A Lagrangian approach for aggregative mean field games of controls with mixed and final constraints

Joseph Frédéric Bonnans, Justina Gianatti, Laurent Pfeiffer

A LAGRANGIAN APPROACH FOR AGGREGATIVE MEAN FIELD GAMES OF CONTROLS WITH MIXED AND FINAL CONSTRAINTS

J. FRÉDÉRIC BONNANS*, JUSTINA GIANATTI†, AND LAURENT PFEIFFER*

Abstract. The objective of this paper is to analyze the existence of equilibria for a class of deterministic mean field games of controls. The interaction between players is due to both a congestion term and a price function which depends on the distributions of the optimal strategies. Moreover, final state and mixed state-control constraints are considered, the dynamics being nonlinear and affine with respect to the control. The existence of equilibria is obtained by Kakutani's theorem, applied to a fixed point formulation of the problem. Finally, uniqueness results are shown under monotonicity assumptions.

Key words. Mean field games of controls, aggregative games, constrained optimal control, Lagrangian equilibria

AMS subject classifications. 49K15, 49N60, 49N80, 91A07, 91A16

1. Introduction. In this article we consider a Nash equilibrium problem involving a large number N of agents, each of them solving a deterministic optimal control problem involving control-affine nonlinear dynamics, final state constraints, and mixed state-control constraints. The agents may only differ from each other by their initial condition. The interaction between the agents is induced by a price variable and a congestion term, which are determined by the collective behavior of the agents. Our mathematical analysis focuses on an equilibrium problem which models the asymptotic limit when N goes to infinity and when each isolated agent is supposed to have no impact on the coupling terms (the price variable and the congestion term). Therefore the problem falls into the class of mean field games (MFGs), which have received considerable attention in the literature since their introduction in the pioneering works by Lasry and Lions [21, 22, 23] and Caines, Huang and Malhamé [18].

The main specificity of our model is the interaction induced by the price variable. In the cost function of each agent, the price penalizes linearly the control variable. It is defined as a monotonic function of some aggregative term that can be interpreted as a demand. Here it is the average value of the controls exerted by all agents. This kind of interaction is similar to the one in Cournot models in economics, where companies without market power compete on the amount of some product. Our model is representative from games in energy markets involving a large number of small storage devices and some endogenous price depending on the average speed of charge of the devices. See for instance [1, 13, 24, 26]. Another specificity of our model is the presence of mixed control-state constraints and final state constraints. They appear naturally in applications in electrical engineering: for example, when the storage devices must be fully (or partially) loaded at the end of the time frame. In the appendix, we motivate the use of mixed constraints with an example involving gas storages.

In most MFG models proposed in the literature, the agents interact only through

*Inria and Centre de Mathématiques Appliquées (CNRS, UMR 7641), Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France (Frederic.Bonnans@inria.fr, laurent.pfeiffer@inria.fr).

†CIFASIS-CONICET-UNR, Ocampo y Esmeralda, S2000EZP, Rosario, Argentina (gianatti@cifasis-conicet.gov.ar).

their position (their state variable). Mean field game models with interaction through the states and controls are now commonly called *MFGs of controls*. The terminologies *extended MFGs* and *strongly coupled MFGs* are also employed. Let us review the articles dedicated to such models. In [15], a stationary second order MFG of controls is studied. A deterministic MFG of controls is considered in [16]. An existence result has been obtained for a quite general MFG model in [11]. A uniqueness result is provided in [4]. The works [19, 20] analyse the existence and uniqueness of classical solutions in the second order case. An existence result is provided in the monograph [12, Section 4.6], for MFGs described by forward backward stochastic differential equations. The particular price interaction investigated in this article has been studied in [6] in the second order case and in [17] in the case of a degenerate diffusion and potential congestion terms.

Most MFG models consist of a coupled system of partial differential equations (PDEs), the Fokker-Planck equation and the Hamilton-Jacobi-Bellman (HJB) equation. The presence of final and mixed constraints in the underlying optimal control problem makes it difficult to characterize the behavior of a representative agent with the classical HJB approach. We therefore rely on a Lagrangian formulation of the problem, rather than on a PDE approach. More precisely, our equilibrium problem is posed on the set of Borel probability measures on the space of state-control trajectories. The Lagrangian approach has been employed in several references dealing with deterministic MFGs. Variational MFGs are studied in [3]. The article [25] deals with minimal-time MFGs. In [9], MFGs with linear dynamics are considered. The three articles [8, 10, 9] deal with state-constrained MFGs and with the connection between the Lagrangian and the PDE formulations.

At a methodological level, the common feature of almost all studies dedicated to MFGs of controls is the introduction of an auxiliary mapping, which allows to put the equilibrium problem in a reduced form that can be handled with a fixed point approach. In the PDE approach, the auxiliary mapping allows to express the control of a representative agent at a given time t in function of its current state x, the equilibrium distribution (of the states) and the gradient of the value function (see for example [12, Lemma 4.60], [11, Lemma 5.2] or [6, Lemma 5]). This relation is in general not explicit, contrary to MFGs with interaction through the state variable only. In the probabilistic approach of [16, Assumption G], the auxiliary mapping depends on t, x, and a pair of random variables (X_t, P_t), whose distribution coincides with the distribution of pairs of state-costate of all agents in the game. In [12, Lemma 4.61], the auxiliary mapping directly depends on the distribution of (X_t, P_t). Our roadmap is the same as the one used in the references mentioned above: we introduce an auxiliary mapping (of the same nature as the one in [12]) which allows to write the equilibrium problem in a reduced form which is then tractable with a fixed point argument. After reformulation, the equilibrium problem is posed on the set of Borel probability measures on the space of state-costate trajectories.

Our article is the first to propose a Lagrangian formulation for an MFG of controls. It is one of the very few publications dealing with first order MFGs of controls. (i) The article of Gomes and Voskanyan [16] is the closest to our work. Their analysis relies in a quite crucial manner on some regularity properties of the value function associated with the underlying optimal control problem (Lipschitz continuity, semiconcavity) which are easily demonstrated in their framework without constraints. Those properties are not needed in the Lagrangian framework. They could probably be established, but under stronger qualification conditions than those in force in the present work. Incidentally, the initial distribution of the agents must have a density in [16], which is not the case in the present work. (ii) Carmona and Delarue have an existence result, for an MFG of controls posed as a forward-backward stochastic differential equation, see [12, Proposition 4.64]. This model relies on Pontryagin's principle, which is a sufficient condition only under convexity assumptions on the underlying optimal control problem (see the assumption SMP [12, page 161]), which we do not need. Let us mention that their other result [12, Proposition 4.64] concerns the second order case. (iii) In a recent work, Graber, Mullenix and Pfeiffer have obtained the existence of a solution for an MFG of controls formulated as a coupled system of possibly degenerate PDEs. This work is restricted to the potential case, when the local congestion term is the derivative of some convex function. It also relies on a periodicity condition on the data functions, which we do not need here.

The paper is organized as follows: In Section 2 we present the problem that we address here, referred to as MFGC. We introduce the main notation and we define the different notions of equilibria that we use throughout this work. In Section 3 we study the optimal control problem associated with an individual player, providing optimality conditions and regularity of solutions. By a fixed point argument, in Section 4 we prove the existence of equilibria. In Section 5, under additional monotonicity assumptions we analyze the uniqueness of solutions.

2. Description of the aggregative MFGC problem.

2.1. Preliminaries. Let (X,d) be a separable metric space. We denote by $\mathcal{P}(X)$ the set of Borel probability measures on X. Given $p\in[1,+\infty)$, it is defined $\mathcal{P}_p(X)$ as the set of probability measures μ on X such that

$$\int_X d(x,x_0)^p \mathrm{d}\mu(x) < +\infty,$$

for some (and thus any) $x_0\in X$. The Monge-Kantorovich distance on $\mathcal{P}_p(X)$ is given by

$$d_p(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \left[\int_X d(x,y)^p \mathrm{d}\pi(x,y)\right]^{\frac{1}{p}},$$

where $\Pi(\mu,\nu)$ denotes the set of probability measures on $X\times X$ with first and second marginals equal to μ and ν respectively. In this paper, we work with $p=1$. For all $\mu,\nu\in\mathcal{P}_1(X)$, we have the following formula (see [14, Theorem 11.8.2]):

$$d_1(\mu,\nu) = \sup\left\{\int_X f(x)\mathrm{d}\mu(x) - \int_X f(x)\mathrm{d}\nu(x) \mid f\colon X\to\mathbb{R} \text{ is 1-Lipschitz}\right\}.$$

We recall the definition of narrow convergence of measures. We say that the sequence $(\mu_n)_{n\in\mathbb{N}}\subset\mathcal{P}(X)$ narrowly converges to $\mu\in\mathcal{P}(X)$ if

$$\lim_{n\to\infty}\int_X f(x)\mathrm{d}\mu_n(x) = \int_X f(x)\mathrm{d}\mu(x), \quad \forall f\in C_b^0(X),$$

where $C_b^0(X)$ denotes the set of all continuous and bounded real functions defined on X. Throughout this work we endow the space $\mathcal{P}_1(X)$ with the narrow topology. As a consequence of [2, Proposition 7.1.5], for any compact set $K\subset X$, we have for all $p\geq 1$, $\mathcal{P}(K)=\mathcal{P}_p(K)$ and d_p metricizes the narrow convergence of probability measures on the set $\mathcal{P}(K)$. In addition, $\mathcal{P}(K)$ is compact.

2.2. MFG equilibria and main notation. We start by defining the optimal control problem that each agent aims to solve, assuming that the price and the distribution of the other players are known. The problem takes the form of a constrained minimization problem parameterized by the initial condition $x_0 \in \mathbb{R}^n$, the agents distribution $m \in C([0,T];\mathcal{P}_1(\mathbb{R}^n))$ and the price $P \in L^\infty(0,T;\mathbb{R}^m)$.

