ON NON-UNIQUENESS AND UNIQUENESS OF SOLUTIONS IN FINITE-HORIZON MEAN FIELD GAMES

MARTINO BARDI AND MARKUS FISCHER

ABSTRACT. This paper presents a class of evolutive Mean Field Games with multiple solutions for all time horizons T and convex but non-smooth Hamiltonian H, as well as for smooth H and T large enough. The phenomenon is analysed in both the PDE and the probabilistic setting. The examples are compared with the current theory about uniqueness of solutions. In particular, a new result on uniqueness for the MFG PDEs with small data, e.g., small T, is proved. Some results are also extended to MFGs with two populations.

CONTENTS

1. INTRODUCTION

In this paper, we study the existence of multiple solutions of Mean Field Games (briefly, MFGs) with finite horizon and of the evolutive system of PDEs associated to them, and compare the results to some uniqueness theorems. The systems of PDEs we consider are backward-forward parabolic and have the form

$$(1.1) \quad \begin{cases} -v_t + H(Dv) = \frac{1}{2}\sigma^2(x)\Delta v + F(x, m(t, \cdot)) & \text{in } (0, T) \times \mathbb{R}^d, \quad v(T, x) = G(x, m(T))), \\ m_t - \text{div}(DH(Dv)m) = \frac{1}{2}\Delta(\sigma^2(x)m) & \text{in } (0, T) \times \mathbb{R}^d, \qquad m(0, x) = \nu(x), \end{cases}$$

in the unknowns (v, m), where $m(t, \cdot)$ is a probability density for each t, and the given running and terminal costs F, G map a subset of $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ into $\mathbb{R}$. In the probabilistic formulation of a Mean

Field Game, the solution is, instead, a pair (u, m) with m as above and u either an open-loop or a feedback control satisfying an optimality and a mean field condition, see the precise definitions in Section 3.1.

There are two regimes under which there is uniqueness of a classical solution to (1.1). The first and best known is the case of H convex and F, G increasing with respect to m in the L^1 sense, which means that imitating the other agents is costly, see [39, 10] for the PDE proof and [1, 14] for probabilistic proofs, which also cover MFGs with a common noise. The second regime is for the length T of the time horizon short enough and H smooth: it was presented in a lecture of Lions on January 9th, 2009 [40], but to our knowledge it does not appear in print anywhere.

The main results in the first part of the paper give counterexamples to uniqueness when either one of these two regimes is violated. More precisely, we describe an explicit class of problems in dimension $d = 1$, with costs F, G describing a mild preference for imitating the other agents, H convex and not differentiable at one point, where multiple classical solutions exist for all $T > 0$. We also provide a variant with smooth H and T larger than a certain threshold. This is done for the PDE problem (1.1) in Section 2, and for its probabilistic formulation in Section 3 under somewhat weaker assumptions (e.g., the volatility σ may vanish).

Explicit examples of finite horizon MFGs with multiple solutions are rare. In the probabilistic literature, based on the FBSDE approach to mean field games, examples of MFGs with multiple solutions that fall into the class of problems considered here are given in Section 5.1 of Carmona, Delarue, Lachapelle [13] and Section 4.2 of Tchuendom [43], where the author shows that adding a common noise restores uniqueness. The "illuminating example" in Lacker [38, Section 3.3] also provides an MFG with explicitly known multiple solutions; cf. Example 3.4 below.

The second part of the paper is devoted to comparing the previous non-uniqueness examples with the assumptions of some uniqueness results. We argue that the monotone regime is close to being sharp, since, for instance, for costs of the form

$$F(x, \mu) = \alpha x \int y \, d\mu(y) + f(\mu), \quad G(x, \mu) = \beta x \int y \, d\mu(y) + g(\mu),$$

with f, g continuous, the costs are increasing in a suitable sense and uniqueness holds if

$$\alpha > 0, \quad \beta \geq 0,$$

whereas there are multiple solutions if

$$\alpha \leq 0, \quad \beta < 0.$$

We recall that the need of a monotonicity condition for having uniqueness for all $T > 0$ and $T \to \infty$ was discussed at length in [40]. Some explicit counterexamples, very different from ours, were shown recently by Briani and Cardaliaguet [9] and Cirant and Tonon [17]. An interesting analysis of multiple oscillating solutions via bifurcations was done very recently by Cirant [16]. For stationary MFG with ergodic cost functional, examples of non-uniqueness are known since the pioneering paper of Lasry and Lions [39], and others more explicit appear in [29, 4, 7, 25].

For the short horizon regime we prove a uniqueness theorem inspired by Lions [40] but under weaker assumptions, different boundary and terminal conditions, and with estimates in different function spaces (see Remark 4.5 for more details on the differences with [40]). It does not require any convexity of H nor monotonicity of F and G. Our example of non-uniqueness fails to satisfy more than one assumption of this theorem, but the crucial one seems to be the smoothness of the Hamiltonian (at least $C^{1,1}$). Some remarkable points of the proof of the uniqueness theorem are the following:
- it is largely self-contained and elementary, since the main estimates are got by energy methods;
- it shows that uniqueness holds also for any T, provided some other data are sufficiently small, such as the Lipschitz constant of DH, or the Lipschitz constants of F and $D_x G$ with respect to the density m (Remarks 4.7 and 4.8);
- it incorporates an a-priori estimate of $\|m(t, \cdot)\|_\infty$ that is proved in the Appendix by a probabilistic argument.
Moreover, it appears flexible enough to be applied in different settings, e.g., Neumann boundary conditions and several populations, see [5]. A different proof of uniqueness for a particular economic

model under a smallness assumption on a parameter is in [28]. More references to papers with related results are in Remark 4.10.

Finally, Section 5 extends some of the preceding results to Mean Field Games describing two populations of homogeneous agents, where the PDE system involves two Hamilton-Jacobi-Bellman and two Kolmogorov-Fokker-Planck equations instead of one. In this case the monotonicity conditions are more restrictive: in addition to a form of aversion to crowd in each population, they require that the costs of intraspecific interactions are larger than the costs of the interactions between the two populations (cfr. [15]). Therefore here there is more distance between the sufficient conditions for uniqueness and the examples of multiple solutions. We refer to [6, 2] for motivations to the study of multipopulations MFGs, to [15, 2] for examples of non-uniqueness in the case of ergodic costs and stationary PDEs, and to [5] for uniqueness when the horizon is short.

We end this introduction with some additional bibliographical remarks. The theory of MFGs started with the independent work of Lasry and Lions [39] and Huang, Caines, and Malhame [31, 33, 32]. It aims at modeling at a macroscopic level non-cooperative stochastic differential games with a very large number N of identical players. The rigorous justification of the PDE (1.1) as limit of the systems of Bellman equations for such games as $N \to \infty$ was proved recently by Cardaliaguet, Delarue, Lasry, Lions [11] using in a crucial way the convexity and monotonicity conditions leading to uniqueness for (1.1). General presentations of the field are [10, 27, 26], and many applications are analysed in [30, 24]. For the probabilistic approach to MFGs, in addition to the above mentioned works [13, 1, 14, 38, 43], we refer to Carmona and Delarue [12], Lacker [37], and Fischer [19].

2. Multiple solutions of the MFG system of PDEs

In this section we consider the backward-forward system of parabolic PDEs

$$(2.1) \quad \begin{cases} -v_t + H(v_x) = \frac{1}{2}\sigma^2(x)v_{xx} + F(x, m(t, \cdot)), & \text{in } (0, T) \times \mathbb{R}, \\ v(T, x) = G(x, m(T)), & \\ \\ m_t - (H'(v_x)m)_x = \frac{1}{2}(\sigma^2(x)m)_{xx} & \text{in } (0, T) \times \mathbb{R}, \\ m(0, x) = \nu(x), \quad m > 0, \quad \int_{\mathbb{R}} m(t, x)dx = 1 \quad \forall t > 0. \end{cases}$$

We will look for solutions such that $m(t, \cdot)$ is the density of a probability measure with finite first moment, that we denote as follows

$$\tilde{\mathcal{P}}_1(\mathbb{R}) := \{\mu \in L^\infty(\mathbb{R}) : \mu \geq 0, \int_{\mathbb{R}} \mu(x)dx = 1, \int_{\mathbb{R}} |y| \, \mu(y) \, dy < +\infty\},$$

and assume that the initial density ν is in $\tilde{\mathcal{P}}_1(\mathbb{R})$. For such functions we denote the mean with

$$M(\mu) := \int_{\mathbb{R}} y \, \mu(y) \, dy.$$

We recall that the Monge-Kantorovich distance between probability measures is

$$d_1(\mu, \nu) = \sup\left\{\int_{\mathbb{R}} \phi(y)(\mu - \nu)(y) \, dy : \phi : \mathbb{R} \to \mathbb{R} \; 1 - \text{Lipschitz}\right\}$$

and that $M(\cdot)$ is continuous for such distance. The running and terminal costs

$$F : \mathbb{R} \times \tilde{\mathcal{P}}_1(\mathbb{R}) \to \mathbb{R}, \quad G : \mathbb{R} \times \tilde{\mathcal{P}}_1(\mathbb{R}) \to \mathbb{R}$$

satisfy the following regularity conditions

(F1): $\mu \mapsto F(x, \mu)$ is continuous for the distance d_1 locally uniformly in x; for all μ the function $x \mapsto F(x, \mu)$ is differentiable with derivative denoted by $D_x F(\cdot, \mu)$, and for some $\alpha \in (0, 1], k \in \mathbb{N}$, there is C_R such that

$$|F(x, \mu) - F(y, \mu)| \leq C_R|x - y|^\alpha, \quad |D_x F(x, \mu) - D_x F(y, \mu)| \leq C_R|x - y|^\alpha$$

for all $|x|, |y| \leq R$, $M(\mu) \leq R$, and

$$|F(x, \mu)| \leq C_R(1 + |x|^k), \quad \forall \, x \in \mathbb{R}, M(\mu) \leq R;$$

(G1): for all μ $x \mapsto G(x, \mu)$ has at most polynomial growth and it is differentiable with continuous derivative denoted by $D_x G(\cdot, \mu)$.

The main qualitative assumption on the costs that allows us to build multiple solutions is the following.

(FG2): for all $x \in \mathbb{R}$ and $\mu \in \tilde{\mathcal{P}}_1(\mathbb{R})$

$$(2.2) \qquad M(\mu)D_x F(x, \mu) \leq 0, \quad M(\mu)D_x G(x, \mu) \leq 0,$$

and for $M(\mu) \neq 0$ either $D_x F(\cdot, \mu) \not\equiv 0$ or $D_x G(\cdot, \mu) \not\equiv 0$.

The meaning of this condition is that it is less costly to move to the right if $M(\mu) > 0$ and to move to the left if $M(\mu) < 0$, so in some sense it is rewarding for a representative agent to imitate the behaviour of the entire population. This is consistent with the known fact that aversion to crowd is related to the monotonicity conditions of Lasry and Lions that imply uniqueness of the solution [39, 30].

Example 2.1. Consider a running cost F of the form

$$F(x, \mu) = f_1\left(x, \int_{\mathbb{R}^d} k(x, y)\mu(y)dy\right) f_2(M(\mu)) + f_3(\mu)$$

with $f_1, k_1 \in C_1(\mathbb{R}^2)$ Lipschitz with Lipschitz derivatives, $f_2 \in C(\mathbb{R})$, and f_3 d_1-continuous. Then (F1) holds. Next assume

$$rf_2(r) \geq 0 \quad \forall r \in \mathbb{R}, \quad f_2(r) \neq 0 \quad \forall r \neq 0,$$

$$\frac{\partial f_1}{\partial x}(x, r) \leq 0, \quad \text{sign}\frac{\partial f_1}{\partial r}(x, r) = -\text{sign}\frac{\partial k}{\partial x}(x, y) \quad \forall x, y, r.$$

Then the condition (2.2) on $D_x F = \left(\frac{\partial f_1}{\partial x} + \frac{\partial f_1}{\partial r}\int_{\mathbb{R}} \frac{\partial k}{\partial x}\mu\, dy\right) f_2(M(\mu))$ in (FG2) is satisfied and $D_x F(\cdot, \mu) \not\equiv 0$ if in addition either $\frac{\partial f_1}{\partial x} \neq 0$ or both $\frac{\partial f_1}{\partial r} \neq 0$ and $\frac{\partial k}{\partial x} \neq 0$. Similar assumptions can be made on G.

About the diffusion coefficient σ we will assume

$$(2.3) \qquad \sigma : \mathbb{R} \to \mathbb{R} \quad \text{Lipschitz}, \qquad \frac{1}{2}\sigma^2(x) \geq \sigma_o > 0 \quad \forall x \in \mathbb{R}.$$

2.1. **Non-uniqueness for any time horizon.** In this section we consider the Hamiltonian

$$(2.4) \qquad H(p) := \max_{a \leq \gamma \leq b}\{-p\gamma\} = \begin{cases} -bp & \text{if } p \leq 0, \\ -ap & \text{if } p \geq 0, \end{cases}$$

so that

$$H'(p) = -b \text{ if } p < 0, \quad H'(p) = -a \text{ if } p > 0.$$

Theorem 2.1. *Assume* (F1), (G1), (FG2), (2.3), *and that H is given by* (2.4) *with $a < 0 < b$. Then, for all ν with $M(\nu) = 0$, there are two classical solutions $(v_1, m_1), (v_2, m_2)$ of* (2.1) *such that $(v_1)_x(t, x) < 0$ and $(v_2)_x(t, x) > 0$ for all $t < T$, $M(m_1(t, \cdot)) = bt$ and $M(m_2(t, \cdot)) = at$ for all t.*

Proof. We begin with the construction of (v_1, m_1) and drop the subscripts. Observe that if $v_x < 0$ the second equation of (2.1) becomes

$$(2.5) \qquad m_t + bm_x = \frac{1}{2}(\sigma^2(x)m)_{xx} \quad \text{in } (0, T) \times \mathbb{R}.$$

A solution of this equation with the initial condition $m(0, x) = \nu(x)$ exists by standard results on parabolic equations [21], and it is the law of the process $X(\cdot)$ solving

$$X(t) = \xi + bt + \int_0^t \sigma(X(s))dW(s), \quad t \in [0, T],$$

where W is a standard one-dimensional Wiener process and $\text{Law}(\xi) = \nu$. Then

$$(2.6) \qquad M(m(t)) = \mathbf{E}[X(t)] = M(\nu) + bt > 0 \quad \forall t \in (0, T].$$

Moreover, there exists $C > 0$ such that

$$d_1(m(t), m(s)) \leq C(b + \|\sigma\|_\infty)\sqrt{|t - s|}, \tag{2.7}$$

see, e.g., [10]. For such m we consider the Cauchy problem

$$-v_t - bv_x = \frac{1}{2}\sigma^2(x)v_{xx} + F(x, m(t)) \quad \text{in } (0, T) \times \mathbb{R}, \quad v(T, x) = G(x, m(T)), \tag{2.8}$$

which has a unique classical solution by standard results on parabolic equations [21], in view of (2.7) and the assumptions (F1), (G1), and (2.3). If we show that $v_x < 0$ this equation coincides with the first PDE in (2.1) and therefore we get the desired solution (v_1, m_1) of (2.1).

We consider $w := v_x$ and by the assumptions on the data can differentiate the equation (2.8) to get

$$-w_t - bw_x = \sigma(x)\sigma(x)_x w_x + \frac{1}{2}\sigma^2(x)w_{xx} + D_x F(x, m(t)) \quad \text{in } (0, T) \times \mathbb{R}, \quad w(T, x) = D_x G(x, m(T)).$$

By (2.6) and (FG2) $D_x F(x, m(t)) \leq 0$ and $D_x G(x, m(T)) \leq 0$, thus the comparison principle implies $w \leq 0$.

By the Strong Maximum Principle, if $w(t, x) = 0$ for some $t < T$ and some x, then $w(s, y) = 0$ for all $t < s < T$ and all y, and so $D_x G(\cdot, m(T)) \equiv 0$. Then we get a contradiction if the condition $D_x G(x, \mu) \not\equiv 0$ for all $M(\mu) \neq 0$ holds in (MF2), and reach the desired conclusion $w(t, \cdot) < 0$ for all $t < T$.

