfrak{f}(t,X_t,h) = \mathscr{F}(t,X_t,h) + \mathscr{R}(t,X_t,h) = \mathscr{F}(t,X_t,h) + \mathcal{O}(h^{1+\varepsilon'}),$$

meaning that only the first term $\mathscr{F}$ matters in the limit, *i.e.* the infinitesimal dynamics involves, as a drift, the mollified version of the initial *distributional* one along the density of the driving noise.

Proof.(i) Coming back to point *(i)* in Section 2.2, the couple $\big((X_t^m,\mathcal{W}_t^m)_{t\in[0,T]}\big)_{m\ge 0}$ 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 $0 \le v < s$. Let $u_m = (u_m^1, \dots, u_m^d)$ where each u_m^k, k in $\{1, \dots, d\}$ is chosen as the solution the Cauchy problem $\mathscr{C}(F_m^k, L^\alpha, 0, x_k, s)$ where x_k is the k^{th} coordinate of $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. We have

$$X_s^m - X_v^m = u_m(s, X_s^m) - u_m(s, X_v^m) = u_m(s, X_s^m) - u_m(v, X_v^m) + u_m(v, X_v^m) - u_m(s, X_v^m).$$

Let us now notice that

$$\mathcal{W}_s^m - \mathcal{W}_v^m = \int_v^s \int_{\mathbb{R}^d \setminus \{0\}} x \tilde{N}^m(dw, dx),$$

so that, from Itô's formula

$$\begin{aligned}
& X_s^m - X_v^m \\
= \; & M_{v,s}^{s,m}(\alpha, u_m, X^m) + [u_m(v, X_v^m) - u_m(s, X_v^m)] \qquad (4.3)\\
= \; & \int_v^s \int_{\mathbb{R}^d \setminus \{0\}} \{u_m(w, X_{w^-}^m + x) - u_m(w, X_{w^-}^m)\} \tilde{N}^m(dw, dx) \\
& + [u_m(v, X_v) - u_m(s, X_v)] \\
= \; & \mathcal{W}_s^m - \mathcal{W}_v^m + [u_m(v, X_v^m) - u_m(s, X_v^m)] \\
& + \int_v^s \int_{|x| \le 1} \{u_m(w, X_{w^-}^m + x) - u_m(w, X_{w^-}^m) - x\} \tilde{N}^m(dw, dx) \\
& + \int_v^s \int_{|x| \ge 1} \{u_m(w, X_{w^-}^m + x) - u_m(w, X_{w^-}^m) - x\} \tilde{N}^m(dw, dx). \\
=: \; & \mathcal{W}_s^m - \mathcal{W}_v^m + [u_m(v, X_v^m) - u_m(s, X_v^m)] + \mathcal{M}_S^m(v, s) + \mathcal{M}_L^m(v, s).
\end{aligned}$$

From the smoothness properties of u_m established in Proposition 8 (in particular $|u_m^s(v, X_v^m) - u_m^s(s, X_v^m)]| \le C(s - v)^{\theta/\alpha}$ and the gradient is uniformly bounded) we have

$$\begin{aligned}
|\mathcal{U}(w, X_{w^-}^m, x)| \; &:= \; \left| u_m(w, X_{w^-}^m + x) - u_m(w, X_{w^-}^m) - x \right| \\
&= \; \Big| \int_0^1 d\lambda (Du_m(w, X_{w^-}^m + \lambda x) - I) \cdot x \Big| \le C(s - w)^{\frac{\theta - 1}{\alpha}} |x|, \qquad (4.4)
\end{aligned}$$

recalling that for all z in $\mathbb{R}^d$, $u_m(s, z) = z$ so that $Du_m(s, z) = I$, and using estimate (2.2). Note that $\left(\mathcal{M}_S^m(v, s)\right)_{0 \le v < s \le T}$ and $\left(\mathcal{M}_L^m(v, s)\right)_{0 \le v < s \le T}$ are respectively $\mathbb{L}^2$ and $\mathbb{L}^{\mathfrak{q}}$ 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

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

where $[\mathcal{M}_L^m]_{(v,s)}$ denotes the corresponding bracket given by the expression

$$\sum_{v \le w \le s} |\mathcal{U}(w, X_{w^-}^m, \Delta \mathcal{W}_w^m)|^2 \mathbf{1}_{|\Delta \mathcal{W}_w^m| \ge 1}.$$

Using the linear growth of $\mathcal{U}$ w.r.t. its third variable (uniformly w.r.t. the second one) from (4.4) together with the fact that $\mathfrak{q}/2 \le 1$ we obtain

$$\begin{aligned}
&\Big(\sum_{v \le w \le s} |\mathcal{U}(w, X_{w^-}^m, \Delta \mathcal{W}_w^m)|^2 \mathbf{1}_{|\Delta \mathcal{W}_w^m| \ge 1} \Big)^{\mathfrak{q}/2} \le C(s - v)^{q \frac{\theta - 1}{\alpha}} \Big(\sum_{v \le w \le s} |\Delta \mathcal{W}_w^m|^2 \mathbf{1}_{|\Delta \mathcal{W}_w^m| \ge 1} \Big)^{\mathfrak{q}/2} \\
&\le C(s - v)^{\mathfrak{q} \frac{\theta - 1}{\alpha}} \sum_{v \le w \le s} |\Delta \mathcal{W}_w^m|^{q} \mathbf{1}_{|\Delta \mathcal{W}_w^m| \ge 1}.
\end{aligned}$$

We then readily get from the compensation formula that

$$\begin{aligned}\mathbb{E}\big[|\mathcal{M}_L^m(v,s)|^{\mathfrak{q}}\big] &\le C(s-v)^{1+\mathfrak{q}\frac{\theta-1}{\alpha}}\int |x|^q \mathbf{1}_{|x|\ge 1}\nu(dx) \le C'(s-v)^{1+\mathfrak{q}\frac{\theta-1}{\alpha}}\\ &\le C'(s-v)^{\frac{q}{\alpha}+\mathfrak{q}\frac{\theta-1}{\alpha}}.\end{aligned}$$

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

$$\begin{aligned}\mathcal{M}_S^m(v,s) &= \mathcal{M}_{S,1}^m(v,s)+\mathcal{M}_{S,2}^m(v,s)\\ &:= \int_v^s\int_{|x|>(s-v)^{\frac{1}{\alpha}}}\mathbf{1}_{|x|\le 1}\mathcal{U}(w,X_{w^-}^m,x)\tilde{N}^m(dw,dx)\\ &\quad+\int_v^s\int_{|x|\le(s-v)^{\frac{1}{\alpha}}}\mathbf{1}_{|x|\le 1}\mathcal{U}(w,X_{w^-}^m,x)\tilde{N}^m(dw,dx).\end{aligned}$$

In the off-diagonal regime (namely for $\mathcal{M}_{S,1}^m(v,s)$), 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 (4.4), and eventually usual convexity arguments together with the compensation formula again to obtain

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

In the diagonal regime (i.e. for $\mathcal{M}_{S,2}^m(v,s)$) we use the BDG inequality and (4.4) to recover integrability w.r.t. the Lévy measure and then use the additional integrability to obtain a better estimate. Namely:

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

Using the above estimates on the $\mathfrak{q}$-moments of $\mathcal{M}_L^m(v,s)$, $\mathcal{M}_{S,1}^m(v,s)$ and $\mathcal{M}_{S,2}^m(v,s)$ the statement follows passing to the limit in m, thanks to Proposition 8.

