ON FULLY NONLINEAR PARABOLIC MEAN FIELD GAMES WITH NONLOCAL AND LOCAL DIFFUSIONS

INDRANIL CHOWDHURY, ESPEN R. JAKOBSEN, AND MIŁOSZ KRUPSKI

Abstract. 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 a new control interpretation involving time change rates of stochastic (Lévy) processes. The main results are existence and uniqueness of solutions under general assumptions. These results are applied to non-degenerate equations — including both local second order and nonlocal with fractional Laplacians. Uniqueness holds under monotonicity of couplings and convexity of the Hamiltonian, but neither monotonicity nor convexity need to be strict. We consider a rich class of nonlocal operators and processes and develop tools to work in the whole space without explicit moment assumptions.

1. Introduction

In this paper we introduce a new model of *mean field games* and analyse it using PDE methods. Mean field games are limits of N-player stochastic games as $N \to \infty$, under certain assumptions allowing for the mean field limit to exist. The Nash equilibria are characterized by a coupled system of PDEs called the *mean field game system*, where the value function of the generic player is given by a backward Hamilton–Jacobi–Bellman equation and the distribution of players by a forward Fokker–Planck equation. The mathematical theory of such problems was introduced by Lasry–Lions [57, 58, 59] and Huang–Caines–Malhamé [41, 40] in 2006, and today this is a large and rapidly expanding field. This research is mostly focused on either PDE or stochastic approaches. Extensive background and recent developments can be found in e.g. [1, 9, 17, 18, 35, 15, 37] and the references therein.

In contrast to the more classical setting, we allow not only the drift of a stochastic process to be controlled but also the diffusion. To be more precise, the players control the time change rate of a Lévy process. If the (diffusion) process is self-similar like a Brownian motion or an α-stable process, this is equivalent to a classical controlled diffusion [32, 66] (see Section 3 and [25] for more details). In our setup the backward equation is fully nonlinear, and the system may be strongly degenerate and local or nonlocal. Problems sharing some of these features have been addressed before. In [16] the authors allow for a degenerate diffusion, but it is not controlled and there are restrictions on its regularity, cf. [71, 11]. There are recent results on mean field games with nonlocal (uncontrolled) diffusion involving Lévy operators [20, 26, 31, 45]. See also [14] for a problem involving fractional time derivatives.

Date: August 30, 2024.

2020 *Mathematics Subject Classification.* 35A01, 35A02, 35D30, 35D40, 35K55, 35K65, 35Q84, 35Q89, 35R09, 47D07, 49L, 49N80, 60G51.

Key words and phrases. Mean field games, Fokker–Planck–Kolmogorov equation, Hamilton–Jacobi–Bellman equation, fully-nonlinear PDEs, nonlocal PDEs, Brownian motion, Lévy processes, controlled diffusion, existence, uniqueness.

Control problems/games have many applications throughout the sciences, engineering, and economics. Controlled diffusions [32] appear e.g. in portfolio optimization in finance, cf. [12, 38, 68]. Despite the many applications in economics [37, 34], see also [3], control of the diffusion is a rare and novel subject in the context of mean field games. So far it has been addressed mostly by stochastic methods: [55] introduces an approach based on relaxed controls and martingale problems to show existence (without uniqueness) of probabilistic solutions to very general local mean field games, see further developments in e.g. [6]. Mean field games of controls are considered in [29], and [8, 7] consider extensions to problems perturbed by bounded nonlocal operators. Some results by PDE methods can be found in [69], as well as [2, 27] for uniformly elliptic (stationary second order) problems. Except for [69], there seem to be no prior uniqueness results for fully nonlinear problems by any methods.

We focus on the case where the generic player has a single control, which addresses most of the novelties. We also explain how to include a separately controlled drift and hence the corresponding first-order terms in the PDEs. To be precise, we mainly study derivation, existence, and uniqueness questions for the mean field game system

$$
\begin{cases}
-\partial_t u = F(\mathcal{L}u) + \mathfrak{f}(m) & \text{on } \mathcal{T}\times\mathbb{R}^d,\\
u(T) = \mathfrak{g}(m(T)) & \text{on } \mathbb{R}^d,\\
\partial_t m = \mathcal{L}^*(F'(\mathcal{L}u)\, m) & \text{on } \mathcal{T}\times\mathbb{R}^d,\\
m(0) = m_0 & \text{on } \mathbb{R}^d,
\end{cases}
\tag{1}
$$

where $\mathcal{T} = (0,T)$ for a fixed $T > 0$. We assume $\mathcal{L}$ to be a Lévy operator with triplet (c, a, ν), an infinitesimal generator of a Lévy process (see [13, §2.1]). The (formal) adjoint $\mathcal{L}^*$ of $\mathcal{L}$ is also a Lévy operator. Typical examples are the Laplacian Δ, $(c, a, \nu) = (0, I, 0)$, the fractional Laplacian $-(-\Delta)^\sigma$, $(c, a, \nu) = (0, 0, \bar{c}\,|z|^{-d-2\sigma}\,dz)$ for $\bar{c} > 0$ and $\sigma \in (0, 1)$, and tempered, nonsymmetric, and even degenerate elliptic operators. We discuss more in Section 2.

A semi-rigorous derivation of problem (1) is given in Section 3, starting with a precise interpretation of the control problem for the generic player in terms of the time change rate of the Lévy process. Our derivation leads to a Hamiltonian F which is convex and non-decreasing, an optimal feedback control $\theta^* = F'(\mathcal{L}u)$, and ultimately to the mean field game system (1) which is then parabolic. It is coupled through the running and terminal costs $\mathfrak{f}$ and $\mathfrak{g}$ and the optimal feedback control. In this paper $\mathfrak{f}$ and $\mathfrak{g}$ are smoothing (nonlocal) couplings.

Our first objective is to study the well-posedness of problem (1) with a (nearly) minimal set of assumptions (A1)–(A5), naturally arising from the analysis in Section 3. This is matched with one of the weakest solution concepts where the feedback control is well-defined: classical solutions u and measure-valued distributional solutions m. Reworking the mean field games arguments to fit our setting, we reduce the question of well-posedness to a set of general conditions (S1)–(S5) describing the properties of solutions of the uncoupled equations making up problem (1): solvability, stability, and regularity of the Hamilton–Jacobi–Bellman equation (4), and uniqueness of the Fokker–Planck equation (5) — see below.

To prove uniqueness, we impose monotonicity assumptions on the couplings. Improving on previous results, we need neither strict convexity of the Hamiltonian nor strict monotonicity of the couplings. Existence for problem (1) holds under much weaker assumptions than uniqueness, in part because we need no uniqueness for the Fokker–Planck equation. We exploit the Kakutani–Glicksberg–Fan fixed point theorem (a generalization of the Schauder theorem), relying on a stability result for sets of solutions of the Fokker–Planck equation. It is based on new tightness arguments which require no moment assumptions in $\mathbb{R}^d$ on m or ν.[1] This approach is of independent interest and has already been exploited in [31, 24, 44].

The second objective is to verify the abstract conditions (S1)–(S5) and hence obtain well-posedness in concrete cases. We consider two non-degenerate problems — local (Section 2.2) and nonlocal (Section 2.3). In an upcoming paper [25], we also show that our findings can be applied to certain strongly degenerate problems. In Section 2.5 we formulate the results for a combination of controlled drift and diffusion where the mean field game system also includes the first-order terms

$$
\begin{cases}
-\partial_t u - H(\nabla u) - F(\mathcal{L}u) = \mathfrak{f}(m) & \text{on } \mathcal{T}\times\mathbb{R}^d,\\
u(T) = \mathfrak{g}(m(T)) & \text{on } \mathbb{R}^d,\\
\partial_t m + \operatorname{div}\left(\nabla H(\nabla u) m\right) - \mathcal{L}^*\left(F'(\mathcal{L}u)\, m\right) = 0 & \text{on } \mathcal{T}\times\mathbb{R}^d,\\
m(0) = m_0 & \text{on } \mathbb{R}^d.
\end{cases}
\tag{2}
$$

The main challenge is to get classical solvability and strong enough a priori regularity estimates for solutions of Hamilton–Jacobi–Bellman equations. In the local case, the key Schauder regularity *and solvability* results are proved in [73]. In the nonlocal case, the Schauder estimates are proved in [30], but we could find no existence result in the literature. To show existence, we adapt the continuity method described in [52], using the a priori estimates of [30], and solvability for *linear* nonlocal equations of [63]. In both non-degenerate cases, uniqueness for the Fokker–Planck equation can be deduced from existing results by adapting the Holmgren method.

To summarise, the main novelties of this paper are:

(*i*) The new model and well-posedness results in Section 2.

(*ii*) The stochastic control interpretation of the Hamilton–Jacobi–Bellman equation (in terms of time-rate change); a heuristic derivation of problem (1) in Section 3.

(*iii*) A theory of mean field games in $\mathbb{R}^d$ without moment assumptions (see Section 2.6); the technical results in Lemmas 4.9, 4.11, and 4.15.

(*iv*) Existence and stability for the Fokker–Planck equation with an arbitrary Lévy operator and a non-negative continuous coefficient in Section 6.1, and their use to prove Theorem 7.5, existence for problem (1).

(*v*) Uniqueness for problem (1) without strong convexity of F or strict monotonicity of $\mathfrak{f}$, $\mathfrak{g}$; the second half of the proof (from (38)) of Theorem 7.7.

The paper is organized as follows: In Section 2 we introduce assumptions, solution concepts, and the (concrete and general) well-posedness results for the fully nonlinear mean field game system (1) along an extension to system (2) with controlled drift. The derivation of PDEs from a stochastic model is given in Section 3. Section 4 discusses both the background material and new results that are needed in the proofs, including tightness and approximations of Lévy operators. Section 5 contains results on Hamilton–Jacobi–Bellman equations, and in Section 6 we discuss well-posedness for Fokker–Planck equations, including existence and stability of solutions under general assumptions. Section 7 contains the proofs of the general

[1] In the mean field game literature, m is usually continuous in the Wasserstein d_1 distance and has two bounded moments, whereas here we work with the Rubinstein–Kantorovich (or bounded–Lipschitz) distance d_0 and no explicit moment bounds. The challenge is to preserve compactness.

existence and uniqueness results for problem (1) and can be read independently. Some technical proofs and auxiliary results are given in the appendices.

2. Main results

We present our setting and the main results regarding well-posedness for problem (1), and their extension to problem (2). We also discuss the lack of moment assumptions for the initial data and the Lévy process.

2.1. Assumptions and solution concepts.

A Lévy measure ν is defined by

$$\nu \text{ is a Radon measure on } \mathbb{R}^d \setminus \{0\}, \quad \nu \geq 0, \quad \int_{\mathbb{R}^d} \left(1 \wedge |z|^2\right) \nu(dz) < \infty. \tag{3}$$

The representation formula for the Lévy operator $\mathcal{L}$ can the be given as:

(L):† $\mathcal{L} : C_b^2(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ is a linear operator with a triplet (c, a, ν), where $c \in \mathbb{R}^d$, $a \in \mathbb{R}^{d\times d}$, ν is a Lévy measure (3), and

$$\mathcal{L}\phi(x) = c \cdot \nabla\phi(x) + \operatorname{tr}\left(aa^T D^2\phi(x)\right) + \int_{\mathbb{R}^d} \Big(\phi(x+z) - \phi(x) - \mathbb{1}_{B_1}(z)\, z \cdot \nabla\phi(x)\Big)\, \nu(dz).$$

Let $\mathcal{P}(\mathbb{R}^d)$ be the set of probability measures on $\mathbb{R}^d$ equipped with the topology of weak convergence of measures. This topology can be metrised by the Rubinstein–Kantorovich norm $\|\cdot\|_0$ defined in Definition 4.4.

In problem (1), we then use the following assumptions:

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

(A2): F is convex;

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

(A4):† $\mathfrak{f} : C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d)) \to C_b(\mathcal{T} \times \mathbb{R}^d)$ and $\mathfrak{g} : \mathcal{P}(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ are continuous, i.e. $\lim\limits_{n\to\infty} \sup_{t\in\mathcal{T}} \|m_n(t) - m(t)\|_0 = 0$ implies

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

(A5): $\mathfrak{f}$ and $\mathfrak{g}$ are monotone operators, namely

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

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

Remark 2.1. (a) $F' \in \mathcal{C}^\gamma(\mathbb{R})$ is needed for uniqueness. For our existence results, $F \in C^1(\mathbb{R})$ and $\gamma = 0$ in (A1) is sufficient.

(b) $F' \geq 0$ in (A1), but $F' \geq \kappa > 0$ in the concrete cases we discuss below in Sections 2.2 and 2.3. However, the general theory of Section 2.4 holds under the weaker assumption $F' \geq 0$. The latter results allow us to handle a class of degenerate mean field games, and are needed for the upcoming paper [25].

(c) By the Legendre–Fenchel transform, for F satisfying (A1) and (A2),

$$F(z) = \sup_{\zeta\in[0,\infty)} \big(z\zeta - F^*(\zeta)\big) \qquad \text{for} \qquad F^*(\zeta) = \sup_{z\in\mathbb{R}} \big(\zeta z - F(z)\big).$$

†These three conditions need to be strengthened for our results to hold in the concrete cases we present, compare the statements of Theorem 2.6, Theorem 2.8, and Theorem 2.9.

Accordingly, *every* such F is the nonlinearity of a Hamilton–Jacobi–Bellman equation from the stochastic control theory [32, 68]. Hence for a fixed m, the first equation in (1) is a Hamilton–Jacobi–Bellman equation.[2] Note that $F^*(\zeta) = \infty$ for $\zeta < 0$. See Appendix B and (10)–(12) for more details.

(d) 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]).

(e) Assumption (A5) is the *standard Lasry–Lions* monotonicity conditions required for uniqueness. The equivalent and more familiar formulation with $\mathfrak{f}$ and $\mathfrak{g}$ non-decreasing [59, 1] is obtained by taking $\widetilde{\mathfrak{g}} = -\mathfrak{g}$, $\widetilde{\mathfrak{f}} = -\mathfrak{f}$, and $\widetilde{u} = -u$, which leads to $-\partial_t \widetilde{u} = -F(-\mathcal{L}\widetilde{u}) + \widetilde{\mathfrak{f}}(m)$ and $\widetilde{u}(T) = \widetilde{\mathfrak{g}}(m)$ in problem (1). Our choice simplifies the notation when nonlinear diffusion is involved.

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

With $(f, g) = \big(\mathfrak{f}(m), \mathfrak{g}(m(T))\big)$, the first pair of equations in problem (1) form a terminal value problem for a fully nonlinear Hamilton–Jacobi–Bellman equation,

$$\begin{cases} -\partial_t u = F(\mathcal{L}u) + f & \text{on } \mathcal{T} \times \mathbb{R}^d, \\ u(T) = g & \text{on } \mathbb{R}^d. \end{cases} \tag{4}$$

In this case the viscosity solution framework applies, but we consider (bounded) classical solutions, where $\partial_t u$ and $\mathcal{L}u$ are continuous functions. Then $\mathcal{L}u$ and the *second pair* of equations in problem (1) are well-defined. With $b = F'(\mathcal{L}u)$ this pair forms an initial value problem for a Fokker–Planck equation,

$$\begin{cases} \partial_t m = \mathcal{L}^*(bm) & \text{on } \mathcal{T} \times \mathbb{R}^d, \\ m(0) = m_0 & \text{on } \mathbb{R}^d. \end{cases} \tag{5}$$

Since $b = F'(\mathcal{L}u)$ need not be very regular and may even degenerate,[3] we consider very weak (measure-valued) solutions of problem (5). Classical solutions m would require even more regularity on u and the data.

Definition 2.2. Suppose $b \in C_b(\mathcal{T} \times \mathbb{R}^d)$. A function $m \in C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ is a *very weak solution* of problem (5) if for every $\phi \in C_c^\infty(\overline{\mathcal{T}} \times \mathbb{R}^d)$ and $t \in \overline{\mathcal{T}}$,

$$\begin{aligned} &\int_{\mathbb{R}^d} \phi(t,x)\, m(t,dx) - \int_{\mathbb{R}^d} \phi(0,x)\, m_0(dx) \\ &\qquad = \int_0^t \int_{\mathbb{R}^d} \big(\partial_t \phi(\tau,x) + b(\tau,x)(\mathcal{L}\phi)(\tau,x)\big)\, m(\tau,dx)\, d\tau. \end{aligned} \tag{6}$$

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

Definition 2.3. A pair (u, m) is a *classical–very weak solution* of problem (1) if u is a bounded classical solution of problem (4) with data $\big(\mathfrak{f}(m), \mathfrak{g}(m(T))\big)$, such that $F'(\mathcal{L}u) \in C_b(\mathcal{T} \times \mathbb{R}^d)$, and m is a very weak solution of problem (5) with initial data m_0 and coefficient $b = F'(\mathcal{L}u)$.

