



**HAL**  
open science

# Multidimensional Stable Driven McKean-Vlasov SDEs with Distributional Interaction Kernel: Critical Thresholds and Related Models

P.-E Chaudru de Raynal, J.-F Jabir, S Menozzi

► **To cite this version:**

P.-E Chaudru de Raynal, J.-F Jabir, S Menozzi. Multidimensional Stable Driven McKean-Vlasov SDEs with Distributional Interaction Kernel: Critical Thresholds and Related Models. 2023. hal-03995199v2

**HAL Id: hal-03995199**

**<https://hal.science/hal-03995199v2>**

Preprint submitted on 24 Feb 2023

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

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

# MULTIDIMENSIONAL STABLE DRIVEN MCKEAN-VLASOV SDES WITH DISTRIBUTIONAL INTERACTION KERNEL: CRITICAL THRESHOLDS AND RELATED MODELS

P.-E. CHAUDRU DE RAYNAL, J.-F. JABIR, AND S. MENOZZI

ABSTRACT. In this work we continue to investigate well-posedness for stable driven McKean-Vlasov SDEs with distributional interaction kernel following the approach introduced in [8]. We specifically focus on the impact of the Besov smoothness of the initial condition and quantify how it affects the corresponding density estimates for the SDE. In particular, we manage to attain some *critical* thresholds allowing to revisit/address in a stable noise setting some concrete physical and biological models.

**Keywords:** McKean-Vlasov SDEs, distributional interaction kernels, stable processes

**AMS Subject classification (2020):** Primary: 60H10, 60H50; Secondary: 35K67, 35Q84.

## 1. Introduction and main results.

**1.1. Framework.** The present work is a follow-up to the previous paper [8] where we investigated well-posedness results -in a weak and strong sense - alongside the distributional regularity of the McKean-Vlasov SDE:

$$X_s^{t,\mu} = \xi + \int_t^s \int b(r, X_r^{t,\mu} - y) \mu_r^{t,\mu}(dy) dr + (\mathcal{W}_s - \mathcal{W}_t), \quad \mu_s^{t,\mu} = \text{Law}(X_s^{t,\mu}), \quad (1)$$

where the characteristic component  $b$  corresponds to a singular interaction kernel lying in a Lebesgue-Besov space of the form

$$b \in L^r((t, T), B_{p,q}^\beta(\mathbb{R}^d, \mathbb{R}^d)) =: L^r(B_{p,q}^\beta), \quad \beta \in [-1, 0], p, q, r \in [1, +\infty]. \quad (\text{A})$$

We refer to Section 2 for a precise definition of these function spaces and related properties or notations. In our model of interest,  $t$  denotes the initial time of the equation,  $\xi$  the initial condition which will be assumed to be distributed according to a given probability measure  $\mu$ , and  $(\mathcal{W}_s)_{s \geq t}$  is a symmetric non-degenerate  $\alpha$ -stable process with  $\alpha \in (1, 2]$  (see Assumption **(UE)** below in the non Brownian case  $\alpha \in (1, 2)$ ). The time  $T$  (possibly infinite) provides the time horizon on which the kernel  $b$  is known.

A natural question which arises consists in deriving conditions which relate the *stable exponent*  $\alpha$ , the *integrability indexes*  $r, p, q$ , the *regularity index*  $\beta$  and the dimension  $d$  to obtain either weak or strong well-posedness for the McKean-Vlasov SDE (1). In the first work [8], we developed an approach which is valid for any initial probability law being viewed as an element of a suitable Besov space (see Lemma 5 in [8] and **(E<sub>3</sub>)** below).

We basically obtained therein that weak uniqueness holds for (1) for any initial probability law provided

$$1 - \alpha + \frac{d}{p} + \frac{\alpha}{r} < \beta, \quad \beta \in (-1, 0]. \quad (\text{C0})$$

We then need the *strengthened* condition

$$2 - \frac{3}{2}\alpha + \frac{d}{p} + \frac{\alpha}{r} < \beta, \quad \beta \in (-1, 0], \quad (\text{C0s})$$

to guarantee strong well-posedness. Let us point out that in the diffusive case  $\alpha = 2$ , both conditions coincide whereas in the pure jump (strictly stable) case, the condition **(C0s)** is indeed stronger than **(C0)**.

The approach we used in the quoted work to derive those results consisted in considering the Fokker-Planck equation associated with a suitable mollification of the coefficients in (1) and in establishing suitable *a priori* estimates that allowed to then pass to the limit. Importantly, see the introduction of

---

*Date:* February 24, 2023.

[8] and Section 1.3 below, the structure of the non-linearity leads to a quadratic like term in the Fokker-Planck equation. Through convolution estimates in Besov norms (see Lemma 4 in [8]) we used what we called a *dequadrification approach*, which on the one hand allowed to get rid of the afore mentioned quadratic dependence and handle any initial probability law, but on the other hand does not allow to consider the *critical* thresholds that naturally appear in some related physical model. A typical set of parameters for such models (like e.g. the Burgers, the 2D incompressible Navier-Stokes or parabolic-elliptic Keller-Segel equations) would be  $\beta = -1, r = p = \infty$  which is not handled by the previous conditions.

The purpose of the current work is therefore to quantify how smooth the initial law  $\mu$  must be in order to establish weak/strong well-posedness for *critical drifts*, beyond the thresholds set in **(C0)** and **(C0<sub>S</sub>)**, which correspond to concrete but peculiar non-linear models. This leads to truly handle the quadratic dependence which will lead to *natural* conditions like well-posedness in short time or global well-posedness for sufficiently small, in an appropriate Besov norm, initial data. Such features are somehow classical in non-linear analysis (see e.g. [32] for the Navier-Stokes equations). We actually manage to provide a *unified* framework to derive/improve known results for some non-linear models established for  $\alpha = 2$  and to extend them to the pure jump case in a systematic way.

On the other hand, in the special case  $\beta = 0$ , which roughly says there is *no* smoothness, we will also investigate how a regularity gain on the initial condition allows to somehow push forward the well-posedness thresholds of the Krylov-Röckner type condition, see [30] for drifts in time-space Lebesgue spaces when  $\alpha = 2$  and [50], [8] for  $\alpha \in (1, 2]$ , which is precisely given by **(C0)** taking therein  $\beta = 0$ .

**Organization of the paper.** We state our main results in the next section. The strategy of the proof is then briefly recalled in Section 1.3. We state in Section 2 some useful properties on Besov spaces that will be used for the proof of the main results. The framework of Besov interaction kernels allows to revisit the classical non linear martingale problem approach for McKean-Vlasov SDEs in a quite systematic way starting from some global density estimates, which are locally stronger than in [8] whenever the measure  $\mu$  lies in some appropriate Besov space. Section 3 is dedicated to those density estimates focusing on the properties of the associated Fokker-Planck equation, and Section 4 to the derivation of the well-posedness -in a weak and strong sense -results. As a by-product of the density estimates obtained in Section 3, the related propagation of chaos for a suitable particle system could be captured. This will specifically concern a future work. Section 5 is dedicated to the connection of our main results with the concrete physical models mentioned above.

We would also like to mention that after a presentation of X. Zhang at the online seminar “Non-local operators, probability and singularities”, a few days prior to this preprint, we realized that he together with Z. Hao and M. Röckner had a paper in preparation with related results, see [24]. We exchanged the current versions of our works and can now specify some differences between them. In [24], the authors address the (wider) kinetic setting for stable driven McKean-Vlasov SDEs. The approaches to derive and quantify regularization effects are yet rather different, multi-scale Littlewood-Paley analysis in [24] whereas we focus on global duality techniques for Besov spaces. Eventually, we try to mainly relate our approach to the probabilistic literature/results on those equations and to the extensions we can provide. The paper [24] is more connected to PDE results.

**1.2. Main results.** For a point  $z \in \mathbb{R}^d$  we write  $z = \zeta\rho$ ,  $(\zeta, \rho) \in \mathbb{S}^{d-1} \times \mathbb{R}_+$  its polar coordinates where  $\mathbb{S}^{d-1}$  stands for the unit sphere of  $\mathbb{R}^d$ . In the pure jump case  $\alpha \in (1, 2)$  we assume the following condition holds.

**Assumptions (UE).** The Lévy measure  $\nu$  of  $\mathcal{W}$  is given by the decomposition  $\nu(dz) = w(d\zeta)/\rho^{1+\alpha}\mathbb{I}_{\rho>0}$  where  $w$  is a symmetric uniformly non-degenerate measure on  $\mathbb{S}^{d-1}$ . Namely,  $w$  satisfies the condition:

$$\kappa^{-1}|\lambda|^\alpha \leq \int_{\mathbb{S}^{d-1}} |\zeta \cdot \lambda|^\alpha w(d\zeta) \leq \kappa|\lambda|^\alpha, \text{ for all } \lambda \in \mathbb{R}^d,$$

for some  $\kappa \geq 1$ . We point out that this condition allows in particular to consider Lévy measures that have a singular spherical part (like e.g. cylindrical processes).

Our main result reads as follows:

**Theorem 1.** Let  $\mu \in \mathcal{P}(\mathbb{R}^d) \cap B_{p_0, q_0}^{\beta_0}$ . Assume without loss of generality (see equation **(E<sub>3</sub>)** p. 6 below) that:

$$\beta_0 \geq 0, p_0 \in [1, +\infty].$$

For  $\beta \in (-1, 0]$  we assume the following condition holds:

$$1 - \alpha + \frac{\alpha}{r} + \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right) \right]_+ < \beta. \quad (\mathbf{C1})$$

If now  $\beta = -1$ , we assume that

$$\frac{\alpha}{r} < \alpha - 1 \text{ and } 1 - \alpha + \frac{\alpha}{r} + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right) < \beta, \quad \frac{d}{p} < \beta_0 + \frac{d}{p'_0}, \quad \operatorname{div}(b) \in L^r(B_{p, q}^\beta). \quad (\mathbf{C2})$$

Then:

- If **(C1)** or **(C2)** holds, the McKean-Vlasov SDE (1) admits, for any  $S$  strictly smaller than a certain time horizon  $\mathcal{T}_1 = \mathcal{T}_1(\alpha, d, b, \mu) \leq T$ , a weak solution such that its marginal laws  $(\mu_s^{t, \mu})_{s \in [t, S]}$  have a density  $\rho_{t, \mu}(s, \cdot)$  for almost any time  $s \in (t, S]$  and

$$\sup_{s \in (t, S]} (s - t)^\theta |\rho_{t, \mu}(s, \cdot)|_{B_{p', 1}^{-\beta + \vartheta \Gamma}} < +\infty, \quad (2)$$

for  $\vartheta \in (0, 1)$  with

$$\Gamma = (\beta(1 + \mathbb{I}_{\beta \in (-1, 0]}) + \alpha - 1 - \frac{\alpha}{r}) + \left( \beta_0 + \frac{d}{p'_0} - \frac{d}{p} \right), \quad \theta = \left( 1 - \frac{1 - \mathbb{I}_{\beta \in (-1, 0]}}{\alpha} - \frac{1}{r} - \eta \right),$$

and  $\eta > 0$  small enough. The solution is unique among those satisfying the property (2).

- Global solutions. If  $|\mu|_{B_{p_0, q_0}^{\beta_0}} \leq c_0$  for a sufficiently small constant  $c_0 := c_0(\alpha, d, b)$ , small initial data, then  $\mathcal{T}_1 = T$  and

$$\sup_{s \in (t, \mathcal{T}_1)} \left( (s - t)^\theta \wedge 1 \right) |\rho_{t, \mu}(s, \cdot)|_{B_{p', 1}^{-\beta + \vartheta \Gamma}} < +\infty.$$

- Under **(C1)** and if

$$\left( 2 - \frac{3}{2}\alpha + \frac{\alpha}{r} + \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right) \right] \right) \vee \left( 1 - \alpha + \frac{\alpha}{r} + \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right) \right]_+ \right) < \beta, \quad (\mathbf{C1s})$$

or

- Under **(C2)** and if

$$2 - \frac{3}{2}\alpha + \frac{\alpha}{r} + \left[ \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right) \right] < \beta = -1, \quad (\mathbf{C2s})$$

then strong well-posedness further holds.

## Comments.

- The setting **(C1)** provides an alternative regime to **(C0)** whose interest consists in specifically specifying how the additional integrability/smoothness of the initial data impacts the previous bounds when  $\beta \in (-1, 0]$ . We see in **(C1)** and **(C1s)** that a key quantity which appears is what we will call from now on the *intrinsic Besov index*  $\beta_0 + \frac{d}{p'_0}$  of the initial distribution.

Let us first give some details about the case  $\beta = 0$  (Krylov and Röckner type framework) in **(C1)**. Observe importantly that when the intrinsic Besov index is sufficiently large, i.e.  $\frac{d}{p} \leq \beta_0 + \frac{d}{p'_0}$ , then weak and strong uniqueness hold as soon as  $\frac{\alpha}{r} < \alpha - 1$ . On the other hand, when the intrinsic Besov index is not large enough, i.e.  $\frac{d}{p} > \beta_0 + \frac{d}{p'_0}$ , condition **(C1)** precisely quantifies how the non-linearity and the smoothness of the initial data allow to weaken the Krylov and Röckner criterion (which reads from **(C0)** taking  $\beta = 0$ ).

If now  $\beta \in (-1, 0)$ , the equilibrium in **(C1)** for the positive part depends on the positivity of  $-\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right)^1$ . The additional term  $-\beta$  here comes from the strategy of the proof we adopt, through the handling of a quadratic term in the related Fokker-Planck equation, see

<sup>1</sup>rewriting this quantity as  $-\beta + \frac{d}{p'} - \left( \beta_0 + \frac{d}{p'_0} \right)$ , we see that it actually corresponds to the difference of the differential/dimension indexes, with the terminology of [44], [45], associated respectively with the spaces  $B_{p', 1}^{-\beta}$ ,  $B_{p_0, q_0}^{\beta_0}$ .

Lemma 5 below. It indeed seems rather natural since, in order to define properly the non-linear drift in (1) when  $b \in L^r(B_{p,q}^\beta)$ , to have estimates on the law  $\mu^{t,\mu}$  in a function space which can be put in duality, at least for the space variable, with the one of the drift. For technical reasons the chosen space will be  $L^\infty(B_{p',1}^{-\beta})$  (and even a slightly more demanding one concerning the regularity parameter, see e.g. the estimates (2) in Theorem 1). The choice of an  $L^\infty$  space in time is performed to iterate the estimates in time whereas taking 1 for the second Besov integrability index instead of the more natural  $q'$  (standing for the conjugate of  $q$ ) gives more flexibility concerning the product laws in Besov spaces (see Theorem 3 below and again the proof of Lemma 5). Anyhow, the corresponding *intrinsic* Besov index for the law then reads  $-\beta + \frac{d}{p}$ . Thus again, if the intrinsic Besov index of the initial condition is greater than the one associated with the function space in which we will estimate the law of the process then weak uniqueness holds under the sole condition  $\beta > 1 - \alpha + \frac{\alpha}{r}$ . With respect to the former condition **(C0)**, valid for any initial probability law, this precisely means that the smoothness of  $\mu$  makes the spatial integrability condition  $\frac{d}{p}$  unnecessary to have weak uniqueness (at least in small time). When  $-\beta + \frac{d}{p} - (\beta_0 + \frac{d}{p_0}) \geq 0$ , i.e. the *intrinsic* Besov index of the law of the process prevails, the condition for weak uniqueness reads as

$$1 - \alpha + \frac{\alpha}{r} - \beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) < \beta \iff \frac{1}{2} \left[ 1 - \alpha + \frac{\alpha}{r} + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) \right] < \beta.$$

In that case, the threshold for weak existence and uniqueness is relaxed, compared to **(C0)**, provided that

$$\frac{1}{2} \left[ 1 - \alpha + \frac{\alpha}{r} + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) \right] < 1 - \alpha + \frac{\alpha}{r} + \frac{d}{p} \iff \alpha - 1 - \frac{\alpha}{r} - \frac{d}{p} < \left(\beta_0 + \frac{d}{p_0}\right).$$

Going to strong uniqueness in this case we see that

$$\begin{aligned} & \left( 2 - \frac{3}{2}\alpha + \frac{\alpha}{r} + \left[ -\beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) \right] \right) \vee \left( 1 - \alpha + \frac{\alpha}{r} + \left[ -\beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) \right]_+ \right) \\ &= \begin{cases} 2 - \frac{3}{2}\alpha + \frac{\alpha}{r} + \left[ -\beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) \right], \\ \text{if } \left[ -\beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) \right]_+ \neq 0 \text{ or } \alpha < 2\left(1 - \beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right)\right) \\ 1 - \alpha + \frac{\alpha}{r}, \text{ if } \left[ -\beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right) \right]_+ = 0 \text{ and } \alpha \geq 2\left(1 - \beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p_0}\right)\right). \end{cases} \end{aligned}$$

Hence, when the initial condition is regular enough and  $\alpha$  is sufficiently large weak and strong uniqueness are implied by the sole condition  $\beta > 1 - \alpha + \frac{\alpha}{r}$ . In the other cases we see that the condition for strong uniqueness **(C0<sub>S</sub>)** is relaxed as soon as  $-\beta - (\beta_0 + \frac{d}{p_0}) < 0$ . Pay attention that this e.g. never occurs if  $p = \infty$  in the considered subcase.

- Let us eventually turn to  $\beta = -1$ . From the additional, and rather strong structure condition that  $\text{div} b \in L^r(B_{p,q}^\beta)$  we see that the term  $-\beta = 1$  disappears in the left hand side of **(C2)** and **(C2<sub>S</sub>)**. This is precisely because the structure condition allows to perform an integration by parts in the analysis of Lemma 5 which somehow leads for the l.h.s. to the case  $\beta = 0$  under **(C1)**. This assumption is strong but can be verified in many settings, one can e.g. think about fluid dynamics problems which involve divergence free drifts, or the Keller-Segel discussed below. In connection with divergence free drifts we can mention the work [51] by Zhang and Zhao who obtained under this additional condition existence for a *linear* SDE beyond the Krylov and Röckner condition.
- We insist that all the above discussion concerning the relaxation of the former condition **(C0)** is valid in *small* time or for an arbitrary final time provided  $|\mu|_{B_{p_0,q_0}^{\beta_0}}$  is small enough. This is one of the drawbacks of the current approach w.r.t. to the one in [8] which yields well-posedness for any initial condition and fixed final time horizon. However, this highly depends on the current approach, which consists in handling the quadratic term deriving from the related Fokker-Planck equations and can be seen as the price to pay to quantify the *global* impact of a smoother initial condition, which could have only been quantified in small time in [8].

**1.3. Mollified SDE and strategy of the proof.** The principal steps of our procedure mainly follow those of our previous work [8]. We briefly recall it for the sake of completeness. The strategy consists

first in establishing the existence of a solution to (1) in terms of a nonlinear martingale problem, through a mollification of the coefficient and a stability argument (the solutions of the non-linear equations with mollified coefficients form a Cauchy sequence in a suitable function space). This actually allows to obtain the well-posedness of (1) directly from the construction of its time-marginal distributions as solution to the nonlinear Fokker-Planck equation related to (1).

The martingale problem approach to the well-posedness of McKean-Vlasov SDEs has been successfully used over the past to handle a wide range of settings from smooth or "quasi"-smooth to singular interacting kernels. We refer to the papers [40], [38] and [26], and again to [12], [46], [36], [16] - among others - and references therein for more particular cases. This approach has been notably successful to validate numerical particle methods.

For convenience we now introduce for a drift  $b$  satisfying condition **(C1)** or **(C2)** and any measure  $\nu$  for which this is meaningful the notation:

$$\mathcal{B}_\nu(s, \cdot) := b(s, \cdot) \star \nu(\cdot),$$

where  $\star$  denotes the spatial convolution. For all  $\varepsilon > 0$  consider a time-space mollified drift  $b^\varepsilon$ , i.e.  $b^\varepsilon$  is smooth and bounded in time and space (see Proposition 2 below for precise properties related to  $b^\varepsilon$  and the proof of this result in [8]). We now write similarly,

$$\mathcal{B}_\nu^\varepsilon(s, \cdot) := b^\varepsilon(s, \cdot) \star \nu(\cdot), \quad (3)$$

which is well defined for any  $\nu \in \mathcal{P}(\mathbb{R}^d)$  since  $b^\varepsilon$  is smooth and bounded.