(iii) Letting $(\mathcal{F}_v^m)_{v\ge 0} := \big(\sigma((X_w^m,\mathcal{W}_w^m)_{0\le w\le v})\big)_{v\ge 0}$, restarting from (4.3) and taking the conditional expectation w.r.t. $\mathcal{F}^m$ yields

$$\mathbb{E}[X_s^m-X_v^m|\mathcal{F}_v^m]=\mathbb{E}[u_m(v,X_v^m)-u_m(s,X_v^m)|\mathcal{F}_v^m]=u_m(v,X_v^m)-X_v^m.$$

Passing to the (weak) limit in m, it can be deduced that from Proposition 8 that

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

where u is the mild solution of $\mathscr{C}(F, L^\alpha, 0, x, s)$. From the mild definition of u in Theorem 2 we obtain that for all $(w, y) \in [s, v] \times \mathbb{R}^d$:

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

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

$$\begin{aligned} & \mathbb{E}[X_s - X_v | \mathcal{F}_v] \\ =\ & u(v, X_v) - u(s, X_v) \\ =\ & \int_v^s dw \int_{\mathbb{R}^d} dy Du(w,y) F(w,y) p_\alpha(w-v, y-X_v) \\ =\ & \int_v^s dw \int_{\mathbb{R}^d} dy F(w,y) p_\alpha(w-v, y-X_v) \\ & + \int_v^s dw \int_{\mathbb{R}^d} dy \int_w^s dw' \int_{\mathbb{R}^d} dy' \big[[Du(w',y') \cdot F(w',y')] \otimes D_y p_\alpha(w'-w, y'-y)\big] F(w,y) \\ & \quad \times p_\alpha(w-v, y-X_v), \end{aligned} \tag{4.5}$$

where we have again plugged the mild formulation of Du from (1.14). Let us first prove that the first term in the above has the right order. Thanks to Lemma 12 (with $\eta = 0$ and $\Psi = \mathrm{Id}$ therein) we obtain that:

$$\begin{aligned} & \mathscr{F}(v, X_v, s-v) \\ :=\ & \Big| \int_v^s dw \int_{\mathbb{R}^d} dy F(w,y) p_\alpha(w-v, y-X_v) \Big| \\ \leq\ & C \|F\|_{\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q})} (s-v)^{1-(\frac{1}{r}+\frac{d}{p\alpha}+\frac{1-\gamma}{\alpha})} \\ \leq\ & C \|F\|_{\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q})} (s-v)^{\frac{1}{\alpha}+\left[1-\frac{1}{\alpha}-(\frac{1}{r}+\frac{d}{p\alpha}+\frac{1-\gamma}{\alpha})\right]} \\ \leq\ & C \|F\|_{\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q})} (s-v)^{\frac{1}{\alpha}+\frac{\theta-1}{\alpha}}. \end{aligned} \tag{4.6}$$

Let us now prove that the second term in the r.h.s. of (4.5) is a negligible perturbation. Setting

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

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:

$$|\mathscr{R}(v, X_v, s-v)| \leq \|F\|_{\mathbb{L}^r([0,T], \mathbb{B}^{-1+\gamma}_{p,q})} \|\psi_{v,\cdot,s}(\cdot)\|_{\mathbb{L}^{r'}([0,T], \mathbb{B}^{1-\gamma}_{p',q'})}. \tag{4.7}$$

Let us now consider the thermic part of $\|\psi_{v,\cdot,s}(\cdot)\|_{\mathbb{L}^{r'}([0,T], \mathbb{B}^{1-\gamma}_{p',q'})}$, which can be split again into a lower and an upper part, as in (3.17). We first deal with the upper part and then with the lower one. With the same previous notations[4]:

$$
\begin{aligned}
&\Big(\mathcal{T}^{1-\gamma}_{p',q'}(\psi_{v,w,s}(\cdot))\Big|_{[(w-v),1]}\Big)^{q'}\\
\le\ & C(w-v)^{-\frac{1-\gamma}{\alpha}q'}\|D\mathfrak{r}(w,\cdot)\|_\infty^{q'}\|p_\alpha(w-v,\cdot,-X_v)\|_{\mathbb{L}^{p'}}^{q'}\\
\le\ & C(w-v)^{-\frac{1-\gamma}{\alpha}q'}(s-w)^{\frac{(\theta-1)}{\alpha}q'}(w-v)^{-\frac{d}{\alpha p}q'}, \qquad (4.8)
\end{aligned}
$$

using (3.32) and (3.19) for the last inequality. Hence,

$$
\Big(\int_v^s dw\Big(\mathcal{T}^{1-\gamma}_{p',q'}(\psi_{v,w,s}(\cdot))\Big|_{[(w-v),1]}\Big)^{r'}\Big)^{1/r'} \le C(s-v)^{\frac{1}{r'}+\frac{\theta-1}{\alpha}-\frac{d}{\alpha p}-\frac{1-\gamma}{\alpha}}. \qquad (4.9)
$$

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 (1.12), we obtain the condition:

$$
\gamma > \frac{3-\alpha+\frac{2d}{p}+\frac{2\alpha}{r}}{2}, \qquad (4.10)
$$

which is precisely (GR-D). 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, as far as the integrability indexes are concerned, i.e. for $p=r=\infty$, (4.10) again corresponds to the condition appearing in Theorem 3.

Let us now deal with the lower part of the thermic characterization. Using a cancellation argument, restarting from (3.34) and (3.1), exploiting as well (3.38) and (3.21), we get for $\beta=\theta-1-\varepsilon$:

$$
\begin{aligned}
&|D\mathfrak{r}(w,y)p_\alpha(w-v,y-x)-D\mathfrak{r}(w,z)p_\alpha(w-v,z-x)| \qquad (4.11)\\
\le\ & C\Bigg[\Bigg(\|D\mathfrak{r}(w,\cdot)\|_{\dot{\mathbb{B}}^\beta_{\infty,\infty}}+\frac{\|D\mathfrak{r}(w,\cdot)\|_{\mathbb{L}^\infty}}{(w-v)^{\frac{\beta}{\alpha}}}\Bigg)\\
&\times(q_\alpha(w-v,y-x)+q_\alpha(w-v,z-x))\Bigg]|y-z|^\beta\\
\le\ & C\Big((s-w)^{\frac{\varepsilon}{\alpha}}+\frac{(s-w)^{\frac{\theta-1}{\alpha}}}{(w-v)^{\frac{\beta}{\alpha}}}\Big)(q_\alpha(w-v,y-x)+q_\alpha(w-v,z-x))\\
&\times|y-z|^\beta,
\end{aligned}
$$

recalling also (3.32) for the last inequality and denoting by $\|\cdot\|_{\dot{\mathbb{B}}^\beta_{\infty,\infty}}$ the homogeneous Besov norm (Hölder modulus of order β). Hence:

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

[4] Pay attention that, in order to absorb some singularities we cannot here directly appeal to Lemma 12 but simply exploit some $\mathbb{L}^\infty$ of $D\mathfrak{r}(t,\cdot)$ in terms of $(T-t)^{\frac{\theta}{\alpha}}$.