We now give the main results of the paper.

---

[2] F has the form of (3.2) in [32, Chapter IV] with control $v = \zeta \in U = [0, \infty)$, (drift) $f = 0$, $a = v$, $L = F^*(v) + \mathfrak{f}(m)$. For $\mathcal{L} = \Delta$, the first equation in (1) is then the HJB equation (3.3) in [32, Chapter IV].

[3] An example of a degenerate model is given in [25].

2.2. **Well-posedness for local second-order mean field games.** Here we assume $2\sigma = 2$ and:

(L′): $\mathcal{L}\phi(x) = \operatorname{tr}\left(aa^T D^2\phi(x)\right)$ where $\det aa^T > 0$.

(R): There are $\alpha \in (0,1]$ and $M \in [0,\infty)$ such that the range[4]

$$\mathcal{R} = \left\{\left(\mathfrak{f}(m), \mathfrak{g}(m(T))\right) : m \in \mathcal{C}^{\frac{1}{2}}(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))\right\}$$

satisfies $\mathcal{R} \subset \mathcal{R}_0(\alpha, M)$, where[5]

$$\begin{aligned}\mathcal{R}_0(\alpha, M) = \Big\{(f,g) :\ & (i) \quad f \in \mathcal{C}_b^{1,\alpha}(\mathcal{T} \times \mathbb{R}^d),\\ & (ii) \quad g \in BUC(\mathbb{R}^d) \quad \text{and} \quad \mathcal{L}g \in L^\infty(\mathbb{R}^d),\\ & (iii) \quad \|f\|_{1,\alpha} + \|\mathcal{L}g\|_\infty + \|g\|_\infty \le M\Big\}.\end{aligned}$$

(A1′): (A1) holds and $F' \ge \kappa$ for some $\kappa > 0$ (i.e. F is strictly increasing).

Under (L′), the operator $\mathcal{L}$ is non-degenerate. For problem (1) to be non-degenerate, we also need to assume (A1′). Solutions m always belong to $\mathcal{C}^{\frac{1}{2}}(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ by Lemma 6.2 *(ii)*. In this setting, we expect interior regularity estimates to hold.

Definition 2.4 (Interior estimates)**.** Assume (L). *Interior* (β,α)*-regularity estimates* hold for problem (4) if for every $f \in \mathcal{C}_b^{\beta,\alpha}(\mathcal{T} \times \mathbb{R}^d)$, and $(t,x) \in \mathcal{T} \times \mathbb{R}^d$, and a viscosity solution u of problem (4),[6] we have

$$[\partial_t u]_{\mathcal{C}^{\beta,\alpha}([0,t]\times B_1(x))} + [\mathcal{L}u]_{\mathcal{C}^{\beta,\alpha}([0,t]\times B_1(x))} \le C(t)\big(\|f\|_{\beta,\alpha} + \|u\|_\infty\big).$$

In view of the comparison principle (Theorem 5.3), the right-hand side can be expressed in terms of $\|f\|_{\beta,\alpha}$ and $\|g\|_\infty$. When F is affine, interior regularity is given by classical Schauder theory (see e.g. [51, 56, 60]). In the fully nonlinear case, such estimates have been proved in [73]. Related results can be found in e.g. [49, 50, 60].

Lemma 2.5. *Assume (L′), $(f,g) \in \mathcal{R}_0(\alpha, M)$ (as in (R)), (A1′), (A2). Then interior $(\alpha/2,\alpha)$-regularity estimates hold for problem (4).*

Proof. The result is stated in a form which is a corollary to [65, Theorem 5.2]. As in [65], it follows from the arguments in [73], in particular Theorems 1.1 and 4.13 and their proofs. (Our case is slightly simpler since $\mathcal{L}$ is translation invariant.) □

Theorem 2.6. *Assume (L′), (R), (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 is a corollary of the more general well-posedness result of Theorem 2.9. The result and an outline of the proof is given in Section 2.4.

2.3. **Well-posedness for nonlocal mean field games.** Here we assume:

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

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

[4] The space $\mathcal{C}^{\frac{1}{2}}(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ is defined as in Definition 4.1, only with respect to the norm $\|\cdot\|_0$.

[5] See Definition 4.2 of spaces $\mathcal{C}_b^{\beta,\alpha}(\mathcal{T} \times \mathbb{R}^d)$; BUC = bounded uniformly continuous.

[6] See Definition 4.2; see Definition 5.1 of viscosity solutions for $a = 0$ (analogous for $a \ne 0$).

where ν is a Lévy measure (see (3)), $\nu|_{B_1}$ is absolutely continuous with respect to the Lebesgue measure, and there exists a function k such that for α as in (R) and $K>0$ (see (15)),

$$\mathbb{1}_{B_1}(z)\,\nu(dz)=\frac{k(z)}{|z|^{d+2\sigma}}\,dz,\quad K^{-1}\le k(z)\le K,\quad [k]_{\mathcal{C}^\alpha(B_1)}<\infty.$$

If $2\sigma=1$, then in addition $\int_{B_1\setminus B_r}\frac{zk(z)}{|z|^{d+1}}\,dz=0$ for every $r\in(0,1)$.[7]

(A1″): (A1′) holds and $F\in C^2(\mathbb{R})$ and $F''\in\mathcal{C}^1(\mathbb{R})$ (see Definition 4.1).

Again $\mathcal{L}$ is non-degenerate and we assume (A1′) to make problem (1) non-degenerate as well. Condition (L″) defines a rich class of nonlocal operators including fractional Laplacians and the nonsymmetric operators in finance. There is no restriction on the tail behaviour of ν other than (3), so underlying Lévy processes and solutions of corresponding Fokker–Planck equations may have no moments.

Despite many related results on interior regularity in the literature (see e.g. [21, 22, 30, 48, 62, 65]), we could not find a statement we could cite. In Appendix A we therefore prove the following.

Lemma 2.7. *Assume (L″), $(f,g)\in\mathcal{R}_0(\alpha,M)$ (as in (R)), (A1″), (A2). Then interior $(\frac{\alpha}{2\sigma},\alpha)$-regularity estimates hold for problem (4).*

We expect the result to be true under weaker regularity assumptions on F.

Theorem 2.8. *Assume (L″), (R), (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 is a corollary of the more general well-posedness result of Theorem 2.9. The result and an outline of the proof is given in Section 2.4.

2.4. General well-posedness theory.

We describe the properties of solutions to problem (4) and problem (5) that lead to well-posedness of problem (1). Let

$$\begin{aligned}\mathcal{S}_{HJB}=\big\{&u\in C_b(\overline{\mathcal{T}}\times\mathbb{R}^d)\text{ is a bounded classical solution of problem (4)}\\&\text{with data }(f,g)=\big(\mathfrak{f}(m),\mathfrak{g}(m(T))\big):m\in\mathcal{C}^{\frac12}(\overline{\mathcal{T}},\mathcal{P}(\mathbb{R}^d))\big\},\\ \mathcal{B}=\big\{&F'\big(\mathcal{L}u\big):u\in\mathcal{S}_{HJB}\big\}.\end{aligned}$$

(S1): For every $m\in\mathcal{C}^{\frac12}(\overline{\mathcal{T}},\mathcal{P}(\mathbb{R}^d))$ there exists a bounded classical solution u of problem (4) with data $(f,g)=\big(\mathfrak{f}(m),\mathfrak{g}(m(T))\big)$.

(S2): 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 $t\in\mathcal{T}$.

(S3): There exists $K_{HJB}\ge0$ such that $\|F'(\mathcal{L}u)\|_\infty\le K_{HJB}$ for every $u\in\mathcal{S}_{HJB}$.

(S4): It holds $\{\partial_t u,\ \mathcal{L}u:u\in\mathcal{S}_{HJB}\}\subset C_b(\mathcal{T}\times\mathbb{R}^d)$.

(S5): 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 (5).

[7]See Remark 4.10 for $2\sigma\in(0,1)$.

Condition (S1) describes existence of solutions of the Hamilton–Jacobi–Bellman equation (4), which are unique by Theorem 5.3, and (S5) describes uniqueness of solutions of the Fokker–Planck equation (5), which exist by Theorem 6.6. Conditions (S2), (S3), (S4) describe various (related) properties of solutions of problem (4). Under (A1), both (S3) and (S4) imply $b = F'(\mathcal{L}u) \in C_b(\mathcal{T} \times \mathbb{R}^d)$ for $u \in \mathcal{S}_{HJB}$.

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.

Proof of Theorem 2.9. The results are proved in Section 7. Existence is addressed in Theorem 7.5 by an application of the Kakutani–Glicksberg–Fan fixed point theorem, which requires a detailed analysis of problem (4) and problem (5). 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 7.7. 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 $\mathfrak{f}$ and $\mathfrak{g}$ in (A5), without further restrictions.

Proofs of Theorems 2.6, 2.8. The well-posedness results for non-degenerate cases (Sections 2.2–2.3) follow by verifying the general conditions (S1)–(S5) and then applying Theorem 2.9. We obtain (S1)–(S5) from Theorem 5.5 and Theorem 6.7 — see Corollary 5.6 and Corollary 6.9.

2.5. **Extensions to include controlled drift.** When the diffusion operator $\mathcal{L}$ is non-degenerate and of order $2\sigma > 1$ (local or nonlocal), the well-posedness results above can easily be extended to include controlled drift. To illustrate this, we consider problem (2) which comes from a model where the drift and the time-rate changes of the driving Lévy process are controlled separately (with separate controls).

Theorem 2.10. *Assume (L′) or (L″) with $\sigma > \frac{1}{2}$, and (R), (A1″), (A2), (A3). Let $H \in C^2(\mathbb{R}^d)$ be strictly convex and $D^2H \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d \times \mathbb{R}^d)$. If in addition*

(i) (A4) holds, then there exists a classical–very weak solution of problem (2);

(ii) (A5) holds, then problem (2) has at most one classical–very weak solution.

We omit the details of the proof. By the assumed regularity and convexity of H, it can be adapted from the arguments we use to prove Theorem 2.6 or Theorem 2.8. With little additional difficulty, most of the effort would involve tedious rewriting of the results of Section 6.1. Importantly, the results we used from [73, 30, 63] to establish the interior regularity estimates for the Hamilton–Jacobi–Bellman equation, as well as uniqueness for the Fokker–Planck equation, still hold in the setting of Theorem 2.10.

2.6. **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 $(\mathcal{P}_1, d_1)$ (or Wasserstein-p for $p > 1$) in the analysis of the Fokker–Planck equations. Here $\mathcal{P}_1$ is the space of probability measures with finite first moments. For compactness, finite $1 + \varepsilon$ moments are typically 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 moments as the tails of their Lévy measures [72, Theorem 25.3], e.g. the Brownian motion has moments of any order, while a 2σ-stable process with $\nu(dz) \approx \frac{dz}{|z|^{d+2\sigma}}$ only has moments of order less than $2\sigma \in (0,2)$. Condition (L″) puts no restriction on $\nu|_{B_1^c}$. The mean field games we consider may thus be driven by processes with unbounded first moments, like the 2σ-stable processes for $2\sigma \leq 1$. This means that we cannot work in $(\mathcal{P}_1, d_1)$, even when the initial distribution m_0 has moments of all orders.

We work in the space $(\mathcal{P}, d_0)$ of probability measures under weak convergence, metrised by d_0, defined by the Rubinstein–Kantorovich norm $\|\cdot\|_0$ (see Section 4.1). The d_0-topology is strictly weaker than the d_1-topology, as it does not require convergence of first moments. The tools we develop can be useful for other problems and have already been used [31, 24, 44]. In the local case they yield results for a larger class of initial distributions m_0 than usually considered. Crucial ingredients are more refined tightness arguments and their interplay with Lévy processes. In particular, the sequence of Lemmas 4.9, 4.11 and 4.15, regarding compact sets in $(\mathcal{P}, d_0)$, and *a priori* estimates for approximations of Lévy operators, leading to Lemma 6.2.

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 [5, 42, 64, 67] with applications e.g. in modelling markets or turbulence [4, 19]. For SDEs driven by self-similar processes, like the Brownian motion or an α-stable process, this type of control coincides with the classical (continuous) control [32, 66].[8] 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 ([5, Definition 1.1])**.** A *random time change* θ_s is an almost surely non-negative, non-decreasing stochastic process which is a finite stopping time for each fixed s.[9] It is *absolutely continuous* if there exists a non-negative $\mathcal{F}_s$-adapted process θ' 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 $X_s^{t,x}$ starting from $X_t^{t,x} = x$ by $X_s^{t,x} = x + X_s - X_t$. Then, for an absolutely continuous random time change θ_s such that $\theta_t = t$, θ'_s is deterministic at $s = t$, 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 $Y_s^{t,x,\theta} = X_{\theta_s}^{t,x}$. It is an inhomogeneous Markov process associated with the families of operators P^θ and transition probabilities p^θ (see [33, §1.1, §1.2 (10)]) given by

$$P_{t,s}^\theta \phi(x) = \int_{\mathbb{R}^d} \phi(y)\, p^\theta(t,x,s,dy) = E\phi\big(Y_s^{t,x,\theta}\big) \tag{7}$$

for $\phi \in C_b(\mathbb{R}^d)$. To compute the "generator" $\mathcal{L}_\theta$ of $Y^{t,x,\theta}$, note that by the Dynkin formula [13, (1.55)], if $\phi \in \mathrm{Dom}(\mathcal{L})$

$$E\phi\big(Y_s^{t,x,\theta}\big) - \phi(x) = E\bigg(\int_t^{\theta_s} \mathcal{L}\phi(X_\tau^{t,x})\, d\tau\bigg),$$

[8] By self-similarity (e.g. for the Brownian motion $B_{ct} = \sqrt{c}B_t$) controlled time change is equivalent to control of the strength of the diffusion (controlled diffusion).

[9] θ_s is a stopping time if $\{\theta_s \leq \tau\} \subset \mathcal{F}_\tau$ for $\tau \geq 0$.

and by a change of variables,

$$\frac{P^\theta_{t+h,t}\phi(x) - \phi(x)}{h} = \frac{E\phi\big(Y^{t,x,\theta}_{t+h}\big) - \phi(x)}{h} = E\Big(\frac{1}{h}\int_t^{t+h} \mathcal{L}\phi\big(X^{t,x}_{\theta_\tau}\big)\theta'_\tau\, d\tau\Big).$$

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

$$\mathcal{L}_\theta\phi(x) = \lim_{h\to 0^+} \frac{P^\theta_{t+h,t}\phi - \phi}{h}(x) = \theta'_t\,\mathcal{L}\phi(x). \tag{8}$$

A proof of a more general result can be found in e.g. [5, Theorem 8.4].

