ON FULLY NONLINEAR PARABOLIC MEAN FIELD GAMES
WITH EXAMPLES OF NONLOCAL AND LOCAL DIFFusions

INDRANIL CHOWDHURY, ESPEN R. JAKOBSEN, AND MIŁOSZ KRUPSKI

Abstract. In this paper we introduce a class of fully nonlinear mean field games posed in $[0,T] \times \mathbb{R}^d$. We justify that they are related to controlled local or nonlocal diffusions, and more generally in our setting, to controlled time change rates of stochastic (Lévy) processes. This control interpretation seems to be new. We prove existence and uniqueness of solutions under general assumptions. Uniqueness follows without strict monotonicity of couplings or strict convexity of Hamiltonians. These results are applied to strongly degenerate equations of order less than one — and non-degenerate equations (including both local second order and nonlocal involving fractional Laplacians). In both cases we consider a rich class of nonlocal operators and corresponding processes. We develop tools to work without explicit moment assumptions, and uniqueness in the degenerate case relies on a new type of argument for the (nonlocal) Fokker–Planck equation.

Contents

1. Introduction 2
   1.1. Statement of the problem 2
   1.2. Concepts of solutions 4
   1.3. Objectives, contributions, and structure of the paper 4
2. Main results – Existence and uniqueness 5
   2.1. Degenerate mean field games of order less than one 5
   2.2. Non-degenerate, local, second-order mean field games 6
   2.3. Non-degenerate, nonlocal mean field games 7
   2.4. General well-posedness theory 8
   2.5. Other contributions 9
3. Derivation of the model 11
   3.1. Time changed Lévy process 11
   3.2. Control problem and Bellman equation 12
   3.3. Optimal control and Fokker–Planck equation 13
   3.4. Heuristic derivation of the mean field game 14
   3.5. Example and a relation to continuous control 14
4. Preliminaries 15
   4.1. Basic notation 15
   4.2. Spaces of measures 15
   4.3. Lévy operators 19
   4.4. Viscosity solutions 22
5. Hamilton–Jacobi–Bellman equation 24
   5.1. The degenerate case 25

Date: May 6, 2022.
2020 Mathematics Subject Classification. 35A01, 35A02, 35D30, 35D40, 35K55, 35K65, 35Q84, 35Q89, 35R11, 47D07, 49L, 49N80, 60G51.
Key words and phrases. Mean field games, Fokker–Planck–Kolmogorov equation, Hamilton–Jacobi–Bellman equation, fully-nonlinear PDEs, degenerate PDEs, nonlocal PDEs, Brownian motion, Lévy processes, controlled diffusion, existence, uniqueness.
5.2. The non-degenerate case 26
6. Fokker–Planck equation — Existence 27
7. Fokker–Planck equation — Uniqueness 33
  7.1. Dual equation in the degenerate case 33
  7.2. Dual equation in the non-degenerate case 40
  7.3. Uniqueness for Fokker–Planck 40
8. Mean field game 41
Appendix A. The fractional Laplacian 45
Appendix B. The Legendre–Fenchel transform 46
Appendix C. Proof of the comparison principle Theorem 5.3 48
Appendix D. Viscosity solutions of dual of Fokker–Planck equation 49
Acknowledgements 50
References 50

1. INTRODUCTION

In this paper we introduce a new model of mean field games and analyse it using PDE methods. The mathematical theory of mean field games was introduced by Lasry–Lions [63, 64, 65] and Huang–Caines–Malhamé [45, 44] in 2006, and today this is a large and rapidly expanding field. This research is mostly focused on either PDE or stochastic approaches. An extensive background and recent developments can be found in e.g. [1, 7, 19, 20, 38, 17, 41] and the references therein.

In contrast to the more classical setting, our problem does not involve controlled drift but rather controlled diffusion, or more specifically, control of the time change rate of a Lévy process. Such a model yields a system of partial differential equations which is fully nonlinear and possibly strongly degenerate. It may be local or nonlocal. In [18] the authors allow for a degenerate diffusion, but it is not controlled and there are restrictions on its regularity, cf. [78, 12]. There are recent results on mean field games with nonlocal (uncontrolled) diffusion involving Lévy operators [22, 27, 33, 50]. See also [16] for a problem involving fractional time derivatives.

Control of the diffusion is a rare and novel subject, mostly addressed by stochastic methods: [61] introduces an approach based on relaxed controls and martingale problems to show existence of probabilistic solutions to very general local mean field games, and [6, 5] consider extensions to problems perturbed by bounded nonlocal terms. Some results by PDE methods can be found in [75], as well as [2, 29] for uniformly elliptic (stationary second order) problems.

1.1. Statement of the problem. We study derivation, existence, and uniqueness questions related to the mean field game system

\[
\begin{aligned}
-\partial_t u &= F(Lu) + f(m) & \text{on } \mathcal{T} \times \mathbb{R}^d, \\
u(T) &= g(m(T)) & \text{on } \mathbb{R}^d, \\
\partial_t m &= L^* (F'(Lu) m) & \text{on } \mathcal{T} \times \mathbb{R}^d, \\
m(0) &= m_0 & \text{on } \mathbb{R}^d,
\end{aligned}
\]

where \( \mathcal{T} = (0, T) \) for a fixed \( T \in (0, \infty) \), and \( \mathcal{T} = [0, T] \). We assume \( L \) to be a Lévy operator, or an infinitesimal generator of a Lévy process (see [14, §2.1]), namely
\((L)\) $\mathcal{L} : C^2_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ is a linear operator with a triplet \((c, a, \nu)\), where $c \in \mathbb{R}^d$, $a \in \mathbb{R}^{d \times d}$, $\nu$ is a Lévy measure (see Definition 4.11), and

$$\mathcal{L}\phi(x) = c \cdot \nabla \phi(x) + \text{tr} \left( aa^T D^2 \phi(x) \right) + \int_{\mathbb{R}^d} \left( \phi(x + z) - \phi(x) - \mathbb{I}_{B_1}(z) \cdot \nabla \phi(x) \right) \nu(dz).$$

By $\mathcal{L}^*$ we denote the (formal) adjoint operator of $\mathcal{L}$, which is also a Lévy operator. Furthermore, we consider the following set of assumptions.

\((A1)\): $F \in C^1(\mathbb{R})$ and $F' \in C^\gamma(\mathbb{R})$ for some $\gamma \in (0, 1]$ (see Definition 4.1), and $F' \geq 0$;

\((A2)\): $F$ is convex;

\((A3)\): $m_0$ is a probability measure on $\mathbb{R}^d$;

\((A4)\): $f : C(\overline{T}, \mathcal{P}(\mathbb{R}^d)) \to C_b(\mathbb{T} \times \mathbb{R}^d)$ and $g : \mathcal{P}(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ are continuous, i.e. $\lim_{n \to \infty} \sup_{t \in \overline{T}} \|m_n(t) - m(t)\| = 0$ (see Definition 4.4) implies

$$\lim_{n \to \infty} \|f(m_n) - f(m)\|_\infty = 0 \quad \text{and} \quad \lim_{n \to \infty} \|g(m_n(T)) - g(m(T))\|_\infty = 0;$$

\((A5)\): \(f\) and \(g\) are monotone operators, namely

$$\int_{\mathbb{B}_1} (g(m_1) - g(m_2))(x)(m_1 - m_2)(dx) \leq 0,$$

$$\int_0^T \int_{\mathbb{R}^d} (f(m_1) - f(m_2))(t,x)(m_1 - m_2)(t,dx) dt \leq 0,$$

for every pair $m_1, m_2$ in $\mathcal{P}(\mathbb{R}^d)$ or $C(\overline{T}, \mathcal{P}(\mathbb{R}^d))$.

**Remark 1.1.** (a) For our existence results, $F \in C^1(\mathbb{R})$ and $\gamma = 0$ in \((A1)\) is sufficient, but not for uniqueness. When $F$ is only locally Lipschitz, the stability in Lemma 6.4 becomes the main obstacle. It seems reasonable that this and other obstacles can be overcome for non-degenerate equations (see also [2, 29]).

(b) Every function $F$ satisfying \((A1)\) and \((A2)\) is the Hamiltonian of a Hamilton–Jacobi–Bellman equation. This follows using the Legendre–Fenchel transform:

$$F(z) = \sup_{\zeta \in (0, \infty)} \left( \zeta z - F^*(\zeta) \right), \quad \text{where} \quad F^*(\zeta) = \sup_{z \in \mathbb{R}} \left( \zeta z - F(z) \right).$$

Notice that \((A1)\) implies $F^*(\zeta) = +\infty$ for $\zeta < 0$. For more details we refer to Appendix B and Section 3.2, in particular equations (9)–(12).

(c) The operators in \((A4)\) are so-called smoothing couplings. Typically they are nonlocal and defined by a convolution with a fixed kernel (see e.g., [1]). The case of local couplings requires a different setup and will not be treated here.

(d) Assumption \((A5)\) is the *standard Lasry–Lions* monotonicity conditions required for uniqueness. The equivalent and more familiar formulation with $\tilde{f}$ and $\tilde{g}$ non-decreasing [65, 1] is obtained by taking $\tilde{f} = -f$, $\tilde{f} = -f$, and $\tilde{u} = -u$, leading to $-\partial_t u = \mathcal{L}(-\tilde{u}) + \tilde{f}(m)$ and $\tilde{u}(T) = \tilde{g}(m)$ in the first part of problem (1). Our choice simplifies the notation when nonlinear diffusion is involved and corresponds to interpreting $u$ as a “maximal gain” and not a “minimal cost” (cf. Section 3.2).

(c) We assume neither strict convexity in \((A2)\) nor strict monotonicity in \((A5)\) and still obtain uniqueness for problem (1).

---

\(^1\)These three conditions need to be strengthened for our results to hold in the concrete cases we present, compare the statements of Theorem 2.1, Theorem 2.5, Theorem 2.8, and Theorem 2.9.
Remark 1.2. Control problems/games have many applications throughout the sciences. Controlled diffusions appear e.g. in portfolio optimization in finance, cf. [13, 42, 74]. Our model is always related to control problems/games with controlled diffusions, see Remark 1.1(b).

1.2. Concepts of solutions. With \((f,g) = (f(m), g(m(T)))\), the first pair of equations in problem (1) form a terminal value problem for a Hamilton–Jacobi–Bellman equation,

\[
\begin{align*}
\partial_t u &= F(Lu) + f & \text{on } T \times \mathbb{R}^d, \\
u(T) &= g & \text{on } \mathbb{R}^d.
\end{align*}
\]

Definition 1.3. A function \(u \in C_b(T \times \mathbb{R}^d)\) is a bounded classical solution of problem (2) with data \((f,g)\), if \(\partial_t u, Lu \in C(T \times \mathbb{R}^d)\) and equations in problem (2) are satisfied pointwise.

With \(b = F'(Lu)\) the second pair of equations in problem (1) form an initial value problem for a Fokker–Planck equation,

\[
\begin{align*}
\partial_t m &= L^*(bm) & \text{on } T \times \mathbb{R}^d, \\
m(0) &= m_0 & \text{on } \mathbb{R}^d.
\end{align*}
\]

We look for very weak solutions of problem (3) when \(b\) is bounded and continuous.

Definition 1.4. Suppose \(b \in C_b(T \times \mathbb{R}^d)\). A function \(m \in C(T, P(\mathbb{R}^d))\) is a very weak solution of problem (3) if for every \(\phi \in C^\infty_c(T \times \mathbb{R}^d)\) and \(t \in T\) we have

\[
m(t)[\phi(t)] = m_0[\phi(0)] + \int_0^t m(\tau) [\partial_t \phi(\tau) + b(\tau)(L\phi)(\tau)] d\tau.
\]

Now we may define the concept of solutions of problem (1).

Definition 1.5. A pair \((u, m)\) is a classical–very weak solution of problem (1) if \(u\) is a bounded classical solution of problem (2) (see Definition 1.3) with data \((f(m), g(m(T)))\), such that \(F'(Lu) \in C_b(T \times \mathbb{R}^d)\), and \(m\) is a very weak solution of problem (3) (see Definition 1.4) with initial data \(m_0\) and coefficient \(b = F'(Lu)\).

Remark 1.6. Problem (2) is fully nonlinear, and the viscosity solutions framework applies. But then, \(Lu\) and \(F'(Lu)\) may not be well-defined. We therefore prefer to work with classical solutions of problem (2). Still, \(b\) need not be very regular and it may be degenerate (i.e. \(b(t,x) = 0\), possibly on a large set), thus we need to consider very weak (measure-valued) solutions of problem (3).

Another benefit of relaxing the solution concept for \(m\), is reduced regularity requirements on the data. In the local case, for \(m\) to be a classical solution, \(u\) and the data would need to have two more derivatives than in our classical–very weak setup. See also the discussion on classical–classical solutions on page 8.

1.3. Objectives, contributions, and structure of the paper. Our first objective is to study problem (1) with a (nearly) minimal set of assumptions (A1)–(A5), naturally arising from the analysis in Section 3. We reduce the question of its well-posedness to a set of general conditions (S1)–(S5) describing the properties of solutions of individual equations (2) and (3) (Section 2.4).

The second objective is to verify these conditions in several concrete cases: We consider a nonlocal degenerate problem (Section 2.1), as well as two non-degenerate problems — local (Section 2.2) and nonlocal (Section 2.3). The last case relies on a well-motivated conjecture in regularity theory, which we leave to the experts in that field. Most of the work in this paper is related to the degenerate case.
The structure of the paper is the following. Section 2 contains our main contributions, the (concrete and general) existence and uniqueness results for the fully nonlinear mean field game (1), along with a list of auxiliary results of independent interest. Another important contribution is given in Section 3: A novel heuristic derivation of the mean field game (1). Section 4 contains both background material and new results that are needed in the proofs, including results on tightness, approximations of Lévy operators, and regularity of viscosity solutions. Section 5 and Section 7 contain results on the Hamilton–Jacobi–Bellman (2) and uniqueness for the Fokker–Planck (3) equations in the concrete cases outlined in Section 2. In Section 6 we derive existence and stability of solutions of the Fokker–Planck equation under general assumptions. Section 8 contains the proofs of the general existence and uniqueness results for the mean field game and can be read independently. Some technical proofs and remarks are given in the appendices.

2. Main results – Existence and uniqueness

We begin the description of our results with a compilation of concrete cases of mean field games with controlled diffusion (a Lévy process), which lead to fully nonlinear systems of problem (1) type.

2.1. Degenerate mean field games of order less than one.

We consider following assumptions.

(LA): Let \(2\sigma \in (0, 1)\) and \(\mathcal{L}\) be given by (see Remark 4.12(b))

\[
\mathcal{L}\phi(x) = \int_{\mathbb{R}^d} \left(\phi(x + z) - \phi(x)\right) \nu(dz),
\]

where the Lévy measure \(\nu\) satisfies

\[
\int_{B_1} \left(1 \wedge \frac{|z|^p}{r^p}\right) \nu(dz) \leq \frac{K}{p - 2\sigma} r^{-2\sigma}
\]

for a constant \(K \geq 0\) and every \(p \in (2\sigma, 1], r \in (0, 1)\).

(RA): There are \(\alpha \in (2\sigma, 1] \text{ and } M \in [0, \infty)\) such that the range \(\mathcal{R} = \{(f(m), g(m(T))): m \in C(\overline{T}, \mathcal{P}(\mathbb{R}^d))\}\) satisfies \(\mathcal{R} \subset \mathcal{R}_A(\alpha, M)\), where\(^1\)

\[
\mathcal{R}_A(\alpha, M) = \{(f, g): \begin{array}{l}
(i) \ f \in UC(T \times \mathbb{R}^d) \cap B(T, \mathcal{C}_0^\alpha(\mathbb{R}^d)), \\
(ii) \ g \in \mathcal{C}_0^\alpha(\mathbb{R}^d), \\
(iii) \ \sup_{t \in T} \|f(t)\|_\alpha + \|g\|_\alpha \leq M \end{array}\}.
\]

For absolutely continuous \(\nu\), (LA) is equivalent to the upper bound \(\frac{d\nu}{dz} \leq \frac{C}{|z|^{2\alpha+1}}\) for \(|z| < 1\), and hence is satisfied for the fractional Laplacian \(\Delta^\sigma\) [14], the nonsymmetric nonlocal operators used in finance [28], and a large class of non-degenerate and degenerate operators. Any bounded Lévy measure (“\(\sigma \downarrow 0\)”), and hence any bounded nonlocal (Lévy) operator, is also included. We refer to Remark A.3 for more details and examples. Note that there is no further restriction on the tail of \(\nu\) (the \(B_1^c\)-part) and hence no explicit moment assumption on the Lévy process and the solution of the Fokker–Planck equation \(m\). See Section 2.5 for more details.

Theorem 2.1. Assume (RA), (LA), (A1), (A3). If in addition

(i) \(A4\) holds, then there exists a classical–very weak solution of problem (1);

\(^1\)See Definition 4.1, Definition 4.3 of spaces \(\mathcal{C}_0^\alpha(\mathbb{R}^d)\) and \(B(T, X)\); \(UC = \) uniformly continuous.
(ii) $(A2), (A5)$ hold and $\frac{2\sigma}{(\alpha-2\sigma)}(1 + \frac{1}{1-2\sigma}) < \gamma$, then problem (1) has at most one classical–very weak solution;

(iii) $(A2), (A5)$ hold, $\nu$ is symmetric at the origin (see Definition 4.11), and $\frac{2\sigma}{(\alpha-2\sigma)}(1 + \frac{1}{\gamma}) < \gamma$, then problem (1) has at most one classical–very weak solution.

This theorem follows as a corollary of the more general existence and uniqueness result, Theorem 2.9 in Section 2.4. For the outline of the proof, see the subsequent explanation therein.

Remark 2.2. (a) Assume $\gamma = \alpha = 1$. Then the condition in Theorem 2.1 (ii) becomes $(1 - 2\sigma)^{-2} < 2$, which leads to $2\sigma < \frac{2 - \sqrt{2}}{2} \approx \frac{3}{10}$. In (iii) we obtain $(1 - 2\sigma)^{-1} < 2$ and then $2\sigma < \frac{7 - \sqrt{33}}{4} \approx \frac{4}{13}$.

(b) See Section 3.5 for an example involving the fractional Laplacian and a strongly degenerate power type nonlinearity $F$.

2.2. Non-degenerate, local, second-order mean field games.

(LB): $\mathcal{L}$ is given by

$$\mathcal{L}\phi(x) = \text{tr} \left( aa^T D^2 \phi(x) \right),$$

where $\det aa^T > 0$.

(RB): There are $\alpha \in (2\sigma, 1]$ and $M \in [0, \infty)$ such that the range

$$\mathcal{R} = \{ (f(m), g(m(T))) : m \in C(T, \mathcal{P}(\mathbb{R}^d)) \}$$

satisfies $\mathcal{R} \subset \mathcal{R}_B(\alpha, M)$, where

$$\mathcal{R}_B(\alpha, M) = \left\{ (f, g) : \begin{array}{ll} (i) & f \in C_{b}^{1,\alpha}(T \times \mathbb{R}^d), \\
(ii) & g \in \text{BUC}(\mathbb{R}^d) \quad \text{and} \quad \mathcal{L}g \in L^\infty(\mathbb{R}^d), \\
(iii) & \|f\|_{1,\alpha} + \|\mathcal{L}g\|_{\infty} + \|g\|_{\infty} \leq M \end{array} \right\}.$$

(FB): $F' \geq \kappa$ for some $\kappa > 0$ (i.e. $F$ is strictly increasing).

Under assumption (LB), operator $\mathcal{L}$ is non-degenerate. For problem (1) to be non-degenerate, we also need to assume (FB). In this setting, we expect interior regularity estimates to hold.

Definition 2.3 (Interior estimates). Assume (L). Interior $(\beta, \alpha)$-regularity estimates hold for problem (2) if for every $f \in C^3_b(\mathcal{T} \times \mathbb{R}^d)$, and $(t, x) \in \mathcal{T} \times \mathbb{R}^d$, and viscosity solution $u$ of problem (2),$^3$ we have

$$[\partial_t u]_{C^3} + f(x) + [\mathcal{L}u]_{C^3} \leq C(t)(\|f\|_{\beta, \alpha} + \|u\|_{\infty}).$$

Note that once we establish the comparison principle (see Definition 4.22 and Theorem 5.1), the right-hand side of the estimate can be expressed in terms of $\|f\|_{\beta, \alpha}$ and $\|g\|_{\infty}$. When $F(z) = az + b$ is affine, interior regularity is given by classical Schauder theory (see e.g. [57, 62, 66]).

In the fully nonlinear case, such estimates have been proved in [86].

Lemma 2.4 (Local case). Assume (LB), (FB), $(f,g) \in \mathcal{R}_B(\alpha, M)$ (as in (RB)), $(A1), (A2)$. Then interior $(\alpha/2, \alpha)$-regularity estimates hold for problem (2).

Proof. The result is stated in a form which is a corollary to [72, Theorem 5.2]. As in [72], the result follows from the arguments in [86], in particular Theorems 1.1 and 4.13 and their proofs. (A translation invariant $\mathcal{L}$ simplifies our case slightly.)

$^2$See Definition 4.2 of spaces $C^3_b(\mathcal{T} \times \mathbb{R}^d)$; BUC = bounded uniformly continuous.

$^3$See footnote 2; see Definition 4.20 of viscosity solutions for $\alpha = 0$; for $\alpha \neq 0$ the definition is analogous.
Related results can be found in e.g. [55, 56, 66].

**Theorem 2.5.** Assume (RB), (FB), (LB), (A1), (A2), (A3). If in addition

(i) (A4) holds, then there exists a classical–very weak solution of problem (1);
(ii) (A5) holds, then problem (1) has at most one classical–very weak solution.

This theorem follows as a corollary of the more general existence and uniqueness result, Theorem 2.9 in Section 2.4. For the outline of the proof, see the subsequent explanation therein.

**Remark 2.6.** Existence (without uniqueness) of probabilistic solutions is shown in [61] under general assumptions. There, it is claimed (in Remark 2.3) that these solutions under various assumptions can be viscosity or classical/(very) weak solutions of mean field game systems like problem (1). However, no proof or explanation is given.

2.3. Non-degenerate, nonlocal mean field games.

**(LC):** Let \(2\sigma \in (0, 2)\) and \(\mathcal{L}\) be given by

\[
\mathcal{L}\phi(x) = \int_{\mathbb{R}^2} \left( \phi(x + z) - \phi(x) - \mathbb{1}_{[1, 2]}(2\sigma) \mathbb{1}_{B_1}(z) \cdot \nabla \phi(x) \right) \nu(dz),
\]

where \(\nu\) is a Lévy measure (see Definition 4.11) whose restriction to \(B_1\) is absolutely continuous with respect to the Lebesgue measure and, for \(\alpha\) as in (RB) and a constant \(K > 0\), \(\nu\) satisfies (see (17))

\[
\mathbb{1}_{B_1}(z) \nu(dz) = \frac{k(z)}{|z|^{d+2\sigma}} \nu(dz), \quad K^{-1} \leq k(z) \leq K, \quad \|k\|_{L^\infty(B_1)} < \infty;
\]

if \(2\sigma = 1\), then in addition \(\int_{B_1 \setminus B_r} \frac{k(z)}{|z|^{d+2\sigma}} \nu(dz) = 0\) for every \(r \in (0, 1)\); see Remark 4.12(b) when \(2\sigma \in (0, 1)\).

As in the previous example, operator \(\mathcal{L}\) is non-degenerate and we assume (FB) to make problem (1) non-degenerate as well.

Condition (LC) defines a rich class of nonlocal operators including fractional Laplacians and the nonsymmetric operators used in finance. There is no additional restrictions on the tail behaviour of \(\nu\).

Here we also expect interior regularity estimates to hold (see Definition 2.3), however no results are available beyond the case when \(F\) is affine (see results for generators of analytic semigroups [67, 82]). Therefore, we pose the following conjecture.

**Conjecture 2.7 (Nonlocal case).** Assume (LC), (FB), \((f, g) \in \mathcal{R}_p(\alpha, M)\) (as in (RB)), (A1), (A2). Then interior \((\alpha/2\sigma, \alpha)\)-regularity estimates hold for problem (2).

Even though there are many related results (see e.g. [23, 24, 32, 54, 69, 72]), the precise setting of Conjecture 2.7 seems to be unsolved and it falls outside the scope of this paper to address it. There is a rich literature for the corresponding elliptic problems as well (see e.g. [15, 53, 79, 81]).

**Theorem 2.8.** Suppose Conjecture 2.7 holds. Assume (LC), (RB), (FB), (A1), (A2), (A3). If in addition

(i) (A4) holds, then there exists a classical–very weak solution of problem (1);
(ii) (A5) holds, then problem (1) has at most one classical–very weak solution.

This theorem follows as a corollary of the more general existence and uniqueness result, Theorem 2.9 in Section 2.4. For the outline of the proof, see the subsequent explanation therein.
On classical–classical solutions. In the non-degenerate setting we expect higher regularity results to hold provided the data and nonlinearities are sufficiently smooth. This can lead to existence of classical–classical solutions for problem (1). Let us explain it in the local case.

Suppose $u$ is a bounded classical solution of problem (2). For $v = \mathcal{L}u$ and $\varphi \in C^\infty_c(\mathcal{T} \times \mathbb{R}^d)$ we obtain

$$
\int_0^T \int_{\mathbb{R}^d} (v \partial_t \varphi - F(v) \mathcal{L} \varphi) \, dx \, dt = \int_0^T \int_{\mathbb{R}^d} f \mathcal{L} \varphi \, dx \, dt,
$$

i.e. $v$ is a distributional solution of the porous medium equation

$$
\begin{cases}
-\partial_t v = LF(v) + Lf & \text{on } \mathcal{T} \times \mathbb{R}^d, \\
v(T) = Lg & \text{on } \mathbb{R}^d.
\end{cases}
$$

In (LB) case, we can write this equation in the quasilinear, divergence form

$$
-\partial_t v = \text{div} \left( F'(v) aa^T \nabla v \right) + Lf.
$$

When we assume $F' \geq \kappa > 0$, $F''' \in C^\gamma(\mathbb{R})$, $Lf \in C^{\alpha,\beta}_c(\mathcal{T} \times \mathbb{R}^d)$, and $Lg \in C^\alpha_c(\mathbb{R}^d)$, we may use [62, §V Theorem 6.1] to establish Hölder regularity of $\partial_t v$ (e.g. by approximations on bounded sets with zero boundary data). It follows that

$$
\partial_t F'(L_u) = \partial_t F'(v) = F'''(v) \partial_t v
$$

is a Hölder-continuous function. We now switch to the Fokker–Planck equation. Let $b = F'(L_u)$ and $\mu = bm$. Note that $b \geq \kappa > 0$ is Hölder-continuous. We have

$$
b \partial_t m = bL(bm) \quad \text{and} \quad \partial_t \mu = bL\mu - \frac{\partial b}{b} \mu.
$$

By applying Schauder theory (this is a linear equation), we find a classical solution. A similar result holds when $L = -(-\Delta)^\sigma$ (see Definition A.1) and $Lf = 0$ [85, Theorem 1.1].

2.4. General well-posedness theory. Let $\mathcal{S}_{HJB} \subset C_b(\mathcal{T} \times \mathbb{R}^d)$ be the set of bounded classical solutions of problem (2) with data $(f, g) = (f(m), g(m(T)))$ for all $m \in C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$. Also denote

$$
\mathcal{B} = \left\{ \int_0^1 F'(sLu_1 + (1-s)Lu_2) \, ds : u_1, u_2 \in \mathcal{S}_{HJB} \right\}.
$$

With these definitions we describe the properties of solutions to problem (2) and problem (3) that lead to well-posedness of problem (1) (see Theorem 2.9).

1. For every $m \in C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$ there exists a bounded classical solution $u$ of problem (2) with data $(f, g) = (f(m), g(m(T)))$.

2. If $\{u_n, u\}_{n \in \mathbb{N}} \subset \mathcal{S}_{HJB}$ are such that $\lim_{n \to \infty} \|u_n - u\|_\infty = 0$, then $\mathcal{L}u_n(t) \to \mathcal{L}u(t)$ uniformly on compact sets in $\mathbb{R}^d$ for every $t \in \mathcal{T}$.

3. There exists a constant $K_{HJB} \geq 0$ such that $\|F'(\mathcal{L}u)\|_\infty \leq K_{HJB}$ for every $u \in \mathcal{S}_{HJB}$.

4. The set $\{\partial_t u, \mathcal{L}u : u \in \mathcal{S}_{HJB}\} \subset C_b(\mathcal{T} \times \mathbb{R}^d)$.

5. For each $b \in \mathcal{B} \cap C_b(\mathcal{T} \times \mathbb{R}^d)$ and initial data $m_0 \in \mathcal{P}(\mathbb{R}^d)$ there exists at most one very weak solution of problem (3).