Let $\Gamma := H^1(0,T;\mathbb{R}^n)$ be equipped with the norm $\|\cdot\|_\infty$. This space is continuously embedded in $C(0,T;\mathbb{R}^n)$ for the supremum norm, denoted by $\|\cdot\|_\infty$. Given $x_0 \in \mathbb{R}^n$, we define $\Gamma[x_0]$ by

$$\Gamma[x_0] = \big\{\gamma \in \Gamma :\, \gamma(0) = x_0\big\}.$$

We take $L^2(0,T;\mathbb{R}^m)$ as the control space, which we denote by $\mathcal{U}$. We denote by $\mathcal{K}[x_0]$ the feasible set that is defined by

$$\mathcal{K}[x_0] = \left\{ (\gamma, v) \in \Gamma \times \mathcal{U} : \quad \begin{array}{ll} \dot\gamma(t) = a(\gamma(t)) + b(\gamma(t))v(t), & \text{for a.e. } t \in (0,T), \\ \gamma(0) = x_0, & \\ c(\gamma(t), v(t)) \le 0, & \text{for a.e. } t \in (0,T), \\ g_1(\gamma(T)) = 0, & \\ g_2(\gamma(T)) \le 0 & \end{array} \right\}.$$

The dynamics coefficients are $a\colon \mathbb{R}^n \to \mathbb{R}^n$ and $b\colon \mathbb{R}^n \to \mathbb{R}^{n\times m}$ (note that $b_i(x) \in \mathbb{R}^n$ will denote the i-th column of $b(x)$). The final equality and inequality constraint functions are, respectively, $g_1\colon \mathbb{R}^n \to \mathbb{R}^{n_{g_1}}$, and $g_2\colon \mathbb{R}^n \to \mathbb{R}^{n_{g_2}}$, and the state-control constraint function is $c\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n_c}$. Now we define the cost functional $J[m,P]\colon \Gamma \times \mathcal{U} \to \mathbb{R}$ as

$$J[m,P](\gamma,v) = \int_0^T \big(L(\gamma(t),v(t)) + \langle P(t), v(t)\rangle + f(\gamma(t), m(t))\big)\,\mathrm{d}t + g_0(\gamma(T), m(T)).$$

Here $L\colon \mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}$ represents the running cost of the agents, $f\colon \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}$, the congestion function, and $g_0\colon \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n) \to \mathbb{R}$ is the final cost. Therefore, the optimal control problem that each agent addresses is

$$\underset{(\gamma,v)\in\mathcal{K}[x_0]}{\operatorname{Min}} J[m,P](\gamma,v). \tag{2.1}$$

The set of optimal trajectories for this minimization problem is denoted by

$$\Gamma[m,P,x_0] = \{\bar\gamma \in \Gamma[x_0] \ :\ \exists \bar v \in \mathcal{U},\, (\bar\gamma,\bar v) \text{ is a solution to (2.1)}\}.$$

2.2.1. Lagrangian MFGC equilibria. In the previous paragraph, we have described the optimization problem, for a particular player, given the price and the agents distribution. We describe now how the price is related to the collective behavior of all agents and give a Lagrangian description of our mean field game.

Let $m_0 \in \mathcal{P}_1(\mathbb{R}^n)$ be the initial distribution of the agents. We also fix a price function $\psi\colon \mathbb{R}^m \to \mathbb{R}^m$. For $t \in [0,T]$, the mapping $e_t : \Gamma \times \mathcal{U} \to \mathbb{R}^n$ is given by $e_t(\gamma, v) = \gamma(t)$. We define the set

$$\mathcal{P}_{m_0}(\Gamma \times \mathcal{U}) = \big\{\eta \in \mathcal{P}_1(\Gamma\times\mathcal{U}) \ :\ e_0\sharp\eta = m_0\big\}.$$

Given $\eta \in \mathcal{P}_{m_0}(\Gamma \times \mathcal{U})$, we define the cost functional $J^\eta = J[m^\eta, P^\eta]$, where the coupling terms $m^\eta\colon t \in [0,T] \mapsto m_t^\eta \in \mathcal{P}_1(\mathbb{R}^n)$ and $P^\eta \in L^\infty(0,T;\mathbb{R}^m)$ are given by

$$m_t^\eta = e_t\sharp\eta \quad\text{and}\quad P^\eta = \Psi\Big(\int_{\Gamma\times\mathcal{U}} v\,\mathrm{d}\eta(\gamma,v)\Big). \tag{2.2}$$

The continuity of the map $t \mapsto m_t^\eta$ will be ensured by Lemma 4.4. In the definition of P^η, $\int_{\Gamma\times\mathcal{U}} v \, \mathrm{d}\eta(\gamma,v)$ is a Bochner integral with value in $L^2(0,T;\mathbb{R}^m)$ (which is well defined since $\eta \in \mathcal{P}_1(\Gamma\times\mathcal{U})$) and the map $\Psi\colon \theta \in L^2(0,T;\mathbb{R}^m) \to \Psi[\theta] \in L^\infty(0,T;\mathbb{R}^m)$ denotes the Nemytskii operator associated with the price function $\psi\colon \mathbb{R}^m \to \mathbb{R}^m$, defined by $\Psi[\theta](t) = \psi(\theta(t))$, for a.e. $t \in (0,T)$.

Given $x_0 \in \mathbb{R}^n$ and $\eta \in \mathcal{P}_{m_0}(\Gamma \times \mathcal{U})$, we denote by $\Gamma^\eta[x_0]$ the set of optimal state-control trajectories associated with the cost J^η and set of constraints $\mathcal{K}[x_0]$:

$$\Gamma^\eta[x_0] = \Big\{ (\bar\gamma,\bar v) \in \mathcal{K}[x_0] : \; J^\eta(\bar\gamma,\bar v) \leq J^\eta(\gamma,v) \;\; \forall (\gamma,v) \in \mathcal{K}[x_0] \Big\}.$$

DEFINITION 2.1. *We call Lagrangian MFGC equilibrium any distribution $\eta \in \mathcal{P}_{m_0}(\Gamma \times \mathcal{U})$ supported on optimal trajectories, i.e.*

$$\operatorname{supp}(\eta) \subset \bigcup_{x \in \operatorname{supp}(m_0)} \Gamma^\eta[x].$$

2.2.2. Auxiliary MFGC equilibria. In order to analyze the existence of Lagrangian MFGC equilibria, we propose here a new notion of equilibrium, that we call auxiliary equilibrium. We set

$$\tilde\Gamma = H^1(0,T;\mathbb{R}^n) \times H^1(0,T;\mathbb{R}^n).$$

We equip $\tilde\Gamma$ with the supremum norm, defined by $\max(\|\gamma\|_\infty, \|p\|_\infty)$ for a given pair $(\gamma,p) \in \tilde\Gamma$. We denote it (by extension) $\|(\gamma,p)\|_\infty$. For any $x_0 \in \mathbb{R}^n$, we define

$$\tilde\Gamma[x_0] = \big\{ (\gamma,p) \in \tilde\Gamma \, : \, \gamma(0) = x_0 \big\}.$$

Given $t \in [0,T]$, we consider the mappings $\tilde e_t \colon \tilde\Gamma \to \mathbb{R}^n$ and $\hat e_t \colon \tilde\Gamma \to \mathbb{R}^n \times \mathbb{R}^n$ defined by $\tilde e_t(\gamma,p) = \gamma(t)$ and $\hat e_t(\gamma,p) = (\gamma(t),p(t))$, for all $(\gamma,p) \in \tilde\Gamma$. We denote

$$\mathcal{P}_{m_0}(\tilde\Gamma) = \big\{ \kappa \in \mathcal{P}_1(\tilde\Gamma) : \, \tilde e_0 \sharp \kappa = m_0 \big\}.$$

Given a distribution $\kappa \in \mathcal{P}_1(\tilde\Gamma)$, we consider the cost functional $\tilde J^\kappa := J[\tilde m^\kappa, \tilde P^\kappa]$, where $\tilde m_t^\kappa = \tilde e_t \sharp \kappa$, for $t \in [0,T]$ and where $\tilde P^\kappa(t)$ is constructed as an auxiliary function of the distribution $\hat e_t \sharp \kappa$ at time t in Lemma 4.1. The well-posedness of $\tilde J^\kappa$ is established in Remark 4.5. Once $\tilde P^\kappa$ has been defined, we can consider the set of optimal trajectories and associated adjoint states $\tilde\Gamma^\kappa[x_0]$ defined by

$$\tilde\Gamma^\kappa[x_0] = \Big\{ (\bar\gamma,p) \in \tilde\Gamma[x_0] \, : \, \bar\gamma \in \Gamma[\tilde m^\kappa, \tilde P^\kappa, x_0] \text{ and } p \text{ costate associated with } \bar\gamma \Big\}.$$

The precise meaning of "associated costate" will be given in Definition 3.7.

DEFINITION 2.2. *A measure $\kappa \in \mathcal{P}_{m_0}(\tilde\Gamma)$ is an auxiliary MFGC equilibrium if*

$$\operatorname{supp}(\kappa) \subset \bigcup_{x \in \operatorname{supp}(m_0)} \tilde\Gamma^\kappa[x].$$

2.3. Assumptions. For a given normed vector space X, we denote by $\bar B_X(R)$ the closed ball of radius R and center 0. When the context is clear, we simply write $\bar B(R)$. Given $R > 0$, we denote $V(R) = \operatorname{conv}\big\{ (x,v) : \, |x| \leq R, \, c(x,v) \leq 0 \big\}$.

We consider the following assumptions:

(H1) *Convexity assumptions*

(i) There exists $C > 0$ such that for all $x \in \mathbb{R}^n$, the map $L(x, \cdot)$ is strongly convex with parameter $1/C$ and such that for all $(x, v) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$L(x, v) \geq (1/C)|v|^2 - C.$$

(ii) For all $x \in \mathbb{R}^n$ and for all $i = 1, ..., n_c$, the map $c_i(x, \cdot)$ is convex.

(iii) The map ψ is the gradient of some C^1 and convex function ϕ.

(H2) *Regularity assumptions*

(i) The maps L and c are twice continuously differentiable.

(ii) The maps a, b, g_0, g_1, and g_2 are continuously differentiable.

(iii) For all $m \in \mathcal{P}_1(\mathbb{R}^n)$, the map $f(\cdot, m)$ is continuously differentiable. The maps f and $D_x f$ are continuous.

(H3) *Boundedness and growth assumptions*

(i) For all $R > 0$, there exists $C(R) > 0$ such that for all $(x, v) \in V(R)$,

$$\begin{aligned} |D_x L(x, v)| \leq\ & C(R)(1 + |v|^2), \\ |D_v L(x, v)| \leq\ & C(R)(1 + |v|). \end{aligned}$$

(ii) For all $R > 0$, there exists $C(R) > 0$ such that for all (x, v) and $(\tilde{x}, \tilde{v})$ in $V(R) + \bar{B}(1)$,

$$\begin{aligned} |D_x c(x, v)| + |D_v c(v, x)| \leq\ & C(R), \\ |D_v c(x, v) - D_v c(\tilde{x}, \tilde{v})| \leq\ & C(R)\big|(x, v) - (\tilde{x}, \tilde{v})\big|. \end{aligned}$$

(iii) There exists $C > 0$ such that for all $x \in \mathbb{R}^n$,

$$|a(x)| \leq C(1 + |x|) \quad \text{and} \quad |b(x)| \leq C(1 + |x|).$$

(iv) The support K_0 of m_0 is bounded.

(v) There exists $C > 0$ such that for all $x_0 \in K_0$, for all $m \in C(0, T; \mathcal{P}_1(\mathbb{R}^n))$ and for all $(\gamma, v) \in \mathcal{K}[x_0]$,

$$\begin{aligned} \int_0^T f(\gamma(t), m(t))\mathrm{d}t + g_0(m(T), \gamma(T)) \geq\ & -C \\ \|D_x f(\gamma(t), m(t))\|_{L^\infty(0,T;\mathbb{R}^n)} \leq\ & C. \end{aligned}$$

(vi) The map ψ is bounded.

(H4) *Feasibility assumptions*

(i) For all $R > 0$, there exists a constant $C(R) > 0$ such that for all $x \in \bar{B}(R)$, there exists $v \in \bar{B}(C(R))$ such that $c(x, v) \leq 0$.

(ii) There exists $C > 0$ such that for all $x_0 \in K_0$, for all $m \in C(0, T; \mathcal{P}(\mathbb{R}^n))$ such that $m(0) = m_0$, and for all $P \in L^\infty(0, T; \mathbb{R}^m)$ such that

$$\|P\|_{L^\infty(0,T;\mathbb{R}^m)} \leq \sup_{\theta \in \mathbb{R}^m} |\psi(\theta)|, \tag{2.3}$$

there exists $(\gamma_0, v_0) \in \mathcal{K}[x]$ such that $J[m, P](\gamma_0, v_0) \leq C$.