The *smoothened* version of (1) is defined by the family of McKean-Vlasov SDEs

$$X_s^{\varepsilon, t, \mu} = \xi + \int_t^s \mathcal{B}_{\mu_r^\varepsilon}^\varepsilon(r, X_r^{\varepsilon, t, \mu}) dr + \mathcal{W}_s - \mathcal{W}_t, \quad \mu_s^{\varepsilon, t, \mu}(s, \cdot) = \text{Law}(X_s^{\varepsilon, t, \mu}), \quad \varepsilon > 0. \quad (4)$$

For every  $\varepsilon > 0$ ,  $\alpha \in (1, 2]$ , the SDE (4) with mollified (i.e. smooth and bounded) interaction kernel  $b^\varepsilon$  admits a unique weak solution whose time marginal distributions  $(\mu_s^{\varepsilon, t, \mu})_{s \in (t, T]}$  are absolutely continuous (see [8] Section 1.3 for details). Namely,

$$\forall A \in \mathcal{B}([t, T]) \otimes \mathcal{B}(\mathbb{R}^d), \quad \mu^{\varepsilon, t, \mu}(A) = \int_A \int_{\mathbb{R}^d} \tilde{\rho}_{t, x, \mu}^\varepsilon(r, y) \mu(dx) dr dy =: \int_A \rho_{t, \mu}^\varepsilon(r, y) dr dy. \quad (5)$$

As a consequence of Itô's formula, the following Duhamel representation holds: for each  $\varepsilon > 0$ ,  $\rho_t^\varepsilon(s, \cdot)$  satisfies for all  $s \in (t, T]$  and all  $(x, y) \in (\mathbb{R}^d)^2$ :

$$\rho_{t, \mu}^\varepsilon(s, y) = p_{s-t}^\alpha \star \mu(y) - \int_t^s dv \left[ \{ \mathcal{B}_{\rho_{t, \mu}^\varepsilon}^\varepsilon(v, \cdot) \rho_{t, \mu}^\varepsilon(v, \cdot) \} \star \nabla p_{s-v}^\alpha \right](y), \quad (6)$$

where  $p^\alpha$  stands for the density of the driving process  $\mathcal{W}$ , and with a slight abuse of notation w.r.t. (3),  $\mathcal{B}_{\rho_{t, \mu}^\varepsilon}^\varepsilon(v, \cdot) = [b^\varepsilon(v, \cdot) \star \rho_{t, \mu}^\varepsilon(v, \cdot)]$ .

Equivalently, see Lemma 3 in [8], for any  $k \geq 1$ ,  $\rho_{t, \mu}^{\varepsilon, k}$  is a *mild* solution of the equation:

$$\begin{cases} \partial_s \rho_{t, \mu}^{\varepsilon, k}(s, y) + \text{div}(\mathcal{B}_{\rho_{t, \mu}^{\varepsilon, k}}^{\varepsilon, k}(s, y) \rho_{t, \mu}^{\varepsilon, k}(s, y)) - L^\alpha \rho_{t, \mu}^{\varepsilon, k}(s, y) = 0, \\ \rho_{t, \mu}^{\varepsilon, k}(t, \cdot) = \mu, \end{cases} \quad (7)$$

where  $L^\alpha$  is the generator of the driving process.

Provided that  $\rho_t^\varepsilon(s, \cdot)$  admits a limit  $\rho_{t, \mu}(s, \cdot)$  in some appropriate function space which precisely allows to take the limit in the Duhamel formulation (6) we derive that the limit satisfies

$$\rho_{t, \mu}(s, y) = p_{s-t}^\alpha \star \mu(y) - \int_t^s dv \left[ \{ \rho_{t, \mu}(v, \cdot) \mathcal{B}_{\rho_{t, \mu}}(v, \cdot) \} \star \nabla p_{s-v}^\alpha \right](y). \quad (8)$$

In other words  $\rho_{t, \mu}(s, y) dy$  is a (mild) solution to the nonlinear Fokker-Planck equation related to (1):

$$\begin{cases} \partial_s \rho_t(s, y) + \text{div}(\rho_{t, \mu}(s, y) \mathcal{B}_{\rho_{t, \mu}}(s, y)) - L^\alpha \rho_{t, \mu}(s, y) = 0 \quad (s, y) \in (t, S] \times \mathbb{R}^d, \\ \rho_t(t, \cdot) = \mu(\cdot). \end{cases} \quad (9)$$

This provides a method to construct a solution to (1) identifying the limit of the martingale problem related to (4). More precisely, from the solution to (4), one can consider the probability measure  $\mathbf{P}_t^\varepsilon$  on the space  $\Omega_\alpha$  (corresponding to the space of càdlàg functions  $\mathbb{D}([t, T]; \mathbb{R}^d)$  if  $\alpha \in (1, 2)$  and to the space

of continuous functions  $\mathcal{C}([t, T]; \mathbb{R}^d)$  if  $\alpha = 2$ ) such that, for  $x(s)$ ,  $t \leq s \leq T$ , the canonical process on  $\Omega_\alpha$ , and for  $\mathbf{P}_t^\varepsilon(s, dx) := \mathbf{P}_t^\varepsilon(x(s) \in dx)$  the family of probability measures induced by  $x(s)$ , we have:  $\mathbf{P}_t^\varepsilon(x(t))$  is equal to  $\mu$  a.s. and for all function  $\phi$  twice continuously differentiable on  $\mathbb{R}^d$ , with bounded derivatives at all order, the process

$$\phi(x(s)) - \phi(x(t)) - \int_t^s \{ \mathcal{B}_{\mathbf{P}_t^\varepsilon}(v, x(v)) \cdot \nabla \phi(x(v)) + L^\alpha(\phi(x(v))) \} dv, \quad t \leq s \leq T,$$

is a martingale. Provided, again, that the time-marginal distributions  $\mathbf{P}_t^\varepsilon(s, dx) = \rho_{t, \mu}^\varepsilon(s, x) dx$  lie in an appropriate space to ensure that  $\mathbf{P}_t^\varepsilon$  is compact in  $\mathcal{P}(\Omega_\alpha)$  any corresponding limit along a converging subsequence defines naturally a solution to the (nonlinear) martingale problem related to (1). From the well-posedness of the limit Fokker-Planck equation one eventually derives uniqueness results for the time-marginal distributions giving in turn the uniqueness of (1). We also emphasize that we here consider the *classical* martingale problem, i.e. the integral of the non-linear drift with singular kernel is well defined (through the estimates established in Section 3.)

## 2. Notation and reminders on some fundamental of Besov spaces.

In this paragraph, we set definitions/notation and remind technical preliminaries - reviewed or directly established in [8] - that will be used throughout the present paper.

From here on, we denote by  $B_{\ell, m}^\gamma$ ,  $\ell, m, \gamma \in \mathbb{R}$  the Besov space with regularity index  $\gamma$  and integrability parameters  $\ell, m$  (see e.g. *et al.* [1], [10], [31] for some related applications and the dedicated monograph [49] by Triebel). We use the thermic characterization for its definition. Namely, denoting by  $\mathcal{S}'(\mathbb{R}^d)$  the dual space of the Schwartz class  $\mathcal{S}(\mathbb{R}^d)$ ,

$$B_{\ell, m}^\gamma = \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : |f|_{B_{\ell, m}^\gamma} := |\mathcal{F}^{-1}(\phi \mathcal{F}(f))| + \mathcal{T}_{\ell, m}^\gamma(f) < \infty \right\},$$

$$\mathcal{T}_{\ell, m}^\gamma(f) := \begin{cases} \left( \int_0^1 \frac{dv}{v} v^{(n-\gamma/\alpha)m} |\partial_v^n \tilde{p}^\alpha(v, \cdot) * f|_{L^\ell}^m \right)^{\frac{1}{m}} & \text{for } 1 \leq m < \infty, \\ \sup_{v \in (0, 1]} \left\{ v^{(n-\gamma/\alpha)} |\partial_v^n \tilde{p}^\alpha(v, \cdot) * f|_{L^\ell} \right\} & \text{for } m = \infty, \end{cases} \quad (10)$$

$n$  being any non-negative integer (strictly) greater than  $\gamma/\alpha$ , the function  $\phi$  being a  $\mathcal{C}_0^\infty$ -function such that  $\phi(0) \neq 0$ , and  $\tilde{p}^\alpha(v, \cdot)$  denoting the density function at time  $v$  of the  $d$ -dimensional isotropic  $\alpha$ -stable process. The thermic characterization of Besov spaces appears rather natural in the current setting since the *a priori* estimates established below are based on Duhamel type representations which themselves involve the density  $p^\alpha$  of the stable driving noise (see again (8) and (6)). The choice of the isotropic stable kernel in (10) follows from the fact that its spatial derivatives enjoy better integrability properties than the heat kernel itself, see e.g. [39]. We now list some properties that we will throughly exploit to establish the density estimates on the solution of (1).

- *Embeddings.*

(i) Between Lebesgue and  $B_{\ell, m}^0$ -spaces:

$$\forall 1 \leq \ell \leq \infty, \quad B_{\ell, 1}^0 \hookrightarrow L^\ell \hookrightarrow B_{\ell, \infty}^0. \quad (\mathbf{E}_1)$$

(ii) Between Besov spaces: For all  $p_0, p_1, q_0, q_1 \in [1, \infty]$  such that  $q_0 \leq q_1$ ,  $p_0 \leq p_1$  and  $s_0 - d/p_0 \geq s_1 - d/p_1$ ,

$$B_{p_0, q_0}^{s_0} \hookrightarrow B_{p_1, q_1}^{s_1}. \quad (\mathbf{E}_2)$$

(iii) For  $\mathcal{P}(\mathbb{R}^d)$ , the space of probability measures on  $\mathbb{R}^d$ ,

$$\mathcal{P}(\mathbb{R}^d) \subset \cap_{\ell \geq 1} B_{\ell, \infty}^{-d/\ell},$$

$$\mathcal{P}(\mathbb{R}^d) \subset \cap_{\ell \geq 1} B_{\ell, m}^{-d/\ell' - \epsilon}, \quad \epsilon > 0, m \in [1, \infty) \text{ where } \ell^{-1} + (\ell')^{-1} = 1. \quad (\mathbf{E}_3)$$

- *Young/Convolution inequality.* Let  $\gamma \in \mathbb{R}$ ,  $\ell$  and  $m$  in  $[1, +\infty]$ . Then for any  $\delta \in \mathbb{R}$ ,  $\ell_1, \ell_2 \in [1, \infty]$  such that  $1 + \ell^{-1} = \ell_1^{-1} + \ell_2^{-1}$  and  $m_1, m_2 \in (0, \infty]$  such that  $m_1^{-1} \geq \max\{m^{-1} - m_2^{-1}, 0\}$

$$|f \star g|_{B_{\ell, m}^\gamma} \leq c_{\mathbf{Y}} |f|_{B_{\ell_1, m_1}^{\gamma-\delta}} |g|_{B_{\ell_2, m_2}^\delta}, \quad (\mathbf{Y})$$

for  $c_{\mathbf{Y}}$  a universal constant depending only on  $d$ .

- *Besov norm of heat kernel* (see [10, Lemma 11]). There exists  $c_{\mathbf{HK}} := C(\alpha, \ell, m, \gamma, d)$  s.t. for all multi-index  $\mathbf{a} \in \mathbb{N}^d$  with  $|\mathbf{a}| \leq 1$ , and  $0 < v < s < \infty$ :

$$|\partial^{\mathbf{a}} p_{s-v}^\alpha|_{B_{\ell, m}^\gamma} \leq \frac{c_{\mathbf{HK}}}{(s-v)^{\frac{\gamma}{\alpha} + \frac{d}{\alpha}(1-\frac{1}{\ell}) + \frac{|\mathbf{a}|}{\alpha}}}. \quad (\mathbf{HK})$$

- *Duality inequality* (see e.g. [31, Proposition 3.6]). For  $\ell, m \in [1, \infty]$ ,  $\gamma \in \mathbb{R}$  and  $(f, g) \in B_{\ell, m}^\gamma \times B_{\ell', m'}^{-\gamma}$ , it holds:

$$\left| \int_{\mathbb{R}^d} f(y)g(y)dy \right| \leq |f|_{B_{\ell, m}^\gamma} |g|_{B_{\ell', m'}^{-\gamma}}. \quad (\mathbf{D})$$

- *Lift operator*. For any  $\gamma \in \mathbb{R}$ ,  $\ell, m \in [1, \infty]$ , there exists  $c_{\mathbf{L}} > 0$  such that

$$|\nabla f|_{B_{\ell, m}^{\gamma-1}} \leq c_{\mathbf{L}} |f|_{B_{\ell, m}^\gamma}. \quad (\mathbf{L})$$

- Smooth approximation of the interaction kernel and associated uniform-control properties:

**Proposition 2.** [8, Proposition 2] *Let  $b \in L^r((t, T], B_{p, q}^\beta)$  and  $\beta \in (-1, 0]$ ,  $1 \leq p, q \leq \infty$ . There exists a sequence of time-space smooth bounded functions  $(b^\varepsilon)_{\varepsilon > 0}$  s.t.*

$$|b - b^\varepsilon|_{L^{\bar{r}}((t, T], B_{p, q}^\beta)} \xrightarrow{\varepsilon \rightarrow 0} 0, \quad \forall \tilde{\beta} < \beta, \quad (11)$$

with  $\bar{r} = r$  if  $r < +\infty$  and for any  $\bar{r} < +\infty$  if  $r = +\infty$ . Moreover, there exists  $c \geq 1$ ,  $\sup_{\varepsilon > 0} |b^\varepsilon|_{L^{\bar{r}}((t, T], B_{p, q}^\beta)} \leq c|b|_{L^{\bar{r}}((t, T], B_{p, q}^\beta)}$ .

If  $p, q, r < +\infty$  it then also holds, see e.g. [31], that

$$|b - b^\varepsilon|_{L^r((t, T], B_{p, q}^\beta)} \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (12)$$

- *Products of Besov spaces and related embeddings*

From the reference [44], we now state some general multiplication principles/paraproduct between elements of Besov spaces of the form  $B_{\ell_1, m_1}^{\gamma_1} \cdot B_{\ell_2, m_2}^{\gamma_2} \hookrightarrow B_{\ell, m}^\gamma$ . Equivalently, for any  $f \in B_{\ell_1, m_1}^{\gamma_1}$ ,  $g \in B_{\ell_2, m_2}^{\gamma_2}$ ,  $f \cdot g$  is an element of  $\mathcal{S}'$  with

$$|f \cdot g|_{B_{\ell, m}^\gamma} \leq c |f|_{B_{\ell_1, m_1}^{\gamma_1}} |g|_{B_{\ell_2, m_2}^{\gamma_2}}$$

for some constant  $c$  independent of  $f$  and  $g$ . These results will be extensively used for estimates on (6) later on. Recall the definition, see [44], Sections 4.2 and 4.3,

$$f \cdot g = \lim_{j \rightarrow \infty} \mathcal{F}^{-1} \left( \phi(2^{-j}\xi) \mathcal{F}(f)(\xi) \right) \times \mathcal{F}^{-1} \left( \phi(2^{-j}\xi) \mathcal{F}(g)(\xi) \right)$$

whenever the limit exists in  $\mathcal{S}'$ . Then, from Theorem 2, p. 177 in [44], the following embeddings hold:

**Theorem 3.** *Let  $\gamma > 0$ ,  $\ell, \ell_1, \ell_2 \in [1, +\infty]$  s.t.  $\frac{1}{\ell} = \frac{1}{\ell_1} + \frac{1}{\ell_2}$ . Then*

$$B_{\ell_1, \infty}^\gamma \cdot B_{\ell_2, 1}^\gamma \hookrightarrow B_{\ell, \infty}^\gamma. \quad (\mathbf{Prod1})$$

In particular there exists  $C$  s.t. for all  $f \in B_{\ell_1, \infty}^\gamma$ ,  $g \in B_{\ell_2, 1}^\gamma$ ,

$$\left| f \cdot g \right|_{B_{\ell, \infty}^\gamma} \leq C \left| f \right|_{B_{\ell_1, \infty}^\gamma} \left| g \right|_{B_{\ell_2, 1}^\gamma}.$$

**Weighted Lebesgue-Besov spaces.** As a preliminary step before stating our main results, we introduce a characteristic class of weighted iterated *Lebesgue-Besov* function spaces. The solutions to the Fokker-Planck equation related to (1) and some associated *a priori* estimates will be sought in those, Bochner type, spaces. Introduce for  $\delta, \ell, m \in [1, +\infty]$ ,  $\gamma, \theta \in \mathbb{R}$ ,  $t \leq S \leq T$ ,

$$L_{w_\theta}^\delta((t, S], B_{\ell, m}^\gamma) := \left\{ f : s \in [t, S] \mapsto f(s, \cdot) \in B_{\ell, m}^\gamma \text{ measurable, s.t. } \int_t^S |f(s, \cdot)|_{B_{\ell, m}^\gamma}^\delta w_\theta(s) ds < +\infty \right\},$$

if  $\delta < +\infty$  and

$$\begin{aligned} & L_{w_\theta}^\infty((t, S], B_{\ell, m}^\gamma) \\ & := \left\{ f : s \in [t, S] \mapsto f(s, \cdot) \in B_{\ell, m}^\gamma \text{ measurable, s.t. } \sup \operatorname{ess}_{s \in (t, S]} \left( w_\theta(s) |f(s, \cdot)|_{B_{\ell, m}^\gamma} \right) < +\infty \right\}, \end{aligned} \quad (13)$$



where the weight function  $w_\theta$  is given by

$$w_\theta(s) = (s - t)^\theta, \quad (\mathbf{W})$$

for some exponent  $\theta \in \mathbb{R}$ .

*Remark 1.* We emphasize that one of the main differences with [8] in the definition of the weighted spaces is that we here consider a weight which involves the initial time  $t$  (or *backward* weight), whereas we previously considered in [8] a weight involving the final time (or *forward* weight). This is mainly due to the fact that we here want to absorb the potentially insufficient smoothing effects of the initial measure in order to address the current critical case. In [8], the forward weights were chosen in order to equilibrate the higher singularities of the gradient of the heat kernel in the Duhamel representation. The approach will be here different since we will not rely exclusively on the heat kernel to absorb the singularities induced by the regularity estimates, but also on the initial condition.

Endowed with the metric

$$|f|_{L_{w_\theta}^\delta((t,S], B_{\ell,m}^\gamma)} = \left( \int_t^S ds |f(s, \cdot)|_{B_{\ell,m}^\gamma}^\delta (s - t)^\theta \right)^{\frac{1}{\Gamma}}, \quad (14)$$

- with the usual modification if  $\delta = +\infty$  -, and recalling that  $(B_{\ell,m}^\gamma, |\cdot|_{B_{\ell,m}^\gamma})$  is a Banach space, the normed space  $(L_{w_\theta}^\delta((t,S], B_{\ell,m}^\gamma, w_\theta), |\cdot|_{L_{w_\theta}^\delta((t,S], B_{\ell,m}^\gamma)})$  is also a Banach space (see e.g. [25, Chapter 1]).

In the case  $\theta = 0$ ,  $L_{w_\theta}^\delta((t,S], B_{\ell,m}^\gamma)$  reduces to  $L^\delta((t,S], B_{\ell,m}^\gamma)$ .

### 3. Estimates on the Fokker-Planck equation (9).

#### 3.1. A priori estimates: a unified approach.

**Proposition 4.** *Let  $\beta \in [-1, 0]$ . Assume that the parameters are such that **(C1)** holds for  $\beta \in (-1, 0]$  (resp. **(C2)** if  $\beta = -1$ ). Then, for any  $(t, \mu)$  in  $[0, T) \times \mathcal{P}(\mathbb{R}^d) \cap B_{p_0, q_0}^{\beta_0}$ , the non-linear Fokker-Planck equation (9) admits a solution which is unique among all the distributional solutions lying in  $L_{w_\theta}^\infty((t, S], B_{p', 1}^{-\beta + \vartheta\Gamma})$ ,  $S < \mathcal{T}_1$  with  $\mathcal{T}_1 \leq T$ ,  $\theta, \Gamma$  as in Theorem 1 and  $\vartheta \in (0, 1)$ .*

*Moreover, for all  $s \in [t, S]$ ,  $\rho_{t,\mu}(s, \cdot)$  belongs to  $\mathcal{P}(\mathbb{R}^d)$ . Eventually, for a.e.  $s$  in  $(t, S]$ ,  $\rho_{t,\mu}(s, \cdot)$  is absolutely continuous w.r.t. the Lebesgue measure and satisfies the Duhamel representation (8).*

Let us sum up the strategy to derive weak and strong well posedness in our non-linear setting under **(C0)** and **(C1)**, **(C2)** respectively. In any case we need to establish *a priori* estimates on the Fokker-Planck equation (7) associated with mollified kernels through its Duhamel representation (6).