\footnote{Note that $F'(\mathcal{L}(u)) \in \mathcal{B}$ for any $u \in \mathcal{S}_{HJB}$. The set $\mathcal{B}$ is only used in the proof of uniqueness of solutions of problem (1), specifically in (63), via (S5).}
Assumption (S1) describes existence of solutions of the Hamilton–Jacobi–Bellman equation, which are unique by Theorem 5.3, and (S5) describes uniqueness of solutions of the Fokker–Planck equation, which exist by Theorem 6.6. Conditions (S2), (S3), (S4) describe various (related) properties of solutions of problem (2). Under assumption (A1), both (S3) and (S4) imply $b = F'(Lu) \in C_b(T \times \mathbb{R}^d)$ for $u \in S_{HJB}$.

Our main result is the following.

**Theorem 2.9.** Assume (L), (A1), (A3). If in addition

(i) $(A4), (S1), (S2), (S3)$ hold, then there exists a classical–very weak solution of problem (1);

(ii) $(A2), (A5), (S4), (S5)$ hold, then problem (1) has at most one classical–very weak solution.

We prove these results in Section 8. Existence is addressed in Theorem 8.5 by an application of the Kakutani–Glicksberg–Fan fixed point theorem, which requires a detailed analysis of problem (2) and problem (3). Of particular interest are the compactness and stability results Lemma 6.2, Corollary 6.3, and Lemma 6.4 for the Fokker–Planck equation.

Uniqueness follows by Theorem 8.7. Note that in contrast to previous work (cf. e.g. [1, (1.24), (1.25)]) we only need (non-strict) convexity of $F$ in (A2) and (non-strict) monotonicity of $f$ and $g$ in (A5), without further restrictions (this remains true for all the particular examples we study).

**Proofs of Theorems 2.1, 2.5, 2.8.** The well-posedness results for degenerate and non-degenerate cases (Sections 2.1–2.3) follow by verifying the general conditions (S1)–(S5) and then applying Theorem 2.9.

For Theorem 2.1, conditions (S1)–(S4) essentially follow from the Hölder regularity of solutions of problem (2), which is a direct consequence of the comparison principle for viscosity solutions (see Theorem 5.1). These results are gathered in Theorem 5.4 and Corollary 5.5. Condition (S5) follows from a Holmgren-type uniqueness proof\(^5\) for problem (3) after a construction of a suitable test function which solves a strongly degenerate dual equation. The construction of such function relies on viscosity solution techniques, a bootstrapping method, and the nonlocal nature of the problem. Put together, our procedure seems to be new. See Theorem 2.10 and Corollary 7.11\(^{i)}\).

For Theorem 2.5 and Theorem 2.8, we obtain (S1)–(S5) from Theorem 5.7 and Theorem 7.10\(^{iii)}, (iv)\) — see Corollary 5.8 and Corollary 7.11\(^{iii)}, (iv)\) (cf. also Remark 5.9). Keep in mind that Theorem 2.8, as stated, also depends on the validity of Conjecture 2.7.

Other approaches to verify (S1)–(S5), avoiding interior regularity estimates for nonlinear equations (as in Definition 2.3), are also possible, at least in some limited scope. We can mention an application of the method of continuity or [67, Theorem 9.2.6]. In an upcoming paper, we plan to investigate the regularity using a natural correspondence between problem (2) and the porous medium equation.

2.5. Other contributions.

**Fokker–Planck equations.** A substantial part of this paper studies the Fokker–Planck equation (3). We prove the following theorem.

---

\(^5\)Holmgren’s method is based on a duality argument and goes back the early 20th century and linear PDEs [83, Chapter 5]. Such arguments are widely used in the modern theory of PDEs, including nonlinear equations, see e.g. [84].
**Theorem 2.10.** Assume \((L), (A3), b \geq 0, and b \in C_b(\mathcal{T} \times \mathbb{R}^d)\). Then there exists a very weak solution of problem \((3)\). This solution is unique if either of the following additional conditions hold:

(i) \((LA)\) and \(b \in B(\mathcal{T}, C^2_b(\mathbb{R}^d))\) for \(\beta > 2\sigma + \frac{2\sigma}{1-2\sigma}\) (see Definition 4.3);

(ii) \((LB)\), \(b \geq \kappa, and b \in B(\mathcal{T}, C^2_b(\mathbb{R}^d)) \cap UC([0,t] \times \mathbb{R}^d)\) for every \(t \in \mathcal{T}\) and constants \(\kappa > 0, \beta \in (0,1]\);

(iii) \((LC)\), \(b \geq \kappa, and b \in B(\mathcal{T}, C^2_b(\mathbb{R}^d))\) for constants \(\kappa > 0, \beta \in (0,1]\).

Existence of solutions is established in **Theorem 6.6.** Uniqueness is addressed in **Theorem 7.10** by a Holmgren-type argument, using existence of smooth solutions of the dual problem \((35)\). The degenerate case in part (i) and its novel proof is of particular interest, see the discussion in Section 2.1.

Note that problem \((3)\) is the forward Kolmogorov equation for the SDE

\[
(5) \quad dZ(t) = b(t, Z(t)) \, dX(t), \quad Z(0) = Z_0 \sim m_0,
\]

where \(X\) is the Lévy process with infinitesimal generator \(\mathcal{L}\). When \(X\) is a Brownian Motion (with drift) and \(\mathcal{L}\) is local, i.e. its triplet is \((c, a, 0)\), we refer e.g. to [1, §3.4] for a brief survey of classical results on well and ill-posedness (see also [11, 68] and the references therein). For general Lévy processes, we mention the recent results of [58] for \(b\) independent of \(t\) and [36, 77] for connections to the Fokker–Planck equation. If \(\mathcal{L} = c \cdot \nabla\) and \(b\) is continuous, but not Lipschitz-continuous, then problem \((3)\) does not have a unique solution (see [12]). Another relevant pathological example is constructed in [47]. Uniqueness for local degenerate equations can be found in [12, 35, 78], while for nonlocal degenerate problems we are not aware of any prior results.

**Mean field games in \(\mathbb{R}^d\) without moment assumptions.** In the mean field game literature (see e.g. [1]), it is common to use the Wasserstein-1 space \((\mathcal{P}_1, d_1)\) in the analysis of the Fokker–Planck equations (or Wasserstein-\(p\) for \(p > 1\)). This is the space \(\mathcal{P}_1\) of probability measures with finite first moments and the metric \(d_1\) describing weak convergence of measures together with convergence of their first moments. To obtain compactness, typically \(1 + \varepsilon\) finite moments are assumed.

Moments of solutions of the Fokker–Planck equation depend on both the driving Lévy process and the initial distribution. Lévy processes have the same kind of moments as the tail of their Lévy measures [80, Theorem 25.3], e.g. the Brownian motion has 0 Lévy measure and moments of any order, while 2\(\sigma\)-stable processes with \(\nu(dz) = c \frac{dz}{|z|^{\sigma+2}}\) only have finite moments of order less than \(2\sigma \in (0, 2)\). Conditions \((LA)\) and \((LC)\) impose no restrictions on the tails of the Lévy measures. Therefore the mean field games we consider may be driven by processes with unbounded first moments, like the \(2\sigma\)-stable processes for \(2\sigma \leq 1\). This means that we cannot work in \((\mathcal{P}_1, d_1)\), even when the initial distribution has finite moments of all orders, or more, is compactly supported.

In this paper we develop tools to analyse mean field games in \(\mathbb{R}^d\) in a setting without any (explicit) moment assumptions. We work in the space \((\mathcal{P}, d_0)\) of probability measures under weak convergence, metrised by \(d_0\), defined from the Rubinstein–Kantorovich norm \(\|\cdot\|_0\) (see Section 4.2). The \(d_0\)-topology is strictly weaker than the \(d_1\)-topology, as it does not require convergence of first moments. The tools developed here are useful for other problems as well and have already been used in the recent papers [33, 26]. In the local case they would give results for a larger class of initial distributions. The crucial ingredient is the more refined tightness results described below and their interplay with Lévy processes.
Tightness of measures and Lévy operators. The Prokhorov theorem [10, Theorem 8.6.2] is a classical result, which equates tightness and pre-compactness of a family of probability measures in the topology of weak convergence. Another useful characterization of tightness may be given in terms of a function with unbounded growth (see [10, Example 8.6.5]). As a slight modification of this result, we prove Lemma 4.9, for Lyapunov functions described in Definition 4.6, which may serve as a convenient tool in studying PDEs with Lévy operators and measure-valued (or \( L^1 \)) solutions on non-compact domains. Specifically, this lemma allows us to avoid moment assumptions on the initial data and the driving Lévy process (the Lévy measure tail), or a restriction to a compact domain like a torus (i.e. the periodic setting).

As an application, we prove several results regarding Lévy operators as defined in (L). In Lemma 4.17 we construct a family of approximate operators \( \{L^\varepsilon\}_{\varepsilon>0} \), which enjoys useful properties akin to discretization. Lemma 4.13 then provides an example where Lemma 4.9 is applied to a set of Lévy measures. Namely, (a) for every Lévy operator \( \mathcal{L} \) there exists a Lyapunov function \( V \) such that \( \|\mathcal{L}V\|_\infty < \infty \) (see Corollary 4.14), (b) this Lyapunov function may then play a role in an a priori estimate (see the proof of Lemma 6.2), and (c) the approximate operators \( \mathcal{L}^\varepsilon \) (uniformly satisfying the same bound by \( V \)) may then be used in a construction of approximate solutions e.g. to problem (3) (see the proof of Theorem 6.6).

3. Derivation of the model

In this section we show heuristically that problem (1) is related to a mean field game where players control the time change rate of a Lévy process. Random time change of SDEs is a well-established technique [4, 46, 71, 73] with applications e.g. in modelling markets or turbulence [3, 21]. For stable (self-similar) processes, including the Brownian motion, we note that this type of control coincides with the classical continuous control (see Section 3.5). However, for other Lévy processes, including compound Poisson and most jump processes used in finance and insurance, this is not the case.

This type of a control problem seems to be new and we plan to analyse it in full detail in a future paper.

3.1. Time changed Lévy process. We start by fixing a Lévy process \( X_t \) and the filtration \( \{\mathcal{F}_t\} \) it generates. The infinitesimal generator \( \mathcal{L} \) of \( X \) is given by (L).

**Definition 3.1** ([4, Definition 1.1]). A random time change \( \theta_s \) is an almost surely non-negative, non-decreasing stochastic process which is a finite stopping time for each fixed \( s \).\(^6\) It is absolutely continuous if there exists a non-negative \( \mathcal{F}_s \)-adapted process \( \theta' \) such that \( \theta(s) - \theta(0) = \int_0^s \theta'(\tau) \, d\tau \).

For \( (t,x) \in \mathcal{T} \times \mathbb{R}^d \) and \( s \geq t \), we define an \( \mathcal{F}_s \)-adapted Lévy process \( X_{t,x}^s \) starting from \( X_t^s = x \) by

\[
X_{t,x}^s = x + X_s - X_t.
\]

Then, for an absolutely continuous random time change \( \theta_s \) such that \( \theta_t = t \), \( \theta'_s \) is deterministic, and \( \theta_{s+h} - \theta_s \) is independent of \( \mathcal{F}_{\theta_s} \) for all \( s,h \geq 0 \), we define a time-changed process

\[
Y_{s,x}^{t,\theta} = X_{\theta_s}^t.
\]

\(^6\) \( \theta_s \) is a stopping time if \( \{\theta_s \leq \tau\} \subset \mathcal{F}_\tau \) for \( \tau \geq 0 \).
It is an inhomogeneous Markov process associated with the families of operators $P^\theta$ and transition probabilities $p^\theta$ (see [37, §1.1, §1.2 (10)]) given by

$$P^\theta_{t,s}\phi(x) = \int_{\mathbb{R}^d} \phi(y)p^\theta(t,x,dy) = E\phi(Y_t^{t,x,\theta})$$

for $\phi \in C_b(\mathbb{R}^d)$. To compute the “generator” $L_\theta$ of $Y_t^{t,x,\theta}$, note that by the Dynkin formula [14, (1.55)],

$$E\phi(Y_t^{t,x,\theta}) - \phi(x) = E\left(\int_t^{t+h} L_\theta \phi(X_{\tau}^{t,x,\theta}) d\tau\right),$$

and by a change of variables,

$$P^\theta_{t+h,t}\phi(x) - \phi(x) = \frac{E\phi(Y_{t+h}^{t,x,\theta}) - \phi(x)}{h} = \frac{1}{h} \int_t^{t+h} L_\theta \phi(X_{\tau}^{t,x,\theta}) d\tau.$$

Under some natural assumptions we can show that $X_{t}^{t,x,\theta} \to x$ as $\tau \to t$ and use the dominated convergence theorem etc. to get that

$$L_\theta \phi(x) = \lim_{h \to 0^+} \frac{P^\theta_{t+h,t}\phi(x) - \phi(x)}{h} = \theta_1 \frac{\partial}{\partial \tau} \phi(x).$$

A proof of a more general result can be found in e.g. [4, Theorem 8.4].

3.2. Control problem and Bellman equation. To control the process $Y_t^{t,x,\theta}$, we introduce a running gain (profit, utility) $\ell$, a terminal gain $g$, and an expected total gain functional

$$J(t,x,\theta) = E\left(\int_t^T \ell(s,Y_s^{t,x,\theta},\theta'_s) ds + g(Y_T^{t,x,\theta})\right).$$

The goal is to find an admissible control $\theta^*$ that maximizes $J$. If such a control exists, the optimally controlled process is given by $Y_t^{t,x,\theta^*}$.

Under a suitable definition of the set of admissible controls $\mathcal{A}$ and standard assumptions on $\ell$ and $g$, $J$ is well-defined. The corresponding value function $u$ (the optimal value of $J$) is given by

$$u(t,x) = \sup_{\theta \in \mathcal{A}} J(t,x,\theta).$$

Let $h > 0$ and $t+h < T$. By the dynamic programming principle,

$$u(t,x) = \sup_{\theta} E\left(\int_t^{t+h} \ell(s,Y_s^{t,x,\theta},\theta'_s) ds + u(t+h,Y_{t+h}^{t,x,\theta})\right),$$

and hence

$$-u(t+h,x) - u(t,x) = \frac{1}{h} \int_t^{t+h} \ell(s,Y_s^{t,x,\theta},\theta'_s) ds.$$
where \( L : [0, \infty) \to \mathbb{R} \cup \{ \infty \} \) is a convex, lower-semicontinuous function. Then the Bellman equation can be expressed in terms of the Legendre–Fenchel transform \( F \) of \( L \), i.e. \( F(\zeta) = \sup_{\zeta \geq 0} (\zeta \zeta - L(\zeta)) \), as

\[
-\partial_t u = F'(Lu) + f(t, x).
\]

By the definitions of \( u \) and \( X^{T,x}_t \) it also follows that

\[
u(T, x) = E_g(X^{T,x}_{\tau^*}) = g(x).
\]

3.3. **Optimal control and Fokker–Planck equation.** By the properties of the Legendre–Fenchel transform, when \( \lim_{\zeta \to \infty} \frac{L(\zeta)}{\zeta} = \infty \) and \( L \) is strictly convex on \( \{ L \neq \infty \} \), the optimal value \( \zeta \) in (9) satisfies \( \zeta = F'(Lu) \) for every \((t, x) \in T \times \mathbb{R}^d \) (see Proposition B.1). We therefore obtain a function

\[
b(t, x) = \zeta = (\theta^*)' = F'(Lu(t, x)).
\]

This is the optimal time change rate in the feedback form. The optimally controlled process and the optimal control in (8) are then implicitly given by

\[
Y^*_s = X^{T,x}_{\theta^*_s}, \quad \theta^*_s = t + \int_t^s b(\tau, Y^*_\tau) \, d\tau.
\]

They are well-defined if \( b \) is e.g. bounded and continuous.

By defining \( p^\theta(t, x, s, A) = \mathbb{P}(Y^*_s \in A) \), if solutions of equations (11)–(12) are unique, we obtain a unique family of transition probabilities \( p^\theta \) (cf. (6)), satisfying the Chapman–Kolmogorov relations. This family, in turn, defines a wide-sense Markov process (see [37, §1.1 Definition 1]). Given an initial condition \( m(0) = m_0 \in \mathcal{P}([\mathbb{R}^d]) \), the (input) distribution \( m \) of this Markov process (see [37, §1.1 Definition 3])\footnote{Alternatively, we may take a random variable \( Z_0 \) with distribution \( m_0 \). Then \( m(t) \) is a distribution of the solution \( Z(t) \) of SDE (5). Moreover, \( Y^*_s = \mathbb{E}[Z(s) | Z(t) = x] \), see [37, §1.2 (9), (10)].} satisfies

\[
\int_{\mathbb{R}^d} \varphi(x) m(t + h, dx) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(y) p^\theta(t, x, t + h, dy) m(t, dx),
\]

for every \( \varphi \in C^\infty_c(\mathbb{R}^d) \) and \( t, h \geq 0 \). Then,

\[
\int_{\mathbb{R}^d} \varphi(t, x) m(t, dx) - \varphi(t + h, x) m(t + h, dx)
= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left( \varphi(x, y) p^\theta(t, x, t, dy) - \varphi(x + h, y) p^\theta(t, x + h, dy) \right) m(t, dx),
\]

and because of (7), (13) and the fact that \( p^\theta(t, x, t, dy) = \delta_x(dy) \), this leads to

\[
\partial_t \int_{\mathbb{R}^d} \varphi(t, x) m(t, dx) = \int_{\mathbb{R}^d} \left( b(t, x) \mathcal{L} \varphi + \partial_x \varphi(t, x) \right) m(t, dx).
\]

Since \( b = F'(Lu) \), by duality (see Definition 1.4) \( m \) is a very weak solution of

\[
\partial_t m = \mathcal{L}^*(F'(Lu) m), \quad m(0) = m_0,
\]

where \( \mathcal{L}^* \) is the formal adjoint of \( \mathcal{L} \).
3.4. Heuristic derivation of the mean field game. A mean field game is a limit of games between identical players as the number of players tends to infinity. In our case, each player controls the time change rate of her own independent copy of the Lévy process $X$, with running and terminal gains depending on the anticipated distribution $\hat{m}$ of the processes controlled (optimally) by the other players (see (A4))

$$f = f(\hat{m}) \quad \text{and} \quad g = g(\hat{m}(T)).$$

By the results of Section 3.2 the corresponding Bellman equation for each player is

$$\begin{cases} -\partial_t u = F(Lu) + f(\hat{m}) & \text{on } T \times \mathbb{R}^d, \\ u(T) = g(\hat{m}(T)) & \text{on } \mathbb{R}^d. \end{cases}$$

Note that the solution $u$ depends on $\hat{m}$, and then so does the optimal feedback control (13). Suppose that the players’ processes start from some known initial distribution $m_0 \in \mathcal{P}(\mathbb{R}^d)$. Then, the actual distribution $m$ of their optimally controlled processes is given by the solution of the Fokker–Planck equation (14), described in Section 3.3.

At a Nash equilibrium we expect $\hat{m} = m$, i.e. the anticipations of the players to be correct. The result is a closed model of coupled equations in the form of problem (1).

3.5. Example and a relation to continuous control. When the Lévy process $X$ is self-similar,\textsuperscript{8} the control of the time change rate can be interpreted as the classical continuous control, i.e. control of the size of the spatial increments of the process.

Let $L$ in (10) be given by $L(\zeta) = \frac{1}{2}\zeta^2$ for $q > 1$, and assume the infinitesimal generator of $X$ is the fractional Laplacian, $\mathcal{L} = (-\Delta)^\sigma$ (see Definition A.1). We have $F(z) = \frac{2-1}{q}(z^+)^{\frac{q}{q+1}}$ (cf. Table 1 in the Appendix) and $\mathcal{L}^* = \mathcal{L}$, hence the mean field game system takes the form

$$\begin{cases} -\partial_t u = \frac{q-1}{q} \left((-\Delta)^\sigma u^+\right)^{\frac{q}{q-1}} + f(m), \\ \partial_t m = (-\Delta)^\sigma \left((-\Delta)^\sigma u^+ m\right)^{\frac{q}{q-1}} . \end{cases}$$

These equations are degenerate and $F$ satisfies (A1) with $\gamma = \frac{1}{q-1}$, as well as (A2). Existence of solutions of problem (15) follows from Theorem 2.1(i) if $2\sigma \in (0,1)$, $m_0$ satisfies (A3), $f, g$ satisfy (A4) and (RA).

If e.g. $\alpha = 1$, $q < q_\sigma(\sigma) = \frac{1+\sigma}{2\sigma(1-\sigma)}$, and (A5) holds, we also have uniqueness (see Theorem 2.1(iii)). Note that $q_\sigma$ is decreasing, $q_\sigma(\frac{1}{2}) = 1$, and $\lim_{\sigma \to 0^+} q_\sigma = \infty$.

Next consider the optimal control problem (8) with non-negative control processes $\lambda$ replacing $\theta$ and controlled process $Y_t$ given by the SDE

$$dY_t = \lambda_s dX_s = \lambda_s \int_{\mathbb{R}^d} z\tilde{N}(dt,dz), \quad \text{and} \quad Y_t = x,$$

where $\tilde{N}$ is the compensated Poisson measure defined from $X$.\textsuperscript{9} This is a classical control problem, and under suitable assumptions it leads to the following Bellman equation (see [8, 39])

$$\begin{equation} -\partial_t u = \sup_{\lambda} \left( \text{p.v.} \int_{\mathbb{R}^d} \left( u(x + \lambda z) - u(x) \right) \frac{c_d \nu}{|z|^d + 2\nu} dz - \tilde{L}(\lambda) + f(s,x) \right), \tag{16} \end{equation}$$

\textsuperscript{8} $X$ is self-similar if there exists $c > 0$ such that for all $a, t > 0$, $a^{-1}X_{at} = X_t$ in distribution.

\textsuperscript{9} $\tilde{N}(dt,dz) = N(dt,dz) - 1_{B_t}(z)\nu(dz) dt$ where is $N$ is the Poisson measure with intensity $\nu$.\textsuperscript{9}
where p.v. denotes the principal value. Self-similarity (seen through ν) then yields
\[
p.v. \int \left( u(x + \lambda z) - u(x) \right) \frac{c_{d,\sigma}}{z^{d+2\sigma}} \, dz = \lambda^{2\sigma} \, p.v. \int \left( u(x + z) - u(x) \right) \frac{c_{d,\sigma}}{z^{d+2\sigma}} \, dz = -\lambda^{2\sigma} (-\Delta)^{\sigma} u(x).
\]
Let \( \lambda^{2\sigma} = \zeta \) and \( \tilde{L}(\lambda) = \frac{\lambda}{2} \lambda^{\frac{d}{2\sigma}} = L(\zeta) \), and \( f = f(m) \). Then the Bellman equations in (15) and (16) coincide. This means that in this case the classical continuous control problem and the original controlled time change rate problem coincide as well.

4. Preliminaries

4.1. Basic notation. By \( K_d = 2\pi^{d/2} \Gamma(d/2)^{-1} \) we denote the surface measure of the \( (d-1) \)-dimensional unit sphere. By \( B_r \) and \( B^c_r \) we denote the ball of radius \( r \) centred at 0 and its complement in \( \mathbb{R}^d \), respectively. Similarly, \( B_r(x) \) denotes a ball centred at \( x \).

**Definition 4.1.** A function \( \phi \) is Hölder-continuous at \( x \in \mathbb{R}^d \) with parameter \( \alpha \in (0, 1] \) if for some \( r > 0 \)
\[
[\phi]_{C^\alpha(B_r(x))} = \sup_{y \in B_r(x) \setminus \{x\}} \frac{|\phi(x) - \phi(y)|}{|x - y|^{\alpha}} < \infty.
\]
The space \( C^\alpha(\mathbb{R}^d) \) consists of functions which are Hölder-continuous at every point in \( \mathbb{R}^d \) with parameter \( \alpha \). Further, define
\[
[\phi]_\alpha = \sup_{x \in \mathbb{R}^d} [\phi]_{C^\alpha(B_r(x))} \quad \text{and} \quad \|\phi\|_\alpha = \|\phi\|_{L^\infty(\mathbb{R}^d)} + [\phi]_\alpha.
\]

We denote \( C^\alpha_0(\mathbb{R}^d) = \{ \phi : \|\phi\|_\alpha < \infty \} \).

Note that the definition of \( C^\alpha_0(\mathbb{R}^d) \) is equivalent to the more standard notation, where the supremum in (17) is taken over \( |x - y| \in \mathbb{R}^d \setminus \{0\} \). The space \( C^1_0(\mathbb{R}^d) \) consists of bounded, Lipschitz-continuous functions. By \( C^1(\mathbb{R}^d) \), \( C^2(\mathbb{R}^d) \) we denote spaces of once or twice continuously differentiable functions.

**Definition 4.2.** For \((t, x) \in T \times \mathbb{R}^d \) and \( \alpha, \beta \in (0, 1] \), define
\[
[\phi]_{C^{\beta,\alpha}([0, t] \times B_r(x))} = \sup_{y \in B_r(x)} |\phi(y)|_{C^{\beta}([0, t])} + \sup_{s \in [0, t]} [\phi(s)]_{C^{\alpha}(B_r(x))}.
\]
We also denote \( C^{\beta,\alpha}_0([0, t] \times \mathbb{R}^d) = \{ \phi : \|\phi\|_{C^{\beta,\alpha}([0, t] \times \mathbb{R}^d)} < \infty \} \), where
\[
\|\phi\|_{C^{\beta,\alpha}([0, t] \times \mathbb{R}^d)} = \|\phi\|_{L^\infty([0, t] \times \mathbb{R}^d)} + \sup_{x \in \mathbb{R}^d} [\phi]_{C^{\beta,\alpha}([0, t] \times B_r(x))}.
\]

**Definition 4.3.** By \( B(T, X) \), where \( X \) is a normed space, we denote the space of bounded functions from \( T \) to \( X \), namely
\[
B(T, X) = \{ u : T \to X : \sup_{t \in T} \|u(t)\|_X < \infty \}.
\]
Note the subtle difference between \( B(T, X) \) and the usual space \( L^\infty(T, X) \).

4.2. Spaces of measures. Let \( \mathcal{P}(\mathbb{R}^d) \) be the space of probability measures on \( \mathbb{R}^d \), a subspace of the space of bounded, signed Radon measures \( M_b(\mathbb{R}^d) = C_0(\mathbb{R}^d)^* \). We denote
\[
m[\phi] = \int_{\mathbb{R}^d} \phi(x) \, m(dx) \quad \text{for every} \ m \in \mathcal{P}(\mathbb{R}^d) \ \text{and} \ \phi \in C_b(\mathbb{R}^d).
\]
The space \( P(\mathbb{R}^d) \) is equipped with the topology of weak convergence of measures,\(^{10}\) namely
\[
\lim_{n \to \infty} m_n = m \quad \text{if and only if} \quad \lim_{n \to \infty} m_n[\phi] = m[\phi] \quad \text{for every} \quad \phi \in C_0(\mathbb{R}^d).
\]
This topology can be metrised by an embedding into a normed space (see [10, §8.3]).

**Definition 4.4.** The Rubinstein–Kantorovich norm \( \| \cdot \|_0 \) on \( M_0(\mathbb{R}^d) \) is given by
\[
\|m\|_0 = \sup \{ m[\psi] : \psi \in C_0(\mathbb{R}^d), \|\psi\|_\infty \leq 1, \|\psi\|_1 \leq 1 \}.
\]

While the space \( (M_0(\mathbb{R}^d), \| \cdot \|_0) \) is not completely metrisable, thanks to [51, Theorems 4.19 and 17.23], both \( P(\mathbb{R}^d) \) and \( C(\overline{\mathbb{T}}, P(\mathbb{R}^d)) \) are complete spaces.

Consider the set \( \mathcal{P}_{ac}(\mathbb{R}^d) = L^1(\mathbb{R}^d) \cap P(\mathbb{R}^d) \), i.e.
\[
\mathcal{P}_{ac}(\mathbb{R}^d) = \{ u \in L^1(\mathbb{R}^d) : \|u\|_{L^1(\mathbb{R}^d)} = 1, \ u \geq 0 \}.
\]
We endow \( \mathcal{P}_{ac}(\mathbb{R}^d) \) with the topology inherited from \( P(\mathbb{R}^d) \).

**Definition 4.5.** A set of measures \( \Pi \subset P(\mathbb{R}^d) \) is tight if for every \( \varepsilon > 0 \) there exists a compact set \( K_\varepsilon \subset \mathbb{R}^d \) such that for every \( m \in \Pi \) we have \( m(K_\varepsilon) \geq 1 - \varepsilon \).

This concept is important because of the Prokhorov theorem, which states that a set \( \Pi \subset P(\mathbb{R}^d) \) is pre-compact if and only if it is tight.