If, instead, only the condition $D_x F(\cdot, \mu) \not\equiv 0$ for all $M(\mu) \neq 0$ holds in (MF2), from $w(s, y) = 0$ for all $t < s < T$ and all y we get a contradiction with the PDE for w in the interval (t, T), because $D_x F(\cdot, m(t)) \not\equiv 0$ and all the other terms in the equation are null. This completes the proof of the existence of (v_1, m_1) with the stated properties.

The solution (v_2, m_2) is built in a symmetric way. We first solve

$$m_t + am_x = \frac{1}{2}(\sigma^2(x)m)_{xx}, \quad m(0, x) = \nu(x),$$

and use that the solution is the law of the process

$$X(t) = \xi + at + \int_0^t \sigma(X(s))dW(s),$$

to see that $M(m(t)) = M(\nu) + at < 0$ for $t > 0$. Next, for such m we solve the Cauchy problem

$$-v_t - av_x = \frac{1}{2}\sigma^2(x)v_{xx} + F(x, m(t)) \quad \text{in } (0, T) \times \mathbb{R}, \quad v(T, x) = G(x, m(T)). \tag{2.9}$$

We differentiate this equation and use the assumption (FG2) and the Strong Minimum Principle as before to show that $v_x(t, x) > 0$ for all x and $t \in (0, T)$. Then for such solution the last equation coincides with the second equation of (2.1), which completes the proof of the existence of (v_2, m_2) with the stated properties. □

Remark 2.1. *Symmetry.* Assume $b = -a$, so that $H(p) = b|p|$, and the data are even, i.e.,

$$F(x, \mu) = F(-x, \mu), \quad G(x, \mu) = G(-x, \mu), \quad \sigma(x) = \sigma(-x), \quad \nu(x) = \nu(-x).$$

Then the solutions (v_1, m_1) and (v_2, m_2) built in the Theorem are even reflection one of the other, i.e.,

$$v_1(t, x) = v_2(t, -x), \quad m_1(t, x) = m_2(t, -x),$$

as it is easy to check in the construction.

Remark 2.2. If $F \equiv 0$ we can drop the sign condition on a and b and consider any initial density ν such that ν with $-bT < M(\nu) < -aT$. Then the same proof produces two solutions with $(v_1)_x(t, x) < 0$ and $(v_2)_x(t, x) > 0$ for all $t < T$, and $M(m_1(T, \cdot)) > 0$, $M(m_2(T, \cdot)) < 0$.

2.2. **Eventual non-uniqueness with smooth Hamiltonian.** In this section we consider Hamiltonians that coincide with the one defined by (2.4) only for $|p| \geq \delta > 0$ and therefore can be smooth, see the Example (2.2). The precise assumption is

(2.10) $\qquad H \in C(\mathbb{R}), \quad \exists \delta, b > 0, \, a < 0 \, : \, H(p) = -bp$ if $p \leq -\delta, \quad H(p) = -ap$ if $p \geq \delta.$

On the other hand the assumptions on $D_x G$ in (FG2) are strengthened a bit by adding

(G3): for some $\varepsilon > 0$, $D_x G(\cdot, \mu) \leq -\delta$ if $M(\mu) \geq \varepsilon b$, and $D_x G(\cdot, \mu) \geq \delta$ if $M(\mu) \leq \varepsilon a.$

Then we get the existence of two distinct solutions if the time horizon T is larger than ε.

Theorem 2.2. *Assume* (F1), (G1), (FG2), (G3), (2.3), *and that H satisfies* (2.10). *Then, for all $T \geq \varepsilon$ and ν with $M(\nu) = 0$, there are two classical solutions $(v_1, m_1), (v_2, m_2)$ of* (2.1) *such that $(v_1)_x(t, x) \leq -\delta$ and $(v_2)_x(t, x) \geq \delta$ for all $0 \leq t \leq T$.*

Proof. The construction is the same as in the proof of Theorem 2.1. Now we have, by (2.6)

$$M(m_1(T)) = bT \geq b\varepsilon,$$

so $D_x G(\cdot, \mu) \leq -\delta$ by (G3). Then $w := (v_1)_x$ satisfies $w(T, x) \leq -\delta$ for all x, and the Maximum Principle implies $(v_1)_x(t, x) = w(t, x) \leq -\delta$ for all $0 \leq t \leq T$. Then, by (2.10), the equation (2.8) for v_1 coincides with the first PDE in (2.1) and the equation (2.5) for m_1 coincides with the second PDE in (2.1). The construction of the second solution is symmetric because $M(m_2(T)) = aT \leq a\varepsilon$. $\qquad \square$

Example 2.2. The Hamiltonian

$$H(p) := \max_{|\gamma| \leq 1} \left\{ -p\gamma + \frac{1}{2}\delta(1 - \gamma^2) \right\} = \begin{cases} \frac{p^2}{2\delta} + \frac{\delta}{2}, & \text{if } |p| \leq \delta, \\ |p|, & \text{if } |p| \geq \delta, \end{cases}$$

satisfies (2.10) with $-a = b = 1$. Note that $H \in C^1(\mathbb{R})$ and H' is Lipschitz and bounded, so it satisfies the assumptions required for the Hamiltonian in the uniqueness result for short time horizon of Section 4.2, see Remark 4.7. A more detailed probabilistic discussion of this example is in Section 3.3.

Example 2.3. The terminal cost

$$G(x, \mu) = -\beta x M(\mu) + g(\mu), \quad \beta > 0$$

satisfies (G1). Since $D_x G(x, \mu) = -\beta M(\mu)$, the inequality $M(\mu) D_x G(x, \mu) \leq 0$ in (FG2) is satisfied and (G3) holds with the choice $\varepsilon := \max\{\frac{\delta}{b\beta}, \frac{\delta}{|a|\beta}\}$.

3. Probabilistic approach to multiple MFG solutions

In this section, we give examples of non-uniqueness analogous to those obtained above, but under slightly different regularity assumptions, based on the probabilistic representation of the mean field game. By this we mean that we work directly with the underlying stochastic dynamics of the controlled state process and the corresponding expected costs. A solution of the mean field game is then a couple of control strategy and flow of probability measures satisfying a certain fixed point property, namely: The strategy is optimal for the control problem associated with the flow of probability measures, which in turn coincides with the flow of marginal distributions of the state process under the control strategy. In subsection 3.1, we give two definitions of solution, differing with respect to the admissible strategies (stochastic open-loop vs. Markov feedback). The second definition, based on Markov feedback strategies, is more closely related to the PDE characterization (2.1) of the mean field game, see Remark 3.3 below.

Denote by $\mathcal{P}(\mathbb{R})$ the set of all probability measures on the Borel sets of $\mathbb{R}$, and set

$$\mathcal{P}_1(\mathbb{R}) \doteq \left\{ \mu \in \mathcal{P}(\mathbb{R}) : \int |x| \mu(dx) < \infty \right\}, \qquad M(\mu) \doteq \int x \, \mu(dx), \qquad \mu \in \mathcal{P}_1(\mathbb{R}).$$

Endow $\mathcal{P}(\mathbb{R})$ with the topology of weak convergence of measures, and $\mathcal{P}_1(\mathbb{R})$ with the topology of weak convergence of measures plus convergence of first absolute moments. Notice that

$$\tilde{\mathcal{P}}_1(\mathbb{R}) \subset \mathcal{P}_1(\mathbb{R})$$

if we identify probability densities with the probability measures they induce. The topology on $\tilde{\mathcal{P}}_1(\mathbb{R})$ generated by the Monge-Kantorovich distance coincides with the topology induced by $\mathcal{P}_1(\mathbb{R})$.

3.1. **The mean field game.** As above, we consider mean field games in dimension one over a finite time horizon $T > 0$, with drift coefficient of the state dynamics equal to the control action and dispersion coefficient $\sigma : \mathbb{R} \to \mathbb{R}$, assumed to be Lipschitz continuous (hence of sublinear growth), but possibly degenerate. Let $\Gamma \subseteq \mathbb{R}$ be a compact interval, the set of control actions, and let $f : \mathbb{R} \times \mathcal{P}(\mathbb{R}) \times \Gamma \to \mathbb{R}$, $g : \mathbb{R} \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ be measurable functions with $f(\cdot, \mu, \gamma)$ of polynomial growth uniformly over compacts in $\mathcal{P}(\mathbb{R}) \times \Gamma$ and $g(\cdot, \mu)$ of polynomial growth uniformly over compacts in $\mathcal{P}(\mathbb{R})$.

Let $\mathcal{U}$ be the set of triples $((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), u, W)$ such that $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ forms a filtered probability space satisfying the usual hypotheses and carries a Γ-valued $(\mathcal{F}_t)$-progressively measurable process u and a one-dimensional $(\mathcal{F}_t)$-Wiener process W. For $\nu \in \mathcal{P}(\mathbb{R})$, let $\mathcal{U}_\nu$ denote the set of quadruples $((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, u, W)$ such that $((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), u, W) \in \mathcal{U}$ and ξ is a real-valued $\mathcal{F}_0$-measurable random variable with $\mathbf{P} \circ \xi^{-1} = \nu$.

Let $\nu \in \mathcal{P}(\mathbb{R})$, and let $u \cong ((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, u, W) \in \mathcal{U}_\nu$. Notice that ξ and W are independent since, by definition of $\mathcal{U}_\nu$, ξ is $\mathcal{F}_0$-measurable, while W is a Wiener process with respect to the filtration $(\mathcal{F}_t)$. The dynamics of the state process are then given by:

$$(3.1) \qquad X(t) = \xi + \int_0^t u(s)ds + \int_0^t \sigma\big(X(s)\big)dW(s), \quad t \in [0, T].$$

Since σ is Lipschitz, the solution $X = X^u$ of Eq. (3.1) is uniquely determined up to $\mathbf{P}$-indistinguishability. Its law is determined by the law $\mathbf{P} \circ (\xi, u, W)^{-1}$.

The costs associated with initial distribution ν, strategy $u \in \mathcal{U}_\nu$, initial time $t \in [0, T]$, and a flow of measures $\mathfrak{m} \in \mathcal{M} \doteq \mathbf{C}([0, T], \mathcal{P}(\mathbb{R}))$ are given by

$$J(t, \nu, u; \mathfrak{m}) \doteq \mathbf{E}\left[\int_0^{T-t} f\left(X^u(s), \mathfrak{m}(t+s), u(s)\right) ds + g\left(X^u(T-t), \mathfrak{m}(T)\right)\right],$$

where X^u is the unique solution of Eq. (3.1) under u, provided the expected value is finite; otherwise set $J(t, \nu, u; \mathfrak{m}) \doteq \infty$. In the above definition of the cost functional, the processes X^u and u always start from time zero, while the flow of measures is shifted according to the initial time. In this way, the set $\mathcal{U}_\nu$ of admissible control bases does not depend on the initial time, as opposed to the more standard, though equivalent, definition used in, for instance, [20].

If $\nu = \delta_x$ for some $x \in \mathbb{R}$, then we can identify $\mathcal{U}_{\delta_x}$ with $\mathcal{U}$. The value function for a flow of measures $\mathfrak{m} \in \mathcal{M}$ is then defined by

$$V(t, x; \mathfrak{m}) \doteq \inf_{u \in \mathcal{U}} J(t, \delta_x, u; \mathfrak{m}), \quad (t, x) \in [0, T] \times \mathbb{R}.$$

Notice that $V(t, x; \mathfrak{m})$ is finite for every $(t, x) \in [0, T] \times \mathbb{R}$ thanks to the growth assumptions on f, g and the boundedness of Γ.

Remark 3.1. Let $t \in [0, T]$, $\nu \in \mathcal{P}(\mathbb{R})$, $\mathfrak{m} \in \mathcal{M}$. If $J(t, \nu, u; \mathfrak{m}) < \infty$ for every $u \in \mathcal{U}_\nu$, then

$$\inf_{\tilde{u} \in \mathcal{U}_\nu} J(t, \nu, \tilde{u}; \mathfrak{m}) = \int_\mathbb{R} V(t, x; \mathfrak{m})\nu(dx).$$

Definition 3.1. *Let $\nu \in \mathcal{P}(\mathbb{R})$. An* open-loop solution of the mean field game *with initial distribution ν is a pair $(u, \mathfrak{m})$ such that*

 (i) $u \cong ((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, u, W) \in \mathcal{U}_\nu$ and $\mathfrak{m} \in \mathcal{M}$;
 (ii) optimality condition: $\infty > J(0, \nu, u; \mathfrak{m}) = \int_\mathbb{R} V(0, x; \mathfrak{m})\nu(dx)$;
(iii) mean field condition: $\mathbf{P} \circ (X^u(t))^{-1} = \mathfrak{m}(t)$ for every $t \in [0, T]$, where X^u is the unique solution of Eq. (3.1) under u.

We are mainly interested in solutions of the mean field game in Markov feedback strategies. To this end, set

$$\mathcal{A} \doteq \{\alpha : [0, T] \times \mathbb{R} \to \Gamma : \alpha \text{ measurable}\}.$$

For $\alpha \in \mathcal{A}$, $\nu \in \mathcal{P}(\mathbb{R})$, $t_0 \in [0,T]$ consider the equation:

$$(3.2) \qquad X(t) = \xi + \int_0^t \alpha\big(t_0 + s, X(s)\big)ds + \int_0^t \sigma\big(X(s)\big)dW(s), \quad t \in [0, T - t_0],$$

where W is a one-dimensional $(\mathcal{F}_t)$-Wiener process on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ satisfying the usual hypotheses and carrying a real-valued $\mathcal{F}_0$-measurable random variable ξ with $\mathbf{P} \circ \xi^{-1} = \nu$.

For $\nu \in \mathcal{P}(\mathbb{R})$, $t_0 \in [0,T]$, let $\mathcal{A}_{\nu,t_0}$ denote the set of all $\alpha \in \mathcal{A}$ such that Eq. (3.2) with initial distribution ν and initial time t_0 possesses a solution that is unique in law.

Remark 3.2. If σ is bounded and such that $\inf_{x \in \mathbb{R}} \sigma(x) > 0$, then, thanks to Girsanov's theorem, $\mathcal{A}_{\nu,t_0} = \mathcal{A}$ for all $t_0 \in [0,T]$, $\nu \in \mathcal{P}(\mathbb{R})$.

If $\alpha \in \mathcal{A}_{\nu,t_0}$, then there exists $u \cong ((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, u, W) \in \mathcal{U}_\nu$ such that

$$(3.3) \qquad \alpha\big(t_0 + s, X^u(s,\omega)\big) = u(s,\omega)$$

for $\mathrm{Leb}_{T-t_0} \otimes \mathbf{P}$-almost all $(s,\omega) \in [0, T - t_0] \times \Omega$, where Leb_{T-t_0} denotes Lebesgue measure on $\mathcal{B}([0, T - t_0])$ and X^u is the unique solution of Eq. (3.1) under u. Moreover, by uniqueness in law, if $\tilde{u} \in \mathcal{U}_\nu$ is any other stochastic open-loop strategy such that

$$\alpha\big(t_0 + s, X^{\tilde{u}}(s,\omega)\big) = \tilde{u}(s,\omega) \text{ for } \mathrm{Leb}_{T-t_0} \otimes \tilde{\mathbf{P}}\text{-a.a. } (s,\omega) \in [0, T - t_0] \times \tilde{\Omega}$$

with $X^{\tilde{u}}$ the unique solution of Eq. (3.1) under control $\tilde{u}$ and $\tilde{\mathbf{P}}$ the probability measure coming with $\tilde{u}$, then

$$\mathbf{P} \circ (X^u, u, W)^{-1} = \tilde{\mathbf{P}} \circ \left(X^{\tilde{u}}, \tilde{u}, \tilde{W}\right)^{-1}.$$

We can therefore define the costs associated with initial time $t \in [0,T]$, initial distribution ν, feedback strategy $\alpha \in \mathcal{A}_{\nu,t}$, and a flow of measures $\mathfrak{m} \in \mathcal{M}$ by setting

$$J(t, \nu, \alpha; \mathfrak{m}) \doteq J(t, \nu, u; \mathfrak{m})$$

for any stochastic open-loop strategy $u \in \mathcal{U}_\nu$ such that Eq. (3.3) holds with respect to α, u and initial time $t_0 = t$.