(H5) *Qualification assumptions*

(i) There exists $C > 0$ such that for all $x_0 \in K_0$, for all $(\gamma, v) \in \mathcal{K}[x_0]$, for all $z_1 \in \mathbb{R}^{n_{g_1}}$, there exists a pair $(y, w) \in H^1(0, T; \mathbb{R}^n) \times L^\infty(0, T; \mathbb{R}^m)$

solution of the linearized state equation

$$\text{(2.4)} \quad \begin{cases} \dot{y}(t) = \big(Da(\gamma(t)) + \sum_{i=1}^{m} Db_i(\gamma(t))v_i(t)\big)y(t) + b(\gamma(t))w(t), \\ y(0) = 0, \end{cases}$$

such that $Dg_1(\gamma(T))y(T) = z_1$, and

$$\|y\|_{H^1(0,T;\mathbb{R}^n)} \leq C|z_1| \quad \text{and} \quad \|w\|_{L^\infty(0,T;\mathbb{R}^n)} \leq C|z_1|.$$

(ii) There exists $C > 0$ such that for all $x_0 \in K_0$, for all $(\gamma, v) \in \mathcal{K}[x_0]$, there exists $(y, w) \in H^1(0,T;\mathbb{R}^n) \times L^\infty(0,T;\mathbb{R}^m)$ satisfying (2.4) such that

$$\begin{cases} Dg_1(\gamma(T))y(T) = & 0, \\ g_2(\gamma(T)) + Dg_2(\gamma(T))y(T) \leq & -1/C, \\ c(\gamma(t), v(t)) + Dc(\gamma(t), v(t))(y(t), w(t)) \leq & -1/C. \end{cases}$$

and such that $\|y\|_{H^1(0,T;\mathbb{R}^n)} \leq C$ and $\|w\|_{L^\infty(0,T;\mathbb{R}^n)} \leq C$.

(iii) There exists $C > 0$ such that for all $(x, v) \in \mathbb{R}^n \times \mathbb{R}^m$ satisfying $c(x, v) \leq 0$, for all $\omega \in \mathbb{R}^{|I(x,v)|}$,

$$|D_v c_{I(x,v)}(x, v)^\top \omega| \geq (1/C)|\omega|,$$

where $I(x, v) = \{i = 1, ..., n_c : \ c_i(x, u) = 0\}$.

(iv) For all $R > 0$, there exist $C(R) > 0$ and $\varepsilon(R) > 0$ such that for all $x \in \bar{B}(R)$ and for all $v \in \mathbb{R}^m$ such that $c(x, v) \leq 0$, there exists $w \in \bar{B}(C(R))$ such that

$$c(x, v) + D_v c(x, v)w \leq -\varepsilon(R).$$

Remark 2.3. For the sake of simplicity in the presentation of this article, we consider time-independent data, but most of the results remain valid if the above assumptions hold uniformly with respect to time.

3. The optimal control problem. In this section, we study the optimal control problem (2.1) that an individual player aims to solve. Throughout this section, we fix a triplet $(m, P, x_0) \in C(0,T;\mathcal{P}_1(\mathbb{R}^n)) \times L^\infty(0,T;\mathbb{R}^m) \times K_0$ such that (2.3) holds.

3.1. Some technical results. The next lemma is a metric regularity property, obtained from the Mangasarian-Fromovitz qualification condition (H5)-(iv), which implies Robinson's qualification condition (see [7, Section 2.3.4]). Thus the lemma is a particular case of the Robinson-Ursescu stability theorem [7, Theorem 2.87].

LEMMA 3.1. *Let $R > 0$. There exist $\delta > 0$ and $C > 0$ such that for all $(x, \tilde{x}, \tilde{v}) \in \bar{B}(R)^2 \times \mathbb{R}^m$ such that $c(\tilde{x}, \tilde{v}) \leq 0$ and $|x - \tilde{x}| \leq \delta$, there exists $v \in \mathbb{R}^m$ such that*

$$c(x, v) \leq 0 \quad \text{and} \quad |v - \tilde{v}| \leq C|x - \tilde{x}|.$$

Moreover, for fixed $\tilde{x}$ and $\tilde{v}$, v can be constructed as a continuous function of x.

Proof. Let $R > 0$. The constant ε used below, as well as all constants $C > 0$, depend only on R. Let $(x, \tilde{x}, \tilde{v}) \in \mathbb{R}^{n+n+m}$ be such that $c(\tilde{x}, \tilde{v}) \leq 0$ and $|x - \tilde{x}| \leq 1$. By Assumption (H5)-(iv), there exist $w \in \mathbb{R}^m$, $C > 0$, and $\varepsilon > 0$ such that

$$c(\tilde{x}, \tilde{v}) + D_v c(\tilde{x}, \tilde{v})w \leq -\varepsilon \quad \text{and} \quad |w| \leq C.$$

Let $\theta \in [0,1]$. Let $v_\theta = \tilde{v} + \theta w$. We have

$$c(x, v_\theta) = c(\tilde{x}, \tilde{v}) + \theta D_v c(\tilde{x}, \tilde{v}) w + a_\theta + b_\theta, \tag{3.1}$$

where

$$\begin{aligned} a_\theta = & \int_0^1 D_x c(\tilde{x} + s(x-\tilde{x}), \tilde{v} + s\theta w)(x - \tilde{x}) \mathrm{d}s, \\ b_\theta = & \ \theta \int_0^1 \big[D_v c(\tilde{x} + s(x-\tilde{x}), \tilde{v} + s\theta w) - D_v(\tilde{x}, \tilde{v}) \big] w \mathrm{d}s. \end{aligned}$$

By Assumption (H3)-(ii), we have

$$|a_\theta| \leq C|x - \tilde{x}| \quad \text{and} \quad |b_\theta| \leq C\theta\big(|x-\tilde{x}| + \theta\big).$$

It follows from (3.1) that

$$\begin{aligned} c(x, v_\theta) &= (1-\theta) c(\tilde{x}, \tilde{v}) + \theta \big[c(\tilde{x}, \tilde{v}) + D_v c(\tilde{x}, \tilde{v}) w \big] + a_\theta + b_\theta \\ &\leq \ -\theta\varepsilon + C|x - \tilde{x}| + C\theta^2. \end{aligned} \tag{3.2}$$

Let us define $\bar{\theta} = \min\left(\frac{\varepsilon}{2C}, 1\right)$ and $\delta = \min\left(\frac{\varepsilon\bar{\theta}}{2C}, 1\right)$, where C is the constant appearing in the right-hand side of (3.2). We assume now that $|x - \tilde{x}| \leq \delta$ and we fix $v = v_\theta$, where $\theta = \frac{2C|x-\tilde{x}|}{\varepsilon}$. It remains to verify that $c(x,v) \leq 0$. Note first that $\theta \leq \frac{2C\delta}{\varepsilon} \leq \bar{\theta}$, by definition of δ. It follows from (3.2) that

$$c(x,v) \leq -\theta\varepsilon + \underbrace{(C\bar{\theta})}_{\leq \varepsilon/2} \theta + C|x - \tilde{x}| \leq -\frac{\theta\varepsilon}{2} + C|x-\tilde{x}| = 0,$$

which concludes the proof. ∎

Lemma 3.2. *(i) For all $x \in \mathbb{R}^n$ and for all $r \in \mathbb{R}^m$, there exists a unique pair $(v, \nu) \in \mathbb{R}^m \times \mathbb{R}^{n_c}$ such that the following holds:*

$$D_v L(x,v)^\top + r + D_v c(x,v)^\top \nu = 0, \quad \nu \geq 0, \quad \textit{and} \quad \langle \nu, c(x,v) \rangle = 0. \tag{3.3}$$

We denote it $(v[x,r], \nu[x,r])$.
(ii) Let $R > 0$. The map $(x,r) \in \bar{B}(R) \mapsto (v[x,r], \nu[x,r])$ is Lipschitz continuous.

Proof. (i) Let $(x,r) \in \mathbb{R}^n \times \mathbb{R}^m$. Consider the optimization problem:

$$\inf_{v \in \mathbb{R}^m} L(x,v) + \langle r, v \rangle, \quad \text{subject to: } c(x,v) \leq 0. \tag{3.4}$$

As a consequence of Assumption (H1)-(i), the above cost function is coercive. By Assumption (H4)-(i), there exists v_0 such that $c(x, v_0) \leq 0$. Therefore, (3.4) possesses a solution v. As a consequence of the qualification assumption (H5)-(iii), the optimality conditions exactly take the form of (3.3). This proves the existence part of the first part of the theorem. Now take a pair (v, ν) satisfying (3.3). Then, by the strong convexity of $L(x, \cdot)$ and by the convexity of the maps $c_i(x, \cdot)$, v is the unique solution to (3.4) and ν is the associated Lagrange multiplier, it is also unique as a consequence of (H5)-(iii).

(ii) Let us prove the Lipschitz continuity of $v[\cdot,\cdot]$, $\nu[\cdot,\cdot]$. We mainly rely on results of [7]. We first reformulate (3.3) as a generalized equation: given (x,r), the pair (v,ν) satisfies (3.3) if and only if:

$$0 \in \Phi(v,\nu;x,r) + N(\nu), \tag{3.5}$$

where $\Phi(v,\nu;x,r) = \Big(D_v L(x,v)^\top + r + D_v c(x,v)^\top \nu, -c(x,v) \Big) \in \mathbb{R}^{m+n_c}$ and where $N(\nu) = \big\{ (0,z) \in \mathbb{R}^{m+n_c} : z \leq 0, \langle z, \nu \rangle = 0 \big\}$, if $\nu \geq 0$, and $N(\nu) = \emptyset$ otherwise. By [7, Proposition 5.38], $(\bar{v},\bar{\nu})$ is a strongly regular solution of (3.5) (in the sense of [7, Definition 5.12]). Note that the required sufficient second-order optimality conditions follow from the strong convexity of $L(x,\cdot)$ and the convexity of $c_i(x,\cdot)$. It follows then from [7, Theorem 5.13] that $v[\cdot,\cdot]$ and $\nu[\cdot,\cdot]$ are locally Lipschitz continuous, and therefore Lipschitz continuous on any compact set, as was to be proved. ∎

Remark 3.3. The twice differentiability of L and c, required in Assumption (H2)-(i) is only used for the application of [7, Proposition 5.38] in the proof of Lemma 3.2. It is sufficient to assume that L and c are continuously differentiable if c does not depend on x (i.e. if we just have control constraints instead of mixed state-control constraints). In that case, the Lipschitz continuity is deduced from [7, Proposition 4.32].

3.2. Estimates for the optimal solutions. The goal of this section is to derive some a priori bounds for solutions $(\bar{\gamma},\bar{v})$ to the optimal control problem (2.1) and for the associated costate and Lagrange multipliers. They will be crucial for the construction of an appropriate set of probability measures on state-costate trajectories. We follow a rather standard methodology. The coercivity of L, together with other feasibility and bound conditions allows to show the existence of a solution and to derive a bound of $\bar{v}$ in $L^2(0,T;\mathbb{R}^m)$. Then we provide first-order necessary optimality conditions and a bound on the associated costate p, with the help of the qualification conditions. We finally obtain bounds of $\bar{\gamma}$ and p in $W^{1,\infty}(0,T;\mathbb{R}^n)$ and $\bar{v}$ in $L^\infty(0,T;\mathbb{R}^m)$.