**Definition 4.6.** A real function \( V \in C^2(\mathbb{R}^d) \) is a Lyapunov function if \( V(x) = V_0(\sqrt{1 + |x|^2}) \) for some subadditive, non-decreasing function \( V_0 : [0, \infty) \to [0, \infty) \) such that \( \|V_0\|_{\infty}, \|V_0''\|_{\infty} \leq 1 \), and \( \lim_{x \to \infty} V_0(x) = \infty \).

**Remark 4.7.** (a) Because \( \|V_0\|_{\infty}, \|V_0''\|_{\infty} \leq 1 \), we also have \( \|\nabla V\|_{\infty}, \|D^2V\|_{\infty} \leq 1 \). Note that the choice of the constant 1 in this condition is arbitrary.

(b) \( 1 + |x|^2 \) is a Lyapunov function (see also Proposition A.4).

(c) If \( m_0 \in P(\mathbb{R}^d) \) has a finite first moment and \( V \) is any Lyapunov function, then \( m_0[V] < \infty \). Indeed, since \( 0 \leq V_0 \leq 1 \), we have \( V(x) = V(0) + |x|, \) thus \( m_0[V] \leq V(0) + \int_{\mathbb{R}^d} |x| \, dm_0 \).

**Proposition 4.8.** If \( V \) is a Lyapunov function, then for every \( r > 0 \) the set
\[
\mathcal{P}_{V,r} = \{ m \in \mathcal{P}(\mathbb{R}^d) : m[V] \leq r \}
\]
is tight and then compact by the Prokhorov theorem.

**Proof.** Notice that the set \( \mathcal{P}_{V,r} \) is closed. Let \( \varepsilon > 0 \). Since \( \lim_{|x| \to \infty} V(x) = \infty \), the set \( K_\varepsilon = \{ x \in \mathbb{R}^d : V(x) \leq \frac{\varepsilon}{r} \} \) is compact. Then it follows from the Chebyshev inequality that for every \( m \in \mathcal{P}_{V,r} \),
\[
m(K_\varepsilon) \leq \frac{\varepsilon}{r} \int_{(V > \frac{\varepsilon}{r})} V \, dm \leq \frac{\varepsilon}{r} m[V] \leq \varepsilon.
\]
Hence the set \( \mathcal{P}_{V,r} \) is tight and thus compact by the Prokhorov theorem. \( \square \)

The reverse statement is also true.

**Lemma 4.9.** If the set \( \Pi \subset P(\mathbb{R}^d) \) is tight, then there exists a Lyapunov function \( V \) such that \( m[V] \leq 1 \) for every \( m \in \Pi \).

\(^{10}\)It is also called narrow, vague or \( w^* \)-convergence.
Proof. We proceed in steps, constructing successive functions, which accumulate properties required by Definition 4.6 and are adequately integrable.

\*Step 1. Integrability, monotonicity, unboundedness. The conclusion of this step is essentially stated in [10, Example 8.6.5 (iii)], but a complete proof is lacking and the precise function \( v_0 \), which we need, cannot be extracted. Let

\[
v(x) = v_0(|x|), \quad \text{where} \quad v_0(t) = \sup_{m \in \Pi} m \{x : |x| \geq t\}.
\]

Then \( v_0 : [0, \infty) \to [0, 1] \) is a non-increasing function such that \( v_0(0) = 1 \). Because \( \Pi \) is tight, we also have \( \lim_{t \to \infty} v_0(t) = 0 \).

For \( m \in \Pi \), let \( \Phi^m(\tau) = m \circ v^{-1}([0, \tau]) \). Then (see Remark 4.10)

\[
\Phi^m(\tau) = m(v^{-1}([0, \tau])) = m \{x : \forall \tilde{m} \in \Pi \; \tilde{m} \{y : |y| \geq |x| \} < \tau \}
\]

\[
\leq m \{x : m\{y : |y| \geq |x| \} < \tau \} \leq \tau.
\]

By “change of variables” [10, Theorem 3.6.1] and integration by parts [10, Exercise 5.8.112], this gives us

\[
\int_{\mathbb{R}} - \log(v(x)) \; m(dx) = \int_0^1 - \log(\tau) \; d\Phi^m(\tau) = \int_0^1 \frac{\Phi^m(\tau)}{\tau} \; d\tau \leq \int_0^1 \frac{1}{\tau} \; d\tau.
\]

Notice that \(- \log(v_0) : [0, \infty) \to [0, \infty] \) is non-decreasing, \( \log(v_0(0)) = 0 \), and \( \lim_{t \to \infty} - \log(v_0(t)) = \infty \).

\[\begin{align*}
&v_1 = -\log(v_0) \\
&\frac{1}{2} < a_1 < a_2 < a_3 \\
&\frac{1}{8} < \frac{1}{2} \text{ and } \frac{1}{3} < a_2 \geq \frac{2}{3}
\end{align*}\]

Figure 1. Comparison of \(- \log(v_0) \) and \( v_1 \).

\*Step 2. Continuity, concavity.\(^{12}\) For \( N \in \mathbb{N} \cup \{\infty\} \) and sequences \( \{a_n\}, \{b_n\} \) to be fixed later, let \( v_1 : [0, \infty) \to [-1, \infty) \) be the piecewise affine function given by (see Figure 1)

\[
v_1(t) = \sum_{n=0}^N l_n(t) \mathbbm{1}_{[a_n, a_{n+1})}(t), \quad \text{where} \quad l_n(t) = 2^{-n}(t - a_n) + b_n.
\]

We set \( a_0 = 0 \). For \( n \in \mathbb{N} \), when \( a_n < \infty \), let \( b_n = - \log(v_0(a_n)) - 2^{-n} \) and

\[
a_{n+1} = \inf A_n, \quad \text{where} \quad A_n = \{t \geq a_n : - \log(v_0(t)) - l_n(t) \leq 2^{-n-1}\}.
\]

We put \( \inf \emptyset = \infty \) and \( N = \sup \{n : a_n < \infty\} \). Note that for every \( n < N + 1 \),

\[
- \log(v_0(a_n)) - v_1(a_n) = - \log(v_0(a_n)) - b_n = 2^{-n}
\]

\(^{11}\)From [10, Exercise 5.8.112 (i)] we get \( \int_1^{\infty} - \log(\tau) \; d\Phi^m(\tau) = \int_1^{\infty} \frac{\Phi^m(\tau)}{\tau} \; d\tau \) for every \( r > 0 \). Then we may pass to the limit \( r \to 0 \) by the monotone convergence theorem, cf. [10, Exercise 5.8.112 (iii)].

\(^{12}\)Concavity serves as an intermediate step to obtain subadditivity.
and on the interval \([a_n, a_{n+1}]\),
\[
- \log (v_0) - v_1 \geq 2^{-n-1} \quad \text{(hence } - \log (v_0(t)) \geq v_1(t) \text{ for every } t \geq 0).\]

To verify continuity, take a sequence \(\{s_k\} \subset A_n\) such that \(\lim_{k \to \infty} s_k = a_{n+1}\). Then, because \(- \log (v_0)\) is non-decreasing and \(l_n\) is continuous,
\[
- \log (v_0(a_{n+1})) - l_n(a_{n+1}) \leq \liminf_{k \to \infty} \left(- \log (v_0(s_k)) - l_n(s_k)\right) \leq 2^{-n-1}.
\]

Thus
\[
- \log (v_0(a_{n+1})) - l_n(a_{n+1}) = 2^{-n-1},
\]
i.e. \(l_{n+1}(a_{n+1}) = b_{n+1} = l_n(a_{n+1})\), which implies that \(v_1\) is continuous. Moreover, \(a_{n+1} - a_n \geq \frac{1}{2}\), since this distance is the shortest when \(\log (v_0)\) is constant on \([a_n, a_{n+1}]\). We have \(v_1(0) = -1\), \(\lim_{t \to \infty} v_1(t) = \infty\), and
\[
v_1' = \sum_{n=0}^{N} 2^{-n} \mathbb{1}_{(a_n, a_{n+1})} \quad \text{(a non-increasing function, see Figure 2)},
\]
which implies that \(v_1\) is concave. In addition, \(v_1(t) \leq t - 1\), hence \(v_1(1) \leq 0\).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2}
\caption{Comparison of \(v_1'\) and \(v_2'\)}
\end{figure}

\(\diamond\) Step 3. Differentiability. Let \(p(t) = \frac{1}{4}(t^3 - 3t + 6)\mathbb{1}_{[-1,1]}(t)\). Then \(p\) acts as a smooth transition between values 2 and 1 on the interval \([-1, 1]\), with vanishing derivatives at the end points. Let \(v_2\) be such that \(v_2(0) = -1\) and (see Figure 2)
\[
v_2'(t) = \mathbb{1}_{[a_1, a_2 - \frac{1}{2}]}(t) + \sum_{n=1}^{N} 2^{-n} \left(p(t - a_n) + \mathbb{1}_{[a_n + \frac{1}{8}, a_{n+1} - \frac{1}{8}]}(t)\right).
\]

Then \(v_2 \in C^2([0, \infty))\), \(v_2\) is concave, increasing, and \(\lim_{t \to \infty} v_2(t) = \infty\). Moreover,
\[
\|v''_2\| \leq \sup_t \left|\frac{1}{2} \frac{d}{dt} p(8t)\right| \leq 3.
\]

Next, we verify that \(v_2 \geq v_1\). Notice that for every \(t \in [-1, 1]\),
\[
\int_{-1}^{1} p(s) \, ds \leq \int_{-1}^{1} 2 \cdot \mathbb{1}_{[-1,0]}(s) + \mathbb{1}_{[0,1]}(s) \, ds, \quad \text{and} \quad \int_{-1}^{1} p(s) \, ds = 3.
\]

By suitable scaling and shifting, for every \(t \in \bigcup_{n=1}^{N} [a_n - \frac{1}{8}, a_n + \frac{1}{8}]\) we obtain \(v_2(t) \leq v_1(t)\), and \(v_2(t) = v_1(t)\) otherwise.

\(\diamond\) Step 4. Subadditivity, bounds on derivatives. Let \(V_0 = \frac{1}{2}(v_2 + 1)\). Then \(V_0 : [0, \infty) \to [0, \infty)\) is concave and hence subadditive. Moreover, \(V_0\) is increasing, \(\lim_{t \to \infty} V_0(t) = \infty\), and \(\|V_0''\|_{\infty} \leq 1\). This proves that \(V(x) = V_0(\sqrt{1 + x^2})\) is a Lyapunov function. By subadditivity and monotonicity we have
\[
V_0(\sqrt{1 + t^2}) \leq V_0(t + 1) \leq V_0(t) + V_0(1),
\]
hence for every \( m \in \Pi \), because \( v_2 \leq v_1 \leq -\log(v_0) \) and by (18),
\[
0 \leq \int_{\mathbb{R}^d} V(x) ma(dx) \leq V_0(1) + \int_{\mathbb{R}^d} V_0(|x|) ma(dx)
\]
\[
\leq \frac{v_2(1) + 1}{3} + \frac{1}{3} - \frac{1}{3} \int_{\mathbb{R}^d} \log(v(x)) ma(dx) \leq \frac{v_1(1)}{3} + \frac{1}{3} + \frac{1}{3} \leq 1.
\]
This shows that \( V \) is a Lyapunov function such that \( m[V] \leq 1 \) for every \( m \in \Pi \). □

**Remark 4.10.** Notice that
\[
\{ x : m\{ y : |y| \geq |x| \} < \tau \} = \{ x : |x| > r_\tau \},
\]
while
\[
\{ x : m\{ y : |y| > |x| \} \geq \tau \} = \{ x : |x| \geq r_\tau \},
\]
where \( r_\tau \) is such that
\[
m\{ x : |x| > r_\tau \} \leq \tau \leq m\{ x : |x| \geq r_\tau \}.
\]
If \( m \) is absolutely continuous with respect to the Lebesgue measure, then the measure \( m \) of both sets is equal to \( \tau \). Choosing the correct inequality in the definition of the function \( v_0 \) is essential to the proof of Lemma 4.9.

### 4.3. Lévy operators.

**Definition 4.11.** A Radon measure \( \nu \) on \( \mathbb{R}^d \setminus \{0\} \) such that \( \int_{\mathbb{R}^d} (1 \wedge |z|^2) \nu(dz) < \infty \) is a Lévy measure. A Lévy measure \( \nu \) is symmetric at the origin if \( \nu(A) = \nu(-A) \) for every \( A \subset B_1 \).

**Remark 4.12.** (a) If \( \nu \) is symmetric at the origin, then we may omit the term \( \mathbb{1}_{B_1}(z) z \cdot \nabla \phi(x) \) under the integral in (L), by considering the principal value integral around \( \{0\} \).

(b) If \( \int_{B_1} |z| \nu(dz) < \infty \), then we may equivalently write
\[
L\phi = \left( c - \int_{B_1} z \nu(dz) \right) \cdot \nabla \phi + \text{tr} \left( a a^T D^2 \phi \right) + \int_{\mathbb{R}^d} \left( \phi(x+z) - \phi(x) \right) \nu(dz).
\]
In particular, we may have \((\int_{B_1} z \nu(dz), 0, \nu)\) as a triplet in (L) (cf. (LA)).

**Lemma 4.13.** Assume (L) and \( V \) is a Lyapunov function. The following are equivalent

(i) \( \int_{B_1} V(z) \nu(dz) < \infty \);

(ii) \( \|LV\|_\infty < \infty \);

(iii) \( \vartheta_1(x) = \int_{B_1} (V(x+z) - V(x)) \nu(dz) \in L^\infty(\mathbb{R}^d) \);

(iv) \( \vartheta_2(x) = \int_{B_1} |V(x+z) - V(x)| \nu(dz) \in L^\infty(\mathbb{R}^d) \).

**Proof.** Let
\[
\vartheta_0(x) = c \cdot \nabla V(x) + \text{tr} \left( a a^T D^2 V(x) \right) + \int_{B_1} \left( V(x+z) - V(x) - z \cdot \nabla V(z) \right) \nu(dz).
\]
Because \( V \) is a Lyapunov function (see Remark 4.7(a)), we have
\[
\|\vartheta_0\|_\infty \leq |c| + |a|^2 + \int_{B_1} |z|^2 \nu(dz).
\]
Observe that \( \|LV\|_\infty - \|\vartheta_0\|_\infty \leq \|\vartheta_1\|_\infty \leq \|\vartheta_2\|_\infty \), hence (iv) \( \Rightarrow \) (iii) \( \Rightarrow \) (ii). We also notice \( \|LV\|_\infty \geq \|\vartheta_1\|_\infty - \|\vartheta_0\|_\infty \) and \( \int_{B_1} V(z) \nu(dz) = \vartheta_1(0) + \nu(B_1) \), thus (ii) \( \Rightarrow \) (iii) \( \Rightarrow \) (i).
Lemma 4.17. Assume (L). For every

\[ |V(y) - V(x)| \leq V_0 \left( |\sqrt{1 + |y|^2} - \sqrt{1 + |x|^2} | \right) \leq V_0 \left( \sqrt{1 + |y - x|^2} \right). \]

Now we may estimate

\[ \int_{B_1^c} |V(x + z) - V(x)| \nu(dz) \leq \int_{B_1^c} V(z) \nu(dz). \]

\[ \Box \]

Corollary 4.14. Assume (L), (A3). There exists a Lyapunov function \( V \) such that \( m_0[V], \|\mathcal{L}V\|_\infty < \infty \).

Proof. Since the measure \( \mathbb{1}_{B_1^c} \nu \) is bounded, the set \( \mathbb{1}_{B_1^c} \nu, m_0 \) is tight. By Lemma 4.9 we can thus find a Lyapunov function such that \( \int_{B_1^c} V(z) \nu(dz) < \infty \) and \( m_0[V] < \infty \). Thanks to Lemma 4.13 (ii) we also have \( \|\mathcal{L}V\|_\infty < \infty \).

Let \( \mathcal{L} \) be a Lévy operator with triplet \((c, a, \nu)\). Denote

\[ \|\mathcal{L}\|_{LK} = |c| + |a|^2 + \frac{1}{2} \int_{B_1} |z|^2 \nu(dz) + 2 \nu(B_1^c). \]

Proposition 4.15. Assume (L). For every \( \phi \in C^2_b(\mathbb{R}^d) \) we have

\[ \|\mathcal{L}\phi\|_\infty \leq \|\mathcal{L}\|_{LK}\|\phi\|_{C^2_b(\mathbb{R}^d)}. \]

Proof. Using the Taylor expansion, we calculate

\[ \|\mathcal{L}\phi\|_\infty \leq |c| \|\nabla \phi\|_\infty + |a|^2 \|D^2 \phi\|_\infty + \int_{\mathbb{R}^d} (\phi(x + z) - \phi(x) - \mathbb{1}_{B_1}(z) z \cdot \nabla \phi(x)) \nu(dz) \]

\[ \leq |c| \|\nabla \phi\|_\infty + |a|^2 \|D^2 \phi\|_\infty + \frac{\|D^2 \phi\|_\infty}{2} \int_{B_1} |z|^2 \nu(dz) + 2 \|\phi\|_\infty \nu(B_1^c). \]

\[ \Box \]

Remark 4.16. The mapping \( \mathcal{L} \mapsto \|\mathcal{L}\|_{LK} \) is a norm on the space (convex cone) of Lévy operators. It dominates the operator norm \( C^2_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d) \), but they are not equivalent.

Lemma 4.17. Assume (L). For \( \varepsilon \in (0, 1) \) there exists a family of operators \( \mathcal{L}^\varepsilon : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d) \) of the form

\[ \mathcal{L}^\varepsilon \mu(x) = \int_{\mathbb{R}^d} (\mu(x + z) - \mu(x)) \nu^\varepsilon(dz), \]

where \( \nu^\varepsilon(\mathbb{R}^d) < \infty \) and \( \text{supp} \nu^\varepsilon \subset \mathbb{R}^d \setminus B_\varepsilon \). Moreover,

(i) \( \|\mathcal{L}^\varepsilon \mu\|_{L^1(\mathbb{R}^d)} \leq (c_\mathcal{L}/\varepsilon^3) \|\mu\|_{L^1(\mathbb{R}^d)} \) for a constant \( c_\mathcal{L} > 0 \);

(ii) \( \lim_{\varepsilon \to 0} \|\mathcal{L}^\varepsilon \varphi - \mathcal{L} \varphi\|_\infty = 0 \) for every \( \varphi \in C_c^\infty(\mathbb{R}^d) \);

(iii) \( \sup_{\varepsilon \in (0, 1)} (\|\mathcal{L}^\varepsilon V\|_\infty + \|\mathcal{L}^\varepsilon\|_{LK}) < \infty \) for every Lyapunov function \( V \) such that \( \|\mathcal{L}V\|_\infty < \infty \).

Proof. Part (i). Let \( (c, a, \nu) \) be the Lévy triplet of \( \mathcal{L} \) and \( a = (a_1, \ldots, a_d) \in \mathbb{R}^{d \times d} \) with \( a_i \in \mathbb{R}^d \). Consider

\[ \nu_c^\varepsilon = \frac{|c|}{\varepsilon^3} \delta_{\varepsilon}, \quad \nu_1^\varepsilon(E) = \nu(E \setminus B_\varepsilon), \]

\[ \nu_a^\varepsilon = \sum_{i=1}^d \frac{|a_i|^2}{\varepsilon^2} (\delta_{E \setminus \varepsilon a_i} + \delta_{-E \setminus \varepsilon a_i}), \quad \nu_2^\varepsilon(E) = \frac{1}{\varepsilon} \nu \left( (B_1 \setminus B_\varepsilon) \cap (-E/\varepsilon) \right), \]

\[ \Box \]
and denote \( \nu^\varepsilon = \nu_1^\varepsilon + \nu_2^\varepsilon + \nu_3^\varepsilon + \nu_4^\varepsilon \). Notice that \( \nu^\varepsilon \) is a bounded, non-negative measure with \( \text{supp} \, \nu^\varepsilon \subset \mathbb{R}^d \setminus B_\varepsilon \) (hence a Lévy measure). Let \( \mathcal{L}^\varepsilon = \mathcal{L}^\varepsilon_{\text{loc}} + \mathcal{L}^\varepsilon_{\text{nonloc}}, \)

where, for \( \mu \in L^1(\mathbb{R}^d), \)

\[
\mathcal{L}^\varepsilon_{\text{loc}} \mu(x) = \int_{\mathbb{R}^d} \left( \mu(x + z) - \mu(x) \right) \nu_1^\varepsilon(dz)
\]

\[
= \left| \frac{c}{\varepsilon} \right| \left( \mu(x + \frac{a_z}{|a_z|}) - \mu(x) \right) + \frac{d}{\varepsilon^2} \left( \mu(x + \frac{a_z}{|a_z|}) + \mu(x) - 2\mu(x) \right),
\]

and

\[
\mathcal{L}^\varepsilon_{\text{nonloc}} \mu = \int_{\mathbb{R}^d} \left( \mu(x + z) - \mu(x) \right) \nu_2^\varepsilon(dz)
\]

\[
= \int_{B_\varepsilon} \left( \mu(x + z) - \mu(x) - \frac{1}{\varepsilon} \mu \left( \frac{z}{|z|} \right) \frac{x - \varepsilon z}{|z|} - \mu(x) \right) \nu(dz).
\]

Note that

\[
\nu_1^\varepsilon(B_1 \setminus B_\varepsilon) + \nu_2^\varepsilon(B_\varepsilon) = (1 + \varepsilon^{-1}) \nu(B_1 \setminus B_\varepsilon) \leq (\varepsilon^{-2} + \varepsilon^{-3}) \int_{B_1} |z|^2 \nu(dz),
\]

and hence

\[
\left\| \mathcal{L}^\varepsilon \mu \right\|_{L^1(\mathbb{R}^d)} \leq \left( \frac{2|c|}{\varepsilon} + \frac{4|a|^2}{\varepsilon^2} + 2
\right) \nu(B_1) \left( \frac{2 + 2\varepsilon}{\varepsilon^3} \int_{B_1} |z|^2 \nu(dz) \right) \left\| \mu \right\|_{L^1(\mathbb{R}^d)}
\]

\[
\leq \frac{4}{\varepsilon^{1/2}} \left( |c| + |a|^2 + \int_{\mathbb{R}^d} (1 + |z|^2) \nu(dz) \right) \left\| \mu \right\|_{L^1(\mathbb{R}^d)}.
\]

This shows that \( \mathcal{L}^\varepsilon : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d) \) and \( \left\| \mathcal{L}^\varepsilon \mu \right\|_{L^1(\mathbb{R}^d)} \leq \left( c\varepsilon^{-1} \right) ||\mu||_{L^1(\mathbb{R}^d)}. \)

\D Part (ii). For every \( \varphi \in C_c^\infty(\mathbb{R}^d) \), by using the Taylor expansion and the Cauchy–Schwarz inequality (for the third-order remainder), we get

\[
\left| \left( \mathcal{L}^\varepsilon_{\text{loc}} - c \cdot \nabla - \text{tr} \left( aa^T D^2(\cdot) \right) \right) \varphi(x) \right| \leq \frac{c}{2} \|D^2 \varphi\|_\infty + 2|a|^2 \|D^3 \varphi\|_\infty.
\]

Let \( \mathcal{L}_{\nu} \varphi(x) = \int_{\mathbb{R}^d} \left( \varphi(x + z) - \varphi(x) - \frac{1}{\varepsilon} \varphi(z) \cdot \nabla \varphi(x) \right) \nu(dz) \). Then

\[
\left( \mathcal{L}^\varepsilon_{\text{nonloc}} - \mathcal{L}_{\nu} \right) \varphi(x) = \left| \int_{B_1 \setminus B_\varepsilon} \left( \varphi(x - \varepsilon z) - \varphi(x) + \varphi(z) \cdot \nabla \varphi(x) \right) \nu(dz) \right|
\]

\[
- \int_{B_\varepsilon} \left( \varphi(x + z) - \varphi(x) - \varphi(z) \cdot \nabla \varphi(x) \right) \nu(dz)
\]

\[
\leq \frac{\varepsilon}{2} ||D^2 \varphi||_\infty \int_{B_1} |z|^2 \nu(dz) + \frac{1}{2} ||D^2 \varphi||_\infty \int_{B_\varepsilon} |z|^2 \nu(dz).
\]

Since \( \lim_{\varepsilon \to 0} \int_{B_\varepsilon} |z|^2 \nu(dz) = 0 \) by the Lebesgue dominated convergence theorem, it follows from (21) and (22) that

\[
\lim_{\varepsilon \to 0} \left( \|\mathcal{L}^\varepsilon - \mathcal{L}\varphi\|_\infty = 0. \right.
\]

\D Part (iii). Let \( V \) be a Lyapunov function such that \( \|\mathcal{L}^\varepsilon V\|_\infty < \infty \). Then also \( \|\mathcal{L}_{\nu} V\|_\infty < \infty \). Because of the definition of \( \mathcal{L}^\varepsilon = \mathcal{L}^\varepsilon_{\text{loc}} + \mathcal{L}^\varepsilon_{\text{nonloc}}, \) in a way similar to (21) and (22), we have

\[
\|\mathcal{L}^\varepsilon V\|_\infty \leq |c| \|\nabla V\|_\infty + |a|^2 \|D^2 V\|_\infty + ||D^2 V||_\infty \int_{B_1} |z|^2 \nu(dz) + ||\mathcal{L}_{\nu} V||_\infty.
\]

Thus \( \sup_{\varepsilon \in (0,1)} \|\mathcal{L}^\varepsilon V\|_\infty < \infty \). Notice that

\[
\int_{B_1} z \nu_1^\varepsilon(dz) = c, \quad \int_{B_1} z \nu_2^\varepsilon(dz) = 0, \quad \text{and} \quad \int_{B_1} (\nu_1^\varepsilon + \nu_2^\varepsilon)(dz) = 0,
\]
thus the Lévy triplet of the operator $\mathcal{L}^c$ is $(c, 0, \nu_c)$ (see Remark 4.12(b)). Hence
\[
\|\mathcal{L}^c\|_{LK} = |c| + \frac{\varepsilon|c|}{2} + |a|^2 + \frac{1}{2} \int_{B_1 \setminus B_r} (1 + \varepsilon)|z|^2 \nu(dz) + 2\nu(B_1^c) \\
\leq (1 + \varepsilon)\|\mathcal{L}\|_{LK}.
\]

Next we prove a result concerning Lévy operators satisfying (LA).

**Proposition 4.18.** Assume (LA) and $\phi \in C_b^p(\mathbb{R}^d)$ for some $p \in (2\sigma, 1]$. Then
\[
\|\mathcal{L}\phi\|_{\infty} \leq \frac{K}{p - 2\sigma}[\phi]_p + 2\|\phi\|_{\infty}\nu(B_1^c)
\]
and
\[
[\mathcal{L}\phi(x)]_{p - 2\sigma} \leq 2\left(\frac{K}{p - 2\sigma} + \nu(B_1^c)\right)[\phi]_p.
\]

Consequently, $\mathcal{L} : C_b^p(\mathbb{R}^d) \to C_b^{p - 2\sigma}(\mathbb{R}^d)$ is a bounded operator.