Definition 3.2. *Let $\nu \in \mathcal{P}(\mathbb{R})$. A Markov feedback solution of the mean field game with initial distribution ν is a pair $(\alpha, \mathfrak{m})$ such that*

   (i) $\alpha \in \mathcal{A}_\nu$ and $\mathfrak{m} \in \mathcal{M}$;
   (ii) optimality condition: $\infty > J(0, \nu, \alpha; \mathfrak{m}) = \int_{\mathbb{R}} V(0, x; \mathfrak{m})\nu(dx)$;
   (iii) mean field condition: $\mathbf{P} \circ (X^u(t))^{-1} = \mathfrak{m}(t)$ for every $t \in [0,T]$, where X^u is the unique solution of Eq. (3.1) under u with $u \in \mathcal{U}_\nu$ such that Eq. (3.3) holds with respect to α and u.

Two solutions in the sense of Definition 3.1 or Definition 3.2 are called *equivalent* if their flows of measures coincide. Notice that, by the mean field condition, equivalent solutions have the same initial distribution.

Remark 3.3. Let the function f for the combined running costs have the form

$$f(x, \mu, \gamma) = l(\gamma) + F(x, \mu)$$

for some continuous function $l \colon \Gamma \to \mathbb{R}$ and some function F as in Section 2. Set

$$H(p) \doteq \max_{\gamma \in \Gamma}\{-l(\gamma) - p\gamma\}, \quad p \in \mathbb{R}.$$

Let $\mathfrak{m} \in \mathcal{M}$, and suppose that $v = v_{\mathfrak{m}}$ is a classical solution of the Hamilton-Jacobi-Bellman equation

$$(3.4) \qquad -v_t + H(v_x) = \frac{1}{2}\sigma^2(x)v_{xx} + F(x, m(t,\cdot)) \text{ in } [0,T) \times \mathbb{R}$$

with terminal condition $v(T, \cdot) = g(\cdot, \mathfrak{m}(T))$. Then $v(t,x) = V(t,x,\mathfrak{m})$ for all $(t,x) \in [0,T] \times \mathbb{R}$. Moreover, if $\alpha \in \mathcal{A}_{\mathfrak{m}(0)}$ is such that the mean field condition of Definition 3.2 holds for α and $\mathfrak{m}$ and

$$\alpha(t,x) \in \mathrm{argmax}_{\gamma \in \Gamma}\{-l(\gamma) - p\gamma\} \text{ for all } (t,x) \in [0,T) \times \mathbb{R},$$

then $(\alpha, \mathfrak{m})$ is a solution in the sense of Definition 3.2.

3.2. **Multiple solutions.** Here, the space of control actions Γ is assumed to be a compact interval; thus, $\Gamma = [a, b]$ for some $a, b \in \mathbb{R}$ with $a < b$. Let $\psi \in \mathbf{C}^2(\mathbb{R})$ be such that

$$\sup_{x \in \mathbb{R}} e^{-c|x|} \cdot \max\{|\psi(x)|, |\psi'(x)|, |\psi''(x)|\} < \infty \text{ for some } c \in (0, \infty),$$

while for all $x \in \mathbb{R}$,

$$(3.5a) \qquad b \cdot \psi'(x) + \frac{1}{2}\sigma^2(x)\psi''(x) > 0,$$

$$(3.5b) \qquad a \cdot \psi'(x) + \frac{1}{2}\sigma^2(x)\psi''(x) < 0.$$

Here are three examples for choices of ψ such that (3.5) holds:

Example 3.1. Let $c > 0$, $d \in \mathbb{R}$. Set $\psi(x) \doteq c \cdot x + d$, $x \in \mathbb{R}$. Then (3.5) holds if $a < 0 < b$.

Example 3.2. Let $\lambda > 0$. Set $\psi(x) \doteq e^{\lambda x} - 1$, $x \in \mathbb{R}$. Then (3.5a) holds if $b > 0$, and (3.5b) holds if σ is bounded and $a < -\frac{\lambda}{2}\sigma^2(x)$ for all $x \in \mathbb{R}$.

Example 3.3. Set $\psi(x) \doteq \tanh(x)$, $x \in \mathbb{R}$. Then (3.5) holds if σ is bounded and $a \leq -\sigma^2(x)$, $\sigma^2(x) \leq b$ for all $x \in \mathbb{R}$.

Set

$$\mathcal{P}_{\exp}(\mathbb{R}) \doteq \left\{ \mu \in \mathcal{P}(\mathbb{R}) : \int e^{cx}\mu(dx) < \infty \text{ for all } c \in \mathbb{R} \right\},$$

$$\mathcal{P}_\psi(\mathbb{R}) \doteq \left\{ \mu \in \mathcal{P}(\mathbb{R}) : \int |\psi(x)|\mu(dx) < \infty \right\}.$$

By the growth assumption on ψ, $\mathcal{P}_{\exp}(\mathbb{R}) \subset \mathcal{P}_\psi(\mathbb{R})$. For $\mu \in \mathcal{P}_\psi(\mathbb{R})$, set

$$M_\psi(\mu) \doteq \int_{\mathbb{R}} \psi(y)\, \mu(dy).$$

If $\psi(x) \doteq x$, $x \in \mathbb{R}$, then we have $\mathcal{P}_\psi(\mathbb{R}) = \mathcal{P}_1(\mathbb{R})$ and $M_\psi(\mu) = M(\mu)$, the mean value of $\mu \in \mathcal{P}_1(\mathbb{R})$, in accordance with the notation introduced in Section 2.

In this subsection, we assume the running costs f to be of the form

$$f(x, \mu, \gamma) = F(x, \mu)$$

for some $F \colon \mathbb{R} \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ measurable such that for every $\mu \in \mathcal{P}_\psi(\mathbb{R})$,

$$F(\cdot, \mu) \text{ is } \begin{cases} \text{decreasing if } M_\psi(\mu) > 0, \\ \text{increasing if } M_\psi(\mu) < 0, \end{cases}$$

and the terminal costs g to be given by

$$g(x, \mu) = G(x, \mu)$$

for some $G \colon \mathbb{R} \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ measurable such that for every $\mu \in \mathcal{P}_\psi(\mathbb{R})$,

$$G(\cdot, \mu) \text{ is } \begin{cases} \text{strictly decreasing if } M_\psi(\mu) > 0, \\ \text{strictly increasing if } M_\psi(\mu) < 0. \end{cases}$$

We set $F(\cdot, \mu) \doteq 0$, $G(\cdot, \mu) \doteq 0$ if $\mu \in \mathcal{P}(\mathbb{R}) \setminus \mathcal{P}_\psi(\mathbb{R})$. Moreover, $F(\cdot, \mu)$, $G(\cdot, \mu)$ are assumed to be of polynomial growth uniformly over compacts in $\mathcal{P}(\mathbb{R})$. Note that these monotonicity conditions on F and G are very similar to the assumption (FG2) of Section 2.

Proposition 3.1. *Grant the hypotheses above, and assume that either σ is bounded or that ψ together with its first two derivatives is of polynomial growth. Let $\nu \in \mathcal{P}_{\exp}(\mathbb{R})$ be such that $M_\psi(\nu) = 0$. Then there exist two non-equivalent open-loop solutions of the mean field game with initial distribution ν, and the associated value functions are different and not constant.*

Proof. Let $((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, W)$ be such that $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ forms a filtered probability space satisfying the usual hypotheses with W a one-dimensional $(\mathcal{F}_t)$-Wiener process and ξ a real-valued $\mathcal{F}_0$-measurable random variable such that $\mathbf{P} \circ \xi^{-1} = \nu$. Let $u_+ \equiv b$ be the constant process equal to b, and let $u_- \equiv a$ be the constant process equal to a. By construction, $((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, u_{\pm}, W) \in \mathcal{U}_\nu$. Let X_+ be the unique solution of Eq. (3.1) with $u = u_+$, and let X_- be the unique solution of Eq. (3.1) with $u = u_-$. Thus, for all $t \in [0, T]$,

$$X_+(t) = \xi + t \cdot b + \int_0^t \sigma\big(X_+(s)\big) dW(s), \quad X_-(t) = \xi + t \cdot a + \int_0^t \sigma\big(X_-(s)\big) dW(s).$$

By the Burkholder-Davis-Gundy inequalities, the sublinear growth of σ, the exponential integrability of ξ and Gronwall's lemma,

$$(3.6) \qquad \mathbf{E}\left[\sup_{t \in [0,T]} |X_{\pm}(t)|^p\right] < \infty \text{ for every } p \geq 1.$$

If σ is bounded, then it also holds that

$$(3.7) \qquad \mathbf{E}\left[\sup_{t \in [0,T]} \exp\left(c \cdot X_{\pm}(t)\right)\right] < \infty \text{ for every } c \in \mathbb{R}.$$

To check (3.7), observe that, thanks to the monotonicity and positivity of the exponential and the Cauchy-Schwarz inequality,

$$\mathbf{E}\left[\sup_{t \in [0,T]} \exp\left(c \cdot X_{\pm}(t)\right)\right] \leq e^{|c| \max\{|a|,|b|\} T} \cdot \sqrt{\mathbf{E}\left[e^{2c\xi}\right]} \cdot \sqrt{\mathbf{E}\left[\sup_{t \in [0,T]} \exp\left(2c \int_0^t \sigma\big(X_{\pm}(s)\big) dW(s)\right)\right]}.$$

The first expected value on the right-hand side above is finite by hypothesis. For the second, we have, by the boundedness of σ,

$$\mathbf{E}\left[\sup_{t \in [0,T]} \exp\left(2c \int_0^t \sigma\big(X_{\pm}(s)\big) dW(s)\right)\right]$$

$$\leq \mathbf{E}\left[\sup_{t \in [0,T]} \exp\left(c \int_0^t \sigma\big(X_{\pm}(s)\big) dW(s) - \frac{c^2}{2} \int_0^t \sigma^2\big(X_{\pm}(s)\big) ds\right)^2\right] \cdot e^{c^2 \|\sigma\|_\infty^2 T}$$

$$\leq 4\, \mathbf{E}\left[\exp\left(2c \int_0^T \sigma\big(X_{\pm}(s)\big) dW(s) - c^2 \int_0^T \sigma^2\big(X_{\pm}(s)\big) ds\right)\right] \cdot e^{c^2 \|\sigma\|_\infty^2 T}$$

$$\leq 4\, \underbrace{\mathbf{E}\left[\exp\left(2c \int_0^T \sigma\big(X_{\pm}(s)\big) dW(s) - \frac{4c^2}{2} \int_0^T \sigma^2\big(X_{\pm}(s)\big) ds\right)\right]}_{=1} \cdot e^{2c^2 \|\sigma\|_\infty^2 T} < \infty,$$

where we have used Doob's maximal inequality (second to third line) as well as the fact that the stochastic exponential of a martingale with bounded quadratic variation process is again a martingale; see, for instance, Section 3.5D in [36].

Set

$$\mathfrak{m}_+(t) \doteq \mathbf{P} \circ (X_+(t))^{-1}, \qquad \mathfrak{m}_-(t) \doteq \mathbf{P} \circ (X_-(t))^{-1}, \qquad t \in [0, T].$$

Clearly, $\mathfrak{m}_+(0) = \nu = \mathfrak{m}_-(0)$. By Eq. (3.6) and the dominated convergence theorem, the mappings $[0, T] \ni t \mapsto \mathfrak{m}_{\pm}(t) \in \mathcal{P}(\mathbb{R})$ are continuous. It follows that $\mathfrak{m}_{\pm} \in \mathcal{M}$ and that $M_\psi(\mathfrak{m}_{\pm}(t))$ is finite for every $t \in [0, T]$.

We are going to show that

$$(3.8) \qquad J(0, \nu, u_+; \mathfrak{m}_+) = \inf_{\tilde{u} \in \mathcal{U}_\nu} J(0, \nu, \tilde{u}; \mathfrak{m}_+).$$

This will imply the optimality condition of Definition 3.1. Since the mean field condition is satisfied by construction of $\mathfrak{m}_+$, it will follow that $(u_+, \mathfrak{m}_+)$ is an open-loop solution of the mean field game with initial distribution ν.

By assumption and construction, $M_\psi(\nu) = 0 = M_\psi(\mathfrak{m}_+(0))$. By Itô's formula, the Fubini-Tonelli theorem, the growth conditions on σ and ψ and its derivatives, and by Eq. (3.6) and Eq. (3.7), respectively, we have for all $t \in [0, T]$,

$$M_\psi(\mathfrak{m}_+(t)) = 0 + \int_0^t \mathbf{E}\left[b \cdot \psi'(X_+(s)) + \frac{1}{2}\sigma^2(X_+(s))\psi''(X_+(s)) \right] ds.$$

By (3.5a), it follows that $M_\psi(\mathfrak{m}_+(t)) > 0$ whenever $t > 0$. As a consequence, the functions $F(\cdot, \mathfrak{m}_+(t))$, $t \in (0, T)$, $G(\cdot, \mathfrak{m}_+(T))$ are decreasing (with strict decrease for the terminal costs).

Let $\tilde{u} \in \mathcal{U}_\nu$. We may assume, since X_+ is measurable with respect to the filtration generated by the initial condition and the Wiener process, that the strategies $\tilde{u}$, u_+ are defined on the same stochastic basis with the same driving Wiener process and the same initial condition. Thus, with a slight abuse of notation, $\tilde{u} \cong ((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, \tilde{u}, W)$, $u_+ \cong ((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P}), \xi, u_+, W)$. Let $\tilde{X}$ be the unique solution of Eq. (3.1) with $u = \tilde{u}$, and let X_+ be the unique (strong) solution of Eq. (3.1) with $u = u_+$, as before. Since $b \geq \tilde{u}(t, \omega)$ for all $(t, \omega) \in [0, T] \times \Omega$, Theorem 1.1 in [34] entails that

(3.9) $$\mathbf{P}\left(X_+(t) \geq \tilde{X}(t) \text{ for all } t \in [0, T] \right) = 1.$$

This implies, in view of the monotonicity of the costs, that

$$J(0, \nu, u_+; \mathfrak{m}_+) \leq J(0, \nu, \tilde{u}; \mathfrak{m}_+),$$

which yields the optimality condition (3.8) for $(u_+, \mathfrak{m}_+)$. The optimality condition for $(u_-, \mathfrak{m}_-)$ is established in a completely analogous way by using (3.5b) instead of (3.5a) and the opposite monotonicity of the costs.

We have shown that $(u_+, \mathfrak{m}_+)$, $(u_-, \mathfrak{m}_-)$ are two open-loop solutions of the mean field game with the same initial distribution. These two solutions are non-equivalent since $M_\psi(\mathfrak{m}_+(t)) > 0$ while $M_\psi(\mathfrak{m}_-(t)) < 0$ whenever $t > 0$. Moreover, the associated value functions $V(\cdot, \cdot, \mathfrak{m}_+)$, $V(\cdot, \cdot, \mathfrak{m}_-)$ are different and non-constant since $G(\cdot, \mathfrak{m}_+(T))$ is strictly decreasing while $G(\cdot, \mathfrak{m}_-(T))$ is strictly increasing. $\qquad\square$

The proof of Proposition 3.1 also shows that, under the hypotheses of the proposition, there exist two non-equivalent Markov feedback solutions of the mean field game with non-constant value function.

In Section 3.3, we will consider a variant with non-zero running costs of the following simple example with zero running costs:

Example 3.4. Choose $\Gamma \doteq [-1, 1]$ for the space of control actions, $f \equiv F \equiv 0$ as the running costs, and $\sigma(\cdot) \equiv \sigma$ for some constant $\sigma \in [0, \infty)$ as dispersion coefficient. Set $\psi(x) \doteq x$, $x \in \mathbb{R}$; thus $M_\psi(\mu) = M(\mu)$ is the mean of $\mu \in \mathcal{P}_\psi(\mathbb{R}) = \mathcal{P}_1(\mathbb{R})$. Set $G(x, \mu) \doteq -M(\mu) \cdot x$, $x \in \mathbb{R}$, if $\mu \in \mathcal{P}_1(\mathbb{R})$. Let $\nu \in \mathcal{P}_{\exp}(\mathbb{R})$ be such that $M_\psi(\nu) = 0$. By Proposition 3.1, there exist two non-equivalent solutions of the mean field game with non-constant value function. In addition, there exists a third solution, namely the trivial solution (open-loop or feedback) corresponding to the constant strategy equal to zero; the value function in this case is constant and equal to zero as well. Observe that these three non-equivalent solutions exist for any time horizon $T > 0$, be it small or large.