1. In [8], under **(C0)**, we used the so-called *de-quadrification* technique. In this case we cannot in some sense benefit from the smoothness of the initial condition since the whole regularity index associated with the Besov norm for which we are estimating the (mollified) density, is felt by the gradient of the stable heat kernel (see as well the related discussion in the proof of Lemma 5 below).
2. In the current work, under **(C1)**-**(C2)**, we use techniques that are more *common* in non-linear analysis and make a quadratic term appear. We will see below that in that case this approach allows to have an extra-margin, what we already called in the comments following Theorem 1 the *intrinsic* regularity index associated with the initial condition  $\mu \in B_{p_0, q_0}^{\beta_0}$ , with respect to the former threshold appearing in **(C0)**. We point out that, for  $\beta = -1$  we anyhow require some additional smoothness properties for the drift to handle this *critical case*.

To derive weak uniqueness, our approach consists in exploiting appropriate estimates on the density so that we can prove that the drift

$$\mathcal{B}_{\rho_{t,\mu}}(s, \cdot) := \int_{\mathbb{R}^d} b(t, y) \rho_{t,\mu}(r, y) dy$$

belongs to the time-space Lebesgue space  $L^\delta - L^\infty$  for an index  $\delta$  which then allows to enter the framework of [10] where, in the linear setting, a parabolic bootstrap result was established for singular drifts.

We insist on the fact that appealing to product rules in Besov spaces precisely allows to derive the highest regularity order, with respect to what could have e.g. been done through easier embeddings, i.e. going

back to Lebesgue spaces and then deteriorating the smoothness indexes. The point is then of course to derive, the maximum regularity index for which the estimates on the equation with mollified coefficients work. We refer to the proof of Lemma 5 below for further details.

From this approach, we will be faced with features which somehow often appear in non-linear analysis, i.e. the results will be valid either in *small time* or for sufficiently small initial data, in the considered setting of Besov spaces.

**About constants.** Introduce the parameter set

$$\Theta := \begin{cases} \{d, \kappa, r, p, q, \beta, |b|_{L^r(B_{p,q}^\beta)}, p_0, q_0, \beta_0, |\mu|_{B_{p_0, q_0}^{\beta_0}}\}, & \text{if } \beta \in (-1, 0], \\ \{d, \kappa, r, p, q, \beta, |b|_{L^r(B_{p,q}^\beta)}, |\operatorname{div}(b)|_{L^r(B_{p,q}^\beta)}, p_0, q_0, \beta_0, |\mu|_{B_{p_0, q_0}^{\beta_0}}\}, & \text{if } \beta = -1. \end{cases} \quad (15)$$

Namely  $\Theta$  gathers the various parameters appearing in Assumption **(UE)** and condition **(C1)** and **(C2)** depending on the considered value of  $\beta$ . In what follows we denote by  $C := C(\Theta)$  a generic constant depending on the parameter set  $\Theta$  that may change from line to line. Other possible dependencies will be explicitly specified.

We first begin with a Lemma giving a control of the Besov norm of the mollified Fokker-Planck equation (7).

**Lemma 5** (A priori estimates on the mollified density). *Assume **(C1)** holds for  $\beta \in (-1, 0]$  (resp. **(C2)** for  $\beta = -1$ ). For*

$$\zeta_0 := \frac{1}{\alpha} \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p_0} \right) \right], \quad (16)$$

set

$$\Gamma = \beta + \alpha - 1 - \frac{\alpha}{r} - \left( -\beta \mathbb{I}_{\beta \in (-1, 0]} + \frac{d}{p} - \left( \beta_0 + \frac{d}{p_0} \right) \right) - \eta = \alpha + \beta \mathbb{I}_{\beta \in (-1, 0]} - 1 - \frac{\alpha}{r} - \alpha \zeta_0 - \eta, \quad (17)$$

with  $\eta > 0$  small enough to have  $\Gamma > 0$ .

Then there exists constants  $c_1, C_1$  depending on  $\Theta$  defined in (15) s.t. for any  $t < S \leq T$ ,  $\varepsilon > 0$ , it holds that

$$\begin{aligned} & \sup_{s \in (t, S]} (s-t)^\theta |\rho_{t, \mu}^\varepsilon(s, \cdot)|_{B_{p', 1}^{-\beta + \Gamma}} \\ & \leq c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}} + C_1 |b|_{L^r(B_{p, q}^\beta)} (S-t)^{\frac{1}{r'} - (\theta + \frac{1 - \beta \mathbb{I}_{\beta \in (-1, 0]}}{\alpha})} \sup_{s \in (t, S]} \left( (s-t)^\theta |\rho_{s, \mu}^\varepsilon|_{B_{p', 1}^{-\beta + \Gamma}} \right)^2, \end{aligned} \quad (18)$$

where

$$\theta = \left( \frac{\Gamma}{\alpha} + \zeta_0 \right) = 1 - \frac{1 - \beta \mathbb{I}_{\beta \in (-1, 0]}}{\alpha} - \frac{1}{r} - \frac{\eta}{\alpha} \quad (19)$$

Note also that for the previous parameters,  $\frac{1}{r'} - (\theta + \frac{1 - \beta \mathbb{I}_{\beta \in (-1, 0]}}{\alpha}) = \frac{\eta}{\alpha} > 0$ .

Recall that under **(C1)**

$$\frac{\alpha}{r} + \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p_0} \right) \right]_+ < \beta + \alpha - 1,$$

so that the parameter  $\theta > 0$  under both conditions **(C1)** and **(C2)**. We point out that if  $\zeta_0 \leq 0$  we could set  $\theta = 0$ , taking  $\Gamma = -\alpha \zeta_0$  instead of the former choice, i.e. in this case one can directly control the  $L^\infty$  norm in time of the spatial Besov norm for a lower regularity index. This occurs if the additional smoothness provided by the initial condition through what we call its *intrinsic* Besov index  $\beta_0 + \frac{d}{p_0}$  is large enough and somehow absorbs the time singularities coming from Besov norms of the heat kernel in the Duhamel formulation. We anyhow chose to obtain the maximum smoothness index in order to derive the best possible thresholds for strong uniqueness, up to a normalizing factor which in the analysis below leads to integrable singularities.

If  $\zeta_0 \leq 0$ , which means that the intrinsic Besov index of the initial condition is not strong enough, w.r.t. to the index  $p'$  of the spatial Besov norm in which we are estimating the density, then a pre-factor in time with index  $\theta > 0$  is needed to control the normalized  $L^\infty$  norm in time. We also point out that we chose here to take the last parameter of the Besov norm equal to one since it allows to use directly the product rule **(Prod1)**. We again point out that the smoothness we manage to derive for the solution of the Fokker-Planck equation involves  $-\beta$ , which is somehow necessary for the non linear drift to be

well-defined, and  $\Gamma$ , which is precisely the *gap* in conditions **(C1)**, **(C2)**. This choice allows to keep integrability properties in the density estimates below.

*Proof.* We start from the Duhamel formulation (6) and apply the norm  $|\cdot|_{B_{p',1}^{-\beta+\Gamma}}$  for  $\Gamma \geq 0$  to be specified later on. For  $s \in (t, T]$ , it follows

$$|\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \leq \left| \mu \star p_{s-t}^\alpha \right|_{B_{p',1}^{-\beta+\Gamma}} + \int_t^s dv \left| \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p',1}^{-\beta+\Gamma}}. \quad (20)$$

We first consider  $\beta \in (-1, 0]$  under **(C1)**. Applying successively **(Y)** (with  $m_1 = 1, m_2 = \infty$ ), **(Prod1)**, **(D)** and finally **(E2)** yields

$$\begin{aligned} & \left| \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq C \left| \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \right|_{B_{p',\infty}^\Gamma} \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta}} \\ & \leq C \left| \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right|_{B_{\infty,\infty}^\Gamma} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^\Gamma} \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta}} \\ & \leq C |b^\varepsilon(v, \cdot)|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',q'}^{-\beta+\Gamma}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^\Gamma} \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta}} \\ & \leq C |b^\varepsilon(v, \cdot)|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}}^2 \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta}}. \end{aligned}$$

We insist that the parameters of the first Young inequality precisely lead to the weakest time singularity for the contribution involving the Besov norm of  $\nabla p_{s-v}^\alpha$  keeping in mind that we also need to make the *homogeneous* contribution  $\left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}}$  appear from the non-linear drift  $\mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot)$ . This seems therefore rather natural. Pay attention that the product rule is here precisely to avoid putting the singularity coming from the smoothness index  $\Gamma$  on the gradient of the heat kernel. Again, it is precisely for the product rule to be valid in any case that we take the second integrability index equal to 1 in the previous computations.

We could have avoided the product rule writing

$$\begin{aligned} & \left| \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq C \left| \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \right|_{B_{p',\infty}^0} \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta+\Gamma}} \\ & \stackrel{\text{(E1)}}{\leq} C \left| \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right|_{B_{\infty,1}^0} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^0} \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta+\Gamma}} \\ & \leq C |b(v, \cdot)|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{-\beta}}^2 \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta+\Gamma}} \\ & \leq C |b(v, \cdot)|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}}^2 \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta+\Gamma}}. \end{aligned}$$

This approach mainly uses the Hölder inequality and the embeddings. Hence, we have the same kind of estimate but with a higher time singularity, see again **(HK)**. Thus, the product rule is to be preferred<sup>2</sup>.

Let us now turn to the initial condition. Applying again **(Y)**,

$$\left| \mu \star p_{s-t}^\alpha \right|_{B_{p',1}^{-\beta+\Gamma}} \leq c_Y |\mu|_{B_{p_0, q_0}^{\beta_0}} \left| p_{s-t}^\alpha \right|_{B_{p(p_0, p'), 1}^{-\beta+\Gamma-\beta_0}},$$

<sup>2</sup>Observe that if the gap  $\Gamma$  is *small* then both approaches are *almost* similar.

where  $[\mathfrak{p}(p_0, p')]^{-1} = 1 + (p')^{-1} - (p_0)^{-1}$ . Finally, the stable-kernel estimate (**HK**) yields:

$$\begin{aligned} & |\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq C |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{\frac{\beta-\Gamma+\beta_0}{\alpha} - \frac{d}{\alpha} (1 - \frac{1}{\mathfrak{p}(p_0, p')})} + C \int_t^s \frac{dv}{(s-v)^{\frac{-\beta+1}{\alpha}}} |b^\varepsilon(v, \cdot)|_{B_{p,q}^\beta} |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}}^2 \\ & \leq C |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{\frac{\beta-\Gamma}{\alpha} - \frac{d}{\alpha p} + \frac{1}{\alpha} (\beta_0 + \frac{d}{p_0})} + C \int_t^s \frac{dv}{(s-v)^{\frac{-\beta+1}{\alpha}}} |b^\varepsilon(v, \cdot)|_{B_{p,q}^\beta} |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}}^2. \end{aligned}$$

Set now as in the statement of the Lemma

$$\zeta_0 := \frac{1}{\alpha} \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p_0} \right) \right].$$

Applying next the  $L^1 : L^r - L^{r'}$ -Hölder inequality in time and from the definition of  $\zeta_0$ , we get:

$$|\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \leq c_1 |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{-\frac{\Gamma}{\alpha} - \zeta_0} + C |b|_{L^r(B_{p,q}^\beta)} \left( \int_t^s \frac{dv}{(s-v)^{\frac{(-\beta+1)r'}{\alpha}}} |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}}^{2r'} \right)^{\frac{1}{r'}}. \quad (21)$$

From this expression we see that two characteristic exponents appear, which should guide us to chose the normalizing condition to be considered. The first one is  $\frac{(-\beta+1)}{\alpha} r'$  for which we assume

$$\frac{(-\beta+1)r'}{\alpha} < 1 \iff \frac{(-\beta+1)}{\alpha} < 1 - \frac{1}{r} \iff \frac{\alpha}{r} < \alpha - 1 + \beta. \quad (22)$$

This condition allows to have an integrable singularity in the previous time integral and readily follows from (**C1**). Observe as well that the critical threshold  $\beta = -1$  is here not attainable with this approach.

The other exponent is  $-\left(\frac{\Gamma}{\alpha} + \zeta_0\right)$  which we are now going to calibrate.

For  $\theta \geq 0$  positive to be specified write:

$$\begin{aligned} & \sup_{s \in (t, S]} \left( (s-t)^\theta |\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \right) \\ & \leq c_1 |\mu|_{B_{p_0,q_0}^{\beta_0}} (S-t)^{\theta - \frac{\Gamma}{\alpha} - \zeta_0} + C |b|_{L^r(B_{p,q}^\beta)} \max_{s \in (t, S]} \left( (s-t)^\theta \left( \int_t^s \frac{dv}{(s-v)^{\frac{1-\beta}{\alpha} r'}} |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}}^{2r'} \right)^{\frac{1}{r'}} \right). \end{aligned}$$

Hence, in order to have a non exploding factor we must in that case choose:

$$\theta \geq \frac{\Gamma}{\alpha} + \zeta_0 \geq 0.$$

On the other hand, for the integral term

$$\begin{aligned} & \int_t^s \frac{dv}{(s-v)^{\frac{1-\beta}{\alpha} r'}} |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}}^{2r'} \\ & = \int_t^s \frac{dv}{(s-v)^{\frac{1-\beta}{\alpha} r'} (v-t)^{2r'\theta}} \left( (v-t)^\theta |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \right)^{2r'} \\ & \leq \sup_{v \in [t, s]} (v-t)^{2\theta r'} |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}}^{2r'} \int_t^s \frac{dv}{(s-v)^{\frac{1-\beta}{\alpha} r'} (v-t)^{2r'\theta}} \\ & = \sup_{v \in [t, s]} (v-t)^{2\theta r'} |\rho_{t,\mu}^\varepsilon(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}}^{2r'} (s-t)^{1 - \frac{1-\beta}{\alpha} r' - 2r'\theta} B \left( 1 - \frac{1-\beta}{\alpha} r', 1 - 2r'\theta \right), \end{aligned}$$

where  $B(\cdot, \cdot)$  stands for the  $\beta$ -function. From the above, it follows that

$$\begin{aligned} & \sup_{s \in (t, S]} (s-t)^\theta |\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq c_1 |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{\theta - \left(\frac{\Gamma}{\alpha} + \zeta_0\right)} + C_1 |b|_{L^r(B_{p,q}^\beta)} (S-t)^{\frac{1}{r'} - \frac{1-\beta}{\alpha} - \theta} \left( \sup_{s \in (t, S]} (s-t)^\theta |\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \right)^2, \end{aligned}$$

where the following condition on  $\theta$  is needed:

$$2r'\theta < 1 \iff \theta < \frac{1}{2r'}$$

for the previous integral to be well defined. Note as well that the condition  $1 > (\frac{1-\beta}{\alpha})r'$  is implied by **(C1)**. For the exponent of the time contribution preceding the quadratic term in the previous inequality to be positive, it is also required that

$$\frac{1}{r'} - \frac{1-\beta}{\alpha} - \theta > 0 \iff \theta < \frac{1}{r'} - \frac{1-\beta}{\alpha}.$$

Thus,

$$\theta < \left( \frac{1}{r'} - \frac{1-\beta}{\alpha} \right) \wedge \frac{1}{2r'} = \left( \frac{1}{r'} - \frac{1-\beta}{\alpha} \right).$$

Observe indeed that

$$\frac{1}{r'} - \frac{1-\beta}{\alpha} \leq \frac{1}{2r'} \iff \frac{1}{2r'} \leq \frac{1-\beta}{\alpha} \iff \alpha(1 - \frac{1}{r}) \leq 2(1-\beta), \quad (23)$$

which is indeed always fulfilled since  $\alpha \in (1, 2]$  and  $r \in [1, +\infty]$ .

To sum up, in order for a  $\theta$  to exist, it is necessary that

$$\zeta_0 = \frac{1}{\alpha} \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p_0} \right) \right] < \left( \frac{1}{r'} - \frac{1-\beta}{\alpha} \right) \iff \frac{\alpha}{r} + -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p_0} \right) < \alpha - 1 + \beta, \quad (24)$$

which is precisely implied by condition **(C1)**.

- If  $\zeta_0 > 0$ , which occurs when the intrinsic Besov index  $(\beta_0 + \frac{d}{p_0})$  is not strong enough to absorb the characteristic index  $-\beta + \frac{d}{p}$  (which reflects the singularity and integrability of the drift), it is natural to take  $\frac{\Gamma}{\alpha}$  corresponding to the *gap* between the right and left hand sides in (24). This yields to consider, for  $\eta > 0$ ,

$$\begin{aligned} \Gamma &= \alpha - 1 + 2\beta - \frac{\alpha}{r} - \left( \frac{d}{p} - \left( \beta_0 + \frac{d}{p_0} \right) \right) - \eta = \alpha - 1 + \beta - \frac{\alpha}{r} - \alpha\zeta_0 - \eta, \\ \theta &= \left( \frac{\Gamma}{\alpha} + \zeta_0 \right) = 1 - \frac{1-\beta}{\alpha} - \frac{1}{r} - \frac{\eta}{\alpha}. \end{aligned}$$

- If  $\zeta_0 \leq 0$ , which occurs when the intrinsic Besov index  $(\beta_0 + \frac{d}{p_0})$  is greater than  $-\beta + \frac{d}{p}$ , the previous choice remains valid but any pair  $\theta \in [0, \frac{1}{r'} - \frac{1-\beta}{\alpha})$  and  $\Gamma = \alpha\theta$  also provide an admissible choice. In particular, taking  $\theta = 0$  induces no normalizing factor in time. This choice will turn out to be useful to derive weak uniqueness.

The lemma is proved under **(C1)**.

Let us now restart from (20) under **(C2)**. We precisely rebalance the gradient through an integration by parts to alleviate the time singularity on the heat kernel. Indeed, it can be seen from (21) and (22) that one cannot directly take  $\beta = -1$  therein. Again we cannot derive estimates from (21) if  $\beta = -1$ . We get

$$\begin{aligned} |\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',1}^{1+\Gamma}} &\leq |\mu \star p_{s-t}^\alpha|_{B_{p',1}^{1+\Gamma}} + \int_t^s dv \left| \left( \operatorname{div} \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right) \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \star p_{s-v}^\alpha \right|_{B_{p',1}^{1+\Gamma}} \\ &\quad + \int_t^s dv \left| \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \nabla \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \star p_{s-v}^\alpha(\cdot) \right|_{B_{p',1}^{1+\Gamma}}. \end{aligned}$$

Applying successively **(Y)** (with  $m_1 = 1, m_2 = \infty$ ), **(Prod1)**, **(D)** and finally **(E2)** yields

$$\begin{aligned} &\left| \operatorname{div} \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right) \rho_{t,\mu}^\varepsilon(v, \cdot) \star p_{s-v}^\alpha \right|_{B_{p',1}^{1+\Gamma}} \\ &\leq C \left| \operatorname{div} \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right) \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',\infty}^\Gamma} \left| p_{s-v}^\alpha \right|_{B_{1,1}^1} \\ &\leq C \left| \operatorname{div} \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right) \right|_{B_{\infty,\infty}^\Gamma} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^\Gamma} \left| p_{s-v}^\alpha \right|_{B_{1,1}^1} \\ &\leq C \left| \operatorname{div}(b^\varepsilon(v, \cdot)) \right|_{B_{p,q}^{-1}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',q'}^{1+\Gamma}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}} \left| p_{s-v}^\alpha \right|_{B_{1,1}^1} \leq C \left| \operatorname{div}(b^\varepsilon(v, \cdot)) \right|_{B_{p,q}^{-1}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}}^2 \left| p_{s-v}^\alpha \right|_{B_{1,1}^1}. \end{aligned}$$

Similarly,

$$\begin{aligned}
& \left| \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \nabla \rho_{t,\mu}^\varepsilon(v, \cdot) \star p_{s-v}^\alpha \right|_{B_{p',1}^{1+\Gamma}} \\
& \leq C \left| \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \nabla \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',\infty}^\Gamma} \left| p_{s-v}^\alpha \right|_{B_{1,1}^1} \\
& \leq C \left| \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right|_{B_{\infty,\infty}^\Gamma} \left| \nabla \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^\Gamma} \left| p_{s-v}^\alpha \right|_{B_{1,1}^1} \\
& \stackrel{(\mathbf{L})}{\leq} C |b^\varepsilon(v, \cdot)|_{B_{p,q}^{-1}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',q'}^{1+\Gamma}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}} \left| p_{s-v}^\alpha \right|_{B_{1,1}^1} \leq C |b^\varepsilon(v, \cdot)|_{B_{p,q}^{-1}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}}^2 \left| p_{s-v}^\alpha \right|_{B_{1,1}^1}.
\end{aligned}$$

Note from the above bounds that then terms  $b^\varepsilon, \operatorname{div}(b^\varepsilon)$  naturally appear with same norm. Also, from **(L)**, estimating  $\left| p_{s-v}^\alpha \right|_{B_{1,1}^1}$  somehow amounts to control  $\left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^0}$ .