We recall that throughout this section the triplet (m,P,x_0) is fixed and satisfies (2.3). Note that all constants C used in this section are independent of (m,P,x_0).

PROPOSITION 3.4. *The optimal control problem* (2.1) *has (at least) one solution. There exist two constants $M_1 > 0$ and $C > 0$, independent of m, P, and x_0, such that for all solutions $(\bar{\gamma},\bar{v})$ to* (2.1)*,*

$$\|\bar{\gamma}\|_{L^\infty(0,T;\mathbb{R}^n)} \leq M_1 \quad \textit{and} \quad \|\bar{v}\|_{L^2(0,T;\mathbb{R}^m)} \leq C. \tag{3.6}$$

Proof. The constants $C > 0$ used in this proof only depend on the data of the problem. Let $(\gamma_0, v_0) \in \mathcal{K}[x_0]$ satisfy Assumption (H4)-(ii). Let $(\gamma_k, v_k)_{k\in\mathbb{N}}$ be a minimizing sequence. Without loss of generality, we can assume that

$$J[m,P](\gamma_k,v_k) \leq J[m,P](\gamma_0,v_0), \quad \forall k \in \mathbb{N}.$$

Using Assumption (H1)-(i), the boundedness of P, and Assumption (H3)-(v), we deduce that

$$\begin{aligned} C &\geq J[m,P](\gamma_0,v_0) \geq J[m,P](\gamma_k,v_k) \\ &\geq \frac{1}{C}\|v_k\|^2_{L^2(0,T)} - C\|v_k\|_{L^1(0,T)} - C \geq \frac{1}{C}\|v_k\|^2_{L^2(0,T)} - C, \end{aligned}$$

for some independent constants C. It follows that v_k in bounded in $L^2(0,T;\mathbb{R}^m)$. By Grönwall's lemma and Assumption (H3)-(iii), there exists a constant $C>0$ such that $\|\gamma_k\|_{L^\infty(0,T;\mathbb{R}^n)} \leq C$. The state equation further implies that $\|\gamma_k\|_{H^1(0,T;\mathbb{R}^n)} \leq C$. Extracting a subsequence if necessary, there exist $(\bar{\gamma},\bar{v}) \in H^1(0,T;\mathbb{R}^n) \times L^2(0,T;\mathbb{R}^m)$ and a $C>0$ such that

$$\|\bar{\gamma}\|_{H^1(0,T;\mathbb{R}^n)} \leq C \quad \text{and} \quad \|\bar{v}\|_{L^2(0,T;\mathbb{R}^m)} \leq C$$

and such that $(\gamma_k, v_k) \rightharpoonup (\bar{\gamma},\bar{v})$ for the weak topology of $H^1(0,T;\mathbb{R}^n) \times L^2(0,T;\mathbb{R}^m)$. Since $H^1(0,T;\mathbb{R}^n)$ is compactly embedded in $L^\infty(0,T;\mathbb{R}^n)$, we deduce that γ_k converges uniformly to $\bar{\gamma}$.

Let us prove that $c(\bar{\gamma}(t),\bar{v}(t)) \leq 0$ for a.e. $t \in (0,T)$. Let $\varphi \in L^\infty(0,T;\mathbb{R}^{n_c})$ be such that $\varphi(t) \geq 0$ for a.e. $t \in (0,T)$. We have

$$\int_0^T \langle \varphi(t), c(\bar{\gamma}(t),\bar{v}(t))\rangle \mathrm{d}t = a_k + b_k + \int_0^T \langle \varphi(t), c(\gamma_k(t), v_k(t))\rangle \mathrm{d}t \leq a_k + b_k,$$

where, skipping the time arguments

$$a_k = \int_0^T \big\langle \varphi, c(\bar{\gamma},\bar{v}) - c(\bar{\gamma}, v_k)\big\rangle \mathrm{d}t \quad \text{and} \quad b_k = \int_0^T \big\langle \varphi, c(\bar{\gamma}, v_k) - c(\gamma_k, v_k)\big\rangle \mathrm{d}t.$$

Note that all these integrals are well-defined as a consequence of Assumption (H3)-(ii). Also by Assumption (H3)-(ii), we easily verify that $D_v c(\bar{\gamma}(\cdot),\bar{v}(\cdot)) \in L^2(0,T;\mathbb{R}^{n_c \times m})$. Therefore, by the convexity of the maps $c_i(x,\cdot)$ in Assumption (H1)-(ii),

$$a_k \leq \int_0^T \langle \varphi(t), D_v c(\bar{\gamma}(t),\bar{v}(t))(\bar{v}(t) - v_k(t))\rangle \mathrm{d}t \underset{k\to\infty}{\longrightarrow} 0.$$

By Assumption (H3)-(ii), we also have

$$|b_k| \leq C\|\varphi\|_{L^\infty(0,T;\mathbb{R}^{n_c})} \|\gamma_k - \bar{\gamma}\|_{L^\infty(0,T;\mathbb{R}^n)} \underset{k\to\infty}{\longrightarrow} 0.$$

It follows that for all $\varphi \geq 0$, $\int_0^T \langle \varphi, c(\bar{\gamma}(t),\bar{v}(t))\rangle \mathrm{d}t \leq 0$. Therefore, $c(\bar{\gamma}(t),\bar{v}(t)) \leq 0$, for a.e. $t \in (0,T)$. With similar arguments, we prove that $(\bar{\gamma},\bar{v})$ is feasible and that

$$J[m,P](\bar{\gamma},\bar{v}) \leq \lim_{k\to\infty} J[m,P](\gamma_k, v_k),$$

which concludes the proof of optimality of $(\bar{\gamma},\bar{v})$. Repeating the above arguments, we show that any solution to (2.1) satisfies the bound (3.6). ∎

We next state optimality conditions for the optimal control problem. The proof of the following proposition is deferred to the appendix in Section A. In the rest of the section, we write $c[t]$ instead of $c(\bar{\gamma}(t),\bar{v}(t))$ (for a specified pair $(\bar{\gamma},\bar{v})$). We use the same convention for a, b, g_0, g_1, and g_2.

Proposition 3.5. *Let $(\bar{\gamma},\bar{v})$ be a solution to (2.1). There exists a quintuplet*

$$(p,\lambda_0,\lambda_1,\lambda_2,\nu) \in W^{1,2}(0,T;\mathbb{R}^n) \times \mathbb{R} \times \mathbb{R}^{n_{g_1}} \times \mathbb{R}^{n_{g_2}} \times L^\infty(0,T;\mathbb{R}^{n_c})$$

such that $(p,\lambda_0) \neq (0,0)$ and such that the adjoint equation

(3.7)
$$\begin{cases} p(T)^\top = \lambda_0 Dg_0[T] + \lambda_1^\top Dg_1[T] + \lambda_2^\top Dg_2[T] \\ -\dot{p}(t)^\top = \lambda_0 D_x L[t] + \lambda_0 D_x f[t] + p(t)^\top \big(Da[t] + \sum_{i=1}^m Db_i[t] v_i(t)\big) + \nu(t)^\top D_x c[t], \end{cases}$$

the stationary condition

$$\lambda_0 D_v L[t] + \lambda_0 P(t)^\top + p(t)^\top b[t] + \nu(t)^\top D_v c[t] = 0, \tag{3.8}$$

and the following sign and complementarity conditions

$$\begin{cases} \lambda_0 \geq 0, & \\ \lambda_2 \geq 0, & \langle \lambda_2, g_2(\bar{\gamma}(T)) \rangle = 0, \\ \nu(t) \geq 0, & \langle c(\bar{\gamma}(t), \bar{v}(t)), \nu(t) \rangle = 0, \quad \text{for a.e. } t \in (0,T) \end{cases} \tag{3.9}$$

are satisfied. Moreover, if $\lambda_0 \neq 0$, then $p \in W^{1,\infty}(0,T;\mathbb{R}^n)$.

The goal of the last two results in this subsection is to obtain uniform bounds for the optimal solutions and their associated multipliers.

PROPOSITION 3.6. *Let $(\bar{\gamma}, \bar{v})$ be a solution to (2.1). There exists a quintuplet $(p, \lambda_0, \lambda_1, \lambda_2, \nu)$ satisfying the optimality conditions of the above proposition and such that $\lambda_0 = 1$. Moreover, for such a quintuplet, we have*

$$\|p\|_{L^\infty(0,T;\mathbb{R}^n)} \leq M_2$$

for some constant M_2 independent of $(\bar{\gamma}, \bar{v})$ and $(p, \lambda_0, \lambda_1, \lambda_2, \nu)$.

Proof. The proof essentially relies on the qualification conditions (H5)-(i) and (H5)-(ii). All constants C used in the proof are independent of $(p, \lambda_0, \lambda_1, \lambda_2, \nu)$ and $(\bar{\gamma}, \bar{v})$. Let (y, w) satisfy the linearized equation (2.4) (for $(\gamma, v) = (\bar{\gamma}, \bar{v})$). By integration by parts we have

$$\begin{aligned} \langle p(T), y(T) \rangle = & \int_0^T \langle \dot{p}(t), y(t) \rangle + \langle p(t), \dot{y}(t) \rangle \mathrm{d}t \\ = & -\lambda_0 \int_0^T (D_x L[t] + D_x f[t]) y(t) \mathrm{d}t \\ & - \int_0^T p(t)^\top \big(Da[t] + \textstyle\sum_{i=1}^m Db_i[t] \bar{v}_i(t) \big) y(t) \mathrm{d}t - \int_0^T \nu(t)^\top D_x c[t] y(t) \mathrm{d}t \\ & + \int_0^T p(t)^\top \big(Da[t] + \textstyle\sum_{i=1}^m Db_i[t] \bar{v}_i(t) \big) y(t) \mathrm{d}t + \int_0^T p(t)^\top b[t] w(t) \mathrm{d}t. \end{aligned}$$

The second and the fourth integral cancel out. Injecting the optimality condition (3.8) in the last integral, we obtain:

$$\begin{aligned} \langle p(T), y(T) \rangle = & -\lambda_0 \int_0^T \Big(DL[t](y(t), w(t)) + \langle P(t), w(t) \rangle + D_x f[t] y(t) \Big) \mathrm{d}t \\ & - \int_0^T \nu(t)^\top Dc[t](y(t), w(t)) \mathrm{d}t. \end{aligned} \tag{3.10}$$

The main feature of this formula is that the right-hand side is independent of p. Let (y, w) satisfy Assumption (H5)-(ii). By (H3)-(i), (H3)-(vi) and (H3)-(v) we have

$$\begin{aligned} & \textstyle\int_0^T \left(DL[t](y(t), w(t)) + \langle P(t), w(t) \rangle + D_x f[t] y(t) \right) \mathrm{d}t \\ & \qquad \leq C \|y\|_\infty \textstyle\int_0^T (1 + |\bar{v}(t)|^2) \mathrm{d}t + C \|w\|_\infty \textstyle\int_0^T (1 + |\bar{v}(t)|) \mathrm{d}t \leq C, \end{aligned}$$

where the last inequality holds by Proposition 3.4 and (H5)-(ii). By (H5)-(ii) and the complementarity conditions (3.9) we obtain

$$\begin{aligned}\int_0^T \nu(t)^\top Dc[t](y(t),w(t))\mathrm{d}t &= \int_0^T \nu(t)^\top \left(c[t] + Dc[t](y(t),w(t))\right)\mathrm{d}t \\ &\leq \; -\frac{1}{C}\|\nu\|_{L^1(0,T;\mathbb{R}^{n_c})}.\end{aligned}$$