**Proof.** Estimate (23) is a simple consequence of (LA). To obtain (24), we write
\[
|\mathcal{L}\phi(x) - \mathcal{L}\phi(y)| \leq \int_{B_1} \left|\left(\phi(x + z) - \phi(x)\right) - \left(\phi(y + z) - \phi(y)\right)\right| \nu(dz)
\]
\[
+ \int_{B_1^c} \left|\left(\phi(x + z) - \phi(y + z) - \phi(x) + \phi(y)\right)\right| \nu(dz) = I_1 + I_2.
\]

For $|x - y| \leq 1$ (cf. Definition 4.1, where $y \in B_1(x)$), we get
\[
I_1 \leq 2[\phi]_p \left(\int_{B_{1|x-y|}} |z|^p \nu(dz) + \int_{B_1 \setminus B_{1|x-y|}} |x - y|^p \nu(dz)\right)
\]
\[
= 2[\phi]_p|x - y|^p \int_{B_1} \left(1 \wedge \frac{|z|^p}{|x - y|^p}\right) \nu(dz) \leq \frac{2K}{p - 2\sigma}[\phi]_p|x - y|^{p - 2\sigma}.
\]

Finally,
\[
I_2 \leq 2\nu(B_1^c)[\phi]_p|x - y|^p \leq 2\nu(B_1^c)[\phi]_p|x - y|^{p - 2\sigma}.
\]

**Remark 4.19.** Note that an operator $\mathcal{L}$ satisfying (LC) with $2\sigma \in (0, 1)$ also satisfies (LA). For $\mathcal{L}$ satisfying (LC) with $2\sigma \in [1, 2)$, we have estimates similar to those in Proposition 4.18. Namely, if $\phi \in C_b^p(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ is such that $\partial_x \phi \in C_b^p(\mathbb{R}^d)$ for every $i = 1, \ldots, d$ with $p \in (2\sigma - 1, 1]$, then
\[
\|\mathcal{L}\phi\|_{\infty} \leq \frac{KK_d}{1 - \sigma} \max_i [\partial_x \phi]_p + 2\nu(B_1^c)||\phi||_{\infty},
\]
\[
|\mathcal{L}\phi(x)|_{p + 1 - 2\sigma} \leq \left(\frac{KK_d}{2 - 2\sigma} + \nu(B_1^c)\right) \max_i [\partial_x \phi]_p + 2\nu(B_1^c)||\nabla\phi||_{\infty}.
\]

### 4.4. Viscosity solutions

Suppose that $(t, x, \ell) \mapsto \mathcal{F}(t, x, \ell)$ and $w_0$ are continuous functions, and $\mathcal{F}$ is non-decreasing in $\ell$. For $\mathcal{L}$ satisfying (L) with $a = 0$, consider the following problem
\[
\begin{cases}
\partial_t w = \mathcal{F}(t, x, (\mathcal{L}w)(t, x)), & \text{on } T \times \mathbb{R}^d, \\
w(0) = w_0, & \text{on } \mathbb{R}^d.
\end{cases}
\]

For $0 \leq r < \infty$ and $p \in \mathbb{R}^d$ we introduce linear operators
\[
\mathcal{L'}(\phi, p)(x) = \int_{B_r} \left(\phi(x + z) - \phi(x) - \mathbb{1}_{B_1}(z) z \cdot p\right) \nu(dz),
\]
\[
\mathcal{L}_c(\phi, x) = \int_{B_r} \left(\phi(x + z) - \phi(x) - \mathbb{1}_{B_1}(z) z \cdot \nabla \phi(x)\right) \nu(dz).
\]
\[\text{We assume } a = 0 \text{ for simplicity and in order to use results of [23]; we need to allow for } c \neq 0 \text{ because of Remark 4.12(b).}\]
Notice that the operator $\mathcal{L}^r$ is well-defined on every bounded semicontinuous function, while the operator $\mathcal{L}_r$ is well-defined on every $C^2$ function. If $2\sigma \in (0, 1)$ in (LA) or (LC), then we may omit the $p$ and $\nabla \phi$ terms.

**Definition 4.20.** A bounded upper-semicontinuous function $u^- : \mathcal{T} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a viscosity subsolution of problem (25) if

(i) $u^-(0, x) \leq w_0(x)$ for every $x \in \mathbb{R}^d$;

(ii) for every $r \in (0, 1)$, test function $\phi \in C^2(\mathcal{T} \times \mathbb{R}^d)$, and a maximum point $(t, x)$ of $u^- - \phi$ we have
\[
\partial_t \phi(t, x) - \mathcal{F}(t, x, (c \cdot \nabla \phi + \mathcal{L}^r(u^-, \nabla \phi(t, x))) + \mathcal{L}_r \phi)(t, x) \leq 0.
\]

A bounded lower-semicontinuous function $u^+ : \mathcal{T} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a viscosity supersolution of problem (25) if

(iii) $u^+(0, x) \geq w_0(x)$ for every $x \in \mathbb{R}^d$;

(iv) for every $r \in (0, 1)$, test function $\phi \in C^2(\mathcal{T} \times \mathbb{R}^d)$, and a minimum point $(t, x)$ of $u^+ - \phi$ we have
\[
\partial_t \phi(t, x) - \mathcal{F}(t, x, (c \cdot \nabla \phi + \mathcal{L}^r(u^+, \nabla \phi(t, x))) + \mathcal{L}_r \phi)(t, x) \geq 0.
\]

A function $w \in C_b(\mathcal{T} \times \mathbb{R}^d)$ is a viscosity solution of problem (25) if it is a subsolution and a supersolution simultaneously.

**Remark 4.21.** (a) A bounded classical solution (in the usual sense, cf. Definition 1.3) is a bounded viscosity solution.

(b) Under some restrictions on the Lévy measure, it possible to consider data and viscosity solutions which are unbounded (with controlled growth).

(c) In Definition 4.20 we could consider $u^-, u^+ : \mathcal{T} \times \mathbb{R}^d \rightarrow \mathbb{R}$ and require inequalities (i), (iii) to hold for upper or lower limits as $t \rightarrow 0$, respectively.

(d) In Definition 4.20 (ii), (iv), we may take a test function $\phi \in \mathcal{X}$, where
\[
\mathcal{X} = \left\{ \psi \in C_b(\mathcal{T} \times \mathbb{R}^d) : \mathcal{L}_1 \psi \in C_b(\mathcal{T} \times \mathbb{R}^d), \partial_t \psi \in C_b(\mathcal{T} \times \mathbb{R}^d) \right\}.
\]

See [34, §10.1.2] for first order PDEs and $\phi \in C^1(\mathcal{T} \times \mathbb{R}^d)$, and the proof for (25) and $\phi \in \mathcal{X}$ is a small modification of this.

(e) In this paper, we only consider two specific examples of the function $\mathcal{F}$

(i) $\mathcal{F}(t, x, \ell) = F(\ell) - f(t, x)$;

(ii) $\mathcal{F}(t, x, \ell) = b(t, x) \ell$;

subjected to further conditions on the regularity of $F$, $f$, and $b$.

**Definition 4.22.** The comparison principle holds for problem (25) if for a subsolution $u^-$ and a supersolution $u^+$, condition $u^-(0, x) \leq u^+(0, x)$ for every $x \in \mathbb{R}^d$ implies $u^-(t, x) \leq u^+(t, x)$ for every $(t, x) \in \mathcal{T} \times \mathbb{R}^d$.

The comparison principle entails uniqueness of viscosity solutions. Complemented with suitable subsolutions and supersolutions, it also implies existence of solutions through the Perron method (cf. [30, Section 4]). If we can show that the viscosity solution is sufficiently regular, then it is a classical solution. We need the following result in this direction.

**Lemma 4.23.** Assume (LA) and let $w$ be a viscosity solution of problem (25). If the comparison principle holds for problem (25) and
\[
w \in B(\mathcal{T}, C^{2\sigma + 1}_b(\mathbb{R}^d)), \quad \sigma \in (0, 1 - 2\varepsilon)
\]
then $\partial_t w \in C_b(\mathcal{T} \times \mathbb{R}^d)$ and $w$ is a bounded classical solution of problem (25).
Proof.  

\begin{enumerate}[\circled{Step 1.}]  
\item We show \(Lw \in C_b(T \times \mathbb{R}^d)\). By Definition 4.20, \(w \in C_b(\mathcal{T} \times \mathbb{R}^d)\), and then since \(w \in B(\mathcal{T}, C^{2\alpha+\varepsilon}_b(\mathbb{R}^d))\), by Proposition 4.18 \(Lw \in B(\mathcal{T}, C_b^2(\mathbb{R}^d))\). Let \((t_n,x_n) \to (t_0,x_0)\), and note that for every \((t,x) \in \mathcal{T} \times \mathbb{R}^d\) we have  
\[
|w(t,x+z) - w(t,x)| \leq 2||w||_B \mathbb{1}_{B_1}(z) + \sup_{s \in \mathcal{T}}|w(s)|2\alpha+\varepsilon \mathbb{1}_{B_2}(z).
\]

The function on the right-hand side is \(\nu\)-integrable in \(z\). Then by (LA), the Lebesgue dominated convergence theorem, and the continuity of \(w\), we get  
\[
\lim_{n \to \infty} Lw(t_n, x_n) = \lim_{n \to \infty} \int_{\mathbb{R}^d} w(t_n, x_n + z) - w(t_n, x_n) \nu(dz) = \int_{\mathbb{R}^d} w(t_0, x_0 + z) - w(t_0, x_0) \nu(dz) = Lw(t_0, x_0).
\]

\item We show \(w\) is a.e. \(t\)-differentiable. Let \(t_0 \in \mathcal{T}\) be fixed and define  
\[
u^+(t,x) = w(t_0,x) + L(t-t_0), \quad L = \|F(t,x,Lw(t,x))\|_{\infty}.
\]

Then \(\nu^+\) and \(\nu^-\) are respectively a viscosity supersolution and a subsolution of problem (25) for \(t \geq t_0\). Therefore, by the comparison principle,  
\[
\nu^-(t,x) \leq w(t,x) \leq \nu^+(t,x), \quad \text{for every } (t,x) \in [t_0,T] \times \mathbb{R}^d.
\]

Hence \(|w(t_0,x) - w(t,x)| \leq L|t-t_0|\), and \(w\) is \(t\)-Lipschitz. Thus, by the theorems of Rademacher [34, §5.8 Theorem 6] and Fubini [10, Theorem 7.6.5], we find that \(w\) is a.e. \(t\)-differentiable in \(\mathcal{T} \times \mathbb{R}^d\).

\item We show \(\partial_t w \in C_b(\mathcal{T} \times \mathbb{R}^d)\). If \(w\) is \(t\)-differentiable at \((t_0,x_0) \in \mathcal{T} \times \mathbb{R}^d\), then there exists a function \(\phi \in X\) (see Remark 4.21(d)) such that \(w - \phi\) has a strict local maximum at \((t_0,x_0)\). Indeed, we can take  
\[
\phi(t,x) = |x - x_0|^{2\alpha+\varepsilon} \left(1 + \sup_{s \in \mathcal{T}}|w(s)|2\alpha+\varepsilon\right) + v(t),
\]

where \(v \in C^1(\mathcal{T})\) is the function we get from [34, §10.1.2, Lemma] (only for \(t\) variable at fixed \(x_0\)) such that \(v(t_0) = w(t_0,x_0)\) and \(v(t,x_0) - v(t)\) attains a maximum at \(t_0\). Hence, since \(\partial_t (w - \phi) = 0\) and \(\mathcal{L}(w - \phi) \leq 0\) at \((t_0,x_0)\), using the definition of a viscosity subsolution (see Remark 4.21(d) again),  
\[
\partial_t w \leq \partial_t \phi \leq F(t,x,L\phi) \leq F(t,x,Lw) \quad \text{at } (t,x) = (t_0,x_0).
\]

A similar argument for supersolutions gives the opposite inequality. Then by a.e. \(t\)-differentiability we get  
\[
\partial_t w(t,x) = F(t,x,Lw(t,x)) \quad \text{a.e. in } \mathcal{T} \times \mathbb{R}^d.
\]

Since \((t,x) \mapsto F(t,x,Lw(t,x)) \in C_b(\mathcal{T} \times \mathbb{R}^d)\), we obtain \(\partial_t w \in C_b(\mathcal{T} \times \mathbb{R}^d)\) and so \(w\) is a bounded classical solution of problem (25). \qed
\end{enumerate}

5. Hamilton–Jacobi–Bellman equation

In this section we discuss problem (2). First, we state a comparison principle for (uniformly continuous) viscosity solutions.

Theorem 5.1. [25, Theorems 6.1, 6.2] Assume (L) with \(a = 0\), (A1), and data \((f,g)\) to be bounded uniformly continuous.

\begin{enumerate}[(i)]
\item The comparison principle (see Definition 4.22) holds for problem (2).
\item There exists a unique viscosity solution of problem (2).
\end{enumerate}

We may also formulate the result in a more useful form.
Lemma 5.2. Assume (L) with $a = 0$ or $\nu = 0$, (A1) and let $u_1, u_2$ be viscosity solutions of problem (2) with bounded uniformly continuous data $(f_1, g_1), (f_2, g_2)$, respectively. Then for every $t \in \mathcal{T}$,
\[
\|u_1(t) - u_2(t)\|_{\infty} \leq (T - t)\|f_1 - f_2\|_{\infty} + \|g_1 - g_2\|_{\infty}.
\]

Proof. For $a = 0$, by the definition of viscosity solutions, we find for $\{i, j\} = \{1, 2\}$, that the function $v_i$ given by
\[
v_i(t, x) = u_j(t, x) - (T - t)\|f_1 - f_2\|_{\infty} - \|g_1 - g_2\|_{\infty}
\]
is a viscosity subsolution of problem (2) with data $(f_i, g_i)$. By Theorem 5.1 (i), we thus obtain $v_1(t, x) \leq u_1(t, x)$ and $v_2(t, x) \leq u_2(t, x)$ for every $(t, x) \in \mathcal{T} \times \mathbb{R}^d$. By combining the inequalities, the result follows.

For $\nu = 0$, see [30] or e.g. [48, Theorem 3.2 (a)]. \(\square\)

For an arbitrary $\mathcal{L}$ satisfying (L), Lemma 5.2 can be proved by combining the arguments of [25] and [49]. However, a full proof would be tedious and such a general statement is not needed in this paper. Since we are only going to use viscosity solutions in detail when (LA) holds, having $a = 0$ is enough. Then, under additional assumptions, we prove existence of bounded classical solutions (see Definition 1.3).

On the other hand, in the proof of Theorem 8.5, where the existence of bounded classical solutions of problem (2) is assumed, we still need a comparison principle. In this setting, it can be obtained by a straightforward application of the maximum principle.

Theorem 5.3. Assume (L), (A1) and let $u_1, u_2$ be bounded classical solutions of problem (2) with continuous data $(f_1, g_1), (f_2, g_2)$, respectively. Suppose there is $K \geq 0$ such that $\|F'(\mathcal{L}u_1)\|_{\infty}, \|F'(\mathcal{L}u_2)\|_{\infty} \leq K$. Then for every $t \in \mathcal{T}$,
\[
\|u_1(t) - u_2(t)\|_{\infty} \leq (T - t)\|f_1 - f_2\|_{\infty} + \|g_1 - g_2\|_{\infty}.
\]

The proof is given in Appendix C.

5.1. The degenerate case.

Theorem 5.4. Assume (LA), (A1) and $(f, g) \in \mathcal{R}_A(\alpha, M)$ (as in (RA)). If $u$ is a viscosity solution of problem (2), then
(i) $u \in B(\mathcal{T}, C^0_b(\mathbb{R}^d))$ with $\max\{\|u(t)\|_{\infty}, [u(t)]_\alpha\} \leq M(T - t + 1)$, and $\mathcal{L}u \in B(\mathcal{T}, C^{\alpha-2\sigma}_b(\mathbb{R}^d))$ with
\[
\|\mathcal{L}u(t)\|_{\alpha-2\sigma} \leq 4\left(\frac{K}{\alpha-2\sigma} + \nu(B_1^\alpha)\right)M(T - t + 1);
\]
(ii) $u$ is a bounded classical solution of problem (2) and $\partial_t u, \mathcal{L}u \in C_0(\mathcal{T} \times \mathbb{R}^d)$;
(iii) if $u_n$ are viscosity solutions of problem (2), with data $(f_n, g_n) \in \mathcal{R}_A(\alpha, M)$ and we have $\lim_{n \to \infty} \|u_n - u\|_{\infty} = 0$, then $\mathcal{L}u_n(t) \to \mathcal{L}u(t)$ uniformly on compact sets in $\mathbb{R}^d$ for every $t \in \mathcal{T}$.

Proof. Part (i). Let $y \in \mathbb{R}^d$ and define $\tilde{u}(t, x) = u(t, x + y)$. Notice that $\tilde{u}$ is a viscosity solution of problem (2) with data $(f, \tilde{g})$, where $f(x) = f(x + y)$ and $\tilde{g}(x) = g(x + y)$. Hence by Lemma 5.2 for every $(t, x, y) \in \mathcal{T} \times \mathbb{R}^d \times \mathbb{R}^d$ we get
\[
|u(t, x + y) - u(t, x)| \leq (T - t)\|f - \tilde{f}\|_{\infty} + \|g - \tilde{g}\|_{\infty}
\]
Since \( ||u(t)||_\infty \leq (T-t)||f||_\infty + ||g||_\infty \leq M(T-t+1) \) by Lemma 5.2 again, it then follows that \( u \in \mathcal{B}(T, C^0_b(\mathbb{R}^d)) \). Then by (LA) and Proposition 4.18 we find that

\[
||Lu(t)||_{\alpha-2\sigma} \leq 4 \left( \frac{K}{n^{2\sigma}} + \nu(B^n_t) \right) M(T-t+1).
\]

\( \circ \) Part (ii). It follows from Part (i) and Lemma 4.23 that \( u \) is a bounded classical solution and \( \partial_t u, Lu \in C_b(T \times \mathbb{R}^d). \)

\( \circ \) Part (iii). By Part (i) and the Arzelà–Ascoli theorem, for every \( t \in T \) there exist a subsequence \( \{u_{n_k}\} \) and a function \( v \in C_b(\mathbb{R}^d) \), such that \( \mathcal{L}u_{n_k}(t) \to v \) uniformly on compact sets in \( \mathbb{R}^d \). On the other hand, by Part (i) and the Lebesgue dominated convergence theorem, \( \lim_{n \to \infty} \mathcal{L}u_{n_k}(t, x) = \mathcal{L}u(t, x) \) for every \( (t, x) \in T \times \mathbb{R}^d \). Hence we find \( \mathcal{L}u_{n_k}(t) \to \mathcal{L}u(t) \) uniformly on compact sets in \( \mathbb{R}^d \) for every \( t \in T \).

\[\square\]

**Corollary 5.5.** Assume (RA), (LA), and (A1). Then conditions (S1), (S2), (S3), (S4) are satisfied.

**Proof.** Conditions (S1) and (S4) are a consequence of Theorem 5.4(ii), while (S3) follows from Theorem 5.4(i) and (A1). We obtain (S2) from Theorem 5.4(iii). \( \square \)

### 5.2. The non-degenerate case

We start with the following auxiliary result.

**Proposition 5.6.** Assume (L) with \( \alpha = 0 \) or \( \nu = 0 \), (A1) and \( u \) is a viscosity solution of problem (2) with bounded uniformly continuous data \( (f, g) \) such that \( \partial_t f \in L^\infty(T \times \mathbb{R}^d) \) and \( \mathcal{L}g \in L^\infty(\mathbb{R}^d) \). Then \( \partial_t u \in L^\infty(T \times \mathbb{R}^d) \) and

\[
||\partial_t u(t)||_\infty \leq (T-t)||\partial_t f||_\infty + ||F(\mathcal{L}g)||_\infty + ||f||_\infty.
\]

**Proof.** Take \( h > 0 \) and \( g_{\ast} = g \ast \rho_{\varepsilon} \), where \( \rho_{\varepsilon} \) is the standard mollifier. Note that \( v_{\varepsilon}(t, x) = g_{\varepsilon}(x) \) is a viscosity (classical) solution of problem (2) with data \( (-F(\mathcal{L}g_{\varepsilon}), g_{\varepsilon}) \), hence by Lemma 5.2,

\[
||u(T-h) - g||_\infty \leq h||F(\mathcal{L}g_{\varepsilon}) + f||_\infty + 2||g_{\varepsilon} - g||_\infty.
\]

By (A1), \( ||F(\mathcal{L}g_{\varepsilon})||_\infty \leq ||F(\mathcal{L}g)||_\infty \), and because \( g \in BUC(\mathbb{R}^d), ||g_{\varepsilon} - g||_\infty \) can be arbitrarily small. Thus,

\[
||u(T-h) - u(T)||_\infty \leq h||F(\mathcal{L}g)||_\infty + ||f||_\infty.
\]

Similarly, \( v_{h}(t, x) = u(t-h, x) \) is a viscosity solution of problem (2) with data \( (f(-h) \cdot h, u(T-h)) \), thus for every \( t \in T \),

\[
||u(t) - v_{h}(t)||_\infty \leq (T-t)||f(-h)\cdot - h ||_\infty + ||u(T-h) - u(T)||_\infty \leq (T-t)||\partial_t f||_\infty h + ||F(\mathcal{L}g) + f||_\infty h.
\]

Hence \( u \) is Lipschitz in time. \( \square \)

**Theorem 5.7.** Assume (FB), (A1), (A2), \( (f, g) \in \mathcal{R}_B(\alpha, M) \) (as in (RB)), and (LB) or (LC). If interior \( (\alpha/2\sigma, \alpha) \)-regularity estimates hold for problem (2) (see Definition 2.3), then

(i) there exists a bounded classical solution \( u \) of problem (2);

(ii) if \( u_n \) are bounded classical solutions of problem (2) with data \( (f_n, g_n) \in \mathcal{R}_B(\alpha, M) \) and \( \lim_{n \to \infty} ||u_n - u||_\infty = 0 \), then \( \mathcal{L}u_n(t) \to \mathcal{L}u(t) \) uniformly on compact sets in \( \mathbb{R}^d \) for every \( t \in T \);

(iii) \( \partial_t u, Lu \in C_b(T \times \mathbb{R}^d) \) and for every \( t \in T \) there is a constant \( C(t, f, g) \) such that

\[
||L u||_{C^{\alpha/2\sigma, \alpha}(0, t) \times \mathbb{R}^d} \leq C(t, f, g).
\]
Theorem 5.1. Because of the interior regularity estimates, we have \( \partial_t u, L u \in C(T \times \mathbb{R}^d) \), hence \( u \) is a bounded classical solution of problem (2).

\( \diamond \) Part (ii). By Part (i) and interior regularity estimates, for every \( t \in T \) and \( r > 0 \), there exists a constant \( C(t, r) > 0 \) such that
\[
\sup_n \left( \| \mathcal{L} u_n(t) \|_{L^\infty(B_r)} + \| \mathcal{L} u_n(t) \|_{C^\alpha(B_{r/2})} \right) \leq C(t, r).
\]

By the Arzelà–Ascoli theorem, for every \( t \in T \) there exist a subsequence \( \{u_{n_k}\} \) and a function \( v \in C_b(\mathbb{R}^d) \) such that \( \mathcal{L} u_{n_k}(t) \to v \) uniformly on compact sets in \( \mathbb{R}^d \). For \( \varphi \in C_c^\infty(\mathbb{R}^d) \), we note that
\[
\lim_{k \to \infty} \int_{\mathbb{R}^d} \mathcal{L} u_{n_k}(t, x) \varphi(x) \, dx = \int_{\mathbb{R}^d} v(x) \varphi(x) \, dx,
\]
and since \( \lim_{n \to \infty} \| u_n - u \|_\infty = 0 \) and \( \mathcal{L}^* \varphi \in L^1(\mathbb{R}^d) \),
\[
\lim_{k \to \infty} \int_{\mathbb{R}^d} \mathcal{L} u_{n_k}(t, x) \varphi(x) \, dx = \lim_{k \to \infty} \int_{\mathbb{R}^d} u_{n_k}(t, x) \mathcal{L}^* \varphi(x) \, dx = \int_{\mathbb{R}^d} u(t, x) \mathcal{L}^* \varphi(x) \, dx.
\]
Hence \( v(x) = \mathcal{L} u(t, x) \), and \( \mathcal{L} u_{n_k}(t) \to \mathcal{L} u(t) \) uniformly on compact sets in \( \mathbb{R}^d \) for every \( t \in T \).

\( \diamond \) Part (iii). By Part (i) and Proposition 5.6, \( \partial_t u \in C_b(T \times \mathbb{R}^d) \). Since \( u \) is a bounded classical solution and \( F' \geq \kappa \), we also have \( \mathcal{L} u = F^{-1}(-\partial_t u - f) \in C_b(T \times \mathbb{R}^d) \). Moreover,
\[
\| \mathcal{L} u \|_\infty \leq F^{-1}\left(T\|\partial_t f\|_\infty + \|F(\mathcal{L} g)\|_\infty + 2\|f\|_\infty\right).
\]

By Theorem 5.3, we have \( \|u\|_\infty \leq T\|f\|_\infty + \|g\|_\infty \). Thus, by interior regularity estimates (which are uniform in \( x \), see Definition 2.3), for every \( t \in T \),
\[
\| \mathcal{L} u \|_{C^{\alpha/2, \alpha}(\{0, t\} \times \mathbb{R}^d)} \leq \| \mathcal{L} u \|_\infty + \sup_{x \in \mathbb{R}^d} \left( \mathcal{L} u \right)_{C^{\alpha/2, \alpha}(\{0, t\} \times B_1(x))} \leq \tilde{C}(t)\left(\|f\|_{\alpha/2, \alpha} + \|\partial_t f\|_\infty + \|\mathcal{L} g\|_\infty + \|g\|_\infty\right). \]

Corollary 5.8. Assume (RB), (FB), (A1), (A2), and (LB) or (LC). If interior \((\alpha/2, \alpha)\)-regularity estimates hold for problem (2), then conditions (S1), (S2), (S3), (S4) are satisfied.

Proof. Condition (S1) follows from Theorem 5.7(i), while (S2) follows from Theorem 5.7(ii), and (S3), (S4) hold by Theorem 5.7(iii).

\( \square \)

Remark 5.9. If instead of (RB) we only assume \( \mathcal{R} \subset C^{\alpha/2, \alpha}_b(\mathbb{T} \times \mathbb{R}^d) \times BUC(\mathbb{R}^d) \) (uniformly bounded in an appropriate way) in Corollary 5.8, then we still obtain (S1) and (S2). We may get (S3) by assuming \( F' \leq K \) (i.e. \( F \) is globally Lipschitz). This is enough for the existence result of Theorem 8.5, but not enough for uniqueness in Theorem 8.7.

6. Fokker–Planck equation — Existence

In this section we prove existence of solutions of problem (3). We consider the following assumption
\[(A1'): b \in C(T \times \mathbb{R}^d) \text{ and } 0 \leq b(t, x) \leq B < \infty \text{ for a constant } B \text{ and every } (t, x) \in T \times \mathbb{R}^d, \]
thus obtain \[ \bigcup \text{Lemma 6.1.} \]

**Lemma 6.1.** Let \( m \in C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d)) \) and \( m(0) = m_0 \). The following are equivalent

(i) \( m \) is a very weak solution of problem (3) (cf. Definition 1.4);

(ii) \( m \) satisfies (4) for every

\[ \phi \in \mathcal{U} = \{ \phi \in C_b(\mathcal{T} \times \mathbb{R}^d) : \partial_t \phi + bL\phi \in C_b(\mathcal{T} \times \mathbb{R}^d) \} \]

(iii) \( m \) satisfies (4) for every \( \lambda \in C^\infty([0, \infty) \times \mathbb{R}^d) : \phi(t) \in C^\infty(\mathbb{R}^d) \) for every \( t \in \mathcal{T} \).

**Proof.** Implications (ii) \( \Rightarrow \) (i) \( \Rightarrow \) (iii) are trivial. By a density argument we get (i) \( \Rightarrow \) (ii). To prove (iii) \( \Rightarrow \) (i), fix \( \varphi \in C^\infty_c(\mathcal{T} \times \mathbb{R}^d) \), \( t \in \mathcal{T} \), and consider a sequence of simple functions \( \varphi^k = \sum_{n=1}^{N_k} \mathbb{1}_{[t_n, t_{n+1})} \varphi(t_n) \to \varphi \) pointwise, where \( \bigcup_{n} [t_n, t_{n+1}) = (0, t) \) for each \( n \in \mathbb{N} \) and \( t_k < t_{k+1} \). Then by (iii) we have

\[ \sum_{n=1}^{N_k} (m(t^k_{n+1}) - m(t^k_n))[\varphi(t^k_n)] = \sum_{n=1}^{N_k} \int_{t^k_n}^{t^k_{n+1}} m(\tau)[b(\tau) L \varphi(t^k_n)] d\tau. \]

Notice that by the Lebesgue dominated convergence theorem we get

\[ \lim_{k \to \infty} \sum_{n=1}^{N_k} \int_{t^k_n}^{t^k_{n+1}} m(\tau)[b(\tau) L \varphi(t^k_n)] d\tau = \lim_{k \to \infty} \int_0^t m(\tau)[b(\tau) L \varphi(\tau)] d\tau = \int_0^t m(\tau)[b(\tau) L \varphi(\tau)] d\tau. \]

We also observe that

\[ \sum_{n=1}^{N_k} (m(t^k_{n+1}) - m(t^k_n))[\varphi(t^k_n)] \]

\[ = m(t)[\varphi(t)] - m_0[\varphi(0)] - \sum_{n=1}^{N_k} \left( m(t^k_{n+1})[\varphi(t^k_{n+1}) - \varphi(t^k_n)] \right). \]

By the Taylor expansion, for some \( \xi^k \in [t^k_n, t^k_{n+1}] \) we have

\[ \varphi(t^k_{n+1}) - \varphi(t^k_n) = \partial_t \varphi(t^k_{n+1})(t^k_{n+1} - t^k_n) - \partial_t^2 \varphi(\xi^k)(t^k_{n+1} - t^k_n)^2 / 2. \]

Since \( m \in C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d)) \), by considering the relevant Riemann integral on \([0, t]\), we thus obtain

\[ \lim_{k \to \infty} \sum_{n=1}^{N_k} \left( m(t^k_{n+1})[\varphi(t^k_{n+1}) - \varphi(t^k_n)] \right) = \int_0^t m(\tau)[\partial_t \varphi(\tau)] d\tau. \]

By combining these arguments we get

\[ m(t)[\varphi(t)] = m_0[\varphi(0)] + \int_0^t m(\tau)[\partial_t \varphi(\tau) + b(\tau)(L \varphi(\tau))] d\tau. \]

**Lemma 6.2.** Assume that triplets \((\mathcal{L}_\lambda, b_\lambda, m_{0, \lambda})_\lambda\) satisfy (L), (A1'), (A3) for each \( \lambda \), and let \( \mathcal{M}_\lambda \) be the sets of very weak solutions of problems

\[ \begin{align*}
\partial_t m_\lambda &= \mathcal{L}^*_\lambda(b_\lambda m_\lambda) \quad \text{on } \mathcal{T} \times \mathbb{R}^d, \\
m_\lambda(0) &= m_{0, \lambda} \quad \text{on } \mathbb{R}^d.
\end{align*} \]

\[ ^{14} \text{In this set functions are constant in time.} \]
If \( \bigcup \{ m_{\alpha, \lambda} \parallel_{B_r^1} \} \) is tight and \( \sup_\lambda (\| b_\lambda \|_\infty + \| L_\lambda \|_{L^1}) < \infty \), then

(i) for every \( \varepsilon > 0 \) there exists a compact set \( K_\varepsilon \subset \mathbb{R}^d \) such that

\[
\sup \left\{ \sup_{t \in \mathcal{T}} m(t)(K_\varepsilon^c) : m \in \bigcup_\lambda \mathcal{M}_\lambda \right\} \leq \varepsilon;
\]

(ii) for every \( m \in \bigcup_\lambda \mathcal{M}_\lambda \) we have

\[
\| m(t) - m(s) \|_0 \leq \sup_\lambda \left( 2 + (2\sqrt{T} + K_d) \| b_\lambda \|_\infty \| L_\lambda \|_{L^1} \right) \sqrt{|t - s|};
\]

(iii) the set \( \bigcup_\lambda \mathcal{M}_\lambda \subset C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d)) \) is pre-compact.