We obtain

$$\begin{aligned}
\left| \rho_{t,\mu}^\varepsilon(s, \cdot) \right|_{B_{p',1}^{1+\Gamma}} & \leq c_Y |\mu|_{B_{p_0,q_0}^{\beta_0}} \left| p_{s-t}^\alpha \right|_{B_{\mathbf{p}(p_0,p),1}^{1+\Gamma-\beta_0}} \\
& \quad + C \int_t^s \frac{dv}{(s-v)^{\frac{1}{\alpha}}} \left( |b^\varepsilon(v, \cdot)|_{B_{p,q}^{-1}} + |\operatorname{div}(b^\varepsilon(v, \cdot))|_{B_{p,q}^{-1}} \right) \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}}^2 \\
& \leq c_2 |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{-\frac{(1+\Gamma)+\beta_0}{\alpha} - \frac{d}{\alpha} \left(1 - \frac{1}{\mathbf{p}(p_0,p)}\right)} \\
& \quad + C \int_t^s \frac{dv}{(s-v)^{\frac{1}{\alpha}}} \left( |b^\varepsilon(v, \cdot)|_{B_{p,q}^{-1}} + |\operatorname{div}(b^\varepsilon(v, \cdot))|_{B_{p,q}^{-1}} \right) \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}}^2,
\end{aligned}$$

where  $C = C(d, \alpha, p)$  and  $[\mathbf{p}(p_0, p)]^{-1} = 2 - 1/p - 1/p_0 = 1 + 1/p'_0 - 1/p$ . Recalling the very definition (16) of  $\zeta_0$ , applying the  $L^r - L^{r'}$  Hölder inequality in time in the above equation, we get

$$\begin{aligned}
& \left| \rho_{t,\mu}^\varepsilon(s, \cdot) \right|_{B_{p',1}^{1+\Gamma}} \\
& \leq c_2 |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{-(\zeta_0 + \frac{\Gamma}{\alpha})} + C \left( |b^\varepsilon|_{L^r(B_{p,q}^{-1})} + |\operatorname{div}(b^\varepsilon)|_{L^r(B_{p,q}^{-1})} \right) \left( \int_t^s \frac{dv}{(s-v)^{\frac{r'}{\alpha}}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}}^{2r'} \right)^{\frac{1}{r'}}. \quad (25)
\end{aligned}$$

We now multiply both sides by  $(s-t)^\theta$  to obtain that

$$\begin{aligned}
& (s-t)^\theta \left| \rho_{t,\mu}^\varepsilon(s, \cdot) \right|_{B_{p',1}^{1+\Gamma}} \leq c_2 |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{\theta - (\zeta_0 + \frac{\Gamma}{\alpha})} \\
& \quad + C \left( |b|_{L^r(B_{p,q}^{-1})} + |\operatorname{div}(b)|_{L^r(B_{p,q}^{-1})} \right) \sup_{t \leq v \leq s} \left( (v-t)^\theta \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{p',1}^{1+\Gamma}} \right)^2 (s-t)^\theta \left( \int_t^s \frac{dv}{(s-v)^{\frac{r'}{\alpha}} (v-t)^{2\theta r'}} \right)^{\frac{1}{r'}}. \quad (26)
\end{aligned}$$

As in the proof under **(C1)** for the above integral to exist, one needs that:

$$2\theta r' < 1 \iff \theta < \frac{1}{2r'},$$

which is guaranteed by **(C2)**, since similarly to (23) and using (19),

$$\theta = 1 - \frac{1}{r} - \frac{1}{\alpha} - \frac{\eta}{\alpha} < \frac{1}{r'} - \frac{1}{\alpha} \leq \frac{1}{2r'}.$$

It then follows that for  $\Gamma$  as in the statement, i.e.  $\theta = \zeta_0 + \frac{\Gamma}{\alpha}$ , one gets:

$$\begin{aligned}
& \sup_{s \in (t, S]} (s-t)^\theta \left| \rho_{t,\mu}^\varepsilon(s, \cdot) \right|_{B_{p',1}^{1+\Gamma}} \leq c_1 |\mu|_{B_{p_0,q_0}^{\beta_0}} \\
& \quad + C_1 \left( |b|_{L^r(B_{p,q}^{-1})} + |\operatorname{div}(b)|_{L^r(B_{p,q}^{-1})} \right) (S-t)^{\frac{1}{r'} - \frac{1}{\alpha} - \theta} \left( \sup_{s \in (t, S]} (s-t)^\theta \left| \rho_{t,\mu}^\varepsilon(s, \cdot) \right|_{B_{p',1}^{1+\Gamma}} \right)^2.
\end{aligned}$$

This completes the proof under **(C2)**.  $\square$

*Remark 2* (About a priori estimates without prefactor). Observe that when  $-\zeta_0 \geq 0$ , i.e. when the intrinsic Besov index is large enough, we could as well take  $\Gamma = -\alpha\zeta_0$  in (21). This choice would give a non exploding power of time, i.e. the exponent  $-\frac{1}{\alpha} - \zeta_0 = 0$  and would also allow to take a supremum norm in time of the Besov norms in (21). Namely,

$$\sup_{s \in (t, S]} |\rho_{t, \mu}^\varepsilon(s, \cdot)|_{B_{p', 1}^{-\beta+\Gamma}} \leq c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}} + C |b|_{L^r(B_{p, q}^\beta)} \sup_{v \in (t, S]} |\rho_{t, \mu}^\varepsilon(v, \cdot)|_{B_{p', 1}^{-\beta+\Gamma}}^2 (S-t)^{\frac{1}{r'} - \frac{1}{\alpha}}. \quad (27)$$

This choice however corresponds to a smaller  $\Gamma$  than the one defined in the statement of the Lemma and this latter will be useful to specify sharp conditions leading to strong uniqueness.

*Remark 3* (About the pre-factor). The computations of Lemma 5 somehow illustrate that, up to the control of the quadratic term, we can hope to derive bounds for suitable time renormalization of appropriate spatial norms. Such kind of norms already appeared e.g. for the Kato solutions to the Navier-Stokes equations, see [32]. In the same spirit we could derive more spatial smoothness through integral controls in time but we do not need them for our current purpose. The next lemma specifies how the quadratic non-linearity appearing in the estimate can be handled. This occurs, either in short time for an arbitrary initial condition or for any time, provided the initial condition is small enough in the corresponding Besov norm. These features are somehow standard in non-linear analysis

*Remark 4* (About the proofs under **(C1)** and **(C2)**). We would like to stress from the previous proof that once the integration by parts has been performed under **(C2)** the proof to derive the a priori estimate is somehow similar to what is done under **(C1)** for  $\beta = 0$ , in terms of the time singularities to handle. This is why in what follows we will provide the proofs under **(C1)** since the arguments under **(C2)** are actually rather similar, again up to the indicated integration by parts (see also again the comments in the previous proof under condition **(C2)**).

We first begin with a local in time control for the density.

**Lemma 6** (A priori control through a Gronwall type inequality with quadratic growth). *Under **(C1)** or **(C2)** and for  $\theta, \Gamma$  as in Lemma 5 and Theorem 1, there exists  $\mathcal{T}_1 = \mathcal{T}_1(\theta, \Theta) > t$  such that, for any  $S < \mathcal{T}_1$ ,*

$$\forall t < s \leq S, \quad \forall \varepsilon > 0, \quad |\rho_{t, \mu}^\varepsilon|_{L_{\omega_\theta}^\infty((t, s], B_{p', 1}^{-\beta+\Gamma})} := \sup_{v \in (t, s]} (v-t)^\theta |\rho_{t, \mu}^\varepsilon(v, \cdot)|_{B_{p', 1}^{-\beta+\Gamma}} \leq 2c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}}. \quad (28)$$

Furthermore, if  $|\mu|_{B_{p_0, q_0}^{\beta_0}}$  is small enough (small initial data), then the solution is global, i.e. one can set  $\mathcal{T}_1 = T$  and there exists  $\bar{c} := \bar{c}(\theta, \Theta)$  s.t.

$$\sup_{v \in (t, \mathcal{T}_1]} \left( (v-t)^\theta \wedge 1 \right) |\rho_{t, \mu}^\varepsilon(v, \cdot)|_{B_{p', 1}^\Gamma} \leq \bar{c}.$$

*Proof.* According to Lemma 5, equation (18) the mapping

$$f_t : s \in [t, +\infty) \mapsto f_t(s) := \sup_{v \in (t, s]} (v-t)^\theta |\rho_{t, \mu}^\varepsilon(v, \cdot)|_{B_{p', 1}^{-\beta+\Gamma}}$$

satisfies an inequality of the form:

$$f_t(s) \leq a + c_t(s)(f_t(s))^2, \quad t < s \leq S, \quad (29)$$

where  $a = c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}}$  and  $c_t(s) = C_1 (s-t)^{\frac{1}{r'} - (\theta + \frac{1-\beta_1 \beta \in (-1, 0] ]}{\alpha})} |b|_{L^r(B_{p, q}^\beta)}$ . We recall that under both **(C1)** and **(C2)**,  $\frac{1}{r'} - (\theta + \frac{1-\beta_1 \beta \in (-1, 0] ]}{\alpha}) > 0$  which implies that  $c_t$  is null at  $s = t$ . Also, since the coefficients are smooth, the controls of Lemma 6 in [8] apply and give that  $f_t$  is continuous (bounded convergence theorem) Hence, the first time  $S_{f_t}$ , at which  $f_t$  reaches the value  $2a$  necessarily occurs after  $t$ . Introducing now  $S_{c_t} := \inf\{s \geq t : c_t(s) = (4c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}})^{-1}\}$ , we have for  $s \in [t, \min(S_{f_t}, S_{c_t})]$ :

$$f_t(s) \leq a + c_t(s) f_t(s)^2 \Rightarrow f_t(s) \leq a + \underbrace{c_t(s)}_{\leq (4a)^{-1}} f_t(s) 2c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}} \Rightarrow f_t(s) \leq 2a = 2c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}}.$$

This gives the first part of the claim which is valid for any given initial condition setting  $\mathcal{T}_1 = \min(S_{f_t}, S_{c_t}) \wedge S = T_1$ .

From  $T_1$  one can now repeat the analysis of Lemma 5. Namely, for  $f_{T_1}(s) = \sup_{v \in (T_1, s]} |\rho_{t, \mu}^\varepsilon(v, \cdot)|_{B_{p', 1}^{-\beta+\Gamma}}$  for  $s \in [T_1, T_1 + \delta)$  with  $\delta$  small enough,  $f_{T_1}$  is continuous and satisfies an inequality of the form (see also (27) in Remark 2):

$$f_{T_1}(s) \leq a_{T_1} + c_{T_1}(s)(f_{T_1}(s))^2, \quad T_1 < s \leq T_1 + \delta, \quad (30)$$

with  $a_T = |\rho_{t, \mu}^\varepsilon(T_1, \cdot)|_{B_{p', 1}^{-\beta+\Gamma}}$  and  $c_{T_1}(s) = C_1(s - T_1)^{\frac{1}{r'} - \frac{1}{\alpha}} |b|_{L^r(B_{p, q}^\beta)}$ . From the first part of the lemma it holds that

$$a_{T_1} \leq 2c_1 |\mu|_{B_{p_0, q_0}^{\beta_0}} T_1^{-\theta}.$$

The condition (30) also implies for  $\delta$  small enough

$$f_{T_1}(s) \leq a_{T_1} + \bar{C}(f_{T_1}(s))^2, \quad t < s \leq T, \quad \bar{C} = C_1 |b|_{L^r(B_{p, q}^\beta)}. \quad (31)$$

Let us consider the equation  $\bar{C}x^2 - x + a_{T_1} = 0$  which actually has positive roots provided  $1 > 4a_{T_1}\bar{C} \iff |\mu|_{B_{p_0, q_0}^{\beta_0}} < \frac{1}{4\bar{C}c_1} \frac{T_1^\theta}{2}$ . By continuity,  $f_{T_1}(s)$  stays below  $\underline{x} := \frac{1 - \sqrt{1 - 4a_{T_1}\bar{C}}}{2\bar{C}}$  (smallest positive roots of the quadratic equation) which then allows to iterate and therefore allow to take in this case  $T_1 = +\infty$ . Together with the previous inequality this gives the second statement of the lemma and completes the proof.  $\square$

**Lemma 7** (Convergence of the mollified densities). *Under the assumptions (C1) or (C2), and with the notations of Lemma 5, for any decreasing sequence  $(\varepsilon_k)_{k \geq 1}$  s.t.  $\varepsilon_k \xrightarrow[k]{} 0$ ,  $(\rho_{t, \mu}^{\varepsilon_k})_{k \geq 1}$  is a Cauchy sequence in  $L_{w_\theta}^\infty((t, S], B_{p', 1}^{-\beta+\vartheta\Gamma})$ , whenever  $S$  is (strictly) smaller than  $T_1$  given as in Lemma 6 for  $\vartheta = 1$  if  $p, q, r < +\infty$  and any  $\vartheta \in (0, 1)$  if  $p \vee q \vee r = +\infty$ . In particular, there exists  $\rho_{t, \mu} \in L_{w_\theta}^\infty((t, S], B_{p', 1}^{-\beta+\vartheta\Gamma})$  s.t.*

$$\sup_{s \in (t, S]} (s - t)^\theta |(\rho_{t, \mu}^{\varepsilon_k} - \rho_{t, \mu})(s, \cdot)|_{B_{p', 1}^{-\beta+\vartheta\Gamma}} + \sup_{s \in (t, S]} |(\rho_{t, \mu}^{\varepsilon_k} - \rho_{t, \mu})(s, \cdot)|_{L^1(\mathbb{R}^d)} \xrightarrow[k]{} 0. \quad (32)$$

*Proof.* Fix  $k, j \in \mathbb{N}$  meant to be large. Assume w.l.o.g. that  $k \geq j$ . We have from the Duhamel representation (6)

$$\rho_{t, \mu}^{\varepsilon_k}(s, y) - \rho_{t, \mu}^{\varepsilon_j}(s, y) = - \int_t^s dv \left[ \{ \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t, \mu}^{\varepsilon_j}(v, \cdot) \} \star \nabla p_{s-v}^\alpha \right](y).$$

From Remark 4 we will restrict to (C1) for the proofs.

We first assume  $p, q, r < +\infty$  for simplicity. Applying successively (Y) (with  $m_1 = 1, m_2 = \infty$ ), (Prod1), (D) and finally (E2) yields

$$\begin{aligned} & \left| \left( \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t, \mu}^{\varepsilon_j}(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p', 1}^{-\beta+\Gamma}} \\ & \leq C \left| \left( \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t, \mu}^{\varepsilon_j}(v, \cdot) \right) \right|_{B_{p', \infty}^\Gamma} \left| \nabla p_{s-v}^\alpha \right|_{B_{1, 1}^{-\beta}} \\ & \leq C \left( \left| \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \right|_{B_{\infty, \infty}^\Gamma} \left| \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p', 1}^\Gamma} + \left| \mathcal{B}_{\rho_{t, \mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \right|_{B_{\infty, \infty}^\Gamma} \left| \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) - \rho_{t, \mu}^{\varepsilon_j}(v, \cdot) \right|_{B_{p', 1}^\Gamma} \right) \\ & \quad \times \left| \nabla p_{s-v}^\alpha \right|_{B_{1, 1}^{-\beta}} \\ & \leq C \left( \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|(v, \cdot) \right|_{B_{p, q}^\beta} \left| \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p', 1}^{-\beta+\Gamma}} + |b^{\varepsilon_j}(v, \cdot)|_{B_{p, q}^\beta} \left| (\rho_{t, \mu}^{\varepsilon_k} - \rho_{t, \mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p', 1}^{-\beta+\Gamma}} \right) \left| \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p', 1}^\Gamma} \\ & \quad + |b^{\varepsilon_j}(v, \cdot)|_{B_{p, q}^\beta} \left| \rho_{t, \mu}^{\varepsilon_j}(v, \cdot) \right|_{B_{p', q'}^{-\beta+\Gamma}} \left| (\rho_{t, \mu}^{\varepsilon_k} - \rho_{t, \mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p', 1}^\Gamma} \right) \left| \nabla p_{s-v}^\alpha \right|_{B_{1, 1}^{-\beta}} \\ & \leq C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|(v, \cdot) \right|_{B_{p, q}^\beta} \left| \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p', 1}^{-\beta+\Gamma}}^2 + |b^{\varepsilon_j}(v, \cdot)|_{B_{p, q}^\beta} \left| (\rho_{t, \mu}^{\varepsilon_k} - \rho_{t, \mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p', 1}^{-\beta+\Gamma}} \left| \rho_{t, \mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p', 1}^{-\beta+\Gamma}} \right) \\ & \quad \times \left| \nabla p_{s-v}^\alpha \right|_{B_{1, 1}^{-\beta}}. \end{aligned} \quad (33)$$



Exploiting now Lemma 6 and **(HK)** we get:

$$\begin{aligned} & \left| \left( \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq c_{\mathbf{HK}} \left( |(b^{\varepsilon_k} - b^{\varepsilon_j})(v, \cdot)|_{B_{p,q}^\beta} \left( (v-t)^{-\theta} 2c_1 |\mu|_{B_{p_0,q_0}^{\beta_0}} \right)^2 \right. \\ & \quad \left. + |b^{\varepsilon_j}(v, \cdot)|_{B_{p,q}^\beta} (v-t)^{-\theta} 2c_1 |\mu|_{B_{p_0,q_0}^{\beta_0}} \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}} \right) (s-v)^{-\frac{1-\beta}{\alpha}}. \end{aligned}$$

Hence, we derive from the Hölder inequality that:

$$\begin{aligned} |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} & \leq C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} + |b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} \sup_{v \in (t,s]} (v-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \right) \\ & \quad \times \left( \int_t^s \frac{dv}{(v-t)^{2r'\theta} (s-v)^{r'\frac{1-\beta}{\alpha}}} \right)^{\frac{1}{r'}}. \end{aligned}$$

We recall that, similarly to the proof of Lemma 5, under **(C1)**, we have  $2\theta r' < 1$  and  $r'(1-\beta)/\alpha < 1$ . Thus,

$$\begin{aligned} & |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} + |b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} \sup_{v \in (t,s]} (v-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \right) (s-t)^{\frac{1}{r'} - \frac{1-\beta}{\alpha} - 2\theta}, \end{aligned}$$

and

$$\begin{aligned} & (s-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} + |b|_{L^r(B_{p,q}^\beta)} \sup_{v \in (t,S]} (v-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \right) (S-t)^{\frac{1}{r'} - \frac{1-\beta}{\alpha} - \theta}. \end{aligned}$$

Hence, since  $\frac{1}{r'} - \frac{1-\beta}{\alpha} - \theta > 0$  for  $S$  small enough we get:

$$\sup_{s \in (t,S]} (s-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} \leq C |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} (S-t)^{\frac{1}{r'} - \frac{1-\beta}{\alpha} - \theta}.$$

If  $p, q, r < +\infty$  then, from equation (12) in Proposition 2,  $|b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} \xrightarrow{j,k} 0$  and  $\rho_{t,\mu}^{\varepsilon_k}$  is thus a Cauchy sequence in  $L_{w_\theta}^\infty((t, S], B_{p',1}^{-\beta+\Gamma})$ .