Therefore,

$$\langle p(T), y(T)\rangle \geq -C\lambda_0 + \frac{1}{C}\|\nu\|_{L^1(0,T;\mathbb{R}^{n_c})}. \tag{3.11}$$

Moreover, we deduce from the terminal condition for p that

$$\begin{aligned}\langle p(T), y(T)\rangle &= \lambda_0 Dg_0[T]y(T) + \langle \lambda_1, Dg_1[T]y(T)\rangle + \langle \lambda_2, Dg_2[T]y(T)\rangle \\ &= \lambda_0 Dg_0[T]y(T) + \langle \lambda_2, g_2[T] + Dg_2[T]y(T)\rangle \\ &\leq C\lambda_0 - \frac{1}{C}|\lambda_2|. \end{aligned}\tag{3.12}$$

The last inequality holds by (H2)-(ii), Proposition 3.4 and (H5)-(ii). It follows from (3.11) and (3.12) that

$$|\lambda_2| \leq C\lambda_0 \quad \text{and} \quad \|\nu\|_{L^1(0,T;\mathbb{R}^{n_c})} \leq C\lambda_0. \tag{3.13}$$

Now, let us consider (y,w) satisfying (H5)-(i) with $z_1 = \lambda_1/|\lambda_1|$. We have

$$\|y\|_{L^\infty(0,T;\mathbb{R}^n)} \leq C \quad \text{and} \quad \|w\|_{L^\infty(0,T;\mathbb{R}^n)} \leq C.$$

Since $Dc[t]$ is bounded in $L^\infty(0,T;\mathbb{R}^{n_c\times(n+m)})$ (by Assumption (H3)-(ii)), we have

$$\|Dc[\cdot](y(\cdot),w(\cdot))\|_{L^\infty(0,T;\mathbb{R}^{n_c})} \leq C. \tag{3.14}$$

Formula (3.10), together with the bound on $\|\nu\|_{L^1(0,T;\mathbb{R}^{n_c})}$ and (3.14) yields

$$\langle p(T), y(T)\rangle \leq C\lambda_0. \tag{3.15}$$

It follows from the terminal condition and the estimate on $|\lambda_2|$ that

$$\begin{aligned}\langle p(T), y(T)\rangle &= \lambda_0 Dg_0[T]y(T) + \langle \lambda_1, Dg_1[T]y(T)\rangle + \langle \lambda_2, Dg_2[T]y(T)\rangle \\ &\geq \; -C\lambda_0 + |\lambda_1|. \end{aligned}\tag{3.16}$$

Combining (3.15) and (3.16), we deduce that

$$|\lambda_1| \leq C\lambda_0. \tag{3.17}$$

If $\lambda_0 = 0$, then $\lambda_1 = 0$, $\lambda_2 = 0$, and $\nu = 0$. Thus $p(T) = 0$ and $\dot{p}(t) = 0$ a.e. and therefore $p = 0$, in contradiction with $(p,\lambda_0) \neq (0,0)$. We deduce that $\lambda_0 > 0$. The optimality conditions being invariant by multiplication of a positive constant, we deduce the existence of a quintuplet satisfying (3.7)-(3.8)-(3.9) and $\lambda_0 = 1$. Bounds of $|\lambda_1|$, $|\lambda_2|$, and $\|\nu\|_{L^1(0,T;\mathbb{R}^{n_c})}$ directly follow from (3.13) and (3.17). Then we obtain a bound of $|p(T)|$ and finally a bound of $\|p\|_{L^\infty(0,T;\mathbb{R}^n)}$ with Grönwall's lemma. ◻

DEFINITION 3.7. *Given a solution $(\bar{\gamma}, \bar{v})$ to (2.1), we call associated costate any p for which there exists $(\lambda_0, \lambda_1, \lambda_2, \nu)$ such that the optimality conditions in Proposition 3.5 hold true and $\lambda_0 = 1$.*

In order to obtain more regularity on $(\bar{\gamma}, \bar{v})$, we need to express the optimal control as an auxiliary function of the state and costate, which is deduced from Lemma 3.2.

LEMMA 3.8. *Let $(\bar{\gamma}, \bar{v})$ and $(p, \lambda_0, \lambda_1, \lambda_2, \nu)$ be as in Proposition 3.6. There exists $C > 0$ independent of $(\bar{\gamma}, \bar{v})$ and $(p, \lambda_0, \lambda_1, \lambda_2, \nu)$ such that*

$$\|\bar{v}\|_{L^\infty(0,T;\mathbb{R}^n)} \leq C, \quad \|\nu\|_{L^\infty(0,T;\mathbb{R}^n)} \leq C.$$

In addition, there exist constants $M_3 > 0$ and $M_4 > 0$, such that

$$\|\dot{\bar{\gamma}}\|_{L^\infty(0,T;\mathbb{R}^n)} \leq M_3, \quad \|\dot{p}\|_{L^\infty(0,T;\mathbb{R}^n)} \leq M_4. \tag{3.18}$$

Proof. It follows from the optimality condition (3.8) and Lemma 3.2 that

$$\begin{aligned} \bar{v}(t) = &\quad v[\bar{\gamma}(t), P(t) + b(\bar{\gamma}(t))^\top p(t)], \\ \nu(t) = &\quad \nu[\bar{\gamma}(t), P(t) + b(\bar{\gamma}(t))^\top p(t)]. \end{aligned} \tag{3.19}$$

Lemma 3.2 further implies that $\|\bar{v}\|_{L^\infty(0,T;\mathbb{R}^m)} \leq C$ and $\|\nu\|_{L^\infty(0,T;\mathbb{R}^{n_c})} \leq C$. The estimates (3.18) follow. ∎

4. Existence of MFGC equilibria. In this section, we prove the main result of the paper. We first construct the auxiliary functions announced in the introduction. Then, applying Kakutani's fixed point theorem, we prove the existence of an auxiliary MFGC equilibrium and finally prove the existence of a Lagrangian equilibrium.

4.1. Auxiliary functions. We set $B = \bar{B}(0, M_1) \times \bar{B}(0, M_2) \subset \mathbb{R}^n \times \mathbb{R}^n$, where M_1 and M_2 are given by Lemma 3.4 and Proposition 3.6, respectively.

LEMMA 4.1. *Let $\mu \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ be such that $\operatorname{supp}(\mu) \subseteq B$. There exists a unique pair $(P, v) \in \mathbb{R}^m \times C(B; \mathbb{R}^m)$ such that*

$$P = \psi\Big(\int_B v(x,q) \mathrm{d}\mu(x,q)\Big) \quad \text{and} \quad v(x,q) = v[x, P + b(x)^\top q] \quad \forall (x,q) \in B, \tag{4.1}$$

where $v[\cdot,\cdot]$ is the auxiliary mapping introduced in Lemma 3.2. Moreover, there exists a constant $C > 0$ independent of μ such that $\|v\|_{L^\infty(B;\mathbb{R}^m)} \leq C$ and such that v is C-Lipschitz continuous.

In the sequel, the unique solution to (4.1) is denoted by $(\mathbf{P}[\mu], \mathbf{v}[\mu])$. It satisfies the identity:

$$\mathbf{P}[\mu] = \psi\Big(\int_B v[x, \mathbf{P}[\mu] + b(x)^\top q] \mathrm{d}\mu(x,q)\Big). \tag{4.2}$$

Proof of Lemma 4.1. Step 1: existence. Let $L^2_\mu(B)$ be the Hilbert space of square summable μ-measurable mappings from B to $\mathbb{R}^m$, equipped with the scalar product

$$\langle v_1, v_2 \rangle_\mu = \int_B \langle v_1(x,q), v_2(x,q) \rangle \mathrm{d}\mu(x,q).$$

Let us denote $\|\cdot\|_{L^2_\mu(B)}$ the associated norm. Consider $J_\mu \colon L^2_\mu(B) \to \mathbb{R}$ defined by

$$J_\mu(v) = \int_B \Big(L(x, v(x,q)) + \langle b(x)^\top q, v(x,q) \rangle \Big) \mathrm{d}\mu(x,q) + \phi\Big(\int_B v(x,q) \mathrm{d}\mu(x,q)\Big)$$

and the following optimization problem:

$$\inf_{v \in L^2_\mu(B)} J_\mu(v), \quad \text{subject to: } c(x, v(x,q)) \leq 0, \quad \text{for } \mu\text{-a.e. } (x,q) \in B. \tag{4.3}$$

The existence of a solution $\bar{v}$ can be established with a rather straightforward adaptation of the arguments used in the proof of Proposition 3.4, thus we do not detail it. In a nutshell, any minimizing sequence $(v_k)_{k\in\mathbb{N}}$ is bounded in $L^2_\mu(B)$, thus possesses an accumulation point for the weak topology of $L^2_\mu(B)$; moreover, the feasible set of (4.3) is weakly closed and J_μ is weakly lower semi-continuous, whence the existence of a solution.

Let us derive optimality conditions. Let us mention that at this stage of analysis, we do not know whether $\bar{v}$ is bounded. It would be therefore difficult to write (4.3) as an abstract optimization problem satisfying an appropriate qualification condition. Thus we use the standard needle perturbation technique for our purpose. Let us introduce some notation. Let $B_\varepsilon(x,q)$ denote the open ball of center (x,q) and radius ε. Let $|B_\varepsilon(x,q)| = \int_{B_\varepsilon(x,q)} \mathrm{d}\mu(x,q)$ denote its volume. Let us fix a representative of $\bar{v}$ (we still denote it $\bar{v}$). We call (abusively) Lebesgue point any pair $(x_0,q_0) \in \operatorname{supp}(\mu)$ such that

$$|B_\varepsilon(x_0,q_0)|^{-1} \int_{B_\varepsilon} |\bar{v}(x,q) - \bar{v}(x_0,q_0)| \mathrm{d}\mu(x,q) \underset{\varepsilon\downarrow 0}{\longrightarrow} 0.$$

By the Lebesgue differentiation theorem, we have that for μ-a.e. $(x_0,q_0) \in B$, (x_0,q_0) is a Lebesgue point. Let (x_0,q_0) be a Lebesgue point, let $v_0 = \bar{v}(x_0,q_0)$. Let $v \in \mathbb{R}^m$ be such that $c(x_0,v) \leq 0$. Let $\delta > 0$ be the constant obtained in Lemma 3.1. The latter implies that there exists a continuous map $\hat{v}\colon B_\delta(x_0) \to \mathbb{R}^m$ such that for some constant C,

$$c(x, \hat{v}(x)) \leq 0, \quad |\hat{v}(x) - v| \leq C|x - x_0|, \quad \forall x \in B_\delta(x_0).$$

For $\varepsilon \in (0,\delta)$, we consider

$$v_\varepsilon(x,q) = \begin{cases} \hat{v}(x) & \text{if } |x - x_0| \leq \varepsilon \\ \bar{v}(x,q) & \text{otherwise.} \end{cases}$$

It can be easily verified that

$$0 \leq \frac{J_\mu(v_\varepsilon) - J_\mu(\bar{v})}{|B_\varepsilon(x_0,q_0)|} \underset{\varepsilon\to 0}{\longrightarrow} L(x_0,v) - L(x_0,v_0) + \langle \tilde{P} + b(x_0)^\top q_0, v - v_0 \rangle,$$

where

$$\tilde{P} = \psi\Big(\int_B \bar{v}(x,q) \mathrm{d}\mu(x,q) \Big).$$

It follows that v_0 is a solution to problem (3.3), taking $(x,r) = (x_0, \tilde{P} + b(x_0)^\top q_0)$. Therefore, we have $\bar{v}(x,q) = v[x, \tilde{P} + b(x)^\top q]$, for μ-a.e. $(x,q) \in B$. Let us define

$$\tilde{v}(x,q) = v[x, \tilde{P} + b(x)^\top q] \quad \forall (x,q) \in B. \tag{4.4}$$

Since $\tilde{v}(x,q) = \bar{v}(x,q)$ for μ-a.e. (x,q), we have

$$\psi\Big(\int_B \tilde{v}(x,q) \mathrm{d}\mu(x,q) \Big) = \psi\Big(\int_B \bar{v}(x,q) \mathrm{d}\mu(x,q) \Big) = \tilde{P}$$

and therefore, $(\tilde{P},\tilde{v})$ satisfies (4.1).