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

$$J(t,x,\theta) = E\bigg(\int_t^T \ell\big(s, Y^{t,x,\theta}_s, \theta'_s\big)\, ds + g\big(Y^{t,x,\theta}_T\big)\bigg).$$

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

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

$$u(t,x) = \sup_{\theta\in\mathcal{A}} J(t,x,\theta). \tag{9}$$

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

$$u(t,x) = \sup_\theta E\bigg(\int_t^{t+h} \ell\big(s, Y^{t,x,\theta}_s, \theta'_s\big)\, ds + u\big(t+h, Y^{t,x,\theta}_{t+h}\big)\bigg),$$

and hence

$$\begin{aligned} &-\frac{u(t+h,x) - u(t,x)}{h} \\ &\quad = \sup_\theta E\bigg(\frac{u\big(t+h, Y^{t,x,\theta}_{t+h}\big) - u(t+h,x)}{h} + \frac{1}{h}\int_t^{t+h} \ell\big(s, Y^{t,x,\theta}_s, \theta'_s\big)\, ds\bigg). \end{aligned}$$

Recalling the definition of $\mathcal{L}_\theta$ in (8), we can (heuristically at least) pass to the limit as $h \to 0$ and find the following dynamic programming — or Bellman — equation

$$-\partial_t u = \sup_{\zeta\geq 0}\Big(\zeta\mathcal{L}u + \ell(t,x,\zeta)\Big), \tag{10}$$

satisfied e.g. in the viscosity sense (see Section 5), where ζ denotes the (deterministic) value of θ'_t to simplify the notation. We now assume that $\ell(t,x,\zeta) = -L(\zeta) + f(t,x)$, 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. $F(z) = \sup_{\zeta\geq 0}\big(\zeta z - L(\zeta)\big)$, as

$$-\partial_t u = F\big(\mathcal{L}u\big) + f(t,x). \tag{11}$$

By the definitions of u and $X^{T,x}_T$ it also follows that

$$u(T,x) = Eg\big(X^{T,x}_T\big) = g(x). \tag{12}$$

3.3. **Optimal control and Fokker–Planck equation.** By the properties of the Legendre–Fenchel transform, when $\lim_{\zeta\to\infty} L(\zeta)/\zeta = \infty$ and L is strictly convex on $\{L \neq \infty\}$, the optimal value ζ in (10) satisfies $\zeta = F'(\mathcal{L}u)$ for every $(t,x) \in \mathcal{T} \times \mathbb{R}^d$ (see Proposition B.1). We therefore obtain a function

$$b(t,x) = \zeta = (\theta^*)'_t = F'\big(\mathcal{L}u(t,x)\big). \tag{13}$$

This is the optimal time change rate in the feedback form. The optimally controlled process and the optimal control in (9) are then implicitly given by

$$Y^*_s = X^{t,x}_{\theta^*_s} \qquad \text{and} \qquad \theta^*_s = t + \int_t^s b(\tau, Y^*_\tau)\, d\tau.$$

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

By defining $p^{\theta^*}(t,x,s,A) = \mathsf{P}(Y^*_s \in A)$, if solutions of equations (11)–(12) are unique, we obtain a unique family of transition probabilities p^{θ^*} (cf. (7)), satisfying the Chapman–Kolmogorov relations. This family, in turn, defines a wide-sense Markov process (see [33, §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 [33, §1.1 Definition 3])[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 $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $t,h \geq 0$. Then,

$$\begin{aligned}&\int_{\mathbb{R}^d} \big(\varphi(t,x)\, m(t,dx) - \varphi(t+h,x)\, m(t+h,dx)\big)\\ &\quad= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \big(\varphi(t,y)\, p^{\theta^*}(t,x,t,dy) - \varphi(t+h,y) p^{\theta^*}(t,x,t+h,dy)\big)\, m(t,dx),\end{aligned}$$

and because of (8), (13) and the fact that $p^{\theta^*}(t,x,t,dy) = \delta_x(dy)$, this leads to

$$\partial_t \int_{\mathbb{R}^d} \varphi(t,x)\, m(t,dx) = \int_{\mathbb{R}^d} \Big(b(t,x)\mathcal{L}\varphi + \partial_t\varphi(t,x)\Big)\, m(t,dx).$$

Since $b = F'(\mathcal{L}u)$, by duality (see Definition 2.2) m is a very weak solution of

$$\partial_t m = \mathcal{L}^*\big(F'(\mathcal{L}u)\, m\big), \quad m(0) = m_0, \tag{14}$$

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 $\widehat{m}$ of the processes controlled (optimally) by the other players (see (A4))

$$f = \mathfrak{f}(\widehat{m}) \qquad \text{and} \qquad g = \mathfrak{g}\big(\widehat{m}(T)\big).$$

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

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

Note that the solution u depends on $\widehat{m}$, 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.

[10]Alternatively, $m(t)$ is the distribution of the solution $Z(t)$ of SDE $dZ(t) = b(t,Z(t))\, dX(t)$, $Z(0) \sim m_0$. Moreover, $Y^*_s = E\big[Z(s)|Z(t) = x\big]$, see [33, §1.2 (9), (10)].

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

4. Preliminaries

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_r^c we denote the ball of radius r centred at 0 and its complement in $\mathbb{R}^d$. Similarly, $B_r(x)$ denotes a ball centred at x.

Definition 4.1. A function ϕ is Hölder-continuous at $x \in \mathbb{R}^d$ with parameter $\alpha \in (0,1]$ if for some $r > 0$

$$[\phi]_{\mathcal{C}^\alpha(B_r(x))} = \sup_{y \in B_r(x)\setminus\{x\}} \frac{|\phi(x) - \phi(y)|}{|x-y|^\alpha} < \infty. \tag{15}$$

The space $\mathcal{C}^\alpha(\mathbb{R}^d)$ consists of functions which are Hölder-continuous at every point in $\mathbb{R}^d$ with parameter α. Further, define $\mathcal{C}_b^\alpha(\mathbb{R}^d) = \{\phi : \|\phi\|_\alpha < \infty\}$, where

$$[\phi]_\alpha = \sup_{x\in\mathbb{R}^d} [\phi]_{\mathcal{C}^\alpha(B_1(x))} \quad \text{and} \quad \|\phi\|_\alpha = \|\phi\|_{L^\infty(\mathbb{R}^d)} + [\phi]_\alpha.$$

Note that the definition of $\mathcal{C}_b^\alpha(\mathbb{R}^d)$ is equivalent to the more standard notation, where the supremum in (15) is taken over $|x-y| \in \mathbb{R}^d \setminus \{0\}$. The space $\mathcal{C}_b^1(\mathbb{R}^d)$ 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 $(t,x) \in \mathcal{T} \times \mathbb{R}^d$ and $\alpha, \beta \in (0,1]$, define

$$[\phi]_{\mathcal{C}^{\beta,\alpha}([0,t]\times B_r(x))} = \sup_{y\in B_r(x)} [\phi(y)]_{\mathcal{C}^\beta([0,t])} + \sup_{s\in[0,t]} [\phi(s)]_{\mathcal{C}^\alpha(B_r(x))}.$$

We also denote $\mathcal{C}_b^{\beta,\alpha}([0,t]\times\mathbb{R}^d) = \{\phi : \|\phi\|_{\mathcal{C}^{\beta,\alpha}([0,t]\times\mathbb{R}^d)} < \infty\}$, where

$$\|\phi\|_{\mathcal{C}^{\beta,\alpha}([0,t]\times\mathbb{R}^d)} = \|\phi\|_{L^\infty([0,t]\times\mathbb{R}^d)} + \sup_{x\in\mathbb{R}^d} [\phi]_{\mathcal{C}^{\beta,\alpha}([0,t]\times B_1(x))}.$$

Definition 4.3. When X is a normed space, $B(\mathcal{T}, X)$ denotes the space of bounded functions from $\mathcal{T}$ to X, i.e. $B(\mathcal{T},X) = \big\{u : \mathcal{T} \to X \ : \ \sup_{t\in\mathcal{T}} \|u(t)\|_X < \infty\big\}$.

Note the subtle difference between $B(\mathcal{T}, X)$ and the usual space $L^\infty(\mathcal{T}, X)$.

4.1. Spaces of measures.

Let $\mathcal{P}(\mathbb{R}^d)$ consist of probability measures on $\mathbb{R}^d$, a subspace of the space of bounded Radon measures $\mathcal{M}_b(\mathbb{R}^d) = C_0(\mathbb{R}^d)^*$. 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 $\mathcal{P}(\mathbb{R}^d)$ is equipped with the topology of weak convergence of measures,[11]

$$\lim_{n\to\infty} m_n = m \quad \text{if and only if} \quad \lim_{n\to\infty} m_n[\phi] = m[\phi] \text{ for every } \phi \in C_b(\mathbb{R}^d).$$

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 $\mathcal{M}_b(\mathbb{R}^d)$ is given by

$$\|m\|_0 = \sup\big\{m[\psi] : \psi \in \mathcal{C}_b^1(\mathbb{R}^d),\ \|\psi\|_\infty \le 1,\ [\psi]_1 \le 1\big\}.$$

While the space $\big(\mathcal{M}_b(\mathbb{R}^d), \|\cdot\|_0\big)$ is not completely metrisable, thanks to [46, Theorems 4.19 and 17.23], both $\mathcal{P}(\mathbb{R}^d)$ and $C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ are complete spaces. Let

$$\mathcal{P}_{ac}(\mathbb{R}^d) = \big\{u \in L^1(\mathbb{R}^d) : \|u\|_{L^1(\mathbb{R}^d)} = 1,\ u \ge 0\ \big\} = L^1(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d).$$

We endow $\mathcal{P}_{ac}(\mathbb{R}^d)$ with the topology inherited from $\mathcal{P}(\mathbb{R}^d)$.

[11] It is also called *narrow*, *vague* or *weak-$*$* convergence.

Definition 4.5. A set of measures $\Pi \subset \mathcal{P}(\mathbb{R}^d)$ 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 $\Pi \subset \mathcal{P}(\mathbb{R}^d)$ 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 $V(x) = V_0\big(\sqrt{1+|x|^2}\big)$ 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 have $\|\nabla V\|_\infty, \|D^2 V\|_\infty \leq 1$. Note that the choice of the constant 1 in this condition is arbitrary.

(b) $\big(1+|x|^2\big)^{a/2}$ for $a \in (0,1]$ and $\log\big(\sqrt{1+|x|^2}+1\big)$ are Lyapunov functions.

(c) If $m_0 \in \mathcal{P}(\mathbb{R}^d)$ 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 $V(x) \leq V(0) + |x|$, 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*

$$\mathcal{P}_{V,r} = \big\{m \in \mathcal{P}(\mathbb{R}^d) : m[V] \leq r\big\}$$

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 $K_\varepsilon = \{x : V(x) \leq \frac{r}{\varepsilon}\}$ is compact. Then it follows from the Chebyshev inequality that for every $m \in \mathcal{P}_{V,r}$,

$$m\big(K_\varepsilon^c\big) \;\leq\; \frac{\varepsilon}{r}\int_{\{V > \frac{r}{\varepsilon}\}} V\, dm \;\leq\; \frac{\varepsilon}{r} m[V] \;\leq\; \varepsilon.$$

Hence the set $\mathcal{P}_{V,r}$ is tight and thus compact by the Prokhorov theorem. □

The reverse statement is also true.

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

This result is crucial for our paper and is the reason why our findings hold without moment assumptions. The proof is given in Appendix A.

4.2. **Lévy operators.** In this section we collect some basic observations on Lévy operators. Recall the representation formula given in (L) in Section 2.1.

Remark 4.10. If $\int_{B_1} |z|\, \nu(dz) < \infty$, then we may equivalently write

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

In particular, we may have $\big(\int_{B_1} z\, \nu(dz), 0, \nu\big)$ as a triplet in (L).

Lemma 4.11. *Assume (L) and V is a Lyapunov function. The following are equivalent*

(i) $\int_{B_1^c} V(z)\, \nu(dz) < \infty$;

(ii) $\|\mathcal{L}V\|_\infty < \infty$;

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

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

Proof. Let

$$\vartheta_0(x) = c\cdot\nabla V(x) + \operatorname{tr}\left(aa^T D^2 V(x)\right) + \int_{B_1}\Big(V(x+z)-V(x)-z\cdot\nabla V(z)\Big)\nu(dz).$$

Because V is a Lyapunov function (see Remark 4.7 (a)), we have

$$\|\vartheta_0\|_\infty \le |c| + |a|^2 + \int_{B_1}|z|^2\,\nu(dz).$$

Observe that $\|\mathcal{L}V\|_\infty - \|\vartheta_0\|_\infty \le \|\vartheta_1\|_\infty \le \|\vartheta_2\|_\infty$, hence (iv) ⇒ (iii) ⇒ (ii). We also notice $\|\mathcal{L}V\|_\infty \ge \|\vartheta_1\|_\infty - \|\vartheta_0\|_\infty$ and $\int_{B_1^c} V(z)\,\nu(dz) = \vartheta_1(0) + \nu(B_1^c)$, thus (ii) ⇒ (iii) ⇒ (i).

It remains to prove (i) ⇒ (iv). Let $V_0\big(\sqrt{1+|x|^2}\big) = V(x)$ as in Definition 4.6 and notice that, because V_0 is subadditive and non-decreasing, we have

$$\big|V(y)-V(x)\big| \le V_0\Big(\Big|\sqrt{1+|y|^2}-\sqrt{1+|x|^2}\Big|\Big) \le V_0\big(\sqrt{1+|y-x|^2}\big).$$

Now we may estimate

$$\int_{B_1^c}\big|V(x+z)-V(x)\big|\,\nu(dz) \le \int_{B_1^c} V(z)\,\nu(dz).$$

□

Corollary 4.12. *Assume (L), (A3). There exists a Lyapunov function V such that $m_0[V], \|\mathcal{L}V\|_\infty < \infty$.*

Proof. Since $\nu|_{B_1^c}$ is a bounded measure, the set $\{\nu|_{B_1^c}, m_0\}$ is tight. Hence, by Lemma 4.9 we may find a Lyapunov function such that $\int_{B_1^c} V(z)\,\nu(dz) < \infty$ and $m_0[V] < \infty$. Thanks to Lemma 4.11 (ii) we also have $\|\mathcal{L}V\|_\infty < \infty$. □

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

$$\|\mathcal{L}\|_{LK} = |c| + |a|^2 + \frac{1}{2}\int_{B_1}|z|^2\,\nu(dz) + 2\nu(B_1^c). \tag{16}$$

Proposition 4.13. *Assume (L), $\phi \in C_b^2(\mathbb{R}^d)$. Then $\|\mathcal{L}\phi\|_\infty \le \|\mathcal{L}\|_{LK}\|\phi\|_{C_b^2(\mathbb{R}^d)}$.*

Proof. Using the Taylor expansion, we calculate

$$\begin{aligned}\|\mathcal{L}\phi\|_\infty &\le |c|\|\nabla\phi\|_\infty + |a|^2\|D^2\phi\|_\infty \\ &\quad + \Big|\int_{\mathbb{R}^d}\big(\phi(x+z)-\phi(x)-\mathbb{1}_{B_1}(z)\,z\cdot\nabla\phi(x)\big)\,\nu(dz)\Big| \\ &\le |c|\|\nabla\phi\|_\infty + |a|^2\|D^2\phi\|_\infty + \frac{\|D^2\phi\|_\infty}{2}\int_{B_1}|z|^2\,\nu(dz) + 2\|\phi\|_\infty\nu(B_1^c).\end{aligned}$$

□

Remark 4.14. The mapping $\mathcal{L} \mapsto \|\mathcal{L}\|_{LK}$ is a norm on the space (convex cone) of Lévy operators. It dominates the operator norm $C_b^2(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$, but they are not equivalent.

Lemma 4.15. *Assume (L). For $\varepsilon \in (0,1)$ there exist $\mathcal{L}^\varepsilon$, ν^ε such that*

$$\mathcal{L}^\varepsilon\mu(x) = \int_{\mathbb{R}^d}\big(\mu(x+z)-\mu(x)\big)\,\nu^\varepsilon(dz), \tag{17}$$

where $\mathcal{L}^\varepsilon : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$, $\nu^\varepsilon(\mathbb{R}^d) < \infty$ and $\operatorname{supp}\nu^\varepsilon \subset \mathbb{R}^d \setminus B_\varepsilon$. Moreover,

(i) $\|\mathcal{L}^\varepsilon\mu\|_{L^1(\mathbb{R}^d)} \le \big(c_{\mathcal{L}}/\varepsilon^3\big)\|\mu\|_{L^1(\mathbb{R}^d)}$ *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) $\sup_{\varepsilon\in(0,1)}\big(\|\mathcal{L}^\varepsilon V\|_\infty + \|\mathcal{L}^\varepsilon\|_{LK}\big) < \infty$ *for every Lyapunov function* V *such that* $\|\mathcal{L}V\|_\infty < \infty$.