$$
\begin{aligned}
&\le \Big(\int_v^s dw (w-v)^{(\frac{\gamma-1+\beta}{\alpha}-\frac{d}{p\alpha})r'}\Big((s-w)^{\frac{\varepsilon}{\alpha}}+\frac{(s-w)^{\frac{\theta-1}{\alpha}}}{(w-v)^{\frac{\beta}{\alpha}}}\Big)^{r'}\Big)^{1/r'}\\
&\le C(s-v)^{\frac{1}{r'}+(\frac{\gamma-1+\beta}{\alpha}-\frac{d}{p\alpha})+\frac{\varepsilon}{\alpha}}=C(s-v)^{\frac{1}{r'}+(\frac{\gamma-1+\theta-1}{\alpha}-\frac{d}{p\alpha})}, \qquad (4.12)
\end{aligned}
$$

which precisely gives a contribution homogeneous to the one of (4.9). We eventually derive that, under the condition (4.10), the remainder in (4.7) is s.t. there exists $\varepsilon' := -1/r + [(\gamma-1+\theta-1)/\alpha] - [d/(p\alpha)] > 0$ for which

$$
|\mathscr{R}(v, X_v, s-v)| \le C(s-v)^{1+\varepsilon'}, \qquad C := C(\|F\|_{\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q})}). \qquad (4.13)
$$

□

4.2 The stochastic non-linear Young integral

Having derived the shape of the drift, 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 $(A(s,t))_{0\le s\le t\le T}$ which will stand, for any $0\le t\le t+h\le t+h'\le T$, for either $(i) A(t,t+h) = X_{t+h} - X_t$ 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: there exists $\varepsilon_0 \in (0, 1-1/\alpha]$, $\varepsilon_1, \varepsilon_1' > 0$ such that for any $1 \le \mathfrak{q} < \alpha$ there exists a constant $C := C(p,q,r,\gamma,\mathfrak{q},T) > 0$ such that

$$
\begin{aligned}
\mathbb{E}^{\frac{1}{\mathfrak{q}}}[|\mathbb{E}[A(t,t+h)|\mathcal{F}_t]|^{\mathfrak{q}}] &\le Ch^{\frac{1}{\alpha}+\varepsilon_0},\\
\mathbb{E}^{\frac{1}{\mathfrak{q}}}[|A(t,t+h)|^{\mathfrak{q}}] &\le Ch^{\frac{1}{\alpha}},\\
\mathbb{E}^{\frac{1}{\mathfrak{q}}}\big[|\mathbb{E}[A(t,t+h)+A(t+h,t+h')-A(t,t+h')|\mathcal{F}_t]|^{\mathfrak{q}}\big] &\le C(h')^{1+\varepsilon_1},\\
\mathbb{E}^{\frac{1}{\mathfrak{q}}}[|A(t,t+h)+A(t+h,t+h')-A(t,t+h')|^{\mathfrak{q}}] &\le C(h')^{\frac{1}{\alpha}+\varepsilon_1'}. \qquad (4.14)
\end{aligned}
$$

Then, we aim at defining for any $T>0$ the stochastic integral $\int_0^T \psi_s A(t,t+dt)$, for processes $(\psi_s)_{s\in[0,T]}$ in $\mathcal{H}^{(1-1/\alpha)-\varepsilon_2}_{\mathfrak{q}'}$ (see (1.18) for the definition) with $\mathfrak{q}'\ge 1$ such that $1/\mathfrak{q}' + 1/\mathfrak{q} = 1/\ell$, for any $1 \le \ell < \alpha$ and $0 < \varepsilon_2 < \varepsilon_0$, as an $\mathbb{L}^\ell$ limit of the associated Riemann sum: for $\Delta = \{0 = t_0 < t_1, \dots, t_N = T\}$

$$
S(\Delta) := \sum_{i=0}^{N-1} \psi_{t_i} A(t_i, t_{i+1}) \xrightarrow{N} \int_0^T \psi_t A(t,t+dt), \quad \text{in } \mathbb{L}^\ell, \qquad (4.15)
$$

which justifies the fact that such an integral is called $\mathbb{L}^\ell$ stochastic non-linear 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:

$$
\begin{aligned}
A(t,t+h) &= A(t,t+h) - \mathbb{E}[A(t,t+h)|\mathcal{F}_t] + \mathbb{E}[A(t,t+h)|\mathcal{F}_t]\\
&:= M(t,t+h) + R(t,t+h), \qquad (4.16)
\end{aligned}
$$

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

Theorem 14 (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 $\pi(\Delta) < 1$,

$$
\|S(\Delta) - S(\Delta')\|_{\mathbb{L}^\ell} \le C \max\{T^{1/\alpha}, T\}(\pi(\Delta))^\eta, \qquad (4.17)
$$

where $\pi(\Delta)$ denotes the step size of the subdivision Δ and with $\eta = \min\{\varepsilon_0 - \varepsilon_2, \varepsilon_1, \varepsilon_1'\}$.

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 13 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 to 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 14, one has

$$\Big\| \int_t^{t+h} \psi_s A(s, s+ds) - \psi_t A(t, t+h) \Big\|_{\mathbb{L}^\ell} \leq C h^{\frac{1}{\alpha}+\eta}, \tag{4.18}$$

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.

4.3 Building the dynamics: proofs of Theorem 4 and Corollary 5

We here follow Section 4.6 of [DD16]. Let us first emphasize that (4.14) hold in all the cases (i), (ii) and (iii) mentioned above from Proposition 13 and Theorem 2 (the two last estimates are equal to 0, since the process A is additive). Note that in case (iii), the third inequality of (4.14) follows from the very definition of $\mathfrak{f}$ as a conditional expectation in the statement of Proposition 13.

We can thus define the process $\big(\int_0^t \psi_s dX_s\big)_{0\leq t\leq T}$ for any progressively measurable $(\psi_s)_{0\leq s\leq T}$ in $\mathcal{H}_{\mathfrak{q}'}^{(1-1/\alpha)-\varepsilon_2}$ (see (1.18)), $1/\mathfrak{q}' + 1/\mathfrak{q} = 1/\ell$, $1 \leq \mathfrak{q}, \ell < \alpha$ with $\varepsilon_2 < (\theta - 1)/\alpha$. Setting

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

the construction of the stochastic non-linear Young integral sketched above (see as well subsections 4.4 and 4.5 of [DD16]) allows to define as well the processes $\big(\int_0^t \psi_s R(s, s+ds)\big)_{0\leq t\leq T}$ and $\big(\int_0^t \psi_s M(s, s+ds)\big)_{0\leq t\leq T}$ and the following relation holds

$$\Big(\int_0^t \psi_s dX_s\Big)_{0\leq t\leq T} = \Big(\int_0^t \psi_s R(s, s+ds)\Big)_{0\leq t\leq T} + \Big(\int_0^t \psi_s M(s, s+ds)\Big)_{0\leq t\leq T}.$$

Thanks to Proposition 13 we have that, actually

$$\big(\int_0^t \psi_s R(s, s+ds)\big)_{0\leq t\leq T} = \big(\int_0^t \psi_s \mathfrak{f}(s, X_s, ds)\big)_{0\leq t\leq T},$$

so that the r.h.s. is well defined. Also, we have that

$$\big(\int_0^t \psi_s\big(R(s, s+ds) - \mathscr{F}(s, X_s, ds)\big)\big)_{0\leq t\leq T} = \big(\int_0^t \psi_s \mathscr{R}(s, X_s, ds)\big)_{0\leq t\leq T}$$

is well defined and is null since the bound appearing in the increment of the l.h.s. is greater than one. Hence,

$$\big(\int_0^t \psi_s \mathfrak{f}(s, X_s, ds)\big)_{0\leq t\leq T} = \big(\int_0^t \psi_s \mathscr{F}(s, X_s, ds)\big)_{0\leq t\leq T}.$$