Step 2: uniqueness. Let (P,v) be such that (4.1) holds. The problem (4.3) being a convex optimization problem, it follows from (4.1) that v is a solution to (4.3). Since L is strictly convex, J_μ is strictly convex, therefore, problem (4.3) has a unique solution in $L^2_\mu(B)$. Thus, $v(x,q)=\tilde{v}(x,q)$ for μ-a.e. $(x,q)\in B$. It follows that

$$P=\psi\Big(\int_B v(x,q)\mathrm{d}\mu(x,q)\Big)=\psi\Big(\int_B \tilde{v}(x,q)\mathrm{d}\mu(x,q)\Big)=\tilde{P}$$

and finally that $v(x,q)=\tilde{v}(x,q)$, for all $(x,q)\in B$.

Step 3: boundedness and Lipschitz continuity. Since ψ is bounded, we directly get a bound for $\tilde{P}$. The Lipschitz continuity follows then from (4.4), Assumption (H2)-(ii), and Lemma 3.2. ∎

LEMMA 4.2. *The map $\mathbf{P}\colon\mathcal{P}(B)\to\mathbb{R}^m$ is uniformly continuous for the d_1 distance.*

Proof. Let μ_1 and μ_2 in $\mathcal{P}(B)$. Let $(P_i,v_i)=(\mathbf{P}[\mu_i],\mathbf{v}[\mu_i])$, for $i=1,2$. Let $\nu_i\colon B\to\mathbb{R}^{n_c}$ be defined by $\nu_i(x,q)=\nu[x,P_i+b(x)^\top q]$, so that

$$D_vL(x,v_i(x,q))^\top+P_i+b(x)^\top q+D_vc(x,v_i(x,q))^\top\nu_i=0,$$

for all $(x,q)\in B$ and for all $i=1,2$. From now on, we omit to mention the dependence of v_i and ν_i with respect to (x,q). We have

$$\begin{aligned}\frac{1}{C}\int_B|v_2-v_1|^2\mathrm{d}\mu_1(x,q)\leq&\int_B\big(D_vL(x,v_2)-D_vL(x,v_1)\big)(v_2-v_1)\mathrm{d}\mu_1(x,q)\\ &=(a)+(b)+(c),\end{aligned}\tag{4.5}$$

where

$$\begin{aligned}(a)&=(-P_2+P_1)^\top\int_B(v_2-v_1)\mathrm{d}\mu_1(x,q),\\ (b)&=\int_B\nu_2^\top D_vc(x,v_2)(v_1-v_2)\mathrm{d}\mu_1(x,q),\\ (c)&=\int_B\nu_1^\top D_vc(x,v_1)(v_2-v_1)\mathrm{d}\mu_1(x,q).\end{aligned}$$

We have

$$\begin{aligned}(a)=&\Big(-\psi(\textstyle\int v_2\mathrm{d}\mu_2)+\psi(\int v_2\mathrm{d}\mu_1)\Big)^\top\displaystyle\int_B(v_2-v_1)\mathrm{d}\mu_1(x,q)\\ &+\Big(-\psi(\textstyle\int v_2\mathrm{d}\mu_1)+\psi(\int v_1\mathrm{d}\mu_1)\Big)^\top\displaystyle\int_B(v_2-v_1)\mathrm{d}\mu_1(x,q).\end{aligned}\tag{4.6}$$

The second term in the right-hand side is non-positive, by monotonicity of ψ (Assumption (H1)-(iii)). By Lemma 4.1, we have

$$\Big|\int_B v_2(x,q)\mathrm{d}\mu_2(x,q)\Big|\leq C\quad\text{and}\quad\Big|\int_B v_1(x,q)\mathrm{d}\mu_2(x,q)\Big|\leq C.$$

For all $\varepsilon>0$, we denote $\omega(\varepsilon)=\sup\{|\psi(z_2)-\psi(z_1)|\ :\ z_1,z_2\in\bar{B}(C),\ \ |z_2-z_1|\leq\varepsilon\}$, the modulus of continuity of ψ on $\bar{B}(C)$. By Assumption (H1)-(iii), ψ is continuous, therefore uniformly continuous on $\bar{B}(C)$ and thus $\omega(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$. Using the Lipschitz continuity of v_2 and Cauchy-Schwarz inequality, we deduce from (4.6) that

$$\begin{aligned} (a) &\leq \omega\Big(\Big|\int_B v_2 \mathrm{d}\mu_2(x,q) - \int_B v_2 \mathrm{d}\mu_1(x,q))\Big|\Big)\Big|\int_B (v_2 - v_1)\mathrm{d}\mu_1(x,q)\Big| \\ &\leq \omega\big(Cd_1(\mu_1,\mu_2)\big)\Big(\int_B |v_2 - v_1|^2 \mathrm{d}\mu_1(x,q)\Big)^{1/2}. \end{aligned} \tag{4.7}$$

We also have the following inequalities, using $\nu_2 \geq 0$, the convexity of the maps c_i, and the complementarity condition:

$$\nu_2^\top D_v c(x,v_2)(v_1 - v_2) \leq \nu_2^\top(c(x,v_1) - c(x,v_2)) = \nu_2^\top c(x,v_1) \leq 0.$$

It follows that $(b) \leq 0$ and by symmetry, we also have $(c) \leq 0$. Combining (4.5), (4.7), and the nonpositivity of (b) and (c), we deduce that

$$\begin{aligned} \Big(\int_B |v_2 - v_1| \mathrm{d}\mu_1(x,q)\Big)^2 \leq & \int_B |v_2 - v_1|^2 \mathrm{d}\mu_1(x,q) \\ & \leq C\omega\big(Cd_1(\mu_1,\mu_2)\big)\int_B |v_2 - v_1|\mathrm{d}\mu_1(x,q). \end{aligned}$$

Therefore, $\int_B |v_2 - v_1|\mathrm{d}\mu_1(x,q) \leq C\omega(Cd_1(\mu_1,\mu_2))$. We finally obtain that

$$\begin{aligned} |P_2 - P_1| &\leq \omega\Big(\int_B v_2 \mathrm{d}\mu_2(x,q) - \int_B v_1 \mathrm{d}\mu_1(x,q)\Big) \\ &\leq \omega\Big(\int_B v_2 \mathrm{d}(\mu_2 - \mu_1)(x,q) + \int_B v_2 - v_1 \mathrm{d}\mu_1(x,q)\Big) \\ &\leq \omega\Big(Cd_1(\mu_1,\mu_2) + C\omega\big(Cd_1(\mu_1,\mu_2)\big)\Big) \longrightarrow 0, \quad \text{as } d_1(\mu_1,\mu_2) \to 0, \end{aligned}$$

which concludes the proof. ∎

4.2. Convergence properties.

LEMMA 4.3. *Let $(\kappa^i)_{i\in\mathbb{N}}$ be a sequence contained in $\mathcal{P}_1(\tilde{\Gamma})$ and let $K \subset \tilde{\Gamma}$ be a compact set such that* $\operatorname{supp}(\kappa^i) \subset K$ *for all $i \in \mathbb{N}$. Assume that $(\kappa^i)_{i\in\mathbb{N}}$ narrowly converges to κ. Then,*

$$\sup_{t\in[0,T]} d_1(\tilde{m}_t^{\kappa^i}, \tilde{m}_t^\kappa) \to 0 \qquad \textit{and} \qquad \sup_{t\in[0,T]} d_1(\mu_t^{\kappa^i}, \mu_t^\kappa) \to 0.$$

Proof. We start proving that for any $\bar{\kappa}, \hat{\kappa} \in \mathcal{P}_1(K)$ we have

$$\sup_{t\in[0,T]} d_1(m_t^{\bar{\kappa}}, m_t^{\hat{\kappa}}) \leq d_1(\bar{\kappa},\hat{\kappa}), \quad \text{and} \quad \sup_{t\in[0,T]} d_1(\mu_t^{\bar{\kappa}}, \mu_t^{\hat{\kappa}}) \leq d_1(\bar{\kappa},\hat{\kappa}), \tag{4.8}$$

We show the result for $\mu_t^{\bar{\kappa}}$ and $\mu_t^{\hat{\kappa}}$, and then the result for $m_t^{\bar{\kappa}}$ and $m_t^{\hat{\kappa}}$ is straightforward. By the Kantorovich-Rubinstein formula, for any $t \in [0,T]$ we have

$$\begin{aligned} d_1(\mu_t^{\bar{\kappa}}, \mu_t^{\hat{\kappa}}) = & \sup_{\varphi \in \mathrm{Lip}_1(\mathbb{R}^n\times\mathbb{R}^n)} \int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi(x,q)\mathrm{d}(\mu_t^{\bar{\kappa}} - \mu_t^{\hat{\kappa}})(x,q) \\ & = \sup_{\varphi \in \mathrm{Lip}_1(\mathbb{R}^n\times\mathbb{R}^n)} \int_{\tilde{\Gamma}} \varphi(\gamma(t),p(t))\mathrm{d}(\bar{\kappa} - \hat{\kappa})(\gamma,p) \;\leq\; d_1(\bar{\kappa},\hat{\kappa}). \end{aligned}$$

In the last inequality we use the fact that given $\varphi \in \mathrm{Lip}_1(\mathbb{R}^n \times \mathbb{R}^n)$, the map $(\gamma, p) \mapsto \varphi(\gamma(t), p(t))$ belongs to $\mathrm{Lip}_1(\tilde{\Gamma})$, for all $t \in [0,T]$.

Since $(\kappa^i)_{i\in\mathbb{N}} \subset \mathcal{P}_1(K)$ narrowly converges to κ, by [2, Proposition 7.1.5], we obtain $d_1(\kappa^i,\kappa) \to 0$. The conclusion follows with (4.8). ∎

In what follows we consider the following compact subset of $\tilde{\Gamma}$,

$$\tilde{\Gamma}_B := \{(\gamma,p) \in \Gamma : \|\gamma\|_\infty \le M_1,\ \|p\|_\infty \le M_2,\ \|\dot\gamma\|_2 \le T^{\frac{1}{2}} M_3,\ \|\dot p\|_2 \le T^{\frac{1}{2}} M_4\}, \tag{4.9}$$

where M_1, M_2, M_3 and M_4 were introduced in Proposition 3.4, Proposition 3.6 and Lemma 3.8.