**Proof.** Part (i). Let \( V(x) = V_0(\sqrt{1 + |x|^2}) \) be a Lyapunov function for which we have \( \sup_\lambda (m_{0, \lambda}[V] + \| L_\lambda V \|_\infty) < \infty \) (see Lemma 4.9, Lemma 4.13, Corollary 4.14). For \( n \in \mathbb{N} \), let \( V_{n,0} \in C_0^1([0, \infty)) \) be such that

\[
V_{n,0}(t) = \begin{cases} V_0(t) & \text{for } t \leq n, \\ V_0(\sqrt{1 + (n+1)^2}) & \text{for } t \geq n + 2,
\end{cases}
\]

and additionally

\[
0 \leq V_{n,0}' \leq V_0' \quad \text{and} \quad |V_{n,0}''| \leq |V_0''|.
\]

Take \( V_n(x) = V_{n,0}(\sqrt{1 + |x|^2}) \). Thanks to Lemma 6.1, for every \( m \in \mathcal{M}_\lambda \),

\[
m(t)[V_n] = m_{0, \lambda}[V_n] + \int_0^t m(\tau)|b_\lambda(\tau) L_\lambda V_n| \, d\tau.
\]

Notice that \( |V_n(x) - V_n(y)| \leq |V(x) - V(y)| \) and

\[
\lim_{n \to \infty} (V_n, \nabla V_n, D^2 V_n)(x) = (V, \nabla V, D^2 V)(x) \quad \text{for every } x \in \mathbb{R}^d.
\]

We now use the formula in (L) with \( \phi = V_n \) and separate the integral part on domains \( B_1 \) and \( B_1^c \). Because of (29), by the Lebesgue dominated convergence theorem — we use Lemma 4.13 (iv) for the integral on \( B_1^c \) and (27) otherwise — we may pass to the limit in (28). For every \( t \in \mathcal{T}, \lambda \), and \( m \in \mathcal{M}_\lambda \) we obtain

\[
m(t)[V] = m_{0, \lambda}[V] + \int_0^t m(\tau)|b_\lambda L_\lambda V| \, d\tau \leq m_{0, \lambda}[V] + \| b_\lambda \|_\infty \| L_\lambda V \|_\infty T.
\]

Thus, because of Proposition 4.8, for every \( \varepsilon > 0 \) there exists a compact set \( K_\varepsilon \) such that

\[
\sup \left\{ m(t)(K_\varepsilon^c) : t \in \mathcal{T}, m \in \bigcup_\lambda \mathcal{M}_\lambda \right\} \leq \varepsilon.
\]

Part (ii). Consider \( \phi_\varepsilon = \phi * \rho_\varepsilon \), where \( \phi \in C_0^1(\mathbb{R}^d) \) is such that \( \| \phi \|_\infty \leq 1 \) and \( [\phi]_1 \leq 1 \), and \( \rho_\varepsilon \) is a standard mollifier. Notice that \( \| \phi - \phi_\varepsilon \|_\infty \leq \varepsilon \) and, by Proposition 4.15, \( \| L_\phi \|_\infty \leq \| L \|_{L^1} \| \phi_\varepsilon \|_{C^2(\mathbb{R}^d)} \). Because of Definition 1.4, for every \( \lambda \) and \( m \in \mathcal{M}_\lambda \) we obtain

\[
|m(t) - m(s)|[\phi] = |(m(t) - m(s))[\phi - \phi_\varepsilon] + (m(t) - m(s))[\phi_\varepsilon]| \\
\leq 2\varepsilon + \int_s^t \int_{\mathbb{R}^d} (L_\lambda \phi_\varepsilon)(x)b_\lambda(t, x) m(t, dx) \, d\tau \\
\leq 2\varepsilon + \| b_\lambda \|_\infty \| L_\lambda \|_{L^1} \| \phi_\varepsilon \|_{C^2(\mathbb{R}^d)} |t - s|.
\]

We also have

\[
\| \phi_\varepsilon \|_{C^2(\mathbb{R}^d)} \leq \left( \| \phi \|_\infty + \| \nabla \phi \|_\infty + \frac{K_d \| \nabla \phi \|_\infty}{\varepsilon} \right) \leq \frac{2\varepsilon + K_d}{\varepsilon}.
\]

\(^{15}\)See (19) for the definition of \( \| \cdot \|_{L^1} \).
By taking $\varepsilon = \sqrt{|t - s|}$, we thus obtain
\[ |m(t) - m(s)|_0 \leq \sup_\lambda \left( 2 + (2\sqrt{T} + K_d) \| b_\lambda \|_\infty \| L_\lambda \|_{L_K} \right) \sqrt{|t - s|}. \]

\[ \diamond \text{Part (iii). It follows from Part (i) that the set } \{ m(t) : m \in \bigcup_\lambda \mathcal{M}_\lambda \} \text{ is equicontinuous for a fixed } t \in \mathcal{T}. \text{ Then, in Part (ii), we showed that the family } \bigcup_\lambda \mathcal{M}_\lambda \text{ is equicon-} \]
\[ \text{tuous in } C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d)). \text{ Hence } \bigcup_\lambda \mathcal{M}_\lambda \subset C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d)) \text{ is pre-compact by the Arzelà-Ascoli theorem } [52, \S 7 \text{ Theorem 17}]. \]
\[ \Box \]

In the general case we are unable to prove uniqueness of solutions of problem (3). However, we can make the following observation about the sets of solutions.

**Corollary 6.3.** Assume (L), (A1'), (A3). If $\mathcal{M} \subset C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$ is the set of solutions of problem (3) corresponding to $(b, m_0)$, then $\mathcal{M}$ is convex, compact, and
\[ \sup_{m \in \mathcal{M}} \sup_{t \in \mathcal{T}} |m(t)|_V \leq c_1, \quad \sup_{m \in \mathcal{M}} \sup_{0 \subset [t-s] \subset \mathcal{T}} \frac{|m(t) - m(s)|_0}{\sqrt{|t - s|}} \leq c_2, \]
for a Lyapunov function $V$ such that $m_0[V], \| L_\infty \|_\infty < \infty$ (see Corollary 4.14), and
\[ c_1 = m_0[V] + T\| b\|_\infty \| L_\infty \|_\infty, \quad c_2 = 2 + (2\sqrt{T} + K_d) \| b\|_\infty \| L_\infty \|_{L_K}. \]

**Proof.** It follows from Definition 1.4 that $\mathcal{M}$ is convex (the equation is linear), as well as that if $\{ m_n \} \subset \mathcal{M}$ and $m_n \to \bar{m}$ in $C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$, then $\bar{m} \in \mathcal{M}$, i.e. the set $\mathcal{M}$ is closed. Hence, by Lemma 6.2 (iii), we obtain that $\mathcal{M} \subset C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$ is compact. The specified bounds follow from Lemma 6.2 (ii) and (30). \[ \Box \]

We now prove a kind of a stability result for solutions (in terms of semicontinuity with respect to upper Kuratowski limits (see [59, §29.III]).

**Lemma 6.4.** Assume (L), (A3), and $\{ a_n, b \} \subset \mathbb{N}$ satisfy (A1'), being uniformly bounded by $B$. Let $\{ \mathcal{M}_n, \mathcal{M}_d \}$ be the corresponding sets of solutions of problem (3) with $m_0$ as initial conditions. If $m_n \in \mathcal{M}_n$ for every $n \in \mathbb{N}$ and $b_n(t) \to b(t)$ uniformly on compact sets in $\mathbb{R}^d$ for every $t \in \mathcal{T}$, then there exists a subsequence $\{ m_{n_k} \}$ and $m \in \mathcal{M}$ such that $m_{n_k} \to m$ in $C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$.

**Proof.** By Lemma 6.2 (iii) the set $\bigcup_\lambda \mathcal{M}_\lambda \subset C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$ is pre-compact, and by Lemma 6.2 (i) for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset \mathbb{R}^d$ such that
\[ \sup_{n \in \mathbb{N}} \sup_{m \in \mathcal{M}_n} \sup_{t \in \mathcal{T}} |m(t)|(K_\varepsilon) \leq \varepsilon. \]

Let $\{ m_{n_k} \} \subset \{ m_n \}$ be a convergente subsequence and $m = \lim_{k \to \infty} m_{n_k}$. Without loss of generality, we may still denote $m_{n_k}$ as $m_n$. For every $\varphi \in C_{\infty}(\mathbb{R}^d)$ we have
\[ \left| \int_0^t (b_n - b)m_n - bm)(\tau)[L\varphi] \, d\tau \right| = \left| \int_0^t \left( (b_n - b)m_n + b(m_n - m) \right)(\tau)[L\varphi] \, d\tau \right|. \]

Since $m_n \to m$ in $C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d))$ and $b \in C_b(\mathcal{T} \times \mathbb{R}^d)$, we notice that
\[ \lim_{n \to \infty} \sup_{\tau \in \mathcal{T}} |m_n(\tau)[b(\tau) - m(\tau)[b(\tau)]| = 0. \]

Next,
\[ \left| \int_0^t (b_n - b)m_n(\tau)[L\varphi] \, d\tau \right| \leq \| L\varphi \|_{\infty} \int_0^t \int_{K_\varepsilon(K_\varepsilon)} |b_n - b(\tau, x)| m_n(\tau, dx) \, d\tau \]
\[ \leq \| L\varphi \|_{\infty} \left( \varepsilon T (\| b\|_\infty + \| b\|_\infty) + \int_0^T \sup_{x \in K_\varepsilon} |b_n(\tau, x) - b(\tau, x)| \, d\tau \right). \]
Thus we observe that for every \( (t, x) \in \mathcal{T} \times \mathbb{R}^d \) and \( b_n(t) \to b(t) \) uniformly on compact sets in \( \mathbb{R}^d \) for every \( t \in \mathcal{T} \), hence

\[
\sup_{x \in K} |b_n(t, x) - b(t, x)| \to 0 \quad \text{pointwise in } t \in \mathcal{T}.
\]

Thus, by Lebesgue dominated convergence theorem,

\[
\sup_{t \in \mathcal{T}} \lim_{n \to \infty} \left| \int_0^t (b_n m_n - bm)(\tau)[\mathcal{L} \varphi] \, d\tau \right| \leq 2 \varepsilon BT \|\mathcal{L} \varphi\|_{\infty}.
\]

Since \( \varepsilon > 0 \) may be arbitrarily small and \( m_n \) are solutions of problem (3), because of Lemma 6.1 (iii),

\[
m(t)[\varphi] - m_0[\varphi] = \lim_{n \to \infty} m_n(t)[\varphi] - m_0[\varphi] = \lim_{n \to \infty} \int_0^t b_n m_n(\tau)[\mathcal{L} \varphi] \, d\tau = \int_0^t bm(\tau)[\mathcal{L} \varphi] \, d\tau.
\]

Thus \( m \) is a solution of problem (3) with parameters \( b \) and \( m_0 \), i.e. \( m \in \mathcal{M} \).

Remark 6.5. When the solutions of problem (3) are unique for each pair \((b, m_0)\), Lemma 6.4 becomes a more standard stability result. Indeed, let \( \{m_n, m\} \) be (the unique) solutions of problem (3) with a fixed initial condition \( m_0 \) and parameters \( \{b_n, b\} \) such that \( b_n \to b \) uniformly on compact sets in \( \mathbb{R}^d \) for every \( t \in \mathcal{T} \). By Lemma 6.4 every subsequence of \( \{m_n\} \) contains a further subsequence convergent to \( m \). Thus \( m_n \to m \) in \( C(\mathcal{T}, \mathcal{P}(\mathbb{R}^d)) \).

Next we show that the set of solutions is non-empty.

**Theorem 6.6.** Assume (L), (A1’), (A3). Problem (3) has a very weak solution.

**Proof.**

\( \diamond \) **Step 1. Approximate problem.** For \( \varepsilon \in (0, 1) \), let \( \mathcal{L}^\varepsilon \) be the sequence of approximations of operator \( \mathcal{L} \) given by Lemma 4.17 and \( \nu^\varepsilon \), \( \mathcal{L}^\varepsilon \ast \) be their Lévy measures and adjoint operators, respectively.

By (20) and the Fubini theorem, for every \( \mu \in L^1(\mathbb{R}^d) \) we have

\[
\int_{\mathbb{R}^d} \mathcal{L}^\varepsilon \ast \mu \, dx = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (\mu(x) - \mu(z)) \, dx \, \nu^\varepsilon(dz) = 0.
\]

Let \( b_\varepsilon = b + \varepsilon \) and \( \mu_0, \varepsilon = m_0 \ast \rho_\varepsilon \), where \( \{\rho_\varepsilon\}_{\varepsilon \in (0, 1)} \) is the sequence of standard mollifiers. We consider the following family of problems

\[
\left\{ \begin{array}{ll}
\partial_t \mu = \mathcal{L}^\varepsilon \ast (b_\varepsilon \mu) & \text{on } \mathcal{T} \times \mathbb{R}^d, \\
\mu(0) = \mu_0, \varepsilon & \text{on } \mathbb{R}^d,
\end{array} \right.
\]

for every \( \varepsilon \in (0, 1) \).

\( \diamond \) **Step 2. Existence of approximate solution \( \mu_\varepsilon \).** For \( \mu \in C(\mathcal{T}, L^1(\mathbb{R}^d)) \), define

\[
\mathcal{G}_\varepsilon(\mu)(t) = \mu(0) + \int_0^t \mathcal{L}^\varepsilon \ast (b_\varepsilon \mu)(\tau) \, d\tau.
\]

We observe that for every \( t_0 \in \mathcal{T} \), because \( \|b_\varepsilon\|_{\infty} < \|b\|_{\infty} + 1 \),

\[
\mathcal{G}_\varepsilon : C([0, t_0], L^1(\mathbb{R}^d)) \to C([0, t_0], L^1(\mathbb{R}^d)) \cap C^1((0, t_0], L^1(\mathbb{R}^d))
\]

is a bounded linear operator.

Let \( \mu_1, \mu_2 \in C(\mathcal{T}, L^1(\mathbb{R}^d)) \) be such that \( \mu_1(0) = \mu_2(0) \) and take

\[
t_\varepsilon = \frac{\varepsilon^3}{4 c_\mathcal{L} \|b_\varepsilon\|_{\infty}}.
\]
where $c_L$ is the constant given by Lemma 4.17. Then, because of Lemma 4.17 (i),
\[
\sup_{t \in [0, t_0]} \left\| G_\varepsilon (\mu_1 - \mu_2)(t) \right\|_{L^1(\mathbb{R}^d)} = \sup_{t \in [0, t_0]} \left\| \int_0^t \mathcal{L}^\varepsilon (b_\varepsilon (\mu_1 - \mu_2)) (\tau) d\tau \right\|_{L^1(\mathbb{R}^d)} \\
\leq t_0 \frac{2c_L}{\varepsilon^3} \sup_{t \in [0, t_0]} \left\| \mu_1 - \mu_2 \right\|_{L^1(\mathbb{R}^d)} \\
\leq \frac{1}{2} \sup_{t \in [0, t_0]} \left\| \mu_1 - \mu_2 \right\|_{L^1(\mathbb{R}^d)}.
\]
Therefore, by the Banach fixed point theorem, problem (32) has a unique solution $\mu_\varepsilon \in C([0, t_0], L^1(\mathbb{R}^d))$ for every $\varepsilon > 0$. Since $t_0 > 0$ is constant for fixed $\varepsilon > 0$, we may immediately extend this solution to the interval $\mathcal{T}$ and conclude that problem (32) has a unique solution in the space $C(\mathcal{T}, L^1(\mathbb{R}^d)) \cap C^1(\mathcal{T}, L^1(\mathbb{R}^d))$.

\[\diamond \text{ Step 3. Compactness of } \{\mu_\varepsilon\} \text{ in } C(\mathcal{T}, P(\mathbb{R}^d)).\]

Because of the regularity of $\mu_\varepsilon$ obtained in Step 2, we have
\[
(34) \quad \partial_t \mu_\varepsilon = \mathcal{L}^\varepsilon (b_\varepsilon \mu_\varepsilon) \quad \text{in } C(\mathcal{T}, L^1(\mathbb{R}^d)).
\]
Consider $\text{sgn}(u)^- = 1_{\{u < 0\}}$. Then by (34),
\[
\int_0^t \int_{\mathbb{R}^d} \partial_t \mu_\varepsilon \text{sgn}(\mu_\varepsilon)^- dx d\tau = \int_0^t \int_{\mathbb{R}^d} \mathcal{L}^\varepsilon (b_\varepsilon \mu_\varepsilon) \text{sgn}(\mu_\varepsilon)^- dx d\tau.
\]
Since $b_\varepsilon > 0$, we have $\text{sgn}(\mu_\varepsilon)^- = \text{sgn}(b_\varepsilon \mu_\varepsilon)^-$ and for arbitrary real functions $u, v$, $v \text{sgn}(u)^- \geq v \text{sgn}(v)^- = -(v)^-$. Therefore
\[
(\mathcal{L}^\varepsilon (b_\varepsilon \mu_\varepsilon) \text{sgn}(\mu_\varepsilon)^-) (x) = \int_{\mathbb{R}^d} (b_\varepsilon \mu_\varepsilon (x-z) - b_\varepsilon \mu_\varepsilon (x)) \text{sgn}(b_\varepsilon \mu_\varepsilon)^- (x) v^\varepsilon (dz)
\geq - \int_{\mathbb{R}^d} (b_\varepsilon \mu_\varepsilon^- (x-z) - b_\varepsilon \mu_\varepsilon^- (x)) v^\varepsilon (dz) = - \mathcal{L}^\varepsilon ((b_\varepsilon \mu_\varepsilon^-) (x).
\]
By (31), $\int_{\mathbb{R}^d} \mathcal{L}^\varepsilon ((b_\varepsilon \mu_\varepsilon^-) dx = 0$. Hence
\[
0 \leq \int_0^t \int_{\mathbb{R}^d} \partial_t \mu_\varepsilon \text{sgn}(\mu_\varepsilon)^- dx d\tau = \int_0^t \int_{\mathbb{R}^d} -\partial_t (\mu_\varepsilon^-) dx d\tau = \\
= \int_{\mathbb{R}^d} (\mu_{0, \varepsilon})^- dx - \int_{\mathbb{R}^d} (\mu_\varepsilon^-) (t) dx.
\]
Since $\mu_{0, \varepsilon} = m_0 * \rho \geq 0$, i.e. $(\mu_{0, \varepsilon})^- = 0$, and $(\mu_\varepsilon^-) \geq 0$, this implies
\[
0 \leq \int_{\mathbb{R}^d} (\mu_\varepsilon^-) (t) dx \leq \int_{\mathbb{R}^d} (\mu_{0, \varepsilon})^- dx = 0.
\]
Therefore $\mu_\varepsilon(t) \geq 0$ for every $t \in \mathcal{T}$.

By Step 2, $\mu_\varepsilon$ is the fixed point of $G_\varepsilon$. Thus, because of (31), (33), and the Fubini-Tonelli theorem, we have
\[
\int_{\mathbb{R}^d} \mu_\varepsilon(t) dx = \int_{\mathbb{R}^d} \mu_{0, \varepsilon} dx + \int_0^t \int_{\mathbb{R}^d} \mathcal{L}^\varepsilon (b_\varepsilon \mu_\varepsilon) dx d\tau = 1.
\]
This, together with $\mu_\varepsilon \geq 0$, means that $\mu_\varepsilon(t) \in P_{ac}(\mathbb{R}^d)$ for every $t \in \mathcal{T}$. Since $\mu_\varepsilon \in C(\mathcal{T}, L^1(\mathbb{R}^d))$, it follows that $\mu_\varepsilon \in C(\mathcal{T}, P_{ac}(\mathbb{R}^d))$.

Notice that $\|b_\varepsilon\|_\infty \leq \|b + 1\|_\infty < B + 1$ by (A1'). Let $V$ be a Lyapunov function such that $m_0[V], \|\mathcal{L}V\|_\infty < \infty$ (see Corollary 4.14). By Definition 4.6,
\[
\mu_{0, \varepsilon}[V] = (m_0 * \rho \mu)[V] \leq m_0[V] + \|\nabla V\|_\infty \int_{B_1} |z| \rho \mu(z) dz \leq m_0[V] + 1.
\]
In combination with Lemma 4.17 (iii) we get
\[
\sup_{\varepsilon \in (0, 1)} (\|b_\varepsilon + \mu_{0, \varepsilon}[V] + \|\mathcal{L}^\varepsilon V\|_\infty + \|\mathcal{L}^\varepsilon \|_{LK}) < \infty.
\]

\[\text{First we show that } \{\mu_\varepsilon\} \subset C(\mathcal{T}, P_{ac}(\mathbb{R}^d)), \text{ then we establish its tightness.}\]
It follows from Lemma 6.2 that the family $\{\mu_\varepsilon\}$ is pre-compact in $C(\overline{T}, P(\mathbb{R}^d))$.

* Step 4. Passing to the limit. Using the result of Step 3, let $\varepsilon_k$ be a sequence such that $\mu_{\varepsilon_k} \to m$ in $C(\overline{T}, P(\mathbb{R}^d))$. By (34), for every $\varepsilon_k, \varphi \in C_c^\infty(\mathbb{R}^d)$ and $s,t \in T$, because $b_{\varepsilon_k} - b = \varepsilon_k$,

$$
\mu_{\varepsilon_k}(t)[\varphi] - \mu_{\varepsilon_k}(s)[\varphi] = \int_s^t \int_{\mathbb{R}^d} (L^{i_\varepsilon} * (\mu_{\varepsilon_k} b_{\varepsilon_k}^i)) \varphi \, dx \, dt
$$

$$
= \varepsilon_k \int_s^t \mu_{\varepsilon_k} [L^{i_\varepsilon} \varphi] \, dt + \int_s^t \mu_{\varepsilon_k} [b(L^{i_\varepsilon} \varphi) - L\varphi] \, dt + \int_s^t \mu_{\varepsilon_k} [bL \varphi] \, dt.
$$

Since $\lim_{k \to \infty} \|L^{i_\varepsilon} \varphi - L\varphi\|_\infty = 0$ by Lemma 4.17(ii) and the Hölder inequality,

$$
m(t)[\varphi] - m(s)[\varphi] = \int_s^t m(\tau)[b(\tau) L \varphi] \, d\tau.
$$

It follows that $m$ is a very weak solution of problem (3) (see Lemma 6.1(iii)). □

7. Fokker–Planck equation — Uniqueness

We show uniqueness of solutions of problem (3) using a Holmgren-type argument. The key step is to construct a test function by means of the “dual” equation

$$
\begin{cases}
\partial_t w - b L w = 0 & \text{on } T \times \mathbb{R}^d, \\
w(0) = \phi & \text{on } \mathbb{R}^d,
\end{cases}
$$

(35)

where $\phi$ is a sufficiently regular function. For the clarity of exposition, we consider problem (35) to “start” at 0 and “go forward” in time. This has to be reversed when we apply those results in Theorem 7.10, the proof of which is the main goal of this section.

In addition to (A1’), we also need to assume that $b$ is Hölder continuous.

(A1’

: $b$ satisfies (A1’); in addition, $b \in B(\mathcal{T}, C_b^\beta(\mathbb{R}^d))$ for some $\beta > 0$.

Our work is divided into two parts. In Section 7.1 we assume (LA) and allow for degenerate equations with $b \geq 0$, but we obtain a restriction on the Hölder exponent $\beta$ (i.e. on $\sigma$). Then in Section 7.2 we have no restrictions on $\beta$ and $\sigma$, but we need to have $b \geq \kappa > 0$, and $\mathcal{L}$ needs to satisfy (LB) or (LC). When (LB) holds, we also assume $b \in UC(\mathcal{T} \times \mathbb{R}^d)$.

7.1. Dual equation in the degenerate case. In this case we employ viscosity solutions techniques to prove existence of solutions of problem (35), and later show that the viscosity solution is a bounded classical solution under certain assumptions.

**Theorem 7.1.** Assume (LA), (A1’), $\phi \in C_b(\mathbb{R}^d)$, $\sigma \in (0, \frac{1}{2})$, and $\beta \in \left[\frac{2\sigma}{1-2\sigma}, 1\right]$. The comparison principle holds for problem (35) (see Definition 4.22).

Before we prove this theorem, we need to introduce notation and establish lemmas concerning the doubling of variables method. For every $\varepsilon, \delta > 0$ let

$$
\psi_{\varepsilon, \delta}(x, y) = \frac{|x - y|^2}{\varepsilon} + \delta (V(x) + V(y)),
$$

(36)

where $V$ is a Lyapunov function such that $\|L V\|_\infty < \infty$ (see Corollary 4.14).

Suppose $\Phi: (\overline{T} \times \mathbb{R}^d)^2 \to \mathbb{R}$ is a bounded upper-semicontinuous function and \{\(a_{\varepsilon, \delta} \}_{\varepsilon, \delta > 0} \subset \mathbb{R} is a bounded set. Let $\eta, \varepsilon, \delta > 0$ and $\psi_{\varepsilon, \delta}$ be given by (36). Define

$$
\Psi_{\eta, \varepsilon, \delta}(t, x, s, y) = \Phi(t, x, s, y) - \psi_{\varepsilon, \delta}(x, y) - \frac{|t - s|^2}{\eta} - a_{\varepsilon, \delta} \frac{t + s}{2}.
$$

17This corresponds to the standard uniform ellipticity assumption, see also (FB).
Lemma 7.2. For every $\eta, \varepsilon, \delta > 0$ the function $\Psi_{\eta, \varepsilon, \delta}$ has a maximum point $(t_*, x_*, s_*, y_*) \in (\mathcal{T} \times \mathbb{R}^d)^2$ such that
\[
\frac{|x_* - y_*|^2}{\varepsilon} + \frac{|t_* - s_*|^2}{\eta} \leq \Phi(t_*, x_*, s_*, y_*) - \frac{\Phi(t_*, x_*, t_*, x_*) + \Phi(s_*, y_*, s_*, y_*)}{2},
\]
(38)
and for every $\delta > 0$ there exist subsequence $\eta_k$ such that for every $\varepsilon > 0$
\[
\lim_{\eta_k \to 0} \frac{|t_* - s_*|^2}{\eta_k} = 0 \quad \text{and} \quad \lim_{\eta_k \to 0} (t_*, x_*, s_*, y_*) = (t_*, x_*, s_*, y_*) = (t_*, x_*, s_*, y_*) = (t_*, x_*, t_*, x_*) \delta > \eta, \varepsilon, \delta.
\]
and a subsequence $\varepsilon_n$ such that
\[
\lim_{\varepsilon_n \to 0} \lim_{\eta_k \to 0} \frac{|x_* - y_*|^2}{\varepsilon_n} = 0 \quad \text{and} \quad \lim_{\varepsilon_n \to 0} \lim_{\eta_k \to 0} (t_*, x_*, s_*, y_*) = (t_*, x_*, t_*, x_*) \delta > \eta, \varepsilon, \delta.
\]
where $(t_*, x_*, s_*, y_*)$ and $(t_*, x_*, t_*, x_*)$ denote the respective limit points.