Let us now turn to the case  $p \vee q \vee r = +\infty$  then one needs to modify (33) in the following way. For any  $\vartheta \in (0, 1)$  write (using again **(Y)**, **(Prod1)**, **(D)** and **(E2)**):

$$\begin{aligned} & \left| \left( \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \\ & \leq C \left( \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{B_{p,q}^{\beta+\Gamma(\vartheta-1)}} | \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) |_{B_{p',1}^{-\beta+\vartheta\Gamma}} \right. \right. \\ & \quad \left. \left. + |b^{\varepsilon_j}(v, \cdot)|_{B_{p,q}^\beta} \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \right| \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \\ & \quad \left. + |b^{\varepsilon_j}(v, \cdot)|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right|_{B_{p',q'}^{-\beta+\Gamma}} \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \right) \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta}}. \end{aligned} \quad (34)$$

We then proceed through Hölder's inequality, using as well Lemma 6, as above to derive:

$$\begin{aligned} & (s-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(s, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \\ & \leq C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^{\bar{r}}(B_{p,q}^{\beta+\Gamma(\vartheta-1)})} (S-t)^{\frac{1}{\bar{r}} - \frac{1-\beta}{\alpha} - \theta} \right. \\ & \quad \left. + |b|_{L^r(B_{p,q}^\beta)} \sup_{v \in (t,S]} (v-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} (S-t)^{\frac{1}{\bar{r}} - \frac{1-\beta}{\alpha} - \theta} \right), \end{aligned}$$

with  $\bar{r} = r$  if  $r < +\infty$  and any  $\bar{r} < +\infty$  otherwise. In the above computations, we precisely need to have a negative exponent  $\Gamma(\vartheta-1)$  in order to invoke (11) in Proposition 2, which allows to reproduce the previous arguments in the current case.

We have thus established that  $(\rho_{t,\mu}^{\varepsilon_k})_k$  is a Cauchy sequence in  $L_{\mathbb{w}\theta}^\infty((t, S], B_{p',1}^{-\beta+\theta\Gamma})$  under **(C1)**. Let us turn to the  $L^1$  norm. We will here only assume for simplicity that  $p, q, r < +\infty$ , the modifications needed otherwise are similar to the previous ones. From the embedding **(E1)** we will actually focus on the  $B_{1,1}^0$  norm of the difference and get similarly to (33):

$$\begin{aligned}
& \left| \left( \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{1,1}^0} \\
\stackrel{(\overline{\mathbf{Y}})}{\leq} & C \left| \left( \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right) \right|_{B_{1,\infty}^0} \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^0} \\
\stackrel{(\mathbf{E}_1)}{\leq} & C \left( \left| \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \right|_{L^\infty} \underbrace{\left| \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) \right|_{L^1}}_{=1} + \left| \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \right|_{L^\infty} \left| \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) - \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right|_{B_{1,1}^0} \right) \\
& \times \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^0} \\
\stackrel{(\mathbf{D}), (\mathbf{E}_2)}{\leq} & C \left( \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|(v, \cdot) \right|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}} + |b^{\varepsilon_j}(v, \cdot)|_{B_{p,q}^\beta} \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}} \right. \\
& \left. + |b^{\varepsilon_j}(v, \cdot)|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}} \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{1,1}^0} \right) \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^0} \\
\stackrel{(\mathbf{HK}), (28)}{\leq} & C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|(v, \cdot) \right|_{B_{p,q}^\beta} + |b^{\varepsilon_j}(v, \cdot)|_{B_{p,q}^\beta} \left[ (v-t)^\theta \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}} \right. \right. \\
& \left. \left. + \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{1,1}^0} \right] \right) \times (v-t)^{-\theta} (s-v)^{-\frac{1}{\alpha}}. \tag{35}
\end{aligned}$$

The fact that we made the  $L^1$  norm of the mollified densities appear spares us a normalization as in the previous computations. Namely, assuming w.l.o.g. that  $S \leq 1$ , we can write from (35) and the Hölder inequality:

$$\begin{aligned}
& |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(s, \cdot)|_{B_{1,1}^0} \\
\leq & C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} + |b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} \left[ \left| \rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j} \right|_{L_{\mathbb{w}\theta}^\infty((t, S], B_{p',1}^{-\beta+\Gamma})} + \sup_{v \in (t, s]} |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot)|_{B_{1,1}^0} \right] \right) \\
& \times \left( \int_t^s \frac{dv}{(v-t)^{r'\theta} (s-v)^{\frac{r'}{\alpha}}} \right)^{\frac{1}{r'}} \\
\leq & C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} + |b|_{L^r(B_{p,q}^\beta)} \left[ \left| \rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j} \right|_{L_{\mathbb{w}\theta}^\infty((t, S], B_{p',1}^{-\beta+\Gamma})} + \sup_{v \in (t, S]} |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot)|_{B_{1,1}^0} \right] \right) \\
& \times (S-t)^{\frac{1}{r'} - (\theta + \frac{1}{\alpha})}.
\end{aligned}$$

Since  $\frac{1}{r'} - (\theta + \frac{1}{\alpha}) > 0$ , we can take the supremum in  $s$  on the above l.h.s. and absorb the corresponding term appearing in the r.h.s. This yields:

$$\sup_{s \in (t, S]} |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(s, \cdot)|_{B_{1,1}^0} \leq C \left( |b^{\varepsilon_k} - b^{\varepsilon_j}|_{L^r(B_{p,q}^\beta)} + \left| \rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j} \right|_{L_{\mathbb{w}\theta}^\infty(B_{p',1}^{-\beta+\Gamma})} \right),$$

which indeed gives, from Proposition 2 and the previous part of the proof, that  $(\rho_{t,\mu}^{\varepsilon_k})_k$  is a Cauchy sequence in  $L^\infty((t, S], L^1)$ . We then derive (32) by completeness. This concludes the proof in the considered case. We already indicated how to modify the previous arguments if  $p \vee q \vee r = +\infty$ .

Under **(C2)**, i.e. for  $\beta = -1$ , let us emphasize that using the structure condition  $\operatorname{div}(b) \in L^r(B_{p,q}^{-1})$ , one could reproduce the previous arguments writing in the case  $p, q, r < +\infty$  instead of (33)

$$\begin{aligned} & \left| \left( \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}^{\varepsilon_k}(v, \cdot) \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_j}}^{\varepsilon_j}(v, \cdot) \rho_{t,\mu}^{\varepsilon_j}(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p',1}^{-\beta+\Gamma}} \\ & \leq C \left( \left[ |b^{\varepsilon_k} - b^{\varepsilon_j}|(v, \cdot) \right]_{B_{p,q}^\beta} + \left| (\operatorname{div}(b^{\varepsilon_k}) - \operatorname{div}(b^{\varepsilon_j}))(v, \cdot) \right|_{B_{p,q}^\beta} \right) \left| \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}}^2 \\ & \quad + \left[ |b^{\varepsilon_j}(v, \cdot)|_{B_{p,q}^\beta} + |\operatorname{div}(b^{\varepsilon_j})(v, \cdot)|_{B_{p,q}^\beta} \right] \left| (\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu}^{\varepsilon_j})(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}} \left| \rho_{t,\mu}^{\varepsilon_k}(v, \cdot) \right|_{B_{p',1}^{-\beta+\Gamma}} \\ & \quad \times \left| p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta}}. \end{aligned}$$

Again, as already indicated in Remark 4, once the integration by parts is performed the analysis involving the singularities corresponds to the previous one for  $\beta = 0$ .  $\square$

Let us now prove that the limit point in the previous lemma is a distributional solution to the Fokker-Planck equation (9) and also satisfies the Duhamel representation (8).

**Lemma 8.** *Assume **(C1)** or **(C2)** is in force and  $S < \mathcal{T}_1$ . Let  $(\varepsilon_k)_{k \geq 1}$  be a decreasing sequence and  $\rho_{t,\mu}$  be the limit point of  $(\rho_{t,\mu}^{\varepsilon_k})_k$  exhibited in Lemma 7, i.e. (32) holds. Then  $\rho_{t,\mu}$  satisfies the Fokker-Planck equation (9) and also enjoys the Duhamel type representation (8). In particular,  $\rho_{t,\mu} \in L_{w_\theta}^\infty(t; S], B_{p',1}^{-\beta+\vartheta\Gamma})$ .*

*Proof.* The claim can be obtained by replicating the arguments of Lemma 9 in [8]. For the sake of completeness, we just draw the essential points of the demonstration, leaving further details to a line-by-line reading of [8]. Let  $\rho_{t,\mu}$  be the cluster point given by Lemma 7.

Starting from weak formulation associated with  $\rho_{t,\mu}^{\varepsilon_k}$ ,  $\rho_{t,\mu}^{\varepsilon_k}$  satisfies, for any  $\varphi \in \mathcal{D}([t, S] \times \mathbb{R}^d)$  and  $k \in \mathbb{N}$ ,

$$\int_t^S \int \rho_{t,\mu}^{\varepsilon_k}(s, x) (\partial_s \varphi + \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}} \cdot \nabla \varphi + L^*(\varphi))(s, x) ds dx = - \int \varphi(t, x) \mu(dx) + \Delta_{\rho_{t,\mu}^{\varepsilon_k}, \rho_{t,\mu}^{\varepsilon_k}}^1 + \Delta_{\rho_{t,\mu}^{\varepsilon_k}, \rho_{t,\mu}^{\varepsilon_k}}^2$$

for

$$\Delta_{\rho_{t,\mu}^{\varepsilon_k}, \rho_{t,\mu}^{\varepsilon_k}}^1 = \int_t^S \int (\rho_{t,\mu}^{\varepsilon_k}(s, x) - \rho_{t,\mu}^{\varepsilon_k}(s, x)) (\partial_s \varphi + L^*(\varphi))(s, x) ds dx$$

and

$$\Delta_{\rho_{t,\mu}^{\varepsilon_k}, \rho_{t,\mu}^{\varepsilon_k}}^2 = \int_t^S \int (\mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}} \rho_{t,\mu}^{\varepsilon_k}(s, x) - \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}^{\varepsilon_k} \rho_{t,\mu}^{\varepsilon_k}(s, x)) \cdot \nabla \varphi(s, x) ds dx.$$

As it is clear that  $(\partial_s - (L^\alpha)^*)\varphi \in L^1([t, S], L^\infty(\mathbb{R}^d))$ , we readily get from (32) that  $|\Delta_{\rho_{t,\mu}^{\varepsilon_k}, \rho_{t,\mu}^{\varepsilon_k}}^1(\varphi)| \xrightarrow[k]{} 0$ .

For the second term  $\Delta_{\rho_{t,\mu}^{\varepsilon_k}, \rho_{t,\mu}^{\varepsilon_k}}^2(\varphi)$ , we simply have to reproduce the computations of the previous Lemma observing that the heat kernel, which induced time singularity is now replaced by a time-space smooth

function. Write indeed,

$$\begin{aligned}
& |\Delta_{\rho_{t,\mu}, \rho_{t,\mu}^{\varepsilon_k}}^2(\varphi)| \\
& \leq \left| \int_t^S \int (\rho_{t,\mu} - \rho_{t,\mu}^{\varepsilon_k})(s, x) (\mathcal{B}_{\rho_{t,\mu}} \cdot \nabla \varphi)(s, x) ds dx \right| \\
& \quad + \left| \int_t^S \int \rho_{t,\mu}^{\varepsilon_k}(s, x) \left( (\mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}} - \mathcal{B}_{\rho_{t,\mu}}) \cdot \nabla \varphi \right)(s, x) ds dx \right| \\
& \leq |\rho_{t,\mu} - \rho_{t,\mu}^{\varepsilon_k}|_{L^\infty((t,S], L^1)} \int_t^S ds |\mathcal{B}_{\rho_{t,\mu}}(s, \cdot) \cdot \nabla \varphi(s, \cdot)|_{L^\infty(\mathbb{R}^d)} \\
& \quad + |\rho_{t,\mu}^{\varepsilon_k}|_{L^\infty((t,S], L^1)} \int_t^S ds |(\mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}} - \mathcal{B}_{\rho_{t,\mu}}) \cdot \nabla \varphi(s, \cdot)|_{L^\infty(\mathbb{R}^d)} \\
& \leq |\nabla \varphi|_{L^\infty} \left( |\rho_{t,\mu} - \rho_{t,\mu}^{\varepsilon_k}|_{L^\infty((t,S], L^1)} \int_t^S ds |b(s, \cdot)|_{B_{p,q}^\beta} |\rho_{t,\mu}(s, \cdot)|_{B_{p',1}^{-\beta}} \right. \\
& \quad \left. + \int_t^S ds \left( |b - b^{\varepsilon_k}(s, \cdot)|_{B_{p,q}^{\beta-\Gamma}} |\rho_{t,\mu}^{\varepsilon_k}(s, \cdot)|_{B_{p',1}^{-\beta+\Gamma}} + |b(s, \cdot)|_{B_{p,q}^\beta} |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu})(s, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \right) \right) \\
& \leq C |\nabla \varphi|_{L^\infty} \left( |\rho_{t,\mu} - \rho_{t,\mu}^{\varepsilon_k}|_{L^\infty((t,S], L^1)} |b|_{L^r(B_{p,q}^\beta)} \left( \int_t^S ds (s-t)^{-\theta r'} \right)^{\frac{1}{r'}} \right. \\
& \quad \left. + |b - b^{\varepsilon_k}|_{L^{\bar{r}}(B_{p,q}^{\beta-\Gamma})} \left( \int_t^S ds (s-t)^{-\theta \bar{r}'} \right)^{\frac{1}{\bar{r}'}} \right. \\
& \quad \left. + |b|_{L^r(B_{p,q}^\beta)} \left( \sup_{s \in (t,S]} (s-t)^\theta |(\rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu})(s, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \right) \left( \int_t^S ds (s-t)^{-\theta r'} \right)^{\frac{1}{r'}} \right)
\end{aligned}$$

using Lemma 6 for the last inequality. We eventually derive from Lemma 7 and Proposition 2 that  $\Delta_{\rho, \rho^{\varepsilon_k}}^2(\varphi) \rightarrow 0$  and we conclude that  $\rho_{t,\mu}$  satisfies (9) in a distributional sense.

The representation (8) could be established similarly. Namely, the point is formally to replace  $\varphi(s, x)$  by  $(s, x) \mapsto p_{s-t}^\alpha(x - y)$  for a given fixed  $y \in \mathbb{R}^d$ . We also refer to Lemmas 3 and 9 of [8] for further details.

The final control stated is also a direct consequence of Lemma 7.  $\square$

**Lemma 9.** *Under the assumptions and with the notations of Lemma 7, the equation (8) admits at most one solution in  $L_{\text{weak}}^\infty((t, S], B_{p',1}^{-\beta+\vartheta\Gamma})$  for any  $S < \mathcal{T}_1$ .*

*Proof.* The proof is here very close to the stability analysis performed in Lemma 7 (see also Lemma 10 in [8] for similar issues). As in the indicated lemma we present the proof under **(C1)**. We refer to Remark 4 and the end of the proof of Lemma 7 for the modifications under **(C2)**.

Assume that  $\rho_{t,\mu}^{(1)}$  and  $\rho_{t,\mu}^{(2)}$  are two possible solutions to (8). Then, for a.e.  $t \leq s \leq S$ ,  $y$  in  $\mathbb{R}^d$ ,

$$\rho_{t,\mu}^{(1)}(s, y) - \rho_{t,\mu}^{(2)}(s, y) = - \int_t^s dv \left[ \mathcal{B}_{\rho_{t,\mu}^{(1)}}(v, \cdot) \rho_{t,\mu}^{(1)}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{(2)}}(v, \cdot) \rho_{t,\mu}^{(2)}(v, \cdot) \right] \star \nabla p_{s-v}^\alpha(y).$$

Similarly to (33)-(34) write

$$\begin{aligned}
& \left| \left( \mathcal{B}_{\rho_{t,\mu}^{(1)}}(v, \cdot) \rho_{t,\mu}^{(1)}(v, \cdot) - \mathcal{B}_{\rho_{t,\mu}^{(2)}}(v, \cdot) \rho_{t,\mu}^{(2)}(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \\
& \leq C \left( |b(v, \cdot)|_{B_{p,q}^\beta} \left| (\rho_{t,\mu}^{(1)} - \rho_{t,\mu}^{(2)})(v, \cdot) \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \left| \rho_{t,\mu}^{(1)}(v, \cdot) \right|_{B_{p',1}^{\vartheta\Gamma}} \right. \\
& \quad \left. + |b(v, \cdot)|_{B_{p,q}^\beta} \left| \rho_{t,\mu}^{(2)}(v, \cdot) \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \left| (\rho_{t,\mu}^{(1)} - \rho_{t,\mu}^{(2)})(v, \cdot) \right|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \right) \left| \nabla p_{s-v}^\alpha \right|_{B_{1,1}^{-\beta}}, \tag{36}
\end{aligned}$$

where  $\vartheta = 1$  if  $p, q, r < +\infty$ ,  $\vartheta \in (0, 1)$  otherwise. From the Hölder inequality and Lemma 6, we thus derive:

$$\begin{aligned} |(\boldsymbol{\rho}_{t,\mu}^{(1)} - \boldsymbol{\rho}_{t,\mu}^{(2)})(s, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} &\leq C|b|_{L^r(B_{p,q}^\beta)} \sup_{v \in (t,s]} (v-t)^\theta |(\boldsymbol{\rho}_{t,\mu}^{(1)} - \boldsymbol{\rho}_{t,\mu}^{(2)})(v, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \\ &\quad \times \left( \int_t^s \frac{dv}{(v-t)^{2r'\theta} (s-v)^{r'\frac{1-\beta}{\alpha}}} \right)^{\frac{1}{r'}} \end{aligned}$$

which in turn yields

$$(s-t)^\theta |(\boldsymbol{\rho}_{t,\mu}^{(1)} - \boldsymbol{\rho}_{t,\mu}^{(2)})(s, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} \leq C|b|_{L^r(B_{p,q}^\beta)} \sup_{v \in (t,S]} (v-t)^\theta |(\boldsymbol{\rho}_{t,\mu}^{(1)} - \boldsymbol{\rho}_{t,\mu}^{(2)})(v, \cdot)|_{B_{p',1}^{-\beta+\vartheta\Gamma}} (S-t)^{\frac{1}{r'} - \frac{1-\beta}{\alpha} - \theta}.$$

Since under **(C1)**,  $\frac{1}{r'} - \frac{1-\beta}{\alpha} - \theta > 0$  and  $S-t$  can be taken small enough, uniqueness follows by taking the supremum in the above l.h.s. and absorbing the r.h.s.  $\square$

#### 4. Well posedness of the non-linear McKean Vlasov SDE. From the Fokker-Planck equation to the non-linear martingale problem.

We will here first focus on the integrability properties of the non-linear drift. Namely, we have the following result:

**Lemma 10.** *Assume that **(C1)** or **(C2)** holds. Then, the mollified non-linear drift  $\mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon$  in (4) is in  $L^{r_0}((t, T], B_{\infty,1}^0)$  with  $r_0 \in (\frac{\alpha}{\alpha - (1-\beta)\mathbb{1}_{\beta \in (-1,0)}}, \frac{r}{1+r\theta})$  and there exists  $C \geq 1$  s.t. for all  $\varepsilon > 0$ :*

$$\forall t \leq S \leq T, |\mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon|_{L^{r_0}((t,S], B_{\infty,1}^0)} \leq C(S-t)^{\frac{1}{r_0} - (\frac{1}{r} + \theta)} |b|_{L^r(B_{p,q}^\beta)}. \quad (37)$$

Importantly, from the definition of  $\theta$  in (19) and the range in which we take  $r_0$ ,  $\Theta := \frac{1}{r_0} - (\frac{1}{r} + \theta) > 0^3$ .