On the other hand, we have that $\big(\int_0^t \psi_s M(s,s+ds)\big)_{0\le t\le T}$ is well defined as well and that $\Big(\int_0^t \psi_s\big(M(s,s+ds)-d\mathcal{W}_s\big)\Big)_{0\le t\le T}=\big(\int_0^t \psi_s \hat M(s,s+ds)\big)_{0\le t\le T}$ 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 $\mathbb{L}^{\mathfrak{q}}$ martingale with $\mathfrak{q}$ moment bounded by $C_{\mathfrak{q}}h^{\mathfrak{q}[\big(1+(\theta-1)\big)/\alpha]}$ (super-diffusive regime) 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 matters. In other words, for any $(\psi_s)_{0\le s\le T}$ in $\mathcal{H}^{1-1/\alpha-\varepsilon_2}_{\mathfrak{q}'}$, with $\varepsilon_2<(\theta-1)/\alpha$,

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

5 Weak formulation and further properties of the drift

5.1 Weak solutions

In this part, we mainly prove Theorem 6. Note first that the existence of a weak solution is a consequence of Theorem 4. It thus only remain to prove weak uniqueness for any $d\ge 1$ to prove Theorem 6-(i) (see the corresponding point below) and pathwise uniqueness for $d=1$ to prove Theorem 6-(ii) (see the corresponding point below as well). In any case, we will need to expand a weak solution along the sequence of classical solutions $(u_m)_{m\ge1}$ of the Cauchy problem $\mathscr{C}(F_m,L^\alpha,f,g,T)$, where F_m is a smooth approximation of F in the sense of Remark 3 and for some smooth functions f,g through Itô's formula. We are therefore led to check whenever the stochastic integrals $\big(\int_0^t Du_m(s,Y_s)dY_s\big)_{0\le t\le T}$ can be defined as $\mathbb{L}^1$-stochastic non-linear Young integrals in the sense of Definition 3. This is the purpose of the next two lemmas. In the first one, we prove that one may define a stochastic calculus w.r.t. the weak solution (*i.e.* w.r.t. quantities in (1.20)), proving thus the last assertion of Theorem 6. In the second one, we prove that one can expand the solution of the mollified PDE from Proposition 8 along the weak solution through Itô's formula.

Lemma 15. Assume that the parameters α,p,q,r and γ satisfy a *good relation for the dynamics* (GR-D). Then, for any weak solution Y to (1.1) in the sense of Definition 4, the processes $\big(\int_0^t\psi_s dY_s\big)_{0\le t\le T}$ and $\big(\int_0^t\psi_s\mathscr{F}(s,Y_s,ds)\big)_{0\le t\le T}$ are well defined for any progressively measurable process $(\psi_s)_{0\le s\le T}$ in $\mathcal{H}^{1-1/\alpha-\varepsilon_2}_{\mathfrak{q}'}$ for all $0<\varepsilon_2<(\theta-1)/\alpha$ and $\mathfrak{q}'\in\big([\alpha/(\alpha-1)],\infty\big]$.

Proof. Starting from the very definition (see Definition 4) of a weak solution $(Y,\mathcal{Z})$, we readily get from (4.6) that

$$\forall 1\le\mathfrak{q}<\alpha,\quad \forall 0\le t<t+h\le T,\ \mathbb{E}^{\frac{1}{\mathfrak{q}}}[|\mathbb{E}[\mathscr{F}(t,Y_t,h)|\mathcal{F}_t]|^{\mathfrak{q}}]+\mathbb{E}^{\frac{1}{\mathfrak{q}}}[|\mathscr{F}(t,Y_t,h)|^{\mathfrak{q}}]\le Ch^{1/\alpha+[(\theta-1)/\alpha]}. \quad (5.1)$$

From the construction in Subsection 4.2, this implies in turn that we can define the process $\big(\int_0^t\psi_s dY_s\big)_{0\le t\le T}$ and thus the process $\big(\int_0^t\psi_s\mathscr{F}(s,Y_s,ds)\big)_{0\le t\le T}$ by the methodology of Subsection 4.3, for any progressively measurable process $(\psi_s)_{0\le s\le T}$ in $\mathcal{H}^{1-1/\alpha-\varepsilon_2}_{\mathfrak{q}'}$ for any $0<\varepsilon_2<(\theta-1)/\alpha$ and $\mathfrak{q}'$ such that $1/\mathfrak{q}'+1/\mathfrak{q}=1$ with $1\le\mathfrak{q}<\alpha$. ☐

Lemma 16. Assume that the parameters α,p,q,r and γ satisfy a *good relation for the dynamics* (GR-D) and let $(Y,\mathcal{Z})$ be a weak solutions of (1.1) in the sense of Definition 4. Then,

$$\left(\int_0^t Du_m(s,Y_s)dY_s\right)_{0\le t\le T}\text{ and so }\left(\int_0^t Du_m(s,Y_s)\mathscr{F}(s,Y_s,ds)\right)_{0\le t\le T},$$

where u_m denotes the solution of the Cauchy problem $\mathscr{C}(F_m, L^\alpha, f, g, T)$ with F_m a smooth approximation of F in the sense of Remark 3 and f, g are smooth functions, are well defined as $\mathbb{L}^1$-stochastic non-linear Young integrals in the sense of Definition 3.

Proof. Thanks to the previous lemma it remains to check that there exists ε_2 in $\big(0, [(\theta-1)/\alpha]\big)$ and $\mathfrak{q}'$ in $\big([\alpha/(\alpha-1)], \infty\big]$ such that $(Du_m(s, Y_s))_{s\in[0,T]}$ belongs to $\mathcal{H}^{1-1/\alpha-\varepsilon_2}_{\mathfrak{q}'}$. From (5.1), we deduce

$$\forall 1 \le \mathfrak{q} < \alpha, \quad \forall s \neq t \in [0,T], \quad \mathbb{E}^{\frac{1}{\mathfrak{q}}}[|Y_t - Y_s|^{\mathfrak{q}}] \le C|s-t|^{\frac{1}{\alpha}}. \tag{5.2}$$

The point is now to notice that, from Proposition 8, we have

$$\begin{aligned}\forall \mathfrak{q}' \ge 1,\ \exists C_{\mathfrak{q}'} > 0:\ \forall t \neq s \in [0,T], \quad & |Du_m(s, Y_s) - Du_m(t, Y_t)|^{\mathfrak{q}'} \\ & \le C_{\mathfrak{q}'} \Big\{ (s-t)^{\mathfrak{q}' \frac{\theta-1}{\alpha}} + |Y_t - Y_s|^{\mathfrak{q}'\rho} \Big\},\end{aligned}$$

for any $\rho < \theta - 1$. Set now $\beta := 1 - 1/\alpha - \varepsilon_2$ where we recall that $\varepsilon_2 \in \big(0, [(\theta-1)/\alpha]\big)$. Thus, if

$$\exists \varepsilon_2, \rho/\alpha \in \big(0, [(\theta-1)/\alpha]\big),\ \exists \mathfrak{q}' \in \big([\alpha/(\alpha-1)], \infty\big] \text{ s.t.} \quad \beta < \rho/\alpha,\ (*) \text{ and } \ \mathfrak{q}'\rho < \alpha,\ (**)$$

we can use (5.2) together with previous estimate to obtain

$$\begin{aligned}&\exists C'_{\mathfrak{q}'} > 0: \\ \|Du_m(\cdot, Y)\|_{\mathcal{H}^\beta_{\mathfrak{q}'}} &:= \sup_{t\neq s\in[0,T]} \left\{ \|Du_m(s, Y_s)\|_{\mathbb{L}^{\mathfrak{q}'}(\tilde\Omega)} + \left\| \frac{|Du_m(s, Y_s) - Du_m(t, Y_t)|}{|s-t|^\beta} \right\|_{\mathbb{L}^{\mathfrak{q}'}(\tilde\Omega)} \right\} \\ &\le C'_{\mathfrak{q}'},\end{aligned} \tag{5.3}$$

from which we deduce, thanks to Lemma 15, that both processes

$$\Big(\int_0^t Du_m(s, Y_s) dY_s\Big)_{0\le t\le T} \text{ and } \Big(\int_0^t Du_m(s, Y_s) \mathscr{F}(s, Y_s, ds)\Big)_{0\le t\le T}$$

are well defined. It thus now remains to prove $(*)$ and $(**)$ to conclude the proof.

