Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 301 (2021) 428–470

Journal of Differential Equations

www.elsevier.com/locate/jde

On fractional and nonlocal parabolic mean field games in the whole space ☆

Olav Ersland, Espen R. Jakobsen *

Department of Mathematical Sciences, Norwegian University of Science and Technology NTNU, Norway

Received 23 April 2021; accepted 24 August 2021
Available online 2 September 2021

Abstract

We study Mean Field Games (MFGs) driven by a large class of nonlocal, fractional and anomalous diffusions in the whole space. These non-Gaussian diffusions are pure jump Lévy processes with some σ-stable like behaviour. Included are σ-stable processes and fractional Laplace diffusion operators $(-\Delta)^{\frac{\sigma}{2}}$, tempered nonsymmetric processes in Finance, spectrally one-sided processes, and sums of subelliptic operators of different orders. Our main results are existence and uniqueness of classical solutions of MFG systems with nondegenerate diffusion operators of order $\sigma \in (1, 2)$. We consider parabolic equations in the whole space with both local and nonlocal couplings. Our proofs use pure PDE-methods and build on ideas of Lions et al. The new ingredients are fractional heat kernel estimates, regularity results for fractional Bellman, Fokker-Planck and coupled Mean Field Game equations, and a priori bounds and compactness of (very) weak solutions of fractional Fokker-Planck equations in the whole space. Our techniques require no moment assumptions and use a weaker topology than Wasserstein.

MSC: 35Q89; 47G20; 35A01; 35A09; 35Q84; 49L12; 45K05; 35S10; 35K61; 35K08

Keywords: Mean field games; Fractional PDE; Existence; Uniqueness; Classical solutions; Nonlocal and local couplings

☆ Both authors were supported by the Toppforsk (research excellence) project Waves and Nonlinear Phenomena (WaNP), grant no. 250070 from the Research Council of Norway.

* Corresponding author.
E-mail addresses: olav.ersland@ntnu.no (O. Ersland), espen.jakobsen@ntnu.no (E.R. Jakobsen).

https://doi.org/10.1016/j.jde.2021.08.026

Contents

1. Introduction

We study parabolic Mean Field Games (MFGs) driven by a large class of nonlocal, fractional and anomalous diffusions in the whole space:

$$\begin{cases} -\partial_t u - \mathcal{L}u + H(x, u, Du) = F(x, m(t)) & \text{in } (0,T)\times\mathbb{R}^d, \\ \partial_t m - \mathcal{L}^* m - div\big(m D_p H(x,u,Du)\big) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ m(0,x) = m_0(x), \qquad u(x,T) = G(x, m(T)), \end{cases} \tag{1}$$

where H is a (nonlinear) Hamiltonian, F and G are source term and terminal condition, and m_0 an initial condition. Furthermore, $\mathcal{L}$ and its adjoint $\mathcal{L}^*$, are non-degenerate fractional diffusion operators of order $\sigma \in (1,2)$ of the form

$$\mathcal{L}u(x) = \int_{\mathbb{R}^d} u(x+z) - u(x) - Du(x)\cdot z 1_{|z|<1}\, d\mu(z), \tag{2}$$

where μ is a nonnegative Radon measure satisfying the Lévy-condition $\int_{\mathbb{R}^d} 1 \wedge |z|^2\, d\mu(z) < \infty$, see (L1) and (L2) below for precise assumptions. The system is uniformly parabolic and consists of a backward in time fractional Hamilton-Jacobi-Bellman (HJB) equation coupled with a forward in time fractional Fokker-Planck (FP) equation.

1.1. Background

MFGs is an emerging field of mathematics with a wide and increasing range of applications in e.g. economy, network engineering, biology, crowd and swarm control, and statistical learning [26,22]. It was introduced more or less at the same time by Lasry and Lions [31,32] and Caines, Huang and Malhamé [27]. Today there is a large and rapidly expanding literature addressing a range of mathematical questions concerning MFGs. We refer to the books and lecture notes [1,12,10,23,7] and references therein for an overview of the theory and the current state of the art. Heuristically a large number of identical players want to minimize some cost depending on their own state and the distribution of the states of the other players, and the mean field game system arises as a characterisation of Nash equilibria when the number of players tends to infinity under certain symmetry assumptions. The optimal MFG feedback control is almost optimal also for finite player games with moderate to large numbers of players, and often provides the only practical way of solving also such games.

In this paper the generic player controls a stochastic differential equation (SDE) driven by a pure jump Lévy process L_t with characteristic triplet $(0, 0, \mu)$ [3],

$$dX_t = \alpha_t\, dt + dL_t, \tag{3}$$

with the aim of minimizing the cost functional

$$\mathbb{E}\left[\int_0^T \Big[L(X_s, \alpha_s) + F(X_s, m(s))\Big]ds + G\left(X_T, m\left(T\right)\right)\right]$$

with respect to the control α_s. Here L is the Legendre transform of H with respect to the second variable, F and G are running and terminal costs, and m the distribution of the states of the other players. If u is the value function of the generic player, then formally the optimal feedback control is $\alpha_t^* = -D_2 H(x, Du)$ and u satisfies the HJB equation in (1). The probability distribution of the optimally controlled process X_t^* then satisfies the FP equation in (1). Since the players are identical, the distribution m of all players will satisfy the same FP equation, now starting from the initial distribution of players m_0. This is a heuristic explanation for (1).

What differs from the standard MFG formulation is the type of noise used in the model. In many real world applications, jump processes or anomalous diffusions will better model the observed noise than Gaussian processes [34,18,38,3]. One example is symmetric σ-stable noise which correspond to fractional Laplacian operators $\mathcal{L} = (-\Delta)^{\frac{\sigma}{2}}$ for $\sigma \in (0, 2)$. In Finance the observed jump processes are not symmetric and σ-stable but rather non-symmetric and tempered. An example is the one-dimensional CGMY process [18] where $\frac{d\mu}{dz}(z) = \frac{C}{|z|^{1+Y}} e^{-Gz^+ - Mz^-}$ for $C, G, M > 0$ and $Y \in (0, 2)$. Our assumptions cover a large class of uniformly elliptic operators $\mathcal{L}$ that includes fractional Laplacians, generators of processes used in Finance, anisotropic operators with different orders σ in different directions, Riesz-Feller operators, and operators with Lévy measures that non-absolutely continuous, spectrally one-sided, have no fractional moments, and a general behaviour at infinity. We refer to Section 4 for a discussion, results, and examples. We also analyse the system in the whole space, while many other papers focus on the compact torus. For control problems and games, the whole space case is usually more natural, but also more technical.

Main results. Under structure and regularity assumptions on $\mathcal{L}, H, F, G, m_0$, we show:

(i) *Existence of smooth solutions of* (1) with nonlocal and local coupling, see Theorems 3.2 and 3.5.
(ii) *Uniqueness of smooth solutions of* (1) with nonlocal and local coupling, see Theorems 3.3 and 3.6.

Our assumptions on H, F, G are fairly standard [33,9,1] (except maybe that the problem is posed on the whole space). For the existence results, we note that the Hamiltonian H (assumptions (A3)–(A5)) can be both nonconvex and noncoercive. Since we consider nondegenerate parabolic problems, the order of the equations has to be greater than one and we do not need or impose semiconcavity assumptions. The proofs of the main results follow from an adaptation of the PDE-approach of Lions [33,9,1], and existence is much more involved than uniqueness. Existence for MFGs with nonlocal coupling is proved using a Schauder fixed point argument and well-posedness, regularity, stability and compactness results for individual fractional HJB and fractional FP equations of the form:

$$\partial_t u - \mathcal{L}u + H(x, u, Du) = f(t, x),$$

$$\partial_t m - \mathcal{L}^* m + div\,(b(t, x)m) = 0.$$

Existence for MFGs with local coupling follows from an approximation argument, the results for nonlocal coupling, and regularity and compactness results, in this case directly for the coupled MFG system.

Secondary results:

(iii) *Fractional heat kernel estimates*, see Theorem 4.3 and Proposition 4.9.
(iv) *Fractional HJB equations:* Regularity, existence, and space-time compactness of derivatives of classical solutions in Theorem 5.5 and Theorem 5.6.
(v) *Fractional FP equations:* Well-posedness, space-time compactness of derivatives, $C(0, T; P(\mathbb{R}^d))$ compactness, and global L^∞ bounds of smooth solutions in Theorem 6.8 (a), Theorem 6.8 (b) and (c), Proposition 6.6, and Lemma 6.7.

For both equations we show new high order regularity results of independent interest. These results are obtained from a Banach fixed point argument using semigroup/Duhamel representation of the solutions and bootstrapping in the spirit of [19,20,28]. Key ingredients are very general fractional heat kernel estimates and global in time Lipschitz bounds for u and L^∞ bounds for m. The heat kernel estimates are based on [25], and we give some extensions, e.g. to operators with general Lévy measures at infinity and sums of subelliptic operators. To show space-time compactness of derivatives, we prove that they are space-time equi-continuous, combining uniform Hölder estimates in space with new time and mixed regularity estimates for the Duhamel representations of the solutions (see Section 5). In the local coupling case, the HJB and FP equations have less regular data, and regularity can no longer be obtained through separate treatment of the equations. Instead we need to work directly on the coupled MFG system and apply a more refined bootstrapping argument based on fractional derivatives. These estimates also require better global in time Lipschitz and L^∞ estimates the HJB and FP equations respectively. Here we use a variant of the Lipschitz bound of [5] and provide a new L^∞-estimate for the FP equation.

For the Schauder fixed point argument to work and give existence for the MFG system, compactness in measure is needed for a family of solutions of the FP-equation. We prove such compactness essentially through an analysis of very weak (distributional) solutions of this equation: We prove preservation of positivity, mass, and L^1-norms, equicontinuity in time, and tightness. Our proof of equicontinuity is simple and direct, without probabilistic SDE-arguments as in e.g. [9,1]. The tightness estimates are new in the fractional MFG setting and more challenging than in the local case.

This paper is the first to consider fractional MFGs in the whole space. To have compactness in measure on non-compact domains, a new ingredient is needed: tightness. Typically tightness is obtained through some moment condition on the family of measures. Such moment bounds depend both on the initial distribution and the generator of the process. In the local case when L_t in (3) is a Brownian motion, then the process X_t and FP solution m have moments of any order, only limited by the number of moments of X_0 and m_0. In the nonlocal/fractional case, X_t and m may have only limited (as for σ-stable processes) or even no fractional power moments at all, even when X_0 and m_0 have moments of all orders. We refer to Section 2.3 for more examples, details, and discussion. Nonetheless it turns out that some generalized moment exists, and tightness and compactness can then be obtained. This relies on Proposition 6.5 (taken from [15]), which gives the existence of a nice "Lyapunov" function that can be integrated against m_0 and $\mu 1_{|z|\geq 1}$.

In this paper we prove tightness and compactness without any explicit moment conditions on the underlying processes X_t or solutions of the FP equations m. This seems to be new for MFGs even in the classical local case. Furthermore, m is typically set in the Wasserstein-1 space $\mathcal{W}_1$ of measures with first moments, and compactness then requires more than one moment to be uniformly bounded. Since our Lévy processes and FP solutions may not have first moments, we can not work in this setting. Rather we work in a weaker setting using a weaker Rubinstein-Kantorovich metric d_0 (defined below) which is equivalent to weak convergence in measure (without moments). This is reflected both in the compactness and stability arguments we use as well as our assumptions on the nonlocal couplings.

1.2. Literature

In the case of Gaussian noise and local MGF systems, this type of MFG problems with nonlocal or local coupling have been studied from the start [31–33,9] and today there is an extensive literature summarized e.g. in [1,23,7] and references therein. For local MFGs with local couplings, there are also results on weak solutions [32,35,11,1], a topic we do not consider in this paper. Duhamel formulas have been used e.g. to prove short-time existence and uniqueness in [17].

In the case of non-Gaussian noise and nonlocal MFGs or MFGs with fractional diffusions, there is already some work. In [13] the authors analyze a stationary MFG system on the torus with fractional Laplace diffusions and both non-local and local couplings. Well-posedness of time-fractional MFG systems, i.e. systems with fractional time-derivatives, are studied in [8]. Fractional parabolic Bertrand and Carnout MFGs are studied in the recent paper [24]. These problems are posed in one space dimension, they have a different and more complicated structure than ours, and the principal terms are the (local) second derivative terms. The nonlocal terms act as lower order perturbations. Moreover, during the rather long preparation of this paper we learned that M. Cirant and A. Goffi were working on somewhat similar problems. Their results have now been published in [16]. They consider time-depending MFG systems on the torus with fractional Laplace diffusions and nonlocal couplings. Since they assume additional convexity and coercivity assumptions to ensure global in time semiconcavity and Lipcshitz bounds on solutions, they consider also fractional Laplacians of the full range of orders $\sigma \in (0, 2)$. Regularity results are given in terms of Bessel potential and Hölder spaces, weak energy solutions are employed when $\sigma \in (0, 1]$, and existence is obtained from the vanishing viscosity method. Our setup is different in many ways, and more general in some (a large class of diffusion operators, less smoothness on the data, problems posed in the whole space, no moment conditions, fixed point arguments), and most of our proofs and arguments are quite different from those in [16]. We also give results for local couplings, which in view of the discussion above is a non-trivial extension.

1.3. Outline of paper

This paper is organized as follows: In section 2 we introduce notation, spaces, and give some preliminary assumptions and results for the nonlocal operators. We state assumptions and give existence and uniqueness results for MFG systems with nonlocal and local coupling in Section 3. To prove these results, we first establish fractional heat kernel estimates in Section 4. Using these estimates and Duhamel representation formulas, we prove regularity results for fractional Hamilton-Jacobi equations in Section 5. In Section 6 we establish results for fractional Fokker-Planck equations, both regularity of classical solutions and $C([0,T], P(\mathbb{R}^d))$ compactness. In Sections 7 and 8 we prove the existence result for nonlocal and local couplings respectively, while uniqueness for nonlocal couplings is proved in Appendix A. Finally we prove a technical space-time regularity lemma in Appendix B.

2. Preliminaries

2.1. Notation and spaces

By C, K we mean various constants which may change from line to line. The Euclidean norm on any $\mathbb{R}^d$-type space is denoted by $|\cdot|$. For any subset $Q \subset \mathbb{R}^N$ and for any bounded, possibly vector valued, function on Q, we define the L^∞ norms by $\|w\|_{L^\infty(Q)} := \text{ess sup}_{y\in Q}|w(y)|$. Whenever $Q = \mathbb{R}^d$ or $Q = [0,T] \times \mathbb{R}^d$, we denote $\|\cdot\|_{L^\infty(Q)} := \|\cdot\|_\infty$. Similarly, the norm in L^p space is denoted by $\|\cdot\|_{L^p(Q)}$ or simply $\|\cdot\|_p$. We use $C_b(Q)$ and $UC(Q)$ to denote the spaces of bounded continuous and uniformly continuous real valued functions on Q, often we denote the norm $\|\cdot\|_{C_b}$ simply by $\|\cdot\|_\infty$. Furthermore, $C_b^k(\mathbb{R}^d)$ or $C_b^{l,m}((0,T)\times\mathbb{R}^d)$ are subspaces of C_b with k bounded derivatives or m bounded space and l bounded time derivatives.

By $P(\mathbb{R}^d)$ we denote the set of Borel probability measure on $\mathbb{R}^d$. The Kantorovich-Rubinstein distance $d_0(\mu_1,\mu_2)$ on the space $P(\mathbb{R}^d)$ is defined as

$$d_0(\mu_1,\mu_2) := \sup_{f\in \text{Lip}_{1,1}(\mathbb{R}^d)} \Big\{ \int_{\mathbb{R}^d} f(x)d(\mu_1-\mu_2)(x)\Big\}, \tag{4}$$

where $\text{Lip}_{1,1}(\mathbb{R}^d) = \Big\{ f : f \text{ is Lipschitz continuous and } \|f\|_\infty, \|Df\|_\infty \le 1\Big\}$. Convergence in d_0 is equivalent to weak convergence of measures (convergence in $(C_b)^*$), and hence tight subsets of $(\mathbf{P}, d_0)$ are precompact by Prokhorov's theorem. We let the space $C([0,T]; P(\mathbb{R}^d))$ be the set of $P(\mathbb{R}^d)$-valued functions on $[0,T]$. It is a metric space with the metric $\sup_{t\in[0,T]} d_0(\mu(t),\nu(t))$, and tight equicontinuous subsets are precompact by the Arzela-Ascoli and Prokhorov theorems.

2.2. Nonlocal operators

Under the Lévy condition

(L1): $\mu \ge 0$ is a Radon measure satisfying $\int_{\mathbb{R}^d} 1 \wedge |z|^2\, d\mu(z) < \infty$,

the operators $\mathcal{L}$ defined in (2) are in one to one correspondence with the generators of pure jump Lévy processes [3]. One example is the symmetric σ-stable processes and the fractional Laplacians,

$$-(-\Delta)^{\frac{\sigma}{2}}\phi(x) = \int_{\mathbb{R}^d} \Big[\phi(x+z) - \phi(x) - z\cdot D\phi(x)1_{|z|<1}\Big]\frac{c_{d,\sigma}\,dz}{|z|^{d+\sigma}}, \quad \sigma\in(0,2).$$

They are well-defined pointwise e.g. on functions in $C_b \cap C^2$ by Taylor's theorem and Fubini:

$$|\mathcal{L}\phi(x)| \le \frac{1}{2}\|D^2\phi\|_{C_b(B(x,1))}\int_{|z|<1}|z|^2 d\mu(z) + 2\|\phi\|_{C_b}\int_{|z|\ge 1} d\mu(z) \quad \text{for} \quad x\in\mathbb{R}^d.$$

Let $\sigma\in[1,2)$. With more precise upper bounds on the integrals of μ near the origin:

$$\text{There is } c>0 \text{ such that} \quad r^\sigma \int_{|z|<1}\frac{|z|^2}{r^2}\wedge 1\, d\mu(z) \le c \quad \text{for } r\in(0,1), \tag{5}$$

or equivalently, $r^{-2+\sigma}\int_{|z|<r}|z|^2 d\mu(z) + r^{-1+\sigma}\int_{r<|z|<1}|z| d\mu(z) + r^\sigma\int_{r<|z|<1} d\mu(z) \le c$ for $r\in(0,1)$, we can have interpolation estimates for the operators $\mathcal{L}$ in L^p.