*Proof.* From the Young inequality **(Y)**, one gets for all  $s \in (t, T]$ :

$$|\mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(s, \cdot)|_{B_{\infty,1}^0} \leq c_{\mathbf{Y}} |b^\varepsilon(s, \cdot)|_{B_{p,q}^\beta} |\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',q'}^{-\beta}}.$$

Take now  $r_0$  as indicated, then  $r > r_0$  and use the Hölder inequality,  $L^{r_0} : L^r - L^{(r_0^{-1} - r^{-1})^{-1}}$  (with the usual convention if  $r = \infty$ ), to derive:

$$\begin{aligned} |\mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon|_{L^{r_0}((t,S], B_{\infty,1}^0)} &\leq c_{\mathbf{Y}} |b^\varepsilon|_{L^r(B_{p,q}^\beta)} \left( \int_t^S ds |\rho_{t,\mu}^\varepsilon(s, \cdot)|_{B_{p',q'}^{-\beta}}^{r_0 \frac{r}{r-r_0}} \right)^{\frac{1}{r_0} - \frac{1}{r}} \\ &\leq c_{\mathbf{Y}} |b^\varepsilon|_{L^r(B_{p,q}^\beta)} |\rho_{t,\mu}^\varepsilon|_{L_{w_\theta}^\infty((t,S], B_{p',1}^{-\beta+\vartheta\Gamma})} \left( \int_t^S ds (s-t)^{-\theta(r_0 \frac{r}{r-r_0})} \right)^{\frac{1}{r_0} - \frac{1}{r}} \\ &\leq C|b|_{L^r(B_{p,q}^\beta)} (S-t)^{\frac{1}{r_0} - \frac{1}{r} - \theta}, \end{aligned}$$

using Proposition 2 and Lemma 6 for the last but one inequality.  $\square$

**Existence results.** We first specify the the canonical space introduced in Section 1.3

$$\Omega_\alpha := \begin{cases} \mathcal{C}([t, S]; \mathbb{R}^d), & \alpha = 2, \\ \mathbb{D}([t, S]; \mathbb{R}^d), & \alpha \in (1, 2). \end{cases}$$

A probability measure  $\mathbf{P}$  on the canonical space  $\Omega_\alpha$  solves the non-linear martingale problem related to (1) on  $[t, S]$  if:

- (i)  $\mathbf{P} \circ x(t)^{-1} = \mu$ ;
- (ii) for a.a.  $s \in (t, S]$ ,  $\mathbf{P} \circ x(s)^{-1}$  is absolutely continuous w.r.t. Lebesgue measure and its density belongs to  $L_{w_\theta}^\infty((t, S], B_{p',1}^{-\beta})$ .

<sup>3</sup>In particular  $r_0 < 1/\theta$ .

(iii) for all  $f$  in  $\mathcal{C}^1([t, S], \mathcal{C}_0^2(\mathbb{R}^d))$ , the process

$$\left\{ f(s, x(s)) - f(t, x(t)) - \int_t^s \left( \partial_v f(v, x(v)) + \mathcal{B}_{\mathbf{P}_{\circ x(v)^{-1}}}(v, x(v)) \cdot \nabla f(v, x(v)) + L^\alpha(f)(v, x(v)) \right) dv \right\}_{t \leq s \leq S}, \quad (\text{MP}_{\text{NL}})$$

is a  $\mathbf{P}$  martingale.

We recall that the smoothness properties required on the marginal laws of the canonical process of  $\mathbf{P}$  allow to define almost everywhere the non-linear drift in  $(\text{MP}_{\text{NL}})$ . Anyhow, this latter might still have time singularities, which prevents from using *standard* results to ensure well-posedness.

From now on, for  $\mathbf{P}(x(v) \in dx) := \mathbf{P}_{t,\mu}(v, dx) = \rho_{t,\mu}(v, x)dx$ , we denote with a slight abuse of notation  $\mathcal{B}_{\mathbf{P}_{\circ x(v)^{-1}}}(r, \cdot) = \mathcal{B}_{\rho_{t,\mu}}(v, \cdot)$ .

**Proposition 11.** *Let (C1) or (C2) be in force. Let  $(\mathbf{P}^\varepsilon)_{\varepsilon > 0}$  denote the solution to the non-linear martingale problem related to (4). Then any limit point of a converging subsequence  $(\mathbf{P}^{\varepsilon_k})_k$ ,  $\varepsilon_k \xrightarrow[k]{} 0$ , in  $\mathcal{P}(\Omega_\alpha)$  equipped with its weak topology, solves the non-linear martingale problem related to (1).*

*Proof of Proposition 11.* We prove tightness and then prove any converging subsequence solves the non-linear martingale.

*Tightness.* From the Aldous tightness criterion (see e.g. [5, Theorem 16.10]) if  $\alpha < 2$  or the Kolmogorov one if  $\alpha = 2$  [5, Theorem 7.3], in the current additive noise setting, the tightness of  $(\mathbf{P}^\varepsilon)_{\varepsilon > 0}$  follows from the uniform (w.r.t.  $\varepsilon$ ) almost-sure continuity of  $s \mapsto \int_t^s \mathcal{B}_{\rho_{t,\mu}^\varepsilon}(v, X_v^{\varepsilon,t,\mu})dv$ . Inequality (37) readily implies this property.

*Limit points.* Let  $(\mathbf{P}^{\varepsilon_k})_k$  be a converging subsequence and denote by  $\mathbf{P}$  its limit. Additionally to the weak convergence of  $(\mathbf{P}^{\varepsilon_k})_k$  towards  $\mathbf{P}$ , Lemma 7 also gives that the marginal distributions  $(\mathbf{P}_{t,\mu}^{\varepsilon_k}(s, dx) = \rho_{t,\mu}^{\varepsilon_k}(s, x) dx)_k$  strongly converge towards  $\mathbf{P}_{t,\mu}(s, dx) = \rho_{t,\mu}(s, x) dx$  in  $L_{w_\theta}^\infty((t, S], B_{p',1}^-)$ . Following the proof of Lemma 10, this strong convergence also yields the convergence of  $(\mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}})_k$  towards  $\mathcal{B}_{\rho_{t,\mu}}$  in  $L^{r_0}((t, S], L^\infty)$ . Indeed, for  $r_0$  as in the quoted lemma:

$$\begin{aligned} & \left| \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}} - \mathcal{B}_{\rho_{t,\mu}} \right|_{L^{r_0}((t,S], L^\infty)} \leq C \left| \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}} - \mathcal{B}_{\rho_{t,\mu}} \right|_{L^{r_0}((t,S], B_{\infty,1}^0)} \\ & \leq C \left( \left| \rho_{t,\mu}^{\varepsilon_k} \right|_{L_{w_\theta}^\infty((t,S]B_{p',1}^{-\beta+\vartheta\Gamma})} |b^{\varepsilon_k} - b|_{L^{\check{r}}(B_{p,q}^{\beta-\vartheta\Gamma})}(T-t)^{\check{\Theta}} \right. \\ & \quad \left. + |b|_{L^{\check{r}}(B_{p,q}^\beta)} \left| \rho_{t,\mu}^{\varepsilon_k} - \rho_{t,\mu} \right|_{L_{w_\theta}^\infty((t,S], B_{p',1}^{-\beta+\vartheta\Gamma})}(T-t)^\Theta \right) \xrightarrow[k]{} 0, \end{aligned}$$

with  $\check{r} = r$ ,  $\check{\Theta} = \Theta$  if  $r < +\infty$  and any finite  $\check{r}$  large enough if  $r = +\infty$  and  $\check{\Theta} = \frac{1}{r_0} - \frac{1}{\check{r}} - \theta$ ,  $\check{r} > r_0$  in that case. Convergence now follows from Proposition 2 and Lemmas 6 and 7.

As a direct consequence of the above bound, we get that for all  $\Phi \in C_0^\infty((t, T) \times \mathbb{R}^d, \mathbb{R}^d)$ ,

$$\lim_{k \rightarrow \infty} \int_t^T \int \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}(s, x) \cdot \Phi(s, x) dx ds = \int_t^T \int \mathcal{B}_{\rho_{t,\mu}}(s, x) \cdot \Phi(s, x) dx ds,$$

which in turn is sufficient to ensure that (e.g. [15, Lemma 5.1]): for any  $0 \leq t_1 \leq \dots \leq t_i \leq \dots \leq t_n \leq t \leq s \leq T$ ,  $\Psi_1, \dots, \Psi_n$  continuous bounded, and for any  $\phi$  of class  $\mathcal{C}_0^2(\mathbb{R}^d, \mathbb{R})$ ,

$$\mathbb{E}_{\mathbf{P}^{\varepsilon_k}} \left[ \prod_{i=1}^n \Psi_i(x(t_i)) \int_t^s \mathcal{B}_{\rho_{t,\mu}^{\varepsilon_k}}(v, x(v)) \cdot \nabla \phi(x(v)) dv \right] \rightarrow_k \mathbb{E}_{\mathbf{P}} \left[ \prod_{i=1}^n \Psi_i(x(t_i)) \int_t^s \mathcal{B}_{\rho_{t,\mu}}(v, x(v)) \cdot \nabla \phi(x(v)) dv \right].$$

This exactly means that  $\mathbf{P}$  solves the non-linear martingale problem related to (1).  $\square$

### Weak uniqueness results.

**Proposition 12** (Uniqueness result). *Under the assumption (C1) or (C2), for any  $\mu$  in  $\mathcal{P}(\mathbb{R}^d) \cap B_{p_0, q_0}^{\beta_0}(\mathbb{R}^d)$  and  $0 \leq t < S < \mathcal{T}_1$ , with  $\mathcal{T}_1$  as in Theorem 1, the SDE (1) admits at most one weak solution s.t. its marginal laws  $(\mu_s^{t,\mu}(\cdot))_{s \in [t, T]}$  have a density for a.e.  $s$  in  $(t, S]$ , i.e.  $\mu_s^{t,\mu}(dx) = \rho_{t,\mu}(s, x)dx$  and  $\rho_{t,\mu} \in L_{w_\theta}^\infty((t, S], B_{p',1}^{-\beta+\vartheta\Gamma})$ ,  $\vartheta, \Gamma$  as in Lemmas 7, 5 respectively.*

*Proof.* Following Proposition 13 in [8] it actually suffices to check that  $\mathcal{B}$  viewed as the drift of a *linear* SDE belongs to  $L^\delta(L^\infty)$  with  $1 \leq \delta \leq \infty$  and

$$\frac{\alpha}{\delta} < \alpha - 1.$$

This is exactly Lemma 14 in the above reference, which appeals to the results of [10] in the linear setting. Now, taking  $r_0 \in \left(\frac{\alpha}{\alpha - (1 - \beta)\mathbb{1}_{\beta \in (-1, 0]}}, \frac{r}{1 + r\theta}\right)$  as in Lemma 10 precisely gives this condition with  $\delta = r_0$ . This completes the proof.  $\square$

### Strong uniqueness results.

**Proposition 13.** *Under the assumption (C1) or (C2), for any  $\mu$  in  $\mathcal{P}(\mathbb{R}^d) \cap B_{p_0, q_0}^{\beta_0}(\mathbb{R}^d)$  and  $0 \leq t < S < \mathcal{T}_1$ , with  $\mathcal{T}_1$  as in Theorem 1, there exists a unique strong solution to (1) such that its law  $\mu^{t, \mu}$  belongs to  $L_{w_\theta}^\infty((t, S], B_{p', 1}^{-\beta + \vartheta\Gamma})$  and such that for a.e.  $s$  in  $(t, S]$ ,  $\mu_s^{t, \mu}(dx) = \rho_{t, \mu}(s, x)dx$  whenever*

- under (C1)

$$\left(2 - \frac{3}{2}\alpha + \frac{\alpha}{r} + \left[-\beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p'_0}\right)\right]\right) \vee \left(1 - \alpha + \frac{\alpha}{r} + \left[-\beta + \frac{d}{p} - \left(\beta_0 + \frac{d}{p'_0}\right)\right]_+\right) < \beta.$$

- under (C2)

$$2 - \frac{3}{2}\alpha + \frac{\alpha}{r} + \left[\frac{d}{p} - \left(\beta_0 + \frac{d}{p'_0}\right)\right] < \beta = -1.$$

*Proof.* Similarly to the weak well-posedness we focus on the integrability properties of the drift viewed as the one of a *linear* version of the McKean-Vlasov SDE (1) where the law is frozen. This therefore amounts to prove that

$$\mathbf{b} : [0, S] \times \mathbb{R}^d \ni (s, x) \mapsto \mathbf{b}(s, x) = \int_{\mathbb{R}^d} b(s, x - y) \mu_s^{t, \mu}(dy) = \int_{\mathbb{R}^d} dy b(s, x - y) \rho_{t, \mu}(s, y), \quad (38)$$

satisfies a Krylov and Röckner type condition, see [30] if  $\alpha = 2$ , or the criterion in [50, Theorem 2.4] if  $\alpha \in (1, 2)$ .

If  $\alpha = 2$  we have already proved in the weak-uniqueness part that  $\mathbf{b} \in L^{r_0}(L^\infty)$  so that the Krylov Röckner criterion  $2/\delta + d/\ell < 1 = \alpha - 1$  actually holds with  $\delta = r_0, \ell = +\infty$ . This gives strong uniqueness in the diffusive case under (C1) or (C2).

Let us turn to  $\alpha \in (1, 2)$  which actually requires some smoothness properties additionally to the integrability conditions. Namely,

- For  $\alpha \in (1, 2)$ , strong well-posedness holds whenever the drift  $\mathbf{b}$  defined in (38) satisfies  $(I - \Delta)^{\gamma/2} \mathbf{b} \in L^\delta(L^\ell)$ , or equivalently,  $\mathbf{b} \in L^\delta(H^{\gamma, \ell})$ , where  $H^{\gamma, \ell}$  stands for the Bessel potential space and  $\gamma, \ell, \delta$  satisfy

$$\gamma \in \left(1 - \frac{\alpha}{2}, 1\right), \quad \ell \in \left(\frac{2d}{\alpha} \vee 2, \infty\right), \quad \delta \in \left(\frac{\alpha}{\alpha - 1}, \infty\right), \quad \frac{\alpha}{\delta} + \frac{d}{\ell} < \alpha - 1. \quad (39)$$

Note that the condition on  $\mathbf{b}$  in (ii) will actually follow if we manage to prove that  $\mathbf{b} \in L^\delta(B_{\ell, 1}^\gamma)$ . We indeed recall  $B_{\ell, 2}^\gamma \hookrightarrow H^{\gamma, \ell}$  when  $\ell \geq 2$  (see e.g. [49, Th. 2.5.6 p.88] and from (E2) for all  $\gamma > 0$ ,  $B_{\ell, 1}^\gamma \hookrightarrow B_{\ell, 2}^\gamma$ ).

From (Y) and for  $\ell$  meant to be large (but finite) write:

$$|\mathbf{b}(s, \cdot)|_{B_{\ell, 1}^\gamma} \leq |b(s, \cdot)|_{B_{p, q}^\beta} |\rho_{t, \mu}(s, \cdot)|_{B_{\ell_2, q'}^{\gamma - \beta}}, \quad \text{with } \ell_2 = p' \frac{1}{\frac{p'}{\ell} + 1}.$$

Note that one can choose any  $\ell_2 < p'$  close to  $p'$ , since again  $\ell$  is arbitrarily large but finite. Write from the Hölder inequality, similarly to the proof of Lemma 10:

$$\int_t^S ds |\mathbf{b}(s, \cdot)|_{B_{\ell, 1}^\gamma}^\delta \leq \left(\int_t^S ds |b(s, \cdot)|_{B_{p, q}^\beta}^{\delta a}\right)^{\frac{1}{a}} \left(\int_t^S ds |\rho_{t, \mu}(s, \cdot)|_{B_{\ell_2, q'}^{\gamma - \beta}}^{\delta a'}\right)^{\frac{1}{a'}}, \quad a^{-1} + (a')^{-1} = 1.$$

Again, since  $b \in L^r((t, S], B_{p, q}^\beta)$  the natural choice consists in taking  $a\delta = r$  giving  $1/a' = 1 - 1/a = 1 - \delta/r \iff a' = r/(r - \delta)$  and  $\gamma = \vartheta\Gamma$ . However, pay attention that, since the integrability index  $\ell_2$

is slightly smaller than  $p'$  we considered for the previous analysis of the density, we need to modify a bit the regularity index and consider a slightly smaller  $\gamma$  than the one indicated above. Namely, we get

$$\begin{aligned} \int_t^S ds |\mathbf{b}(s, \cdot)|_{B_{\ell_1,1}^\gamma}^\delta &\leq |b|_{L^r(B_{p,q}^\beta)}^\delta \left( \int_t^S ds |\boldsymbol{\rho}_{t,\mu}(s, \cdot)|_{B_{\ell_2,q'}^{\gamma-\beta}}^{\frac{\delta r}{r-\delta}} \right)^{\frac{1}{a'}} \\ &\leq |b|_{L^r(B_{p,q}^\beta)}^\delta |\boldsymbol{\rho}_{t,\mu}(s, \cdot)|_{L_{\bar{\omega}_\theta}^\infty((t,S], B_{\ell_2,1}^{-\beta+\theta\bar{\Gamma}})}^\delta \left( \int_t^S ds (s-t)^{-\bar{\theta}(\frac{\delta r}{r-\delta})} \right)^{\frac{1}{a'}} \\ &\leq C |b|_{L^r(B_{p,q}^\beta)}^\delta (S-t)^{\frac{1}{a'} - \bar{\theta}\delta} = C |b|_{L^r(B_{p,q}^\beta)}^\delta (S-t)^{\delta(\frac{1}{\delta} - (\frac{1}{r} + \bar{\theta}))}, \end{aligned} \quad (40)$$

for any  $\ell_2 < p'$  with  $\bar{\Gamma} = \Gamma - \bar{\eta}$ ,  $\bar{\theta} = \theta - \bar{\eta}$ ,  $\bar{\eta} := \bar{\eta}(p' - \ell_2) > 0$ , going to 0 when  $\ell_2$  goes to  $p'$  and with  $\Gamma, \theta$  as in (17), (19) respectively. Indeed, we can reproduce the previous steps, starting from the proof of Lemma 5 for the equation with mollified coefficients, in order to take into consideration a slightly smaller integration index.

Write indeed:

- Under **(C1)**

$$\left| \left( \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} \leq C \left| \left( \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \right|_{B_{\rho_1,\infty}^\Gamma} \left| \nabla p_{s-v}^\alpha \right|_{B_{\rho_2,1}^{-\beta}}.$$

- Under **(C2)**

$$\begin{aligned} &\left| \left( \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} \\ &\leq C \left[ \left| \left( \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \nabla \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \right|_{B_{\rho_1,\infty}^\Gamma} + \left| \left( \operatorname{div}(\mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot)) \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \right|_{B_{\rho_1,\infty}^\Gamma} \right] \left| p_{s-v}^\alpha \right|_{B_{\rho_2,1}^{-\beta}}, \end{aligned}$$

with  $1 + (\ell_2)^{-1} = \rho_1^{-1} + \rho_2^{-1}$ .

Apply now the product rule (**Prod1**) from Theorem 3 . We get:

- Under **(C1)**:

$$\left| \left( \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} \leq C \left| \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right|_{B_{\ell_1,\infty}^\Gamma} \left| \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^\Gamma} \left| \nabla p_{s-v}^\alpha \right|_{B_{\rho_2,1}^{-\beta}}.$$

- Under **(C2)**

$$\begin{aligned} &\left| \left( \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} \\ &\leq C \left[ \left| \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \right|_{B_{\ell_1,\infty}^\Gamma} \left| \nabla \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^\Gamma} + \left| \operatorname{div}(\mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot)) \right|_{B_{\ell_1,\infty}^\Gamma} \left| \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^\Gamma} \right] \left| p_{s-v}^\alpha \right|_{B_{\rho_2,1}^{-\beta}}, \end{aligned}$$

with  $\frac{1}{\rho_1} = \frac{1}{\ell_1} + \frac{1}{\ell_2}$ .