Proof. Notice that the function
\[
(t, x) \mapsto u(t, x) = (v(s_*, y_*) + \psi_{t, \delta}(x, y_*) - \frac{|t_* - s_*|^2}{\eta} - \alpha_{t, \delta} \frac{t_* + s_*}{2})
\]
attains a maximum at the point $(t_*, x_*)$, while
\[
(s, y) \mapsto v(s, y) = (u(t_*, x_*) - \psi_{x, \delta}(x_*, y) - \frac{|t_* - s_*|^2}{\eta} - \alpha_{t, \delta} \frac{t_* + s_*}{2})
\]
has a minimum point at $(s_*, y_*)$. Because $t_*, s_* > 0$, by Definition 4.20 we have
\[
\frac{\alpha_{t, \delta}}{2} + \frac{2(t_* - s_*)}{\eta} - b(t_*, x_*) \left( \mathcal{L}^r u(t_*, x_*) + \mathcal{L}^r \psi_{x, \delta}(\cdot, y_*)(x_*) \right) \leq 0,
\]
\[
\frac{-\alpha_{t, \delta}}{2} + \frac{2(t_* - s_*)}{\eta} - b(s_*, y_*) \left( \mathcal{L}^r v(s_*, y_*) - \mathcal{L}^r \psi_{x, \delta}(x_*, \cdot)(y_*) \right) \geq 0.
\]
We add these two inequalities and obtain
\[
\alpha_{t, \delta} \leq b(t_*, x_*) \left( \mathcal{L}^r u(t_*, x_*) + \mathcal{L}^r \psi_{x, \delta}(\cdot, y_*)(x_*) \right)
\]
\[- b(s_*, y_*) \left( \mathcal{L}^r v(s_*, y_*) - \mathcal{L}^r \psi_{x, \delta}(x_*, \cdot)(y_*) \right)
\]
(39)
\[
= b(s_*, y_*) \mathcal{L}^r \psi_{x, \delta}(x_*, \cdot)(y_*) + b(t_*, x_*) \mathcal{L}^r \psi_{x, \delta}(\cdot, y_*)(x_*)
\]
\[+ b(t_*, x_*) \left( (\mathcal{L}^r u)(t_*, x_*) - (\mathcal{L}^r v)(s_*, y_*) \right)
\]
\[+ b(t_*, x_*) \left( (\mathcal{L}^r v)(s_*, y_*) \right).
\]
Observe that for every $z \in \mathbb{R}^d$ we have
\[
\Psi_{\eta, x, \delta}(t_*, x_*, z, s_*, y_*, y_*) \leq \Psi_{\eta, x, \delta}(t_*, x_*, s_*, y_*)
\]
which implies
\[ u(t_\ast, x_\ast + z) - v(s_\ast, y_\ast + z) - u(t_\ast, x_\ast) + v(s_\ast, y_\ast) \]
\[ \leq \delta \left( V(x_\ast + z) + V(y_\ast + z) - V(x_\ast) - V(y_\ast) \right). \]
Therefore, because of (LA), for every \( r \in (0, 1) \)
\begin{equation}
(\mathcal{L} u)(t_\ast, x_\ast) - (\mathcal{L} v)(s_\ast, y_\ast) \leq 2\delta \left( \|\nabla V\|_\infty + \|\nabla V\|_\infty \int_{B_r} |z| \nu (dz) \right).
\end{equation}
We also find that
\begin{equation}
(\mathcal{L} v)(t_\ast, y_\ast) \leq 2\|v\|_\infty \left( \nu(B_1^\ast) + \int_{B_1 \setminus B_r} \nu (dz) \right).
\end{equation}
By using (40), (41), (LA), and Proposition D.1 in inequality (39), we obtain
\[
a_{\varepsilon, \delta} \leq \frac{2K \|b\|_\infty}{1 - 2\sigma} \left( 2\delta + \frac{2|x_\ast - y_\ast| + r}{\varepsilon} \right)^{1 - 2\sigma} + 2\|v\|_\infty \|L V\|_\infty
\]
\[
+ 2\|v\|_\infty |b(t_\ast, x_\ast) - b(s_\ast, y_\ast)| \left( \nu(B_1^\ast) + \frac{K}{1 - 2\sigma} r^{-2\sigma} \right).
\]
We now take the limit (see Lemma 7.2) as \( \eta_k \to 0 \), and recall that
\[
\lim_{\eta_k \to 0} t_\ast = \lim_{\eta_k \to 0} s_\ast = t_\varepsilon, \delta \quad \text{and} \quad \lim_{\eta_k \to 0} (x_\ast, y_\ast) = (x_\varepsilon, \delta, y_\varepsilon, \delta),
\]
which for an adequate \( C > 0 \) gives us
\[
a_{\varepsilon, \delta} \leq C \left( \frac{r + |x_\varepsilon, \delta - y_\varepsilon, \delta|}{\varepsilon} r^{1 - 2\sigma} + |b(t_\varepsilon, \delta, x_\varepsilon, \delta) - b(t_\varepsilon, \delta, y_\varepsilon, \delta)| \left( 1 + r^{-2\sigma} \right) + \delta \right).
\]
Because of (A1’’), for another constant \( C \), we have
\[
a_{\varepsilon, \delta} \leq C \left( \frac{r + |x_\varepsilon, \delta - y_\varepsilon, \delta|}{\varepsilon} r^{1 - 2\sigma} + |x_\varepsilon, \delta - y_\varepsilon, \delta| \left( 1 + r^{-2\sigma} \right) + \delta \right). \quad \Box
\]
Now we are in a position to prove Theorem 7.1.

Proof of Theorem 7.1. We continue to use the notation introduced earlier in this Section. Denote
\[
M_0 = \sup_{x \in \mathbb{R}^d} \left( u(0, x) - v(0, x) \right) \quad \text{and} \quad M = \sup_{t \in \mathcal{T}} \sup_{x \in \mathbb{R}^d} \left( u(t, x) - v(t, x) \right),
\]
and assume by contradiction that \( M_0 \leq 0 \) and \( M > 0 \). Because the functions \( u \) and \( v \) are bounded, we also have \( M < \infty \).

In definition (37) of the function \( \Psi_{\eta, \varepsilon, \delta} \), let \( a_{\varepsilon, \delta} = \frac{M}{27} \) for every \( \varepsilon, \delta > 0 \) and \( \Phi(t, x, s, y) = u(t, x) - v(s, y) \) as in Lemma 7.3. Let the points \( (t_\ast, x_\ast, s_\ast, y_\ast) \) and sequences \( \{\eta_k\}, \{\varepsilon_n\} \) be given by Lemma 7.2.

Suppose \( \lim_{\varepsilon_n, \eta_k \to 0} t_\ast = t_\delta = 0 \). Then for every \( \delta > 0 \) we have
\begin{equation}
\limsup_{\varepsilon_n, \eta_k \to 0} \sup_{t \in \mathcal{T}} \Psi_{\eta_k, \varepsilon_n, \delta}(t_\ast, x_\ast, s_\ast, y_\ast) \leq \Phi(0, x_\delta, 0, x_\delta) \leq M_0 \leq 0.
\end{equation}
On the other hand, by the definition of \( M \), there exists a point \( (t_M, x_M) \) such that \( \Phi(t_M, x_M, t_M, x_M) \geq \frac{4}{9} M \). Take \( \delta > 0 \) such that \( \delta V(x_M) \leq \frac{1}{16} M \). Then we get
\[
\Psi_{\eta_k, \varepsilon_n, \delta}(t_\ast, x_\ast, s_\ast, y_\ast) \geq \Phi(t_M, x_M, t_M, x_M) - 2\delta V(x_M) - \frac{1}{2} M \geq M \left( \frac{4}{9} - \frac{1}{5} - \frac{1}{2} \right) = \frac{1}{6} M > 0.
\]
This contradicts (42) and shows that \( t_\delta > 0 \) for \( \delta \leq \frac{M}{\inf V(x_M)} \). Hence, without loss of generality, we may assume \( t_\ast, s_\ast > 0 \).
We now use Lemma 7.3 to obtain
\begin{equation}
\frac{M}{2CT} \leq \varepsilon_n r + \frac{|x_{\varepsilon_n, \delta} - y_{\varepsilon_n, \delta}|}{\varepsilon_n} r^{1-2\sigma} + \frac{|x_{\varepsilon_n, \delta} - y_{\varepsilon_n, \delta}|}{\varepsilon_n} \frac{\beta(1 + r^{-2\sigma}) + \delta}{1 + r^{-2\sigma}}.
\end{equation}
We put \( r^{2\sigma} = \varepsilon_n^{\beta/2} \) (see Remark 7.4(a)) and get
\begin{equation}
\frac{M}{2CT} \leq \varepsilon_n^{\beta(1-\sigma)/2} + \frac{|x_{\varepsilon_n, \delta} - y_{\varepsilon_n, \delta}|^2}{\varepsilon_n} \frac{\beta(1-\sigma)}{4\sigma} \frac{1}{\frac{\beta(1-\sigma)}{4\sigma}} + \frac{|x_{\varepsilon_n, \delta} - y_{\varepsilon_n, \delta}|^2}{\varepsilon_n} + |x_{\varepsilon_n, \delta} - y_{\varepsilon_n, \delta}|^{\beta + \delta}.
\end{equation}
Assumption \( \beta \geq \frac{2\sigma}{1-2\sigma} \) is equivalent to \( \beta(1-\sigma)/4\sigma \geq 1/2 \geq 0 \) (see Remark 7.4(b)). By Lemma 7.2 we have \( \varepsilon_n^{1}|x_{\varepsilon_n, \delta} - y_{\varepsilon_n, \delta}|^2 \rightarrow 0 \), thus the expression on the right-hand side converges to \( \delta \) as \( \varepsilon_n \rightarrow 0 \). Since \( \delta \) is arbitrary, we obtain \( M \leq 0 \), which is a contradiction.

Remark 7.4. Eventually want to put \( b = F'(L \alpha) \). In the best case, both \( u \) and \( F' \) may be Lipschitz and then \( \beta = 1 - 2\sigma \). Our aim is to obtain the most lenient estimate on \( M \) in terms of \( \beta \).

(a) To this end, we cannot do better than substituting \( r^{2\sigma} = \varepsilon_n^{\beta/2} \). If \( a \) is such that \( r = \varepsilon_n^{\sigma} \), then we need to have \( \beta(1-\sigma)/4\sigma \geq 0 \) and \( (1-2\sigma)a - \frac{\beta(1-\sigma)}{4\sigma} \geq 0 \). By combining both inequalities we obtain \( \frac{\beta(1-\sigma)}{4\sigma} \geq a \geq \frac{\beta(1-\sigma)}{4\sigma} \). Hence, we get \( \beta \geq \frac{2\sigma}{1-2\sigma} \). When \( \beta = 1 - 2\sigma \), this translates to \( \sigma \in (0, \frac{3-\sqrt{5}}{4}) \), and \( \frac{3-\sqrt{5}}{4} \approx 0.236 \).

(b) When \( \nu \) is symmetric at the origin (see Definition 4.11), we may obtain a better estimate on \( \sigma \). Using Proposition D.1 for the symmetric case in the last lines of the proof of Lemma 7.3, allows us to replace (43) with
\begin{equation}
\frac{M}{2CT} \leq \varepsilon_n^{-1} r^{2-2\sigma} + |x_{\varepsilon_n, \delta} - y_{\varepsilon_n, \delta}|^{\beta(1 + r^{-2\sigma})} + \delta.
\end{equation}
Under the same scaling \( r^{2\sigma} = \varepsilon_n^{\beta/2} \), the dominant exponent is then \( \beta(1-\sigma)/4\sigma \). It has to be strictly positive, hence \( \beta > \frac{2\sigma}{1-2\sigma} \). When \( \beta = 1 - 2\sigma \), this translates to \( \sigma \in (0, \frac{5-\sqrt{22}}{4}) \), and \( \frac{5-\sqrt{22}}{4} \approx 0.236 \).

Corollary 7.5. Assume (LA), (A1’’), \( \phi \in C_b(\mathbb{R}^d) \), \( \sigma \in (0, \frac{1}{2}) \), and \( \beta \in \left( \frac{2\sigma}{1-2\sigma}, 1 \right] \).
There exists a viscosity solution of problem (35).
Proof. Notice that \( u \equiv -\|\phi\|_{\infty} \) is a subsolution of problem (35), while \( v \equiv \|\phi\|_{\infty} \) is a supersolution. Using Theorem 7.1, existence of a (unique) bounded continuous viscosity solution follows by the Perron method (cf. e.g. the proof of [9, Theorem 2.3] for a similar result).

We next show the Hölder-continuity of the viscosity solution.

Theorem 7.6. Assume (LA), (A1’’), \( \phi \in C^1_b(\mathbb{R}^d) \), \( \sigma \in (0, \frac{1}{2}) \), and \( \beta \in \left( \frac{2\sigma}{1-2\sigma}, 1 \right] \).
If \( w \) is a viscosity solution of problem (35), then \( w \in C_0^{\beta_0}(\mathbb{R}^d) \), where \( \beta_0 = \beta - \frac{2\sigma}{1-2\sigma} \).
Proof. \( \diamond \) Step 1. For every \( \varepsilon, \delta > 0 \) we define
\[
M_0^{\varepsilon, \delta} = \sup_{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d} \left\{ w(0, x) - w(0, y) - \psi_{\varepsilon, \delta}(x, y) \right\}
\]
and
\[
M^{\varepsilon, \delta} = \sup_{t \in T} \sup_{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d} \left\{ w(t, x) - w(t, y) - \psi_{\varepsilon, \delta}(x, y) \right\}.
\]
Note that we have
\begin{equation}
0 \leq (M_{\varepsilon,\delta} - M_{\varepsilon,\delta}^0) \leq 4\|w\|_{\infty},
\end{equation}
and for every \((t, x, y) \in \mathcal{T} \times \mathbb{R}^d \times \mathbb{R}^d\) it holds that
\begin{equation}
w(t, x) - w(t, y) \leq M_{\varepsilon,\delta} + \psi_{\varepsilon,\delta}(x, y).
\end{equation}

We have \(w(0) = \phi \in C_1^1(\mathbb{R}^d)\) and without loss of generality (since the equation is linear) we may assume \(|\phi|_1 \leq 1\). Then
\[
 w(0, x) - w(0, y) - \psi_{\varepsilon,\delta}(x, y) \leq |x - y| - \frac{|x - y|^2}{\varepsilon} \leq \frac{\varepsilon}{4},
\]
and thus \(M_{\varepsilon,\delta}^0 \leq \varepsilon/4\). We now consider the function \(\Psi_{\eta,\varepsilon,\delta}\) given by (37) with
\[
\Phi(t, x, s, y) = w(t, x) - w(s, y) \quad \text{and} \quad a_{\varepsilon,\delta} = \frac{M_{\varepsilon,\delta} - M_{\varepsilon,\delta}^0}{2T}.
\]
Let \((t_*, x_*, s_*, y_*)\) denote the maximum points of \(\Psi_{\eta,\varepsilon,\delta}\) given by Lemma 7.2, together with sequences \(\{\eta_k\}\) and \(\{\varepsilon_n\}\). Recall that
\[
\lim_{\eta_k \to 0} (t_*, x_*, s_*, y_*) = (t_{\varepsilon,\delta}, x_{\varepsilon,\delta}, t_{\varepsilon,\delta}, y_{\varepsilon,\delta}).
\]

Let us fix \(\varepsilon, \delta > 0\). If \(t_{\varepsilon,\delta} = 0\), then
\[
\lim_{\eta_k \to 0} \Psi_{\eta,\varepsilon,\delta}(t_*, x_*, s_*, y_*) = w(0, x_{\varepsilon,\delta}) - w(0, y_{\varepsilon,\delta}) - \psi_{\varepsilon,\delta}(x_{\varepsilon,\delta}, y_{\varepsilon,\delta}) \leq M_{\varepsilon,\delta}^0.
\]

Notice that for every \((t, x, y) \in \mathcal{T} \times \mathbb{R}^d \times \mathbb{R}^d\)
\[
\Psi_{\eta,\varepsilon,\delta}(t_*, x_*, s_*, y_*) \geq w(t, x) - w(t, y) - \psi_{\varepsilon,\delta}(x, y) - \frac{M_{\varepsilon,\delta} - M_{\varepsilon,\delta}^0}{2}.
\]

Thus for every \(\eta > 0\) we have
\[
\Psi_{\eta,\varepsilon,\delta}(t_*, x_*, s_*, y_*) \geq M_{\varepsilon,\delta} - \frac{M_{\varepsilon,\delta} - M_{\varepsilon,\delta}^0}{2} = \frac{M_{\varepsilon,\delta} + M_{\varepsilon,\delta}^0}{2}.
\]

It then follows that
\[
M_{\varepsilon,\delta} \leq 2 \lim_{\eta_k \to 0} \Psi_{\eta,\varepsilon,\delta}(t_*, x_*, s_*, y_*) - M_{\varepsilon,\delta}^0 \leq M_{\varepsilon,\delta},
\]
and by (44) we get \(M_{\varepsilon,\delta} = M_{\varepsilon,\delta}^0\). Because of (45), for every \((t, x, y) \in \mathcal{T} \times \mathbb{R}^d \times \mathbb{R}^d\) we thus have
\begin{equation}
w(t, x) - w(t, y) \leq M_{\varepsilon,\delta}^0 + \psi_{\varepsilon,\delta}(x, y) \leq \frac{\varepsilon}{4} + \frac{|x - y|^2}{\varepsilon} + \delta\left(V(x) + V(y)\right).
\end{equation}

In turn, if \(t_{\varepsilon,\delta} > 0\), then without loss of generality we may assume that \(t_*, s_0 > 0\). By Lemma 7.3 we therefore obtain
\begin{equation}
\frac{M_{\varepsilon,\delta} - M_{\varepsilon,\delta}^0}{2CT} \leq \frac{r + |x_{\varepsilon,\delta} - y_{\varepsilon,\delta}|}{\varepsilon}^{1-2\sigma} + |x_{\varepsilon,\delta} - y_{\varepsilon,\delta}|^{\beta}(1 + r^{-2\sigma}) + \delta.
\end{equation}

We also have \(|x_{\varepsilon,\delta} - y_{\varepsilon,\delta}|^2 \leq 2\varepsilon\|w\|_{\infty}\), thanks to (38). Thus, by combining (45) and (47), we get
\[
w(t, x) - w(t, y) \leq M_{\varepsilon,\delta}^0 + \psi_{\varepsilon,\delta}(x, y) + 2CT\delta
\]
\[
+ 2CT \left(\frac{r + |x_{\varepsilon,\delta} - y_{\varepsilon,\delta}|}{\varepsilon}^{1-2\sigma} + |x_{\varepsilon,\delta} - y_{\varepsilon,\delta}|^{\beta}(1 + r^{-2\sigma})\right)
\leq \frac{\varepsilon}{4} + \frac{|x - y|^2}{\varepsilon} + \delta \left(V(x) + V(y) + 2CT\right)
\]
\[
+ 2CT \varepsilon^{-2\sigma} \left(\frac{r^2}{\varepsilon} + \sqrt{\frac{2\|w\|_{\infty}r^2}{\varepsilon}} + 2\varepsilon^{\beta/2}\varepsilon^{\beta/2}(1 + r^{2\sigma})\right).
\]
We stress that the above inequality holds for all \((t, x, y) \in T \times \mathbb{R}^d \times \mathbb{R}^d\), and since the right-hand side dominates the right-hand side in (46), it also holds for every \(\varepsilon, \delta > 0\).

By taking \(\delta \to 0\) for fixed \(t, x, y\), and \(\varepsilon\) we thus get

\[
 w(t, x) - w(t, y) \leq \frac{\varepsilon}{4} + \frac{|x - y|^2}{\varepsilon} + c_w r^{-2\sigma} \left( \frac{r^2}{\varepsilon} + \sqrt{\frac{r^2}{\varepsilon} + \varepsilon^{\beta/2}} \right),
\]

where \(c_w = 8CT\max\{1, \sqrt{\|w\|_{\infty}}\}\). To balance the second and the third terms in the parenthesis, we put \(r^2 = \varepsilon^{\beta+1}\) for every \(\varepsilon \in (0, 1)\). Since \(\frac{\beta - 2\sigma(\beta + 1)}{2} \in (0, 1)\) as \(\beta \in \left(\frac{2\sigma}{1-2\sigma}, 1\right)\) and \(\sigma \in (0, \frac{1}{2})\), this gives us

\[
 w(t, x) - w(t, y) \leq \frac{\varepsilon}{4} + \frac{|x - y|^2}{\varepsilon} + 3c_w \varepsilon^{-2n(\beta+1)}
\]

\[
 \leq C_1 \left( \frac{|x - y|^2}{\varepsilon} + \varepsilon^{-2\sigma(\beta+1)(1-2\sigma-1)} \right),
\]

where \(C_1 = \max\{1, 4c_w\}\). When \(|x - y| < 1\), we let \(\varepsilon = |x - y|^{\omega_1}\) for

\[
 \omega_1 = \frac{4}{(\beta + 1)(1 - 2\sigma) + 1} = \frac{4}{(\beta + 2)(1 - 2\sigma) + 2\sigma}.
\]

Notice that \(\omega_1 \in \left(\frac{3}{4}, 2\right)\), since \(\beta \in \left(\frac{2\sigma}{1-2\sigma}, 1\right)\) and \(\sigma \in (0, \frac{1}{2})\). Then

\[
 w(t, x) - w(t, y) \leq 2C_1|x - y|^{2-\omega_1}
\]

for every \((t, x, y) \in T \times \mathbb{R}^d \times \mathbb{R}^d\) such that \(|x - y| < 1\) (cf. Definition 4.1).

\(\diamond\) \textbf{Step 2.} We “bootstrap” the argument of \textbf{Step 1} to improve the Hölder exponent.

By combining (38) and (49), after passing to the limit \(\eta_k \to 0\), we get

\[
 \frac{|x_{e, \delta} - y_{e, \delta}|^2}{\varepsilon} \leq w(t_{e, \delta}, x_{e, \delta}) - w(t_{e, \delta}, y_{e, \delta}) \leq 2C_1|x_{e, \delta} - y_{e, \delta}|^{2-\omega_1}.
\]

It follows that \(|x_{e, \delta} - y_{e, \delta}|^{\omega_1} \leq 2C_1 \varepsilon\).

Now we go back to (47) and follow the subsequent arguments, using the new bound. We obtain

\[
 w(t, x) - w(t, y) \leq \frac{\varepsilon}{4} + \frac{|x - y|^2}{\varepsilon} + c_1 r^{-2\sigma} \left( \frac{r^2}{\varepsilon} + \varepsilon^{\omega_1 - 1} r + \varepsilon^{\beta / \omega_1} \right),
\]

where \(c_1 = 8CTC_1^{\omega_1}\) (note that \(C_1, \omega_1 \geq 1\)). To balance the second and the third terms in the parenthesis, for every \(\varepsilon \in (0, 1)\) we put \(r^2 = \varepsilon^{\beta+\omega_1-1}\) (see \textbf{Remark 7.7 (a)}). Since \(\frac{\beta - 2\sigma(\beta + \omega_1 - 1)}{\omega_1} < 1\) (because \(\beta \leq 1 < \omega_1\)) is then the dominant exponent in (50), this gives us

\[
 w(t, x) - w(t, y) \leq C_2 \left( \frac{|x - y|^2}{\varepsilon} + \varepsilon^{-2\sigma(\beta+\omega_1-1)} \right),
\]

where \(C_2 = \max\{1, 4c_1\}\). Choosing \(\varepsilon = |x - y|^{\omega_2}\) for \(\omega_2 = \frac{2\omega_1}{(\beta + \omega_1)(1 - 2\sigma) + 2\sigma}\) gives us

\[
 w(t, x) - w(t, y) \leq 2C_2|x - y|^{2-\omega_2}.
\]

By repeating this procedure, we obtain recursive formulas

\[
 \begin{align*}
 \omega_0 &= 2, & \omega_{n+1} &= \frac{2\omega_n}{(\beta + \omega_n)(1 - 2\sigma) + 2\sigma}, \\
 C_0 &= \max\{1, \|w\|_{\infty}\}, & C_{n+1} &= \max\{1, 32CTC_n^{\omega_n}\},
\end{align*}
\]

for \(n \in \mathbb{N}\).
Notice that $\beta + \omega_0 > \frac{2\sigma}{1-2\sigma} + 2 = \frac{2-2\sigma}{1-2\sigma}$. Now, assume $\beta + \omega_n > \frac{2-2\sigma}{1-2\sigma}$ for some $n \in \mathbb{N}$. Then,

$$
\beta + \omega_{n+1} = \frac{\beta (\beta + \omega_n)(1-2\sigma) + 2\beta \sigma + 2\omega_n}{(\beta + \omega_n)(1-2\sigma) + 2\sigma} > \frac{2}{(\beta + \omega_n)(1-2\sigma) + 2\sigma} = \frac{2}{(1-2\sigma) + \frac{2\sigma}{\beta + \omega_n}} > \frac{2}{(1-2\sigma) + \frac{2\sigma}{\frac{2-2\sigma}{1-2\sigma}}} = \frac{2}{1} - \frac{2\sigma}{1-2\sigma}.
$$

By the principle of induction, we get $\beta + \omega_n > \frac{2-2\sigma}{1-2\sigma} > 2$ for every $n \in \mathbb{N}$. Then,

$$
\frac{\omega_{n+1}}{\omega_n} = \frac{2}{(\beta + \omega_n)(1-2\sigma) + 2\sigma} < \frac{2}{1-2\sigma} = 1,
$$

i.e. $\omega_{n+1} < \omega_n$. This also implies $\omega_n \in (1, 2]$, since $\omega_0 = 2$ and $2 - \omega_n < \beta \leq 1$.

Passing to the limit in (52) we then find that $\lim_{n \to \infty} \omega_n = \frac{2-2\sigma}{1-2\sigma} = \beta = \omega_\infty$.

By (52), notice that $C_n \geq 1$ for every $n \in \mathbb{N}$. Moreover, if $32CT \leq 1$ and $C_{n_0} = 1$ for some $n_0 \in \mathbb{N}$, then $C_n = 1$ for every $n \geq n_0$. In any other case, $C_{n+1} = 32CTC_n^\frac{1}{\omega_n}$ for every $n \in \mathbb{N}$. Then

$$
C_{n+1} = (32CT)^{\Sigma_n} C_0^{\Pi_n}, \quad \text{where } \Pi_n = \prod_{k=1}^n \frac{1}{\omega_k} \text{ and } \Sigma_n = \Pi_n + \sum_{k=1}^n \Pi_k - \Pi_k.
$$

We observe that $\lim \Pi_n = 0$ because $\omega_n \geq \omega_\infty > 1$ (since $\beta \leq 1$ and $\sigma > 0$) and $\lim \Sigma_n = \sum_{k=0}^{\infty} \frac{1}{\omega_k} = 1 + \frac{1-2\sigma}{1-2\sigma} = \infty$ since $\beta \leq 1$. In either case, $\lim_{n \to \infty} C_n = \infty$.

By writing (51) for every $n$ and then passing to the limit $n \to \infty$, we get $w \in B(T, C^2_b(\mathbb{R}^d))$, where

$$
\beta_0 = \lim_{n \to \infty} (2 - \omega_n) = \beta - \frac{2\sigma}{1-2\sigma}.
$$

\(\square\)

**Remark 7.7.** (a) Our aim is to obtain the best Hölder regularity. The choice of scaling $r^2 = \varepsilon^3 + 1$ in (48) is clearly optimal. When we repeat this argument in (50), we want the **lowest** of the three exponents to be the **highest** possible. If $r = \varepsilon^a$, then the exponents are

$$(2 - 2\sigma) a - 1, \quad (1 - 2\sigma) a + 1/\omega_n - 1, \quad -2\sigma a + 1/\omega_n,$$

which are affine functions of $a$. The first two are increasing, and the third is decreasing, hence the optimal choice is at the intersection of either 1st and 3rd, or 2nd and 3rd lines, which corresponds to $a = \max \left\{ \frac{2-2\sigma-1}{\omega_n}, \frac{2-2\sigma}{2\omega_n} \right\}$. We have $a = \frac{1}{\beta + \omega_n - 1}$, since $\beta + \omega_n > 2$.

(b) If the Lévy measure $\nu$ is symmetric at the origin (see Definition 4.11), the proof of Theorem 7.6 leads to $w \in B(T, C^2_b(\mathbb{R}^d))$, where $\beta_0 = \beta - \frac{2\sigma}{1-2\sigma}$ (cf. Remark 7.4 (b)).

**Theorem 7.8.** Assume (LA), (A1′′), $\phi \in C^1_b(\mathbb{R}^d)$, $\sigma \in (0, \frac{3-\sqrt{5}}{2})$, $\beta \in (2\sigma + \frac{2\sigma}{1-2\sigma}, 1]$. There exists a bounded classical solution of problem (35).