From the very definition of β, $(*)$ rewrites $\beta - \rho/\alpha < 0 \Leftrightarrow [(\alpha-1)/\alpha] - \varepsilon_2 - \rho/\alpha < 0$. Hence, a sufficient condition for $(*)$ to hold is to prove that there exists $\tilde\rho \in \big(0, 2[(\theta-1)/\alpha]\big)$ such that $[(\alpha-1)/\alpha] - \tilde\rho < 0$ (and thus to choose $\varepsilon_2 = \rho/\alpha = \tilde\rho/2$). Notice now from the very definition of θ in (1.12) and the *good relation for the dynamics* (GR-D) that we have $2[(\theta-1)/\alpha] > [(\alpha-1)/\alpha]$. Hence, for any choice of the parameters satisfying (GR-D), one can find such a $\tilde\rho$. More precisely, for any choice of the parameters satisfying (GR-D), there exists $0 < \varepsilon \ll 1$ such that $\tilde\rho = 2[(\theta-1)/\alpha] - 2\varepsilon$ implies that $(*)$ holds. This allows to choose $\varepsilon_2 = \rho/\alpha = [(\theta-1)/\alpha] - \varepsilon$. It remains to check wether such a choice allows to obtain $\mathfrak{q}' \in \big([\alpha/(\alpha-1)], \infty\big]$ so that $(**)$ holds. Choose $0 < \eta < [(d/p + \alpha/r + \varepsilon)/(\alpha-1)]$ and let $\mathfrak{q}' = [\alpha/(\alpha-1)] + \eta$, then, $(**)$ holds. Indeed, we have $\mathfrak{q}'\rho < \alpha \Leftrightarrow \mathfrak{q}'\rho/\alpha - 1 < 0$ and $\mathfrak{q}'\rho/\alpha - 1 = [(\theta-\alpha)/(\alpha-1)] - \varepsilon[\alpha/(\alpha-1)] + \eta\rho/\alpha$. As $[(\theta-\alpha)/(\alpha-1)] = [(\gamma - 1 - d/p - \alpha/r)/(\alpha-1)] < [(-d/p - \alpha/r)/(\alpha-1)]$ we obtain $\mathfrak{q}'\rho/\alpha - 1 < [(-d/p - \alpha/r)/(\alpha-1)] - \varepsilon[\alpha/(\alpha-1)] + \eta\rho/\alpha < 0$, from the previous choice of η since $\theta < 2, \alpha > 1$ and therefore $\rho/\alpha < 1$. This concludes the claim. □

(i) Weak uniqueness in any dimension: proof of point (i) of Theorem 6. Having this result at hand, one can now expand any weak solution of the *formal* SDE (1.1) along the solution of the Cauchy problem $\mathscr{C}(F_m, L^\alpha, g, f, T)$ through Itô's formula for any smooth f, g to obtain for any t in $[0,T]$,

$$\begin{aligned}u_m(t,Y_t) =& u_m(0,x)+\int_0^t f(s,Y_s)ds+\int_0^t Du_m(s,Y_s)[\mathscr{F}(s,Y_s,ds)-F_m(s,Y_s)ds]\\ &+M_{0,t}(\alpha,u_m,Y)\end{aligned} \tag{5.4}$$

where $M_{0,t}(\alpha,u_m,Y)$, defined by (2.3) (up to the substitution of X^m by Y therein), is a true martingale thanks to Proposition 8. From Proposition 7 and again Proposition 8 one can pass to the (weak) limit in (5.4) to get that for any t in $[0,T]$,

$$u(t,Y_t)-u(0,x)-\int_0^t f(s,Y_s)ds = M_{0,t}(\alpha,u,Y), \tag{5.5}$$

where u is the solution of the Cauchy problem $\mathscr{C}(F,L^\alpha,g,f,T)$ with $g\in\mathcal{C}^1(\mathbb{R}^d,\mathbb{R})$ with $Dg\in\mathbb{B}^{\theta-1}_{\infty,\infty}(\mathbb{R}^d,\mathbb{R}^d)$, where θ is given by (1.12) and $M_{0,t}(\alpha,u,Y)$ is again a true martingale.

Taking now the expectation, one gets that for any t in $[0,T]$,

$$-u(0,x)=-\mathbb{E}[u(t,Y_t)]+\mathbb{E}[\int_0^t f(s,Y_s)ds].$$

Choosing $g\equiv 0$ and $t=T$ we obtain that the left hand side does not depend on the specific choice of Y, so that uniqueness in law follows. This concludes the proof of point (i) in Theorem 6. □

Remark 17. Observe that the right hand side of (5.5) is a $\mathbb{P}$-martingale and one can use usual arguments to build a probability measure on the associated canonical space Ω_α from $\mathbb{P}$ and Y for which finite dimensional marginal coincide. Hence, the Martingale formulation, in the sense of Definition 2, holds. In other words, existence of a weak solution implies the existence of a Martingale solution. Also, as a consequence of the previous arguments, it is plain to check that uniqueness of the Martingale solution implies weak uniqueness.

(ii) Pathwise uniqueness in dimension one: proof of point (ii) of Theorem 6. The aim of this part is to prove Theorem 6-(ii), adapting to this end the proof of Proposition 2.9 in [ABM20] 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 the *formal* SDE (1.1) in the sense of Definition 4. Let also u_m be the solution of the Cauchy problem $\mathscr{C}(F_m,L^\alpha,F_m,0,T)$. Thanks to Lemma 16, one can apply Itô's formula on $(X^i_t-u_m(t,X^i_t))$, $i\in\{1,2\}$ to obtain for any t in $[0,T]$ the two corresponding Itô-Zvonkin transforms

$$\begin{aligned}X^{Z,m,i}_t &:= X^i_t-u_m(t,X^i_t)\\ &= x-u_m(0,x)+\mathcal{W}_t-M_{0,t}(\alpha,u_m,X^i)+R_{0,t}(\alpha,F_m,\mathscr{F},X^i),\ i\in\{1,2\},\end{aligned}$$

where $M_{0,t}(\alpha,u_m,X)$ is as in (2.3) with X instead of X^m therein and $R_{0,t}(\alpha,F_m,\mathscr{F},X):=\int_0^t\mathscr{F}(s,X_s,ds)-F_m(s,X_s)ds$.