Proof. $\diamond$ *Part* (i). Let (c, a, ν) be the Lévy triplet of $\mathcal{L}$ and $a = (a_1, \dots, a_d) \in \mathbb{R}^{d\times d}$ with $a_i \in \mathbb{R}^d$. Consider $\nu^\varepsilon = \nu_c^\varepsilon + \nu_a^\varepsilon + \nu_1^\varepsilon + \nu_2^\varepsilon$, where

$$\nu_c^\varepsilon = \frac{|c|}{\varepsilon}\delta_{\varepsilon\frac{c}{|c|}}, \qquad \nu_1^\varepsilon(E) = \nu(E\setminus B_\varepsilon),$$

$$\nu_a^\varepsilon = \sum_{i=1}^d \frac{|a_i|^2}{\varepsilon^2}\big(\delta_{\varepsilon\frac{a_i}{|a_i|}} + \delta_{-\varepsilon\frac{a_i}{|a_i|}}\big), \qquad \nu_2^\varepsilon(E) = \frac{1}{\varepsilon}\nu\Big(\big(B_1\setminus B_\varepsilon\big)\cap(-E/\varepsilon)\Big),$$

Notice that ν^ε is a bounded, non-negative measure with $\operatorname{supp}\nu^\varepsilon \subset \mathbb{R}^d\setminus B_\varepsilon$ (hence a Lévy measure). Let $\mathcal{L}^\varepsilon = \mathcal{L}^\varepsilon_{\text{loc}} + \mathcal{L}^\varepsilon_{\text{nloc}}$, where, for $\mu\in L^1(\mathbb{R}^d)$,

$$\begin{aligned}\mathcal{L}^\varepsilon_{\text{loc}}\,\mu(x) &= \int_{\mathbb{R}^d}\big(\mu(x+z)-\mu(x)\big)\,(\nu_c^\varepsilon+\nu_a^\varepsilon)(dz)\\ &= \frac{|c|}{\varepsilon}\Big(\mu\big(x+\varepsilon\tfrac{c}{|c|}\big)-\mu(x)\Big)+\sum_{i=1}^d\frac{|a_i|^2}{\varepsilon^2}\Big(\mu\big(x+\varepsilon\tfrac{a_i}{|a_i|}\big)+\mu\big(x-\varepsilon\tfrac{a_i}{|a_i|}\big)-2\mu(x)\Big).\end{aligned}$$

and

$$\begin{aligned}\mathcal{L}^\varepsilon_{\text{nloc}}\,\mu &= \int_{\mathbb{R}^d}\big(\mu(x+z)-\mu(x)\big)\,(\nu_1^\varepsilon+\nu_2^\varepsilon)(dz)\\ &= \int_{B_\varepsilon^c}\bigg(\mu(x+z)-\mu(x)+\mathbb{1}_{B_1}(z)\frac{\mu(x-\varepsilon z)-\mu(x)}{\varepsilon}\bigg)\,\nu(dz).\end{aligned}$$

Note that

$$\nu_1^\varepsilon(B_1\setminus B_\varepsilon)+\nu_2^\varepsilon(\mathbb{R}^d) = (1+\varepsilon^{-1})\nu(B_1\setminus B_\varepsilon)\le(\varepsilon^{-2}+\varepsilon^{-3})\int_{B_1}|z|^2\,\nu(dz),$$

and hence

$$\begin{aligned}\|\mathcal{L}^\varepsilon\mu\|_{L^1(\mathbb{R}^d)} &\le \bigg(\frac{2|c|}{\varepsilon}+\frac{4|a|^2}{\varepsilon^2}+2\nu(B_1^c)+\frac{2+2\varepsilon}{\varepsilon^3}\int_{B_1}|z|^2\,\nu(dz)\bigg)\|\mu\|_{L^1(\mathbb{R}^d)}\\ &\le \frac{4}{\varepsilon^3}\bigg(|c|+|a|^2+\int_{\mathbb{R}^d}\big(1\wedge|z|^2\big)\,\nu(dz)\bigg)\|\mu\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

This shows that $\mathcal{L}^\varepsilon : L^1(\mathbb{R}^d)\to L^1(\mathbb{R}^d)$ and $\|\mathcal{L}^\varepsilon\mu\|_{L^1(\mathbb{R}^d)}\le\big(c_{\mathcal{L}}/\varepsilon^3\big)\|\mu\|_{L^1(\mathbb{R}^d)}$.

$\diamond$ *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

$$\Big|\Big(\mathcal{L}^\varepsilon_{\text{loc}} - c\cdot\nabla - \operatorname{tr}\big(aa^TD^2(\,\cdot\,)\big)\Big)\varphi(x)\Big| \le \varepsilon\bigg(\frac{|c|}{2}\|D^2\varphi\|_\infty + |a|^2\|D^3\varphi\|_\infty\bigg). \tag{18}$$

Let $\mathcal{L}_\nu\varphi(x) = \int_{\mathbb{R}^d}\big(\varphi(x+z)-\varphi(x)-\mathbb{1}_{B_1}(z)\,z\cdot\nabla\varphi(x)\big)\,\nu(dz)$. Then

$$\begin{aligned}\Big|\big(\mathcal{L}^\varepsilon_{\text{nloc}}-\mathcal{L}_\nu\big)\varphi(x)\Big| &= \bigg|\int_{B_1\setminus B_\varepsilon}\bigg(\frac{\varphi(x-\varepsilon z)-\varphi(x)}{\varepsilon}+z\cdot\nabla\varphi(x)\bigg)\,\nu(dz)\\ &\qquad -\int_{B_\varepsilon}\Big(\varphi(x+z)-\varphi(x)-z\cdot\nabla\varphi(x)\Big)\,\nu(dz)\bigg|\\ &\le \frac{\varepsilon}{2}\|D^2\varphi\|_\infty\int_{B_1}|z|^2\,\nu(dz)+\frac{1}{2}\|D^2\varphi\|_\infty\int_{B_\varepsilon}|z|^2\,\nu(dz).\end{aligned} \tag{19}$$

Since $\lim_{\varepsilon\to0}\int_{B_\varepsilon}|z|^2\,\nu(dz) = 0$ by the Lebesgue dominated convergence theorem, it follows from (18) and (19) that $\lim_{\varepsilon\to0}\|(\mathcal{L}^\varepsilon-\mathcal{L})\varphi\|_\infty = 0$.

$\diamond$ *Part (iii)*. Let V be a Lyapunov function such that $\|\mathcal{L}V\|_\infty < \infty$. Then also $\|\mathcal{L}_\nu V\|_\infty < \infty$. By the definition of $\mathcal{L}^\varepsilon = \mathcal{L}^\varepsilon_{\text{loc}} + \mathcal{L}^\varepsilon_{\text{nloc}}$, in a way similar to (18), (19),

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

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

$$\int_{B_1} z\,\nu^\varepsilon_c(dz) = c, \qquad \int_{B_1} z\,\nu^\varepsilon_a(dz) = 0, \quad \text{and} \quad \int_{B_1} z\,(\nu^\varepsilon_1 + \nu^\varepsilon_2)(dz) = 0,$$

thus the Lévy triplet of the operator $\mathcal{L}^\varepsilon$ is $(c, 0, \nu_\varepsilon)$ (see Remark 4.10). Hence

$$\begin{aligned}\|\mathcal{L}^\varepsilon\|_{LK} &= |c| + \frac{\varepsilon|c|}{2} + |a|^2 + \frac{1}{2}\int_{B_1\setminus B_\varepsilon} (1+\varepsilon)|z|^2\,\nu(dz) + 2\nu(B_1^c)\\ &\le (1+\varepsilon)\|\mathcal{L}\|_{LK}.\end{aligned}$$

□

5. Hamilton–Jacobi–Bellman equations

In this section we define viscosity solutions and give results for problem (4). Let $(t, x, \ell) \mapsto \mathcal{F}(t, x, \ell)$ and w_0 be continuous functions, and $\mathcal{F}$ be non-decreasing in ℓ. For $\mathcal{L}$ satisfying (L) with $a = 0$,[12] consider the following problem

$$\begin{cases} \partial_t w = \mathcal{F}\big(t, x, (\mathcal{L}w)(t,x)\big), & \text{on } \mathcal{T}\times\mathbb{R}^d,\\ w(0) = w_0, & \text{on } \mathbb{R}^d. \end{cases} \tag{20}$$

For $0 \le r < \infty$ and $p \in \mathbb{R}^d$ we introduce linear operators

$$\begin{aligned}\mathcal{L}^r(\phi, p)(x) &= \int_{B_r^c} \Big(\phi(x+z) - \phi(x) - \mathbb{1}_{B_1}(z)\,z\cdot p\Big)\,\nu(dz),\\ \mathcal{L}_r\phi(x) &= \int_{B_r} \Big(\phi(x+z) - \phi(x) - \mathbb{1}_{B_1}(z)\,z\cdot\nabla\phi(x)\Big)\,\nu(dz),\end{aligned}$$

defined for bounded semicontinuous and C^2 functions respectively.

Definition 5.1. A bounded upper-semicontinuous function $u^- : \overline{\mathcal{T}}\times\mathbb{R}^d \to \mathbb{R}$ is a *viscosity subsolution* of problem (20) if $u^-(0, x) \le w_0(x)$ for every $x \in \mathbb{R}^d$ and for every $r \in (0,1)$, test function $\phi \in C^2\big(\mathcal{T}\times\mathbb{R}^d\big)$, and a maximum point (t, x) of $u^- - \phi$,

$$\partial_t\phi(t,x) - \mathcal{F}\Big(t, x, \big(c\cdot\nabla\phi + \mathcal{L}^r\big(u^-, \nabla\phi(t,x)\big) + \mathcal{L}_r\phi\big)(t,x)\Big) \le 0.$$

A supersolution is defined similarly, replacing max, upper-semicontinuous, and "≤" by min, lower-semicontinuous, and "≥". A viscosity solution is a sub- and supersolution at the same time. Note that bounded classical solutions are also bounded viscosity solutions.

Definition 5.2. The comparison principle holds for problem (20) if any subsolution u^- and supersolution u^+ satisfy $u^-(t,x) \le u^+(t,x)$ for every $(t,x) \in \overline{\mathcal{T}}\times\mathbb{R}^d$.

We have the following uniqueness, stability, and existence result for viscosity solutions of problem (4).

Theorem 5.3. *Assume (L), (A1), and (f, g) are bounded and continuous.*

(i) The comparison principle (see Definition 5.2) holds for problem (4).

[12] We take $a = 0$ for simplicity and to use the results of [23].

(ii) *Let u_1, u_2 be viscosity solutions of problem (4) with bounded uniformly continuous data (f_1, g_1), (f_2, g_2), respectively. Then for every $t \in \overline{\mathcal{T}}$,*

$$\|u_1(t) - u_2(t)\|_\infty \leq (T-t)\|f_1 - f_2\|_\infty + \|g_1 - g_2\|_\infty.$$

(iii) *There exists a unique viscosity solution of problem (4).*

Proof. ⋄ *Part (i).* In the nonlocal case ($a = 0$) with uniformly continuous f, u_0, this is [23, Theorem 6.1]. In the general case the result follows from a standard but long and tedious combination of the arguments of [23] and [43]. We omit this proof.

⋄ *Part (ii).* Note that for $\{i, j\} = \{1, 2\}$,

$$v_i(t, x) = u_j(t, x) - (T-t)\|f_1 - f_2\|_\infty - \|g_1 - g_2\|_\infty$$

is a viscosity subsolution of problem (4) with data (f_i, g_i). The result then follows from the comparison principle in Part *(i)*.

⋄ *Part (iii).* Part *(i)* entails uniqueness of viscosity solutions. It also implies existence of solutions through the Perron method (cf. [28, Section 4]). See also [23, Theorems 6.2] for the result when $a = 0$.

□

Now we give results that are specific for non-degenerate cases of problem (4), which correspond to the setting of Sections 2.2 and 2.3.

Proposition 5.4. *Assume (L), (A1), and u is a viscosity solution of problem (4) with bounded uniformly continuous data (f, g) such that $\partial_t f \in L^\infty(\mathcal{T} \times \mathbb{R}^d)$ and $\mathcal{L}g \in L^\infty(\mathbb{R}^d)$. Then $\partial_t u \in L^\infty(\mathcal{T} \times \mathbb{R}^d)$ 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 $g_\varepsilon = g * \rho_\varepsilon$, where ρ_ε is the standard mollifier. Note that $v_\varepsilon(t, x) = g_\varepsilon(x)$ is a viscosity (classical) solution of problem (4) with data $(-F(\mathcal{L}g_\varepsilon), g_\varepsilon)$, hence by Theorem 5.3 *(ii)*,

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

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,

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

Similarly, $v_h(t, x) = u(t-h, x)$ is a viscosity solution of problem (4) with data $(f(\,\cdot - h), u(T-h))$, thus for every $t \in \mathcal{T}$,

$$\begin{aligned}\|u(t) - v_h(t)\|_\infty &\leq (T-t)\|f(\,\cdot\,) - f(\,\cdot - h)\|_\infty + \|u(T-h) - u(T)\|_\infty \\ &\leq (T-t)\|\partial_t f\|_\infty h + \|F(\mathcal{L}g) + f\|_\infty h.\end{aligned}$$

Hence u is Lipschitz in time. □

Theorem 5.5. *Assume (L′) or (L″), $(f, g) \in \mathcal{R}_0(\alpha, M)$ (as in (R)), (A1′), (A2), and interior $(\frac{\alpha}{2\sigma}, \alpha)$-regularity estimates (Definition 2.4) hold for problem (4).*

(i) *There exists a bounded classical solution u of problem (4).*

(ii) *If u_n are bounded classical solutions of problem (4) with data $(f_n, g_n) \in \mathcal{R}_0(\alpha, M)$ 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 $t \in \mathcal{T}$.*

(iii) $\partial_t u,\ \mathcal{L}u \in C_b(\mathcal{T}\times\mathbb{R}^d)$ *and for every* $t\in\mathcal{T}$ *there is a constant* $C(t,f,g)$ *such that* $\|\mathcal{L}u\|_{\mathcal{C}^{\alpha/2\sigma,\alpha}([0,t]\times\mathbb{R}^d)} \le C(t,f,g)$.

Proof. ⋄ *Part* (i). There exists a bounded viscosity solution by Theorem 5.3 *(iii)*. Because of the interior regularity estimates, we have $\partial_t u,\ \mathcal{L}u \in C(\mathcal{T}\times\mathbb{R}^d)$, hence u is a bounded classical solution of problem (4).

⋄ *Part* (ii). By Part (i) and interior regularity estimates, for every $t\in\mathcal{T}$ and $r>0$, there exists a constant $C(t,r)>0$ such that

$$\sup_n \Big(\|\mathcal{L}u_n(t)\|_{L^\infty(B_r)} + [\mathcal{L}u_n(t)]_{\mathcal{C}^\alpha(B_r)} \Big) \le C(t,r).$$

By the Arzelà–Ascoli theorem, for every $t\in\mathcal{T}$ 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 $\mathcal{L}^*\varphi\in L^1(\mathbb{R}^d)$,

$$\begin{aligned}\lim_{k\to\infty}\int_{\mathbb{R}^d}\mathcal{L}u_{n_k}(t,x)\varphi(x)\,dx &= \lim_{k\to\infty}\int_{\mathbb{R}^d} u_{n_k}(t,x)\mathcal{L}^*\varphi(x)\,dx\\ &= \int_{\mathbb{R}^d} u(t,x)\mathcal{L}^*\varphi(x)\,dx = \int_{\mathbb{R}^d}\mathcal{L}u(t,x)\varphi(x)\,dx.\end{aligned}$$

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 $t\in\mathcal{T}$.

⋄ *Part* (iii). By Part (i) and Proposition 5.4, $\partial_t u\in C_b(\mathcal{T}\times\mathbb{R}^d)$. Since u is a bounded classical solution and $F'\ge\kappa$, we also have $\mathcal{L}u = F^{-1}(-\partial_t u - f)\in C_b(\mathcal{T}\times\mathbb{R}^d)$. Moreover, $\|\mathcal{L}u\|_\infty \le F^{-1}\big(T\|\partial_t f\|_\infty + \|F(\mathcal{L}g)\|_\infty + 2\|f\|_\infty\big)$.

By Theorem 5.3 (ii), we have $\|u\|_\infty\le T\|f\|_\infty + \|g\|_\infty$. Thus, by interior regularity estimates (which are uniform in x, see Definition 2.4), for every $t\in\mathcal{T}$,

$$\begin{aligned}\|\mathcal{L}u\|_{\mathcal{C}^{\alpha/2\sigma,\alpha}([0,t]\times\mathbb{R}^d)} &\le \|\mathcal{L}u\|_\infty + \sup_{x\in\mathbb{R}^d}\Big([\mathcal{L}u]_{\mathcal{C}^{\alpha/2\sigma,\alpha}([0,t]\times B_1(x))}\Big)\\ &\le \widetilde{C}(t)\Big(\|f\|_{\alpha/2\sigma,\alpha} + \|\partial_t f\|_\infty + \|\mathcal{L}g\|_\infty + \|g\|_\infty\Big).\end{aligned}$$

□

Corollary 5.6. *Assume (L′) or (L″), and (R), (A1′), (A2). If interior* $(\frac{\alpha}{2\sigma},\alpha)$*-regularity estimates hold for problem (4), then (S1), (S2), (S3), (S4) are satisfied.*

Proof. Condition (S1) follows from Theorem 5.5 (i), while (S2) follows from Theorem 5.5 (ii), and (S3), (S4) hold by Theorem 5.5 (iii). □

Remark 5.7. If instead of (R) we only assume $\mathcal{R}\subset\mathcal{C}_b^{\alpha/2\sigma,\alpha}(\mathcal{T}\times\mathbb{R}^d)\times BUC(\mathbb{R}^d)$ (uniformly bounded in an appropriate way) in Corollary 5.6, then we still obtain (S1) and (S2). We may get (S3) by assuming $F'\le K$ (i.e. F is globally Lipschitz). This is enough for existence in Theorem 7.5, but not for uniqueness in Theorem 7.7.