Lemma 2.1. *(L^p-bounds). Assume (L1), (5) with $\sigma\in[1,2)$, and $u\in C_b^2$. Then for all $p\in[1,\infty]$, and $r\in(0,1]$,*

$$\|\mathcal{L}u\|_{L^p(\mathbb{R}^d)} \le C\Big(\|D^2u\|_{L^p} r^{2-\sigma} + \|Du\|_{L^p}\Gamma(\sigma,r) + \|u\|_{L^p}\mu(B_1^c)\Big) \tag{6}$$

where

$$\Gamma(\sigma,r) = \begin{cases} |\ln r|, & \sigma = 1,\\ r^{1-\sigma}-1, & 1<\sigma<2. \end{cases}$$

Proof. For $p\in[1,\infty)$ we split $\mathcal{L}u$ into three parts, $L_1 = \int_{B_r} u(x+z)-u(x)-Du(x)\cdot z\,d\mu(z)$, $L_2 = \int_{B_1\setminus B_r} u(x+z)-u(x)-Du(x)\cdot z\,d\mu(z)$, and $L_3 = \int_{\mathbb{R}^d\setminus B_1} u(x+z)-u(x)\,d\mu(z)$. Using Taylor expansions, Minkowski's integral inequality, and (5),

$$\|L_1\|_{L^p(\mathbb{R}^d)} \le \Big(\int_{\mathbb{R}^d}|D^2u(x)|^p\,dx\Big)^{1/p}\int_{B_r}|z|^2\,d\mu(z) \le C\|D^2u\|_{L^p(\mathbb{R}^d)} r^{2-\sigma},$$

$$\|L_2\|_{L^p(\mathbb{R}^d)} \le 2\Big(\int_{\mathbb{R}^d}|Du(x)|^p\,dx\Big)^{1/p}\int_{B_1\setminus B_r}|z|\,d\mu(z) \le C\|Du\|_{L^p(\mathbb{R}^d)}\Gamma(\sigma,r),$$

$$\|L_3\|_{L^p(\mathbb{R}^d)} \le 2\Big(\int_{\mathbb{R}^d}|u(x)|^p\,dx\Big)^{1/p}\Big(\int_{\mathbb{R}^d\setminus B_1}\Big)d\mu(z) \le 2\|u\|_{L^p(\mathbb{R}^d)}\mu(B_1^c).$$

Summing these estimates we obtain (2.1). The case $p=\infty$ is similar, so we omit it. □

Similar estimates are given e.g. in Section 2.5 in [21]. Note that assumption (5) holds for $-(-\Delta)^{\beta/2}$ for any $\beta\in(0,\sigma]\setminus\{1\}$ and is related to the order of $\mathcal{L}$.

Remark 2.2. (a) When μ is symmetric, $\int_{B_1\setminus B_r} Du(x)\cdot z\, d\mu(z)=0$,

$$\|L_2\|_{L^p} \le 2\|u\|_p \int\limits_{r<|z|<1} d\mu(z) \le C\|u\|_p r^{-\sigma},$$

and $\|\mathcal{L}u\|_{L^p(\mathbb{R}^d)} \le C\big(\|D^2u\|_{L^p} r^{2-\sigma} + \|u\|_{L^p} r^{-\sigma}\big)$. Minimizing w.r.t. r then yields

$$\|\mathcal{L}u\|_{L^p} \le C\|D^2u\|_p^{\sigma/2}\|u\|_p^{1-\sigma/2}.$$

This result holds for the fractional Laplacian $\mathcal{L}=(-\Delta)^{\sigma/2}$ when $\sigma\in(1,2)$.

(b) When $\sigma\in(0,1)$, a similar argument shows that

$$\|\mathcal{L}u\|_{L^p} \le C\big(\|Du\|_{L^p} r^{1-\sigma} + \|u\|_{L^p} r^{-\sigma}\big),$$

and we find that $\|(-\Delta)^{\sigma/2}u\|_{L^p(\mathbb{R}^d)} \le C\|Du\|_p^{\sigma}\|u\|_p^{1-\sigma}$ for $\sigma\in(0,1)$.

We define the adjoint of $\mathcal{L}$ in the usual way.

Definition 2.3. (Adjoint). The adjoint of $\mathcal{L}$ is the operator $\mathcal{L}^*$ such that

$$\langle \mathcal{L}f, g\rangle_{L^2(\mathbb{R}^d)} = \langle f, \mathcal{L}^*g\rangle_{L^2(\mathbb{R}^d)} \qquad \text{for all} \qquad f,g\in C_c^2(\mathbb{R}^d).$$

The $\mathcal{L}^*$ operator has the same form as $\mathcal{L}$, with the "antipodal" Lévy measure μ^*:

Lemma 2.4. *Assume (L1) holds. The adjoint operator $\mathcal{L}^*$ is given by*

$$\mathcal{L}^*u(x) = \int\limits_{\mathbb{R}^d} u(x+z) - u(x) - Du(x)\cdot z 1_{|z|<1}\, d\mu^*(z),$$

where $\mu^*(B)=\mu(-B)$ for all Borel sets $B\subset\mathbb{R}^d$.

This result is classical (see e.g. Section 2.4 in [21]). *Hence all assumptions and results in this paper for μ and $\mathcal{L}$ automatically also hold for μ^* and $\mathcal{L}^*$ (and vice versa).*

2.3. Moments of Lévy-measures, processes and FP equations

Consider the solution X_t of the SDE (3) (e.g. with $X_0=x\in\mathbb{R}^d$) and the corresponding FP equation for its probability distribution m, $m_t+\operatorname{div}(\alpha m)-\mathcal{L}^*m=0$. If $\alpha\in L^\infty$ and (L1) holds, then it follows that X_t (and m) has $s>0$ moments if and only if $\mu 1_{|z|>1}$ has s moments [3]:

$$E|X_t|^s = \int\limits_{\mathbb{R}^d} |x|^s m(dx,t) < \infty \quad \Longleftrightarrow \quad \int\limits_{|z|>1} |z|^s d\mu(z) < \infty.$$

The symmetric σ-stable processes have finite s-moments for any $s \in (0, \sigma)$. It is well-known that smoothing properties of $\mathcal{L}$ only depend on the (moment) properties of $\mu 1_{|z|<1}$, and hence is completely independent of the number of moments of $\mu 1_{|z|>1}$, X_t and $m(t)$. This fact is reflected in the ellipticity assumption (L2') in the next section, and follows e.g. from simple heat kernel considerations in section 4, see Remark 4.8.

In this paper we will be as general as possible and assume no explicit moment assumptions on $\mu 1_{|z|>1}$, X_t, and $m(t)$. The only condition we impose on $\mu 1_{|z|>1}$ is (L1).

Note however, that we will still always have some sort of generalized moments, but maybe not of power type, and these "moments" will be important for tightness and compactness for the FP equations. We refer to section 6 and Proposition 6.5 for more details.

3. Existence and uniqueness for fractional MFG systems

Here we state our assumptions and the existence and uniqueness results for classical solutions of the system (1) both with nonlocal and local couplings.

3.1. Assumptions on the fractional operator $\mathcal{L}$ in (2)

We assume (L1) and

(L2'): (Uniform ellipticity) There are constants $\sigma \in (1, 2)$ and $C > 0$ such that

$$\frac{1}{C}\frac{1}{|z|^{d+\sigma}} \le \frac{d\mu}{dz} \le C\frac{1}{|z|^{d+\sigma}} \quad \text{for} \quad |z| \le 1.$$

These assumptions are satisfied by generators $\mathcal{L}$ of pure jump processes whose infinite activity part is close to α-stable. But scale invariance is not required nor any restrictions on the tail of μ except for (L1). Some examples are α-stable processes, tempered α-stable processes, and the nonsymmetric CGMY process in Finance [18,3]. Note that the upper bound on $\frac{d\mu}{dz}$ implies that (5) holds. A much more general condition than (L2') is:

(L2): There is $\sigma \in (1, 2)$, such that
(i) μ satisfies the upper bound (5).
(ii) There is $\mathcal{K} > 0$ such that the heat kernels K_σ and K^*_σ of $\mathcal{L}$ and $\mathcal{L}^*$ satisfy for $K = K_\sigma, K^*_\sigma$: $K \ge 0$, $\|K(t, \cdot)\|_{L^1(\mathbb{R}^d)} = 1$, and

$$\|D^\beta K(t, \cdot)\|_{L^p(\mathbb{R}^d)} \le \mathcal{K} t^{-\frac{1}{\sigma}(|\beta|+(1-\frac{1}{p})d)} \quad \text{for } t \in (0, T)$$

and any $p \in [1, \infty)$ and multi-index $\beta \in \mathbb{N}_0^d$ where D is the gradient in $\mathbb{R}^d$.

The heat kernel is a transition probability/fundamental solution. Under (L2) Lévy measures need not be absolutely continuous, e.g. $\mathcal{L} = -\Big(-\frac{\partial^2}{\partial x_1^2}\Big)^{\sigma_1/2} - \cdots - \Big(-\frac{\partial^2}{\partial x_d^2}\Big)^{\sigma_d/2}$ for $\sigma_1, \ldots, \sigma_d \in (1, 2)$ satisfies (L2) with $\sigma = \min_i \sigma_i$ and $d\mu(z) = \sum_{i=1}^d \frac{dz_i}{|z_i|^{1+\sigma_i}} \Pi_{j\neq i}\delta_0(dz_j)$. See Section 4 for precise definitions, a proof that (L2') implies (L2), more examples and extensions.

In the local coupling case, we need in addition to (L2) also the following assumption:

(L3): Let the cone $\mathcal{C}_{\eta,r}(a) := \{z \in B_r : (1-\eta)|z||a| \leq |\langle a, z\rangle|\}$. There is $\beta \in (0,2)$ such that for every $a \in \mathbb{R}^d$ there exist $0 < \eta < 1$ and $C_\nu > 0$, and for all $r > 0$,

$$\int_{\mathcal{C}_{\eta,r}(a)} |z|^2 \nu(dz) \geq C_\nu \eta^{\frac{d-1}{2}} r^{2-\beta}.$$

This assumption is introduced in [5] to prove Lipschitz bounds for fractional HJB equations. It holds e.g. for fractional Laplacians [5, Example 1] and then also if the inequality of (L2') holds for all $z \in \mathbb{R}^d$. Since the assumption is in integral form, it also holds for non-absolutely continuous Lévy measures, spectrally one-sided processes, sums of operators etc.

3.2. Fractional MFGs with nonlocal coupling

We consider the MFG system

$$\begin{cases} -\partial_t u - \mathcal{L}u + H(x,u,Du) = F(x, m(t)) & \text{in} \quad (0,T)\times\mathbb{R}^d, \\ \partial_t m - \mathcal{L}^* m - div\left(m D_p H(x,u,Du)\right) = 0 & \text{in} \quad (0,T)\times\mathbb{R}^d, \\ m(x,0) = m_0(x), \qquad u(x,T) = G(x, m(T)) & \text{in} \quad \mathbb{R}^d, \end{cases} \tag{7}$$

where the functions $F, G : \mathbb{R}^d \times P\left(\mathbb{R}^d\right) \to \mathbb{R}$ are non-local coupling functions, and $H : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is the Hamiltonian. We impose fairly standard assumptions on the data and nonlinearities [33,9,1] (but note we use the metric d_0 and not Wasserstein-1):

(A1): There exists a $C_0 > 0$ such that for all $(x_1, m_1), (x_2, m_2) \in \mathbb{R}^d \times P\left(\mathbb{R}^d\right)$:

$$|F(x_1,m_1) - F(x_2,m_2)| + |G(x_1,m_1) - G(x_2,m_2)| \leq C_0(|x_1 - x_2| + d_0(m_1,m_2)).$$

(A2): There exist constants $C_F, C_G > 0$, such that

$$\sup_{m\in P(\mathbb{R}^d)} \|F(\cdot,m)\|_{C_b^2(\mathbb{R}^d)} \leq C_F \quad \text{and} \quad \sup_{m\in P(\mathbb{R}^d)} \|G(\cdot,m)\|_{W^{3,\infty}(\mathbb{R}^d)} \leq C_G.$$

(A3): For every $R > 0$ there is $C_R > 0$ such that for $x \in \mathbb{R}^d, u \in [-R, R], p \in B_R, \alpha \in \mathbb{N}_0^N$, $|\alpha| \leq 3$,

$$|D^\alpha H(x,u,p)| \leq C_R.$$

(A4): For every $R > 0$ there is $C_R > 0$ such that for $x, y \in \mathbb{R}^d, u \in [-R, R], p \in \mathbb{R}^d$:

$$|H(x,u,p) - H(y,u,p)| \leq C_R(|p|+1)|x-y|.$$

(A5): There exists $\gamma \in \mathbb{R}$ such that for all $x \in \mathbb{R}^d, u, v \in \mathbb{R}, u \leq v, p \in \mathbb{R}^d$,

$$H(x,v,p) - H(x,u,p) \geq \gamma(v-u).$$

(A6): $m_0 \in W^{2,\infty}\left(\mathbb{R}^d\right) \cap \mathbf{P}(\mathbb{R}^d)$.

Note that convexity or coercivity of H is not assumed at this point and that we identify probability measures and their density functions (see (A6)).

Definition 3.1. (Classical solution) A classical solution of (7) is a pair (u, m) such that (i) $u, m \in C(\mathbb{R}^d \times [0, T])$, (ii) $m \in C([0, T]; P(\mathbb{R}^d))$, (iii) $Du, D^2u, \mathcal{L}u, u_t, Dm, \mathcal{L}^*m, m_t \in C(\mathbb{R}^d \times (0, T))$, and (iv) (u, m) solves (7) at every point.

Theorem 3.2. *(Existence of classical solutions) Assume (L1), (L2), (A1)–(A6). Then there exists a classical solution (u, m) of (7) such that $u \in C_b^{1,3}((0, T) \times \mathbb{R}^d)$ and $m \in C_b^{1,2}((0, T) \times \mathbb{R}^d) \cap C([0, T]; P(\mathbb{R}^d))$.*

The proof will be given in Section 7. It is an adaptation of the fixed point argument of P.-L. Lions [33,9,1] and requires a series of a priori, regularity, and compactness estimates for fractional HJB and fractional FP equations given in Sections 5 and 6.

For uniqueness, we add the following assumptions:

(A7): F and G satisfy monotonicity conditions:

$$\int_{\mathbb{R}^d} (F(x, m_1) - F(x, m_2))\, d(m_1 - m_2)(x) \geq 0 \qquad \forall m_1, m_2 \in P(\mathbb{R}^d),$$

$$\int_{\mathbb{R}^d} (G(x, m_1) - G(x, m_2))\, d(m_1 - m_2)(x) \geq 0 \qquad \forall m_1, m_2 \in P(\mathbb{R}^d).$$

(A8): The Hamiltonian $H = H(x, p)$ and is uniformly convex with respect to p:

$$\exists C > 0, \quad \frac{1}{C} I_d \leq D_{pp}^2 H(x, p) \leq C I_d.$$

Theorem 3.3. *Assume (L1), (A1)-(A8). Then there is at most one classical solution of the MFG system (7).*

Since $\mathcal{L}$ and $\mathcal{L}^*$ are adjoint operators, the proof of uniqueness is essentially the same as the proof in the College de France lectures of P.-L. Lions [33,9,1]. For the readers convenience we give the proof in Appendix A.

Example 3.4. (a) $F(x, m) = (\rho * m)(x)$ satisfies (A1) and (A2) if $\rho \in C_b^2(\mathbb{R}^d)$.

(b) $F(x, m) = \int_{\mathbb{R}^d} \Phi(z, (\rho * m)(z))\rho(x - z)dz$ satisfies (A1) and (A2) if $\rho \in C_b^2$ and $\Phi \in C^1$.

(c) Both functions satisfy (A7) if $\rho \geq 0$ and Φ is nondecreasing in its second argument.

3.3. Fractional MFG with local coupling

We consider the MFG system

$$\begin{cases} -\partial_t u - \mathcal{L}u + H(x, Du) = f(x, m(t,x)) & \text{in } (0,T) \times \mathbb{R}^d \\ \partial_t m - \mathcal{L}^* m - div\left(m D_p H(x, Du)\right) = 0 & \text{in } (0,T) \times \mathbb{R}^d \\ m(0) = m_0, u(x,T) = g(x), \end{cases} \tag{8}$$

where the coupling term f are local and only depends on the value of m at (x,t). Again we impose fairly standard assumptions on f, g and H [33,9]:

(A2'): (Regularity) $f \in C^2(\mathbb{R}^d \times [0, +\infty))$ with $\|f(\cdot, k)\|_{C_b^2} \leq C_k$, and $g \in C_b^3(\mathbb{R}^d)$.

(A2"): (Uniform bound f) $\|f\|_{C_b} \leq K_f$ for $K_f > 0$.

(A3'): (Lipschitz bound H) $\|D_p H\|_\infty \leq L_H$ for $L_H > 0$.

Theorem 3.5. *Assume (L1)–(L3), (A3)-(A6), (A2'), and either (A2") or (A3'). Then there exists a classical solution (u,m) of (8) such that $u \in C_b^{1,3}((0,T) \times \mathbb{R}^d)$ and $m \in C_b^{1,2}((0,T) \times \mathbb{R}^d) \cap C([0,T]; P(\mathbb{R}^d))$.*

The proof of this result is given in Section 8. The idea is to approximate by a MFG system with nonlocal coupling and use the compactness and stability results to pass to the limit. These results rely on new regularity results. As opposed to the case of nonlocal coupling, it not enough to consider the HJB and FP equations separately, in this local coupling case, regularity has to be obtained directly for the coupled system. This requires the use of fractional regularity and bootstrap arguments.

For uniqueness we follow [33,9] and look at the more general MFG system

$$\begin{cases} -\partial_t u - \mathcal{L}u + H(x, Du, m) = 0 & \text{in } \mathbb{R}^d \times (0,T) \\ \partial_t m - \mathcal{L}^* m - div\left(m D_p H(x, Du(t,x), m)\right) = 0 & \text{in } \mathbb{R}^d \times (0,T) \\ m(0) = m_0\ ,\ u(x,T) = G(x), \end{cases} \tag{9}$$

where $H = H(x, p, m)$ is convex in p and

(A9): $$\begin{bmatrix} m\partial_{pp}^2 H & \frac{1}{2} m \partial_{pm}^2 H \\ \frac{1}{2} m \left(\partial_{pm}^2 H\right)^T & -\partial_m H \end{bmatrix} > 0 \quad \text{for all } (x, p, m) \text{ with } m > 0.$$

Note that when $H(x, p, m) = \tilde{H}(x, p) - F(x, m)$, we recover assumption (A8).

Theorem 3.6. *Assume (L1), (A9), and $H = H(x, p, m) \in C^2$. Then (8) has at most one classical solution.*

We skip the proof which in view of adjointness of $\mathcal{L}$ and $\mathcal{L}^*$ is the same as in [33,9]. The minor adaptations needed can be extracted from the uniqueness proof for nonlocal couplings given in Appendix A.