3.3. **Simple example with regular Hamiltonian.** In this section, we work with the following data:

- set of control actions $\Gamma = [-1, 1]$;
- running costs f given by $f(x, \mu, \gamma) \doteq c_0 \cdot |\gamma|^2$ for some constant $c_0 > 0$;
- terminal costs $g(x, \mu) \doteq -M(\mu) \cdot x$;
- dispersion coefficient $\sigma(\cdot) \equiv \sigma$ for some constant $\sigma \in [0, \infty)$.

The associated Hamiltonian H is therefore

$$H(p) \doteq \max_{\gamma \in [-1,1]} \left\{ -c_0\gamma^2 - p\gamma \right\} = -c_0 \min\left\{ \frac{p^2}{4c_0^2}, 1 \right\} + |p| \cdot \min\left\{ \frac{|p|}{2c_0}, 1 \right\}, \quad p \in \mathbb{R}.$$

Note that it differs from Example 2.2 only by a constant.

For this simple example, the value function associated with a given flow of measures can be computed explicitly. For $M \in \mathbb{R}$, let $\alpha_M \in \mathcal{A}$ denote the constant Markov feedback strategy given by

$$\alpha_M \equiv \operatorname{sgn}(M) \cdot \min\left\{\frac{|M|}{2c_0}, 1\right\}.$$

In particular, $\alpha_0 \equiv 0$.

Let $\mathfrak{m} \in \mathcal{M}$. Set $M \doteq M(\mathfrak{m}(T))$. Then, for all $(t, x) \in [0, T] \times \mathbb{R}$,

$$J(t, \delta_x, \alpha_M; \mathfrak{m}) = c_0 \int_0^{T-t} \left(\min\left\{\frac{|M|}{2c_0}, 1\right\}\right)^2 dt - M \cdot \left(x + \int_0^{T-t} \operatorname{sgn}(M)\left(\min\left\{\frac{|M|}{2c_0}, 1\right\}\right) dt\right)$$

$$= -M \cdot x + (T - t)\left(c_0 \min\left\{\frac{M^2}{4c_0^2}, 1\right\} - |M| \cdot \min\left\{\frac{|M|}{2c_0}, 1\right\}\right).$$

The function $(t, x) \mapsto J(t, \delta_x, \alpha_M; \mathfrak{m})$ satisfies the Hamilton-Jacobi-Bellman equation (3.4) with terminal condition $g(\cdot, \mathfrak{m}(T))$. It follows that $V(t, x, \mathfrak{m}) = J(t, \delta_x, \alpha_M; \mathfrak{m})$; cf. Remark 3.3. In particular, $V(\cdot, \cdot, \mathfrak{m}) \equiv 0$ if $M = 0$.

Let $\nu \in \mathcal{P}_1(\mathbb{R})$. Let X^{α_M} be a solution of Eq. (3.2) with feedback strategy $\alpha = \alpha_M$, initial distribution ν and initial time $t_0 = 0$. Then

$$\mathbf{E}\left[X^{\alpha_M}(T)\right] = M(\nu) + \operatorname{sgn}(M) \cdot T \cdot \min\left\{\frac{|M|}{2c_0}, 1\right\}.$$

Observe that X^{α_M} is the unique in law optimal process for the flow of measures $\mathfrak{m}$, where $M = M(\mathfrak{m}(T))$. Define the flow of measures $\mathfrak{m}_{\nu, M}$ by

$$\mathfrak{m}_{\nu, M}(t) \doteq \operatorname{Law}\left(X^{\alpha_M}(t)\right), \quad t \in [0, T].$$

Notice that $\mathfrak{m}_{\nu, M}$ depends only on ν and the value of M. In view of the mean field condition in Definition 3.2 (or Definition 3.1), it follows that $\mathfrak{m}$ is the flow of measures of a solution of the mean field game with initial distribution ν if and only if $\mathfrak{m} = \mathfrak{m}_{\nu, M}$ and

(3.10) $$M(\nu) = M - \operatorname{sgn}(M) \cdot T \cdot \min\left\{\frac{|M|}{2c_0}, 1\right\}.$$

In this case, the solution is given by $(\alpha_M, \mathfrak{m}) = (\alpha_M, \mathfrak{m}_{\nu, M})$. The number of solutions of the mean field game with initial distribution ν is therefore equal to the number of solutions of Eq. (3.10) as M varies over $\mathbb{R}$. We distinguish the following cases:

a) Eq. (3.10) has at most one solution with $M = 0$, namely one if $M(\nu) = 0$, else zero.
b) $|M| \in (0, 2c_0)$: In this domain, Eq. (3.10) has at most one solution if $T \neq 2c_0$, namely one if $M(\nu) = (1 - \frac{T}{2c_0})M$, else zero. If $T = 2c_0$, then Eq. (3.10) has infinitely many solutions if $M(\nu) = 0$, else zero.
c) $|M| \geq 2c_0$: In this domain, Eq. (3.10) has at most one solution, namely one if $M(\nu) = M - \operatorname{sgn}(M)T$, else zero.

The number of solutions of the mean field game thus depends on the time horizon T; the critical value is $T = 2c_0$:

Small time horizon: $T < 2c_0$. In this case, the mean field game possesses exactly one solution with initial distribution ν. The unique solution is given by $(\alpha_M, \mathfrak{m}_{\nu, M})$ with

$$M = \begin{cases} M(\nu)/\left(1 - \frac{T}{2c_0}\right) & \text{if } |M(\nu)| \in [0, 2c_0 - T), \\ M(\nu) + \operatorname{sgn}(M(\nu))T & \text{if } |M(\nu)| \geq 2c_0 - T. \end{cases}$$

Critical time horizon: $T = 2c_0$. In this case, the mean field game possesses either one or infinitely many solutions, depending on the initial distribution ν: There are infinitely many solutions if $M(\nu) = 0$, namely the solutions corresponding to any $M \in [-2c_0, 2c_0]$; there is exactly one solution if $M(\nu) \neq 0$, namely the solution corresponding to $M = M(\nu) + \operatorname{sgn}(M(\nu))T$.

Large time horizon: $T > 2c_0$. In this case, the mean field game possesses one, two, or three non-equivalent solutions, depending on the initial distribution ν: There are exactly three solutions if $|M(\nu)| < T - 2c_0$, namely the solutions corresponding to $M \in \{M(\nu)/(1 - \frac{T}{2c_0}), M(\nu) + T, M(\nu) - T\}$. There are exactly two solutions if $|M(\nu)| = T - 2c_0$, namely the solutions corresponding to $M \in \{M(\nu) + T, M(\nu) - T\}$. There is exactly one solution if $|M(\nu)| > T - 2c_0$, namely the solution corresponding to $M = M(\nu) + \mathrm{sgn}(M(\nu))T$.

4. COMPARING THE EXAMPLES OF NON-UNIQUENESS WITH SOME UNIQUENESS RESULTS

In this section we assume for simplicity that σ is constant, and then we can consider without loss of generality

$$\frac{1}{2}\sigma^2(x) = 1.$$

First we state a uniqueness result under the classical monotonicity condition on the costs of Lasry and Lions [39] and see that such condition is essentially sharp in view of the results of Sections 2 and 3. In the following section we prove a uniqueness theorem under general assumptions on the costs and merely smoothness of the Hamiltonian, provided the time horizon T is short enough, and compare it with our examples of multiple solutions for all horizons $T > 0$.

4.1. **Uniqueness under monotonicity conditions.** In the next theorem we will assume, for all measures $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ admitting a C^2 density

$$(4.1) \qquad \int_{\mathbb{R}} (G(x, \mu) - G(x, \nu))d(\mu - \nu)(x) \geq 0,$$

$$(4.2) \qquad \int_{\mathbb{R}} (F(x, \mu) - F(x, \nu))d(\mu - \nu)(x)) > 0, \quad \text{if } M(\mu) \neq M(\nu),$$

$$(4.3) \qquad \text{for any fixed } \mu, \ x \mapsto F(x, \mu) \ \text{and} \ x \mapsto G(x, \mu) \ \text{depend only on } M(\mu),$$

$$(4.4) \qquad |F(x, \mu)| + |G(x, \mu)| \leq C_R(1 + |x|), \quad \text{if } \int_{\mathbb{R}} |x| d\mu(x) \leq R.$$

We call the property (4.1) *weak monotonicity* of G.

Theorem 4.1. *Assume* (4.1), (4.2), (4.3), (4.4), σ *constant,* H *convex and Lipschitz, and* $\nu \, dx \in \mathcal{P}_1(\mathbb{R})$. *Let* $(v_1, m_1), (v_2, m_2)$ *be two classical solutions of* (2.1) *such that* $v := v_1 - v_2$ *has* v_t, v_x, *and* v_{xx} *bounded by* $C(1 + |x|)$. *Then* $v_1 = v_2$ *and* $m_1 = m_2$ *in* $[0, T] \times \mathbb{R}$.

Proof. We only make some little variants to the proof of Theorem 3.6 of [10]. Using the representation of m_i as the law of a diffusion process with drift bounded by $\|H'\|_\infty$, as in the proof of Theorem 2.1, by standard properties of the Ito integral we get

$$\int_{\mathbb{R}} |x| dm_i(t, x) \leq \int_{\mathbb{R}} |x| d\nu(x) + T\|H'\|_\infty + |\sigma|\sqrt{T}.$$

Then $m_i(t, x) \, dx \in \mathcal{P}_1(\mathbb{R})$ for all t.

We set $m := m_1 - m_2$. We write the equation for m

$$m_t - m_{xx} - \mathrm{div}(m_1 H'((v_1)_x) - m_2 H'((v_2)_x)) = 0,$$

multiply it by $v\zeta_r$, where ζ_r is a cutoff function in space with $\zeta_r(x) = 1$ for $|x| \leq r$, and integrate by parts. Then we take the equation for v

$$-v_t + H((v_1)_x) - H((v_2)_x) = v_{xx} + F(x, m_1) - F(x, m_2),$$

multiply it by m, integrate in time and space and add it to the previous one. Next we let $r \to \infty$, which is possible by (4.4) and the growth conditions on the derivatives of v. As in [10], by (4.1) and the convexity of H we get

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

Then (4.2) implies $M(m_1(t)) = M(m_2(t))$ for all t. Now by (4.3) and the uniqueness results for HJB equations we obtain $v_1 = v_2$. We plug this into the KFP equation and finally get $m_1 = m_2$. □

Remark 4.1. Some variants of this theorem and more details on the proof can be found in [18].

Corollary 4.2. *Assume the Hamiltonian is given by* (2.4), σ *is constant, and the costs are*

(4.5) $$F(x,\mu) = \alpha x M(\mu) + f(\mu), \quad G(x,\mu) = \beta x M(\mu) + g(\mu),$$

with $f,g : \tilde{\mathcal{P}}_1(\mathbb{R}) \to \mathbb{R}$ *continuous for* d_1. *Then*

   i) there is at most one classical solution of (2.1) *with derivatives bounded by* $C(1+|x|)$ *if* $\nu\, dx \in \mathcal{P}_1(\mathbb{R})$ *and*
$$\alpha > 0, \quad \beta \geq 0\,;$$

   ii) there are two distinct classical solutions of (2.1) *with derivatives bounded by* $C(1+|x|)$ *if* $\nu \in \tilde{\mathcal{P}}_1(\mathbb{R})$, $M(\nu) = 0$, *and*
$$\alpha \leq 0, \quad \beta < 0.$$

Proof. i) First observe that
$$\int_{\mathbb{R}} (f(\mu) - f(\nu))(\mu - \nu)(x)\, dx = 0$$
because $\int_{\mathbb{R}} \mu(x)\, dx = \int_{\mathbb{R}} \nu(x)\, dx = 1$. Then for the costs of the form (4.5) we compute
$$\int_{\mathbb{R}} (F(x,\mu) - F(x,\nu))d(\mu - \nu)(x) = \alpha(M(\mu) - M(\nu)) \int_{\mathbb{R}} x d(\mu - \nu) = \alpha(M(\mu) - M(\nu))^2,$$
so (4.2) is satisfied if $\alpha > 0$. Similarly, G satisfies (4.1) if $\beta \geq 0$. Then we get the conclusion by Theorem 4.1.

   ii) This comes from Theorem 2.1 in view of Example 2.1. The growth condition on the solutions v_i follows from the equations (2.8) and (2.9), by (4.4) and the boundedness of $M(m_i(t))$. □

Remark 4.2. The last result can be modified to cover possibly smooth Hamiltonians satisfying merely (2.10) instead of (2.4). In fact, statement *i)* remains true if H is convex, whereas *ii)* holds for $T > \max\{\frac{\delta}{b\beta}, \frac{\delta}{|a|\beta}\}$ by Theorem 2.2 and Example 2.3.

Remark 4.3. The assumptions in the two cases of the Corollary have the following simple interpretation in terms of aversion or not to the crowd. The term $xM(\mu)$ is positive if the position x of a representative individual in the population is on the same side of the origin as the expected value of the population distribution. Then $\alpha > 0$ and $\beta > 0$ mean that the individual pays a cost for such a position, whereas he has a gain if he stays on the opposite side. This models a form of aversion to crowd, whereas the case $\alpha < 0$ and $\beta < 0$ corresponds to a reward for having positions on the same side as the mean position of the population. The fact that the monotonicity conditions for uniqueness are related to crowd-aversion is well known [30], and there are examples of multiple solutions when imitation is rewarding in stationary MFG [29, 4]. The present example seems to be the first for evolutive MFG PDEs, together with those in the very recent papers [9] and [17].

Remark 4.4. Note that in the local case with costs of the form $F(x,\mu) = f(x)\mu(x)$ the crucial quantity for uniqueness is the sign of f ($f > 0$ implies uniqueness), whereas in the non-local case $F(x,\mu) = f(x)M(\mu)$ what seems to count is the sign of f', because $f' < 0$ implies non-uniqueness.

4.2. **Uniqueness for short time horizon.** In this section we assume $\frac{1}{2}\sigma^2(x) = 1$. On the other hand we allow any space dimension $d \geq 1$. We consider the MFG system

(4.6) $$\begin{cases} -v_t + H(Dv) = \Delta v + F(x, m(t, \cdot)) & \text{in } (0,T) \times \mathbb{R}^d, \quad v(T,x) = G(x, m(T))), \\ \\ m_t - \text{div}(DH(Dv)m) = \Delta m & \text{in } (0,T) \times \mathbb{R}^d, \quad m(0,x) = \nu(x), \end{cases}$$

where $Dv = \nabla_x v$ denotes the gradient of v with respect to the space variables, Δ is the Laplacian with respect to the space variables x, and $DH = \nabla_p H$ is the gradient of the Hamiltonian. Our

main assumptions are the smoothness of the Hamiltonian and a Lipschitz continuity of the costs in the norm $\|\cdot\|_2$ of $L^2(\mathbb{R}^d)$ that we state next. Define

$$\tilde{\mathcal{P}}(\mathbb{R}^d) := \{\mu \in L^\infty(\mathbb{R}^d) : \mu \geq 0, \int_\mathbb{R} \mu(x)dx = 1\},$$

and note that $\tilde{\mathcal{P}}(\mathbb{R}^d) \subseteq L^2(\mathbb{R}^d)$. We will assume $F, G : \tilde{\mathcal{P}}(\mathbb{R}^d) \to \mathbb{R}$ satisfy, for all μ, ν,

(4.7)
$$\|F(\cdot, \mu) - F(\cdot, \nu)\|_2^2 \leq L_F \|\mu - \nu\|_2^2,$$

(4.8)
$$\|DG(\cdot, \mu) - DG(\cdot, \nu)\|_2^2 \leq L_G \|\mu - \nu\|_2^2,$$

The next theorem is a variant and an extension of a result presented by Lions in [40], see Remark 4.5 for a comparison.