6. Fokker–Planck equations

6.1. **Existence.** In this section we prove existence for problem (5). We assume:

(B): $b\in C(\mathcal{T}\times\mathbb{R}^d)$ and $b(t,x)\in[0,B]$ for fixed $B\in[0,\infty)$ and every $(t,x)\in\mathcal{T}\times\mathbb{R}^d$.

For $b=F'(\mathcal{L}u)$ this is a consequence of (A1) and either (S3) or (S4) when u is a bounded classical solution of problem (4).

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

(i) m is a very weak solution of problem (5) (cf. Definition 2.2);

(ii) m satisfies (6) for every
$\phi \in \mathcal{U} = \{\phi \in C_b(\overline{\mathcal{T}} \times \mathbb{R}^d) : \ \partial_t \phi + b\mathcal{L}\phi \in C_b(\mathcal{T} \times \mathbb{R}^d)\}$;

(iii) m satisfies (6) for every[13]
$\phi \in \{\phi \in C_c^\infty(\overline{\mathcal{T}} \times \mathbb{R}^d) : \phi(t) = \psi \in C_c^\infty(\mathbb{R}^d)$ *for every* $t \in \overline{\mathcal{T}}\}$.

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

$$\sum_{n=1}^{N_k} \big(m(t_{n+1}^k) - m(t_n^k)\big)[\varphi(t_n^k)] = \sum_{n=1}^{N_k} \int_{t_n^k}^{t_{n+1}^k} m(\tau)\big[b(\tau)\mathcal{L}\varphi(t_n^k)\big]\, d\tau.$$

Notice that by the Lebesgue dominated convergence theorem we get

$$\lim_{k\to\infty} \sum_{n=1}^{N_k} \int_{t_n^k}^{t_{n+1}^k} m(\tau)\big[b(\tau)\mathcal{L}\varphi(t_n^k)\big]\, d\tau = \lim_{k\to\infty} \int_0^t m(\tau)\big[b(\tau)\mathcal{L}\varphi^k(\tau)\big]\, d\tau = \int_0^t m(\tau)\big[b(\tau)\mathcal{L}\varphi(\tau)\big]\, d\tau.$$

We also observe that

$$\sum_{n=1}^{N_k} \big(m(t_{n+1}^k) - m(t_n^k)\big)[\varphi(t_n^k)] = m(t)[\varphi(t)] - m_0[\varphi(0)] - \sum_{n=1}^{N_k} \Big(m(t_{n+1}^k)[\varphi(t_{n+1}^k) - \varphi(t_n^k)]\Big).$$

By the Taylor expansion, for some $\xi_n^k \in [t_n^k, t_{n+1}^k]$ we have

$$\varphi(t_{n+1}^k) - \varphi(t_n^k) = \partial_t \varphi(t_{n+1}^k)(t_{n+1}^k - t_n^k) - \partial_t^2 \varphi(\xi_n^k) \frac{(t_{n+1}^k - t_n^k)^2}{2}.$$

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

$$\lim_{k\to\infty} \sum_{n=1}^{N_k} \Big(m(t_{n+1}^k)[\varphi(t_{n+1}^k) - \varphi(t_n^k)]\Big) = \int_0^t m(\tau)[\partial_t \varphi(\tau)]\, d\tau.$$

By combining these arguments we obtain

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

□

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

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

[13]In this set functions are constant in time.

If $\bigcup_\lambda \{m_{0,\lambda}, (\nu_\lambda)|_{B_1^c}\}$ *is tight and* $\sup_\lambda \big(\|b_\lambda\|_\infty + \|\mathcal{L}_\lambda\|_{LK}\big) < \infty$,[14] *then*

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

$$\sup\Big\{\sup_{t\in\overline{\mathcal{T}}} m(t)(K_\varepsilon^c) : m \in \textstyle\bigcup_\lambda \mathcal{M}_\lambda\Big\} \le \varepsilon;$$

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

$$\|m(t) - m(s)\|_0 \le \sup_\lambda \Big(2 + \big(2\sqrt{T} + K_d\big)\|b_\lambda\|_\infty \|\mathcal{L}_\lambda\|_{LK}\Big)\sqrt{|t-s|};$$

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

Proof. ⋄ *Part* (i). Let $V(x) = V_0\big(\sqrt{1+|x|^2}\big)$ be a Lyapunov function for which we have $\sup_\lambda \big(m_{0,\lambda}[V] + \|\mathcal{L}_\lambda V\|_\infty\big) < \infty$ (see Lemma 4.9, Lemma 4.11, Corollary 4.12). For $n \in \mathbb{N}$, let $V_{n,0} \in C_b^2\big([0,\infty)\big)$ be such that

$$V_{n,0}(t) = \begin{cases} V_0(t) & \text{for } t \le n, \\ V_0\big(\sqrt{1+(n+1)^2}\big) & \text{for } t \ge n+2, \end{cases}$$

and additionally

$$0 \le V_{n,0}' \le V_0' \qquad \text{and} \qquad |V_{n,0}''| \le |V_0''|. \tag{21}$$

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

$$m(t)[V_n] = m_{0,\lambda}[V_n] + \int_0^t m(\tau)[b_\lambda(\tau)\mathcal{L}_\lambda V_n]\,d\tau. \tag{22}$$

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

$$\lim_{n\to\infty} \big(V_n, \nabla V_n, D^2 V_n\big)(x) = \big(V, \nabla V, D^2 V\big)(x) \quad \text{for every } x \in \mathbb{R}^d. \tag{23}$$

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 (23), by the Lebesgue dominated convergence theorem — we use Lemma 4.11 *(iv)* for the integral on B_1^c and (21) otherwise — we may pass to the limit in (22). For every $t \in \overline{\mathcal{T}}$, λ, and $m \in \mathcal{M}_\lambda$ we obtain

$$m(t)[V] = m_{0,\lambda}[V] + \int_0^t m(\tau)[b_\lambda \mathcal{L}_\lambda V]\,d\tau \le m_{0,\lambda}[V] + \|b_\lambda\|_\infty \|\mathcal{L}_\lambda V\|_\infty T. \tag{24}$$

Thus, by Proposition 4.8, for every $\varepsilon > 0$ there exists a compact set K_ε such that

$$\sup\Big\{m(t)(K_\varepsilon^c) : t \in \overline{\mathcal{T}},\ m \in \textstyle\bigcup_\lambda \mathcal{M}_\lambda\Big\} \le \varepsilon.$$

⋄ *Part* (ii). Consider $\phi_\varepsilon = \phi * \rho_\varepsilon$, where $\phi \in \mathcal{C}_b^1(\mathbb{R}^d)$ is such that $\|\phi\|_\infty \le 1$ and $[\phi]_1 \le 1$, and ρ_ε is a standard mollifier. Then $\|\phi - \phi_\varepsilon\|_\infty \le \varepsilon$ and, by Proposition 4.13, $\|\mathcal{L}\phi_\varepsilon\|_\infty \le \|\mathcal{L}\|_{LK}\|\phi_\varepsilon\|_{C_b^2(\mathbb{R}^d)}$. By Definition 2.2, for every λ and $m \in \mathcal{M}_\lambda$,

$$\begin{aligned} \big|\big(m(t) - m(s)\big)[\phi]\big| &= \big|\big(m(t) - m(s)\big)[\phi - \phi_\varepsilon] + \big(m(t) - m(s)\big)[\phi_\varepsilon]\big| \\ &\le 2\varepsilon + \left|\int_s^t \int_{\mathbb{R}^d} (\mathcal{L}_\lambda \phi_\varepsilon)(x) b_\lambda(\tau, x)\, m(\tau, dx)\, d\tau\right| \\ &\le 2\varepsilon + \|b_\lambda\|_\infty \|\mathcal{L}_\lambda\|_{LK} \|\phi_\varepsilon\|_{C_b^2(\mathbb{R}^d)} |t-s|. \end{aligned}$$

We also have

$$\|\phi_\varepsilon\|_{C_b^2(\mathbb{R}^d)} \le \left(\|\phi\|_\infty + \|\nabla\phi\|_\infty + \frac{K_d\|\nabla\phi\|_\infty}{\varepsilon}\right) \le \frac{2\varepsilon + K_d}{\varepsilon}.$$

[14]See (16) for the definition of $\|\cdot\|_{LK}$.

By taking $\varepsilon = \sqrt{|t-s|}$, we thus obtain

$$\|m(t)-m(s)\|_0 \le \sup_\lambda \Big(2 + \big(2\sqrt{T}+K_d\big)\|b_\lambda\|_\infty \|\mathcal{L}_\lambda\|_{LK}\Big)\sqrt{|t-s|}.$$

⋄ *Part (iii).* It follows from Part (*i*) that the set $\{m(t) : m \in \bigcup_\lambda \mathcal{M}_\lambda\}$ is pre-compact for a fixed $t \in \overline{\mathcal{T}}$. Then, in Part (*ii*), we showed that the family $\bigcup_\lambda \mathcal{M}_\lambda$ is equicontinuous in $C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$. Hence $\bigcup_\lambda \mathcal{M}_\lambda \subset C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ is pre-compact by the Arzelà–Ascoli theorem [47, §7 Theorem 17].

□

In the general case we are unable to prove uniqueness of solutions of problem (5). However, we can make the following observation about the sets of solutions.

Corollary 6.3. *Assume (L), (B), (A3). If $\mathcal{M} \subset C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ is the set of solutions of problem (5) corresponding to (b, m_0), then $\mathcal{M}$ is convex, compact, and*

$$\sup_{m\in\mathcal{M}}\sup_{t\in\overline{\mathcal{T}}} m(t)[V] \le c_1, \qquad \sup_{m\in\mathcal{M}}\sup_{0<|t-s|\le T} \frac{\|m(t)-m(s)\|_0}{\sqrt{|t-s|}} \le c_2,$$

for a Lyapunov function V such that $m_0[V], \|\mathcal{L}V\|_\infty < \infty$ (see Corollary 4.12), and

$$c_1 = m_0[V] + T\|b\|_\infty \|\mathcal{L}V\|_\infty, \qquad c_2 = 2 + \big(2\sqrt{T}+K_d\big)\|b\|_\infty\|\mathcal{L}\|_{LK}.$$

Proof. It follows from Definition 2.2 that $\mathcal{M}$ is convex (the equation is linear), as well as that if $\{m_n\} \subset \mathcal{M}$ and $m_n \to \widehat{m}$ in $C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$, then $\widehat{m} \in \mathcal{M}$, i.e. the set $\mathcal{M}$ is closed. Hence, by Lemma 6.2 (*iii*), we obtain that $\mathcal{M} \subset C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ is compact. The specified bounds follow from Lemma 6.2 (*ii*) and (24). □

We now prove a kind of a stability result for solutions (in terms of semicontinuity with respect to upper Kuratowski limits (see [53, §29.III]).

Lemma 6.4. *Assume (L), (A3), and $\{b_n, b\}_{n\in\mathbb{N}}$ satisfy (B) with a uniform bound by B. Let $\{\mathcal{M}_n, \mathcal{M}\}$ be the corresponding sets of solutions of problem (5) 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 $C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$.*

Proof. By Lemma 6.2 (*iii*) the set $\bigcup_n \mathcal{M}_n \subset C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ 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

$$\sup_{n\in\mathbb{N}}\sup_{m\in\mathcal{M}_n}\sup_{t\in\overline{\mathcal{T}}} m(t)(K_\varepsilon^c) \le \varepsilon.$$

Let $\{m_{n_k}\} \subset \{m_n\}$ be a convergent 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 $\varphi \in C_c^\infty(\mathbb{R}^d)$ we have

$$\left|\int_0^t (b_n m_n - bm)(\tau)[\mathcal{L}\varphi]\,d\tau\right| = \left|\int_0^t \Big((b_n-b)m_n + b(m_n-m)\Big)(\tau)[\mathcal{L}\varphi]\,d\tau\right|.$$

Since $m_n \to m$ in $C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ and $b \in C_b(\mathcal{T}\times\mathbb{R}^d)$, we notice that

$$\lim_{n\to\infty}\sup_{\tau\in\mathcal{T}} \big|m_n(\tau)[b(\tau)] - m(\tau)[b(\tau)]\big| = 0.$$

Next,

$$\left|\int_0^t (b_n-b)m_n(\tau)[\mathcal{L}\varphi]\,d\tau\right| \le \|\mathcal{L}\varphi\|_\infty \int_0^T \int_{K_\varepsilon\cup K_\varepsilon^c} |b_n-b|(\tau,x)\,m_n(\tau,dx)\,d\tau$$
$$\le \|\mathcal{L}\varphi\|_\infty\bigg(\varepsilon T\big(\|b_n\|_\infty + \|b\|_\infty\big) + \int_0^T \sup_{x\in K_\varepsilon}\big|b_n(\tau,x)-b(\tau,x)\big|\,d\tau\bigg).$$

We have $\big|b_n(t,x)-b(t,x)\big| \le 2B$ 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 $\sup_{x\in K_\varepsilon}\big|b_n(t,x)-b(t,x)\big| \to 0$ pointwise in $t\in\mathcal{T}$. Thus, by Lebesgue dominated convergence theorem,

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

Since $\varepsilon>0$ may be arbitrarily small and m_n are solutions of problem (5), because of Lemma 6.1 *(iii)*,

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

Thus m is a solution of problem (5) with parameters b and m_0, i.e. $m\in\mathcal{M}$. □

Remark 6.5. When the solutions of problem (5) are unique, Lemma 6.4 is a standard stability result. Indeed, let $\{m_n, m\}$ be (the unique) solutions of problem (5) 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 $\{m_n\}$ has a further subsequence convergent to m. Thus $m_n\to m$ in $C(\overline{\mathcal{T}},\mathcal{P}(\mathbb{R}^d))$.

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

Theorem 6.6. *Assume (L), (B), (A3). Problem (5) has a very weak solution.*

Proof. ⋄ *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 4.15 and ν^ε, $\mathcal{L}^{\varepsilon\,*}$ be their Lévy measures and adjoint operators, respectively.

By (17) and the Fubini theorem, for every $\mu\in L^1(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d}\mathcal{L}^{\varepsilon\,*}\mu\,dx = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\big(\mu(x-z)-\mu(x)\big)\,dx\,\nu^\varepsilon(dz) = 0. \tag{25}$$

Let $b_\varepsilon = b+\varepsilon$ and $\mu_{0,\varepsilon} = m_0 * \rho_\varepsilon$, where $\{\rho_\varepsilon\}_{\varepsilon\in(0,1)}$ is the sequence of standard mollifiers. For $\varepsilon\in(0,1)$ we consider the following family of problems

$$\begin{cases}\partial_t\mu = \mathcal{L}^{\varepsilon\,*}(b_\varepsilon\mu) & \text{on } \mathcal{T}\times\mathbb{R}^d,\\ \mu(0) = \mu_{0,\varepsilon} & \text{on } \mathbb{R}^d.\end{cases} \tag{26}$$

⋄ *Step 2. Existence of approximate solution μ_ε.* For $\mu\in C(\overline{\mathcal{T}},L^1(\mathbb{R}^d))$, define

$$\mathcal{G}_\varepsilon(\mu)(t) = \mu(0) + \int_0^t \mathcal{L}^{\varepsilon\,*}\big(b_\varepsilon\mu\big)(\tau)\,d\tau. \tag{27}$$

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

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

is a bounded linear operator.

Let $\mu_1,\mu_2\in C(\overline{\mathcal{T}},L^1(\mathbb{R}^d))$ be such that $\mu_1(0)=\mu_2(0)$ and take $t_\varepsilon = \frac{\varepsilon^3}{4\,c_{\mathcal{L}}\|b_\varepsilon\|_\infty}$, where $c_{\mathcal{L}}$ is the constant given by Lemma 4.15. Then, because of Lemma 4.15 *(i)*,

$$\begin{aligned}\sup_{t\in[0,t_\varepsilon]}\|\mathcal{G}_\varepsilon(\mu_1-\mu_2)(t)\|_{L^1(\mathbb{R}^d)} &= \sup_{t\in[0,t_\varepsilon]}\left\|\int_0^t\mathcal{L}^{\varepsilon\,*}\big(b_\varepsilon(\mu_1-\mu_2)\big)(\tau)\,d\tau\right\|_{L^1(\mathbb{R}^d)}\\ &\le t_\varepsilon\frac{2\,c_{\mathcal{L}}\|b_\varepsilon\|_\infty}{\varepsilon^3}\sup_{t\in[0,t_\varepsilon]}\|\mu_1-\mu_2\|_{L^1(\mathbb{R}^d)} \le \frac{1}{2}\sup_{t\in[0,t_\varepsilon]}\|\mu_1-\mu_2\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

Therefore, by the Banach fixed point theorem, problem (26) has a unique solution $\mu_\varepsilon\in C\big([0,t_\varepsilon],L^1(\mathbb{R}^d)\big)$ for every $\varepsilon>0$. Since $t_\varepsilon>0$ is constant for fixed $\varepsilon>0$, we may immediately extend this solution to the interval $\overline{\mathcal{T}}$ and conclude that problem (26) has a unique solution in the space $C(\overline{\mathcal{T}},L^1(\mathbb{R}^d))\cap C^1\big(\mathcal{T},L^1(\mathbb{R}^d)\big)$.