4. Fractional heat kernel estimates

Here we introduce fractional heat kernels and prove L^1-estimates of their spatial derivatives. These estimates are used for the regularity results of Sections 5, 6, and 8. The heat kernel of an elliptic operator $\mathcal{A}$ is the fundamental solution of $\partial_t u - \mathcal{A}u = 0$, or $u = \mathcal{F}^{-1}(e^{t\hat{\mathcal{A}}})$, where $\hat{\mathcal{A}}$ is the Fourier multiplier defined by $\mathcal{F}(\mathcal{A}u) = \hat{\mathcal{A}}\hat{u}$. Taking the Fourier transform of (2), a direct calculation (see [3]) shows that

$$\mathcal{F}\big(\mathcal{L}u\big) = \hat{\mathcal{L}}(\xi)\hat{u}(\xi),$$

where

$$\hat{\mathcal{L}}(\xi) = \int_{\mathbb{R}^d} \big(e^{ix\cdot\xi} - 1 - i\xi \cdot z 1_{|z|<1}\big)\, d\mu(z). \tag{10}$$

We can split $\hat{\mathcal{L}}$ into a singular and a non-singular part,

$$\hat{\mathcal{L}}(\xi) = \Bigg(\int_{|z|<1} + \int_{|z|\geq 1}\Bigg)\big(e^{ix\cdot\xi} - 1 - i\xi \cdot z 1_{|z|<1}\big)\, d\mu(z) = \hat{\mathcal{L}}_s(\xi) + \hat{\mathcal{L}}_n(\xi). \tag{11}$$

Note that since $\mu \geq 0$, $\mathrm{Re}\, \hat{\mathcal{L}} = \int \big(\cos(z\cdot\xi) - 1\big)\, d\mu \leq 0$.

We will need the heat kernels K_σ and $\tilde{K}_\sigma$ of $\mathcal{L}$ and $\mathcal{L}_s$:

$$K_\sigma(t,x) = \mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}(\cdot)}\big) \qquad \text{and} \qquad \tilde{K}_\sigma(t,x) = \mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}_s(\cdot)}\big). \tag{12}$$

By the Lévy-Kinchine theorem (Theorem 1.2.14 in [3]), K_σ and $\tilde{K}_\sigma$ are probability measures for $t > 0$:

$$K_\sigma, \tilde{K}_\sigma \geq 0 \qquad \text{and} \qquad \int_{\mathbb{R}^d} K_\sigma(x,t)\, dx = 1 = \int_{\mathbb{R}^d} \tilde{K}_\sigma(x,t)\, dx.$$

When (L2') holds, $\mathrm{Re}\, \hat{\mathcal{L}}$ and $\mathrm{Re}\, \hat{\mathcal{L}}_s \leq -c|\xi|^\sigma$ for $|\xi| \geq 1$, and K_σ and $\tilde{K}_\sigma$ are absolutely continuous since $|e^{t\hat{\mathcal{L}}(\cdot)}|$ decays exponentially at infinity. An immediate consequence of assumption (L2) is existence for the corresponding fractional heat equation.

Proposition 4.1. *Assume (L1), (L2), $u_0 \in L^\infty(\mathbb{R}^d)$, and let $u(t,x) = K_\sigma(t,\cdot) * u_0(x)$. Then $u \in C^\infty\big((0,T)\times\mathbb{R}^d\big)$ and u is a classical solution of*

$$\partial_t u - \mathcal{L}u = 0 \quad in \quad \mathbb{R}^d \times (0,T), \qquad u(0,x) = u_0(x) \quad in \quad \mathbb{R}^d.$$

We first show that sums of operators $\mathcal{L}_i$ satisfying (L1) and (L2) also satisfy (L1) and (L2). Let

$$\mathcal{L} = \mathcal{L}_1 + \cdots + \mathcal{L}_N \quad \text{where} \quad \mathcal{L}_i u(x) = \int_{Z_i} u(x+z) - u(x) - Du(x) \cdot z 1_{|z|<1} \, d\mu_i(z), \tag{13}$$

Z_i is a d_i-dimensional subspace, $\oplus_{i=1}^N Z_i = \mathbb{R}^d$, and $\mathcal{L}_i$ satisfy (L1) and (L2) in Z_i:

(L1"): (i) $Z_i \simeq \mathbb{R}^{d_i}$ is a subspace for $i = 1, \ldots, N$, and $\oplus_{i=1}^M Z_i = \mathbb{R}^d$ for $M \leq N$.
(ii) $\mu_i \geq 0$ is a Radon measure on Z_i satisfying $\int_{Z_i} 1 \wedge |z|^2 \, d\mu_i(z) < \infty$.
(L2"): (i) μ_i satisfy the upper bound (5) with $\sigma = \min_i \sigma_i$.
(ii) There are $\sigma_i \in (1, 2)$ and $c_i > 0$ such that the heat kernels K_i and K_i^* of $\mathcal{L}_i$ and $\mathcal{L}_i^*$ satisfy for $p \in [1, \infty)$, $\beta \in \mathbb{N}_0^{d_i}$, $i = 1, \ldots, M$, and $t \in (0, T)$,

$$\|D_{z_i}^\beta K_i(t, \cdot)\|_{L^p(Z_i)} + \|D_{z_i}^\beta K_i^*(t, \cdot)\|_{L^p(Z_i)} \leq c_i t^{-\frac{1}{\sigma_i}(|\beta| + (1-\frac{1}{p})d)}.$$

First observe that here $\mu = \sum_i \mu_i \delta_{0, Z_i^\perp}$ where $\delta_{0, Z_i^\perp}$ is the delta-measure in $Z_i^\perp$ centered at 0. It immediately follows that (L1") and (L2") imply (L1) and (L2) (i).

Theorem 4.2. *Assume (L1"), (L2") (ii), and $\mathcal{L}$ is defined in (13). Then the heat kernel K and K^* of $\mathcal{L}$ and $\mathcal{L}^*$ belongs to C^∞ and satisfy (L2) (ii) with $\sigma = \min_i \sigma_i$, i.e.*

$$\|D_x^\beta K(t, \cdot)\|_{L^p(\mathbb{R}^d)} + \|D_x^\beta K^*(t, \cdot)\|_{L^p(\mathbb{R}^d)} \leq c_{\beta, T} t^{-\frac{1}{\sigma_i}(|\beta| + (1-\frac{1}{p})d)} \quad \text{for} \quad t \in (0, T), \quad \beta \in \mathbb{N}_0^d.$$

Proof. First note that in this case $K(t) = \mathcal{F}^{-1}(e^{t\hat{\mathcal{L}}_1} \cdots e^{t\hat{\mathcal{L}}_N}) = \mathcal{K}_1(t) * \cdots * \mathcal{K}_N(t)$ where

$$\mathcal{K}_i(t) := \mathcal{F}_{\mathbb{R}^d}^{-1}(e^{t\hat{\mathcal{L}}_i}) = K_i(t)\delta_{0, Z_i^\perp}, \quad K_i(t) = \mathcal{F}_{Z_i}^{-1}(e^{t\hat{\mathcal{L}}_i}).$$

For $t \in (0, T)$, (L2") (ii) implies that

$$\|D_{z_i}^\beta \mathcal{K}_i(t)\|_{L^p(\mathbb{R}^d)} = \|D_{z_i}^\beta K_i(t, \cdot)\|_{L^p(Z_i)} \leq c_i t^{-\frac{1}{\sigma_i}(|\beta| + (1-\frac{1}{p})d)} \leq c_T t^{-\frac{1}{\sigma}(|\beta| + (1-\frac{1}{p})d)} \; (\sigma \leq \sigma_i)$$

for some constant $c_T > 0$. Since K_i is a probability measure by the Lévy-Kinchine theorem [3, Thm 1.2.14], $\|\mathcal{K}_j(t)\|_{L^1(\mathbb{R}^d)} = \|K_j(t)\|_{L^1(Z_j)} = 1$. By properties of mollifiers and Young's inequality for convolutions it then follows that

$$\|D_{z_i}^\beta K(t, \cdot)\|_{L^p} = \|\mathcal{K}_1 * \cdots * D_{z_i}^\beta \mathcal{K}_i * \cdots * \mathcal{K}_N\|_{L^p} \leq 1 \cdot \|D_{z_i}^\beta K_i\|_{L^p(Z_i)} \leq c_T t^{-\frac{1}{\sigma}(|\beta| + (1-\frac{1}{p})d)}.$$

Since $i = 1, \ldots, M$ was arbitrary and $\oplus_{i=1}^M Z_i = \mathbb{R}^d$, the proof for K is complete. The proof for K^* is similar. □

It is easy to check that (L2') implies (L2)(i). We then check that (L2') implies (L2)(ii).

Theorem 4.3. *Assume (L1), (L2'), and $\mathcal{L}$ is defined in (2). Then the heat kernels K and K^* of $\mathcal{L}$ and $\mathcal{L}^*$ belong to C^∞ and satisfies (L2)(ii): For $p \in [1, \infty)$, $\beta \in \mathbb{N}_0^d$,*

$$\|D_x^\beta K(t, \cdot)\|_{L^p(\mathbb{R}^d)} + \|D_x^\beta K^*(t, \cdot)\|_{L^p(\mathbb{R}^d)} \leq c_{\beta, T} t^{-\frac{1}{\sigma}(|\beta| + (1-\frac{1}{p})d)} \quad \text{for} \quad t \in (0, T).$$

Example 4.4. In view of Theorems 4.2 and 4.3, assumption (L2) is satisfied by e.g.

$$\mathcal{L}_1 = -(-\Delta_{\mathbb{R}^d})^{\sigma_1/2} - (-\Delta_{\mathbb{R}^d})^{\sigma_2/2},$$

$$\mathcal{L}_2 = -\Big(-\frac{\partial^2}{\partial x_1^2}\Big)^{\sigma_1/2} - \cdots - \Big(-\frac{\partial^2}{\partial x_d^2}\Big)^{\sigma_d/2},$$

$$\mathcal{L}_3 u(x) = \int_{\mathbb{R}} u(x+z) - u(x) - u'(x)z1_{|z|<1}\frac{Ce^{-Mz^+-Gz^-}}{|z|^{1+Y}},$$

where $C, G, M > 0$, $Y \in (0,2)$, [CGMY model in Finance]

$$\mathcal{L}_4 = \mathcal{L} + L \quad \text{where } \mathcal{L} \text{ satisfy (L2) and } L \text{ is any other Lévy operator.}$$

We can even take L to be any local Lévy operator (e.g. Δ) if we relax the definition of $\mathcal{L}_i$ to $\mathcal{L}_i u(x) = \mathrm{tr}[a_i D^2 u] + b_i \cdot Du + \int_{Z_i} u(x+z) - u(x) - Du(x)\cdot z1_{|z|<1}\, d\mu_i(s)$ for $a_i \geq 0$.

Remark 4.5. (a) (L2) holds also for very non-symmetric operators where μ has support in a cone in $\mathbb{R}^d$. Examples are Riesz-Feller operators like

$$\mathcal{L}_3 u(x) = \int_{z>0} u(x+z) - u(x) - u'(x)z1_{z<1}\frac{dz}{z^{1+\alpha}}, \quad \alpha \in (0,2).$$

We refer to [2] for results and discussion, see e.g. Lemma 2.1 (G7) and Proposition 2.3.
(b) More general conditions implying (L2) can be derived from the very general results on derivatives of heat semigroups in [36] and heat kernels in [25]. Such conditions could include more non absolutely continuous and non-symmetric Lévy measures.

We will now prove Theorem 4.3 and start by proving the result for $\tilde{K}_\sigma$, the kernel of $\mathcal{L}_s$.

Lemma 4.6. *Assume (L1) and (L2'). Then $\tilde{K}_\sigma \in C^\infty$, and for all $\beta \in \mathbb{N}_o^d$ and $p \in [1,\infty)$, there is $c > 0$ such that*

$$\|D_x^\beta \tilde{K}_\sigma(\cdot,t)\|_{L^p(\mathbb{R}^d)} \leq ct^{-\frac{1}{\sigma}(|\beta|+(1-\frac{1}{p})d)} \qquad \text{for all} \qquad t>0.$$

Remark 4.7. (a) When $p = 1$, the bound simplifies to $\|D_x^\beta \tilde{K}_\sigma(\cdot,t)\|_{L^1(\mathbb{R}^d)} \leq ct^{-\frac{|\beta|}{\sigma}}$.
(b) When $|\beta| = 1$, the bound is locally integrable in t when $1 \leq p < p_0 := \frac{d}{1+d-\sigma}$. Note that $p_0 > 1$.

Proof. We verify the conditions of Theorem 5.6 of [25]. By (L2'), assumption (5.5) in [25] holds with

$$\nu_0(|x|) = \begin{cases} \frac{1}{|x|^{d+\sigma}}, & |x| < 1, \\ 0, & |x| \geq 0. \end{cases}$$

Then we compute the integral h_0,

$$h_0(r) := \int_{\mathbb{R}^d} 1 \wedge \frac{|x|^2}{r^2} \nu_0(|x|)\, dx = \begin{cases} c_d(\frac{1}{2-\sigma} + \frac{1}{\sigma}) r^{-\sigma} - \frac{c_d}{\sigma}, & r < 1, \\ c_d \frac{1}{2-\sigma} r^{-2}, & r \geq 1, \end{cases}$$

where c_d is the area of the unit sphere. Note that h_0 is positive, strictly decreasing, and that $h_0(r) \leq \lambda^\sigma h_0(\lambda r)$ for $0 < \lambda \leq 1$ and every $r > 0$. Hence the scaling condition (5.6) in [25] holds with $C_{h_0} = 1$ for any $\theta_{h_0} > 0$. The inverse is given by

$$h_0^{-1}(\rho) = \begin{cases} \left(\frac{(2-\sigma)\rho}{c_d}\right)^{-\frac{1}{2}}, & \rho \leq \frac{c_d}{2-\sigma}, \\ \left(\frac{\rho}{c_d} + \frac{1}{\sigma}\right)^{-\frac{1}{\sigma}} \left(\frac{\sigma(2-\sigma)}{2}\right)^{-\frac{1}{\sigma}}, & \rho \geq \frac{c_d}{2-\sigma}. \end{cases}$$

In both cases $t \leq (2-\sigma)/c_d$ and $t \geq (2-\sigma)/c_d$, we then find that $h_0^{-1}(1/t) \leq (\tilde{c}t)^{1/\sigma}$, where $\tilde{c}$ only depends on σ and d.

At this point we can use Theorem 5.6 in [25] to get the following heat kernel bound:

$$\left|\partial_x^\beta p(t, x + t b_{[h_0^{-1}(1/t)]})\right| \leq C_0 [h_0^{-1}(1/t)]^{-|\beta|} Y_t(x) = C_{0,\sigma} t^{-\frac{|\beta|}{\sigma}} Y_t(x),$$

for any $t > 0$, where b_r does not depend on x,

$$Y_t(x) = [h_0^{-1}(1/t)]^{-d} \wedge \frac{t K_0(|x|)}{|x|^d},$$

and

$$K_0(r) := r^{-2} \int_{|x|<r} |x|^2 \nu_0(|x|) dx = \frac{c_d}{2-\sigma} \cdot \begin{Bmatrix} r^{-\sigma}, & r < 1 \\ r^{-2}, & r \geq 1 \end{Bmatrix} \leq \frac{c_d}{2-\sigma} r^{-\sigma}.$$

An integration in x then yields for $p \in [1, \infty)$,

$$\|\partial_x^\beta p(t, \cdot)\|_{L^p(\mathbb{R}^d)}^p \leq C_{0,\sigma}^p \tilde{c}^p t^{-\frac{p|\beta|}{\sigma}} \int_{R^d} Y_t(x)^p\, dx. \tag{14}$$

We compute $\|Y_t\|_{L^p(\mathbb{R}^d)}$. Since $h_0^{-1}(1/t) \leq \tilde{c} t^{1/\sigma}$ and $K_0(r) \leq \frac{c_d}{2-\sigma} r^{-\sigma}$, we can compute the minimum to find a constant $c_{\sigma,d} > 0$ such that

$$0 \leq Y_t(x) \leq \begin{cases} (\tilde{c}t)^{-d/\sigma}, & \text{for } |x| < c_{\sigma,d} t^{1/\sigma} \\ \frac{c_d}{2-\sigma} \frac{t}{|x|^{d+\sigma}}, & \text{otherwise.} \end{cases}$$

A direct computation then shows that

$$\int_{\mathbb{R}^d} Y_t(x)^p\, dx \leq c_{d,\sigma,p} t^{-\frac{(p-1)d}{\sigma}},$$

where $c_{d,\sigma,p} > 0$ is a constant not depending on t. Combining this estimate with (14) concludes the proof of the Lemma. □

Proof of Theorem 4.3. Result for K_σ follows by Lemma 4.6 and a simple computation:

$$\begin{aligned}\|D_x^\beta K_\sigma\|_{L^p} &= \|D_x^\beta \mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}_s}e^{t\hat{\mathcal{L}}_n}\big)\|_{L^p} = \Big\|\Big(D_x^\beta \mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}_s}\big)\Big) * \mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}_n}\big)\Big\|_{L^p}\\ &\le \|D_x^\beta \mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}_s}\big)\|_{L^p}\int_{\mathbb{R}^d}\mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}_n}\big) \le ct^{-\frac{1}{\sigma}(|\beta|+(1-\frac{1}{p})d)}\cdot 1.\end{aligned}$$

The last integral is 1 since $\mathcal{F}^{-1}\big(e^{t\hat{\mathcal{L}}_n}\big)$ is a probability by e.g. Theorem 1.2.14 in [3]. Since $\mathcal{L}^*$ is an operator of the same type as $\mathcal{L}$ with a Lévy measure μ^* also satisfying (L1) and (L2') (cf. Lemma 2.4), the computations above show that K_σ^* also satisfy the same bound as K_σ. □

Remark 4.8. From this proof it follows that the smoothing properties of $\mathcal{L}$ and K_σ are independent of $\hat{\mathcal{L}}_n$ and then also $\mu 1_{|z|>1}$.

By interpolation we obtain estimates for fractional derivatives of the heat kernel.

Proposition 4.9. *Assume (L1), (L2), $t \in [0, T]$, $s, \sigma \in (0, 2)$, and $|D|^s := (-\Delta)^{s/2}$. Then*

$$\||D|^s K_\sigma(t)\|_{L^1(\mathbb{R}^d)} \le ct^{-\frac{s}{\sigma}},$$

and if $s \in (0, 1)$, then

$$\||D|^s \partial_x K_\sigma(t)\|_{L^1(\mathbb{R}^d)} \le ct^{-\frac{s+1}{\sigma}}.$$

Proof. By Remark 2.2 (a) with $p = 1$ and (L2), we find that

$$\int ||D|^s K^\sigma(t)|\,dx \le c\|D^2 K^\sigma(t)\|_{L^1}^{\frac{s}{2}}\|K^\sigma\|_{L^1}^{1-\frac{s}{2}} \le \big(ct^{-\frac{2}{\sigma}}\big)^{s/2}1^{1-s/2} \le ct^{-\frac{s}{\sigma}}.$$

The proof of the second part follows in a similar way from Remark 2.2 (b). □

5. Fractional Hamilton-Jacobi-Bellman equations

Here we prove regularity and well-posedness for solutions of the fractional Hamilton-Jacobi equation. In our proof we use heat kernel estimates (Section 4), a Duhamel formula, and a fixed point argument as in [28,19]. The fractional Hamilton-Jacobi equation is given by

$$\begin{cases}\partial_t u - \mathcal{L}u + H(x, u, Du) = f(t, x) & \text{in } (0, T)\times\mathbb{R}^d,\\ u(0, x) = u_0(x) & \text{in } \mathbb{R}^d,\end{cases} \tag{15}$$

where f is the source term and u_0 initial condition. We assume

(B1): $u_0 \in C_b(\mathbb{R}^d)$ and $f \in C_b([0, T]\times\mathbb{R}^d)$.