LEMMA 4.4. *Let $\kappa \in \mathcal{P}_{m_0}(\tilde{\Gamma})$ with $\mathrm{supp}(\kappa) \subset \tilde{\Gamma}_B$. Then $\tilde{m}_t^\kappa$ and μ_t^κ are $\frac{1}{2}$-Hölder continuous w.r.t. $t \in [0,T]$.*

Proof. Recalling that $B = \overline{B}(0,M_1) \times \overline{B}(0,M_2) \subset \mathbb{R}^n \times \mathbb{R}^n$, since $\mathrm{supp}(\kappa) \subset \tilde{\Gamma}_B$, we obtain $\mathrm{supp}(\mu_t^\kappa) \subset B$ and $\mathrm{supp}(\tilde{m}_t^\kappa) \subset \bar{B}(0,M_1)$ for all $t \in [0,T]$.

For all $s,t \in [0,T]$ we have

$$\begin{aligned} d_1(\mu_t^\kappa,\mu_s^\kappa) &= \sup_{\varphi \in \mathrm{Lip}_1(\mathbb{R}^n\times\mathbb{R}^n)} \int_B \varphi(x,q)(\mathrm{d}\mu_t^\kappa - \mathrm{d}\mu_s^\kappa)(x,q) \\ &= \sup_{\varphi \in \mathrm{Lip}_1(\mathbb{R}^n\times\mathbb{R}^n)} \int_\Gamma [\varphi(\gamma(t),p(t)) - \varphi(\gamma(s),p(s))]\mathrm{d}\kappa(\gamma,p) \\ &\le \int_\Gamma \max\{|\gamma(t)-\gamma(s)|, |p(t)-p(s)|\}\mathrm{d}\kappa(\gamma,p) \\ &\le T^{\frac{1}{2}} \max\{M_3,M_4\}|t-s|^{\frac{1}{2}}. \end{aligned}$$

The last inequality holds by the assumption $\mathrm{supp}(\kappa) \subset \tilde{\Gamma}_B$. ∎

Remark 4.5. Given $\kappa \in \mathcal{P}(\tilde{\Gamma}_B)$, by Lemma 4.3, we obtain $\tilde{m}^\kappa \in C([0,T];\mathcal{P}_1(\mathbb{R}^n))$. Setting $\tilde{P}^\kappa(t) = \mathbf{P}[\mu_t^\kappa]$, where $\mathbf{P}$ is defined in Lemma 4.1, by Lemma 4.2 and Lemma 4.3, we obtain $\tilde{P}^\kappa \in C(0,T;\mathbb{R}^m)$. Therefore, the definition of $\tilde{J}^\kappa$ in section 2.2.2 makes sense.

LEMMA 4.6. *Let $(\kappa^i)_{i\in\mathbb{N}} \subset \mathcal{P}_{m_0}(\tilde{\Gamma})$, $\kappa \in \mathcal{P}_{m_0}(\tilde{\Gamma})$ be such that $\mathrm{supp}(\kappa^i) \subset \tilde{\Gamma}_B$ for all i and $\mathrm{supp}(\kappa) \subset \tilde{\Gamma}_B$. Assume that κ^i narrowly converges to $\bar{\kappa}$. Let $(x_i)_{i\in\mathbb{N}} \subset K_0$ be a sequence such that $x_i \to \bar{x}$ and let $(\gamma_i,p_i)_{i\in\mathbb{N}} \subset \tilde{\Gamma}^{\kappa^i}[x_i]$ (again defined in section 2.2.2) be a sequence such that $(\gamma_i,p_i) \to (\bar\gamma,\bar p)$ uniformly on $[0,T]$. Then $(\bar\gamma,\bar p) \in \tilde{\Gamma}^{\bar\kappa}[\bar x]$.*

Proof. We have to prove that there exists $\bar v \in L^2(0,T;\mathbb{R}^m)$ such that $(\bar\gamma,\bar v) \in \mathcal{K}[\bar x]$ and

$$\tilde{J}^{\bar\kappa}(\bar\gamma,\bar v) \le \tilde{J}^{\bar\kappa}(\gamma,v) \quad \forall (\gamma,v) \in \mathcal{K}[\bar x].$$

In addition, we have to prove that $\bar p$ is the costate associated with $(\bar\gamma,\bar v)$, in the sense of Definition 3.7.

Since $(\gamma_i,p_i)_{i\in\mathbb{N}} \subset \tilde{\Gamma}^{\kappa^i}[x_i]$, there exists for all $i \in \mathbb{N}$ a control $v_i \in L^2(0,T;\mathbb{R}^m)$ such that $(\gamma_i,v_i) \in \mathcal{K}[x_i]$ and (γ_i,v_i) is optimal for $\tilde{J}^{\kappa^i}$. By Proposition 3.4, since $(x_i)_{i\in\mathbb{N}} \subset K_0$, we have $\|\gamma_i(t)\|_\infty \le M_1$ and $\|v_i\|_2 \le C$, for all $i \in \mathbb{N}$. Therefore, there exists $\bar v \in L^2(0,T;\mathbb{R}^m)$ such that, up to a subsequence, $v_i \rightharpoonup \bar v$. By Lemma 3.8, the sequence $(\gamma_i)_{i\in\mathbb{N}}$ is a bounded sequence in $H^1(0,T;\mathbb{R}^n)$, since $\gamma_i \to \bar\gamma$ in $C(0,T;\mathbb{R}^n)$, it follows that $\bar{\gamma} \in H^1$ and $\dot{\gamma}_i \rightharpoonup \dot{\bar{\gamma}}$ in $L^2(0,T;\mathbb{R}^n)$. In addition, by (H2)-(ii), (H3)-(iii), the uniform convergence of γ_i to $\bar{\gamma}$ and the weak convergence of v_i to $\bar{v}$ we obtain

$$a(\gamma_i) + b(\gamma_i)v_i \rightharpoonup a(\bar{\gamma}) + b(\bar{\gamma})\bar{v}, \quad \text{in } L^2(0,T;\mathbb{R}^n),$$

which implies that $\dot{\bar{\gamma}}(t) = a(\bar{\gamma}(t)) + b(\bar{\gamma}(t))\bar{v}(t)$, for a.e. $t \in (0,T)$. It is clear that $\bar{\gamma}(0) = \bar{x}$.

Furthermore, for all $i \in \mathbb{N}$ there exists $(\lambda_1^i, \lambda_2^i, \nu_i) \in \mathbb{R}^{n_{g_1}} \times \mathbb{R}^{n_{g_2}} \times L^\infty(0,T;\mathbb{R}^{n_c})$ such that (3.7), (3.8) and (3.9) hold for (γ_i, v_i, p_i) and $\lambda_0^i = 1$. By the proof of Proposition 3.6, we obtain that $(\lambda_1^i, \lambda_2^i)_{i\in\mathbb{N}}$ is bounded, then there exists a subsequence, still denoted $(\lambda_1^i, \lambda_2^i)_{i\in\mathbb{N}}$, that converges to $(\bar{\lambda}_1, \bar{\lambda}_2)$.

By Lemma 3.2 and (3.8), we deduce

$$v_i(t) = v\left[\gamma_i(t), \tilde{P}^{\kappa^i}(t) + b(\gamma_i(t))^\top p_i(t)\right], \quad \nu_i(t) = \nu\left[\gamma_i(t), \tilde{P}^{\kappa^i}(t) + b(\gamma_i(t))^\top p_i(t)\right].$$

By our assumptions, Lemma 4.2 and Lemma 4.3, the sequences (γ_i), (p_i) and $(\tilde{P}^{\kappa^i})$ are bounded and they converge to $\bar{\gamma}$, $\bar{p}$ and $\tilde{P}^{\bar{\kappa}}$, uniformly over $[0,T]$. By Lemma 3.2, the maps $v[\cdot,\cdot]$ and $\nu[\cdot,\cdot]$ are Lipschitz continuous over bounded sets, then

$$v_i(t) \to \bar{v}(t) = v\left[\bar{\gamma}(t), \tilde{P}^{\bar{\kappa}}(t) + b(\bar{\gamma}(t))^\top \bar{p}(t)\right], \tag{4.10}$$

and

$$\nu_i(t) \to \bar{\nu}(t), \text{ where } \bar{\nu}(t) = \nu\left[\bar{\gamma}(t), \tilde{P}^{\bar{\kappa}}(t) + b(\bar{\gamma}(t))^\top \bar{p}(t)\right],$$

uniformly over $[0,T]$. In addition by Lemma 4.3, $\sup_{t\in[0,T]} d_1(\tilde{m}_t^{\kappa^i}, \tilde{m}_t^{\bar{\kappa}}) \to 0$. Therefore by (H2) and (H3), we can pass to the limit in (3.7). By similar arguments we can pass to the limit in (3.8) and (3.9). Finally we can conclude that $(\bar{p}, 1, \bar{\lambda}_1, \bar{\lambda}_2, \bar{\nu})$ satisfies the adjoint equation, the stationary condition and the complementarity condition for $(\bar{\gamma}, \bar{v})$.

Now, we prove the optimality of $(\bar{\gamma}, \bar{v})$ for $\tilde{J}^{\bar{\kappa}}$. First we prove that

$$\tilde{J}^{\bar{\kappa}}(\bar{\gamma}, \bar{v}) = \lim_{i\to\infty} \tilde{J}^{\kappa^i}(\gamma_i, v_i). \tag{4.11}$$

By the uniform convergence of the sequence (γ_i), Lemma 4.3 and (H2) we have

$$\int_0^T f\left(\gamma_i(t), \tilde{m}_t^{\kappa^i}\right) \mathrm{d}t \to \int_0^T f\left(\bar{\gamma}(t), \tilde{m}_t^{\bar{\kappa}}\right) \mathrm{d}t \text{ and } g_0\left(\gamma_i(T), \tilde{m}_T^{\kappa^i}\right) \to g_0\left(\bar{\gamma}(T), \tilde{m}_T^{\bar{\kappa}}\right).$$

Skipping the time arguments, we have

$$\int_0^T \left[\langle \tilde{P}^{\kappa^i}, v_i\rangle - \langle \tilde{P}^{\bar{\kappa}}, \bar{v}\rangle\right] \mathrm{d}t = \int_0^T \left[\langle \tilde{P}^{\kappa^i} - \tilde{P}^{\bar{\kappa}}, v_i\rangle + \langle \tilde{P}^{\bar{\kappa}}, v_i - \bar{v}\rangle\right] \mathrm{d}t.$$

By Lemma 4.2, Lemma 4.3, the uniform convergence in (4.10) and the boundedness of the sequences $(\tilde{P}^{\kappa^i})$ and (v_i) we conclude that $\int_0^T \langle \tilde{P}^{\kappa^i}, v_i\rangle \mathrm{d}t \to \int_0^T \langle \tilde{P}^{\bar{\kappa}}, \bar{v}\rangle \mathrm{d}t$. By (H2)-(i) and the uniform convergence of (γ_i) and (v_i) to $\bar{\gamma}$ and $\bar{v}$, respectively, we deduce that $\int_0^T L(\gamma_i, v_i)\mathrm{d}t \to \int_0^T L(\bar{\gamma}, \bar{v})\mathrm{d}t$. Combining the above estimates, (4.11) follows.