◇ *Step 3. Compactness of* $\{\mu_\varepsilon\}$ *in* $C(\overline{\mathcal{T}},\mathcal{P}(\mathbb{R}^d))$.[15] Because of the regularity of μ_ε obtained in Step 2, we have

$$\partial_t\mu_\varepsilon = \mathcal{L}^{\varepsilon\,*}(b_\varepsilon\mu_\varepsilon) \quad \text{in} \quad C\big(\mathcal{T},L^1(\mathbb{R}^d)\big). \tag{28}$$

Consider $\operatorname{sgn}(u)^- = \mathbb{1}_{\{u<0\}}$. Then, by (28),

$$\int_0^t\int_{\mathbb{R}^d}\partial_t\mu_\varepsilon\operatorname{sgn}(\mu_\varepsilon)^-\,dx\,d\tau = \int_0^t\int_{\mathbb{R}^d}\mathcal{L}^{\varepsilon\,*}(b_\varepsilon\mu_\varepsilon)\operatorname{sgn}(\mu_\varepsilon)^-\,dx\,d\tau.$$

Since $b_\varepsilon>0$, we have $\operatorname{sgn}(\mu_\varepsilon)^- = \operatorname{sgn}(b_\varepsilon\mu_\varepsilon)^-$ and for arbitrary real functions u,v, $v\operatorname{sgn}(u)^-\geq v\operatorname{sgn}(v)^- = -(v)^-$. Therefore

$$\begin{aligned}\big(\mathcal{L}^{\varepsilon\,*}(b_\varepsilon\mu_\varepsilon)\operatorname{sgn}(\mu_\varepsilon)^-\big)(x) &= \int_{\mathbb{R}^d}\big(b_\varepsilon\mu_\varepsilon(x-z)-b_\varepsilon\mu_\varepsilon(x)\big)\operatorname{sgn}(b_\varepsilon\mu_\varepsilon)^-(x)\,\nu^\varepsilon(dz)\\ &\geq -\int_{\mathbb{R}^d}\big((b_\varepsilon\mu_\varepsilon)^-(x-z)-(b_\varepsilon\mu_\varepsilon)^-(x)\big)\,\nu^\varepsilon(dz) = -\mathcal{L}^{\varepsilon\,*}\big((b_\varepsilon\mu_\varepsilon)^-\big)(x).\end{aligned}$$

By (25), $\int_{\mathbb{R}^d}\mathcal{L}^{\varepsilon\,*}\big((b_\varepsilon\mu_\varepsilon)^-\big)\,dx = 0$. Hence

$$\begin{aligned}0 &\leq \int_0^t\int_{\mathbb{R}^d}\partial_t\mu_\varepsilon\operatorname{sgn}(\mu_\varepsilon)^-\,dx\,d\tau = \int_0^t\int_{\mathbb{R}^d}-\partial_t(\mu_\varepsilon)^-\,dx\,d\tau\\ &= \int_{\mathbb{R}^d}(\mu_{0,\varepsilon})^-\,dx - \int_{\mathbb{R}^d}(\mu_\varepsilon)^-(t)\,dx.\end{aligned}$$

Since $\mu_{0,\varepsilon} = m_0*\rho_\varepsilon\geq 0$, 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 $t\in\overline{\mathcal{T}}$.

By Step 2, μ_ε is the fixed point of $\mathcal{G}_\varepsilon$. Thus, because of (25), (27), 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 $\mu_\varepsilon(t)\in\mathcal{P}_{ac}(\mathbb{R}^d)$ for every $t\in\overline{\mathcal{T}}$. Since $\mu_\varepsilon\in C(\overline{\mathcal{T}},L^1(\mathbb{R}^d))$, it follows that $\mu_\varepsilon\in C\big(\overline{\mathcal{T}},\mathcal{P}_{ac}(\mathbb{R}^d)\big)$.

Notice that $\|b_\varepsilon\|_\infty\leq\|b+1\|_\infty<B+1$ by (B). Let V be a Lyapunov function such that $m_0[V],\|\mathcal{L}V\|_\infty<\infty$ (see Corollary 4.12). By Definition 4.6,

$$\mu_{0,\varepsilon}[V] = (m_0*\rho_\varepsilon)[V]\leq m_0[V]+\|\nabla V\|_\infty\int_{B_1}|z|\,\rho_\varepsilon(z)\,dz\leq m_0[V]+1.$$

In combination with Lemma 4.15 *(iii)* we get

$$\sup_{\varepsilon\in(0,1)}\big(b_\varepsilon+\mu_{0,\varepsilon}[V]+\|\mathcal{L}^\varepsilon V\|_\infty+\|\mathcal{L}^\varepsilon\|_{LK}\big)<\infty.$$

It follows from Lemma 6.2 that the family $\{\mu_\varepsilon\}$ is pre-compact in $C(\overline{\mathcal{T}},\mathcal{P}(\mathbb{R}^d))$.

◇ *Step 4. Passing to the limit.* Using the result of Step 3, let ε_k be a sequence such that $\mu_{\varepsilon_k}\to m$ in $C(\overline{\mathcal{T}},\mathcal{P}(\mathbb{R}^d))$. By (28), for every ε_k, $\varphi\in C_c^\infty(\mathbb{R}^d)$ and $s,t\in\overline{\mathcal{T}}$, because $b_{\varepsilon_k}-b=\varepsilon_k$,

$$\begin{aligned}\mu_{\varepsilon_k}(t)[\varphi]-\mu_{\varepsilon_k}(s)[\varphi] &= \int_s^t\int_{\mathbb{R}^d}\big(\mathcal{L}^{\varepsilon_k\,*}(\mu_{\varepsilon_k}b_{\varepsilon_k})\big)\varphi\,dx\,d\tau\\ &= \varepsilon_k\int_s^t\mu_{\varepsilon_k}[\mathcal{L}^{\varepsilon_k}\varphi]\,d\tau+\int_s^t\mu_{\varepsilon_k}\big[b(\mathcal{L}^{\varepsilon_k}\varphi-\mathcal{L}\varphi)\big]\,d\tau+\int_s^t\mu_{\varepsilon_k}[b\mathcal{L}\varphi]\,d\tau.\end{aligned}$$

[15] First we show that $\{\mu_\varepsilon\}\subset C(\mathcal{T},\mathcal{P}_{ac}(\mathbb{R}^d))$, then we establish its tightness.

Since $\lim_{k\to\infty} \|\mathcal{L}^{\varepsilon_k}\varphi - \mathcal{L}\varphi\|_\infty = 0$, by Lemma 4.15 (*ii*) and the Hölder inequality,

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

It follows that m is a very weak solution of problem (5) (see Lemma 6.1 (*iii*)).

□

6.2. **Uniqueness.** Uniqueness for problem (5) holds when b is more regular:

(B′): b satisfies (B); in addition, $b \in B(\mathcal{T}, \mathcal{C}_b^\beta(\mathbb{R}^d))$ for some $\beta > 0$.

This condition is valid for $b = F'(\mathcal{L}u)$ when $F' \in \mathcal{C}^\gamma(\mathbb{R})$ with $\gamma > 0$ and $\mathcal{L}u$ is smooth (Theorem 5.5 (*iii*)).

Theorem 6.7. *Assume (B′) on $[0,t]$ for every $t \in \mathcal{T}$ and (A3). If either*

(i) (L′), $b \geq \kappa$ for some $\kappa > 0$, and $b \in UC\big([0,t]\times\mathbb{R}^d\big)$ for every $t \in \mathcal{T}$; or

(ii) (L″) and $b \geq \kappa$ for some $\kappa > 0$;

then problem (5) has precisely one very weak solution.

We show uniqueness of solutions of problem (5) using a Holmgren-type argument. The idea is to use a solution of the "dual" equation,

$$\begin{cases} \partial_t w - b\mathcal{L}w = 0 & \text{on } \mathcal{T}\times\mathbb{R}^d, \\ w(0) = \phi & \text{on } \mathbb{R}^d, \end{cases} \tag{29}$$

as a test function in Definition 2.2. For simplicity we consider a forward-in-time problem and then reverse time in the proof of Theorem 6.7. We need sufficient regularity of solutions (see Lemma 6.1) when ϕ is taken from a dense subset of $C(\mathbb{R}^d)$. Because of the non-degeneracy of the operator $\mathcal{L}$ and the standard uniform ellipticity assumption $b \geq \kappa > 0$, existing results suffice to conclude.[16]

Lemma 6.8. *Under the assumptions of Theorem 6.7 there exists a bounded classical solution of problem (29).*

Proof. ⋄ *Part* (*i*). The statement follows from [61, Theorem 5.1.9] (see [61, page 175] for relevant notation).

⋄ *Part* (*ii*). Because $\phi \in C_c^\infty(\mathbb{R}^d)$, we have $\mathcal{L}\phi \in C_b^\infty(\mathbb{R}^d)$ and thus by (B′) we get $b\mathcal{L}\phi \in C_b(\mathcal{T}, \mathcal{C}_b^\beta(\mathbb{R}^d))$. Notice that w is a bounded classical solution of problem (29) if and only if $v = w - \phi$ is a bounded classical solution of

$$\begin{cases} \partial_t v - b\mathcal{L}v = b\mathcal{L}\phi & \text{on } \mathcal{T}\times\mathbb{R}^d, \\ v(0) = 0 & \text{on } \mathbb{R}^d. \end{cases} \tag{30}$$

We study problem (30) using the results in [63]. We write $b\mathcal{L} = A + B$, where

$$(A\phi)(t,x) = \int_{\mathbb{R}^d} \big(\phi(t,x+z) - \phi(t,x) - \mathbb{1}_{[1,2)}(2\sigma)\, z\cdot\nabla\phi(x)\big) b(t,x) \frac{\widetilde{k}(z)}{|z|^{d+2\sigma}}\, dz, \tag{31}$$

$\widetilde{k}(z) = \mathbb{1}_{B_1} k(z) + \mathbb{1}_{B_1^c} k(\frac{z}{|z|})$ is a normal extension of k (defined in (L″)) to $\mathbb{R}^d$, and $B = b\mathcal{L} - A : C_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ is a bounded operator (with Lévy measure supported on B_1^c). We check the assumptions for operators A and B given by (31). Assumption **A** in [63] is satisfied, because we assume (L″), (B′), and $b \geq \kappa > 0$. To verify assumptions **B1** and **B2** in [63], we choose $c(t,x,v) = v$, $U_n = B_1$, and $\pi = \nu|_{B_1^c}$ (in the notation of [63]) and again use (L″), (B′).

[16] In [25] we prove uniqueness for a degenerate case of the Fokker–Planck equation (5).

By [63, Theorem 4] there exists a unique solution v of problem (30) such that $\mathcal{L}v \in B(\mathcal{T}, \mathcal{C}_b^\beta(\mathbb{R}^d))$ and $\partial_t v \in C_b(\mathcal{T} \times \mathbb{R}^d)$ (see [63, Definition 3]). Thus $w = v - \phi$ is a bounded classical solution of problem (29).

□

Proof of Theorem 6.7. Existence of a very weak solution follows by Theorem 6.6. Fix arbitrary $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $t_0 \in (0, T]$, and take $\widetilde{b}(t) = b(t_0 - t)$ for every $t \in [0, t_0]$. Replace b by $\widetilde{b}$ in problem (29). Then there exists a bounded classical solution $\widetilde{w}$ of problem (29) — by Lemma 6.8. Let $w(t) = \widetilde{w}(t_0 - t)$ for $t \in [0, t_0]$. Then $\partial_t w, \mathcal{L}w \in C\big((0, t_0) \times \mathbb{R}^d\big)$, and w is a bounded classical solution of

$$\begin{cases} \partial_t w(t) + b(t)\mathcal{L}w(t) = 0 & \text{in } (0, t_0) \times \mathbb{R}^d, \\ w(t_0) = \varphi. \end{cases} \tag{32}$$

Suppose m and $\widehat{m}$ are two very weak solutions of problem (5) with the same initial condition m_0 and coefficient b. By Definition 2.2 (see Lemma 6.1 (*ii*)) and (32),

$$\big(m(t_0) - \widehat{m}(t_0)\big)[\varphi] = \int_0^{t_0} \big(m(\tau) - \widehat{m}(\tau)\big)\big[\partial_t w + b\mathcal{L}(w)\big]\, d\tau = 0.$$

Hence, for every $t \in (0, T]$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$, $(m(t) - \widehat{m}(t))[\varphi] = 0$, which means that $m(t) = \widehat{m}(t)$ in $\mathcal{P}(\mathbb{R}^d)$. □

Corollary 6.9. *Assume (A2), (A3), (R). Condition (S5) is satisfied if either*

(i) (A1′) and (L′) *or* *(ii) (A1″) and (L″)*

Proof. Let $u_1, u_2 \in \mathcal{S}_{HJB}$ and $v_1 = \mathcal{L}u_1$, $v_2 = \mathcal{L}u_2$. Since $F' \in \mathcal{C}^\gamma(\mathbb{R})$ by (A1), we may consider

$$b(t, x) = \int_0^1 F'\big(sv_1(t, x) + (1 - s)v_2(t, x)\big)\, ds.$$

Because $u_1, u_2 \in \mathcal{S}_{HJB}$ and $F' \geq 0$, we have $b \in C\big(\mathcal{T} \times \mathbb{R}^d\big)$ and $b \geq 0$.

⋄ *Part (i).* By Lemma 2.5 and Theorem 5.5 (*iii*), $v_1, v_2 \in C_b(\mathcal{T} \times \mathbb{R}^d)$ and $v_1, v_2 \in B\big([0, t], \mathcal{C}_b^\alpha(\mathbb{R}^d)\big) \cap UC\big([0, t] \times \mathbb{R}^d\big)$ for every $t \in \mathcal{T}$. Thus b satisfies (B′) on $[0, t]$ with $\beta = \gamma\alpha$ and $b \in UC\big([0, t] \times \mathbb{R}^d\big)$. Since $F' \geq \kappa > 0$, we have $b \geq \kappa > 0$ and (S5) follows from Theorem 6.7 (*i*).

⋄ *Part (ii).* By Lemma 2.7 and Theorem 5.5 (*iii*), $v_1, v_2 \in C_b(\mathcal{T} \times \mathbb{R}^d)$ and $v_1, v_2 \in B\big([0, t], \mathcal{C}_b^\alpha(\mathbb{R}^d)\big)$ for every $t \in \mathcal{T}$. Thus b satisfies (B′) on $[0, t]$ with $\beta = \gamma\alpha$. Since $F' \geq \kappa > 0$, we have $b \geq \kappa > 0$ and (S5) follows from Theorem 6.7 (*ii*).

□

7. The Mean Field Game system

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

Definition 7.1. 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 7.2. A set-valued map $\mathcal{K} : X \to 2^Y$ is upper-semicontinuous if, for each open set $A \subset Y$, the set $\mathcal{K}^{-1}(2^A) = \{x : \mathcal{K}(x) \subset A\}$ is open.

Theorem 7.3 (Kakutani–Glicksberg–Fan [36, §7 Theorem 8.4])**.** *Let $\mathcal{S}$ be a convex subset of a normed space and $\mathcal{K} : \mathcal{S} \to 2^{\mathcal{S}}$ be a compact set-valued map. If $\mathcal{K}$ is upper-semicontinuous with non-empty compact convex values, then $\mathcal{K}$ has a fixed point, i.e. there exists $x \in \mathcal{S}$ such that $x \in \mathcal{K}(x)$.* □

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

Lemma 7.4 ([54, §43.II Theorem 1])**.** *Let X be a Hausdorff space and Y a compact metric space. A set-valued compact map $\mathcal{K} : X \to 2^Y$ is upper-semicontinuous if and only if the conditions*

$$x_n \to x \text{ in } X, \qquad y_n \to y \text{ in } Y, \qquad \text{and} \qquad y_n \in \mathcal{K}(x_n)$$

imply $y \in \mathcal{K}(x)$. □

Theorem 7.5. *Assume (L), (A1), (A3), (A4), (S1), (S2), (S3). Then there exists a classical–very weak solution of problem (1).*

Proof. Let $X = \big(C(\overline{\mathcal{T}}, \mathcal{M}_b(\mathbb{R}^d)), \sup_t \|\cdot\|_0\big)$ (see Definition 4.4). 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 $\mathcal{K} : \mathcal{S} \to 2^{\mathcal{S}}$ on a certain compact, convex set $\mathcal{S} \subset X$. Then the map $\mathcal{K}$ is automatically compact and we may use Lemma 7.4 to obtain upper-semicontinuity.