In order to repeat the previous procedure, one needs to take  $\bar{\ell}_2 = \ell_2$ . On the other hand, the previous choice yields

$$1 + \frac{1}{\ell_2} - \frac{1}{\rho_2} = \frac{1}{\rho_1} = \frac{1}{\ell_2} + \frac{1}{\ell_1} \implies 1 - \frac{1}{\rho_2} = \frac{1}{\ell_1}.$$

To fit the previous estimates it then remains to apply the Young inequality (**Y**):

- Under **(C1)**

$$\begin{aligned} \left| \left( \mathcal{B}_{\boldsymbol{\rho}_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} &\leq C |b^\varepsilon(v, \cdot)|_{B_{p,q}^\beta} \left| \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} \left| \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^\Gamma} \left| \nabla p_{s-v}^\alpha \right|_{B_{\rho_2,1}^{-\beta}} \\ &\leq C |b^\varepsilon(v, \cdot)|_{B_{p,q}^\beta} \left| \boldsymbol{\rho}_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^{-\beta+\Gamma}}^2 \left| \nabla p_{s-v}^\alpha \right|_{B_{\rho_2,1}^{-\beta}}. \end{aligned}$$



- Under **(C2)**

$$\begin{aligned} & \left| \left( \mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot) \cdot \rho_{t,\mu}^\varepsilon(v, \cdot) \right) \star \nabla p_{s-v}^\alpha \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} \\ & \leq C \left( |b^\varepsilon(v, \cdot)|_{B_{p,q}^\beta} + \left| \operatorname{div}(\mathcal{B}_{\rho_{t,\mu}^\varepsilon}^\varepsilon(v, \cdot)) \right|_{B_{p,q}^\beta} \right) \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^{-\beta+\Gamma}}^2 \left| p_{s-v}^\alpha \right|_{B_{\rho_2,1}^{-\beta}}, \end{aligned}$$

with parameters  $1 + \frac{1}{\ell_1} = \frac{1}{p} + \frac{1}{\ell_2}$ .

We thus deduce that  $\frac{1}{\ell_1} = 1 - \frac{1}{\rho_2} = \frac{1}{p} + \frac{1}{\ell_2} - 1 \iff \frac{1}{\rho_2} = 2 - \left(\frac{1}{p} + \frac{1}{\ell_2}\right) < 1$  since  $\ell_2 < p'$ . The procedure can be summed up in the following way: in order to slightly decrease the integrability index in the density estimate, we can slightly increase, through  $\rho_2$  the integrability which is asked on the gradient of the heat kernel. Using **(HK)** we would get similarly to (21) under **(C1)** and (25) under **(C2)**,

$$\begin{aligned} & \left| \rho_{t,\mu}^\varepsilon(s, \cdot) \right|_{B_{\ell_2,1}^{-\beta+\Gamma}} \\ & \leq c_1 |\mu|_{B_{p_0,q_0}^{\beta_0}} (s-t)^{-\frac{\Gamma}{\alpha} - \bar{\zeta}_0} + C |b|_{L^r(B_{p,q}^\beta)} \left( \int_t^s \frac{dv}{(s-v)^{r' \left( \frac{1-\beta \mathbb{I}_{\beta \in (-1,0]}}{\alpha} + \frac{d}{\alpha} \left(1 - \frac{1}{\rho_2}\right)\right)}} \left| \rho_{t,\mu}^\varepsilon(v, \cdot) \right|_{B_{\ell_2,1}^{-\beta+\Gamma}}^{2r'} \right)^{\frac{1}{r'}}, \end{aligned}$$

with

$$\bar{\zeta}_0 := \frac{1}{\alpha} \left[ -\beta + \frac{d}{\ell_2} - \left( \beta_0 + \frac{d}{p'_0} \right) \right], \quad \frac{1}{\ell_2} + \frac{1}{\ell'_2} = 1.$$

Following, up to the previous modifications, the arguments of Lemmas 5 and 7 then yields to (40).

The previous computations give that we can actually take  $\ell_2$  as close as we want to  $p'$  but in order to have  $\ell < +\infty$ . Also, for  $\ell_2$  close to  $p'$ ,  $\rho_2$  is close to one. Similarly, we can indeed take  $\gamma = \bar{\vartheta}\Gamma$  with  $\bar{\vartheta} \in (0, 1)$ , hence as close to 1 as desired.

It thus remains, to determine the conditions that will guarantee that strong well posedness holds, to check when the inequality

$$\Gamma > 1 - \frac{\alpha}{2}$$

is true.

Recall now that since  $\zeta_0 := \frac{1}{\alpha} \left[ -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right) \right]$ , and from (17)

$$\Gamma = \alpha + \beta \mathbb{I}_{\beta \in (-1,0]} - 1 - \frac{\alpha}{r} - \left( -\beta + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right) \right) - \eta = \alpha + \beta \mathbb{I}_{\beta \in (-1,0]} - 1 - \frac{\alpha}{r} - \alpha \zeta_0 - \eta, \eta > 0.$$

The previous condition thus rewrites

$$\beta(1 + \mathbb{I}_{\beta \in (-1,0]}) > 2 - \frac{3}{2}\alpha + \frac{\alpha}{r} + \frac{d}{p} - \left( \beta_0 + \frac{d}{p'_0} \right).$$

□

## 5. Connection with some physical and biological models.

We discuss in this section some specific applications of Theorem 1. We consider models related to turbulence theory and particle methods in Computational Fluid Dynamics as well as systems arising from recent trends in Biology. We particularly focus on the three following equations:

- The (scalar) Burgers equation,
- The two dimensional vortex equation for the incompressible Navier-Stokes equations,
- The parabolic-elliptic Keller-Segel equations for chemotaxis.

The common feature of those three equations is that they can be written as scalar valued singular non-linear transport-diffusion PDEs of the following form:

$$\partial_s u(s, \cdot) + \operatorname{div} \left( u(s, x) F(u(s, \cdot)) \right) - L^\alpha u(s, \cdot) = 0, \quad u(0, \cdot) = u^0, \quad (41)$$

where the driving  $F$  writes  $F(s, u) = K \star u(x)$ , for  $K$  a time-homogeneous strongly concentrating kernel. Importantly the systematic interpretation of the solution  $u$  - understood in a distributional sense - as the time marginal distributions  $\mu^{t,\mu}$  of (1) - taking therein  $t = 0$  - requires some preliminary considerations which start with the initial condition  $u^0$ . Due to the conservative form of (41), to ensure that  $u(s, \cdot)$  can

be viewed as a probability measure, this property needs to be fulfilled by the initial condition  $u(0, \cdot)$ . While this naturally restricts the physical interpretation of the model involved, beyond this situation, the McKean-Vlasov interpretation of (41) becomes trickier. We refer to [27] or [33] for related issues, see also Section 5.2 below.

The focus on the Burgers and Navier-Stokes equations will allow us to briefly revisit some predominant literature from the eighties and nineties. For the sake of clarity, points of comparison with the literature will be essentially focused on probabilistic models, leaving purposely aside a more complete survey on the PDE analysis of (41). For the same reasons, precise comparisons on the smallness of the initial condition and possible range of the time-horizons  $\mathcal{T}_1$  will be left aside.

**5.1. The Burgers equation.** In its most popular formulation, the Burgers equation corresponds to the scalar non-linear PDE:

$$\partial_s u(s, \cdot) + \frac{1}{2} \partial_x (u(s, \cdot))^2 - \nu \Delta u(s, \cdot) = 0, \quad s > 0, \quad u(0, \cdot) = u^0(\cdot), \quad (42)$$

where the solution  $u$  models the speed motion of a viscous fluid evolving on the real line  $\mathbb{R}$  under the joint action of a nonlinear transport operator  $\frac{1}{2} \partial_x (u(s, \cdot))^2 = u(s, \cdot) \partial_x u(s, \cdot)$  and the viscous dissipation  $\Delta u(s, \cdot)$  - for consistency with (1), the kinematic viscosity  $\nu$  has to be set to  $1/2$ . While (42) initially depicts a one-dimensional pressure-less model of Navier-Stokes equation, the Burgers equation nowadays applies in various disciplines such as aerodynamics, molecular biology, cosmology and traffic modelling. Its fractional version<sup>4</sup>:

$$\partial_s u(s, \cdot) + \frac{1}{2} \partial_x (u(s, \cdot))^2 - L^\alpha u(s, \cdot) = 0, \quad s > 0, \quad u(0, \cdot) = u^0(\cdot), \quad (43)$$

which substitutes the characteristic *heat* dissipation operator  $\Delta$  with the fractional one  $L^\alpha$ , presents a particular interest for hydrodynamics and statistical turbulence (we refer the interested reader to [4] and references therein for a brief exposure of the physical interest of (43) and to [2] for the impact of modified fractional dissipativity on recovering some characteristic scaling laws in turbulence).

From (9), it is easily seen that the McKean-Vlasov model related to the Burgers corresponds to an interaction kernel given by  $\frac{1}{2} \delta_{\{0\}}$  where  $\delta_{\{0\}}$  denotes the Dirac mass at 0.

Properly, the resulting McKean-Vlasov model formulates as

$$X_s^{0,\mu} = \xi + \frac{1}{2} \int_0^s \boldsymbol{\rho}(v, X_v^{0,\mu}) dv + \mathcal{W}_s, \quad \xi \sim u^0, \quad \boldsymbol{\rho}(v, \cdot) = \text{p.d.f of Law}(X_v), \quad (44)$$

or as

$$X_s^{0,\mu} = \xi + \frac{1}{2} \int_0^s \tilde{\mathbb{E}}[\delta_{\{X_v^{0,\mu} - \tilde{X}_v^{0,\mu}\}}] dv + \mathcal{W}_s, \quad (45)$$

where  $(\tilde{X}_s^{0,\mu})_{s \geq 0}$  has under  $\tilde{\mathbb{P}}$  the same law as  $(X_s^{0,\mu})_{s \geq 0}$ , provided the law of  $X_s^{0,\mu}$  is absolutely continuous at all time  $s \in (0, T]$ . For the case  $\alpha = 2$ , these formulations have been formally introduced in the seminal paper [34] (together with the interpretation of a model of the Boltzmann equation).

While [23], [7] and [42] focused on the particle approximation and associated propagation of chaos properties through analytic techniques, existence and uniqueness of a solution to (44) was, to the best of our knowledge, firstly established in [46]. The author obtained therein the existence and uniqueness of a strong solution, with  $\boldsymbol{\rho}$  in  $L^2((0, T] \times \mathbb{R}^d)$  for any arbitrary time horizon  $T$  (see Theorems 2.5 and 4.1 therein) under the condition  $u^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ .

In the fractional case  $\alpha \in (1, 2)$ , similar weak existence results have been successively established in [22] in the case of a symmetric stable noise with  $1 < \alpha < 2$  and  $u^0$  lying in the Sobolev space  $H^1(\mathbb{R}) = W^{1,2}(\mathbb{R})$ . Uniqueness is only established for one-time marginal distributions in the class  $\boldsymbol{\rho} \in L^\infty((0, T] \times L^2(\mathbb{R})) \cap L^2((0, T] \times H^1(\mathbb{R}))$  (see Theorems 2.1 and 3.1 of the indicated reference).

Recall from **(E<sub>3</sub>)** that a Dirac measure belongs to the Besov  $B_{p,\infty}^{-d/p'}$  for  $p \in [1, \infty]$ . In particular, the interaction kernel in (44) lies in the space  $L^\infty(B_{p,\infty}^{-1/p'})$ . Accordingly, we derive from condition **(C1)** that weak well-posedness holds if

$$1 - \alpha + \left[ 1 - \left( \beta_0 + \frac{1}{p'_0} \right) \right]_+ < -\frac{1}{p'}. \quad (46)$$

<sup>4</sup>also called fractal in some related papers, see [22]

Meanwhile,  $(\mathbf{C1}_S)$  becomes

$$2 - \frac{3}{2}\alpha + \left[1 - \left(\beta_0 + \frac{1}{p'_0}\right)\right]_+ < -\frac{1}{p'}. \quad (47)$$

The model (44) enters the setting of Assumption  $(\mathbf{C1})$  for a variety of situations, provided  $\mu$  is *more than a probability measure*, the threshold  $-\beta_0 \neq \frac{1}{p'_0}$  in (46) and (47) preventing for instance atomic initial state in (44). Illustratively, taking  $p = 1 = d$  and therefore  $\beta = -\frac{1}{p'} = 0$ ,  $r = q = \infty$ , one can derive from  $(\mathbf{C1})$ , (46) that weak uniqueness will hold as soon as either  $p_0 > 1$ ,  $\beta_0 = 0$  or  $\beta_0 > 0, p_0 = 1$ . From  $(\mathbf{C1}_S)$ , (47), we can observe that strong uniqueness will hold as soon as  $\beta_0 + \frac{1}{p'_0} > 3(1 - \frac{\alpha}{2})$ . This situation specifically reduces in the Brownian case, to  $\beta_0 + \frac{1}{p'_0} > 0$ .

Let us now turn to the specific case  $p = 2$  for which our approach allows to complete some results in the literature. From the above conditions (46), (47), the weak wellposedness of (44) provided  $\beta_0 + \frac{1}{p'_0} > \frac{5}{2} - \alpha$  and extends into a strong wellposedness result provided  $\beta_0 + \frac{1}{p'_0} > \frac{5}{2} - \frac{3}{2}\alpha + 1$ . In particular, from  $(\mathbf{E}_1)$ , we cover the previous setting of [46] in this case. In particular, from  $(\mathbf{E}_1)$  which implies  $L^1 \cap L^\infty \hookrightarrow L^{p_0} \hookrightarrow B_{p_0, \infty}^0$  for  $p_0 \in [1, \infty]$ , we cover - slightly extending the initial distribution to  $B_{p_0, \infty}^0$  - the class of strong solutions in [46] whenever  $p_0 > 2 \Leftrightarrow 1 - \frac{1}{p_0} = \frac{1}{p'_0} > \frac{1}{2} = \frac{5}{2} - \frac{3}{2}\alpha + 1$ . In the fractal case,  $\alpha \in (1, 2)$ , identify  $W^{1,2} = B_{2,2}^1$  (see e.g. [49], Theorem 2.5.6, p. 88, and Theorem 2.3.9, p. 61), we recover the existence result of [22], and add to this result, weak uniqueness of the corresponding SDE provided  $\frac{3}{2} < \alpha < 2$ , and strong uniqueness provided  $\frac{5}{3} < \alpha < 2$ .

**5.2. The vortex equation in dimension 2.** The vortex equations (or vorticity equations) model the rotational properties of an incompressible Newtonian turbulent fluid flow. The motions of such a fluid are described, at each time  $s$ , and each point  $x$  of  $\mathbb{R}^d$ ,  $d \in \{2, 3\}$ , through their macroscopic velocity  $u(s, x) = (u^{(1)}(s, x), \dots, u^{(d)}(s, x))$ , and their evolution, characterized by the incompressible Navier-Stokes equations:

$$\begin{aligned} \partial_s u(s, x) + (u(s, x) \cdot \nabla)u(s, x) &= -\nabla p + \frac{1}{2}\Delta u(s, x), \quad \nabla \cdot u(s, x) = 0, \quad s \geq 0, x \in \mathbb{R}^d, \\ u(0, x) &= u^0(x). \end{aligned}$$

For simplicity, we purposely focus the presentation of the equations on *function*-solutions  $u$  rather than distributional solutions, and remain in the classical dissipative case  $\alpha = 2$ . We also, again, set the kinematic viscosity to  $1/2$ . The divergence free constraint  $\nabla \cdot u(s, x) = 0$  reflects the incompressibility of the flow and  $\nabla p$  stands for the gradient of the pressure acting on the fluid.

In the case of a two-dimensional ( $d = 2$ ) flow, the vorticity  $w(s, x) := \text{curl}(u)(s, x) = \nabla \times u(s, x)$  is a scalar function driven by the equation

$$\begin{aligned} \partial_s w(s, x) + (u(s, x) \cdot \nabla_x)w(s, x) &= \frac{1}{2}\Delta w(s, x), \quad s \geq 0, x \in \mathbb{R}^2, \\ w(0, x) &= \nabla \times u^0(x). \end{aligned} \quad (48)$$

This case is thus simpler than the three dimensional case,  $d = 3$ , for which the vorticity is a field, i.e. the previous equation must then be understood as a system of 3 equations. The vorticity equation somehow decouples the non-linearity and allows to avoid to directly handle the pressure term (the curl of a gradient is zero). The original velocity  $u$  can be recovered, up to an additive constant, from the vorticity  $w$  using

$$\text{the identity } \Delta u = \begin{cases} \nabla^\perp w = \begin{pmatrix} -\partial_2 w \\ \partial_1 w \end{pmatrix}, & d = 2, \\ -\nabla \times w, & d = 3 \end{cases} \quad \text{which follows from the incompressibility property}$$

and *formally* leads to the identity

$$u(s, x) = K * w(s, x),$$

where  $K$  stands for the Biot-Savart kernel. Its expression actually appears applying the adjoint of the operator  $-\nabla^\perp$  for  $d = 2$ , resp.  $\nabla \times$  for  $d = 3$ , to the Poisson kernel  $\mathfrak{P}$ , i.e.  $u = \begin{cases} (-\Delta)^{-1}(-\nabla^\perp w), \\ (-\Delta)^{-1}\nabla \times w, & d = 3 \end{cases}$ .

Namely,

$$K(x) = \begin{cases} -\nabla^\perp \mathfrak{P}(x), & d = 2, \\ \nabla \mathfrak{P}(x) \times, & d = 3 \end{cases} \quad \text{with} \quad \mathfrak{P}(x) = \begin{cases} -\frac{\log(|x|)}{(2\pi)} & \text{if } d = 2, \\ \frac{1}{4\pi|x|^{d-2}} & \text{if } d = 3, \end{cases} \quad (49)$$

using as well the convention  $K * w(x) = \int_{\mathbb{R}^3} \nabla \mathfrak{P}(x-y) \times w(y) dy$  when  $d = 3$ . As for the Burgers equation, the *fractal/fractional* version of the Navier-Stokes equations, and by extension of (48):

$$\begin{aligned} \partial_s w(s, x) + (u(s, x) \cdot \nabla_x) w(s, x) &= L^\alpha w(s, x), \quad s \geq 0, \quad x \in \mathbb{R}^2, \\ w(0, x) &= \nabla \times u^0(x), \end{aligned} \quad (50)$$

presents a particular physical interest - we again refer to [2], and to the exhaustive presentation in [28].

Within the diffusive setting  $\alpha = 2$ , Chorin exploited in [11] the vorticity equations to develop particle methods -commonly known today as vortex methods- for the simulation of turbulent fluid flows. In [33], Marchioro and Pulvirenti addressed the link between the two-dimensional vortex equation and the McKean-Vlasov model. They later exploited this link to validate Chorin's particle method, introducing a smoothed mean-field particle approximation of (48) where the Poisson kernel  $\mathfrak{P}$  is regularized at the neighborhood of 0. Then, the authors established that the time marginal empirical measures propagate chaos toward the solution to (48), even in the zero viscosity limit. Osada [41] established a similar result for a non vanishing viscosity and without any smoothing of  $K$ . While these results apply to a peculiar probabilistic interpretation of (48), Méléard [36, 37] considered a McKean-Vlasov representation of (48) of the form:

$$X_s^{0,\mu} = \xi + \int_0^s \tilde{\mathbb{E}}[h_0(\tilde{X}_0^{0,\mu}) K(X_r^{0,\mu} - \tilde{X}_r^{0,\mu})] dr + \mathcal{W}_s, \quad (51)$$

where again  $(\tilde{X}_s^{0,\mu})_{s \geq 0}$  has under  $\tilde{\mathbb{P}}$  the same law as  $(X_s^{0,\mu})_{s \geq 0}$ , and  $h_0$  is a pre-factor which derives from the formulation of the problem in the McKean setting, i.e. solutions are sought as density functions whereas the initial condition  $w(0, \cdot)$  is not necessarily one. We again refer to [27] for additional related details.

The well-posedness of (51)- along a quantitative particle approximation - was first established -in terms of a nonlinear martingale problem- in [36] (see Theorems 1.2 and 2.4 therein), for a non-negative initial condition  $w_0$  in  $L^1 \cap L^\infty$  (see Theorems 1.2 and 2.4 of the same reference), and later extended in [37] (Theorems 1.2 and 3.4), to the case of  $w(0, \cdot)$  being a Radon measure of the form  $\nabla \times u^0$  where  $u^0$  belongs to a suitable Lorentz space.