We point out that we here use the mollified PDE, keeping therefore the remainder term and dependence in m for the martingale part. Of course, we will have to control the remainders, which is precisely possible from Proposition 7. 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, with

$$V_n:\mathbb{R}\ni x\mapsto\begin{cases}|x|,\ |x|\geq\frac1n,\\ \frac{3}{8n}+\frac34 nx^2-\frac18 n^3x^4,\ |x|\leq\frac1n,\end{cases}$$

the smooth approximation $V_n(X_t^{Z,m,1}-X_t^{Z,m,2})$ of $|X_t^{Z,m,1}-X_t^{Z,m,2}|$. For fixed m,n, thanks to Lemma 16, we can apply Itô's formula to obtain:

$$\begin{aligned}
& V_n(X_t^{Z,m,1}-X_t^{Z,m,2})\\
= \quad & V_n(0)+\int_0^t V_n'(X_s^{Z,m,1}-X_s^{Z,m,2})\big[\mathscr{F}(s,X_s^1,ds)-F_m(s,X_s^1)ds\\
& -(\mathscr{F}(s,X_s^2,ds)-F_m(s,X_s^2)ds)\big]\\
& +\int_0^t\int_{\mathbb{R}\backslash\{0\}}[V_n(X_s^{Z,m,1}-X_s^{Z,m,2}+h_m(s,X_s^1,X_s^2,r))-V_n(X_s^{Z,m,1}-X_s^{Z,m,2})]\\
& \quad\times\tilde N(ds,dr)\\
& +\int_0^t\int_{|r|\geq 1}\psi_n(X_s^{Z,m,1}-X_s^{Z,m,2},h_m(s,X_s^1,X_s^2,r))\nu(dr)ds\\
& +\int_0^t\int_{|r|\leq 1}\psi_n(X_s^{Z,m,1}-X_s^{Z,m,2},h_m(s,X_s^1,X_s^2,r))\nu(dr)ds\\
=: \quad & \frac{3}{8n}+\Delta\mathcal{R}_{0,t}^{m,n}+\Delta M_{0,t}^{m,n}+\Delta C_{0,t,L}^{m,n}+\Delta C_{0,t,S}^{m,n},
\end{aligned} \tag{5.6}$$

recalling that $X_0^{Z,m,1}=X_0^{Z,m,2}$, using the definition of V_n and denoting for all $(s,x_1,x_2,r)\in[0,t]\times\mathbb{R}^3$:

$$\begin{aligned}
h_m(s,x_1,x_2,r) &= u_m(s,x_1+r)-u_m(s,x_1)-[u_m(s,x_2+r)-u_m(s,x_2)], \quad (5.7)\\
\psi_n(x_1,r) &= V_n(x_1+r)-V_n(x_1)-V_n'(x_1)r.
\end{aligned}$$

The point is now to take the expectations in (5.6). 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 Proposition 7 that:

$$\mathbb{E}[|\Delta\mathcal{R}_{0,t}^{m,n}|]\underset{m}{\to}0. \tag{5.8}$$

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

$$\begin{aligned}
\mathbb{E}[|\Delta C_{0,t,L}^{m,n}|] &\leq 2\|V_n'\|_\infty\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}\int_0^t\mathbb{E}[|X_s^1-X_s^2|]ds\\
&\leq C\int_0^t\mathbb{E}[|X_s^1-X_s^2|]ds,
\end{aligned} \tag{5.9}$$

observing that $|h_m(s,x_1,x_2,r)|\leq 2\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}|x_1-x_2|$. Also, from Proposition 9, $\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}\leq C_T\underset{T\to 0}{\longrightarrow}0$ 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

$$\begin{aligned}
|x_1-u_m(t,x_1)-(x_2-u_m(t,x_2))| &\geq |x_1-x_2|-|u_m(t,x_1)-u_m(t,x_2)|\\
&\geq |x_1-x_2|(1-\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)})\\
&\geq \frac{3}{4}|x_1-x_2|.
\end{aligned} \tag{5.10}$$

Hence,

$$|h_m(s,X_s^1,X_s^2,r)|\leq 2\|Du_m\|_{\mathbb{L}^\infty(\mathbb{L}^\infty)}|X_s^1-X_s^2|\leq\frac{2}{3}|X_s^{Z,m,1}-X_s^{Z,m,2}|. \tag{5.11}$$

Therefore, if $|X_s^{Z,m,1} - X_s^{Z,m,2}| \geq 3/n$, we have for any r either that $X_s^{Z,m,1} - X_s^{Z,m,2} + h_m(s, X_s^1, X_s^2, r) \geq 1/n$ if $X_s^{Z,m,1} - X_s^{Z,m,2} \geq 3/n$, or $X_s^{Z,m,1} - X_s^{Z,m,2} + h_m(s, X_s^1, X_s^2, r) \leq -1/n$ if $X_s^{Z,m,1} - X_s^{Z,m,2} \leq -3/n$. It is thus readily seen that $\psi_n(X_s^{Z,m,1} - X_s^{Z,m,2}, h_m(s, X_s^1, X_s^2, r)) = 0$. We thus have:

$$\begin{aligned} &|\mathbb{E}[\Delta C_{0,t,S}^{m,n}]| \\ = \ & \Big|\mathbb{E}[\int_0^t \int_{|r|\leq 1} \mathbb{I}_{|X_s^{Z,m,1} - X_s^{Z,m,2}|\leq \frac{3}{n}} \psi_n(X_s^{Z,m,1} - X_s^{Z,m,2}, h_m(s, X_s^1, X_s^2, r))\nu(dr)ds]\Big| \\ \leq \ & Cn\mathbb{E}[\int_0^t \int_{|r|\leq 1} \mathbb{I}_{|X_s^{Z,m,1} - X_s^{Z,m,2}|\leq \frac{3}{n}} |h_m(s, X_s^1, X_s^2, r)|^2 \nu(dr)ds], \end{aligned} \tag{5.12}$$

using for the last inequality the definition of V_n which gives that there exists C s.t. for all $x, y \in \mathbb{R}$, $|\psi_n(x, y)| \leq Cn|y|^2$. 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 (5.7) and usual interpolation techniques (see e.g. Lemma 5.5 in [ABM20] or Lemma 4.1 in [Pri12]) we get:

$$|h_m(s, X_s^1, X_s^2, r)| \leq \|u_m\|_{\mathbb{L}^\infty(\mathbb{B}^\eta_{\infty,\infty})} |X_s^1 - X_s^2|^{\eta_1} r^{\eta_2},$$