⋄ *Step 1.* Let V be a Lyapunov function such that $m_0[V], \|\mathcal{L}V\|_\infty < \infty$ (see Corollary 4.12). Define

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

where m_0 is fixed and satisfies (A3), and

$$c_1 = m_0[V] + TK_{HJB}\|\mathcal{L}V\|_\infty, \qquad c_2 = 2 + \big(2\sqrt{T} + K_d\big)K_{HJB}\|\mathcal{L}\|_{LK}.$$

The set $\mathcal{S}$ is clearly convex. In addition, $\mathcal{S}$ is compact because of Proposition 4.8, the assumed equicontinuity in time, and the Arzelà–Ascoli theorem.

⋄ *Step 2.* Take $\mu \in \mathcal{S}$ and let $f = \mathfrak{f}(\mu)$ and $g = \mathfrak{g}\big(\mu(T)\big)$. We define a map $\mathcal{K}_1 : \mathcal{S} \to C_b(\overline{\mathcal{T}} \times \mathbb{R}^d)$ by $\mathcal{K}_1(\mu) = u$, where u is the unique bounded classical solution of problem (4), corresponding to data (f, g). The map $\mathcal{K}_1$ is well-defined because of (S1), (S3), and Theorem 5.3. By (A1) we find that $b = F'(\mathcal{L}u)$ satisfies (B).

We define a set-valued map $\mathcal{K}_2$ by $\mathcal{K}_2(u) = \mathcal{M}$, where $\mathcal{M}$ is the set of very weak solutions of problem (5) corresponding to $b = F'(\mathcal{L}u)$. The set $\mathcal{M} \subset \mathcal{S} \subset C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ is convex, compact, and non-empty because of Corollary 6.3 and Theorem 6.6. Now we define the fixed point map $\mathcal{K}(\mu) = \mathcal{K}_2(\mathcal{K}_1(\mu)) = \mathcal{M}$. Because of its construction, $\mathcal{K} : \mathcal{S} \to 2^{\mathcal{S}}$ is a compact map with non-empty compact convex values.

⋄ *Step 3.* It remains to show that the map $\mathcal{K} : \mathcal{S} \to 2^{\mathcal{S}}$ is upper-semicontinuous. Let $\{\mu_n, \mu\}_{n\in\mathbb{N}} \subset \mathcal{S}$ be such that $\lim_{n\to\infty} \mu_n = \mu$ and let $\{u_n, u\} = \{\mathcal{K}_1(\mu_n), \mathcal{K}_1(\mu)\}$ be the corresponding solutions of problem (4), and $\{\mathcal{M}_n, \mathcal{M}\} = \{\mathcal{K}(\mu_n), \mathcal{K}(\mu)\}$ be the corresponding sets of solutions of problem (5).

Since $\lim_{n\to\infty} \mu_n = \mu$, by (A4), Theorem 5.3 *(ii)*[17], 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 $b_n = F'(\mathcal{L}u_n)$

[17]This result can be obtained directly for classical solutions (under (S3)) by a straightforward application of the maximum principle property of the Lévy operator $\mathcal{L}$.

and $b = F'(\mathcal{L}u)$, then by (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 (B) and are uniformly bounded, by (S3).

Consider a sequence $m_n \in \mathcal{M}_n$ and suppose it converges to some $\widehat{m} \in \mathcal{S}$. Then we use Lemma 6.4 to say that $\widehat{m} \in \mathcal{M}$. This proves that the map $\mathcal{K}$ is upper-semicontinuous by Lemma 7.4.

⋄ *Step 4.* We now use Theorem 7.3 to get a fixed point $\widehat{m} \in \mathcal{S}$ of the map $\mathcal{K}$. Because of how $\mathcal{K}$ is defined, we have $\widehat{m} \in \mathcal{K}(\widehat{m}) = \mathcal{K}_2(\mathcal{K}_1(\widehat{m}))$. Thus there exists $\widehat{u} = \mathcal{K}_1(\widehat{m})$, which is a bounded classical solution of problem (4) with $f = \mathfrak{f}(\widehat{m})$ and $g = \mathfrak{g}(\widehat{m}(T))$, and $\|F'(\mathcal{L}\widehat{u})\|_\infty \le K_{HJB}$ by (S3). Note that $\widehat{m}$ is a very weak solution of problem (5) with $\widehat{m}(0) = m_0$ and $b = F'(\mathcal{L}\widehat{u})$. This, in turn, means that the pair $(\widehat{u}, \widehat{m})$ is a classical–very weak solution of problem (1) (see Definition 2.3).

□

Remark 7.6. Adding assumption (S5) to Theorem 7.5, yields singleton-valued maps $\mathcal{K}_2 : \mathcal{S}_{HJB} \to 2^{\mathcal{S}}$ and $\mathcal{K} : \mathcal{S} \to 2^{\mathcal{S}}$, and hence both are continuous (see Step 3, Remark 6.5). To conclude we may then use the classical Schauder theorem [36, §6 Theorem 3.2] (a special case of the Kakutani–Glicksberg–Fan theorem, cf. Lemma 7.4).

Theorem 7.7. *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 2.3), and take $u = u_1 - u_2$, and $m = m_1 - m_2$. 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 2.3, u_1, u_2 are bounded classical solutions of problem (4), and by (S4), $\{\partial_t u_1, \partial_t u_2, \mathcal{L}u_1, \mathcal{L}u_2\} \subset C_b(\mathcal{T} \times \mathbb{R}^d)$. By (A1), $F'(v_1), F'(v_2) \in C_b(\mathcal{T} \times \mathbb{R}^d)$, thus $u \in \mathcal{U}$, where $\mathcal{U}$ is defined in Lemma 6.1 *(ii)*. Further, m_1, m_2 are very weak solutions of problem (5) and satisfy (6) for every $\phi \in \mathcal{U}$ by Lemma 6.1 *(ii)*. Hence,

$$
\begin{aligned}
&m(T)\big[u(T)\big] - m(0)\big[u(0)\big] \\
&\quad= \big(m_1(T) - m_2(T)\big)\big[u_1(T) - u_2(T)\big] - \big(m_1(0) - m_2(0)\big)\big[u_1(0) - u_2(0)\big] \\
&\quad= \int_0^T \Big(m_1\big[\partial_t u + F'(v_1)v\big] - m_2\big[\partial_t u + F'(v_2)v\big]\Big)(\tau)\, d\tau.
\end{aligned} \tag{33}
$$

As $m_1(0) = m_2(0) = m_0$, we have $m(0)\big[u(0)\big] = 0$ and, thanks to (A5),

$$m(T)[u(T)] = \big(m_1(T) - m_2(T)\big)\big[\mathfrak{g}\big(m_1(T)\big) - \mathfrak{g}\big(m_2(T)\big)\big] \le 0.$$

Hence by (33) we get

$$\int_0^T \Big(m_1\big[\partial_t u + F'(v_1)v\big] - m_2\big[\partial_t u + F'(v_2)v\big]\Big)(\tau)\, d\tau \le 0. \tag{34}$$

We further notice that $\partial_t u + F(v_1) - F(v_2) = \mathfrak{f}(m_2) - \mathfrak{f}(m_1)$. Then, by integrating this expression with respect to the measure m, we obtain

$$\int_0^T m\left[\partial_t u + F(v_1) - F(v_2)\right](\tau)\, d\tau = \int_0^T m\left[\mathfrak{f}(m_2) - \mathfrak{f}(m_1)\right] d\tau. \tag{35}$$

From (A2) we know that F is convex, thus

$$F(v_1) - F(v_2) \le F'(v_1)\, v \quad \text{and} \quad F(v_1) - F(v_2) \ge F'(v_2)\, v, \tag{36}$$

and since $m_1, m_2 \in C(\overline{\mathcal{T}}, \mathcal{P}(\mathbb{R}^d))$ are non-negative measures, by (35), (36) and (A5),

$$\begin{aligned}\int_0^T m_1\big[\partial_t u + F'(v_1)\,v\big](\tau)\,d\tau - \int_0^T m_2\big[\partial_t u + F'(v_2)\,v\big](\tau)\,d\tau \\ \geq \int_0^T (m_1-m_2)[\mathfrak{f}(m_2)-\mathfrak{f}(m_1)](\tau)\,d\tau \geq 0.\end{aligned} \tag{37}$$

Combining (34) and (37), we find that

$$\int_0^T m_1\big[\partial_t u + F'(v_1)\,v\big](\tau)\,d\tau - \int_0^T m_2\big[\partial_t u + F'(v_2)\,v\big](\tau)\,d\tau = 0.$$

Then, taking into account (35), we get

$$\begin{aligned}0 = \int_0^T \int_{\mathbb{R}^d} \big(F'(v_1)\,v - F(v_1) + F(v_2)\big)\,m_1(\tau,dx)\,d\tau \\ + \int_0^T \int_{\mathbb{R}^d} \big(F(v_1) - F(v_2) - F'(v_2)\,v\big)\,m_2(\tau,dx)\,d\tau.\end{aligned}$$

By (36), both functions under the integrals are non-negative and continuous, thus in particular

$$F(v_1) - F(v_2) - F'(v_1)(v_1-v_2) = 0 \quad \text{on } \operatorname{supp} m_1, \tag{38}$$

where by $\operatorname{supp} m_1$ we understand the support of m_1 taken as a measure on $\overline{\mathcal{T}} \times \mathbb{R}^d$.

Let $(t,x) \in \operatorname{supp} m_1$. If $v_1(t,x) \neq v_2(t,x)$, then by (38)

$$F'(v_1(t,x)) = \frac{F(v_1(t,x)) - F(v_2(t,x))}{v_1(t,x) - v_2(t,x)}.$$

This means that the tangent line to the graph of F at $v_1(t,x)$ and the secant line joining $F(v_1(t,x))$ and $F(v_2(t,x))$ coincide. By (A1) and (A2) both lines also coincide with the tangent at $v_2(t,x)$, thus $F'(v_1(t,x)) = F'(v_2(t,x))$. Of course if $v_1(t,x) = v_2(t,x)$, then $F'(v_1(t,x)) = F'(v_2(t,x))$ as well. Therefore m_1 can be written as a solution of problem (5) with $F'(v_2)$ in place of $F'(v_1)$. By (S5) we get $m_1 = m_2$. Then also $u_1 = u_2$ by Theorem 5.3. □

Appendix A. Proofs of some technical results

Outline of proof of Lemma 2.7. We employ the method of continuity, following the scheme of the proof of [52, Theorem 13.9.1], using the Schauder estimates of [30, Theorem 1.3], and the existence results for linear problems in [63, Theorem 4].

For $s \in [0,1]$ consider the family of problems

$$\begin{cases} -\partial_t u = (1-s)\mathcal{L}u + sF(\mathcal{L}u) + f, & f \in \mathcal{C}_b^{\alpha/2\sigma,\alpha}(\mathcal{T}\times\mathbb{R}^d) \\ u(T) = g. \end{cases} \tag{P_s}$$

with regularized initial data g. Let

$$\begin{aligned}S = \{s : P_s \text{ has a classical solution satisfying} \\ \text{the interior } (\tfrac{\alpha}{2\sigma},\alpha)\text{-regularity estimates}\} \subset [0,1].\end{aligned}$$

We have $0 \in S$ by [63, Theorem 4][18] and S is closed by [30, Theorem 1.3]. Next we show that S is open. Let u_0 be a solution of problem P_{s_0}. Consider the map

$$\Psi^s : \mathcal{C}_b^{\sigma+\alpha/2\sigma,2\sigma+\alpha}(\mathcal{T}\times\mathbb{R}^d) \to \mathcal{C}_b^{\sigma+\alpha/2\sigma,2\sigma+\alpha}(\mathcal{T}\times\mathbb{R}^d)$$

[18][63] shows well-posedness of a strong solution with the correct spatial regularity uniform in time. Given our assumptions, time regularity can then be obtained from the equation in the standard way.

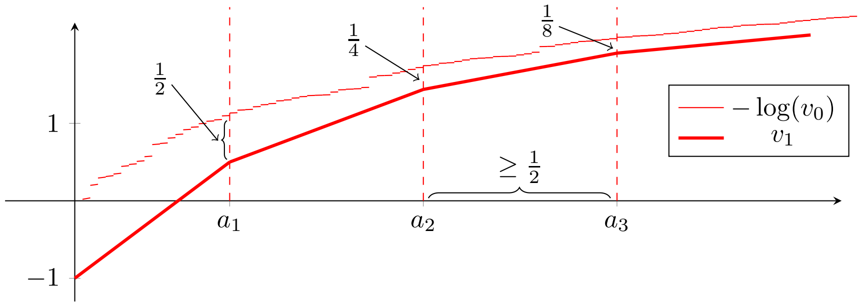

FIGURE 1. Comparison of $-\log(v_0)$ and v_1.

given by $\Psi^s(w) = v$, where v is a solution to the linear problem

$$\begin{cases} -\partial_t v = (1-s)\mathcal{L}v + sF'(\mathcal{L}u_0)\mathcal{L}v + s\big(F(\mathcal{L}w) - F'(\mathcal{L}u_0)\mathcal{L}w\big) + f, \\ v(T) = g. \end{cases}$$

It is well-defined because of [63, Theorem 4]. We use a second-order approximation as in [52, Theorem 13.9.1], and again [63, Theorem 4], to get that Ψ^s is a self-map on a certain neighbourhood of u_0. Then we show that Ψ^s is a contraction on this set. The fixed point given by the Banach theorem is a solution of problem P_s for $0 < |s_0 - s| < \varepsilon$. Note that the computations are essentially the same as in [52] because the problems depend linearly on the time derivatives.

The case of general data g follows by an approximation argument. □

Proof of Lemma 4.9. We proceed in steps, constructing successive functions that 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 (*ii*)], but a complete proof is lacking and the precise function v_0, which we need, cannot be extracted. Let

$$v(x) = v_0\big(|x|\big), \qquad \text{where} \qquad v_0(t) = \sup_{m\in\Pi} m\{x : |x| \geq t\}.$$

Then $v_0 : [0,\infty) \to [0,1]$ is non-increasing and $v_0(0) = 1$. Because Π is tight, we have $\lim_{t\to\infty} v_0(t) = 0$. Thus, $-\log(v_0) : [0,\infty) \to [0,\infty]$ is non-decreasing, $\log(v_0(0)) = 0$, and $\lim_{t\to\infty} -\log(v_0(t)) = \infty$. For $m \in \Pi$, let $\Phi^m(\tau) = m \circ v^{-1}\big([0,\tau)\big)$. Then,[19]

$$\begin{aligned}\Phi^m(\tau) = m\big(v^{-1}\big([0,\tau)\big)\big) &= m\big\{x : \forall\, \widehat{m} \in \Pi \quad \widehat{m}\{y : |y| \geq |x|\} < \tau\big\} \\ &\leq m\big\{x : m\{y : |y| \geq |x|\} < \tau\big\} \leq \tau.\end{aligned}$$

Integrating *by substitution* [10, Theorem 3.6.1] and *by parts* [10, Exercise 5.8.112],[20]

$$\int_{\mathbb{R}^d} -\log\big(v(x)\big)\, m(dx) = \int_0^1 -\log(\tau)\, d\Phi^m(\tau) = \int_0^1 \frac{\Phi^m(\tau)}{\tau}\, d\tau \leq \int_0^1 d\tau. \tag{39}$$

[19] Notice that $\{x : m\{y : |y| \geq |x|\} < \tau\} = \{x : |x| > r_\tau\}$, while $\{x : m\{y : |y| > |x|\} \leq \tau\} = \{x : |x| \geq r_\tau\}$, where r_τ 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 τ. Choosing the correct inequality in the definition of the function v_0 is essential.

[20] From [10, Exercise 5.8.112 (*i*)] we get $\int_r^1 -\log(\tau)\, d\Phi^m(\tau) = \int_r^1 \frac{\Phi^m(\tau)}{\tau}\, d\tau$ 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*)].