Theorem 4.3. *Assume $H \in C^2(\mathbb{R}^d)$, (4.7), (4.8), $\nu \in \tilde{\mathcal{P}}(\mathbb{R}^d)$, and $(v_1, m_1), (v_2, m_2)$ are two classical solutions of (4.6) such that $D(v_1 - v_2) \in L^2([0, T] \times \mathbb{R}^d)$. Suppose also that either*
(i) $|DH| \leq C_H$ and $|D^2 H| \leq \bar{C}_H$, or
(ii) $|Dv_i| \leq K$, $i = 1, 2$.
Then there exists $\bar{T} > 0$ such that, if $T < \bar{T}$, $v_1(t, \cdot) = v_2(t, \cdot)$ and $m_1(t, \cdot) = m_2(t, \cdot)$ for all $t \in [0, T]$. Moreover $\bar{T}$ depends only on $d, L_F, L_G, \|\nu\|_\infty, C_H$, and $\bar{C}_H$ in case (i), and only on $d, L_F, L_G, \|\nu\|_\infty, \sup_{|p| \leq K} |DH(p)|$, and $\sup_{|p| \leq K} |D^2 H(p)|$ in case (ii).

Proof. Step 0. We first use the density estimate of Proposition A.1 in the Appendix to get that $m_i(t, \cdot) \in \tilde{\mathcal{P}}(\mathbb{R}^d)$ for all t and

(4.9)
$$\|m_i(t, \cdot)\|_\infty \leq C \|\nu\|_\infty, \quad i = 1, 2, \ \forall t \in [0, T],$$

where the constant $C = C(T, \|DH(Dv_i)\|_\infty, d)$ is bounded for bounded entries. By known properties of the Fokker-Planck equation for m_i, we have that $m_i(t, \cdot) \in \tilde{\mathcal{P}}$ for all t and then

$$\sup_{0 \leq t \leq T} \|m_i(t, \cdot)\|_2 \leq \sup_{0 \leq t \leq T} \|m_i(t, \cdot)\|_\infty < +\infty.$$

Step 1. Define $v := v_1 - v_2$, $m := m_1 - m_2$, $B(t, x) := \int_0^1 DH(Dv_2 + s(Dv_1 - Dv_2))ds$, and observe that $B \in L^\infty((0, T) \times \mathbb{R}^d)$ and v satisfies

$$\begin{cases} -v_t + B(t, x) \cdot Dv = \Delta v + F(x, m_1) - F(x, m_2) & \text{in } (0, T) \times \mathbb{R}^d, \\ v(T, x) = G(x, m_1(T)) - G(x, m_2(T)). \end{cases}$$

Now set $w := Dv$, $\bar{F}(t, x) := F(x, m_1) - F(x, m_2)$, and differentiate the equation to get the system of parabolic PDEs

(4.10)
$$\begin{cases} -\frac{\partial w_j}{\partial t} + (B(t, x) \cdot w)_{x_j} = \Delta w_j + (\bar{F}(t, x))_{x_j} & \text{in } (0, T) \times \mathbb{R}^d, \quad j = 1, \ldots, d, \\ w(T, x) = D_x G(x, m_1(T)) - D_x G(x, m_2(T)). \end{cases}$$

Step 2. Take a smooth cutoff function $\zeta_R \geq 0$ satisfying $|D\zeta_R| \leq C$, $\zeta_R(x) = 1$ for $|x| \leq R$ and $\zeta_R(x) = 0$ for $|x| \geq 2R$. Multiply the j-th equation of (4.10) by $w_j \zeta_R^2$, integrate by parts in space and in the interval $[t, T]$ in time to get

$$\int_t^T \frac{d}{dt} \int_{\mathbb{R}^d} \frac{w_j^2}{2} \zeta_R^2 dx ds + \int_t^T \int_{\mathbb{R}^d} B \cdot w \left(\frac{\partial w_j}{\partial x_j} \zeta_R^2 + w_j \frac{\partial \zeta_R^2}{\partial x_j}\right) dx ds =$$
$$\int_t^T \int_{\mathbb{R}^d} \left(|Dw_j|^2 \zeta_R^2 + w_j Dw_j \cdot D\zeta_R^2\right) dx ds + \int_t^T \int_{\mathbb{R}^d} \bar{F} \left(\frac{\partial w_j}{\partial x_j} \zeta_R^2 + w_j \frac{\partial \zeta_R^2}{\partial x_j}\right) dx ds,$$

so that

$$(4.11) \quad \int_{\mathbb{R}^d} \frac{w_j^2(t, \cdot)}{2} \zeta_R^2 dx - \int_{\mathbb{R}^d} \frac{w_j^2(T, \cdot)}{2} \zeta_R^2 dx + \int_t^T \int_{\mathbb{R}^d} |Dw_j|^2 \zeta_R^2 dx ds \leq$$

$$\int_t^T \int_{\mathbb{R}^d} |w_j| |Dw_j| 2\zeta_R |D\zeta_R| dx ds + \|B\|_\infty \int_t^T \int_{\mathbb{R}^d} \left(|w| \left| \frac{\partial w_j}{\partial x_j} \right| \zeta_R^2 + |w||w_j| \left| \frac{\partial \zeta_R^2}{\partial x_j} \right| \right) dx ds +$$

$$\int_t^T \int_{\mathbb{R}^d} |\bar{F}| \left(\left| \frac{\partial w_j}{\partial x_j} \right| \zeta_R^2 + |w_j| \left| \frac{\partial \zeta_R^2}{\partial x_j} \right| \right) dx ds.$$

We estimate

$$\int_t^T \int_{\mathbb{R}^d} |w_j| |Dw_j| 2\zeta_R |D\zeta_R| dx ds \leq \varepsilon \int_t^T \int_{\mathbb{R}^d} |Dw_j|^2 \zeta_R^2 dx ds + \frac{1}{\varepsilon} \int_t^T \int_{\mathbb{R}^d} w_j^2 |D\zeta_R|^2 dx ds$$

and observe that all integrals involving derivatives of ζ_R vanish as $R \to \infty$ because $D\zeta_R \to 0$ a.e. and $w, \bar{F} \in L^2([0,T] \times \mathbb{R}^d)$. Next we estimate

$$\int_t^T \int_{\mathbb{R}^d} |w| \left| \frac{\partial w_j}{\partial x_j} \right| \zeta_R^2 dx ds \leq \int_t^T \int_{\mathbb{R}^d} \left(\frac{1}{2\varepsilon} |w|^2 \zeta_R^2 + \frac{\varepsilon}{2} \left| \frac{\partial w_j}{\partial x_j} \right|^2 \zeta_R^2 \right) dx ds,$$

$$\int_t^T \int_{\mathbb{R}^d} |\bar{F}| \left| \frac{\partial w_j}{\partial x_j} \right| \zeta_R^2 dx ds \leq \int_t^T \int_{\mathbb{R}^d} \left(\frac{1}{2\varepsilon} |\bar{F}|^2 \zeta_R^2 + \frac{\varepsilon}{2} \left| \frac{\partial w_j}{\partial x_j} \right|^2 \zeta_R^2 \right) dx ds.$$

We plug these inequalities in (4.11) with ε satisfying $1 = (\|B\|_\infty + 3)\varepsilon/2$, so that the terms involving Dw_j cancel out. Now we can let $R \to \infty$, sum over j, and use the terminal condition on $w = Dv$ with the assumption (4.8) to get

$$\|w(t, \cdot)\|_2^2 \leq L_G \|m(T, \cdot)\|_2^2 + \int_t^T \frac{d}{\varepsilon} \|\bar{F}(s, \cdot)\|_2^2 ds + \frac{d\|B\|_\infty}{\varepsilon} \int_t^T \|w(s, \cdot)\|_2^2 ds.$$

Gronwall inequality gives, for all $0 \leq t \leq T$,

$$\|w(t, \cdot)\|_2^2 \leq \left(L_G \|m(T, \cdot)\|_2^2 + \frac{d}{\varepsilon} \int_t^T \|\bar{F}(s, \cdot)\|_2^2 ds \right) e^{d\|B\|_\infty T/\varepsilon}.$$

We set $c_o := d(\|B\|_\infty + 3)/2 = d/\varepsilon$ and use (4.7) to get

$$(4.12) \quad \|Dv(t, \cdot)\|_2^2 \leq \left(L_G \|m(T, \cdot)\|_2^2 + c_o L_F \int_t^T \|m(s, \cdot)\|_2^2 ds \right) e^{c_o \|B\|_\infty T}.$$

Step 3. Observe that m satisfies

$$\begin{cases} m_t - \operatorname{div}(DH(Dv_1)m) = \Delta m + \operatorname{div}((DH(Dv_1) - DH(Dv_2))m_2) & \text{in } (0,T) \times \mathbb{R}^d, \\ m(0,x) = 0. \end{cases}$$

Define $\tilde{B}(t,x) := DH(Dv_1)$, the matrix

$$A(t,x) := m_2 \int_0^1 D^2 H(Dv_2 + s(Dv_1 - Dv_2)) ds$$

and $\tilde{F}(t,x) := A(t,x)(Dv_1 - Dv_2)$. Then the PDE for m reads

$$m_t - \operatorname{div}(\tilde{B}m) = \Delta m + \operatorname{div}\tilde{F},$$

with $\tilde{B}$ and A bounded by the assumption *(i)* or *(ii)* and the estimate (4.9). As in Step 2 we multiply the equation by $m\zeta_R^2$ and integrate by parts. Now $m \in L^2([0,T] \times \mathbb{R}^d)$ by Step 0 and $\tilde{F} \in L^2([0,T] \times \mathbb{R}^d)$ by (4.12). Then we can estimate as in Step 2 and let $R \to \infty$ to get

$$\|m(t, \cdot)\|_2^2 \leq \frac{1}{\varepsilon} \int_0^t \|\tilde{F}(s, \cdot)\|_2^2 ds + \frac{\|\tilde{B}\|_\infty}{\varepsilon} \int_0^t \|m(s, \cdot)\|_2^2 ds,$$

where we used the initial condition $m(0, x) = 0$ and chose $\varepsilon = 2/(\|\tilde{B}\|_\infty + 3) =: 1/c_1$. Then Gronwall inequality gives, for all $0 \le t \le T$,

$$(4.13) \qquad \|m(t, \cdot)\|_2^2 \le c_1 e^{c_1 \|\tilde{B}\|_\infty T} \int_0^t \|\tilde{F}(s, \cdot)\|_2^2 ds \le c_1 e^{c_1 \|\tilde{B}\|_\infty T} \|A\|_\infty^2 \int_0^t \|Dv(s, \cdot)\|_2^2 ds.$$

Step 4. Now we set $\phi(t) := \|Dv(t, \cdot)\|_2^2$ and combine (4.12) and (4.13) to get

$$(4.14) \qquad \phi(t) \le C_1 C_3 \int_0^T \phi(s) ds + C_2 C_3 \int_t^T \int_0^\tau \phi(s) ds \, d\tau,$$

for suitable explicit constants C_i depending only on the quantities listed in the statement of the theorem. Then $\Phi := \sup_{0 \le t \le T} \phi(t)$ satisfies

$$\Phi \le \Phi(T C_1 C_3 + T^2 C_2 C_3 / 2)$$

which implies $\Phi = 0$ if $T < \bar{T} := (C_1 + \sqrt{C_1^2 + 2 C_2 / C_3}) / C_2$. Therefore for such T we conclude that $Dv_1(t, x) = Dv_2(t, x)$ for all x and $0 \le t \le T$. By the uniqueness of solution for the KFP equation we deduce $m_1 = m_2$ and then, by the uniqueness of solution of the HJB equation, $v_1 = v_2$. $\qquad \square$

Remark 4.5. The same result holds for solutions $\mathbb{Z}^d$-periodic in the space variable x in the case that F and G are $\mathbb{Z}^d$-periodic in x, with the same proof (and no need of cutoff). In such case of periodic boundary conditions a uniqueness result for short T was presented by Lions in [40] for regularizing running cost F and for terminal cost G independent of m. He used estimates in L^1 norm for m and in L^∞ norm for Dv, instead of the L^2 norms we used here in (4.12) and (4.13). Our main contributions are the replacement of the hard estimates stated by Lions with more direct ones obtained by energy methods that require less assumptions on F, and the consideration of a cost G depending on the terminal density $m(T)$.

Remark 4.6. If the terminal cost G satisfies (4.8) with $L_G = 0$, i.e., DG does not depend on $m(T)$, then $C_1 = 0$ and $\bar{T} = \sqrt{2/(C_2 C_3)}$. By a simple and elegant argument of Lions [40] based on the inequality (4.14), the condition for uniqueness $T \le \bar{T}$ can be improved in this case to $T \le \pi/(2\sqrt{C_2 C_3})$.

Remark 4.7. The proof of the theorem shows also that there is uniqueness for any $T > 0$ if *(i)* holds with $\bar{C}_H$ sufficiently small, or if *(ii)* holds with $\sup_{|p| \le K} |D^2 H(p)|$ sufficiently small. In fact, $\|A\|_\infty$ becomes small as this quantities get small, and so the constant C_3 can be made as small as we want. This generalizes a recent result in [44] for nonconvex H with both C_H and $\bar{C}_H$ small (and periodic boundary conditions).

Remark 4.8. The proof of the theorem shows also the uniqueness of the solution for any $T > 0$ if the Lipschitz constants L_F and L_G of the running cost F and of $D_x G$ are small enough, because the constants C_1 and C_2 can be made small. See also [3] for existence and uniqueness results under smallness assumptions on the data.

Remark 4.9. The estimate (4.9), obtained by probabilistic methods in the Appendix, could be replaced by

$$\|m_i(t, \cdot)\|_\infty \le C_d \left(\|\nu\|_\infty + T(1 + \|DH(Dv_i)\|_\infty)^{d+4} \right), \quad i = 1, 2, \ \forall \, t \in [0, T],$$

which can be deduced from Corollary 7.3.8 of [8], p. 302. Another alternative estimate, obtained by a simple application of the comparison principle, gives

$$m_i(t, x) \le e^{t \sup(\mathrm{div} DH(Dv_i))^+} \sup \nu,$$

where, however, the right hand side may depend on the second derivatives of v_i.

Remark 4.10. A similar structure of proof, based on combining backward and forward integral estimates, was also employed for finite state MFGs by [23]. An existence and uniqueness result under a "small data" condition was proved in [32] for Linear-Quadratic-Gaussian MFGs using a contraction mapping argument to solve the associated system of Riccati differential equations. For different classes of linear-quadratic MFGs, a contraction mapping argument is also used in [45, 41] to obtain existence and uniqueness of solutions under certain implicitly "small data"

assumptions. The work [1] mentioned in the introduction provides existence and uniqueness for MFGs with common noise under a weak monotonicity condition in the spirit of Lasry-Lions. The proof relies on the representation of the MFG in terms of a forward-backward stochastic differential equation. Existence of a unique solution is first established for small times by a contraction mapping argument. The monotonicity assumption then allows to extend the solution to any given time horizon. Finally, our assumption $D(v_1 - v_2) \in L^2([0,T] \times \mathbb{R}^d)$ can be replaced by $(v_1 - v_2)_t \in L^2([0,T] \times \mathbb{R}^d)$, see [18].

Remark 4.11. *On the smoothness of H.* The assumption $H \in C^2(\mathbb{R}^d)$ can be relaxed to $H \in C^1(\mathbb{R}^d)$ with DH locally Lipschitz. Then the statement of the Theorem remains true with the following changes: in case *(i)* DH is assumed globally Lipschitz and $\bar{C}_H$ is redefined as its Lipschitz constant, in case *(ii)* the time $\bar{T}$ depends on the Lipschitz constant of DH on the ball $\{p : |p| \leq K\}$ instead of $\sup_{|p| \leq K} |D^2 H(p)|$. The proof is the same, after changing Step 3 with the observation that there exists a measurable and locally bounded matrix valued function $\tilde{A}$ such that

$$DH(p) - DH(q) = \tilde{A}(p,q)(p-q),$$

see [5] for a proof of this fact.