(B2): There is an $L > 0$ such that for all $x, y \in \mathbb{R}^d$, $t \in [0, T]$,

$$|f(t,x) - f(t,y)| + |u_0(x) - u_0(y)| \leq L|x - y|.$$

assumptions (L1), (A3)–(A5), (B1)–(B2) implies that there exists a bounded x-Lipschitz continuous viscosity solution u of (15) (cf. e.g. [29,30,6,28]).

Theorem 5.1 *(Comparison principle). Assume (L1), (A3)–(A5), (B1)–(B2) and u, v are viscosity sub- and supersolutions of (15) with bounded continuous initial data u_0, v_0. If $u_0 \leq v_0$ in $\mathbb{R}^d$, then $u \leq v$ in $\mathbb{R}^d \times (0, T)$.*

Outline of proof. If u and v are uniformly continuous, then the proof is essentially the same as the proof of Theorem 2 in [28]. When u and v are not uniformly continuous, the limit (13) in [28] no longer holds because (in the notation of [28]) $\frac{|\bar{x}-\bar{y}|^2}{\varepsilon} \not\to 0$. However, this can be fixed under our assumptions, loosely speaking because we can remove all $O(\frac{|\bar{x}-\bar{y}|^2}{\varepsilon})$-terms before taking limits by modifying the test function. The modification consists in introducing an exponential factor in the quadratic term: $\frac{e^{Ct}}{\varepsilon}|x-y|^2$ for C large enough. □

Remark 5.2. We drop a complete proof here for two reasons: (i) it is long and rather standard, and (ii) we only apply the result in cases where u and v are uniformly continuous and an argument like in [28] is sufficient.

Theorem 5.3 *(Well-posedness). Assume (L1), (A3)–(A5), and (B1)–(B2).*
(a) There exists a (unique) bounded continuous viscosity solution u of (15) in $(0, T) \times \mathbb{R}^d$ such that $u(0, x) = u_0(x)$.

(b) $\|u\|_\infty \leq \|u_0\|_\infty + C_0 T$ where $C_0 := \|H(\cdot, 0, 0)\|_\infty + \|f\|_\infty$ is finite by (A3) and (B1).

(c) If also $u_0 \in W^{1,\infty}(\mathbb{R}^d)$, then

$$\|Du(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq M_T,$$

where $M_T = \mathrm{e}^{2C_R T}\left(\frac{1}{2}C_R + T^2\|D_x f\|_\infty^2 + \|Du_0\|_\infty^2\right)^{1/2}$, with C_R from (A4) and $R = \|u\|_\infty$.

Proof. The proof of (a) is quite standard and almost identical to the proof of Theorem 3 in [28]. Part (b) follows from comparison, Theorem 5.1, and the proof of part (c) is similar to the proof of Lemma 2 in [28]. □

Using parabolic regularity (in the form of (L3) [5]) and the method of Ishii-Lions, it is possible to obtain Lipschitz bounds that only depend on the C_b-norm of f:

Theorem 5.4. *Assume (L1), (L3), (A3)–(A5), $f \in C_b([0, T] \times \mathbb{R}^d)$ and $u_0 \in W^{1,\infty}(\mathbb{R}^d)$. Then the viscosity solution u of (15) is Lipschitz continuous in x and there is a constant $M > 0$ such that for all $t \in [0, T]$,*

$$\|Du(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq M,$$

where M depends on $\|u\|_\infty$, $\|f\|_\infty$, d, and the quantities in (A3)–(A5).

Proof. In the periodic case this result is a direct consequence of Corollary 7 in [5]. The original proof is for a right-hand side f not depending on t. For continuous $f = f(x,t)$, the proof is exactly the same. The result also holds in the whole space case, and this is explained in section 5.1 (see Theorem 6 for the stationary case). □

Similar parabolic results for the whole space are also given in [14]. To have classical solutions we make further regularity assumptions on the data:

(B3): $\|f(t,\cdot)\|_{C_b^2(\mathbb{R}^d)} \leq C$ for all $t \in [0,T]$.
(B4): $u_0 \in C_b^3(\mathbb{R}^d)$.

Note that f needs less spatial regularity than H in (A3).

Theorem 5.5 *(Classical solutions). Assume (L1)–(L2), (A3)–(A5), and (B1)–(B4). Then (15) has a unique classical solution u such that $\partial_t u, u, Du, \cdots, D^3u \in C_b\big((0,T)\times\mathbb{R}^d\big)$ with*

$$\|\partial_t u\|_{L^\infty} + \|u\|_{L^\infty} + \|Du\|_{L^\infty} + \ldots + \|D^3u\|_{L^\infty} \leq c,$$

where c is a constant depending only on σ, T, d, and quantities from (L1)–(B4).

To have space-time uniform continuity (and compactness) of derivatives, we assume:

(B5): There is a modulus of continuity ω_f such that for all $x, y \in \mathbb{R}^d$, $t, s \in [0,T]$,

$$|f(s,x) - f(t,y)| \leq \omega_f(|s-t| + |x-y|).$$

Theorem 5.6 *(Uniform continuity). Assume (L1)–(L2), (A3)–(A5), and (B1)–(B5). Then the unique classical solution u of (15) also satisfies*

$$\begin{aligned} &|u(t,x) - u(s,y)| + |Du(t,x) - Du(s,y)| + |D^2u(t,x) - D^2u(s,y)| \\ &+ |\partial_t u(t,x) - \partial_t u(s,y)| + |\mathcal{L}u(t,x) - \mathcal{L}u(s,y)| \leq \omega(|t-s| + |x-y|), \end{aligned} \tag{16}$$

where ω only depends on σ, T, d, and quantities from (L1)–(B5).

Remark 5.7. Imbert shows in [28] that when $\mathcal{L} = -(-\Delta)^{\sigma/2}$, $f \equiv 0$, and $u_0 \in W^{1,\infty}(\mathbb{R}^d)$, there exists a classical solution u such that $\|u\|_{C_b} + \|Du\|_{C_b} + \|t^{1/\sigma}D^2u\|_{C_b} \leq c$. He goes on to show that when $H = H(p) \in C^\infty$, then $u \in C^\infty$. In this paper we prove results for a much larger class of equations and nonlocal operators. Our results are also more precise: We need and prove time-space uniform continuity of all derivatives appearing in the equation, see Theorem 5.6. To do we need a finer analysis of the regularity in time. A final difference is that our estimates do not blow up as $t \to 0^+$. Note that it is easy to adapt our proofs and obtain even higher order regularity, e.g. treat the case $H = H(x,u,p) \in C^\infty$.

To prove Theorem 5.5 and 5.6, we first restrict ourselves to a short time interval.

5.1. Short time regularity by a Duhamel formula

Let K be the fractional heat kernel defined in (12). A solution v of (15) is formally given by the Duhamel formula

$$\begin{aligned} v(t,x) &= \psi(v)(t,x) \\ &:= K(t,\cdot) * v_0(\cdot)(x) - \int_0^t K(t-s,\cdot) * (H(s,\cdot,v(s,\cdot),Dv(s,\cdot)) - f(s,\cdot))(x)\,ds, \end{aligned} \tag{17}$$

where $*$ is convolution in $\mathbb{R}^d$. Note that solutions of this equation are fixed points of ψ.

Proposition 5.8 *(Spatial regularity). Assume (L1)–(L2), (A3)–(A5), (B1)–(B3), and $k \in \{2,3\}$. For $R_0 \geq 0$, let $R_1 = (1+\mathcal{K})R_0 + 1$ with $\mathcal{K}$ defined in (L2).*

(a) If $v_0 \in W^{k-1,\infty}(\mathbb{R}^d)$ with $\|v_0\|_{W^{k-1,\infty}} \leq R_0$, then there is $T_0 \in (0,T)$ such that ψ in (17) has a unique fixed point $v \in C_b^{k-1}([0,T_0]\times\mathbb{R}^d)$ with $t^{1/\sigma}D^k v \in C_b([0,T_0]\times\mathbb{R}^d)$ and

$$\|v\|_{L^\infty} + \cdots + \|D^{k-1}v\|_{L^\infty} + \|t^{1/\sigma}D^k v\|_{L^\infty} \leq R_1.$$

(b) If $v_0 \in W^{k,\infty}(\mathbb{R}^d)$ with $\|v_0\|_{W^{k,\infty}} \leq R_0$, then there is $T_0 \in (0,T)$ such that ψ in (17) has a unique fixed point $v \in C_b^k([0,T_0]\times\mathbb{R}^d)$ and

$$\|v\|_{L^\infty} + \cdots + \|D^k v\|_{L^\infty} \leq R_1.$$

In both cases T_0 only depends on σ and the quantities in (L1)–(B3).

Proof. (a) We will use the Banach fixed point theorem in the Banach (sub) space

$$X = \left\{v : v, \ldots, D^{k-1}v, t^{1/\sigma}D^k v \in C_b([0,T_0]\times\mathbb{R}^d) \text{ and } |||v|||_k \leq R_1\right\},$$

where $|||v|||_k = \|v\|_{k-1} + \sum_{|\beta|=k}\|t^{1/\sigma}D_x^\beta v\|_\infty$ and $\|v\|_k = \sum_{0\leq|\beta|\leq k}\|D_x^\beta v\|_\infty$.

Let $v \in X$. For $i = 1,\ldots,d$ and $\beta \in \mathbb{N}^d$, $|\beta| \leq k-2$,

$$\partial_x^\beta \partial_{x_i}\psi(v) = K(t) * \partial_x^\beta \partial_{x_i} v_0(x) - \int_0^t \partial_{x_i}K\big(t-s\big) * \partial_x^\beta\Big(H\big(\cdot,\cdot,v,Dv\big) - f\Big)(s,x)\,ds, \tag{18}$$

while for $|\beta| = k-1$,

$$\begin{aligned} t^{1/\sigma}\partial_x^\beta\partial_{x_i}\psi(v) &= t^{1/\sigma}\partial_{x_i}K(t) * \partial_x^\beta v_0(x) \\ &\quad - t^{1/\sigma}\int_0^t \partial_{x_i}K(t-s) * \partial_x^\beta\Big(H\big(\cdot,\cdot,v,Dv\big) - f\Big)(s,x)\,ds. \end{aligned} \tag{19}$$

If w and F are bounded functions, then $K(t,\cdot) * w$ and $\int_0^t \partial_x K(t-s,\cdot) * F(s,\cdot)\,ds$ are well-defined, bounded and continuous by (L2) and an argument like in the proof of [19, Proposition 3.1]. It follows by (A3) and (B3), that the derivatives of $\psi(v)$ in (18) and (19) are well-defined, bounded, and continuous. In particular by (L2), for $t \in (0,T)$,

$$\|t^{1/\sigma} \partial_{x_i} K(t) * \partial_x^\beta v_0\|_{C_b} \leq \mathcal{K} \|\partial_x^\beta v_0\|_{C_b}.$$

Let $u, v \in X$ and $t \in [0, T_0]$. By (A3) and (B3) there is a constant $C_{R_1} > 0$, such that

$$\left|\partial_x^\beta\left[H(s,x,u(s,x),Du(s,x))\right]\right| + \left|\partial_x^\beta f(s,x)\right| \leq \begin{cases} C_{R_1}, & 0 \leq |\beta| \leq k-2, \\ C_{R_1}\left(1+s^{-\frac{1}{\sigma}}\right), & |\beta| = k-1, \end{cases} \tag{20}$$

$$\left|\partial_x^\beta\left[H(s,x,u,Du)\right] - \partial_x^\beta\left[H(s,x,v,Dv)\right]\right| \leq \begin{cases} C_{R_1}\|u-v\|_{|\beta|+1}, & 0 \leq |\beta| \leq k-2, \\ C_{R_1}\left(1+s^{-\frac{1}{\sigma}}\right)|||u-v|||_3, & |\beta| = k-1. \end{cases} \tag{21}$$

By (L2) $\int_0^t \int_{\mathbb{R}^d} |K(t-s,x)|\,dx\,ds \leq T_0$, $\int_0^t \int_{\mathbb{R}^d} |\partial_{x_i} K(t-s,x)|\,dx\,ds \leq k(\sigma)\,T_0^{1-1/\sigma}$, and

$$\int_0^t s^{-1/\sigma} \int_{\mathbb{R}^d} |\partial_{x_i} K(t-s,x)|\,dx\,ds \leq \gamma(\sigma)\,T_0^{1-1/\sigma},$$

where $k(\sigma) = \mathcal{K}\frac{\sigma}{\sigma-1}$ and $\gamma(\sigma) = \mathcal{K}\int_0^1 (1-\tau)^{-1/\sigma}\,\tau^{-1/\sigma}\,d\tau$. From these considerations and Young's inequality for convolutions on (18) and (19), we compute the norm in X,

$$\begin{aligned}
&\|\psi(v)\|_\infty + \sum_{i=1}^d \Big(\|\partial_i \psi(v)\|_\infty + \sum_{1\leq|\beta|=k-2} \|\partial_x^\beta \partial_i \psi(v)\|_\infty + \sum_{|\beta|=k-1} \|t^{1/\sigma}\partial_x^\beta \partial_i \psi(v)\|_\infty\Big) \\
&\leq (1+\mathcal{K})\,R_0 \\
&+ \underbrace{C_{R_1}\Big(T_0 + \sum_{i=1}^d \Big(k(\sigma)\,T_0^{1-1/\sigma} + \sum_{1\leq|\beta|=k-2} k(\sigma)\,T_0^{1-1/\sigma} + \sum_{|\beta|=k-1} k(\sigma)\,T_0 + \gamma(\sigma) T_0^{1-1/\sigma}\Big)\Big)}_{=:c(T_0)}.
\end{aligned}$$

Taking $T_0 \in (0,T)$ such that $c(T_0) \leq 1/2$, ψ maps X into itself: By the definition of R_1,

$$|||\psi(v)|||_k \leq (1+\mathcal{K})R_0 + \frac{1}{2} \leq R_1.$$

It is also a contraction on X. By (21) and $\|u\|_1 \leq \|u\|_{k-1} \leq |||u|||_k$,

$$\begin{aligned}
&|||\psi(u) - \psi(v)|||_k \\
&\quad \leq C_{R_1}\Big(T_0\|u-v\|_1 + \sum_{i=1}^d \Big(k(\sigma)\,T_0^{1-1/\sigma}\|u-v\|_1 + \sum_{1\leq|\beta|\leq k-2} k(\sigma)\,T_0^{1-1/\sigma}\|u-v\|_{|\beta|+1}
\end{aligned}$$

$$+\sum_{|\beta|=k-1}\left(k(\sigma)T_0+\gamma(\sigma)T_0^{1-1/\sigma}\right)|||u-v|||_{|\beta|+1}\Big)\Big)$$

$$\leq c(T_0)|||u-v|||_k \leq \frac{1}{2}|||u-v|||_k.$$

An application of Banach's fixed point theorem in X now concludes the proof of part (a).

(b) We define the Banach (sub) space

$$X=\left\{v: v, Dv, \ldots, D^k v \in C_b((0,T_0)\times\mathbb{R}^d) \text{ and } \|v\|_k \leq R_1\right\},$$

where $\|v\|_k=\sum_{0\leq|\beta|\leq k}\|D_x^\beta v\|_\infty$. We use (18) with $|\beta|\leq k-1$, and only the first parts of (20) and (21). The rest of the proof is then similar to the proof of part (a). □

We proceed to prove time and mixed time-space regularity results. As a consequence, the solution of (17) is a classical solution of (15).

Proposition 5.9. *Assume $T_0>0$, (L1)–(L2), (A3)–(A5), (B1)–(B3), v satisfies (17), and $v, Dv, D^2v \in C_b([0,T_0]\times\mathbb{R}^d)$. Then*

(a) $\partial_t v \in C_b([0,T_0]\times\mathbb{R}^d)$ and $\|\partial_t v\|_\infty \leq c$, where c depends only on σ, T_0, d, the quantities in assumptions (L1)–(B3), and $\|D^k v\|_\infty$ for $k=0,1,2$.

Assume in addition $D^3 v\in C_b([0,T_0]\times\mathbb{R}^d)$.

(b) $v, Dv, \mathcal{L}v, D^2v \in UC([0,T_0]\times\mathbb{R}^d)$ with modulus $\omega(t-s,x-y)=C(|t-s|^{\frac{1}{2}}+|x-y|)$, where $C>0$ only depends on σ, T_0, d, the quantities in assumptions (L1)–(B3), and $\|D^k v\|_\infty$ for $k=0,1,2,3$.
(c) If also (B5), then $\partial_t v\in UC((0,T_0]\times\mathbb{R}^d)$, where the modulus only depends on T_0, σ, T_0, d, the quantities in assumptions (L1)–(B5), and the moduli of $v, Dv, \mathcal{L}v, D^2v$.

Corollary 5.10. *Under the assumptions of Proposition 5.9 (a), v is a classical solution of (15) on $(0,T_0)\times\mathbb{R}^d$.*

Follows by differentiating formula (17). To prove Theorem 5.9 we use the Duhamel formula

$$v(t,x)=K(t,\cdot)*v_0(\cdot)(x)-\int_0^t K(t-s,\cdot)*g(s,\cdot)(x)\,ds, \tag{22}$$

corresponding to the equation

$$\partial_t v(t,x)-\mathcal{L}v(t,x)+g(t,x)=0. \tag{23}$$

The following technical lemma is proved in Appendix B.

Lemma 5.11. *Assume (L1)–(L2), $g, \nabla g \in C_b\big((0,T)\times\mathbb{R}^d\big)$, and let*

$$\Phi(g)(t,x) = \int_0^t K(t-s,\cdot) * g(s,\cdot)(x)ds.$$

(a) $\Phi(g)(t,x)$ is C^1 w.r.t. $t \in (0,T)$ and $\partial_t \Phi(g)(t,x) = g(t,x) + \mathcal{L}[\Phi(g)](t,x)$.