Let us now discuss what can be derived from the approach developed in the current work in this setting. Observe first from (49) that for the two-dimensional vortex equation

$$K(x) = \frac{(-x_2, x_1)}{2\pi(x_1^2 + x_2^2)}.$$

Observe that in the distributional sense  $\operatorname{div} K = 0$ . Since  $|K(x)| \leq C/|x|$ ,  $K$  can be viewed as an element of  $L_{\text{loc}}^{2-\epsilon}$  for  $\epsilon > 0$  arbitrary. Essentially the singularity of  $K$  is localized in the neighborhood of 0, with  $K$  being smooth and bounded outside this region. We can so focus our attention on a truncated version of the kernel given by the drift  $b(x) = K(x) \mathbb{I}_{B(0,R)}^*(x)$ , where the cut-off  $\mathbb{I}_{B(0,R)}^*$  stands for a mollification of the indicator function of the ball  $B(0, R)$  with support in  $B(0, R+1)$  for a given radius  $R > 0$ . As such,  $b \in L^{2-\epsilon}$  and, since the embeddings  $(\mathbf{E}_1)$  and  $(\mathbf{E}_2)$  yield  $L^{2-\epsilon} \hookrightarrow B_{2-\epsilon, \infty}^0 \hookrightarrow B_{p(\epsilon), \infty}^{-1}$ , it also holds that  $b \in B_{p(\epsilon), \infty}^{-1}$  with  $p(\epsilon) = (4 - 2\epsilon)\epsilon^{-1}$  increasing as  $\epsilon \rightarrow 0$ . Observe that the truncation destroys the divergence free property of  $K$ , as  $\operatorname{div}(b)(x) = \left( \sum_{i \in \{1,2\}} K_i \partial_{x_i} \mathbb{I}_{B(0,R)}^*(x) \right) \neq 0$  for  $|x|$  in  $[R, R+1]$ . Nevertheless  $\operatorname{div}(b) \in L^{2-\epsilon}$  and therefore both  $b$  and  $\operatorname{div}(b)$  can be viewed simulatenuously as elements of  $B_{p(\epsilon), \infty}^{-1}$ .

Hence, the localized drift  $b$  associated with the two-dimensional vortex equation enters the setting of  $(\mathbf{C2})$  and  $(\mathbf{C2}_S)$  with  $r = q = \infty, p = p(\epsilon), d = 2$  provided  $\frac{d}{p(\epsilon)} < \beta_0 + \frac{d}{p'_0}$ . The corresponding inequalities then read respectively as

$$1 - \alpha + \frac{\epsilon}{2 - \epsilon} - \left( \beta_0 + \frac{2}{p'_0} \right) < -1, \quad \frac{\epsilon}{2 - \epsilon} < \beta_0 + \frac{2}{p'_0} \quad (52)$$

and

$$2 - \frac{3}{2}\alpha + \frac{\epsilon}{2 - \epsilon} - (\beta_0 + \frac{2}{p_0}') < -1. \quad (53)$$

This precisely implies that weak and strong uniqueness can respectively be derived under the conditions

$$\beta_0 + \frac{2}{p_0}' > 2 - \alpha \text{ and } \beta_0 + \frac{2}{p_0}' > 3 - \frac{3}{2}\alpha,$$

since a corresponding  $\epsilon > 0$  can then always be found. In the case  $\alpha = 2$ , this means that as soon as  $\beta_0 + \frac{2}{p_0}' > 0$  then weak and strong well-posedness hold. In particular, these results hold for  $\beta_0 > 0, p_0 = 1$  or  $\beta_0 = 0, p_0 > 1$ .

We essentially focused above on a *localized* version of the Biot-Savart kernel. Up to some refinement of the analysis leading to Theorem 1, in order to handle the tails of the kernel (which are bounded, smooth but do not satisfy the same integrability conditions observed near 0) we believe the results should extend to the complete Biot-Savart kernel. We would therefore recover the results of [36] and [37] for  $\alpha = 2$  and provide some tractable conditions in the strictly stable case.

To complete this section, we may briefly mention the three-dimensional case. The McKean-Vlasov model related to the three dimensional vortex equation was (formally) introduced in [14]. The corresponding well-posedness and the validity of a smoothed particle approximation were then obtained in [17, 18]. The increased complexity of the three dimensional case naturally extend to its McKean-Vlasov interpretation, the corresponding model involving an intricate systems of coupled nonlinear dynamics. Despite this complexity, we are confident that we could also recover those results (as well as a stable extension) from our current approach.

**5.3. The Keller-Segel model.** The Keller-Segel equations are a system of second order PDEs describing the joint evolution of the distribution  $\nu_s(dx)$  of cells (e.g. bacteria) and the concentration of chemo-attractant  $c = c(s, x)$ , which induces a significant force field in the cell evolution. In its parabolic-elliptic form, and assuming that the cell distribution has a density, i.e.  $\nu_s(dx) = u(s, x)dx$ , the equations write as

$$\begin{aligned} \partial_s u(s, x) + \chi \nabla \cdot (u(s, x) \nabla c(s, x)) - \frac{1}{2} \Delta u(s, x) &= 0, \\ - \Delta c(s, x) &= u(s, x), \\ u(0, x) = u^0(x), \quad c(0, x) = c^0(x) &\text{ given.} \end{aligned} \quad (54)$$

The coefficient  $\chi$  modulates the intensity of the action of the concentration. Formally, writing again  $c(s, x) = (-\Delta)^{-1} u(s, x)$  we have  $c(s, x) = \int_{\mathbb{R}^d} \mathfrak{P}(x-y) u(s, y) dy$  so that  $\nabla c(s, x) = \int_{\mathbb{R}^d} \nabla \mathfrak{P}(x-y) u(s, y) dy := (K * u(s, \cdot))(x)$  for  $K(z) = -z/|z|^d c_d$  (where the constant  $c_d$  depends on the considered dimension). This leads to consider a kernel of the form  $b(x) = \chi K(x)$  to derive the corresponding McKean-Vlasov interpretation. The kernel is strongly attractive, and compared to the vortex equations is *not* divergence-free. In particular this can lead to blow-up phenomena, i.e. the cells aggregate at 0, leading to a degeneracy of  $\nu_s$  to the Dirac measure  $\delta_{\{0\}}$ . The way these blow-up phenomena emerge is inherently related to the smallness of  $u^0$  and  $\chi$ . In dimension 2, it is known, see e.g. the monograph of Biler [3], that global well-posedness will hold provided  $\chi < 8\pi$ . The motivation to model diffusion through a (non-local) fractional diffusion comes from the fact that organisms may adopt Lévy flight search strategies for their nutriment. In that setting, dispersal is then better modeled by non-local operators ([13], [6]).

To show how (54) enters in the framework of the assumptions of Theorem 1, we may proceed as in the case of the vortex equation, observe first that the kernel  $K$  belongs to  $L_{\text{loc}}^{\mathbf{p}}$  for  $\mathbf{p} < \frac{d}{d-1}$ . Hence, setting  $b = \chi K \mathbb{I}_{B(0,R)}^*$ ,  $R > 0$  where as above  $\mathbb{I}_{B(0,R)}^*$  stands for a mollification of the indicator function of the ball  $B(0, R)$ , we get  $b \in L^p$  for  $p < \frac{d}{d-1}$ . Using again the embeddings **(E<sub>1</sub>)** and **(E<sub>2</sub>)** we thus derive that for  $\epsilon > 0$ ,  $b \in B_{p_d(\epsilon), \infty}^{\beta_d(\epsilon)}$  with  $0 - \frac{d}{d-1-\epsilon} = \beta_d(\epsilon) - \frac{d}{p_d(\epsilon)}$ . In order to first enter the setting of **(C2)** this leads to consider  $\beta_d(\epsilon) = -1$  and consequently  $p_d(\epsilon) = \frac{d}{d-2+\frac{\epsilon(d-1)}{d}} \times (1 - \frac{\epsilon(d-1)}{d})$ . It now remains to check that the parameters can be tailored so that  $\text{div}(b)$  also belongs to this function space. In the distributional sense  $\text{div}(K) = \delta_{\{0\}}$  and  $K_i \partial_{x_i} (\mathbb{I}_{B(0,R)}^*) \in L^p$ ,  $\mathbf{p} < \frac{d}{d-1}$ . Recall from **(E<sub>3</sub>)** that  $\delta_{\{0\}} \in B_{\ell, \infty}^{-d/\ell'}$ ,  $\ell \in [1, +\infty]$ . Therefore, to concur with the setting one must take  $\ell' = d$  and therefore  $\ell = \frac{d}{d-1}$ . Consequently  $\epsilon = \frac{1}{d-1}$

and so  $\ell = p_d(\epsilon)$ . These choices precisely ensure that  $b, \operatorname{div}(b) \in B_{\frac{d}{d-1}, \infty}^{-1}$ . It can be pointed out that for the Keller-Segel model, it is the Dirac concentration of  $\operatorname{div}(K)$  which imposes the functional setting of the kernel - as opposed to the Vortex case for which the kernel was predominant (because of its divergence free property).

Assumption **(C2)** for the weak well-posedness of (1) for  $r = q = \infty, p = \frac{d}{d-1}$  gives<sup>5</sup>

$$1 - \alpha + d < \beta_0 + \frac{d}{p'_0}$$

and **(C2<sub>S</sub>)**,

$$2 - \frac{3}{2}\alpha + d < \beta_0 + \frac{d}{p'_0}$$

In particular, in the the Brownian regime, the condition for the (local) *weak/strong* reads as

$$d - 1 - \left(\beta_0 + \frac{d}{p'_0}\right) < 0.$$

This will be fulfilled if e.g.  $\beta_0 = 0$  (no *a priori* smoothness of the initial data) and  $p_0 > d$  or at the other extreme, (imposing no specific integrability properties) provided  $\beta_0 > d - 1$  for  $p_0 = 1$ .

As for the vortex case, despite the localization of the kernel, the above conditions should be the ones under which *weak and strong* well-posedness for the McKean-Vlasov SDE associated with the Keller-Segel system (54) hold.

## References.

- [1] S. Athreya, O. Butkovsky, and L. Mytnik. Well-posedness of stochastic heat equation with distributional drift and skew stochastic heat equation. *ArXiv preprint: arXiv:2011.13498*, 2020.
- [2] C. Bardos, U. Frisch, M. Lesieur and A. Brissaud. *A Markovian random coupling model for turbulence*, Journal of Fluid Mechanics, 65(1): 145-152, 1979.
- [3] P. Biler. *Singularities of Solutions to Chemotaxis Systems*, De Gruyter, 2019.
- [4] P. Biler, T. Funaki, and W.A. Woyczuński. Fractal Burgers equations. *Journal of Differential Equations*, 148: 9 -46, 1998.
- [5] P. Billingsley. *Convergence of Probability Measures, Second Edition*, Wiley Intersciences Publications, 1999.
- [6] Bournaveas, N., and Calvez, V., *The one-dimensional Keller-Segel model with fractional diffusion of cells*, Nonlinearity, 23(4): 923-935, 2010. Link
- [7] P. Calderoni and M. Pulvirenti. *Propagation of chaos for Burgers' equation*, Annales de l'I.H.P., section A, 39(1): 85-97, 1983.
- [8] P.-E. Chaudru de Raynal and J.-F. Jabir and S. Menozzi. *McKean-Vlasov SDEs with distributional interaction kernel and stable additive noise: well-posedness and smoothing effects*. Preprint, 2022.
- [9] P.-E. Chaudru de Raynal and N. Frikha. *Well-Posedness for some non-linear diffusion processes and related PDEs on the Wasserstein space*. *Journal de Mathématiques Pures et Appliquées*, 159:1-167, 2022.
- [10] P.-E. Chaudru de Raynal and S. Menozzi. *On multidimensional stable-driven stochastic differential equations with Besov drift*. *Electron. J. Probab.* 27: 1-52 (2022). DOI: 10.1214/22-EJP864.
- [11] A. J. Chorin. *Vorticity and Turbulence*, Springer, 1994.
- [12] D. A. Dawson. *Critical dynamics and fluctuations for a mean field model of cooperative behavior*, J. Statist. Phys. 31: 29-85, 1983.
- [13] C. Escudero. *The fractional Keller-Segel model*, Nonlinearity, 19: 2909-2918, 2006.
- [14] R. Esposito, R. Marra, M. Pulvirenti, and C. Sciarretta. *A stochastic lagrangian picture for the three dimensional Navier-Stokes equation*, Commun. in Partial Differential Equations, 13(12): 1601-1610, 1988.
- [15] S. N. Ethier and T. G. Kurtz. *Markov Processes - Characterization and Convergence*, Willey, 1985.
- [16] J. Fontbona. *Nonlinear martingale problems involving singular integrals*, Journal of Functional Analysis, 200(1):198-236, 2003.
- [17] J. Fontbona. *A probabilistic interpretation and stochastic particle approximations of the 3-dimensional Navier-Stokes equations*, Probab. Theory Relat. Fields 136: 102-156, 2006.
- [18] J. Fontbona. *Stochastic vortex method for forced three-dimensional Navier-Stokes equations and pathwise convergence rate*, The Annals of Applied Probability, 20(5): 1761-1800, 2010.
- [19] N. Fournier, and B. Jourdain. *Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes*, Ann. Appl. Probab. 5, 2807-2861, 2017.
- [20] A. V. Friedman. *Partial Differential Equations of Parabolic Type*, Prentice-Hall, 1964.
- [21] N. Frikha, V. Konakov, and S. Menozzi. *Well-posedness of some non-linear stable driven SDEs*, Discrete Contin. Dyn. Syst., 41(2): 849-898, 2021.

---

<sup>5</sup>We carefully mention that the condition  $\frac{d}{p} = d - 1 < \beta_0 + \frac{d}{p'_0}$  appearing in **(C2)** is actually implied by the following one since  $\alpha \in (1, 2]$ .

- [22] T. Funaki, and W.A. Woyczyński. Interacting particle approximation for fractal Burgers equations. In *Stochastic Processes and Related Topics, A Volume in Memory of Stamatis Cambanis*, Birkhäuser-Boston, 141-166, 1998.
- [23] E. Gutkin, and M. Kac. Propagation of Chaos and the Burgers Equation. *SIAM Journal on Applied Mathematics*, 43(4): 971–980, 1983.
- [24] Z. Hao, M. Röckner, X. Zhang. Second order fractional mean-field SDEs with singular kernels and measure initial data. *arXiv:2302.04392*
- [25] T. Hytönen, J. van Neerven, M. Veraar and L. Weis. *Analysis in Banach Spaces*, Springer, 2016.
- [26] B. Jourdain. *Diffusions with a nonlinear irregular drift coefficient and probabilistic interpretation of generalized Burgers equations*, ESAIM Probability and Statistics, 1: 339-355, 1997.
- [27] B. Jourdain. *Diffusion processes associated with nonlinear evolution equations for signed measures*, Methodology and Computing in Applied Probability, 2(1):69-91, 2000.
- [28] M. K. Kavvas and A. Ercan. *Generalizations of incompressible and compressible Navier-Stokes equations to fractional time and multi=fractional space*, Sci. Rep. 12, 19337, 2022
- [29] V. N. Kolokoltsov. Symmetric Stable Laws and Stable-Like Jump-Diffusions. *Proceedings of the London Mathematical Society*, 80(3):725–768, May 2000.
- [30] N. Krylov and M. Röckner. *Strong solutions of stochastic equations with singular time dependent drift*, Prob. Theory Rel. Fields, 131: 154-196, 2005.
- [31] P.-G. Lemarié-Rieusset. *Recent Developments in the Navier-Stokes Problem*, Chapman & Hall/CRC Research Notes in Mathematics, 2002.
- [32] P.-G. Lemarié-Rieusset. *The Navier-Stokes Problem in the 21<sup>st</sup> century*, Chapman & Hall/CRC Research Notes in Mathematics, 2016.
- [33] C. Marchioro and M. Pulvirenti. *Hydrodynamics in two dimensions and vortex theory*, Commun. Math. Phys. 84: 483-503, 1982.
- [34] H. P. McKean. *A class of markov processes associated with nonlinear parabolic equations*, Proceedings of the National Academy of Sciences of the United States of America, 56(6): 1907-1911, 1966.
- [35] S. Méléard. *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, In Probabilistic Models for Nonlinear Partial Differential Equations, Lectures of the 1st Session of the C.I.M.E., Springer, 1995.
- [36] S. Méléard. *A trajectorial proof of the Vortex method for the two-dimensional Navier–Stokes equation*, The Annals of Applied Probability, 10(4): 1197-1211, 2000.
- [37] S. Méléard. *Monte-Carlo approximations for 2D Navier-Stokes equations with measure initial data*, Probab. Theory Relat. Fields, 121: 367–388, 2001.
- [38] S. Méléard and S. Roelly-Coppoletta. *A propagation of chaos result for a system of particles with moderate interaction* Stochastic Process.Appl. 26: 317–332, 1987.
- [39] S. Menozzi and X. Zhang. *Heat kernel of supercritical SDEs with unbounded drifts*. *Journal de l’Ecole Polytechnique*, 9:537–579, 2022.
- [40] K. Oelschläger. *A Martingale Approach to the Law of Large Numbers for Weakly Interacting Stochastic Processes*, Ann. Probab. 12(2): 458-479, 1984.
- [41] H. Osada. *Propagation of chaos for the two dimensional Navier-Stokes equation*, Proc. Japan Acad., 62, Ser. A, 1986.
- [42] H. Osada and S. Kotani. *Propagation of chaos for the Burgers equation*, J. Math. Soc. Japan, 37(2): 275-294, 1985.
- [43] M. Röckner and X. Zhang. *Well-posedness of distribution dependent SDEs with singular drifts*, Bernoulli Journal, 27(2): 1131-1158, 2021.
- [44] T. Runst and W. Sickel. *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations*, Walter de Gruyter & Co, 1997.
- [45] Y. Sawano. *Theory of Besov Spaces*, Springer, 2018.
- [46] A.-S. Sznitman. *A Propagation of Chaos Result for Burgers’ Equation*, Probab. Th. Rel. Fields, 71: 581-613, 1986.
- [47] A.-S. Sznitman. *Topics in propagation of chaos*, in Ecole d’Eté de Probabilités de Saint-Flour XIX -pp 165-251 - 1989.
- [48] H. Tanaka with K. Kawazu, H. P. McKean, M. Nagasawa, Y. Saisho and T. Shiga, *Stochastic Processes - Selected papers of Hiroshi Tanaka*, World Scientific, 2002.
- [49] H. Triebel. *Theory of Function Spaces*, Modern Birkhäuser Classics, 2010 reprint of the 1983 edition.
- [50] L. Xie and X. Zhang. *Ergodicity of stochastic differential equations with jumps and singular coefficients*. Ann. Inst. Henri Poincaré Probab. Stat. 56 (2020), n°1, 175-229.
- [51] X. Zhang and G. Zhao. *Stochastic Lagrangian path for Leray’s solutions of 3D Navier-Stokes equations*. Comm. Math. Phys. 381 (2021), n°2, 491-525.

LABORATOIRE DE MATHÉMATIQUES JEAN LERAY, UNIVERSITY OF NANTES, 2, RUE DE LA HOUSSINIÈRE BP 92208 F-44322 NANTES CEDEX 3, FRANCE.

*Email address:* [pe.deraynal@univ-nantes.fr](mailto:pe.deraynal@univ-nantes.fr)

LABORATORY OF STOCHASTIC ANALYSIS, HSE UNIVERSITY, POKROVSKY BLVD, 11, MOSCOW, RUSSIAN FEDERATION.

*Email address:* [jjabir@hse.ru](mailto:jjabir@hse.ru)

LAMME, UNIVERSITÉ D’EVRY VAL D’ESSONNE, UNIVERSITÉ PARIS-SACLAY, CNRS (UMR 8071), 23 BOULEVARD DE FRANCE 91037 EVRY, FRANCE & LABORATORY OF STOCHASTIC ANALYSIS, HSE UNIVERSITY, POKROVSKY BLVD, 11, MOSCOW, RUSSIAN FEDERATION.

*Email address:* [stephane.menzio@univ-evry.fr](mailto:stephane.menzio@univ-evry.fr)