$(\eta_1, \eta_2) \in (0, 1)^2, \eta_1 + \eta_2 = \eta < \theta$. The point is now to apply the above identity with η_1 large enough in order to get rid of the explosive term in (5.12) (i.e. $\eta_1 > 1/2$) and with η_2 sufficiently large in order to guarantee the integrability of the Lévy measure (i.e. $\eta_2 > \alpha/2$). From the very definition of θ, see (1.12), the constraint $1/2 + \alpha/2 < \theta$ is satisfied as soon as $\gamma > [3 - \alpha + 2d/p + 2\alpha/r]/2$, which is precisely the condition ensured when the parameters satisfy a *good relation for the dynamics*, see (GR-D). In such a case, for any choice of the parameters α, p, q, r, γ, one can find $0 < \tilde{\varepsilon} \ll 1$ such that $\eta_1 = 1/2 + \tilde{\varepsilon}/2$, $\eta_2 = \alpha/2 + \tilde{\varepsilon}/2$ and $\eta_1 + \eta_2 < \theta$. Hence,

$$\begin{aligned} |\mathbb{E}[\Delta C_{0,t,S}^{m,n}]| \ &\leq \ Cn\mathbb{E}\left[\int_0^t \int_{|r|\leq 1} \mathbb{I}_{|X_s^{Z,m,1} - X_s^{Z,m,2}|\leq \frac{3}{n}} |X_s^1 - X_s^2|^{1+\tilde{\varepsilon}} r^{\alpha+\tilde{\varepsilon}} \frac{dr}{r^{1+\alpha}} ds\right] \\ &\leq \ Cn\mathbb{E}\left[\int_0^t \mathbb{I}_{|X_s^{Z,m,1} - X_s^{Z,m,2}|\leq \frac{3}{n}} |X_s^{Z,m,1} - X_s^{Z,m,2}|^{1+\tilde{\varepsilon}} ds\right] \\ &\leq \ Cn^{-\tilde{\varepsilon}}, \end{aligned} \tag{5.13}$$

using (5.10) and the definition of $(X^{Z,m,i})_{i\in\{1,2\}}$ for the last but one inequality. Plugging (5.13), (5.9) into (5.6) (taking therein the expectations) and recalling that $\mathbb{E}[\Delta M_{0,t}^{m,n}] = 0$, eventually yields:

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

Passing to the limit, first in m recalling that $\mathbb{E}[|\Delta \mathcal{R}_{0,t}^{m,n}|] \underset{m}{\to} 0$ uniformly in n, gives (from the smoothness properties of $(u_m)_{m\geq 1}$ in Proposition 8, see also point *(ii)* in Section 2.2):

$$\begin{aligned} \mathbb{E}[V_n(X_t^{Z,1} - X_t^{Z,2})] \quad &\leq \quad \frac{3}{8n} + C \int_0^t \mathbb{E}[|X_s^1 - X_s^2|]ds + \frac{C}{n^{\tilde{\varepsilon}}}, \\ X_t^{Z,i} \quad &:= \quad X_t^i - u(t, X_t^i), \ i \in \{1, 2\}. \end{aligned}$$

Take now the limit in n and write from (5.10) (which also holds replacing u_m by u):

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

which readily gives from the Gronwall Lemma $\mathbb{E}[|X_t^1 - X_t^2|] = 0$. This concludes the proof for T small enough. One may then iterate the argument on small time intervals to extend the result for any arbitrary $T > 0$ and then on the whole positive real line. $\square$

5.2 Further properties of the drift

We here give the proof of Proposition 7. For both the martingale and weak solution, proof of point (i) is obvious from the definition of the dynamics. Point (ii) follows from the application of Itô's formula to the sequence of classical solution $(u_m)_{m\geq 0}$ of the Cauchy problem $\mathscr{C}(F_m, L^\alpha, -F_m, 0, T)$, where F_m is a smooth approximation of F in the sense of Remark 3, along the Martingale or the weak solutions, which is licit from Lemma 16 (note that this Lemma only uses the very definition of the dynamics, so that it holds for the Martingale solution as well). We can then conclude this point by passing to the limit in m together with Proposition 9. We now prove points (iii) and (iv).

Let $(F_m)_{m\in\mathbb{N}^*}$ satisfying

$$\lim_{m\to\infty} \|F - F_m\|_{\mathbb{L}^r([0,T],\mathbb{B}^{-1+\gamma}_{p,q}(\mathbb{R}^d))} = 0^5.$$

We aim at proving that, for either the Martingale or the weak solutions one has for all t in $[0,T]$,

$$\lim_{m\to\infty} \left\| \int_0^t \psi_s \mathscr{F}(s, X_s, ds) - \int_0^t \psi_s F_m(s, X_s) ds \right\|_{\mathbb{L}^\ell} = 0, \tag{5.14}$$

with $\ell = 1$ in the case of a weak solution (while not being recalled, this last fact will be implicitly assumed in the following). We want to investigate:

$$\lim_{m\to\infty} \mathbb{E} \left| \int_0^t \psi_s \mathscr{F}(s, X_s, ds) - \int_0^t \psi_s F_m(s, X_s) ds \right|^\ell, \tag{5.15}$$

for any $\psi \in \mathcal{H}^{1-1/\alpha-\varepsilon_2}_{\mathfrak{q}'}$, $\varepsilon_2 \in \big(0, [(\theta-1)/\alpha]\big)$, $\mathfrak{q}' \in \big([\alpha/(\alpha-1)], \infty\big]$. Coming back to the definition of such integrals, this means that we want to control

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

We have the following decomposition:

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

[5] See Remark 3 for the cases when p and/or r are/is $+\infty$.

with the previous notations. Note that $\lim_{m\to\infty}\|S^1_m(\Delta)\|_{\mathbb{L}^\ell}=0$, uniformly w.r.t. Δ and that from estimate (4.6) and by construction (see Subsection 4.2) for each m, $S^1_m(\Delta)$ tends to some S^1_m in $\mathbb{L}^\ell$ as $\pi(\Delta)\to 0$. One can hence swap both limits and therefore deduce that

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

For the second term, we note that, using Minkowski's and then Hölder's inequalities due to the regularity of F_m (using e.g. its $\mathbb{B}^\beta_{\infty,\infty}(\mathbb{B}^\beta_{\infty,\infty})$ norm, for some $\beta\in(0,\alpha)$) that

$$\begin{aligned}
&\mathbb{E}^{1/\ell}\left|\sum_{i=0}^{N-1}\psi_{t_i}\int_{t_i}^{t_{i+1}}ds\int dy[F_m(s,y)-F_m(t_i,X_{t_i})]p_\alpha(s-t_i,y-X_{t_i})\right|^\ell\\
\le\ & C_{m,\beta}\sum_{i=0}^{N-1}(t_{i+1}-t_i)^{(1+\frac{\beta}{\alpha})},
\end{aligned}$$

so that $\lim_{m\to\infty}\lim_{\pi(\Delta)\to 0}\|S^2_m(\Delta)\|_{\mathbb{L}^\ell}=0$. This proves (iii). To prove (iv), it suffices to notice that in such a case, the term $\|S^1_m(\Delta)\|_{\mathbb{L}^\ell}$ defined above is 0. □

A Proof of Lemma 12

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:

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

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

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

From Lemma 11 and similarly to (3.19), we have

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

We thus obtain

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

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

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

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 $(\mathbb{R}^d)^2$:

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

This readily gives, using $\mathbb{L}^1-\mathbb{L}^{p'}$ convolution estimates and (3.19), that

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

Putting together estimates (A.2) and (A.5) into (A.1) yields the estimate (3.12) in Lemma 12.