(b) If $\beta \in (\sigma-1,1)$ and $g \in UC((0,T)\times\mathbb{R}^d)$, then

$$\begin{aligned}&|\partial_t\Phi(g)(t,x) - \partial_t\Phi(g)(s,y)| + |\mathcal{L}\Phi(g)(t,x) - \mathcal{L}\Phi(g)(s,y)|\\ &\le 2(1+c)\|g\|_{C_{b,t}C^1_{b,x}}|x-y|^{1-\beta}\\ &\quad + 2(1+c)\|g\|^{\beta}_{C_{b,t}C^1_{b,x}}\omega_g(|t-s|)^{1-\beta} + \tilde{c}\|g\|_{C_b}|t-s|^{\frac{\sigma-1}{\sigma}},\end{aligned}$$

where $c = \frac{\sigma}{\sigma-1}T^{\frac{\sigma-1}{\sigma}}\mathcal{K}\int_{|z|<1}|z|^{1+\beta}d\mu(z) + 4T\int_{|z|\ge 1}d\mu(z)$,

$$\tilde{c} = 2\frac{\sigma}{\sigma-1}\mathcal{K}\int_{|z|<1}|z|^{1+\beta}d\mu(z)K + 2T^{\frac{1}{\sigma}}\int_{|z|\ge 1}d\mu(z),$$

and $K = \max_{s,t\in[0,T]}\left|t^{\frac{\sigma-1}{\sigma}} - s^{\frac{\sigma-1}{\sigma}}\right|/|t-s|^{\frac{\sigma-1}{\sigma}}$.

Note that c, $\tilde{c}$, and K are finite. We have the following results for (22) and (23).

Lemma 5.12. *Assume (L1)–(L2), v satisfies (22), and $v, \nabla v, D^2 v, g, \nabla g \in C_b\left([0,T]\times\mathbb{R}^d\right)$.*
(a) $\partial_t v \in C_b\left((0,T)\times\mathbb{R}^d\right)$, and v solves equation (23) pointwise.
(b) If in addition $g \in UC([0,T]\times\mathbb{R}^d)$, then $\partial_t v$ and $\mathcal{L}v$ are uniformly continuous and for any $x, y \in \mathbb{R}^d$, $t, s \in [0,T]$, $k = 0, 1, 2$,

$$|\partial_t v(t,x) - \partial_t v(s,y)| + |\mathcal{L}v(t,x) - \mathcal{L}v(s,y)| \le \omega(|t-s| + |x-y|), \tag{24}$$

where ω only depends on ω_g, $\|g\|_\infty$, $\|g\|_{C_{b,t}C^1_{b,x}}$, $\|Dv_0\|_\infty$, $\|D^2 v_0\|_\infty$, σ, T, and μ.

Proof. (a) By the assumptions and Proposition 4.1 and Lemma 5.11 (a), we can differentiate the right hand side of (22). Differentiating and using the two results then leads to

$$\begin{aligned}\partial_t v &= \partial_t\left(K(t) * v_0\right) - \partial_t\int_0^t K(t-s) * g(s)\,ds\\ &= \mathcal{L}\left(K(t) * v_0\right) - g(t) - \mathcal{L}\int_0^t K(t-s) * g(s)\,ds\end{aligned}$$

$$= -g(t) + \mathcal{L}\left(K(t) * v_0 - \int_0^t K(t-s) * g(s)\, ds\right)$$
$$= -g(t) + \mathcal{L}v(t).$$

Thus we end up with (23) and the proof of (a) is complete.
(b) By (22), v is the sum of two convolution integrals. The regularity of the second integral follows from Lemma 5.11 (b). The regularity of the first integral follows by similar but much simpler arguments, this time with no derivatives on the kernel K (and hence two derivatives on v_0). We omit the details. □

Proof of Proposition 5.9. (a) In view of the assumptions, the result follows from Lemma 5.12(a) with $g(t,x) = H\big(x, v(t,x), Dv(t,x)\big) - f(t,x)$.
(b) By (a) and Corollary 5.10, v solve (15). We show that $D^2v \in UC([0,T] \times \mathbb{R}^d)$. Let $w = \partial^2_{x_i x_j} v$ and $w^\epsilon = w * \rho_\epsilon$ for a standard mollifier ρ_ϵ. Convolving (15) with ρ_ϵ and then differentiating twice ($\partial_{x_i}\partial_{x_j}$), we find that

$$\partial_t w^\epsilon - \mathcal{L}w^\epsilon + \partial^2_{x_i x_j}\big(H(t,x,v,Dv) * \rho_\epsilon\big) = \partial_{x_i x_j} f * \rho_\epsilon.$$

By Lemma 2.1 $\|\mathcal{L}w^\epsilon\|_\infty \le c\|w^\epsilon\|_{C^2_b}$, and then by properties of convolutions,

$$\|\mathcal{L}w^\epsilon\|_\infty \le c\sum_{k=2}^{4} \|D^k v^\epsilon\|_\infty \le \frac{c}{\epsilon}\|D\rho\|_{L^1}\|D^3 v\|_\infty + c(\|D^3 v\|_\infty + \|D^2 v\|_\infty).$$

It follows that $|\partial_t w^\epsilon| \le \frac{\tilde{c}}{\epsilon} + K$, where $\tilde{c} = c\|D\rho\|_{L^1}\|D^3 v\|_\infty$ and $K > 0$ is a constant only depending on $\|v\|_\infty, \|Dv\|_\infty, \|D^2v\|_\infty, \|D^3v\|_\infty, \|D^2 f\|_\infty$ and $C_R > 0$ from (A3), with $R = \max(\|v\|_\infty, \|Dv\|_\infty)$. We find that

$$\|w(t) - w(s)\|_\infty \le \|w^\epsilon(t) - w(t)\|_\infty + \|w^\epsilon(t) - w^\epsilon(s)\|_\infty + \|w^\epsilon(s) - w(s)\|_\infty$$
$$\le 2\|Dw\|_\infty \cdot \epsilon + \|\partial_t w^\epsilon\|_\infty |t-s| \le 2\|D^3 v\|_\infty \cdot \epsilon + (\frac{\tilde{c}}{\epsilon} + K)|t-s| \le C|t-s|^{\frac{1}{2}} + K|t-s|,$$

where we took $\epsilon = |t-s|^{\frac{1}{2}}$. Since w is bounded, this implies Hölder 1/2 regularity in time. The spatial continuity follows from $|w(t,x) - w(t,y)| \le \|D^3 v\|_\infty |x-y|$. In total, we get (recalling that $w = \partial_{x_i}\partial_{x_j} v$),

$$|D^2 v(s,x) - D^2 v(t,y)| \le C(|t-s|^{\frac{1}{2}} + |x-y|),$$

where $C > 0$ is only dependent on T_0, σ, T, d, the quantities in (L1)–(B3), and $\|D^k v\|_\infty$ for $k = 0, 1, 2, 3$. The results for v and Dv follow by simpler similar arguments. Since v, Dv and $D^2 v$ are uniformly continuous, by Taylor's theorem (as in the proof Lemma 2.1) $\mathcal{L}v$ is uniformly continuous with a modulus only depending on the moduli of v, Dv and $D^2 v$.
(c) By (B5) and the results from (b), $\partial_t v \in UC((0, T_0) \times \mathbb{R}^d)$ by the equation (15). □

5.2. Global regularity and proofs of Theorem 5.5 and 5.6

From the local in time estimates, we construct a classical solution u of (15) on the whole interval $(0, T) \times \mathbb{R}^d$. By Theorem 5.3, there is a unique viscosity solution u of (15) on $(0, T)$. To show that this solution is smooth, we proceed in steps.

1) By Lemma 5.8 (b) we find a $T_0 > 0$ and a unique solution v of (17) satisfying

$$v, Dv, D^2v, D^3v \in C_b([0, T_0] \times \mathbb{R}^d) \text{ and } v(0) = u_0,$$

and by Corollary 5.10, v is a classical solution of (15) on $(0, T_0)$. Since classical solutions are viscosity solutions, v coincides with the unique viscosity solution u on $(0, T_0)$.

2) Fix $t_0 \in [0, T)$ and take the value of the viscosity solution u of (15) as initial condition for (17) at $t = t_0$. Then $v(t_0, x) = u(t_0, x)$ and by Lemma 5.3,

$$\|v(t_0, \cdot)\|_{W^{1,\infty}(\mathbb{R}^d)} \leq M_T. \tag{25}$$

We apply Lemma 5.8 (a) with $k = 2$ (translate time $t \to t - t_0$, apply the theorem, and translate back) to obtain a $T_1 > 0$, independent of t_0, such that on

$$(t_0, t_0 + T_1),$$

we have a unique solution v of (17) satisfying $v, \nabla v, (t - t_0)^{1/\sigma} D^2v \in C_b$. Then

$$v, \nabla v, D^2v \in C_b\left((t_0 + \delta_1, t_0 + T_1) \times \mathbb{R}^d\right)$$

for any $\delta_1 \in (0, T_1)$. Let $\delta_1 \leq \frac{1}{4}\min(T_0, T_1)$, and take $v(t_0 + \delta_1, \cdot)$ as initial condition. By Lemma 5.8 (a) again we find a $T_2 > 0$ such that on the interval

$$(t_0 + \delta_1, t_0 + \delta_1 + T_2)$$

there exists a unique solution v of (17) such that for any $\delta_2 \in (0, T_2)$,

$$v, \nabla v, D^2v, t^{1/\sigma} D^3v \in C_b((t_0 + \delta_1 + \delta_2, t_0 + \delta_1 + T_2)).$$

We define $\tilde{T} := \min(T_0, T_1, T_2)$, and let $\delta_2 \leq \frac{1}{8}\tilde{T}$. Defining $\delta := \delta_1 + \delta_2 \leq \frac{1}{2}\tilde{T}$, we find that

$$v, Dv, \dots, D^3v \in C_b((t_0 + \delta, t_0 + \delta + \tilde{T}) \times \mathbb{R}^d).$$

By Proposition 5.9, $\partial_t v \in C_b$, and v is a classical solution of (15) on $(t_0 + \delta, t_0 + \delta + \tilde{T})$, therefore coinciding with u on this time interval. Note that $\tilde{T} > 0$ can be chosen independently of t_0 by (A3), (B3), (B4), and (25).

3) We cover all of $(0, T)$ by intervals from step 1) and 2), repeatedly taking $t_0 = 0, \frac{1}{2}\tilde{T}, \tilde{T}, \frac{3}{2}\tilde{T}, \dots, \frac{N-1}{2}\tilde{T}$ with $\frac{N}{2}\tilde{T} \geq T$. We then find that the viscosity solution u is a classical solution with bounded derivatives on $(0, T)$ and the proof of Theorem 5.5 is complete.

4) Theorem 5.6 follows from Theorem 5.5 and Proposition 5.9.

6. Fractional Fokker-Planck equations

Here we prove the existence of smooth solutions of the fractional Fokker-Planck equation, along with C_b, L^1, tightness, and time equicontinuity in L^1 a priori estimates. The equation is given by

$$\begin{cases} \partial_t m - \mathcal{L}^* m + div\,(b(t,x)m) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ m\,(0,x) = m_0\,(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{26}$$

where $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$, and $\mathcal{L}$ (and hence also $\mathcal{L}^*$) satisfies (L1), (L2).

We first show preservation of positivity and a first C_b-bound for bounded solutions.

Proposition 6.1. *Assume (L1) and $b, Db \in C_b((0,T)\times\mathbb{R}^d)$ and m is a bounded classical solution of (26).*

(a) If $m_0 \geq 0$, then $m(x,t) \geq 0$ for $(x,t) \in [0,T]\times\mathbb{R}^d$.

(b) If $m_0 \in C_b(\mathbb{R}^d)$, then $\|m(t,\cdot)\|_\infty \leq e^{\|(\operatorname{div} b)^+\|_\infty t}\|m_0\|_\infty$.

In fact this result also holds for bounded viscosity solutions, but this is not needed here. The result is an immediate consequence of the following lemma.

Lemma 6.2. *Assume (L1) and $b, Db \in C_b((0,T)\times\mathbb{R}^d)$ and m is a bounded classical subsolution of (26). Then for $t \in [0,T]$,*

$$\|m(t,\cdot)^+\|_\infty \leq e^{\|(\operatorname{div} b)^+\|_\infty t}\|m_0^+\|_\infty \tag{27}$$

Proof of Proposition 6.1. (a) Apply Lemma 6.2 on $-m$ (which still is a solution) and note that $(-m_0)^+ = 0$. (b) Apply Lemma 6.2 on m and $-m$. □

Proof of Lemma 6.2. In non-divergence form we get (the linear!) inequality

$$\partial_t m - \mathcal{L}^* m + b\cdot Dm + (\operatorname{div} b)\, m \leq 0,$$

with C_b coefficients by the assumptions. The proof is then completely standard and we only sketch the case that $\operatorname{div} b < 0$. Let

$$a = \sup_{(x,t)\in Q_T} m(x,t)^+ - \|m_0^+\|_\infty,$$

$\chi_R(x) = \chi(\frac{x}{R})$ where $0 \leq \chi \in C_c^2$ such that $\chi = 1$ in B_1 and $= 0$ in B_2^c, and

$$\Psi(x,t) = m(x,t) - \|m^+\|_\infty - at - \|m^+\|_\infty \chi_R(x).$$

We must show that $a \leq 0$. Assume by contradiction that $a > 0$. Then there exists a max point $(\bar{x},\bar{t})$ of Ψ such that $\bar{t} > 0$. At this max point $m > 0$ (since $a > 0$) and

$$m_t \geq a, \quad Dm = D\chi_R, \quad \text{and} \quad \mathcal{L}^* m \leq \mathcal{L}^* \chi_R.$$

Hence using the subsolution inequality at this point and $\operatorname{div} b < 0$, we find that

$$a \le m_t \le \mathcal{L}^* m + b \cdot Dm + (\operatorname{div} b)\, m \le \|m^+\|_\infty \Big(\mathcal{L}^* \chi_R + b \cdot D\chi_R\Big).$$

An easy computation shows that all χ_R-terms converge to zero as $R \to \infty$. Hence we pass to the limit and find that $a \le 0$, a contradiction to $a > 0$. The result follows. □

The Fokker-Planck equation (26) is mass and positivity preserving (it preserves pdfs) and therefore may preserve the L^1-norm in time. We will now prove a sequence of a priori estimates for L^1 solutions of (26), using a "very weak" (distributional) formulation of the equation.

Lemma 6.3. *Assume (L1), $m_0 \in L^1_{\mathrm{loc}}$, $b, Db \in C_b$, and m is a classical solution of (26) such that $m, Dm, D^2m \in C_b$. Then for every $\phi \in C_c^\infty(Q_T)$, $0 \le s < t \le T$,*

$$\int_{\mathbb{R}^d} m\phi(x,t)\,dx = \int_{\mathbb{R}^d} m\phi(x,s)\,dx + \int_s^t \int_{\mathbb{R}^d} m\big(\phi_t + \mathcal{L}\phi - b \cdot D\phi\big)(x,r)\,dx\,dr. \tag{28}$$

Proof. Note that $\mathcal{L}\phi \in C([0,T]; L^1(\mathbb{R}^d))$ by Lemma 2.1 with $p = 1$. Multiply (26) by ϕ, integrate in time and space, and integrate by parts. The proof is completely standard, after noting that $\int \mathcal{L}^* m\,\phi\,dx = \int m\,\mathcal{L}\phi\,dx$ in view of the assumptions of the Lemma. □

Remark 6.4. If in addition $m \in C([0,T]; L^1(\mathbb{R}^d))$, then a density argument shows that (28) holds for any $\phi \in C_b^\infty$.

Next we prove mass preservation, time-equicontinuity, and tightness for positive solutions in L^1. For tightness we need the following result:

Proposition 6.5. *Assume (L1) and $m_0 \in P(\mathbb{R}^d)$. There exists a function $0 \le \psi \in C^2(\mathbb{R}^d)$ with $\|D\psi\|_\infty, \|D^2\psi\|_\infty < \infty$, and $\lim_{|x|\to\infty} \psi(x) = \infty$, such that*

$$\int_{\mathbb{R}^d} \Psi(x)\, m_0(dx) < \infty, \qquad \int_{\mathbb{R}^d \setminus B_1} \Psi(x)\,\mu(dx) < \infty \tag{29}$$

Proof. We let $\mu_0 = \frac{\mu(dx)\mathbf{1}_{|x|\ge 1}}{\int_{B_1^c} \mu(dx)}$ and $\Pi = \{m_0, \mu_0\}$ and apply [15, Proposition 3.8]. □

Proposition 6.6. *Assume (L1), $m_0 \in C_b$, $b, Db \in C_b$, and m is a classical solution of (26) such that $m, Dm, D^2m \in C_b$. We also assume $m \in C([0,T]; L^1(\mathbb{R}^1))$, $m_0 \ge 0$, and $\int_{\mathbb{R}^d} m_0\,dx = 1$.*
(a) $m \ge 0$ and $\int_{\mathbb{R}^d} m(x,t)\,dx = 1$ for $t \in [0,T]$.
(b) There exists a constant $c_0 > 0$ such that

$$d_0(m(t), m(s)) \le c_0(1 + \|b\|_\infty)|t-s|^{\frac{1}{\sigma}} \qquad \forall s,t \in [0,T].$$

(c) For ψ defined in Proposition 6.5 there is $c > 0$ such that for $t \in [0, T]$,

$$\int_{\mathbb{R}^d} m(x,t)\psi(|x|)\,dx \le \int_{\mathbb{R}^d} m_0\psi(|x|)\,dx \tag{30}$$

$$+ 2\|\psi'\|_{C_b} + cT\|\psi'\|_{C_b}\Big(\|b\|_{C_b} + \int_{|z|<1} |z|^2 d\mu(z)\Big) + T\int_{|z|>1} \psi(|z|)\,d\mu(z).$$

Proof. (a) By Proposition 6.1, $m \ge 0$. Let $R > 1$ and $\chi_R(x) = \chi(\frac{x}{R})$ for $\chi \in C_c^\infty$ such that $0 \le \chi \le 1$ and $\chi = 1$ in B_1 and $= 0$ in B_2^c. We will apply Lemma 6.3 with $\phi = \chi_R$ and $s = 0$ and pass to the limit as $R \to \infty$. To do that, we write $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}^1 = \int_{|z|<1}\cdots + \int_{|z|>1}\cdots$, and note that by Lemma 2.1 with $p = \infty$ and $\mu(B_1^c) = 0$,

$$\|\mathcal{L}_1\chi_R\|_{C_b} \le C \inf_{r\in(0,1)}\Big(r^{2-\sigma}\frac{1}{R^2}\|D^2\chi\|_{C_b} + (r^{1-\sigma} - 1)\frac{1}{R}\|D\chi\|_{C_b}\Big) \le C\frac{1}{R^2}\|\chi\|_{C_b^2},$$

and then

$$\|\partial_t\chi_R + \mathcal{L}_1\chi_R - b\cdot D\chi_R\|_{C_b} \le \frac{1}{R}\Big(\|\chi\|_{C_b^2} + \|b\|_{C_b}\|D\chi\|_{C_b}\Big) \underset{R\to\infty}{\longrightarrow} 0.$$

Also note that $\|\mathcal{L}^1\phi_R\|_{C_b} \le 2\mu(B_1^c)$ and $\mathcal{L}^1\phi_R(x) \to 0$ for every $x \in \mathbb{R}^d$. Since $m \in C([0,T];L^1)$ by assumption, it follows by the dominated convergence theorem that,

$$\int_0^t\int_{R^d} m\,\mathcal{L}^1\chi_R\,dx\,dr \underset{R\to\infty}{\longrightarrow} 0.$$

Now we apply Lemma 6.3 with $\phi = \chi_R$ and $s = 0$ and pass to the limit in (28) as $R \to \infty$:

$$\lim_{R\to\infty}\int_{\mathbb{R}^d} m(x,t)\chi_R(x)\,dx = \lim_{R\to\infty}\int_{\mathbb{R}^d} m_0\chi_R(x)\,dx + 0.$$

The result now follows from the dominated convergence theorem since $\chi_R \to 1$ pointwise and $\int m_0\,dx = 1$.