Example 4.1. *Regularizing costs.* Consider F and G of the form

$$F(x,\mu) = F_1\left(x, \int_{\mathbb{R}^d} k_1(x,y)\mu(y)dy\right), \quad G(x,\mu) = g_1(x)\int_{\mathbb{R}^d} k_2(x,y)\mu(y)dy + g_2(x)$$

with $F_1 : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ measurable and $k_1, k_2 \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$. Then

$$\| \int_{\mathbb{R}^d} k_i(\cdot,y)\mu(y)dy - \int_{\mathbb{R}^d} k_i(\cdot,y)\nu(y)dy\|_2 \leq \|k_i(\cdot,\cdot)\|_2 \|\mu - \nu\|_2, \quad i = 1,2.$$

About F we suppose

$$|F_1(x,r) - F_1(x,s)| \leq L_1|r-s| \quad \forall x \in \mathbb{R}^d, r,s \in \mathbb{R}$$

and get (4.7) with $L_F = L_1^2 \|k_1(\cdot,\cdot)\|_2^2$. About G we assume $g_1, g_2 \in C^1(\mathbb{R}^d)$, Dg_1 bounded, $D_x k_2 \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$. Then

$$DG(x,\mu) = Dg_1(x)\int_{\mathbb{R}^d} k_2(x,y)\mu(y)dy + g_1(x)\int_{\mathbb{R}^d} D_x k_2(x,y)\mu(y)dy + Dg_2(x)$$

satisfies

$$\|DG(\cdot,\mu) - DG(\cdot,\nu)\|_2 \leq (\|Dg_1\|_\infty \|k_2(\cdot,\cdot)\|_2 + \|g_1\|_\infty \|D_x k_2(\cdot,\cdot)\|_2)\|\mu - \nu\|_2,$$

which implies (4.8).

Example 4.2. *Local costs.* Take $G = G(x)$ independent of $m(T)$ and F of the form

$$F(x,\mu) = F_l(x,\mu(x))$$

with $F_l : \mathbb{R}^d \times [0,+\infty) \to \mathbb{R}$ such that

$$|F_l(x,r) - F_l(x,s)| \leq L_l|r-s| \quad \forall x \in \mathbb{R}^d, r,s \geq 0.$$

Then F satisfies (4.7) with $L_F = L_l^2$.

Remark 4.12. The functional $M(\mu) = \int_{\mathbb{R}^d} y\mu(y)dy$ is Lipschitz in L^2 among densities of measures with support contained in a given compact set, by Example 4.1. This is the case, for instance, of periodic boundary conditions where the measures are on the torus $\mathbb{T}^d$ (see Remark 4.5). In such a case F of the form $F(x, M(\mu))$ satisfies (4.7) if it is Lipschitz in the second entry uniformly with respect to x.

Remark 4.13. If we compare the uniqueness Theorem 4.3 with the non-uniqueness Theorem 2.1 there are several different assumptions. For instance, in Theorem 2.1 DH is discontinuous in 0 and $v_1 - v_2 \notin L^2(\mathbb{R})$, and the costs in Example 2.1 do not satisfy (4.7) and (4.8). So it is not clear which of these conditions is mostly responsible of the lack of uniqueness for short T. From the proof of Theorem 4.3 we guess that at least the continuity of DH is indispensable for the short-horizon uniqueness.

5. Uniqueness and non-uniqueness for two populations

Consider the system of MFG PDEs corresponding to two populations

(5.1)
$$\begin{cases} -\partial_t v_i + H_i(\partial_x v_i) = \frac{1}{2}\sigma_i^2(x)\partial_{xx}v_i + F_i(x, m_1(t, \cdot), m_2(t, \cdot)) & \text{in } (0, T) \times \mathbb{R}, \\[2mm] v_i(T, x) = G_i(x, m_1(T), m_2(T)), \\[2mm] \partial_t m_i - (H_i'(\partial_x v_i)m_i)_x = \frac{1}{2}\sigma_i^2(x)\partial_{xx}m_i & \text{in } (0, T) \times \mathbb{R}, \\[2mm] m_i(0, x) = \nu_i(x), & i = 1, 2, \end{cases}$$

where the Hamiltonians are

(5.2)
$$H_i(p) := \max_{a_i \le \gamma \le b_i}\{-p\gamma\} = \begin{cases} -b_i p & \text{if } p \le 0, \\ -a_i p & \text{if } p \ge 0, \end{cases} \qquad a_i < 0 < b_i,$$

so that

$$H_i'(p) = -b_i \text{ if } p < 0, \quad H_i'(p) = -a_i \text{ if } p > 0,$$

σ_i satisfy (2.3), and $F_i, G_i : \mathbb{R} \times \tilde{\mathcal{P}}_1(\mathbb{R}) \times \tilde{\mathcal{P}}_1(\mathbb{R}) \to \mathbb{R}$ verify the same regularity conditions as F, G in Section 2.

We give two different sets of qualitative assumptions that produce examples of nonuniqueness. The first is the following, for $i = 1, 2$,

(5.3)
$$D_x F_i(x, \mu_1, \mu_2) \begin{cases} \le 0 & \text{if } M(\mu_1), M(\mu_2) > 0, \\ \ge 0 & \text{if } M(\mu_1), M(\mu_2) < 0. \end{cases}$$

(5.4)
$$D_x G_i(x, \mu_1, \mu_2) \begin{cases} \le 0 & \text{and } \not\equiv 0 & \text{if } M(\mu_1), M(\mu_2) > 0, \\ \ge 0 & \text{and } \not\equiv 0 & \text{if } M(\mu_1), M(\mu_2) < 0. \end{cases}$$

Proposition 5.1. *Assume (5.2), F_i satisfy (F1) and (5.3), G_i satisfy (G1) and (5.4). Then, for all ν_i with $M(\nu_i) = 0$, there is a classical solution of (5.1) with $\partial_x v_1(t, x), \partial_x v_2(t, x) < 0$ for all $0 < t < T$ and a classical solution with $\partial_x v_1(t, x), \partial_x v_2(t, x) > 0$ for all $0 < t < T$.*

Proof. For the solution with $\partial_x v_1(t, x), \partial_x v_2(t, x) < 0$ we make the ansatz that m_i solves

(5.5)
$$\partial_t m_i + b_i \partial_x m_i = \frac{1}{2}\sigma_i^2(x)\partial_{xx}m_i \quad \text{in } (0, T) \times \mathbb{R}, \quad m_i(0, x) = \nu_i(x),$$

so that $M(m_1(t, \cdot)) = b_1 t > 0$ and $M(m_2(t, \cdot)) = b_2 t > 0$. We solve the equations

(5.6) $\quad -\partial_t v_i - b_i \partial_x v_i = \frac{1}{2}\sigma_i^2(x)\partial_{xx}v_i + F_i(x, m_1(t, \cdot), m_2(t, \cdot)), \quad v_i(T, x) = G_i(x, m_1(T), m_2(T)),$

for $i = 1, 2$. Then $w_i := \partial_x v_i$ solves
(5.7)
$$-\partial_t w_i - b_i \partial_x w_i - \sigma_i(x)\sigma_i(x)_x \partial_x w_i - \frac{1}{2}\sigma_i^2(x)\partial_{xx}w_i = D_x F_i(x, m_1(t, \cdot), m_2(t, \cdot)) \le 0, \text{ in } (0, T) \times \mathbb{R},$$

(5.8)
$$w_i(T, x) = D_x G_i(x, m_1(T, \cdot), m_2(T, \cdot)) < 0,$$

and so $w_i < 0$ for all $t < T$ by the Strong Maximum Principle. On the other hand, if $\partial_x v_i < 0$ the equations for m_i in (5.1) are (5.5) and the equations for v_i in (5.1) are (5.6), as we guessed.

The construction of a solution with $\partial_x v_1(t, x), \partial_x v_2(t, x) > 0$ is done in a symmetric way, starting with the ansatz that m_i solves

(5.9)
$$\partial_t m_i + a_i \partial_x m_i = \frac{1}{2}\sigma_i^2(x)\partial_{xx}m_i \quad \text{in } (0, T) \times \mathbb{R}, \quad m_i(0, x) = \nu_i(x),$$

so now $M(m_1(t, \cdot)) = a_1 t < 0$ and $M(m_2(t, \cdot)) = a_2 t < 0$. We proceed as before by solving a linear equation for v_i like (5.6) but with the term $b_i \partial_x v_i$ replaced by $a_i \partial_x v_i$. Now $D_x F_i \ge 0$ in the equation for $w_i := \partial_x v_i$ and $D_x G_i > 0$ in the terminal conditions, so $w_i > 0$ for all $t < T$. $\qquad \square$

The second set of assumptions is the following.

$$(5.10) \qquad D_x F_1(x, \mu_1, \mu_2) \begin{cases} \leq 0 & \text{if } M(\mu_1) > 0, M(\mu_2) < 0 \\ \geq 0 & \text{if } M(\mu_1) < 0, M(\mu_2) > 0. \end{cases}$$

$$(5.11) \qquad D_x F_2(x, \mu_1, \mu_2) \begin{cases} \leq 0 & \text{if } M(\mu_1) < 0, M(\mu_2) > 0, \\ \geq 0 & \text{if } M(\mu_1) > 0, M(\mu_2) < 0. \end{cases}$$

$$(5.12) \qquad D_x G_1(x, \mu_1, \mu_2) \begin{cases} \leq 0 & \text{and } \not\equiv 0 & \text{if } M(\mu_1) > 0, M(\mu_2) < 0, \\ \geq 0 & \text{and } \not\equiv 0 & \text{if } M(\mu_1) < 0, M(\mu_2) > 0. \end{cases}$$

$$(5.13) \qquad D_x G_2(x, \mu_1, \mu_2) \begin{cases} \leq 0 & \text{and } \not\equiv 0 & \text{if } M(\mu_1) < 0, M(\mu_2) > 0, \\ \geq 0 & \text{and } \not\equiv 0 & \text{if } M(\mu_1) > 0, M(\mu_2) < 0. \end{cases}$$

Proposition 5.2. *Assume* (5.2), F_i *satisfy* (F1), (5.10), *and* (5.11), G_i *satisfy* (G1), (5.12), *and* (5.13), $i = 1, 2$. *Then, for all* ν_i *with* $M(\nu_i) = 0$, *there is a solution of* (2.1) *with* $\partial_x v_1(t, x) < 0, \partial_x v_2(t, x) > 0$ *for all* $0 < t < T$ *and a solution with* $\partial_x v_1(t, x) > 0, \partial_x v_2(t, x) < 0$ *for all* $0 < t < T$.

Proof. This is a variant of the preceding proof, so we only explain the changes. For the solution with $\partial_x v_1 < 0, \partial_x v_2 > 0$ we start solving (5.5) for $i = 1$ and (5.9) for $i = 2$, so that $M(m_1(t)) > 0$ and $M(m_2(t)) < 0$. Then for $i = 1$ we solve (5.6) and get (5.7) and (5.8), by the assumptions (5.10) and (5.12), so $\partial_x v_1 < 0$. For $i = 2$ we solve (5.6) with the term $b_i \partial_x v_i$ replaced by $a_2 \partial_x v_2$. Then $D_x F_2 \geq 0$ in the equation for $w_2 := \partial_x v_2$ and $D_x G_2 > 0$ in the terminal conditions, by the assumptions (5.11) and (5.13). Thus $\partial_x v_2 > 0$. The construction of the solution with $\partial_x v_1 > 0, \partial_x v_2 < 0$ is symmetric. $\qquad \square$

Remark 5.1. In Proposition 5.1 the assumption $D_x G_i \not\equiv 0$ in (5.4) can be dropped and replaced by $D_x F_i \not\equiv 0$ in (5.4), by the argument of Theorem 2.1. The same variant can be done on the assumptions of Proposition 5.2.

Next we use the last two propositions and the example to check the sharpness of some sufficient conditions for uniqueness. A natural generalization to systems with two populations of the monotonicity conditions (4.1) (4.2) is given the following result, see also [15] for stationary equations.

Theorem 5.3. *Assume* H_i *are convex and Lipschitz,* $\sigma_i > 0$ *are constant,* $\nu_i \, dx \in \mathcal{P}_1(\mathbb{R})$, *the functions* $x \mapsto F_i(x, \mu_1, \mu_2)$ *and* $x \mapsto G_i(x, \mu_1, \mu_2)$ *depend only on* $M(\mu_1), M(\mu_2)$, *they grow at most linearly in* x *for bounded* $\int_{\mathbb{R}} |x| d\mu_i(x)$, *and for some* $\lambda_i > 0$ *and all* $(\mu_1, \mu_2), (\bar{\mu}_1, \bar{\mu}_2) \in \mathcal{P}_1(\mathbb{R})^2$ *with a* C^2 *density*

$$(5.14)$$
$$\int_{\mathbb{R}} \sum_{i=1}^2 \lambda_i [F_i(x, \mu_1, \mu_2) - F_i(x, \bar{\mu}_1, \bar{\mu}_2)] d(\mu_i - \bar{\mu}_i)(x) > 0, \text{ if } M(\mu_1) \neq M(\bar{\mu}_1) \text{ or } M(\mu_2) \neq M(\bar{\mu}_2),$$

$$(5.15) \qquad \int_{\mathbb{R}} \sum_{i=1}^2 \lambda_i [G_i(x, \mu_1, \mu_2) - G_i(x, \bar{\mu}_1, \bar{\mu}_2)] d(\mu_i - \bar{\mu}_i)(x) \geq 0.$$

Then there is at most one classical solution (v_1, v_2, m_1, m_2) *of the problem* (5.1) *such that each* v_i *and its derivatives are bounded by* $C(1 + |x|)$.

Proof. We follow the proof of Theorem 4.1. Let $(\bar{v}_1, \bar{v}_2, \bar{m}_1, \bar{m}_2)$ be a second solution. We multiply the equations of the i-th population by λ_i and add them over i. After using the convexity of H_i,

as in the proof of Theorem 4.1, we reach

$$\int_{\mathbb{R}} \sum_{i=1}^{2} \lambda_i [G_i(x, m_1(T), m_2(T)) - G_i(x, \bar{m}_1(T), \bar{m}_2(T))] d(m_i(T) - \bar{m}_i(T))(x) +$$

$$\int_0^T \int_{\mathbb{R}} \sum_{i=1}^{2} \lambda_i [F_i(x, m_1(t), m_2(t)) - F_i(x, \bar{m}_1(t), \bar{m}_2(t))] d(m_i(t) - \bar{m}_i(t))(x) \, dt \le 0.$$

Then (5.14) and (5.15) imply $M(m_1(t)) = M(\bar{m}_1(t))$ and $M(m_2(t)) = M(\bar{m}_2(t))$ for all t. Since F_i and G_i depend only on $M(\mu_1), M(\mu_2)$, from the HJB equations we get $v_i = \bar{v}_i$ and finally the KFP equations give $m_i = \bar{m}_i$, $i = 1, 2$. $\qquad\square$

We consider the following example

(5.16) $$F_i(x, \mu_1, \mu_2) = \alpha_i x M(\mu_1) + \beta_i x M(\mu_2) + f_i(\mu_1, \mu_2), \quad i = 1, 2,$$

(5.17) $$G_i(x, \mu_1, \mu_2) = \gamma_i x M(\mu_1) + \delta_i x M(\mu_2) + g_i(\mu_1, \mu_2), \quad i = 1, 2,$$

with $\alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R}$, $f_i, g_i : \mathcal{P}_1(\mathbb{R})^2 \to \mathbb{R}$.

Corollary 5.4. *Let H_i be given by (5.2) and F_i, G_i of the form (5.16), (5.17) with f_i, g_i d_1-continuous. Then the problem (5.1) has at most one solution with derivatives bounded by $C(1+|x|)$, if $\nu_i \, dx \in \mathcal{P}_1(\mathbb{R})$ and there exists $\lambda > 0$ such that the matrices*

$$M_1 := \begin{pmatrix} \lambda\alpha_1 & \lambda\beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix}, \qquad M_2 := \begin{pmatrix} \lambda\gamma_1 & \lambda\delta_1 \\ \gamma_2 & \delta_2 \end{pmatrix}$$

are, respectively, positive definite and positive semi-definite; on the other hand, it has at least two solutions with derivatives bounded by $C(1 + |x|)$ if $\nu_i \in \tilde{\mathcal{P}}_1(\mathbb{R})$, $M(\nu_i) = 0$, and either

(5.18) $$\alpha_i, \beta_i, \gamma_i, \delta_i \le 0, \quad \gamma_i + \delta_i < 0, \quad i = 1, 2,$$

or

(5.19) $$\alpha_1, \beta_2, \gamma_1, \delta_2 \le 0, \quad \alpha_2, \beta_1, \gamma_2, \delta_1 \ge 0, \quad \gamma_1 < \delta_1, \quad \gamma_2 > \delta_2.$$

Proof. It is easy to compute the integral in (5.14) and get, after normalizing λ_2 to 1 and setting $\lambda_1 = \lambda$,

$$\lambda\alpha_1(M(m_1) - M(\bar{m}_1))^2 + (\lambda\beta_1 + \alpha_2)(M(m_1) - M(\bar{m}_1))(M(m_2) - M(\bar{m}_2)) + \beta_2(M(m_2) - M(\bar{m}_2))^2,$$

so condition (5.14) is satisfied if the matrix M_1 is positive definite. Similarly, condition (5.15) is satisfied if the quadratic form associated to M_2 is positive semi-definite.