**Proof.** The condition on $\sigma$ ensures that $2\sigma + \frac{2\sigma}{1-2\sigma} < 1$. Consider the viscosity solution $w$ of problem (35) given by Corollary 7.5. By Theorem 7.6 we have $w \in B(T, C^2_b(\mathbb{R}^d))$, where $\beta_0 = \beta - \frac{2\sigma}{1-2\sigma}$. Since $\beta_0 > 2\sigma$, by Proposition 4.18 we have $Cw \in B(T, C^3_b(\mathbb{R}^d))$ and by Lemma 4.23, $w$ is a bounded classical solution of problem (35). \(\square\)
7.2. Dual equation in the non-degenerate case. When (LC) holds, we write \( bL = A + B \), where

\[
(A \phi)(t, x) = \int_{\mathbb{R}^d} (\phi(t, x + z) - \phi(t, x) - \mathbf{1}_{|z| \leq 2\sigma}(2\sigma) z \cdot \nabla \phi(x)) b(t, x) \frac{\tilde{k}(z)}{|z|^{d+2\sigma}} dz,
\]

\( \tilde{k}(z) = \mathbf{1}_{B_1} k(z) + \mathbf{1}_{B_1^c} k(\frac{z}{|z|}) \) is a normal extension of \( k \) (defined in (LC)) to \( \mathbb{R}^d \), and \( B = bL - A : C_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d) \) is a bounded operator (with Lévy measure supported on \( B_1^c \)).

**Theorem 7.9.** Assume (A1’’), \( b \geq \kappa \) for some \( \kappa > 0 \), and \( \phi \in C_c^\infty(\mathbb{R}^d) \). If either

(i) (LB) and \( b \in UC(T \times \mathbb{R}^d) \);

(ii) (LC);

holds, then problem (35) has a bounded classical solution and \( Lw \in B(T, C_b^\alpha(\mathbb{R}^d)) \).

**Proof.** Part (i). The statement follows from [67, Theorem 5.1.9] (see [67, page 175] for relevant notation).

Part (ii). Because \( \phi \in C_c^\infty(\mathbb{R}^d) \), we have \( L\phi \in C_b^\infty(\mathbb{R}^d) \) and thus by (A1’’') we get \( bL\phi \in C_b(\mathcal{T}, C_b^\alpha(\mathbb{R}^d)) \). Notice that \( w \) is a bounded classical solution of problem (35) if and only if \( v = w - \phi \) is a bounded classical solution of

\[
\begin{cases}
\partial_t v - bL \phi = bL \phi & \text{on } T \times \mathbb{R}^d, \\
v(0) = 0 & \text{on } \mathbb{R}^d.
\end{cases}
\]

We study problem (54) using the results proved in [70]. We check the assumptions for operators \( A \) and \( B \) given by (53). Assumption A in [70] is satisfied, because we assume (LC), (A1’’’), and \( b \geq \kappa > 0 \). To verify assumptions B1 and B2 in [70], we choose \( c(t, x, v) = v \), \( U_n = B_1 \), and \( \pi = \nu^1_{B_1} \) (in the notation of [70]) and again use (LC), (A1’’’).

By [70, Theorem 4] there exists a unique solution \( v \) of problem (54) such that \( L \phi \in B(T, C_b^\alpha(\mathbb{R}^d)) \) (see Remark 4.19) and \( \partial_t v \in C_b(\mathcal{T} \times \mathbb{R}^d) \) (see [70, Definition 3]). Thus \( w = v - \phi \) is a bounded classical solution of problem (35). \( \square \)

7.3. Uniqueness for Fokker–Planck. We are now in a position to prove uniqueness of solutions of problem (3). Since \( T \) is arbitrary, we shall apply Theorem 7.8 and Theorem 7.9 on \([0, t)\) for \( t < T \) in place of \( T \). We also consider condition (A1’’’) on this smaller interval.

**Theorem 7.10.** Assume (A1’’’) on \([0, t] \) for every \( t \in \mathcal{T} \), (A3), and any of the following

(i) (LA) and \( \beta > 2\sigma + \frac{2\sigma}{1-2\sigma} \);

(ii) (LA) with \( \nu \) symmetric at the origin and \( \beta > 2\sigma + \frac{2\sigma}{1-\sigma} \);

(iii) (LB), \( b \geq \kappa \) for some \( \kappa > 0 \), and \( b \in UC([0, t] \times \mathbb{R}^d) \) for every \( t \in \mathcal{T} \);

(iv) (LC) and \( b \geq \kappa \) for some \( \kappa > 0 \).

Then problem (3) has precisely one very weak solution.

**Proof.** Existence of a very weak solution follows from Theorem 6.6. Fix arbitrary \( \varphi \in C_c^\infty(\mathbb{R}^d) \) and \( t_0 \in (0, T] \), and take \( \tilde{b}(t) = b(t_0 - t) \) for every \( t \in [0, t_0] \). Replace \( b \) by \( \tilde{b} \) in problem (35). Then there exists a bounded classical solution \( \tilde{w} \) of problem (35) — by Theorem 7.8 if (i) or (ii) holds (see Remark 7.7 (b) in case (ii)) and by Theorem 7.9 if (iii) or (iv) holds.
Let \( w(t) = \tilde{w}(t_0 - t) \) for \( t \in [0, t_0] \). Then \( w \) is a bounded classical solution of
\begin{equation}
\begin{aligned}
\partial_t w(t) + b(t) L w(t) &= 0 \quad \text{in } (0, t_0) \times \mathbb{R}^d, \\
 w(t_0) &= \varphi.
\end{aligned}
\end{equation}

In particular, \( \partial_t w, Lw \in C((0, t_0) \times \mathbb{R}^d) \). Suppose \( m \) and \( \tilde{m} \) are two very weak solutions of problem (3) with the same initial condition \( m_0 \) and coefficient \( b \). By Definition 1.4 (see Lemma 6.1 (ii)) and (55),
\[
(m(t_0) - \tilde{m}(t_0))[\varphi] = \int_0^{t_0} (m(\tau) - \tilde{m}(\tau))[\partial_t w + b L w] \, d\tau = 0.
\]
Hence, for every \( t \in (0, T) \) and \( \varphi \in C_0^\infty(\mathbb{R}^d) \),
\[
(m(t) - \tilde{m}(t))[\varphi] = 0,
\]
which means that \( m(t) = \tilde{m}(t) \) in \( \mathcal{P}(\mathbb{R}^d) \).
\( \square \)

**Corollary 7.11.** Assume (A1), (A4). Condition (S5) is satisfied if any of the following sets of assumptions holds:
\begin{enumerate}[label=(\roman*)]
\item (RA), (LA), and \( \frac{2\sigma}{\alpha - 2\sigma} (1 + \frac{1}{1 - \sigma}) < \gamma \);
\item (RA), (LA) with \( \nu \) symmetric at the origin, and \( \frac{2\sigma}{\alpha - 2\sigma} (1 + \frac{1}{1 - \sigma}) < \gamma \);
\item (RB), (FB), (LB), (A2);
\item (RB), (FB), (LC), (A2), and Conjecture 2.7 is true.
\end{enumerate}

**Proof.** Let \( u_1, u_2 \in S_{RB} \) and \( v_1 = \mathcal{L} u_1, v_2 = \mathcal{L} u_2 \). Since \( F' \in C(\mathbb{R}) \) by (A1), we may consider
\[
b(t, x) = \int_0^1 F'(sv_1(t, x) + (1-s)v_2(t, x)) \, ds.
\]
Because \( u_1, u_2 \in S_{RB} \) and \( F' \geq 0 \), we have \( b \in C(\mathcal{T} \times \mathbb{R}^d) \) and \( b \geq 0 \).

\( \diamond \) **Part (i).** We have \( v_1, v_2 \in B(\mathcal{T}, C_0^{\alpha-2\sigma}(\mathbb{R}^d)) \), because of Theorem 5.4 (i). Thus \( b \) satisfies (A1") with \( \beta = \gamma (\alpha - 2\sigma) \). Since \( \frac{2\sigma}{\alpha - 2\sigma} (1 + \frac{1}{1 - \sigma}) < \gamma \), we have \( \beta > 2\sigma + \frac{2\sigma}{\alpha - 2\sigma} \) and (S5) follows from Theorem 7.10 (i).

\( \diamond \) **Part (ii).** We proceed as in Part (i) and use Theorem 7.10 (ii).

\( \diamond \) **Part (iii).** By Lemma 2.4 and Theorem 5.7 (iii), \( v_1, v_2 \in C_b(\mathcal{T} \times \mathbb{R}^d) \) and \( v_1, v_2 \in B([0, \hat{t}], C_0^\alpha(\mathbb{R}^d)) \cap UC([0, \hat{t}] \times \mathbb{R}^d) \) for every \( t \in \mathcal{T} \). Thus \( b \) satisfies (A1") on \( [0, \hat{t}] \) with \( \beta = \gamma \alpha \) and \( b \in UC([0, \hat{t}] \times \mathbb{R}^d) \). Since \( F' \geq \kappa > 0 \), we have \( b \geq \kappa > 0 \) and (S5) follows from Theorem 7.10 (iii).

\( \diamond \) **Part (iv).** By Conjecture 2.7 and Theorem 5.7 (iii), \( v_1, v_2 \in C_b(\mathcal{T} \times \mathbb{R}^d) \) and \( v_1, v_2 \in B([0, \hat{t}], C_0^\alpha(\mathbb{R}^d)) \) for every \( t \in \mathcal{T} \). Thus \( b \) satisfies (A1") on \( [0, \hat{t}] \) with \( \beta = \gamma \alpha \). Since \( F' \geq \kappa > 0 \), we have \( b \geq \kappa > 0 \) and (S5) follows from Theorem 7.10 (iv).

8. **Mean field game**

In this section we prove existence and uniqueness of solutions of problem (1) under general assumptions. These results yield a proof of Theorem 2.9. For the proof of existence, we need to recall some terminology concerning set-valued maps, in order to use the Kakutani–Glicksberg–Fan fixed point theorem.

**Definition 8.1.** A set-valued map \( \mathcal{K} : X \to 2^Y \) is compact if the image \( \mathcal{K}(X) = \bigcup \{ \mathcal{K}(x) : x \in X \} \) is contained in a compact subset of \( Y \).

**Definition 8.2.** A set-valued map \( \mathcal{K} : X \to 2^Y \) is upper-semicontinuous if, for each open set \( A \subseteq Y \), the set \( \mathcal{K}^{-1}(2^A) = \{ x : \mathcal{K}(x) \subseteq A \} \) is open.
Theorem 8.3 (Kakutani–Glicksberg–Fan [40, §7 Theorem 8.4]). Let $S$ be a convex subset of a normed space and $K : S \rightarrow 2^S$ be a compact set-valued map. If $K$ is upper-semicontinuous with non-empty compact convex values, then $K$ has a fixed point, i.e. there exists $x \in S$ such that $x \in K(x)$. □

In addition, the following lemma lets us express upper-semicontinuity in terms of sequences, which are easier to handle (cf. Lemma 6.4).

Lemma 8.4 ([40, §43.II Theorem 4]). Let $X$ be a Hausdorff space and $Y$ be a compact metric space. A set-valued compact map $K : X \rightarrow 2^Y$ is upper-semicontinuous if and only if the conditions

$$x_n \rightarrow x \text{ in } X, \quad y_n \rightarrow y \text{ in } Y,$$

implies $y \in K(x)$. □

Theorem 8.5. Assume (L), (A1), (A3), (A4), (S1), (S2), (S3). There exists a classical very weak solution of problem (1).

Proof. Let $X = \{ C(\overline{T}, M_0[\mathbb{R}^d]) \}$. We want to find a solution of problem (1) in $X$ by applying the Kakutani–Glicksberg–Fan fixed point theorem. To this end, we shall define a map $K : S \rightarrow 2^S$ on a certain compact, convex set $S \subset X$. Then the map $K$ is automatically compact and we may use Lemma 8.4 to obtain upper-semicontinuity.

- Step 1. Let $V$ be a Lyapunov function such that $m_0[V] \|LV\|_\infty < \infty$ (see Corollary 4.14). Define

$$S = \left\{ \mu \in C(\overline{T}, \mathcal{P}(\mathbb{R}^d)) : \mu(0) = m_0, \sup_{t \in \overline{T}} \mu(t)[V] \leq c_1, \sup_{0 \leq |t-s| \leq T} \frac{\|\mu(t) - \mu(s)\|_0}{\sqrt{|t-s|}} \leq c_2 \right\},$$

where $m_0$ is fixed and satisfies (A3), and

$$c_1 = m_0[V] + TK_{HJB}\|LV\|_\infty, \quad c_2 = 2 + (2\sqrt{T} + K_4)K_{HJB}\|L\|_{LK}.$$ The set $S$ is clearly convex. In addition, $S$ is compact because of Proposition 4.8, the assumed equicontinuity in time, and the Arzelà–Ascoli theorem.

- Step 2. Take $\mu \in S$ and let

$$f = f(\mu) \quad \text{and} \quad g = g(\mu(T)).$$

We define a map $K_1 : S \rightarrow C_0(\overline{T} \times \mathbb{R}^d)$ by $K_1(\mu) = u$, where $u$ is the unique bounded classical solution of problem (2), corresponding to data $(f, g)$. The map $K_1$ is well-defined because of (S1), (S3), and Theorem 5.3. By (A1) we find that $b = F'(Lu)$ satisfies (A1').

We define a set-valued map $K_2$ by $K_2(u) = M$, where $M$ is the set of very weak solutions of problem (3) corresponding to $b = F'(Lu)$. The set $M \subset C(\overline{T}, \mathcal{P}(\mathbb{R}^d))$ is convex, compact, and non-empty because of Corollary 6.3 and Theorem 6.6. Now we define the fixed point map

$$K(\mu) = K_2(K_1(\mu)) = M.$$ Because of its construction, $K : S \rightarrow 2^S$ is a compact map with non-empty compact convex values.

- Step 3. It remains to show that the map $K : S \rightarrow 2^S$ is upper-semicontinuous. Let $\{\mu_n, \mu\} \subset S$ be such that $\lim_{n \to \infty} \mu_n = \mu$ and let $\{\mu_n, u_n\} = \{K_1(\mu_n), K_1(\mu)\}$ be the corresponding solutions of problem (2), and $\{M_n, M\} = \{K(\mu_n), K(\mu)\}$ be the corresponding sets of solutions of problem (3).
Since \( \lim_{n \to \infty} \mu_n = \mu \), by (A4), Theorem 5.3, and (S2), we obtain \( \mathcal{L}u_n \to \mathcal{L}u \) uniformly on compact sets in \( \mathbb{R}^d \) for every \( t \in \mathcal{T} \). Hence, if we let
\[
b_n = F'(\mathcal{L}u_n) \quad \text{and} \quad b = F'(\mathcal{L}u),
\]
then, because of (A1), \( b_n \to b \) uniformly on compact sets in \( \mathbb{R}^d \) for every \( t \in \mathcal{T} \). Moreover, the functions \( b_n \) and \( b \) satisfy (A1') and are uniformly bounded, because of (S3).

Consider a sequence \( m_n \in \mathcal{M}_n \) and suppose it converges to some \( \hat{m} \in \mathcal{S} \). Then we use Lemma 6.4 to say that \( \hat{m} \in \mathcal{M} \). This proves that the map \( \mathcal{K} \) is upper-semicontinuous by Lemma 8.4.

\[\diamond \text{Step 4.}\] We now use Theorem 8.3 to get a fixed point \( \hat{m} \in \mathcal{S} \) of the map \( \mathcal{K} \).

Because of how \( \mathcal{K} \) is defined, we have
\[
\hat{m} \in \mathcal{K}(\hat{m}) = \mathcal{K}_2(\mathcal{K}_1(\hat{m})).
\]

Thus there exists \( \hat{u} = \mathcal{K}_1(\hat{m}) \), which is a bounded classical solution of problem (2) with \( f = f(\hat{m}) \) and \( g = g(\hat{m}(T)) \), and \( \|F'(\mathcal{L}\hat{u})\|_\infty \leq K_{HJB} \) by (S3). Note that \( \hat{m} \) is a very weak solution of problem (3) with \( \hat{m}(0) = m_0 \) and \( b = F'(\mathcal{L}\hat{u}) \). This, in turn, means that the pair \((\hat{u}, \hat{m})\) is a classical–very weak solution of problem (1) (see Definition 1.5).

\[\square\]

Remark 8.6. When we add assumption (S5) to Theorem 8.5,\(^\dagger\) we can say that the values of the maps \( \mathcal{K}_2 : \mathcal{S}_{HJB} \to 2^S \) and \( \mathcal{K} : \mathcal{S} \to 2^S \) are singletons, and hence both are continuous (see Step 3, Remark 6.5). Instead of the Kakutani–Glicksberg–Fan theorem, we may then use the Schauder theorem [40, §6 Theorem 3.2]. In fact, the former simply reduces to the latter in this setting (cf. Lemma 8.4).

Theorem 8.7. Assume (L), (A1), (A2), (A3), (A5), (S4), (S5). Then problem (1) has at most one solution.

Proof. Suppose \((u_1, m_1)\) and \((u_2, m_2)\) are classical–very weak solutions of problem (1) (see Definition 1.5), and take
\[
u = u_1 - u_2, \quad m = m_1 - m_2.
\]

To shorten the notation further, let \( \mathcal{L}u_1 = v_1, \mathcal{L}u_2 = v_2, \) and \( v = v_1 - v_2 \).

By Definition 1.5, \( u_1, u_2 \) are bounded classical solutions of problem (2), and by (S4), \( \{\partial_t u_1, \partial_t u_2, \mathcal{L}u_1, \mathcal{L}u_2\} \subseteq C_b(T \times \mathbb{R}^d) \). Thus, \( F'(v_1), F'(v_2) \in C_b(T \times \mathbb{R}^d) \) because of (A1) and \( u \in \mathcal{U} \) where \( \mathcal{U} \) is defined in Lemma 6.1 (ii). Further, \( m_1, m_2 \) are very weak solutions of problem (3), hence they satisfy (4) for every \( \phi \in \mathcal{U} \) by Lemma 6.1 (ii). Therefore,
\[
m(T)[u(T)] - m(0)[u(0)] = (m_1(T) - m_2(T))[u_1(T) - u_2(T)] - (m_1(0) - m_2(0))[u_1(0) - u_2(0)]
\]
\[
= \int_0^T \left( m_1[\partial_t u + F'(v_1)v] - m_2[\partial_t u + F'(v_2)v]\right)(\tau) \, d\tau.
\]

As \( m_1(0) = m_2(0) = m_0 \), we have
\[
m(0)[u(0)] = (m_1(0) - m_2(0))[u(0)] = 0
\]
and, thanks to (A5),
\[
m(T)[u(T)] = (m_1(T) - m_2(T))[g(m_1(T)) - g(m_2(T))] \leq 0.
\]

Hence by (56) we get
\[
\int_0^T \left( m_1[\partial_t u + F'(v_1)v] - m_2[\partial_t u + F'(v_2)v]\right)(\tau) \, d\tau \leq 0.
\]

\[\dagger\]Here it would be sufficient to consider (S5) for a smaller set \( \mathcal{B} \) constructed with \( u_1 = u_2 \).
We further notice that
\[ \partial_t u + F(v_1) - F(v_2) = f(m_2) - f(m_1). \]

Then, by integrating this expression with respect to the measure \( m \), we obtain
\[ \int_0^T (m_1 - m_2)[\partial_t u + F(v_1) - F(v_2)](\tau) \, d\tau = \int_0^T (m_1 - m_2)[f(m_2) - f(m_1)] \, d\tau. \tag{58} \]

From (A2) we know that \( F \) is convex, thus
\[ F(v_1) - F(v_2) \leq F'(v_1) v \quad \text{and} \quad F(v_1) - F(v_2) \geq F'(v_2) v, \]
and since \( m_1, m_2 \in C(\overline{T}, P(\mathbb{R}^d)) \) are non-negative measures, by (58), (59) and (A5),
\[ \int_0^T m_1[\partial_t u + F'(v_1) v](\tau) \, d\tau - \int_0^T m_2[\partial_t u + F'(v_2) v](\tau) \, d\tau \]
\[ \geq \int_0^T (m_1 - m_2)[f(m_2) - f(m_1)](\tau) \, d\tau \geq 0. \tag{60} \]

Combining (57) and (60), we find that
\[ \int_0^T m_1[\partial_t u + F'(v_1) v](\tau) \, d\tau - \int_0^T m_2[\partial_t u + F'(v_2) v](\tau) \, d\tau \]
\[ = \int_0^T (m_1 - m_2)[f(m_2) - f(m_1)](\tau) \, d\tau = 0. \]

Then taking into account (58), we find that
\[ 0 = \int_0^T \int_{\mathbb{R}^d} (F'(v_1) v - F(v_1) + F(v_2)) \, m_1(\tau, dx) \, d\tau \]
\[ + \int_0^T \int_{\mathbb{R}^d} (F(v_1) - F(v_2) - F'(v_2) v) \, m_2(\tau, dx) \, d\tau. \]

By (57), both functions under the integrals are non-negative and continuous, thus
\[ F(v_1) - F(v_2) - F'(v_1)(v_1 - v_2) = 0 \quad \text{on} \quad \text{supp} \, m_1, \tag{61} \]
\[ F(v_1) - F(v_2) - F'(v_2)(v_1 - v_2) = 0 \quad \text{on} \quad \text{supp} \, m_2, \]
where by \( \text{supp} \, m_i \), we understand the support of \( m_i \) taken as a measure on \( T \times \mathbb{R}^d \).
Let us define
\[ b(t, x) = \begin{cases} \frac{F(v_1(t, x)) - F(v_2(t, x))}{v_1(t, x) - v_2(t, x)}, & \text{if } v_1(t, x) \neq v_2(t, x), \\ F'(v_1(t, x)), & \text{if } v_1(t, x) = v_2(t, x). \end{cases} \tag{62} \]

Because of (A1), we may also write (62) as
\[ b(t, x) = \int_0^1 F'(sv_1(t, x) + (1 - s)v_2(t, x)) \, ds. \tag{63} \]

Because of (S4), \( b \in C_b(T \times \mathbb{R}^d) \). Notice that if \( v_1 \neq v_2 \), the following identities are equivalent
\[ F(v_1) - F(v_2) - F'(v_1)(v_1 - v_2) = 0 \quad \Leftrightarrow \quad \frac{F(v_1) - F(v_2)}{v_1 - v_2} = F'(v_1). \]

We can make a similar observation for \( F'(v_2) \). Thus by (61),
\[ b(t, x) = F'(v_1)(t, x) \quad \text{when} \ (t, x) \in \text{supp} \, m_1, \]
\[ b(t, x) = F'(v_2)(t, x) \quad \text{when} \ (t, x) \in \text{supp} \, m_2. \]
Since \((u_1, m_1)\) and \((u_2, m_2)\) are classical–very weak solutions of problem (1), it now follows that both \(m_1\) and \(m_2\) are very weak solutions of problem (3) with initial condition \(m_0\) and coefficient \(b\) given by (62). By (63) and (S5) we get \(m_1 = m_2\). Then also \(u_1 = u_2\) by Theorem 5.3.

**Appendix A. The fractional Laplacian**

**Definition A.1.** The fractional Laplacians are Lévy operators given by

\[-(\Delta)^\sigma u(x) = c_{d,\sigma} \text{ p.v.} \int_{\mathbb{R}^d} \frac{u(x + z) - u(x)}{|z|^{d+2\sigma}} \, dz,
\]

where p.v. denotes the principal value, \(\sigma \in (0, 1)\), and

\[c_{d,\sigma} = \sigma(1 - \sigma) \frac{2^{2\sigma+1}\Gamma(\sigma)}{K_d B(\frac{d}{2}, \sigma)\Gamma(2 - \sigma)}, \quad \text{and } B \text{ is the beta function.}
\]

**Proposition A.2.** The fractional Laplacians satisfy (LC) and, for \(2\sigma \in (0, 1)\), also satisfy (LA).

**Proof.** The Lévy measure of a fractional Laplacian is \(c_{d,\sigma}|z|^{-d-2\sigma} \, dz\). In fact,

\[
\int (1 + |z|^2) \frac{1}{|z|^{d+2\sigma}} \, dz = \int_{B_1} \frac{1}{|z|^{d-2\sigma}} \, dz + \int_{B_1} \frac{1}{|z|^{d+2\sigma}} \, dz = \frac{K_d}{\sigma(1 - \sigma)}.
\]

Since this measure is also symmetric at the origin (see Definition 4.11 and Remark 4.12 (a)), the fractional Laplacian satisfies (LC). Suppose \(2\sigma \in (0, 1)\). For every \(r \in (0, 1)\) and \(p \in (2\sigma, 1]\) we have

\[
\int_{B_1} (1 + |z|^p) \frac{1}{|z|^{d+2\sigma}} \, dz = r^{-p} \int_{B_r} \frac{1}{|z|^{d-p+2\sigma}} \, dz + \int_{B_1 \setminus B_r} \frac{1}{|z|^{d+2\sigma}} \, dz
\]

\[
= K_d \left( \frac{r^{-2\sigma}}{p - 2\sigma} + \frac{r^{-2\sigma} - 1}{2\sigma} \right) < \frac{K_d}{\sigma(p - 2\sigma)r^{-2\sigma}}.
\]

This shows that the fractional Laplacian satisfies (LA) when \(2\sigma \in (0, 1)\). □

**Remark A.3.** A similar proof can be given for an anisotropic and/or subelliptic operator \(-\sum_{i=1}^d c_i(\Delta x_i)^{\sigma_i}\), which is a sum of one-dimensional fractional Laplacians taken with (possibly) different values \(\sigma_i\) in each coordinate direction and weights \(c_i \geq 0\). The Lévy measure is then concentrated entirely on the axes — it is not absolutely continuous with respect to the Lebesgue measure on \(\mathbb{R}^d\) when \(d \geq 2\) and if some \(c_i = 0\), not on all of the axes.

Other examples in the same spirit would be the CGMY operator on \(\mathbb{R}\) with density \(\nu(dz) = k(z) \, dz = C \frac{1}{|z|^{2\sigma}} \left( e^{-G|z|^\alpha} I_{\{z < 0\}} + e^{-M|z|^\alpha} I_{\{z > 0\}} \right) \, dz\) and other non-local operators used in finance (cf. e.g. [28]), one-sided operators with the same asymptotic behaviour as the fractional Laplacian (e.g. Riesz–Feller operators), any convex combinations of such, and “embdeddings” into higher dimensions.

**Proposition A.4.** Let \(V(x) = \log \left( \sqrt{1 + |x|^2} + 1 \right)\). Then \(V\) is a Lyapunov function and \(\|(-\Delta)^\sigma V\|_\infty \leq \frac{c_{d,\sigma} K_d}{\sigma} \left( \log(2) + \frac{\pi}{2\sin(\sigma\pi)} \right)\) for every \(\sigma \in (0, 1)\).

**Proof.** We have \(V \in C^2(\mathbb{R}^d)\) and \(V(x) = V_0(\sqrt{1 + |x|^2})\), where \(V_0(x) = \log(x + 1)\). Note that \(V_0 : [0, \infty) \to [0, \infty)\) is concave, and hence subadditive. Moreover, \(0 \leq -V_0'' \leq 1\) and \(\lim_{x \to \infty} V_0(x) = \infty\). Hence \(V\) is a Lyapunov function.

We have \(|V(x + z) - V(x)| \leq V(z)\) (see the proof of Lemma 4.13) and

\[V(z) = \log \left( \sqrt{1 + |z|^2} + 1 \right) \leq \log(1 + |z|^2) + \log(2).
\]
By using the Taylor expansion and the fact that $(|z|^2 + 1)^{-1} \geq \frac{1}{2}$ for $|z| \leq 1$ we thus get

$$\left| \frac{1}{c_{d, \sigma}} (-\Delta)^{\sigma} V(x) \right| \leq \text{p.v.} \int_{\mathbb{R}^d} \frac{V(x+z) - V(x)}{|z|^{d+2\sigma}} \, dz$$

$$\leq \int_{B_1} \frac{|z|^2}{(|z|^2 + 1)|z|^{d+2\sigma}} \, dz + \int_{B_1^c} \frac{\log(|z|^2 + 1)}{|z|^{d+2\sigma}} \, dz + \frac{K_d}{2\sigma} \log(2).$$

Using polar coordinates, substitution, and integration by parts we obtain

$$\int_{B_1} \frac{|z|^2}{(|z|^2 + 1)|z|^{d+2\sigma}} \, dz + \int_{B_1^c} \frac{\log(|z|^2 + 1)}{|z|^{d+2\sigma}} \, dz = \frac{K_d}{2\sigma} \left( \sigma \int_0^1 \frac{t^{-\sigma}}{t+1} \, dt + \int_1^{\infty} \frac{t^{-\sigma}}{t+1} \, dt + \log(2) \right).$$

We have $\sigma \in (0, 1)$ and by using formulas 6.2.1, 6.2.2, and 6.1.17 in [34], we calculate

$$\int_0^{\infty} \frac{t^{-\sigma}}{t+1} \, dt = B(1-\sigma, \sigma) = \frac{\pi}{\sin(\pi \sigma)},$$

where $B$ is the beta function. Hence

$$\|(-\Delta)^{\sigma} V\|_{\infty} \leq \frac{c_{d, \sigma} K_d}{2\sigma} \left( \log(4) + \frac{\pi}{\sin(\pi \sigma)} \right). \quad \Box$$

**Remark A.5.** Notice that $\lim_{\sigma \to 1} \frac{\pi(1-\sigma)}{\sin(\pi \sigma)} = 1$ and $\sup_{\sigma \in (0, 1)} \frac{2^\sigma \Gamma(\sigma)}{B(\frac{1}{2}, \sigma) \Gamma(2-\sigma)} < \infty$, thus the Lyapunov function $V$ in Proposition A.4 satisfies

$$\sup_{\sigma \in (0, 1)} \|(-\Delta)^{\sigma} V\|_{\infty} \leq \infty \quad \text{for every } \varepsilon \in (0, 1).$$

It is impossible to put $\varepsilon = 0$, since the relevant family of Lévy measures restricted to $B_1^c$ is not tight (see Proposition 4.8).