(b) Fix a $\mathrm{Lip}_{1,1}$ function $\phi(x)$. For $\epsilon \in (0,1)$, let $\phi_\epsilon \in C_b^\infty$ be an approximation (e.g. by mollification) such that

$$\|\phi - \phi_\epsilon\|_{C_b} \le \epsilon\|D\phi\|_{C_b} \qquad \text{and} \qquad \|D^k\phi_\epsilon\|_{C_b} \le c\epsilon^{(k-1)^+}\|\phi\|_{C_b^1}, \quad k \ge 0. \tag{31}$$

Applying Lemma 6.3 and Remark 6.4 with $\phi = \phi_\epsilon(x)$, then leads to

$$\int_{\mathbb{R}^d}(m(x,t) - m(x,s))\phi_\epsilon(x)\,dx = \int_s^t\int_{\mathbb{R}^d} m\big(0 + \mathcal{L}\phi_\epsilon - b\cdot D\phi_\epsilon\big)(x,r)\,dx\,dr.$$

By Lemma 2.1 with $p = \infty$ and (31),

$$\|\mathcal{L}\phi_\epsilon\|_{C_b} \le c \inf_{r\in(0,1)} \Big(r^{2-\sigma}\|D^2\phi_\epsilon\|_{C_b} + r^{1-\sigma}\|D\phi_\epsilon\|_{C_b} + \|\phi_\epsilon\|_{C_b}\Big)$$
$$\le c \inf_{r\in(0,1)} \big(r^{2-\sigma}\frac{1}{\epsilon} + r^{1-\sigma} + 1\big)\|\phi\|_{C_b^1} \le C\epsilon^{1-\sigma}\|\phi\|_{C_b^1},$$

and hence

$$\int_{\mathbb{R}^d} (m(x,t) - m(x,s))\phi_\epsilon(x)\,dx \le C|t-s|\epsilon^{1-\sigma}(1 + \|b\|_{C_b})\|\phi\|_{C_b^1}\|m\|_{C(0,T;L^1)}.$$

Then by adding and subtracting $(m(x,t) - m(x,s))\phi_\epsilon(x)$ terms, we find that

$$\int_{\mathbb{R}^d} (m(x,t) - m(x,s))\phi(x)\,dx$$
$$\le \int_{\mathbb{R}^d} (m(x,t) - m(x,s))\phi_\epsilon(x)\,dx + 2\|m\|_{C(0,T;L^1)}\|\phi - \phi_\epsilon\|_{C_b}$$
$$\le C(|t-s|\epsilon^{1-\sigma} + \epsilon)(1 + \|b\|_{C_b})\|\phi\|_{C_b^1}\|m\|_{C(0,T;L^1)}.$$

Since $\|m\|_{C(0,T;L^1)} = 1$ by part (a), and $\|\phi\|_{C_b^1} \le 2$ for $\mathrm{Lip}_{1,1}$-functions, the result follows from the definition of the d_0 distance in (4) after a minimization in ϵ.
(c) Let $\psi_R(r) = \rho_1 * (\psi \wedge R)(r)$ for $r \ge 1$, where $0 \le \rho_1 \in C_c^\infty((-1,1))$ is symmetric and has $\int \rho_1\,dx = 1$ (a mollifier). We note that $\rho_1 * \psi \le \psi$ and that $\psi \wedge R$ is nondecreasing, concave, and $\nearrow \psi$. Standard arguments then show that $\psi_R \in C_b^\infty([1,\infty))$,

$$0 \le \psi_R \le R, \quad 0 \le \psi_R' \le \psi', \quad \psi_R'' \le 0, \quad \|\psi_R''\|_{C_b} \le \|\rho_1'\|_{L^1}\|\psi'\|_{C_b}, \tag{32}$$
$$\psi_R \nearrow \rho_1 * \psi\ (\le \psi) \quad \text{as} \quad R \to \infty. \tag{33}$$

The convergence as $R \to \infty$ is pointwise. We apply Lemma 6.3 and Remark 6.4 with

$$\phi(x,t) = \phi_R(x) := \psi_R(\sqrt{1 + |x|^2}).$$

Let $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}^1$ as in the proof of part (a), and note that (using also (32) and Lemma 2.1 with $r = 1$),

$$\|D\phi_R\|_{C_b} \le c\|\psi'\|_{C_b},\ \|D^2\phi_R\|_{C_b} \le c\|\rho_1'\|_{L^1}\|\psi'\|_{C_b},\ \|\mathcal{L}_1\phi_R\| \le c\|\rho_1'\|_{L^1}\|\psi'\|_{C_b}\int_{|z|<1} |z|^2\,d\mu.$$

Next since ψ_R is nonnegative, nondecreasing, and subadditive,[1] we observe that

[1] Nonnegative concave functions h on $[0,\infty)$ are subadditive: $h(a+b) \le h(a) + h(b)$ for $a, b \ge 0$.

$$|\psi_R(r)-\psi_R(s)|\leq\psi_R(r-s)\quad\text{for all}\quad r,s\geq 0.$$

Hence we find that

$$|\mathcal{L}^1\phi_R(x)|\leq\int\limits_{|z|>1}\left|\psi_R(\sqrt{1+|x+z|^2})-\psi_R(\sqrt{1+|x|^2})\right|d\mu(z)$$
$$\leq\int\limits_{|z|>1}\psi_R(|z|)\,d\mu(z)\leq\int\limits_{|z|>1}\psi(|z|)\,d\mu(z).$$

From the estimates above we conclude that

$$\left|\partial_t\phi_R+\mathcal{L}\phi_R-b\cdot D\phi_R\right|\leq c\|\psi'\|_{C_b}\Big(\|b\|_{C_b}+\|\rho_1'\|_{C_b}\int\limits_{|z|<1}|z|^2\,d\mu\Big)+\int\limits_{|z|>1}\psi(|z|)\,d\mu.$$

Inserting this estimate into (28) with $\phi=\phi_R$, along with $m\geq 0$, $\int m(x,t)\,dx=1$ (by part (a)), and $\phi_R(x)\leq\psi(\sqrt{1+|x|^2})$, we get

$$\int\limits_{\mathbb{R}^d}m(x,t)\phi_R(x)\,dx\leq\int\limits_{\mathbb{R}^d}m_0(x)\psi(\sqrt{1+|x|^2})\,dx$$
$$+Tc\|\psi'\|_{C_b}\Big(\|b\|_{C_b}+\|\rho_1'\|_{C_b}\int\limits_{|z|<1}|z|^2\,d\mu\Big)+T\int\limits_{|z|>1}\psi(|z|)\,d\mu.$$

By the monotone convergence theorem and (33),

$$\lim_{R\to\infty}\int\limits_{\mathbb{R}^d}m(x,t)\phi_R(x)\,dx=\int\limits_{\mathbb{R}^d}m(x,t)\,\rho_1*\psi(\sqrt{1+|x|^2})\,dx.$$

To conclude that (30) holds, we note that $\rho_1*\psi\geq\psi-\|\psi'\|_{C_b}$ and

$$\psi(|x|)\leq\psi(\sqrt{1+|x|^2})\leq\psi(|x|)+\|\psi'\|_{C_b}.$$

The proof of (c) is complete. □

Solutions in L^1 also have a better C_b bound than the one in Proposition 6.1. This bound is needed in the local coupling case – see Section 8.

Lemma 6.7. *Assume (L1), (L2) (ii), $b\in C_b$, $0\leq m_0\in C_b$, and $0\leq m\in C_b(Q_T)$ is a classical solution of (26). If $m\in C(0,T;L^1(\mathbb{R}^d))$, then there exists a constant $C>0$ only dependent on d,q,σ,T, such that for any $1<p<p_0:=\frac{d}{d+1-\sigma}$,*

$$\|m\|_{C_b}\leq 1\vee\Big[\|m_0\|_{C_b}+CT^{\frac{d-p(1+d-\sigma)}{p\sigma}}\|b\|_{C_b}\Big]^{\frac{p}{p-1}}.$$

Proof. (Inspired by [4, Proposition 2.2]) For any $y \in \mathbb{R}^d$, let $\phi(s,x) = K(t-s, y-x)$ where K is the heat kernel of Section 4. Then $\phi \geq 0$ is smooth, $\int_{\mathbb{R}^d} \phi(x,s)dx = 1$, and ϕ solves the backward heat equation

$$\begin{cases} -\partial_t \phi - \mathcal{L}\phi = 0, & s < t, \\ \phi(x,t) = \delta_y(x), \end{cases} \tag{34}$$

where the δ-measure δ_y has support in y. Multiply (26) by ϕ, integrate in time and space, and integrate by parts to get

$$\int m\phi(x,t)dx - \int m\phi(x-y,0)dx = \int_0^t \int m(x,s)[\phi_t + \mathcal{L}\phi - b\cdot D\phi](x-y,s)\,dx\,ds$$

or

$$m(y,t) = m * K(\cdot,t)(y) + \int_0^t \int (bm)(\cdot,s) * DK(\cdot,t-s)\,dx\,ds.$$

Then by the heat kernel estimates of (L2) (ii), $\|DK(s,\cdot)\|_{L^p} \leq Cs^{\frac{d-p(1+d)}{p\sigma}}$, the Hölder and Young's inequalities, the properties of K, and $\|m(\cdot,t)\|_{L^q}^q \leq \|m\|_{C_b}^{q-1}\|m(\cdot,t)\|_{L^1} = \|m\|_{C_b}^{q-1}$,

$$\begin{aligned} |m(y,t)| &\leq \|m_0\|_{C_b} + \|b\|_{C_b} \int_0^t \|DK(\cdot,t-s)\|_{L^p} \|m(\cdot,s)\|_{L^{p'}}\,dt \\ &\leq \|m_0\|_\infty + Ct^{\frac{d-p(1+d)+p\sigma}{p\sigma}} \|b\|_{C_b} \|m\|_{C_b}^{1-\frac{1}{p'}}, \end{aligned}$$

for $1 \leq p \leq \frac{d}{1+d-\sigma}$ where $\frac{1}{p} + \frac{1}{p'} = 1$. Since y is arbitrary, we get after taking the supremum and dividing both sides by $\|m\|_{C_b}^{\frac{1}{p}}$ that

$$\|m\|_{C_b} \leq 1 \vee \left[\|m_0\|_{C_b} + CT^{\frac{d-p(1+d-\sigma)}{p\sigma}} \|b\|_{C_b}\right]^{p'}.$$

This concludes the proof. □

Finally, we state the main result of this section, the existence of classical solutions of (26) that are positive and mass-preserving.

Proposition 6.8. *Assume (L1), (L2), $b, Db, D^2b \in C_b\big((0,T)\times\mathbb{R}^d\big)$, $0 \leq m_0 \in C_b^2(\mathbb{R}^d)$, and* $\int_{\mathbb{R}^d} m_0\,dx = 1$.

(a) There exists a unique classical solution m of (26) satisfying $m \geq 0$, $\int_{\mathbb{R}^d} m(x,t)\,dx = 1$ for $t \in [0,T]$, and

$$\|m\|_{L^\infty} + \|Dm\|_{L^\infty} + \|D^2 m\|_{L^\infty} + \|\partial_t m\|_{L^\infty} \le c,$$

where c is a constant depending only on σ, T, d, and $\|D^k b\|_\infty$ for $k = 0, 1, 2$.

(b) There exists a modulus $\tilde{\omega}$ only depending on $\|D^k m\|_\infty$, $\|D^k b\|_\infty$ for $k = 0, 1, 2$, and (L1), such that for $s, t \in [0, T]$ and $x, y \in \mathbb{R}^d$,

$$|m(t,x) - m(s,y)| + |Dm(t,x) - Dm(s,y)| \le \tilde{\omega}(|t-s| + |x-y|).$$

(c) If in addition $b, Db \in UC((0,T)\times\mathbb{R}^d)$, then there exists a modulus ω only depending on $\tilde{\omega}$, ω_b, ω_{Db}, $\|Db\|_\infty$, m_0, T, σ, and d, such that for $s, t \in [0, T]$ and $x, y \in \mathbb{R}^d$,

$$|\mathcal{L}^* m(x,t) - \mathcal{L}^* m(s,y)| + |\partial_t m(x,t) - \partial_t m(s,y)| \le \omega(|s-t| + |x-y|).$$

Proof. (a) The proof uses a Banach fixed point argument based on the Duhamel formula

$$\begin{aligned} m(t,x) &= \tilde{\psi}(m)(t,x) \\ &:= K^*(t,\cdot) * m_0(\cdot)(x) - \sum_{i=1}^{d} \int_0^t \partial_{x_i} K^*(t-s,\cdot) * (b_i m)(s,\cdot)\, ds, \end{aligned} \tag{35}$$

and is similar to the proof of Theorem 5.5. Here K^* is the heat kernel of $\mathcal{L}^*$. It is essentially a corollary to Proposition 5.1 in [19] (but in our case the we have more regular initial conditions and hence no blowup of norms when $t \to 0^+$).

Similar to the corresponding proof for the HJB equation, we first show short-time C^1-regularity using the Duhamel formula. Let $R_0 = 1 + \|m_0\|_\infty$, $R_1 = (2 + d\mathcal{K})R_0 + 1$, and the Banach (sub) space

$$X = \Big\{ m : m, t^{1/\sigma} Dm \in C_b\big((0,T_0)\times\mathbb{R}^d\big),\ m \in C([0,T]; L^1(\mathbb{R}^d)),\ \text{and } \|m\| \le R_1 \Big\}, \tag{36}$$

where $\|m\| = \|m\|_{C([0,T];L^1)} + \|m\|_\infty + \sum_{i=1}^d \|t^{1/\sigma}\partial_{x_i} m\|_\infty$. Then if $k(\sigma)$ and $\gamma(\sigma)$ are defined in the proof of Proposition 5.8 (a), we find from (35) that for $p \in \{1, \infty\}$,

$$\begin{aligned} \|\tilde{\psi}(m)(t,x)\|_{L^p} &\le \|K^*\|_{L^1}\|m_0\|_p + \sum_{i=1}^d \int_0^t \|\partial_{x_i} K^*(t-s,\cdot)\|_{L^1}\|b_i\|_\infty \|m(s)\|_p\, ds \\ &\le R_0 + d k(\sigma) T_0^{1-\frac{1}{\sigma}} \|b\|_\infty R_1, \end{aligned}$$

and

$$\begin{aligned} &|t^{1/\sigma}\partial_{x_j}\tilde{\psi}(m)(t,x)| \\ &\le t^{1/\sigma}\|\partial_{x_j} K^*\|_{L^1}\|m_0\|_\infty + \sum_{i=1}^d t^{1/\sigma} \int_0^t \|\partial_{x_i} K^*(t-s,\cdot)\|_{L^1} \|(m\partial_j b_i + b_i \partial_j m)\|_\infty ds \end{aligned}$$

$$\leq \mathcal{K}R_0 + \sum_{i=1}^{d} t^{1/\sigma} \int_0^t \mathcal{K}(t-s)^{-1/\sigma}\Big[\|m\|_\infty\|\partial_j b_i\|_\infty + s^{-1/\sigma}\|b_i\|_\infty\|s^{1/\sigma}\partial_j m\|_\infty\Big]ds$$
$$\leq \mathcal{K}R_0 + \Big[k(\sigma)T_0\|Db\|_\infty + \gamma(\sigma)T_0^{1-1/\sigma}\|b\|_\infty\Big]dR_1,$$

Computing the full norm, we get

$$\|\tilde{\psi}(m)\|$$
$$\leq (2+d\mathcal{K})R_0 + \underbrace{\Big[2dk(\sigma)T_0^{1-\frac{1}{\sigma}}\|b\|_\infty + d^2\Big[k(\sigma)T_0\|Db\|_\infty + \gamma(\sigma)T_0^{1-1/\sigma}\|b\|_\infty\Big]\Big]}_{:=c(T_0)}R_1.$$

We take $T_0 > 0$ so small that $c(T_0) \leq 1/2$. Then it follows that $\tilde{\psi}$ maps X into itself by the definition of R_1. It is also a contraction since for $m_1, m_2 \in X$, it easily follows that

$$\|\tilde{\psi}(m_1) - \tilde{\psi}(m_2)\| \leq c(T_0)\|m_1 - m_2\|.$$

An application of Banach's fixed point theorem in X then concludes the proof. Note that we only needed $m_0 \in C_b$ and $b, Db \in C_b$ to obtain the result.

We can now repeatedly differentiate the Duhamel formula (17) and use similar contraction arguments to conclude that if $b, Db, ..., D^k b \in C_b((0,T)\times\mathbb{R}^d)$, then there exists a solution $m \in X$ such that

$$D^2 m, ..., D^{k-1}m, t^{\frac{1}{\sigma}} D^k m \in C_b((0,T_0)\times\mathbb{R}^d) \quad \text{for} \quad T_0 > 0 \text{ sufficiently small.}$$

In a similar way as in Proposition 5.9 (a) and Corollary 5.10 for the HJB-equation, it now follows that m is a classical solution to (26). By Lemma 6.1 and Lemma 6.6 (a), we then have global in time bounds m in $C_b \cap C([0,T];L^1)$. We can therefore extend the local existence and the derivative estimates to all of $[0,T]$. The argument is very similar to the proof in Section 5.9 and we omit it. Finally, by Lemma 6.6 (a) again, we get that $m \geq 0$ and $\int_{\mathbb{R}^d} m(x,t)\,dx = 1$.

(b) Part (b) follows in a similar way as part (b) in Theorem 5.9. We omit the details.