Remark 18 (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 (3.19), 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 $(x,y,z)\in(\mathbb{R}^d)^2$,

$$\begin{aligned} & |\mathscr{D}^\eta p_\alpha(s-t,z-x)-\mathscr{D}^\eta p_\alpha(s-t,y-x)| \\ \le\ & \frac{C}{(s-t)^{\frac{\beta'+\eta}{\alpha}}}|z-y|^{\beta'}\Big(q_\alpha(s-t,z-x)+q_\alpha(s-t,y-x)\Big). \end{aligned} \tag{A.6}$$

Indeed, (A.6) is direct if $|z-y|\ge[1/2](s-t)^{1/\alpha}$ (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 $\big(|z-y|/(s-t)^{1/\alpha}\big)^{\beta'}\ge 1$. 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 11 (see in particular (3.5)), for all $\lambda\in[0,1]$:

$$\begin{aligned} & |\mathscr{D}^\eta Dp_M(s-t,y-x+\lambda(y-z))| \\ \le\ & \frac{C_m}{(s-t)^{\frac{\eta+1}{\alpha}}}p_{\bar M}(s-t,y-x-\lambda(y-z)) \\ \le\ & \frac{C_m}{(s-t)^{\frac{\eta+1+d}{\alpha}}}\frac{1}{\Big(1+\frac{|y-x-\lambda(z-y)|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m} \\ \le\ & \frac{C_m}{(s-t)^{\frac{\eta+1+d}{\alpha}}}\frac{1}{\Big(\frac12+\frac{|y-x|}{(s-t)^{\frac{1}{\alpha}}}\Big)^m}\le 2\frac{C_m}{(s-t)^{\frac{\eta+1}{\alpha}}}p_{\bar M}(s-t,y-x). \end{aligned} \tag{A.7}$$

Therefore, in the diagonal case (A.6) follows from (A.7) and (3.5) writing $|\mathscr{D}^\eta p_\alpha(s-t,z-x)-\mathscr{D}^\eta p_\alpha(s-t,y-x)|\le\int_0^1 d\lambda|\mathscr{D}^\eta Dp_\alpha(s-t,y-x+\lambda(y-z))\cdot(y-z)|\le 2C_m(s-t)^{-[(\eta+1)/\alpha]}q_\alpha(s-t,y-x)|z-y|\le\tilde C_m(s-t)^{-[(\eta+\beta')/\alpha]}q_\alpha(s-t,y-x)|z-y|^{\beta'}$ for all $\beta'\in(0,1]$ (exploiting again that $|z-y|\le[1/2](s-t)^{1/\alpha}$ for the last inequality). 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

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

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

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

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 $(\mathbb{R}^d)^2$:

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

Repeating the computations from (A.3) to (A.5) and using the above estimate, we obtain that:

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

provided

$$\beta+\gamma>1. \qquad \text{(A.9)}$$

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

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

Setting:

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

it now remains to control this term. Precisely,

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

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

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

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

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

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

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). □

References

[ABLM21] S. Athreya, O. Butkovsky, K. Lê, and L. Mytnik. Well-posedness of stochastic heat equation with distributional drift and skew stochastic heat equation. *arXiv:2011.13498*, 2021.

[ABM20] S. Athreya, O. Butkovsky, and L. Mytnik. Strong existence and uniqueness for stable stochastic differential equations with distributional drift. *Ann. Probab.*, 48–1:178–210, 2020. MR4079434

[AH96] D. R. Adams and L. I. Hedberg. *Function spaces and Potential Theory*. Springer, 1996. MR1411441

[Bas11] R. F. Bass. *Stochastic processes*, volume 33. Cambridge series in statistical and probabilistic mathematics, 2011. MR2856623

[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. MR1867429

[BC03] R. F. Bass and Z.-Q. Chen. Brownian motion with singular drift. *Ann. Probab.*, 31–2:791–817, 2003. MR1964949

[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. MR3444003

[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. MR3785598

[CdRHM21] P-E. Chaudru de Raynal, I. Honoré, and S. Menozzi. Sharp Schauder Estimates for some Degenerate Kolmogorov Equations. *Ann. Sci. Scuola Normale Superiore di Pisa*, 22–3:989–1089, 2021. MR4334312

[CdRHM22] P.-E. Chaudru de Raynal, I. Honoré, and S. Menozzi. Strong regularization by Brownian noise propagating through a weak Hörmander structure. *Prob. Theory and Rel. Fields*, 184, n° 1–2:1–83, 2022. MR4498506

[CdRMP20] P. E. Chaudru de Raynal, S. Menozzi, and E. Priola. Schauder estimates for drifted fractional operators in the supercritical case. *Journal of Functional Analysis*, 278–8:108425, 57p, 2020. MR4056997

[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. MR3505229

[CZZ17] Z.-Q. Chen, X. Zhang, and G. Zhao. Well-posedness of supercritical SDE driven by Lévy processes with irregular drifts. *Trans. Amer. Math. Soc.*, 374, n°11, 7621–7655, 2021. MR4328678

[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. MR3500267

[DGI19] T. De Angelis, M. Germain, and E. Issoglio. A numerical scheme for stochastic differential equations with distributional drift. *Stoch. Proc. and Appl.*, 154, 55–90, 2022. MR4490482

[EK86] S. Ethier and T.G. Kurtz. *Characterization and convergence*. John Wiley & Sons, Inc., New York,, 1986.

[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. MR3581216

[FOT10] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes*. De Gruyter, 2010. MR2778606

[FRW03] F. Flandoli, F. Russo, and J. Wolf. Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.*, 40:493D542, 2003. MR1988703

[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. MR2065168

[HM16] L. Huang and S. Menozzi. A Parametrix Approach for some Degenerate Stable Driven SDEs. *Annales Instit. H. Poincaré*, 52(4):1925–1975, 2016. MR3573301

[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. MR3906169

[HWZ20] Z. Hao, M. Wu, and X. Zhang. Schauder estimates for nonlocal kinetic equations and applications. *J. Math. Pures Appl. (9)*, 140:139–184, 2020. MR4124429

[IR22] E. Issoglio, F. Russo. SDEs with singular coefficients: The martingale problem view and the stochastic dynamics view. *arXiv*.

[Kol00] V. N. Kolokoltsov. Symmetric Stable Laws and Stable-Like Jump-Diffusions. *Proceedings of the London Mathematical Society*, 80(3):725–768, May 2000. MR1744782

[Kol11] V. N. Kolokoltsov. *Markov processes, semigroups and generators*, volume 38. De Gruyter, Studies in Mathematics, 2011. MR2780345

[KP22] H. Kremp and N. Perkowski. Multidimensional SDE with distributional drift and Lévy noise. *Bernoulli*, 28–3, 1757–1783, 2022. MR4411510

[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. MR2117951

[Kur10] T.G. Kurtz. Equivalence of stochastic equations and martingale problems. *In Crisan D. (eds) Stochastic Analysis*, 2010. MR2789081

[LR02] P.-G. Lemarié-Rieusset. *Recent developments in the Navier-Stokes problem*. CRC Press, 2002. MR1938147

[LZ22] C. Ling and G. Zhao. Non-local elliptic equation in Hölder space and the martingale problem. *Journal of Diff. Equations*, 314:653–699, 2022. MR4369182

[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. MR3201992

[Pri12] E. Priola. Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.*, 49–2:421–447, 2012. MR2945756

[Sat99] K. Sato. *Lévy processes and Infinitely divisible Distributions*. Cambridge University Press, 1999. MR1739520

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979. MR0532498

[Szt10] P. Sztonyk. Estimates of tempered stable densities. *J. Theoret. Probab.*, 23(1):127–147, 2010. MR2591907

[Tri83] H. Triebel. *Theory of function spaces, II*. Birkhauser, 1983. MR0781540

[Wat07] T. Watanabe. Asymptotic estimates of multi-dimensional stable densities and their applications. *Transactions of the American Mathematical Society*, 359(6):2851–2879, 2007. MR2286060

[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. MR0336813

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

Acknowledgments. The first author thanks the Centre Henri Lebesgue ANR-11-LABX-0020-01 for creating an attractive mathematical environment. This work was initiated in 2019 and has been funded by the Russian Science Foundation project (project N° 20-11-20119).