As for non-uniqueness, the first statement follows from Prop. 5.1 and the second from Prop. 5.2. $\qquad\square$

Remark 5.2. The conditions for the definiteness of M_1

(5.20) $$\alpha_1 > 0, \qquad \beta_2 > 0, \qquad \lambda\alpha_1\beta_2 > (\lambda\beta_1 + \alpha_2)^2/4,$$

require not only the form of crowd-aversion within each population explained in Remark 4.3 for the case of a single population, but also that the costs for intraspecific interactions are dominant over the costs of the interactions of a population with the other. The same holds for the conditions of semi-definiteness of M_2, i.e.,

$$\gamma_1 \ge 0, \qquad \delta_2 \ge 0, \qquad \lambda\gamma_1\delta_2 \ge (\lambda\delta_1 + \gamma_2)^2/4.$$

On the other hand, the hypotheses of the examples of non-uniqueness (5.18) or (5.19) hold only if the intraspecific costs are null or imitation is rewarding within each population.

Note also that the gap between the sufficient conditions for uniqueness and for non-uniqueness is larger here than in the case of a single population.

Example 5.1. Consider a terminal cost G as in Corollary 5.4 and satisfying

$$\gamma_1, \delta_2 \geq 0, \qquad \delta_1 = \gamma_2 = 0,$$

so that, in particular, there is no cost or gain for interspecific interactions at the terminal time T. Then the sufficient condition for uniqueness reduces to (5.20) for some $\lambda > 0$, and this holds under the simple conditions

(5.21) $$\alpha_1, \beta_2 > 0, \quad \alpha_1\beta_2 > \beta_1\alpha_2,$$

as it can be easily seen by choosing $\lambda = \frac{2\alpha_1\beta_2 - \beta_1\alpha_2}{\beta_1^2}$ if $\beta_1 \neq 0$, and $\lambda = \frac{2\alpha_1\beta_2 - \beta_1\alpha_2}{\alpha_2^2}$ if $\alpha_2 \neq 0$.

Remark 5.3. If Theorem 5.3 is specialized to stationary equations, it improves slightly upon the uniqueness result in [15] because it allows to choose the parameters λ_i. This can be seen in the case of local and linear costs

$$F_i(x, \mu_1, \mu_2) = \alpha_i\mu_1(x) + \beta_i\mu_2(x), \quad i = 1, 2,$$

In fact, the integrand of the integral in (5.14) is

$$\lambda\alpha_1(m_1 - \bar{m}_1)(x)^2 + (\lambda\beta_1 + \alpha_2)(m_1 - \bar{m}_1)(x)(m_2 - \bar{m}_2)(x) + \beta_2(m_2 - \bar{m}_2)(x)^2,$$

and then the condition (5.14) is satisfied again if the matrix M_1 is positive definite. Hence (5.21) is a sufficient condition for uniqueness of the stationary MFG equations for this local case, more general than the assumption in [15].

Remark 5.4. A uniqueness result with assumptions of short time horizon and smooth Hamiltonian replacing the convexity of H and monotonicity of the costs, similar to Theorem 4.3, can be proved also for systems with several population such as (5.1). This is done in [5] for Neumann boundary conditions in bounded domains.

Appendix A. Density estimate

Let $\sigma > 0$. Let $b \colon [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ be bounded and measurable. Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ be such that ν is absolutely continuous with respect to Lebesgue measure with bounded density, that is,

$$\frac{d\nu}{d\lambda_d}(.) = m_0(.)$$

for some bounded and measurable $m_0 \colon \mathbb{R}^d \to [0, \infty)$ with $\int m_0(x)dx = 1$.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ be a standard filtered probability space carrying a d-dimensional $(\mathcal{F}_t)$-Wiener process, an $\mathbb{R}^d$-valued $\mathcal{F}_0$-measurable random variable ξ with distribution $\mathbf{P} \circ (\xi)^{-1} = \nu$ and an $\mathbb{R}^d$-valued continuous $(\mathcal{F}_t)$-adapted process X such that

(A.1) $$X(t) = \xi + \int_0^t b(s, X(s))\, ds + \sigma W(t), \quad t \in [0, T].$$

We also assume—as we may—that $(\Omega, \mathcal{F})$ is a Borel space.

Let $\mathfrak{m}$ be the flow of marginal distributions of X:

$$\mathfrak{m}(t) \doteq \mathbf{P} \circ (X(t))^{-1}, \quad t \in [0, T].$$

Lastly, for $t \in (0, T]$, let p_t denote the density of the d-variate Gaussian distribution with mean zero and covariance matrix $t \operatorname{Id}_d$:

$$p_t(y) \doteq (2\pi t)^{-d/2} e^{-\frac{|y|^2}{2t}}, \quad y \in \mathbb{R}^d.$$

For the following L^∞-estimate on the Lebesgue densities of the flow of measures $\mathfrak{m}$, we use the Girsanov transformation and a conditioning argument in the spirit of Exercise 7.4 in [22, p. 170]. In [42], sharp estimates on the transition probability densities are obtained through more sophisticated probabilistic methods.

Proposition A.1. *Under the above assumptions, the marginal distribution of X at any time $t \in [0, T]$ is absolutely continuous with respect to Lebesgue measure with bounded density. More precisely, for every $t \in (0, T]$, there exists a bounded measurable function $m_t \colon \mathbb{R}^d \to [0, \infty)$ with $\int m_t(x)dx = 1$ such that*

$$\frac{d\mathfrak{m}(t)}{d\lambda_d}(.) = m_t(.) \qquad and \qquad \|m_t\|_\infty \le \hat{C}_t \cdot \|m_0\|_\infty,$$

where $\hat{C}_t = \hat{C}_t(\sigma, \|b\|_\infty, d)$ is a finite constant that need not be greater than

$$\left(e^{\frac{8\|b\|_\infty^2}{\sigma^2} t} + \frac{4\|b\|_\infty}{\sigma} \sqrt{2\pi t} \cdot e^{\frac{16\|b\|_\infty^2}{\sigma^2} t} \right)^{d/4} \cdot e^{\frac{\|b\|_\infty^2}{2\sigma^2} t} \cdot \int_{\mathbb{R}^d} e^{\frac{\|b\|_\infty}{\sigma^2} |x|} p_{\sigma^2 t}(x) dx.$$

Proof. For $x \in \mathbb{R}^d$, define the process $Y^x = (Y_1^x, \dots, Y_d^x)$ through

$$Y^x(t) \doteq x + \sigma W(t), \quad t \in [0, T],$$

and a process Z^x over $[0, T]$ by

$$Z^x(t) \doteq \exp\left(\frac{1}{\sigma} \int_0^t b\left(s, Y^x(s)\right) \cdot dW(s) - \frac{1}{2\sigma^2} \int_0^t |b\left(s, Y^x(s)\right)|^2 ds \right)$$

$$= \exp\left(\frac{1}{\sigma} \sum_{i=1}^d \int_0^t b_i\left(s, Y^x(s)\right) dW_i(s) - \frac{1}{2\sigma^2} \int_0^t \sum_{i=1}^d |b_i\left(s, Y^x(s)\right)|^2 ds \right).$$

Girsanov's theorem (for instance, Chapter 7 in [22] or Sections 3.5 and 5.3.B in [36]) and the fact that ξ and W are independent yield, for $t \in [0, T]$, every bounded measurable function $g \colon \mathbb{R}^d \to \mathbb{R}$,

$$\mathbf{E}\left[g\left(X(t)\right)\right] = \int_{\mathbb{R}^d} \mathbf{E}\left[g\left(Y^x(t)\right) Z^x(t)\right] \nu(dx).$$

This implies (taking indicator functions for g) that

(A.2) $$\mathfrak{m}(t, B) = \int_{\mathbb{R}^d} \mathbf{E}\left[\mathbf{1}_B\left(Y^x(t)\right) Z^x(t)\right] \nu(dx) \text{ for all } B \in \mathcal{B}(\mathbb{R}^d).$$

By construction, $\mathfrak{m}(0) = \nu$. Fix $t \in (0, T]$.

For $y \in \mathbb{R}^d$ define a process $\tilde{W}^{t,y}$ over $[0, t]$ by

$$\tilde{W}^{t,y}(s) \doteq \begin{cases} \frac{s}{t} y + (t - s) \int_0^s \frac{1}{t-r} dW(r) & \text{if } s \in [0, t) \\ y & \text{if } s = t. \end{cases}$$

Then $\tilde{W}^{t,y}$ is a d-dimensional Brownian bridge (with respect to $\mathbf{P}$) from 0 to y over time interval $[0, t]$. The process $\tilde{W}^{t,y}$ is continuous on $[0, t]$ $\mathbf{P}$-almost surely; cf. Section 5.6.B in [36, pp. 358-360]; its Itô differential is given by

(A.3) $$d\tilde{W}^{t,y}(s) = \frac{y}{t} ds - \left(\int_0^s \frac{1}{t-r} dW(r) \right) ds + dW(s), \quad s \in [0, t].$$

That (A.3) holds if $s < t$ is clear by Itô's formula and the definition of $\tilde{W}^{t,y}$. To see that it also holds if $s = t$, apply Fubini's theorem, then the Cauchy-Schwarz inequality followed by the Itô isometry, to find that

$$\mathbf{E}\left[\int_0^t \left| \int_0^s \frac{1}{t-r} dW(r) \right| ds \right] = \int_0^t \mathbf{E}\left[\left| \int_0^s \frac{1}{t-r} dW(r) \right| \right] ds$$

$$\le \int_0^t \sqrt{\mathbf{E}\left[\int_0^s \frac{dr}{(t-r)^2} \right]} ds = \int_0^t \sqrt{\frac{s}{t(t-s)}} ds \le \int_0^t \frac{ds}{\sqrt{t-s}} < \infty.$$

Using also the continuity of $\tilde{W}^{t,y}$, it follows that with $\mathbf{P}$-probability one, for all $t_1, t_2 \in [0, t]$ such that $t_1 \le t_2$,

$$\tilde{W}^{t,y}(t_2) - \tilde{W}^{t,y}(t_1) = \frac{t_2 - t_1}{t} y - \int_{t_1}^{t_2} \left(\int_0^s \frac{1}{t-r} dW(r) \right) ds + W(t_2) - W(t_1),$$

which is equivalent to (A.3). In particular, $\tilde{W}^{t,y}$ is a vector of continuous semimartingales on $[0,t]$, with the same cross-variation processes as W. Moreover (for instance, Exercise 5.6.17 in [36, p. 361]), we have for all $\hat{g}\colon \mathbf{C}([0,t],\mathbb{R}^d)\to\mathbb{R}$ bounded and measurable,

$$(\mathrm{A.4}) \qquad \mathbf{E}\left[\hat{g}(W)\right] = \int_{\mathbb{R}^d} \mathbf{E}\left[\hat{g}(\tilde{W}^{t,y})\right] p_t(y)dy,$$

where p_t is the density of the d-variate Gaussian distribution with mean zero and covariance matrix $t\,\mathrm{Id}_d$. Formula (A.4) corresponds to conditioning the Wiener process W on its values at time t. To be more precise, we choose a regular conditional distribution of $\mathbf{P}$ given $W(t)$; cf. Theorem 6.3 in [35, p. 107]. Since $(\Omega,\mathcal{F})$ is Borel, there exists a probability kernel $\kappa_t\colon \mathbb{R}^d\times\mathcal{F}\to[0,1]$ such that for every $A\in\mathcal{F}$,

$$\kappa_t(W(t),A) = \mathbf{E}\left[\mathbf{1}_A\,\middle|\,W(t)\right] \quad \mathbf{P}\text{-almost surely.}$$

The probability measures $\kappa_t(y,\cdot)$, $y\in\mathbb{R}^d$, are uniquely determined Lebesgue almost everywhere. In view of (A.4), we have

$$(\mathrm{A.5}) \qquad \kappa_t(y,\cdot)\circ(W)^{-1} = \mathbf{P}\circ(\tilde{W}^{t,y})^{-1} \text{ for Lebesgue almost every } y\in\mathbb{R}^d.$$

We are going to apply the above regular conditional distribution to representation (A.2). By construction, for all $B\in\mathcal{B}(\mathbb{R}^d)$, $x\in\mathbb{R}^d$,

$$\mathbf{E}\left[\mathbf{1}_B\left(Y^x(t)\right)Z^x(t)\right] = \mathbf{E}\left[\mathbf{1}_B\left(x+\sigma W(t)\right)\exp\left(\frac{1}{\sigma}\int_0^t b\left(s,x+\sigma W(s)\right)\cdot dW(s)\right.\right.$$
$$\left.\left. -\frac{1}{2\sigma^2}\int_0^t |b\left(s,x+\sigma W(s)\right)|^2\,ds\right)\right].$$

Recalling that p_t is the density of the law of $W(t)$, we find that

$$\mathbf{E}\left[\mathbf{1}_B\left(Y^x(t)\right)Z^x(t)\right] = \int_{\mathbb{R}^d} \mathbf{E}_{\kappa_t(y,\cdot)}\left[\mathbf{1}_B\left(Y^x(t)\right)Z^x(t)\right] p_t(y)dy.$$

Setting

$$\Psi_t(x,y) \doteq \mathbf{E}\left[\exp\left(\frac{1}{\sigma}\int_0^t b\left(s,x+\sigma\tilde{W}^{t,y}(s)\right)\cdot d\tilde{W}^{t,y}(s) - \frac{1}{2\sigma^2}\int_0^t \left|b\left(s,x+\sigma\tilde{W}^{t,y}(s)\right)\right|^2 ds\right)\right]$$

we have, by (A.5),

$$(\mathrm{A.6}) \qquad \mathbf{E}\left[\mathbf{1}_B\left(Y^x(t)\right)Z^x(t)\right] = \int_{\mathbb{R}^d} \mathbf{1}_B\left(x+\sigma y\right)\Psi_t(x,y)p_t(y)dy.$$

Thanks to Eq. (A.3), $\Psi_t(x,y)$ can be expressed as

$$\Psi_t(x,y) = \mathbf{E}\left[\exp\left(\frac{1}{\sigma t}\int_0^t y\cdot b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)ds\right)\right.$$
$$\cdot\exp\left(-\frac{1}{\sigma}\int_0^t b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)\cdot\left(\int_0^s \frac{1}{t-r}dW(r)\right)ds\right)$$
$$\left.\cdot\exp\left(\frac{1}{\sigma}\int_0^t b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)\cdot dW(s) - \frac{1}{2\sigma^2}\int_0^t \left|b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)\right|^2 ds\right)\right].$$

Below, we will show that for every $c\in[0,\infty)$,

$$(\mathrm{A.7}) \qquad \mathbf{E}\left[\exp\left(c\sum_{i=1}^d \int_0^t \left|\int_0^s \frac{1}{t-r}dW_i(r)\right|ds\right)\right] \le \left(e^{2c^2t} + 2c\sqrt{2\pi t}\cdot e^{4c^2t}\right)^{d/2} < \infty.$$