(c) From part (a), (b), and the assumptions, the function $g(t,x) = \mathrm{div}(mb)$ satisfies $g, \nabla g \in C_b((0,T)\times\mathbb{R}^d)$ and $g \in UC((0,T)\times\mathbb{R}^d)$. Lemma 5.11 (b) (with K^* instead of K) then gives that $\partial_t\Phi(g), \mathcal{L}^*\Phi(g) \in UC((0,T)\times\mathbb{R}^d)$ with modulus ω only dependent on $\sigma, T, d, \|g\|_\infty, \|\nabla g\|_\infty$ and ω_g. A similar, but simpler argument shows that $\partial_t K_t^* * m_0 = \mathcal{L}^* K_t^* * m_0 \in UC((0,T)\times\mathbb{R}^d)$. Since $m = K_t^* * m_0 - \Phi(g)$, this concludes the proof. □

7. Existence for MFGs with nonlocal coupling – proof of Theorem 3.2

We adapt [33,9,1] and use the Schauder fixed point theorem. We work in $C([0,T],\mathbf{P}(\mathbb{R}^d))$ with metric $d(\mu,\nu) = \sup_{t\in[0,T]} d_0(\mu(t),\nu(t))$ and the subset

$$\mathcal{C} := \left\{ \mu \in C([0,T], \mathbf{P}(\mathbb{R}^d)) : \sup_{t\in[0,T]} \int_{\mathbb{R}^d} \psi(|x|)\, \mu(dx,t) \le C_1,\ \sup_{s\neq t} \frac{d_0(\mu(s),\mu(t))}{|s-t|^{\frac{1}{\sigma}}} \le C_2 \right\}, \tag{37}$$

where ψ is defined in Proposition 6.5 and the constants $C_1, C_2 > 0$ are to be determined. For $\mu \in \mathcal{C}$, define $S(\mu) := m$ where m is the classical solution of the fractional FPK equation

$$\begin{cases} \partial_t m - \mathcal{L}^* m - div\big(D_p H(x,u,Du)m\big) = 0, \\ \qquad\qquad m(0,\cdot) = m_0(\cdot), \end{cases} \tag{38}$$

and u is the classical solution of the fractional HJB equation

$$\begin{cases} -\partial_t u - \mathcal{L}u + H(x,u,Du) = F(x,\mu), \\ \qquad\qquad u(x,T) = G(x,\mu(T)). \end{cases} \tag{39}$$

Let $\mathcal{U} := \{u : u \text{ solves (39) for } \mu \in \mathcal{C}\}$ and $\mathcal{M} := \{m : m \text{ solves (38) for } u \in \mathcal{U}\}$.

1. ($\mathcal{C}$ convex, closed, compact). The subset $\mathcal{C}$ is convex and closed in $C([0,T], \mathbf{P}(\mathbb{R}^d))$ by standard arguments. It is compact by the Prokhorov and Arzèla-Ascoli theorems.

2. ($S : \mathcal{C} \to \mathcal{C}$ is well-defined). By (L1), (L2), (A1)–(A6), Theorem 5.5 and 5.6, there is a unique solution u of (39) with

$$\begin{aligned} &\|u\|_\infty, \|Du\|_\infty, \cdots, \|D^3u\|_\infty, \|\partial_t u\|_\infty \le U_1, \\ &\partial_t u, u, Du, D^2u, \mathcal{L}u \quad \text{equicontinuous with modulus } \omega, \end{aligned} \tag{40}$$

where U_1 depends on d, σ and the spatial regularity of F, G and H. The modulus ω depends in addition on C_2 in (37). By the uniform bound in (A2), U_1 is independent of μ. By Proposition 6.8 part (a)–(c), for any $u \in \mathcal{U}$ there is a unique m solving (38) such that

$$\begin{aligned} &\|m\|_\infty, \|Dm\|_\infty, \|D^2m\|_\infty, \|\partial_t m\|_\infty \le M_1, \\ &\partial_t m, m, Dm, \mathcal{L}^* m \quad \text{are equicontinuous with modulus } \bar{\omega}, \end{aligned} \tag{41}$$

where M_1 depends on U_1 and the local regularity of H but not on μ. The modulus $\bar{\omega}$ depends in addition on ω. By Lemma 6.6 (b)–(c),

$$\begin{aligned} &d_0(m(s), m(t)) \le c_0(1 + \|D_p H(\cdot, Du)\|_\infty)|s-t|^{\frac{1}{\sigma}}, \\ &\int_{\mathbb{R}^d} m(x,t)\psi(|x|)\,dx \le \int_{\mathbb{R}^d} m_0 \psi(|x|)\,dx \\ &\qquad + 2\|\psi'\|_{C_b} + cT\|\psi'\|_{C_b}\Big(\|D_p H(\cdot, Du)\|_{C_b} + \int_{|z|<1} |z|^2 d\mu(z)\Big) + T \int_{|z|>1} \psi(|z|)\, d\mu(z). \end{aligned}$$

By (40) and (A3), $\|D_pH(x, Du)\|_\infty \leq \tilde{C}$, where $\tilde{C}$ is independent of μ. Hence, we take $C_1 = \int_{\mathbb{R}^d} m_0\psi(|x|)\,dx + 2\|\psi'\|_{C_b} + cT\|\psi'\|_{C_b}\tilde{C} + \int_{|z|<1}|z|^2 d\mu(z)\big) + T\int_{|z|>1}\psi(|z|)\,d\mu(z)$, and $C_2 = c_0(1+\tilde{C})$ and get that S maps $\mathcal{C}$ into itself.

3. (S is continuous). We use the well-known result:

Lemma 7.1. *Let (X, d) a metric space, $K \subset\subset X$ a compact subset and $(x_n) \subset K$ a sequence such that all convergent subsequences have the same limit $x^* \in K$. Then $x_n \to x^*$.*

Define $X_1 := \{f : f, Df, D^2f, \partial_t f, \mathcal{L}f \in C_b\}$ and $X_2 := \{f : f, Df, \partial_t f, \mathcal{L}^* f \in C_b\}$, equipped with the metric of local uniform convergence, taken at all the derivatives. Then X_1 and X_2 are complete metric spaces. By (40), (41), Arzela-Ascoli, and a diagonal (covering) argument $\mathcal{U}$ and $\mathcal{M}$ are compact in X_1 and X_2, respectively.

Let $\mu_n \to \mu \in \mathcal{C}$, and let (u_n, m_n) be the corresponding solutions of (39) and (38). Take a convergent subsequence $(u_n) \supset u_{n_k} \to \tilde{u} \in \mathcal{U}$ and let $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}^1 = \big(\int_{|z|<1} + \int_{|z|\geq 1}\big)(\ldots)$. By uniform convergence $\mathcal{L}_1 u_{n_k}(t,x) \to \mathcal{L}_1\tilde{u}(t,x)$, and by dominated convergence $\mathcal{L}^1 u_{n_k}(t,x) \to \mathcal{L}^1\tilde{u}(t,x)$. By (A1), (A3) and for any $(t,x) \in (0,T)\times\mathbb{R}^d$:

$$
\begin{aligned}
&\big| -\partial_t\tilde{u}(t,x) - \mathcal{L}\tilde{u}(t,x) + H(x, D\tilde{u}(t,x)) - F(x,\mu)\big| \\
&\leq \big|\partial_t u_{n_k}(t,x) - \partial_t\tilde{u}(t,x)\big| + \big|\mathcal{L}u_{n_k}(t,x) - \mathcal{L}\tilde{u}(t,x)\big| \\
&+ \big|H(x, Du_{n_k}) - H(x, D\tilde{u})\big| + \big|F(x, \mu_{n_k}(t)) - F(x, \mu(t))\big| \\
&\to 0,
\end{aligned}
$$

and $\big|\tilde{u}(T,x) - G(x,\mu(T))\big| \leq |\tilde{u}(T,x) - u_{n_k}(T,x)| + \big|G(x,\mu_{n_k}(T)) - G(x,\mu(T))\big| \to 0$. This shows that $\tilde{u}$ solves (39) with μ as input. By uniqueness of the HJB equation, compactness of $\mathcal{U}$ in X_1, and Lemma 7.1, we conclude that $u_n \to u$ in X_1.

A similar argument shows that $m_n \to m \in X_2$. By compactness of $\mathcal{C}$ in part 2, uniqueness of solutions, and Lemma 7.1, we also find that $m_n \to m$ in $C([0,T], \mathbf{P}(\mathbb{R}^d))$. The map $S : \mathcal{C} \to \mathcal{C}$ is therefore continuous.

4. (Fixed point). By Schauder fixed point theorem there then exists a fixed point $S(m) = m$, and this fixed point is a classical solution of (7) and the proof of Theorem 3.2 is complete.

8. Existence for MFGs with local coupling – proof of Theorem 3.5

1. (Approximation) We follow Lions [33,9], approximating by a system with non-local coupling and passing to the limit. Let $\epsilon > 0$, $0 \leq \phi \in C_c^\infty$ with $\int_{\mathbb{R}^d}\phi = 1$, $\phi_\epsilon := \frac{1}{\epsilon^d}\phi(x/\epsilon)$, and for $\mu \in P(\mathbb{R}^d)$ let $F_\epsilon(x,\mu) := f(x, \mu * \phi_\epsilon(x))$. For each fixed $\epsilon > 0$, F_ϵ is a nonlocal coupling function satisfying (A1)–(A2), since $\|D^\beta(\mu * \phi_\epsilon)\|_\infty \leq \|\mu\|_1\|D^\beta\phi_\epsilon\|_\infty = \|D^\beta\phi_\epsilon\|_\infty$. assumptions (L1)–(L2), (A1)–(A6) then hold for the approximate system

$$
\begin{cases}
-\partial_t u_\epsilon - \mathcal{L}u_\epsilon + H(x, Du_\epsilon) = F_\epsilon(x, m_\epsilon(t)) & \text{in } (0,T)\times\mathbb{R}^d, \\
\partial_t m_\epsilon - \mathcal{L}^* m_\epsilon - div(m_\epsilon D_p H(x, Du_\epsilon)) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\
m(0) = m_0, \qquad u(x,T) = g(x),
\end{cases}
\tag{42}
$$

and by Theorem 3.2 there exists a classical solution (u_ϵ, m_ϵ) of this system.

2. (Uniform bounds) Since either (A3') or (A2") holds, $F_\epsilon(x, m_\epsilon(t))$ is uniformly bounded in ϵ. In the case of (A3') this follows from Lemma 6.7 and the estimate

$$\|m_\epsilon\|_{C_b} \leq 1 \vee \left[\|m_0\|_{C_b} + CT^{\frac{d-p(1+d-\sigma)}{p\sigma}}\|D_pH(\cdot, Du_\epsilon)\|_\infty\right]^{\frac{p}{p-1}} \leq K \tag{43}$$

for K independent of ϵ. By Theorem 5.3 (b) and (A3) we then have

$$\|u_\epsilon\|_\infty \leq \|g\|_\infty + (T-t)(\|F_\epsilon(\cdot, m_\epsilon(t))\|_\infty + \|H(\cdot,0)\|_\infty) \leq \tilde{K} \tag{44}$$

for $\tilde{K} > 0$ independent of ϵ, and since F_ϵ is also continuous, by Theorem 5.4

$$\|Du_\epsilon\|_\infty \leq C \tag{45}$$

for $C \geq 0$ independent of ϵ (C depends on F_ϵ only through its C_b-norm). Under (A3') m is bounded and satisfies (43), and this is still true if (A3') is replaced by (A2") in view of the uniform bound on Du_ϵ in (45).

3. (Improvement of regularity) The Duhamel formulas for m_ϵ and Du_ϵ are given by

$$m_\epsilon(t) = K_\sigma^*(t) * m_0 - \sum_{i=1}^d \int_0^t \partial_i K_\sigma^*(t-s) * m_\epsilon[D_pH(\cdot, Du_\epsilon(s))]_i ds, \tag{46}$$

$$Du_\epsilon(t) = K_\sigma(t) * Du_0 - \int_0^t D_xK_\sigma(t-s) * (H(\cdot, Du_\epsilon(s)) - F_\epsilon(\cdot, m_\epsilon(s,\cdot)))ds, \tag{47}$$

where $K_\sigma(t) = K_\sigma(t,x)$ and $K_\sigma^*(t) = K_\sigma^*(t,x)$ are the fractional heat kernels in $\mathbb{R}^d$ corresponding to $\mathcal{L}$ and $\mathcal{L}^*$. Fractional differentiations of these will lead to improved regularity.

Assume that for $k \in \{0,1,2\}$ and $\alpha \in [0,1)$, there is $C \geq 0$ independent of ϵ such that for all $t \in [0,T]$,

$$\|m_\epsilon(t)\|_{C^{k,\alpha}(\mathbb{R}^d)} + \|Du_\epsilon(t)\|_{C^{k,\alpha}(\mathbb{R}^d)} \leq C. \tag{48}$$

We will show that for any $\delta \in (0,\alpha)$ and $s \in (0, \sigma-1)$ there is $\tilde{C} \geq 0$ independent of ϵ and t such that

$$\begin{cases} \|m_\epsilon(t)\|_{C^{k,s+\alpha-\delta}(\mathbb{R}^d)} + \|Du_\epsilon(t)\|_{C^{k,s+\alpha-\delta}(\mathbb{R}^d)} \leq \tilde{C}, & \text{for } s+\alpha-\delta \leq 1, \\ \|m_\epsilon(t)\|_{C^{k+1,s+\alpha-\delta-1}(\mathbb{R}^d)} + \|Du_\epsilon(t)\|_{C^{k+1,s+\alpha-\delta-1}(\mathbb{R}^d)} \leq \tilde{C}, & \text{for } s+\alpha-\delta > 1. \end{cases} \tag{49}$$

Assume first $\alpha \in (0,1)$ and consider the m_ϵ-estimate. When (48) holds, then $m_\epsilon D_pH(x, Du_\epsilon) \in C^{k,\alpha}(\mathbb{R}^d)$ by the chain rule and (A3), and $|D|^{\alpha-\delta}D^k\big(m_\epsilon D_pH(x,Du_\epsilon)\big) \in C_b^{0,\delta}(\mathbb{R}^d)$ for $\delta \in (0,\alpha)$ by [37, Proposition 2.7]. Let $s \in (0,\sigma-1)$ and apply $|D|^s|D|^{\alpha-\delta}D^k$ to (46),

$$|D|^s|D|^{\alpha-\delta}D^k m_\epsilon = K_\sigma^*(t) * |D|^{s+\alpha-\delta}D^k m_0 - \sum_{i=1}^d \int_0^t |D|^s DK_\sigma^*(t-s) * |D|^{\alpha-\delta}D^k\big[m_\epsilon D_p H(\cdot, Du_\epsilon)\big]_i ds.$$

By Young's inequality and Proposition 4.9 (heat kernel estimates),

$$\||D|^{s+\alpha-\delta}D^k m_\epsilon\|_\infty \leq \||D|^{s+\alpha-\delta}D^k m_0\|_\infty + c\frac{T^{1-\frac{1+s}{\sigma}}}{1-\frac{1+s}{\sigma}}\||D|^{\alpha-\delta}D^k(m_\epsilon D_p H(\cdot, Du_\epsilon))\|_\infty,$$

and taking $\delta < \alpha/2$, we get uniform in ϵ Hölder estimates by [37, Proposition 2.9],

$$m_\epsilon(t) \in \begin{cases} C_b^{k,s+\alpha-2\delta}(\mathbb{R}^d), & \text{for } s+\alpha-2\delta \leq 1, \\ C_b^{k+1,s+\alpha-2\delta-1}(\mathbb{R}^d), & \text{for } s+\alpha-2\delta > 1. \end{cases}$$

The case $\alpha = 0$ follows in a similar but more direct way differentiating (46) by $|D|^s D^k$ instead of $|D|^s|D|^{\alpha-\delta}D^k$ as above. The estimates on Du_ϵ follow similarly.

4. (Iteration and compactness) Starting from (43), (44), and (A2') and (A3), we iterate using (49) to find that

$$\|u_\epsilon(t)\|_{C_b^3(\mathbb{R}^d)} + \|m_\epsilon(t)\|_{C_b^2(\mathbb{R}^d)} \leq C$$

independent of ϵ and $t \in [0, T]$. By Proposition 5.9 and Proposition 6.8, we then find that

$$\|\partial_t u_\epsilon\|_\infty \leq U \quad \text{and} \quad \partial_t u_\epsilon, u_\epsilon, Du_\epsilon, D^2u_\epsilon, \mathcal{L}u_\epsilon \quad \text{equicontinuous with modulus } \omega,$$
$$\|\partial_t m_\epsilon\|_\infty \leq M \quad \text{and} \quad \partial_t m_\epsilon, m_\epsilon, Dm_\epsilon, \mathcal{L}^* m_\epsilon \quad \text{equicontinuous with modulus } \bar{\omega},$$

where U, ω, M and $\bar{\omega}$ are independent of ϵ. As in the proof of Theorem 3.2, these bounds imply compactness of (m_ϵ, u_ϵ) in $X_1 \times X_2$ (see below Lemma 7.1 for the definitions).

5. (Passing to the limit) We extract a convergent subsequence, $(u_{\epsilon_k}, m_{\epsilon_k}) \to (u, m)$ in $X_1 \times X_2$. By a direct calculation the limit (u, m) solves equation (8). This concludes the proof of Theorem 3.5.

Appendix A. Uniqueness of solutions of MFGs – proof of Theorem 3.3

The proof of uniqueness is essentially the same as the proof in the College de France lectures of P.-L. Lions [33,9]. Let (u_1, m_1) and (u_2, m_2) be two classical solutions, and set $\tilde{u} = u_1 - u_2$ and $\tilde{m} = m_1 - m_2$. By (7) and integration by parts,

$$\frac{d}{dt}\int_{\mathbb{R}^d} \tilde{u}\tilde{m}\, dx = \int_{\mathbb{R}^d} \frac{\partial}{\partial t}(\tilde{u}\tilde{m})\, dx = \int_{\mathbb{R}^d} (\partial_t \tilde{u})\,\tilde{m} + \tilde{u}\,(\partial_t \tilde{m})\, dx$$
$$= \int_{\mathbb{R}^d} \Big[\Big(-\mathcal{L}\tilde{u} + H(x, Du_1) - H(x, Du_2) - F(x, m_1) + F(x, m_2)\Big)\tilde{m}$$

$$+\tilde{u}\mathcal{L}^*\tilde{m} - \langle D\tilde{u}, m_1 D_p H(x, Du_1) - m_2 D_p H(x, Du_2)\rangle \Big] dx.$$

By the definition of the adjoint, $\int_{\mathbb{R}^d} (\mathcal{L}\tilde{u})\,\tilde{m} - \tilde{u}\,(\mathcal{L}^*\tilde{m})\,dx = 0$, and from (A7) we get

$$\int_{\mathbb{R}^d} (-F(x, m_1) + F(x, m_2))\, d(m_1 - m_2)(x) \ge 0 \qquad \forall m_1, m_2 \in P(\mathbb{R}^d).$$

For the remaining terms on the right hand side, we use a Taylor expansion and (A8),

$$\begin{aligned}&\int_{\mathbb{R}^d} \Big[-m_1\Big(H(x, Du_1) - H(x, Du_2) - \langle D_p H(x, Du_1), Du_2 - Du_1\rangle\Big)\\ &\qquad - m_2\Big(H(x, Du_2) - H(x, Du_1) - \langle D_p H(x, Du_2), Du_1 - Du_2\rangle\Big)\Big] dx\\ &\le -\int_{\mathbb{R}^d} \frac{m_1 + m_2}{2C} |Du_2 - Du_1|^2\, dx.\end{aligned}$$

Integrating from 0 to T, using the fact that $\tilde{m}(t=0) = 0$ and $\tilde{u}(t=T) = G(x, m_1(T)) - G(x, m_2(T))$,

$$\int_0^T \frac{d}{dt} \int_{\mathbb{R}^d} \tilde{u}\tilde{m}\, dx\, dt = \int_{\mathbb{R}^d} (G(x, m_1(T)) - G(x, m_2(T)))\,(m_1(x, T) - m_2(x, T))\, dx \ge 0,$$

where we used (A7) again. Combining all the estimates we find that

$$0 \le -\int_0^T \int_{\mathbb{R}^d} \frac{m_1 + m_2}{2C} |Du_1 - Du_2|^2\, dx\, dt$$

Hence since the integrand is nonnegative it must be zero and $Du_1 = Du_2$ on the set $\{m_1 > 0\} \cup \{m_2 > 0\}$. This means that m_1 and m_2 solve the same equation (the divergence terms are the same) and hence are equal by uniqueness. Then also u_1 and u_2 solve the same equation and $u_1 = u_2$ by standard uniqueness for nonlocal HJB equations (see e.g. [29]). The proof is complete.