**Appendix B. The Legendre–Fenchel Transform**

For a comprehensive treatment of the Legendre–Fenchel transform we refer to [76, 43]. Below we gather the properties needed to derive the model in Section 3.

**Proposition B.1.** Let $L : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ be a lower-semicontinuous function such that $L \not\equiv \infty$ and define $F(z) = \sup_{\zeta \in [0, \infty]} (z \zeta - L(\zeta))$. Then $F$ is convex and non-decreasing. In addition,

(i) if $\lim_{\zeta \to \infty} L(\zeta)/\zeta = \infty$, then $F$ has finite values and is locally Lipschitz-continuous;

(ii) if $L$ is convex and is strictly convex on $\{L \not\equiv \infty\}$, then $F$ is differentiable on $\{F \not\equiv \infty\}$ and $\zeta \mapsto z \zeta - L(\zeta)$ achieves its supremum at $\zeta = F'(z)$;

(iii) let $L$ be convex, $\lim_{\zeta \to \infty} L(\zeta)/\zeta = \infty$ and $\partial L$ be the subdifferential of $L$. If for every $\zeta_1, \zeta_2 \in [0, \infty)$ and $z_1 \in \partial L(\zeta_1)$ there exists $c_{z_1} > 0$ such that for every $z_2 \in \partial L(\zeta_2)$ satisfying $|z_1 - z_2| \leq 1$ we have

$$(z_1 - z_2)(\zeta_1 - \zeta_2) \geq c_{z_1} |\zeta_1 - \zeta_2|^{1+\frac{1}{2}},$$

then $F' \in C^1(\mathbb{R})$.

---

\[19\] When $\gamma = 1$ and $c_{z_1}$ is in fact independent of $z_1$, this corresponds to the usual strong convexity of $L$ (see [43, Theorem D.6.1.2]); if $\gamma_1 < \gamma_2$, the condition with $\gamma_1$ allows for a flatter (less non-affine) function $L$ than the one with $\gamma_2$. 

---
Proof. The function \( F \) is convex as a supremum of convex (affine) functions. For \( \zeta, h \geq 0 \) and \( z \in \mathbb{R} \) we have \( (z + h)\zeta - L(\zeta) \geq z\zeta - L(\zeta) \) and thus
\[
F(z + h) = \sup_{\zeta \in [0, \infty)} ((z + h)\zeta - L(\zeta)) \geq \sup_{\zeta \in [0, \infty)} (z\zeta - L(\zeta)) = F(z).
\]

\( \diamond \) Part (i). Because \( \lim_{\zeta \to \infty} L(\zeta)/\zeta = \infty \), for every \( z \in \mathbb{R} \),
\[
\lim_{\zeta \to \infty} \left( z - \frac{L(\zeta)}{\zeta} \right) \zeta = -\infty.
\]
Since \( L \) is lower-semicontinuous and \( L \neq \infty \), there exists \( \zeta_0 < \infty \) such that
\[
L(\zeta_0) < \infty \quad \text{and} \quad \sup_{\zeta \in [0, \infty)} \left( z - \frac{L(\zeta)}{\zeta} \right) = z\zeta_0 - L(\zeta_0).
\]
As a convex function with finite values, \( F \) is then locally Lipschitz-continuous.

\( \diamond \) Part (ii). Since \( L \) is lower-semicontinuous, the statement follows from [76, Theorem 23.5, Corollary 23.5.1, Theorem 26.3, page 52].

\( \diamond \) Part (iii). Note that (iii) implies (i) and (ii) (cf. [43, Theorem D.6.1.2]) hence \( F \) has finite values on \( \mathbb{R} \) and \( F' \) exists everywhere. If \( z_i \in \partial L(\zeta_i) \), then \( \zeta_i = F'(z_i) \) by [76, Theorem 23.5]. For \( |z_1 - z_2| \leq 1 \) we thus have
\[
|z_1 - z_2|\frac{F'(z_1) - F'(z_2)}{\zeta} \geq c_{z_1}|F'(z_1) - F'(z_2)|^{1 + \frac{1}{q}}.
\]
which gives us \( F' \in C^1(\mathbb{R}) \) (see (17) in Definition 4.1). \( \square \)

| \( L : [0, \infty) \to \mathbb{R} \cup \{\infty\} \) | \( F : \mathbb{R} \to \mathbb{R} \) |
|---|---|
| (a) | \( \chi_{(\zeta)} \) | \( \kappa z \) |
| (b) | \( \chi_{[0,\zeta]} \) | \( \kappa z^+ \) |
| (c) | \( \chi_{[0,\zeta]} + \varepsilon \left( \frac{1}{\zeta} z^2 - \zeta \right) \) | \( \Pi_{[-\varepsilon,\varepsilon]}(z) \frac{1}{\zeta^2} (z + \varepsilon)^2 + \Pi_{[\varepsilon,\infty]}(z) \kappa z \) |
| (d) | \( \frac{1}{q} q^q \) | \( \frac{2 - 1}{q} (z^+) \frac{1}{q+1} \) |
| (e) | \( \zeta \log(\zeta) - \zeta \) | \( e^z \) |
| (f) | \( \chi_{[\zeta, \infty]}(\zeta) + 1 \) \( L_0(\zeta) - \kappa \) | \( F_0(z) + \kappa z \) |

Table 1. Pairs of Legendre–Fenchel conjugate functions (see Remark B.2). Here \( \chi_A(x) = \infty \) for \( x \not\in A \), \( \chi_A(x) = 0 \) for \( x \in A \); and \( z^+ = \max(z, 0) \); \( L_0, F_0 \) is an arbitrary conjugate pair.

Remark B.2 (On Table 1). For convenience we provide some prototypical examples of gain functionals \( L \) (see Section 3) that can (or cannot) be used to derive problem (1), and their corresponding Legendre–Fenchel transforms, i.e. Hamiltonians.

(a) Players are forced to always choose the same control, \( \kappa \). Mean field game reduces to a pair of linear heat equations.

(b) Players can choose a control between 0 and \( \kappa \), for a constant (zero) gain. Hamiltonian \( F \) is not differentiable, (A1) fails.

(c) Players can choose a control between 0 and \( \kappa \), and the gain \( L \) is strongly convex on \([0, \kappa]\). Hamiltonian \( F \in C^1(\mathbb{R}) \), with \( F' \) Lipschitz-continuous, is sufficiently regular, while not strictly convex (note how this can be used to approximate the previous case). Assumptions (A1), (A2) are satisfied, but not (FB).

(d) The “standard” linear-quadratic control adapted to the fractional Laplacian (see Section 3.5).

(e) A gain functional resulting in \( F \in C^\infty(\mathbb{R}) \); \( F > 0 \), but (FB) is not satisfied.

(f) This template can be used to modify any conjugate pair \( (L_0, F_0) \) in order to adjust the lower bound on the derivative, e.g. to satisfy condition (FB).
Consider an operator
\[
P(v_1, v_2) = \partial_t (v_1 - v_2) + F(\mathcal{L}v_1) - F(\mathcal{L}v_2)
\]
and a constant
\[
A = \sup \left\{ F'(z) : -\infty < z < \sup_{x \in \mathbb{R}^d} \mathcal{L}u_1(x) + \| \mathcal{L}V \|_\infty \right\} \| \mathcal{L}V \|_\infty + 1,
\]
where \( \mathcal{L} \) is a Lyapunov function such that \( \| \mathcal{L}V \|_\infty < \infty \) (see Corollary 4.14). Note that \( \| F'(\mathcal{L}u_1) \|_\infty \leq K \) means that either \( F \) is globally Lipschitz or \( \mathcal{L}u_1 \) is bounded from above (since \( F' \geq 0 \) by (A1)). In both cases, \( A \in [0, \infty) \).

Let
\[
u_1^\varepsilon = u_1 - \varepsilon V + (t - T)(\| f_1 - f_2 \|_\infty + \varepsilon A) - \frac{\varepsilon}{2} t
\]
for \( \varepsilon \in (0, 1) \).

Since \( P(u_1, u_2) = f_2 - f_1 \), for every \( (t, x) \in \mathcal{T} \times \mathbb{R}^d \) by (65) we have
\[
P(u_1^\varepsilon, u_2) = \partial_t (u_1 - u_2) + F(\mathcal{L}u_1) - F(\mathcal{L}u_2) + \| f_1 - f_2 \|_\infty + \varepsilon A + \frac{\varepsilon}{2t}
\]
\[
+ F(\mathcal{L}(u_1 - \varepsilon V)) - F(\mathcal{L}u_1)
\]
\[
\geq \varepsilon \left( A + \frac{1}{2t} - \mathcal{L}V \int_0^1 F'(\mathcal{L}(u_1 - \varepsilon \tau V)) d\tau \right) > 0.
\]

Let \( u^\varepsilon = u_1^\varepsilon - u_2 \). Since \( u_1, u_2 \) are bounded, for \( t \to 0 \) or \( |x| \to \infty \) we find \( u^\varepsilon \to -\infty \), and \( u^\varepsilon \) is clearly bounded from above. Therefore, \( u^\varepsilon \) attains a maximum within \( (\mathcal{T} \cup \{ T \}) \times \mathbb{R}^d \) for each \( \varepsilon > 0 \).

Let \( (t_\varepsilon, x_\varepsilon) \) be a maximum point of \( u^\varepsilon \). If \( (t_\varepsilon, x_\varepsilon) \in \mathcal{T} \times \mathbb{R}^d \), then \( \partial_t u^\varepsilon = 0 \) and, because of (A1), the fundamental theorem of calculus, and the maximum principle (64),
\[
P(u_1^\varepsilon, u_2)(t_\varepsilon, x_\varepsilon) = F(\mathcal{L}u_1^\varepsilon(t_\varepsilon, x_\varepsilon)) - F(\mathcal{L}u_2(t_\varepsilon, x_\varepsilon))
\]
\[
= \mathcal{L}u^\varepsilon(t_\varepsilon, x_\varepsilon) \int_0^1 F'(\mathcal{L}(\tau u_1^\varepsilon + (1 - \tau)u_2)) d\tau \leq 0.
\]
This contradicts (67) and shows that for every \( \varepsilon > 0 \), the maximum values of \( u^\varepsilon \) are only attained within the set \( \{ T \} \times \mathbb{R}^d \). Thus for every \( (t, x) \in \mathcal{T} \times \mathbb{R}^d \) and \( \varepsilon > 0 \)
\[
(u_1^\varepsilon - u_2)(t, x) \leq (u_1^\varepsilon - u_2)(T, x) = (g_1 - g_2)(x) - \varepsilon V(x) - \frac{\varepsilon}{T}.
\]
By (66) we get
\[
(u_1 - u_2)(t, x) \leq \varepsilon \left( \frac{1}{t} + TA \right) + (T - t)\| f_1 - f_2 \|_\infty + \| g_1 - g_2 \|_\infty,
\]
and therefore
\[
u_1 - u_2 \leq (T - t)\| f_1 - f_2 \|_\infty + \| g_1 - g_2 \|_\infty.
\]
A symmetric argument then completes the proof. \( \square \)
Appendix D. Viscosity solutions of dual of Fokker–Planck equation

Proposition D.1. Assume (LA). Let \( L_r \) be given by (26) and \( \psi_{\varepsilon, \delta} \) by (36). For every \( r \in (0, 1), \varepsilon, \delta > 0 \) and \( x, y \in \mathbb{R}^d \) we have

\[
(L_r \psi_{\varepsilon, \delta}(x, \cdot))(y) \leq \frac{K}{1 - 2\sigma} \left( \delta + \frac{2|x - y| + r}{\varepsilon} \right)^{1 - 2\sigma}.
\]

If the Lévy measure \( \nu \) is symmetric at the origin (see Definition 4.11), then

\[
(L_r \psi_{\varepsilon, \delta}(x, \cdot))(y) \leq \frac{K}{1 - 2\sigma} \left( \frac{r}{\varepsilon} \right)^{1 - 2\sigma}.
\]

Proof. Notice that because of (LA) and the Cauchy–Schwarz inequality,

\[
(L_r \psi_{\varepsilon, \delta}(x, \cdot))(y) = \int_{|z| \leq r} \frac{|x - y|^2 - |x - y|^2}{\varepsilon} \nu(dz) + \delta L_r V(y)
\]

\[
\leq \int_{|z| \leq r} \left( \frac{|z|^2 + 2(x - y) \cdot z}{\varepsilon} + \delta ||\nabla V||_{\infty} |z| \right) \nu(dz)
\]

\[
\leq \frac{K}{1 - 2\sigma} \left( \frac{r}{\varepsilon} \right)^{1 - 2\sigma}.
\]

If \( \nu \) is symmetric at the origin then \( \int_{|z| \leq r} ((x - y) \cdot z) \nu(dz) = 0 \). The result follows. \( \square \)

Proof of Lemma 7.2. Notice that

\[
\lim_{|x|, |y| \to \infty} \Psi_{\eta, \varepsilon, \delta}(t, x, y, s) = -\infty \quad \text{for every } (t, s) \in \mathcal{T} \times \mathcal{T}
\]

and hence, because \( \Phi \) is bounded and upper-semicontinuous, the function \( \Psi_{\eta, \varepsilon, \delta} \) reaches a maximum at some point \((t_*, x_*, s_*, y_*)\), which depends on \( \eta, \varepsilon, \) and \( \delta \). Moreover, for each \( \delta > 0 \) there exists a compact set \( \Omega_\delta \) such that for every \( \eta, \varepsilon > 0 \) and \( (t_*, x_*, s_*, y_*) \in \Omega_\delta \)

We may also write

\[
\Phi(t_*, x_*, t, s_* + s) + \Phi(s_*, y_*, s_*, y_*) - 2\delta(V(x_*) + V(y_*)) - a_{\varepsilon, \delta}(t_* + s_*
\]

\[
= \Psi_{\eta, \varepsilon, \delta}(t_*, x_*, t_*, x_*) + \Psi_{\eta, \varepsilon, \delta}(s_*, y_*, s_*, y_*) \leq 2\Psi_{\eta, \varepsilon, \delta}(t_*, x_*, s_*, y_*)
\]

from which (38) follows:

\[
\frac{|x_* - y_*|^2}{\varepsilon} + \frac{|t_* - s_*|^2}{\eta}
\]

\[
\leq \Phi(t_*, x_*, s_*, y_*) - \Phi(t_*, x_*, t_*, x_*) + \Phi(s_*, y_*, s_*, y_*)
\]

It then implies

\[
|t_* - s_*| \leq \sqrt{2\eta \|\Phi\|_{\infty}}, \quad |x_* - y_*| \leq \sqrt{2\varepsilon \|\Phi\|_{\infty}}.
\]

Recall that \((t_*, x_*, s_*, y_*) \in \Omega_\delta\), which is a compact set. Thus for every \( \varepsilon, \delta > 0 \) there exists a subsequence \( \eta_k \) such that

\[
\lim_{\eta_k \to 0} (t_*^k, x_*^k, s_*^k, y_*^k) = (t_\varepsilon, x_\varepsilon, s_\varepsilon, y_\varepsilon),
\]

where \( \lim_{\eta_k \to 0} t_*^k = \lim_{\varepsilon \to 0} t_*^k \) follows from the first part of (68). By selecting another subsequence and using the second part of (68), for every \( \delta > 0 \) we get

\[
\lim_{\varepsilon \to 0} (t_\varepsilon, x_\varepsilon, s_\varepsilon, y_\varepsilon) = (t_\delta, x_\delta, s_\delta, y_\delta).
\]

Let

\[
\xi_{\varepsilon, \delta} = \sup_{(t, x) \in \mathcal{T} \times \mathbb{R}^d} \left( \Phi(t, x, t, s) - 2\delta V(x) - a_{\varepsilon, \delta}t \right)
\]
Then for every \((t, x) \in \mathcal{T} \times \mathbb{R}^d\) we have
\[
\Phi(t, x, t, x) - 2\delta V(x) \leq \liminf_{\varepsilon_n \to 0} \left( a_{\varepsilon_n, \delta} t + \xi_{\varepsilon_n, \delta} \right).
\]

Because \((t_*, x_*, s_*, y_*)\) is a maximum point of \(\Psi_{\eta, \varepsilon, \delta}\), we also have
\[
\xi_{\varepsilon, \delta} \leq \Psi_{\eta, \varepsilon, \delta}(t_*, x_*, s_*, y_*) \quad \text{for every } \eta > 0.
\]
Hence, by rearranging the terms, we obtain
\[
0 \leq \frac{|x_* - y_*|^2}{\varepsilon_n} + \frac{|t_* - s_*|^2}{\eta_k}
\leq \Phi(t_*, x_*, s_*, y_*) - \delta(V(x_*) + V(y_*)) - a_{\varepsilon_n, \delta} \frac{t_* + s_*}{2} - \xi_{\varepsilon_n, \delta}.
\]
Now we use the upper-semicontinuity of \(\Phi\) and continuity of \(V\) to pass to the upper-limits
\[
\limsup_{\varepsilon_n \to 0} \limsup_{\eta_k \to 0} \left( \Phi(t_*, x_*, s_*, y_*) - \delta(V(x_*) + V(y_*)) - a_{\varepsilon_n, \delta} \frac{t_* + s_*}{2} - \xi_{\varepsilon_n, \delta} \right)
\leq \Phi(\delta, x_*, t_*, s_*) - 2\delta V(x_*) - \liminf_{\varepsilon_n \to 0} \left( a_{\varepsilon_n, \delta}(t_{\varepsilon_n, \delta} - t_*) + a_{\varepsilon_n, \delta} s_* + \xi_{\varepsilon_n, \delta} \right).
\]
By (69) and because \(\{a_{\varepsilon_n, \delta}\}\) is a bounded sequence, we get
\[
0 \leq \limsup_{\varepsilon_n \to 0} \limsup_{\eta_k \to 0} \left( \frac{|x_* - y_*|^2}{\varepsilon_n} + \frac{|t_* - s_*|^2}{\eta_k} \right) \leq 0
\]
and then immediately
\[
\lim_{\varepsilon_n \to 0} \lim_{\eta_k \to 0} \frac{|x_* - y_*|^2}{\varepsilon_n} = 0 \quad \text{and} \quad \lim_{\eta_k \to 0} \frac{|t_* - s_*|^2}{\eta_k} = 0.
\]

Acknowledgements

IC and ERJ were supported by the Toppforsk (research excellence) project Waves and Nonlinear Phenomena (WaNP), grant no. 250070 from the Research Council of Norway. MK was supported by the Polish NCN grant 2016/23/B/ST1/00434. The main part of the research behind this paper was conducted when IC and MK were fellows of the ERCIM Alain Bensoussan Programme at NTNU.

References

[1] Y. Achdou, P. Cardaliaguet, F. Delarue, A. Porretta, and F. Santambrogio, Mean field games, CIME Lecture Notes in Mathematics, Springer, 2020.
[2] P. D. S. Andrade and E. A. Pimentel, Stationary fully nonlinear mean-field games, J. Anal. Math., 145 (2021), pp. 335–356.
[3] O. E. Barndorff-Nielsen and J. Schmiegel, Time change, volatility, and turbulence, in Mathematical control theory and finance, Springer, 2008, pp. 29–53.
[4] O. E. Barndorff-Nielsen and A. Shiryaev, Change of time and change of measure, World Scientific Publishing, 2nd ed., 2015.
[5] C. Benazzoli, L. Campi, and L. Di Persio, \(\varepsilon\)-Nash equilibrium in stochastic differential games with mean-field interaction and controlled jumps, Statist. Probab. Lett., 154 (2019), pp. 108522, 8.
[6] ———, Mean field games with controlled jump-diffusion dynamics: existence results and an illiquid interbank market model, Stochastic Process. Appl., 130 (2020), pp. 6927–6964.
[7] A. Bensoussan, J. Frehse, and P. Yam, Mean field games and mean field type control theory, Springer, 2013.
[8] I. H. Biswas, On zero-sum stochastic differential games with jump-diffusion driven state: a viscosity solution framework, SIAM J. Control Optim., 50 (2012), pp. 1825–1858.
[9] I. H. Biswas, E. R. Jakobsen, and K. H. Karlsen, Viscosity solutions for a system of integro-PDEs and connections to optimal switching and control of jump-diffusion processes, Appl. Math. Optim., 62 (2010), pp. 47–80.
[10] V. I. Bogachev, Measure theory, Vol. I, II, Springer, 2007.
[11] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, Fokker–Planck–Kolmogorov equations, American Mathematical Society, 2015.
[12] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, Uniqueness problems for degenerate Fokker–Planck–Kolmogorov equations, J. Math. Sci. (N.Y.), 207 (2015), pp. 147–165.
[13] V. S. Borkar, Controlled diffusion processes, Probab. Surv., 2 (2005), pp. 213–244.
[14] B. Bouddig, R. Schilling, and J. Wang, Lévy matters. III, Springer, 2013.
[15] L. Caffarelli and L. Silvestre, The Evans–Krylov theorem for nonlocal fully nonlinear equations, Ann. of Math. (2), 174 (2011), pp. 1163–1187.
[16] F. Camilli and R. De Maio, A time-fractional mean field game, Adv. Differential Equations, 24 (2019), pp. 531–554.
[17] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, The master equation and the convergence problem in mean field games, Princeton University Press, 2019.
[18] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon, Second order mean field games with degenerate diffusion and local coupling, NoDEA Nonlinear Differential Equations Appl., 22 (2015), pp. 1287–1317.
[19] R. Carmona and F. Delarue, Probabilistic theory of mean field games with applications. I, Springer, 2018.
[20] ———, Probabilistic theory of mean field games with applications. II, Springer, 2018.
[21] P. Carr and L. Wu, Time-changed Lévy processes and option pricing, Journal of Financial Economics, 71 (2004), pp. 113 – 141.
[22] A. Cesaroni, M. Cirant, S. Dipierro, M. Novaga, and E. Valdinoci, On stationary fractional mean field games, J. Math. Pures Appl. (9), 122 (2019), pp. 1–22.
[23] H. Chang–Lara and G. Davila, Regularity for solutions of non local parabolic equations, Calc. Var. Partial Differential Equations, 49 (2014), pp. 139–172.
[24] ———, Regularity for solutions of nonlocal parabolic equations. II, J. Differential Equations, 256 (2014), pp. 130–156.
[25] E. Chasseigne and E. R. Jakobsen, On nonlocal quasilinear equations and their local limits, J. Differential Equations, 262 (2017), pp. 3759–3804.
[26] I. Chowdhury, O. Esler, and E. R. Jakobsen, On Numerical approximations of fractional and nonlocal Mean Field Games, arXiv e-prints, arXiv:2105.00073 (2021).
[27] M. Cirant and A. Goffi, On the existence and uniqueness of solutions to time-dependent fractional MFG, SIAM J. Math. Anal., 51 (2019), pp. 913–954.
[28] R. Cont and P. Tankov, Financial modelling with jump processes, Chapman & Hall/CRC, 2004.
[29] J. C. Correia and E. A. Pimentel, A Hessian-dependent functional with free boundaries and applications to mean-field games, arXiv e-prints, (2021), p. arXiv:2107.00743.
[30] M. G. Chandalli, H. Ishii, and P.-L. Lions, User’s guide to viscosity solutions of second order partial differential equations, Bull. Amer. Math. Soc. (N.S.), 27 (1992), pp. 1–67.
[31] P. J. Davis, Gamma function and related functions, in Handbook of Mathematical Functions, M. Abramowitz and I. A. Stegun, eds., U.S. GPO, 10th ed., 1972, ch. 6, pp. 253–293.
[32] H. Dong, T. Jin, and H. Zhang, Dini and Schauder estimates for nonlocal fully nonlinear parabolic equations with drifts, Anal. PDE, 11 (2018), pp. 1487–1534.
[33] O. Esler and E. R. Jakobsen, On fractional and nonlocal parabolic mean field games in the whole space, J. Differential Equations, 301 (2021), pp. 428–470.
[34] L. C. Evans, Partial differential equations, American Mathematical Society, 2nd ed., 2010.
[35] A. Figalli, Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, J. Funct. Anal., 254 (2008), pp. 109–153.
[36] N. Fournier and L. Xu, On the equivalence between some jumping SDEs with rough or degenerate coefficients and some non-local PDEs, Ann. Inst. Henri Poincaré Probab. Stat., 55 (2019), pp. 1163–1178.
[37] I. I. Gikhman and A. V. Skorokhod, The theory of stochastic processes. II, Springer, 2004. Reprint of the 1975 edition.
[38] D. A. Gomes, E. A. Pimentel, and V. Voskanjan, Regularity theory for mean-field game systems, Springer, 2016.
[39] R. Gong, C. Mou, and A. Święch, Stochastic representations for solutions to parabolic Dirichlet problems for nonlocal Bellman equations, Ann. Appl. Probab., 29 (2019), pp. 3271–3310.
[40] A. Granas and J. Dugundji, Fixed point theory, Springer, 2003.
[41] O. Guéant, J.-M. Lasry, and P.-L. Lions, Mean field games and applications, in Paris–Princeton Lectures on Mathematical Finance 2010, Springer, 2011, pp. 205–266.
[42] F. B. Hanson, Applied stochastic processes and control for jump-diffusions: Modeling, analysis, and computation, Society for Industrial and Applied Mathematics (SIAM), 2007.
[74] H. Pham, Continuous-time stochastic control and optimization with financial applications, Springer, 2009.
[75] M. Ricciardi, Mean Field Games PDE with Controlled Diffusion, in Some Advances in Mean Field Games Theory, Rome–Paris, 2020, pp. 133–173. PhD Thesis.
[76] R. T. Rockafellar, Convex analysis, Princeton University Press, 1970.
[77] M. Röckner, L. Xie, and X. Zhang, Superposition principle for non-local Fokker–Planck–Kolmogorov operators, Probab. Theory Related Fields, 178 (2020), pp. 699–733.
[78] M. Röckner and X. Zhang, Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients, C. R. Math. Acad. Sci. Paris, 348 (2010), pp. 435–438.
[79] X. Ros-Oton and J. Serra, Regularity theory for general stable operators, J. Differential Equations, 260 (2016), pp. 8675–8715.
[80] K. Sato, Lévy processes and infinitely divisible distributions, Cambridge University Press, 2013.
[81] J. Serra, $C^{\alpha + \epsilon}$ regularity for concave nonlocal fully nonlinear elliptic equations with rough kernels, Calc. Var. Partial Differential Equations, 54 (2015), pp. 3571–3601.
[82] E. Sinestrari, On the abstract Cauchy problem of parabolic type in spaces of continuous functions, J. Math. Anal. Appl., 107 (1985), pp. 16–66.
[83] J. Smoller, Shock waves and reaction-diffusion equations, vol. 258, Springer, second ed., 1994.
[84] J. L. Vázquez, The porous medium equation: Mathematical theory, The Clarendon Press, Oxford University Press, 2007.
[85] J. L. Vázquez, A. de Pablo, F. Quirós, and A. Rodríguez, Classical solutions and higher regularity for nonlinear fractional diffusion equations, J. Eur. Math. Soc. (JEMS), 19 (2017), pp. 1949–1975.
[86] L. Wang, On the regularity theory of fully nonlinear parabolic equations, II, Comm. Pure Appl. Math., 45 (1992), pp. 141–178.

I. Chowdhury
Prirodoslovno–matematički fakultet, Sveučilište u Zagrebu,
Horvatovac 102A, 10000 Zagreb, Croatia
Email address: indranil.chowdhury@math.hr, indranill2011@gmail.com

E. R. Jakobsen
Institutt for matematiske fag, NTNU,
7491 Trondheim, Norway
Email address: espen.jakobsen@ntnu.no

M. Krupski
Instytut Matematyczny, Uniwersytet Wrocławski,
pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland
Prirodoslovno–matematički fakultet, Sveučilište u Zagrebu,
Horvatovac 102A, 10000 Zagreb, Croatia
Email address: mirosz.krupski@uwr.edu.pl