Thus, Ψ_t is well defined as a (measurable) function $\mathbb{R}^d\times\mathbb{R}^d\to(0,\infty)$. By inequality (A.7), the boundedness of b, the Cauchy-Schwarz inequality and the (super-)martingale property of the

stochastic exponential of a martingale, we obtain

$$\Psi_t(x,y) \le e^{\frac{\|b\|_\infty}{\sigma}|y|} \cdot \mathbf{E}\left[\exp\left(\frac{2\|b\|_\infty}{\sigma}\sum_{i=1}^{d}\int_0^t \left|\int_0^s \frac{1}{t-r}dW_i(r)\right|ds\right)\right]^{1/2}$$

$$\cdot \mathbf{E}\left[e^{\frac{1}{\sigma}\int_0^t 2b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)\cdot dW(s) - \frac{1}{4\sigma^2}\int_0^t \left|2b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)\right|^2 ds}\right]^{1/2}.$$

$$\le e^{\frac{\|b\|_\infty}{\sigma}|y|} \cdot \left(e^{\frac{8\|b\|_\infty^2}{\sigma^2}t} + \frac{4\|b\|_\infty}{\sigma}\sqrt{2\pi t}\cdot e^{\frac{16\|b\|_\infty^2}{\sigma^2}t}\right)^{d/4}$$

$$\cdot \underbrace{\mathbf{E}\left[e^{\frac{1}{\sigma}\int_0^t 2b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)\cdot dW(s) - \frac{1}{2\sigma^2}\int_0^t \left|2b\left(r,x+\sigma\tilde{W}^{t,y}(s)\right)\right|^2 ds}\right]^{1/2}}_{=1} \cdot e^{\frac{\|b\|_\infty^2}{2\sigma^2}t},$$

hence

(A.8) $$\sup_{x\in\mathbb{R}^d}\Psi_t(x,y) \le C_t \cdot e^{\frac{\|b\|_\infty}{\sigma}|y|}, \quad y\in\mathbb{R}^d,$$

where the finite constant $C_t = C_t(\sigma,\|b\|_\infty,d)$ is given by

$$C_t(\sigma,\|b\|_\infty,d) \doteq \left(e^{\frac{8\|b\|_\infty^2}{\sigma^2}t} + \frac{4\|b\|_\infty}{\sigma}\sqrt{2\pi t}\cdot e^{\frac{16\|b\|_\infty^2}{\sigma^2}t}\right)^{d/4}\cdot e^{\frac{\|b\|_\infty^2}{2\sigma^2}t}.$$

Recalling (A.2), (A.6) and the hypothesis that ν has density m_0, we see that for all $B\in\mathcal{B}(\mathbb{R})$,

$$\mathfrak{m}(t,B) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}_B\left(x+\sigma y\right)\Psi_t(x,y)p_t(y)m_0(x)dxdy$$

$$= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}_B(z)\Psi_t\left(x,\frac{z-x}{\sigma}\right)p_{\sigma^2 t}(z-x)m_0(x)dxdz.$$

It follows that $\mathfrak{m}(t)$ possesses a density with respect to Lebesgue measure:

$$\frac{d\mathfrak{m}(t)}{d\lambda_d}(z) = m_t(z) \doteq \int_{\mathbb{R}^d}\Psi_t\left(x,\frac{z-x}{\sigma}\right)p_{\sigma^2 t}(z-x)m_0(x)dx, \quad z\in\mathbb{R}^d.$$

Thanks to (A.8), the density is bounded:

$$\|m_t\|_\infty \le C_t(\sigma,\|b\|_\infty,d)\cdot \sup_{z\in\mathbb{R}^d}\int_{\mathbb{R}^d}e^{\frac{\|b\|_\infty}{\sigma^2}|z-x|}p_{\sigma^2 t}(z-x)m_0(x)dx$$

$$\le C_t(\sigma,\|b\|_\infty,d)\cdot\|m_0\|_\infty\cdot\int_{\mathbb{R}^d}e^{\frac{\|b\|_\infty}{\sigma^2}|x|}p_{\sigma^2 t}(x)dx < \infty.$$

It remains to prove inequality (A.7). Let $c\in[0,\infty)$. Since $W_1,\ldots,W_d$ are independent one-dimensional Wiener processes, we have

$$\mathbf{E}\left[\exp\left(c\sum_{i=1}^{d}\int_0^t\left|\int_0^s\frac{1}{t-r}dW_i(r)\right|ds\right)\right] = \mathbf{E}\left[\exp\left(c\int_0^t\left|\int_0^s\frac{1}{t-r}dW_1(r)\right|ds\right)\right]^d.$$

Let $\mathrm{sgn}\colon\mathbb{R}\to\{-1,1\}$ denote the left-continuous version of the sign function, and denote by Leb_t Lebesgue measure on $[0,t]$. By (A.3) and the definition of $\tilde{W}^{t,0}$, we have with $\mathbf{P}$-probability one,

(A.9) $$\int_0^t\left|\int_0^s\frac{1}{t-r}dW_1(r)\right|ds = -\int_0^t\mathrm{sgn}\left(\tilde{W}_1^{t,0}(s)\right)d\tilde{W}_1^{t,0}(s) + \int_0^t\mathrm{sgn}\left(\tilde{W}_1^{t,0}(s)\right)dW_1(s)$$

since, for $\mathrm{Leb}_t\otimes\mathbf{P}$-almost all $(s,\omega)\in[0,t]\times\Omega$,

$$\left|\int_0^s\frac{1}{t-r}dW_1(r)\right| = \mathrm{sgn}\left(\int_0^s\frac{1}{t-r}dW_1(r)\right)\left(\int_0^s\frac{1}{t-r}dW_1(r)\right)$$

$$= \mathrm{sgn}\left(\tilde{W}_1^{t,0}(s)\right)\left(\int_0^s\frac{1}{t-r}dW_1(r)\right).$$

Let $\tilde{L}_1^{t,0}(0)$ be the local time at the origin of the continuous semimartingale $\tilde{W}_1^{t,0}$ according to Theorem 3.7.1 in [36, p. 218]. In particular, $s \mapsto \tilde{L}_s^{t,0}(0)$ is a continuous non-decreasing non-negative process with $\tilde{L}_0^{t,0}(0) = 0$ such that, for every $s \in [0, t]$,

$$\lim_{\epsilon \searrow 0} \frac{1}{4\epsilon} \int_0^s \mathbf{1}_{[-\epsilon,\epsilon]} \left(\tilde{W}_1^{t,0}(r) \right) dr = \tilde{L}_s^{t,0}(0) \quad \mathbf{P}\text{-almost surely;}$$

cf. Theorem 3.7.1(iii) and Problem 3.7.6 in [36]. By the Itô-Tanaka-Meyer formula applied to the absolute value function (see Eq. (3.7.9) in [36, p. 220]), we have

$$- \int_0^t \mathrm{sgn}\left(\tilde{W}_1^{t,0}(s) \right) d\tilde{W}_1^{t,0}(s) = |\tilde{W}^{t,0}(0)| - |\tilde{W}^{t,0}(t)| + 2\tilde{L}_t^{t,0}(0) = 2\tilde{L}_t^{t,0}(0).$$

This, together with (A.9), yields

$$\mathbf{E}\left[\exp\left(c \sum_{i=1}^d \int_0^t \left| \int_0^s \frac{1}{t-r} dW_i(r) \right| ds \right) \right]$$
$$= \mathbf{E}\left[\exp\left(c \left(2\tilde{L}_t^{t,0}(0) + \int_0^t \mathrm{sgn}\left(\tilde{W}_1^{t,0}(s) \right) dW_1(s) \right) \right) \right]^d.$$

Using again the Cauchy-Schwarz inequality and the (super-)martingale property of the stochastic exponential of a martingale, we find that

$$\mathbf{E}\left[\exp\left(c \sum_{i=1}^d \int_0^t \left| \int_0^s \frac{1}{t-r} dW_i(r) \right| ds \right) \right]$$
$$\leq \mathbf{E}\left[\exp\left(4c\tilde{L}_t^{t,0}(0) \right) \right]^{d/2} \cdot \mathbf{E}\left[\exp\left(2c \int_0^t \mathrm{sgn}\left(\tilde{W}_1^{t,0}(s) \right) dW_1(s) - \frac{4c^2}{2}t + \frac{4c^2}{2}t \right) \right]^{d/2}$$
$$\leq \mathbf{E}\left[\exp\left(4c\tilde{L}_t^{t,0}(0) \right) \right]^{d/2} \cdot e^{c^2 t \cdot d}.$$

The distribution of $2\tilde{L}_t^{t,0}(0)$ is known explicitly. Let $(L_s(0))_{s \geq 0}$ be the local time at the origin of the one-dimensional Wiener process W_1. Then, as a consequence of (A.5), we have that the distribution of $2\tilde{L}_t^{t,0}(0)$ coincides with the conditional distribution of $2L_t(0)$ given $W_1(t) = 0$. The joint distribution of $W_1(t)$ and $2L_t(0)$ is known to be absolutely continuous with respect to two-dimensional Lebesgue measure with density given by

$$\frac{\mathbf{P} \circ (W_1(t), 2L_t(0))^{-1}}{d\lambda_2}(w, l) = \underbrace{\mathbf{1}_{(0,\infty)}(l) \cdot \frac{l + |w|}{\sqrt{2\pi t^3}} \exp\left(-\frac{(l+|w|)^2}{2t} \right)}_{\doteq \phi(w,l)}, \quad w, l \in \mathbb{R},$$

see, for instance, Problem 6.3.4 in [36, p. 420]. Conditioning on $W_1(t) = 0$, we obtain

$$\frac{\mathbf{P} \circ (2\tilde{L}_t^{t,0}(0))^{-1}}{d\lambda_1}(l) = \frac{\phi(0,l)}{\int \phi(0,r)dr} = \mathbf{1}_{(0,\infty)}(l) \cdot \frac{l}{t} \exp\left(-\frac{l^2}{2t} \right), \quad l \in \mathbb{R}.$$

We therefore have

$$\mathbf{E}\left[\exp\left(4c\tilde{L}_t^{t,0}(0) \right) \right] = \int_0^\infty \exp\left(2c \cdot l \right) \frac{l}{t} \exp\left(-\frac{l^2}{2t} \right) dl = 1 + 2c \int_0^\infty \exp\left(2c \cdot l - \frac{l^2}{2t} \right) dl$$
$$= 1 + 2c \cdot e^{2c^2 t} \int_0^\infty \exp\left(-\frac{(l - 2tc)^2}{2t} \right) dl \leq 1 + 2c\sqrt{2\pi t} \cdot e^{2c^2 t}.$$

It follows that

$$\mathbf{E}\left[\exp\left(c \sum_{i=1}^d \int_0^t \left| \int_0^s \frac{1}{t-r} dW_i(r) \right| ds \right) \right] \leq \left(1 + 2c\sqrt{2\pi t} \cdot e^{2c^2 t} \right)^{d/2} \cdot e^{c^2 t \cdot d},$$

which proves (A.7).

$\square$

References

[1] S. Ahuja. Wellposedness of mean field games with common noise under a weak monotonicity condition. *SIAM J. Control Optim.*, 54(1):30–48, 2016.

[2] Y. Achdou, M. Bardi, M. Cirant: Mean Field Games models of segregation, Math. Models Methods Appl. Sci. 27 (2017), 75–113,

[3] D.M. Ambrose: Strong solutions for time-dependent mean field games with non-separable Hamiltonians, arXiv:1605.01745

[4] M. Bardi: Explicit solutions of some Linear-Quadratic Mean Field Games, Netw. Heterog. Media 7 (2012), 243–261

[5] M. Bardi, M. Cirant, Uniqueness of solutions in Mean Field Games with several populations and Neumann conditions, preprint 2017, arXiv:1709.02158, to appear in "PDE models for multi-agent phenomena" (P. Cardaliaguet, A. Porretta, F. Salvarani eds.), Springer INdAM Series

[6] M. Bardi, E. Feleqi: Nonlinear elliptic systems and mean field games, NoDEA Nonlinear Differential Equations Appl. 23 (2016), 23–44

[7] M. Bardi, F.S. Priuli: Linear-Quadratic N-person and Mean-Field Games with Ergodic Cost, SIAM J. Control Optim. 52 (2014), 3022–3052.

[8] V.I. Bogachev, N.V. Krylov, M. Röckner, S.V. Shaposhnikov: Fokker-Planck-Kolmogorov equations. Mathematical Surveys and Monographs, 207. American Mathematical Society, Providence, RI, 2015

[9] A. Briani, P. Cardaliaguet: Stable solutions in potential mean field game systems, preprint 2016, arXiv:1612.01877, to appear in NoDEA.

[10] P. Cardaliaguet: Notes on Mean Field Games (from P-L. Lions' lectures at Collège de France), 2010.

[11] P. Cardaliaguet, F. Delarue, J.-M. Lasry, P.-L. Lions: The master equation and the convergence problem in mean field games, preprint 2015, arXiv:1509.02505

[12] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. *SIAM J. Control Optim.*, 51(4): 2705–2734, 2013.

[13] R. Carmona, F. Delarue, and A. Lachapelle. Control of McKean-Vlasov dynamics versus mean field games. *Math. Financ. Econ.*, 7(2):131–166, 2013.

[14] R. Carmona, F. Delarue, and D. Lacker. Mean field games with common noise. *Ann. Probab.*, 44(6):3740–3803, 2016.

[15] M. Cirant: Multi-population mean field games systems with Neumann boundary conditions, J. Math. Pures Appl. 103 (2015), 1294–1315.

[16] M. Cirant: On the existence of oscillating solutions in non-monotone Mean-Field Games, preprint 2017, arXiv:1711.08047

[17] M. Cirant, D. Tonon: Time-dependent focusing Mean-Field Games: the sub-critical case, preprint 2017, arXiv:1704.04014

[18] S. Farinelli: Mean Field Games systems of partial differential equations: uniqueness of solutions, Master Thesis, University of Padua, 2018.

[19] M. Fischer. On the connection between symmetric N-player games and mean field games. *Ann. Appl. Probab.*, 27(2):757–810, 2017.

[20] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions.* Springer, New York, 2nd edition, 2006.

[21] A. Friedman: Partial differential equations of parabolic type. Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964

[22] A. Friedman. *Stochastic Differential Equations and Applications*, volume 1. Academic Press, New York, 1975.

[23] D.A. Gomes, J. Mohr, and R.R. Souza. Continuous time finite state mean field games. Appl. Math. Optim. 68 (2013), 99–143.

[24] D. Gomes, L. Nurbekyan and E. Pimentel: Economic models and mean-field games theory, IMPA Mathematical Publications, Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 2015.

[25] D.A. Gomes, L. Nurbekyan, M. Prazeres: One-Dimensional Stationary Mean-Field Games with Local Coupling, Dyn Games Appl (2017). https://doi.org/10.1007/s13235-017-0223-9

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

[27] D. Gomes, J. Saude: Mean field games models: a brief survey. Dyn. Games Appl. 4 (2014), 110–154

[28] P.J. Graber, A. Bensoussan: Existence and uniqueness of solutions for Bertrand and Cournot mean field games, Applied Math. Optim. (2016). https://doi.org/10.1007/s00245-016-9366-0

[29] O. Guéant, A reference case for mean field games models, J. Math. Pures Appl. (9), 92 (2009), 276–294.

[30] O. Guéant, J.-M. Lasry, P.-L. Lions: Mean field games and applications, in "Paris-Princeton Lectures on Mathematical Finance 2010" (eds. R. A. Carmona, et al.), Lecture Notes in Math., 2003, Springer, Berlin, (2011), 205–266.

[31] M. Huang, P.E. Caines and R.P. Malhamé, Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria, *IEEE Trans. Automat. Control* **52** (2007) 1560–1571.

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

[33] 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) 221–251.

[34] N. Ikeda and S. Watanabe. A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka J. Math.*, 14(3):619–633, 1977.

[35] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York, 2nd edition, 2001.

[36] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer, New York, 2nd edition, 1991.

[37] D. Lacker. Mean field games via controlled martingale problems: Existence of Markovian equilibria. *Stoch. Processes Appl.*, 125(7):2856–2894, 2015.

[38] D. Lacker. A general characterization of the mean field limit for stochastic differential games. *Probab. Theory Related Fields*, 165(3):581–648, 2016.

[39] J.-M. Lasry, P.-L. Lions: *Mean field games*, Jpn. J. Math. **2** (2007), 229–260.

[40] P.-L. Lions: Lectures at Collège de France 2008-9.

[41] J. Moon and T. Başar. Linear quadratic risk-sensitive and robust mean field games. *IEEE Trans. Automat. Control*, 62(3):1062–1077, 2016.

[42] Z. Qian, F. Russo, and W. Zheng. Comparison theorem and estimates for transition probability densities of diffusion processes. *Probab. Theory Relat. Fields*, 127(3):388–406, 2003.

[43] R. F. Tchuendom. Uniqueness for linear-quadratic mean field games with common noise. *Dyn. Games Appl.*, online, 2016.

[44] H.V. Tran: A note on nonconvex Mean Field Games , preprint 2016, arXiv:1612.04725

[45] B.-C. Wang and J.-F. Zhang. Mean field games for large-population multiagent systems with Markov jump parameters. *SIAM J. Control Optim.*, 50(4):2308–2334, 2012.

Department of Mathematics, University of Padova, Via Trieste 63, 35121 Padova, Italy
E-mail address: bardi@math.unipd.it, fischer@math.unipd.it