Appendix B. Proof of Lemma 5.11

a) The proof is exactly the same as in [28]. The difference is that f only needs to be C^1 in space, since $D_x K$ is integrable in t.

b) *Part 1:* Uniform continuity in x for $\mathcal{L}\Phi(f)$ and $\partial_t \Phi(f)$. By the definition of $\mathcal{L}$,

$$\mathcal{L}[\Phi(f)](t, x) = \int_0^t \mathcal{L}K(t - s, \cdot) * f(s, \cdot)(x)\, ds$$

$$= \int_0^t \int_{\mathbb{R}^d} \Big[\int_{\mathbb{R}^d} K(t-s, y+z) - K(t-s, y) - \nabla_x K(t-s, y) \cdot z 1_{|z|<1} d\mu(z) \Big] f(s, x-y) dy ds$$

$$= \int_0^t \int_{\mathbb{R}^d} \int_{|z|<1} \Big(\cdots \Big) + \int_0^t \int_{\mathbb{R}^d} \int_{|z|>1} \Big(\cdots \Big) =: I_1(t,x) + I_2(t,x).$$

After a change of variables and $\|K(t,\cdot)\|_{L^1} = 1$,

$$|I_2(t,x_1) - I_2(t,x_2)| \leq \int_0^t \int_{|z|\geq 1} \int_{\mathbb{R}^d} K(t-s,y) \Big[f(s, x_1 - y + z) - f(s, x_1 - y) - f(s, x_2 - y + z) + f(s, x_2 - y) \Big] dy d\mu(z) ds$$

$$\leq 2t \|f\|_{C_{b,t}C^1_{b,x}} |x_1 - x_2| \int_{|z|\geq 1} d\mu(z).$$

Then since and $\|I_2(t,\cdot)\|_{C_b} \leq 2t\|f\|_{C_{b,t}C^1_{b,x}} \int_{|z|\geq 1} d\mu(z)$,

$$|I_2(t,x_1) - I_2(t,x_2)| \leq (2\|I_2(t,\cdot)\|_{C_b})^{\beta} |I_2(t,x_2) - I_2(t,x_2)|^{1-\beta}$$
$$\leq 4t\|f\|_{C_{b,t}C^1_{b,x}} \int_{|z|\geq 1} d\mu(z) |x_1 - x_2|^{1-\beta}.$$

By the fundamental theorem, Fubini, and a change of variables,

$$I_1(t,x) = \int_0^t \int_{|z|<1} \Big[\int_{\mathbb{R}^d} \int_0^1 \nabla_x K(t-s, y+\sigma z) - \nabla_x K(t-s,y) \Big] \cdot z f(s, x-y) d\sigma dy d\mu(z) ds,$$

$$= \int_0^t \int_0^1 \int_{\mathbb{R}^d} \int_{|z|<1} \nabla_x K(t-s,y) \cdot z \Big[f(s, x-y+\sigma z) - f(s, x-y) \Big] d\mu(z) dy d\sigma ds.$$

It follows that

$$I_1(t,x_1) - I_1(t,x_2) = \int_0^t \int_0^1 \int_{\mathbb{R}^d} \nabla_x K(t-s,y) \cdot \int_{|z|<1} z \Big[f(s, x_1 - y + \sigma z) - f(s, x_2 - y + \sigma z) - \big(f(s, x_1 - y) - f(s, x_2 - y) \big) \Big] d\mu(z) dy d\sigma ds.$$

Since

$$
\begin{aligned}
&|f(x_1+\sigma z)-f(x_1)-f(x_2+\sigma z)-f(x_2)|^{1-\beta+\beta}\\
&\le 2\|f\|^{1-\beta}_{C_{b,t}C^1_{b,x}}|x_1-x_2|^{1-\beta}\|f\|^{\beta}_{C_{b,t}C^1_{b,x}}|\sigma z|^{\beta},
\end{aligned}
$$

we see by Theorem 4.3 and (L1) that

$$
\begin{aligned}
&|I_1(t,x_1)-I_1(t,x_2)|\\
&\le \int_0^t\int_{\mathbb{R}^d}|\nabla_x K(t-s,y)|dyds\, 2\|f\|^{1-\beta}_{C_{b,t}C^1_{b,x}}|x_1-x_2|^{1-\beta}\|f\|^{\beta}_{C_{b,t}C^1_{b,x}}\int_{|z|<1}|z|^{\beta+1}d\mu(z)\\
&\le \mathcal{K}\frac{\sigma}{\sigma-1}T^{\frac{\sigma-1}{\sigma}}\int_{|z|<1}|z|^{\beta+1}d\mu(z)\|f\|_{C_{b,t}C^1_{b,x}}|x_1-x_2|^{1-\beta}.
\end{aligned}
$$

Combining the above two estimates, we conclude that

$$
|\mathcal{L}[\Phi(f)](t,x_1)-\mathcal{L}[\Phi(f)](t,x_2)|\le c\|f\|_{C_{b,t}C^1_{b,x}}|x_1-x_2|^{1-\beta},
$$

with $c=\frac{\sigma}{\sigma-1}T^{\frac{\sigma-1}{\sigma}}\mathcal{K}\int_{|z|<1}|z|^{1+\beta}d\mu(z)+4T\int_{|z|\ge 1}d\mu(z)$. By part a), $\partial_t\Phi(f)(t,x)=f(t,x)+\mathcal{L}[\Phi(f)](t,x)$. Since

$$
|f(t,x)-f(t,y)|\le (2\|f\|_{C_b})^{\beta}|f(t,x)-f(t,y)|^{1-\beta}\le 2\|f\|_{C_{b,t}C^1_{b,x}}|x-y|^{1-\beta},
$$

we then also get that

$$
|\partial_t\Phi[f](t,x_1)-\partial_t\Phi[f](t,x_2)|\le (2+c)\|f\|_{C_{b,t}C^1_{b,x}}|x_1-x_2|^{1-\beta}.
$$

b) *Part 2:* Uniform continuity in time. First note that

$$
\begin{aligned}
&\mathcal{L}\Phi[f](t,x)-\mathcal{L}\Phi[f](s,x)=\int_0^t\mathcal{L}K(\tau,\cdot)*f(t-\tau,\cdot)d\tau-\int_0^s\mathcal{L}K(\tau,\cdot)*f(s-\tau,\cdot)d\tau\\
&=\int_0^s\mathcal{L}K(\tau,\cdot)*\big(f(t-\tau,\cdot)-f(s-\tau,\cdot)\big)d\tau+\int_s^t\mathcal{L}K(\tau,\cdot)*f(t-\tau,\cdot)d\tau.
\end{aligned}
$$

Now we do as before: Split the z-domain in two parts, use the fundamental theorem and a change of variables to get

$$
\begin{aligned}
&\mathcal{L}K(\tau,\cdot)*\big(f(t-\tau,\cdot)-f(s-\tau,\cdot)\big)\\
&=\int_0^1\int_{\mathbb{R}^d}\int_{|z|<1}\nabla_x K(\tau,x-y)\cdot z\big[f(t-\tau,y+\sigma z)-f(t-\tau,y)
\end{aligned}
$$

$$- f(s-\tau, y+\sigma z) + f(s-\tau, y)\big]d\mu(z)dyd\sigma.$$

$$+\int_{\mathbb{R}^d}\int_{|z|\geq 1} K(\tau, x-y)[f(t-\tau, y+z) - f(t-\tau, y)$$

$$- f(s-\tau, y+z) + f(s-\tau, y)]d\mu(z)dy.$$

Then we apply the trick

$$\begin{aligned}&|f(t-\tau, y+\sigma z) - f(t-\tau, y) - f(s-\tau, y+\sigma z) + f(s-\tau, y)|\\ &\leq 2\omega_f(|t-s|)^{1-\beta}(\|f\|_{C_{b,t}C^1_{b,x}}|z|)^\beta \quad \text{or} \quad 4\omega_f(|t-s|)^{1-\beta}\|f\|^\beta_{C_b},\end{aligned}$$

and find using Theorem 4.3 and (L1) that

$$\begin{aligned}&\Big|\int_0^s \mathcal{L}K(\tau,\cdot) * \big(f(t-\tau,\cdot) - f(s-\tau,\cdot)\big)d\tau\Big|\\ &\leq \Big[\frac{\sigma}{\sigma-1}s^{\frac{\sigma-1}{\sigma}}\mathcal{K}\int_{|z|<1}|z|^{1+\beta}d\mu(z) + 4s\int_{|z|\geq 1}d\mu(z)\Big]\|f\|^\beta_{C_{b,t}C^1_{b,x}}\omega_f(|t-s|)^{1-\beta}.\end{aligned}$$

In a similar way we find that

$$\begin{aligned}&\Big|\int_s^t \mathcal{L}K(\tau,\cdot) * f(t-\tau,\cdot)d\tau\Big|\\ &\leq \Big[2\frac{\sigma}{\sigma-1}(t^{\frac{\sigma-1}{\sigma}} - s^{\frac{\sigma-1}{\sigma}})\mathcal{K}\int_{|z|<1}|z|^{1+\beta}d\mu(z) + 2(t-s)\int_{|z|\geq 1}d\mu(z)\Big]\|f\|_{C_b}\\ &\leq c_1\|f\|_{C_b}|t-s|^{\frac{\sigma-1}{\sigma}}.\end{aligned}$$

Combining all above estimates leads to

$$\Big|\mathcal{L}\Phi[f](t,x) - \mathcal{L}\Phi[f](s,x)\Big| \leq c\|f\|^\beta_{C_{b,t}C^1_{b,x}}\omega_f(|t-s|)^{1-\beta} + \tilde{c}\|f\|_{C_b}|t-s|^{\frac{\sigma-1}{\sigma}},$$

where c is defined above and in the Lemma and

$$\tilde{c} = 2\frac{\sigma}{\sigma-1}\mathcal{K}\int_{|z|<1}|z|^{1+\beta}d\mu(z)\max_{s,t\in[0,T]}\frac{|t^{\frac{\sigma-1}{\sigma}} - s^{\frac{\sigma-1}{\sigma}}|}{|t-s|^{\frac{\sigma-1}{\sigma}}} + 2T^{\frac{1}{\sigma}}\int_{|z|\geq 1}d\mu(z).$$

Note that $\tilde{c}$ is finite. Then since

$$\partial_t\Phi[f](t,x) - \partial_t\Phi[f](s,x) = f(t,x) - f(s,x) + \mathcal{L}\Phi[f](t,x) - \mathcal{L}\Phi[f](s,x),$$

and $|f(t,x) - f(s,x)| \leq (2\|f\|_{C_b})^\beta\omega_f(|t-s|)^{1-\beta}$, the continuity estimate for $\partial_t\Phi[f]$ follows.

c) The proof follows by writing

$$\partial_{x_i}\Phi(g)(t,x)=\int_0^t \partial_{x_i}K(\tau,z)g(t-\tau,x-z)dzd\tau,$$

and then directly compute the difference $|\partial_{x_i}\Phi(g)(t,x)-\partial_{x_i}\Phi(g)(s,y)|$.

The proof is complete.

References

[1] Y. Achdou, P. Cardaliaguet, F. Delarue, A. Porretta, F. Santambrogio, Mean field games, in: CIME - Subseries, in: Lectures Notes in Mathematics, vol. 2281, Springer, 2020.

[2] F. Achleitner, C. Kuehn, Traveling waves for a bistable equation with nonlocal diffusion, Adv. Differ. Equ. 20 (9–10) (2015) 887–936.

[3] David Applebaum, Lévy Processes and Stochastic Calculus, Cambridge University Press, 2009.

[4] M. Bardi, M. Cirant, Uniqueness of solutions in mean field games with several populations and Neumann conditions, in: PDE Models for Multi-Agent Phenomena, in: Springer INdAM Ser., vol. 28, Springer, 2018, pp. 1–20.

[5] G. Barles, E. Chasseigne, A. Ciomaga, C. Imbert, Lipschitz regularity of solutions for mixed integro-differential equations, J. Differ. Equ. 252 (11) (2012) 6012–6060.

[6] G. Barles, C. Imbert, Second-order elliptic integro-differential equations: viscosity solutions' theory revisited, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 25 (3) (2008) 567–585.

[7] A. Bensoussan, J. Frehse, P. Yam, Mean Field Games and Mean Field Type Control Theory, SpringerBriefs in Mathematics, vol. 101, Springer, 2013.

[8] F. Camilli, R. De Maio, et al., A time-fractional mean field game, Adv. Differ. Equ. 24 (9/10) (2019) 531–554.

[9] P. Cardaliaguet, Notes on mean field games (from P.-L. Lions' lectures at college de France), Preprint, https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf, 2013.

[10] P. Cardaliaguet, F. Delarue, J.-M. Lasry, P.-L. Lions, The Master Equation and the Convergence Problem in Mean Field Games, Annals of Mathematics Studies, vol. 201, Princeton University Press, 2019.

[11] P. Cardaliaguet, P.J. Graber, A. Porretta, D. Tonon, Second order mean field games with degenerate diffusion and local coupling, NoDEA Nonlinear Differ. Equ. Appl. 22 (5) (2015) 1287–1317.

[12] R. Carmona, F. Delarue, Probabilistic Theory of Mean Field Games with Applications i and II, Springer Verlag, 2018.

[13] A. Cesaroni, M. Cirant, S. Dipierro, M. Novaga, E. Valdinoci, On stationary fractional mean field games, J. Math. Pures Appl. 122 (9) (2019) 1–22.

[14] E. Chasseigne, O. Ley, T.T. Nguyen, A priori Lipschitz estimates for solutions of local and nonlocal Hamilton-Jacobi equations with Ornstein-Uhlenbeck operator, Rev. Mat. Iberoam. 35 (5) (2019) 1415–1449.

[15] I. Chowdhury, E.R. Jakobsen, K. Milosz, On fully nonlinear parabolic mean field games with examples of nonlocal and local diffusions, Preprint, 2020.

[16] M. Cirant, A. Goffi, On the existence and uniqueness of solutions to time-dependent fractional MFG, SIAM J. Math. Anal. 51 (2) (2019) 913–954.

[17] M. Cirant, D. Tonon, Time-dependent focusing mean-field games: the sub-critical case, J. Dyn. Differ. Equ. 31 (1) (2019) 49–79.

[18] R. Cont, P. Tankov, Financial Modelling with Jump Processes, Chapman & Hall/CRC, 2004.

[19] J. Droniou, T. Gallouët, J. Vovelle, Global solution and smoothing effect for a non-local regularization of a hyperbolic equation, J. Evol. Equ. 3 (3) (2003) 499–521.

[20] J. Droniou, C. Imbert, Fractal first-order partial differential equations, Arch. Ration. Mech. Anal. 182 (2) (2006) 299–331.

[21] M.G. Garroni, J.L. Menaldi, Second Order Elliptic Integro-Differential Problems, Chapman & Hall/CRC, 2002.

[22] D.A. Gomes, L. Nurbekyan, E.A. Pimentel, Economic Models and Mean-Field Games Theory, IMPA Mathematical Publications, IMPA, Rio de Janeiro, 2015.

[23] D.A. Gomes, E.A. Pimentel, V. Voskanyan, Regularity Theory for Mean-Field Game Systems, Springer, Berlin, 2016.

[24] P.J. Graber, V. Ignazio, A. Neufeld, Nonlocal Bertrand and Cournot mean field games with general nonlinear demand schedule, J. Math. Pures Appl. 148 (2021) 150–198.

[25] T. Grzywny, K. Szczypkowski, Estimates of heat kernels of non-symmetric Lévy processes, Preprint, arXiv:1710.07793, 2017.
[26] O. Gueant, J.-M. Lasry, P.-L. Lions, Mean field games and applications, in: Paris-Princeton Lectures on Mathematical Finance, Springer, 2010, pp. 205–266.
[27] M. Huang, R.P. Malhamé, P.E. Caines, et al., Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, Commun. Inf. Syst. 6 (3) (2006) 221–252.
[28] C. Imbert, A non-local regularization of first order Hamilton–Jacobi equations, J. Differ. Equ. 211 (1) (2005) 218–246.
[29] E.R. Jakobsen, K.H. Karlsen, Continuous dependence estimates for viscosity solutions of integro-pdes, J. Differ. Equ. 212 (2) (2005) 278–318.
[30] E.R. Jakobsen, K.H. Karlsen, A "maximum principle for semicontinuous functions" applicable to integro-partial differential equations, NoDEA Nonlinear Differ. Equ. Appl. 13 (2) (2006) 137–165.
[31] J.-M. Lasry, P.-L. Lions, Jeux à champ moyen. II. Horizon fini et contrôle optimal, C. R. Math. Acad. Sci. Paris 343 (10) (2006) 679–684 (French).
[32] J.-M. Lasry, P.-L. Lions, Mean field games, Jpn. J. Math. 2 (1) (2007) 229–260.
[33] P.-L. Lions, Courses in college de France on mean field games, https://www.college-de-france.fr/site/en-pierre-louis-lions/_course.htm, 2006–2012.
[34] R. Metzler, J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, Phys. Rep. 339 (1) (2000) 77.
[35] A. Porretta, Weak solutions to Fokker-Planck equations and mean field games, Arch. Ration. Mech. Anal. 216 (1) (2015) 1–62.
[36] R.L. Schilling, P. Sztonyk, J. Wang, Coupling property and gradient estimates of Lévy processes via the symbol, Bernoulli 18 (4) (2012) 1128–1149.
[37] L. Silvestre, Regularity of the obstacle problem for a fractional power of the Laplace operator, Commun. Pure Appl. Math. 60 (1) (2007) 67–112.
[38] W.A. Woyczyński, Lévy processes in the physical sciences, in: Lévy Processes, Birkhäuser, 2001, pp. 241–266.