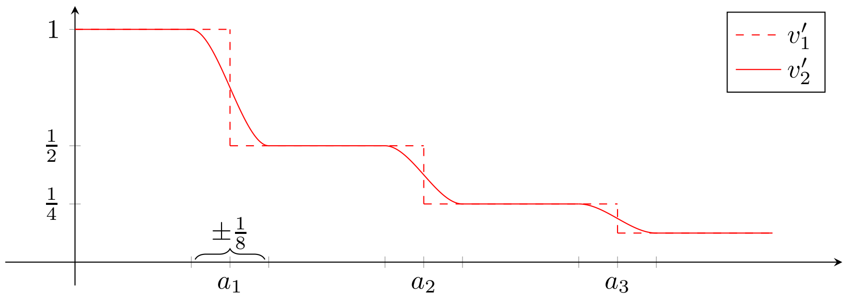

FIGURE 2. Comparison of v_1' and v_2'

◇ *Step 2. Continuity, concavity.*[21] 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)

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

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

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

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

$$-\log\big(v_0(a_n)\big) - v_1(a_n) = -\log\big(v_0(a_n)\big) - b_n = 2^{-n}$$

and on the interval $[a_n, a_{n+1}]$,

$$-\log(v_0) - v_1 \geq 2^{-n-1} \qquad \big(\text{hence } -\log\big(v_0(t)\big) \geq v_1(t) \text{ for every } t \geq 0\big).$$

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,

$$-\log\big(v_0(a_{n+1})\big) - l_n(a_{n+1}) \leq \liminf_{k\to\infty} \Big(-\log\big(v_0(s_k)\big) - l_n(s_k) \Big) \leq 2^{-n-1}.$$

Thus $-\log\big(v_0(a_{n+1})\big) - l_n(a_{n+1}) = 2^{-n-1}$, 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

$$v_1' = \sum_{n=0}^{N} 2^{-n}\mathbb{1}_{[a_n,a_{n+1})} \qquad (\text{a non-increasing function, see Figure 2}),$$

which implies that v_1 is concave. In addition, $v_1(t) \leq t-1$, hence $v_1(1) \leq 0$.

◇ *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)

$$v_2'(t) = \mathbb{1}_{[0,a_1-\frac{1}{8})}(t) + \sum_{n=1}^{N} 2^{-n}\Big(p\big(8(t-a_n)\big) + \mathbb{1}_{[a_n+\frac{1}{8},a_{n+1}-\frac{1}{8})}(t) \Big).$$

[21] Concavity serves as an intermediate step to obtain subadditivity.

Then $v_2 \in C^2\big([0,\infty)\big)$, v_2 is concave, increasing, and $\lim_{t\to\infty} v_2(t) = \infty$. Moreover,

$$\|v_2''\| \le \sup_t \left|\frac{1}{2}\frac{d}{dt}p(8t)\right| \le 3.$$

Next, we verify that $v_2 \le v_1$. Notice that for every $t \in [-1,1]$,

$$\int_{-1}^{t} p(s)\,ds \le \int_{-1}^{t} 2\cdot \mathbb{1}_{[-1,0]}(s) + \mathbb{1}_{[0,1]}(s)\,ds, \quad \text{and} \quad \int_{-1}^{1} p(s)\,ds = 3.$$

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

⋄ *Step 4. Subadditivity, bounds on derivatives.* Let $V_0 = \frac{1}{3}(v_2 + 1)$. 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 $\|V_0'\|_\infty, \|V_0''\|_\infty \le 1$. This proves that $V(x) = V_0\big(\sqrt{1+|x|^2}\big)$ is a Lyapunov function. By subadditivity and monotonicity,

$$V_0\big(\sqrt{1+t^2}\big) \le V_0(t+1) \le V_0(t) + V_0(1),$$

hence for every $m \in \Pi$, because $v_2 \le v_1 \le -\log(v_0)$ and by (39),

$$\begin{aligned} 0 &\le \int_{\mathbb{R}^d} V(x)\,m(dx) \le V_0(1) + \int_{\mathbb{R}^d} V_0\big(|x|\big)\,m(dx) \\ &\le \frac{v_2(1)+1}{3} + \frac{1}{3} - \frac{1}{3}\int_{\mathbb{R}^d} \log(v(x))\,m(dx) \le \frac{v_1(1)}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} \le 1. \end{aligned}$$

This shows that V is a Lyapunov function such that $m[V] \le 1$ for every $m \in \Pi$.

□

Appendix B. The Legendre–Fenchel transform

For a comprehensive treatment of the Legendre–Fenchel transform we refer to [70, 39]. Below we gather the particular properties of cost functions L needed to derive the model in Section 3, and corresponding to Hamiltonians F satisfying (A1) and (A2). These properties are expected, but in the setting we consider, we could not find the proofs in the literature.

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 $F(z) = \sup_{\zeta\in[0,\infty)}(z\zeta - L(\zeta))$.*[22] *Then F is convex and non-decreasing. In addition,*

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

(ii) if L is convex and is strictly convex on $\{L \ne \infty\}$, then F is differentiable on $\{F \ne \infty\}$ 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 ∂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 $|z_1 - z_2| \le 1$ we have

$(z_1 - z_2)(\zeta_1 - \zeta_2) \ge c_{z_1}|\zeta_1 - \zeta_2|^{1+\frac{1}{\gamma}}$,[23]

[22] Taking the supremum over $[0,\infty)$ is consistent with extending L by $L(\zeta) = \infty$ for $\zeta < 0$ and taking the supremum over all of $\mathbb{R}$ as is usual. Conversely, if $\lim_{z\to-\infty} F(z) \ne \infty$, taking $L(\zeta) = \sup_{z\in\mathbb{R}}\big(\zeta z - F(z)\big)$ results in $L(\zeta) = \infty$ for $\zeta < 0$. This is logical since ζ stands for the time rate (see Section 3) and the cost of going back in time should be prohibitive.

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

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

$$F(z+h) = \sup_{\zeta \in [0,\infty)} \big((z+h)\zeta - L(\zeta)\big) \geq \sup_{\zeta \in [0,\infty)} \big(z\zeta - L(\zeta)\big) = F(z).$$

◇ *Part* (*i*). Because $\lim_{\zeta \to \infty} L(\zeta)/\zeta = \infty$, for every $z \in \mathbb{R}$, $\lim_{\zeta \to \infty} \big(z - L(\zeta)/\zeta\big)\zeta = -\infty$. Since L is lower-semicontinuous and $L \not\equiv \infty$, there exists $\zeta_0 < \infty$ such that

$$L(\zeta_0) < \infty \qquad \text{and} \qquad \sup_{\zeta \in [0,\infty)} \Big(\big(z - \tfrac{L(\zeta)}{\zeta}\big)\zeta\Big) = z\zeta_0 - L(\zeta_0).$$

As a convex function with finite values, F is then locally Lipschitz-continuous.

◇ *Part* (*ii*). Since L is lower-semicontinuous, the statement follows from [70, Theorem 23.5, Corollary 23.5.1, Theorem 26.3, page 52].

◇ *Part* (*iii*). Note that (*iii*) implies (*i*) and (*ii*) (cf. [39, 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 [70, Theorem 23.5]. For $|z_1 - z_2| \leq 1$ we thus have

$$|z_1 - z_2||F'(z_1) - F'(z_2)| \geq c_{z_1}|F'(z_1) - F'(z_2)|^{1+\frac{1}{\gamma}}.$$

which gives us $F' \in \mathcal{C}^\gamma(\mathbb{R})$ (see (15) in Definition 4.1).

□

Acknowledgements

IC was supported by the INSPIRE faculty fellowship (IFA22-MA187). ERJ received funding from the Research Council of Norway under Grant Agreement No. 325114 "IMod. Partial differential equations, statistics and data: An interdisciplinary approach to data-based modelling". MK was supported by the Polish NCN grant 2016/23/B/ST1/00434 and Croatian Science Foundation grant IP-2018-01-2449. 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] R. J. Aumann, *Markets with a continuum of traders*, Econometrica, 32 (1964), pp. 39–50.

[4] O. E. Barndorff-Nielsen and J. Schmiegel, *Time change, volatility, and turbulence*, in Mathematical control theory and finance, Springer, 2008, pp. 29–53.

[5] O. E. Barndorff-Nielsen and A. Shiryaev, *Change of time and change of measure*, World Scientific Publishing, 2nd ed., 2015.

[6] A. Barrasso and N. Touzi, *Controlled diffusion mean field games with common noise and McKean-Vlasov second order backward SDEs*, Theory Probab. Appl., 66 (2022), pp. 613–639. Translation of Teor. Veroyatn. Primen. **66** (2021), 774–805.

[7] C. Benazzoli, L. Campi, and L. Di Persio, *ε-Nash equilibrium in stochastic differential games with mean-field interaction and controlled jumps*, Statist. Probab. Lett., 154 (2019), pp. 108522, 8.

[8] ———, *Mean field games with controlled jump-diffusion dynamics: existence results and an illiquid interbank market model*, Stochastic Process. Appl., 130 (2020), pp. 6927–6964.

[9] A. Bensoussan, J. Frehse, and P. Yam, *Mean field games and mean field type control theory*, Springer, 2013.

[23] When $\gamma = 1$ and c_{z_1} is in fact independent of z_1, this corresponds to the usual strong convexity of L (see [39, Theorem D.6.1.2]); if $\gamma_1 < \gamma_2$, the condition with γ_1 allows for a *flatter* (less non-affine) function L than the one with γ_2.

[10] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer, 2007.

[11] 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.

[12] V. S. Borkar, *Controlled diffusion processes*, Probab. Surv., 2 (2005), pp. 213–244.

[13] B. Böttcher, R. Schilling, and J. Wang, *Lévy matters. III*, Springer, 2013.

[14] F. Camilli and R. De Maio, *A time-fractional mean field game*, Adv. Differential Equations, 24 (2019), pp. 531–554.

[15] 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.

[16] 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.

[17] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications. I*, Springer, 2018.

[18] ———, *Probabilistic theory of mean field games with applications. II*, Springer, 2018.

[19] P. Carr and L. Wu, *Time-changed Lévy processes and option pricing*, Journal of Financial Economics, 71 (2004), pp. 113 – 141.

[20] 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.

[21] H. Chang-Lara and G. Dávila, *Regularity for solutions of non local parabolic equations*, Calc. Var. Partial Differential Equations, 49 (2014), pp. 139–172.

[22] ———, *Regularity for solutions of nonlocal parabolic equations. II*, J. Differential Equations, 256 (2014), pp. 130–156.

[23] E. Chasseigne and E. R. Jakobsen, *On nonlocal quasilinear equations and their local limits*, J. Differential Equations, 262 (2017), pp. 3759–3804.

[24] I. Chowdhury, O. Ersland, and E. R. Jakobsen, *On numerical approximations of fractional and nonlocal mean field games*, Found. Comput. Math., 23 (2023), pp. 1381–1431.

[25] I. Chowdhury, E. R. Jakobsen, and M. Krupski, *A strongly degenerate fully nonlinear mean field game with nonlocal diffusion*, in preparation, (2024).

[26] 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.

[27] J. C. Correa and E. A. Pimentel, *A Hessian-dependent functional with free boundaries and applications to mean-field games*, J. Geom. Anal., 34 (2024), pp. Paper No. 95, 21.

[28] M. G. Crandall, 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.

[29] M. F. Djete, *Mean field games of controls: On the convergence of Nash equilibria*, Ann. Appl. Probab., 33 (2023), pp. 2824 – 2862.

[30] 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.

[31] O. Ersland and E. R. Jakobsen, *On fractional and nonlocal parabolic mean field games in the whole space*, J. Differential Equations, 301 (2021), pp. 428–470.

[32] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Springer, 2nd ed., 2006.

[33] I. I. Gikhman and A. V. Skorokhod, *The theory of stochastic processes. II*, Springer, 2004. Reprint of the 1975 edition.

[34] D. A. Gomes, L. Nurbekyan, and E. A. Pimentel, *Economic models and mean-field games theory*, Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2015.

[35] D. A. Gomes, E. A. Pimentel, and V. Voskanyan, *Regularity theory for mean-field game systems*, Springer, 2016.

[36] A. Granas and J. Dugundji, *Fixed point theory*, Springer, 2003.

[37] 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.

[38] F. B. Hanson, *Applied stochastic processes and control for jump-diffusions: Modeling, analysis, and computation*, Society for Industrial and Applied Mathematics (SIAM), 2007.

[39] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of convex analysis*, Springer, 2001.

[40] M. Huang, P. E. Caines, and R. P. Malhamé, *An invariance principle in large population stochastic dynamic games*, J. Syst. Sci. Complex., 20 (2007), pp. 162–172.

[41] M. Huang, R. P. Malhamé, and P. E. Caines, *Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6 (2006), pp. 221–251.

[42] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland Publishing Co; Kodansha, 2nd ed., 1989.

[43] E. R. Jakobsen and K. H. Karlsen, *Continuous dependence estimates for viscosity solutions of integro-PDEs*, J. Differential Equations, 212 (2005), pp. 278–318.

[44] E. R. Jakobsen and A. Rutkowski, *The master equation for mean field game systems with fractional and nonlocal diffusions*, arXiv:2305.18867, (2023).

[45] P. Jameson Graber, V. Ignazio, and A. Neufeld, *Nonlocal Bertrand and Cournot mean field games with general nonlinear demand schedule*, J. Math. Pures Appl. (9), 148 (2021), pp. 150–198.

[46] A. S. Kechris, *Classical descriptive set theory*, Springer, 1995.

[47] J. L. Kelley, *General topology*, Springer, 1975. Reprint of the 1955 edition [Van Nostrand].

[48] Y.-C. Kim and K.-A. Lee, *The Evans–Krylov theorem for nonlocal parabolic fully nonlinear equations*, Nonlinear Anal., 160 (2017), pp. 79–107.

[49] N. V. Krylov, *Boundedly inhomogeneous elliptic and parabolic equations*, Izv. Akad. Nauk SSSR Ser. Mat., 46 (1982), pp. 487–523, 670.

[50] ———, *Nonlinear elliptic and parabolic equations of the second order*, D. Reidel Publishing Co., 1987.

[51] ———, *Lectures on elliptic and parabolic equations in Hölder spaces*, American Mathematical Society, 1996.

[52] ———, *Sobolev and viscosity solutions for fully nonlinear elliptic and parabolic equations*, American Mathematical Society, 2018.

[53] K. Kuratowski, *Topology. Vol. I*, Academic Press; Polskie Wydawnictwo Naukowe, 1966.

[54] ———, *Topology. Vol. II*, Academic Press; Polskie Wydawnictwo Naukowe, 1968.

[55] D. Lacker, *Mean field games via controlled martingale problems: existence of Markovian equilibria*, Stochastic Process. Appl., 125 (2015), pp. 2856–2894.

[56] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, American Mathematical Society, 1968.

[57] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. I. Le cas stationnaire*, C. R. Math. Acad. Sci. Paris, 343 (2006), pp. 619–625.

[58] ———, *Jeux à champ moyen. II. Horizon fini et contrôle optimal*, C. R. Math. Acad. Sci. Paris, 343 (2006), pp. 679–684.

[59] ———, *Mean field games*, Jpn. J. Math., 2 (2007), pp. 229–260.

[60] G. M. Lieberman, *Second order parabolic differential equations*, World Scientific Publishing, 1996.

[61] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Birkhäuser, 2013. Reprint of the 1995 original.

[62] R. Mikulevičius and H. Pragarauskas, *Nonlinear potentials of the Cauchy–Dirichlet problem for the Bellman integro-differential equation*, Liet. Mat. Rink., 36 (1996), pp. 178–218.

[63] ———, *On the Cauchy problem for integro-differential operators in Hölder classes and the uniqueness of the martingale problem*, Potential Anal., 40 (2014), pp. 539–563.

[64] I. Monroe, *Processes that can be embedded in Brownian motion*, Ann. Probability, 6 (1978), pp. 42–56.

[65] C. Mou, *Remarks on Schauder estimates and existence of classical solutions for a class of uniformly parabolic Hamilton–Jacobi–Bellman integro-PDEs*, J. Dynam. Differential Equations, 31 (2019), pp. 719–743.

[66] B. Ø ksendal and A. Sulem, *Applied stochastic control of jump diffusions*, Universitext, Springer, Berlin, second ed., 2007.

[67] B. Øksendal, *When is a stochastic integral a time change of a diffusion?*, J. Theoret. Probab., 3 (1990), pp. 207–226.

[68] H. Pham, *Continuous-time stochastic control and optimization with financial applications*, Springer, 2009.

[69] 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.

[70] R. T. Rockafellar, *Convex analysis*, Princeton University Press, 1970.

[71] 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.

[72] K. Sato, *Lévy processes and infinitely divisible distributions*, Cambridge University Press, 2013.

[73] L. Wang, *On the regularity theory of fully nonlinear parabolic equations. II*, Comm. Pure Appl. Math., 45 (1992), pp. 141–178.

I. Chowdhury

Indian Institute of Technology - Kanpur,
Department of Mathematics and Statistics,
Kalyanpur, Kanpur - 208016, India

Email address: indranil@iitk.ac.in

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: milosz.krupski@uwr.edu.pl