Now, let $(\hat{\gamma}, \hat{v}) \in \mathcal{K}[\bar{x}]$ be an optimal solution for $\tilde{J}^{\bar{\kappa}}$ and initial condition $\bar{x}$. By (H5)(i)-(ii), Robinson's constraint qualification (see [7, (2.163)]) holds at $\hat{v}$. By

[7, Theorem 2.87] and (H2) we conclude that there exists a sequence $(\hat{v}_i)_{i\in\mathbb{N}} \subset L^\infty(0,T;\mathbb{R}^m)$ such that $\|\hat{v}_i - \hat{v}\|_\infty \to 0$, and the sequence $(\hat{\gamma}_i)_{i\in\mathbb{N}}$, given by

$$\begin{cases} \dot{\hat{\gamma}}_i(t) = & a(\hat{\gamma}_i(t)) + b(\hat{\gamma}_i(t))\hat{v}_i(t), \quad \text{for a.e. } t \in [0,T] \\ \hat{\gamma}_i(0) = & x_i \end{cases}$$

is such that $(\hat{\gamma}_i, \hat{v}_i) \in \mathcal{K}[x_i]$. In addition, by our assumptions and Grönwall's Lemma we deduce that $(\hat{\gamma}_i)_{i\in\mathbb{N}}$ is uniformly bounded in $L^\infty(0,T;\mathbb{R}^n)$ and $\|\hat{\gamma}_i - \hat{\gamma}\|_\infty \to 0$.

By the optimality of $(\gamma_i, p_i) \in \tilde{\Gamma}^{\kappa^i}[x_i]$, we have

$$\tilde{J}^{\kappa^i}(\gamma_i, v_i) \le \tilde{J}^{\kappa^i}(\hat{\gamma}_i, \hat{v}_i) \quad \forall i \in \mathbb{N}. \tag{4.12}$$

Since $\|\hat{v}_i - \hat{v}\|_\infty \to 0$ and $\|\hat{\gamma}_i - \hat{\gamma}\|_\infty \to 0$, arguing as above we obtain,

$$\lim_{i\to\infty} \tilde{J}^{\kappa^i}(\hat{\gamma}_i, \hat{v}_i) = \tilde{J}^{\bar{\kappa}}(\hat{\gamma}, \hat{v}). \tag{4.13}$$

By (4.11), (4.12) and (4.13), we have

$$\tilde{J}^{\bar{\kappa}}(\bar{\gamma}, \bar{v}) = \lim_{i\to\infty} \tilde{J}^{\kappa^i}(\gamma_i, v_i) \le \lim_{i\to\infty} \tilde{J}^{\kappa^i}(\hat{\gamma}_i, \hat{v}_i) = \tilde{J}^{\bar{\kappa}}(\hat{\gamma}, \hat{v}).$$

Then, $(\bar{\gamma}, \bar{v})$ is optimal, which finally proves that $(\bar{\gamma}, \bar{p}) \in \tilde{\Gamma}^{\bar{\kappa}}[\bar{x}]$. ∎

4.3. Existence results. In this section, we characterize auxiliary MFGC equilibria as fixed points of a set-valued map. Applying Kakutani's fixed point theorem, we prove the existence of such equilibria. Then we show how an auxiliary equilibrium defines a Lagrangian one.

By [2, Theorem 5.3.1] (Disintegration Theorem), for any $\kappa \in \mathcal{P}_{m_0}(\Gamma)$, there exists a m_0-a.e. uniquely determined Borel measurable family $\{\kappa_x\}_{x\in\mathbb{R}^n} \subset \mathcal{P}(\tilde{\Gamma})$ such that

$$\mathrm{supp}(\kappa_x) \subset \tilde{\Gamma}[x], \quad m_0\text{−a.e. } x \in \mathbb{R}^n,$$

and for any Borel map $\varphi\colon \tilde{\Gamma} \to [0,+\infty]$,

$$\int_{\tilde{\Gamma}} \varphi(\gamma,p)\mathrm{d}\kappa(\gamma,p) = \int_{\mathbb{R}^n} \Big(\int_{\tilde{\Gamma}[x]} \varphi(\gamma,p)\mathrm{d}\kappa_x(\gamma,p) \Big) \mathrm{d}m_0(x).$$

Following the lines of [8], we define the set-valued map $E : \mathcal{P}_{m_0}(\Gamma) \rightrightarrows \mathcal{P}_{m_0}(\Gamma)$ as

$$E(\kappa) = \{\hat{\kappa} \in \mathcal{P}_{m_0}(\Gamma) : \mathrm{supp}(\hat{\kappa}_x) \subset \tilde{\Gamma}^\kappa[x],\ m_0\text{−a.e. } x \in \mathbb{R}^n\}.$$

It follows that κ is an auxiliary MFGC equilibrium if and only if $\kappa \in E(\kappa)$.

THEOREM 4.7. *There exists at least one auxiliary MFGC equilibrium.*

Proof. Arguing as in [8, Lemma 3.5], for any $\kappa \in \mathcal{P}_{m_0}(\Gamma)$ the set $E(\kappa)$ is a nonempty convex set. By Proposition 3.4, Proposition 3.6 and Lemma 3.8 we have

$$E(\kappa) \subset \mathcal{P}_{m_0}(\tilde{\Gamma}_B), \quad \forall \kappa \in \mathcal{P}_{m_0}(\Gamma),$$

where $\tilde{\Gamma}_B$ was introduced in (4.9). By Lemma 4.6, and [8, Lemma 3.6], we conclude that the map $E\colon \mathcal{P}_{m_0}(\Gamma) \rightrightarrows \mathcal{P}_{m_0}(\Gamma)$ has closed graph.

Finally, since the set $\tilde{\Gamma}_B$ is a compact subset of $\tilde{\Gamma}$, we obtain that $\mathcal{P}_{m_0}(\tilde{\Gamma}_B)$ is a nonempty compact convex set. Then, we can apply Kakutani's fixed point theorem, to deduce that there exists $\hat{\kappa} \in \mathcal{P}_{m_0}(\tilde{\Gamma}_B)$ such that $\hat{\kappa} \in E(\hat{\kappa})$. ∎

Given $\kappa \in \mathcal{P}(\tilde{\Gamma}_B)$, let $V^\kappa \colon \tilde{\Gamma} \to \mathcal{U}$ be defined by

$$V^\kappa(\gamma, p) = v\left[\gamma, \tilde{P}^\kappa + b(\gamma)^\top p\right],$$

where the r.h.s. is the Nemytskii operator associated with the auxiliary mapping introduced in Lemma 4.1. Let $\pi_1 \colon \tilde{\Gamma} \to \Gamma$ be such that $\pi_1(\gamma, p) = \gamma$. Then, we define $\eta[\kappa] \in \mathcal{P}_1(\Gamma \times \mathcal{U})$ by

$$\eta[\kappa] = (\pi_1, V^\kappa) \sharp \kappa.$$

COROLLARY 4.8. *Let $\kappa \in \mathcal{P}(\tilde{\Gamma})$ be an auxiliary MFGC equilibrium. Then, $\eta[\kappa] \in \mathcal{P}_1(\Gamma \times \mathcal{U})$ is a Lagrangian MFGC equilibrium.*

Proof. For the sake of simplicity we note η instead of $\eta[\kappa]$. The main point is to prove that $\tilde{P}^\kappa = P^\eta$, where P^η was introduced in (2.2). By the definition of η, it is supported on regular curves, then

$$\begin{aligned} P^\eta(t) &= \psi\Big(\textstyle\int_{\Gamma \times \mathcal{U}} v(t) \mathrm{d}\eta(\gamma, v)\Big) \\ &= \psi\Big(\textstyle\int_{\tilde{\Gamma}} v\left[\gamma(t), \tilde{P}^\kappa(t) + b(\gamma(t))^\top p(t)\right] \mathrm{d}\kappa(\gamma, p)\Big) \\ &= \psi\Big(\textstyle\int_{\mathbb{R}^n \times \mathbb{R}^n} v\left[x, \tilde{P}^\kappa(t) + b(x)^\top q\right] \mathrm{d}\mu_t^\kappa(x, q)\Big) = \tilde{P}^\kappa(t). \end{aligned}$$

The last equality follows from (4.2). It is clear that $\tilde{m}^\kappa = m^\eta$, then $J^\eta = \tilde{J}^\kappa$. Since κ is an auxiliary MFGC equilibrium, any $(\bar{\gamma}, \bar{p}) \in \operatorname{supp}(\kappa)$ defines an optimal pair $(\bar{\gamma}, V^\kappa(\bar{\gamma}, \bar{p}))$ for J^η. We conclude that η is a Lagrangian MFGC equilibrium. ◻

Remark 4.9. Let us comment on the impossibility to employ a similar fixed point approach directly based on the notion of Lagrangian equilibria (Definition 2.1). Consider a probability distribution η of state-control trajectories. From the definition of P^η, there is no regularity property (with respect to time) to expect, since the controls in problem (2.1) are taken in $L^2(0, T; \mathbb{R}^m)$. Consequently, it is not possible to use relation (3.19) to derive any regularity property for the optimal controls with respect to the criterion $J[m^\eta, P^\eta]$ and thus it does not seem possible to construct an appropriate compact set of probability distributions of state-control trajectories, on which some fixed point relation could be defined.

5. Uniqueness.

As usual in the MFG theory, by adding some monotonicity assumptions we can obtain uniqueness results.

DEFINITION 5.1. *A function $\varphi \colon \mathbb{R}^n \times \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}$ is monotone if*

$$\int_{\mathbb{R}^n} (\varphi(x, m_1) - \varphi(x, m_2)) \, \mathrm{d}\left(m_1 - m_2\right)(x) \geq 0, \quad \forall m_1,\ m_2 \in \mathcal{P}(\mathbb{R}^n).$$

It is strictly monotone if it is monotone and

$$\int_{\mathbb{R}^n} (\varphi(x, m_1) - \varphi(x, m_2)) \, \mathrm{d}\left(m_1 - m_2\right)(x) = 0,$$

if and only if $\varphi(x, m_1) = \varphi(x, m_2)$ for all $x \in \mathbb{R}^n$.

An example of strictly monotone function can be found in [8].

THEOREM 5.2. *Assume that f and g_0 are strictly monotone and ϕ is strictly convex. Let $\eta_1, \eta_2 \in \mathcal{P}_{m_0}(\Gamma \times \mathcal{U})$ be Lagrangian MFGC equilibria for m_0, then $P^{\eta_1} = P^{\eta_2}$ and $J^{\eta_1} = J^{\eta_2}$.*

Proof. Let us define $u^i(x) = \inf_{(\gamma,v)\in\mathcal{K}[x]} J^{\eta_i}(\gamma, v)$, $i = 1, 2$. Let $(\gamma, v) \in \operatorname{supp}(\eta_1)$, then

$$u^1(\gamma(0)) = \int_0^T (L(\gamma(t), v(t)) + \langle P^{\eta_1}(t), v(t)\rangle + f(\gamma(t), m_t^{\eta_1}))\,\mathrm{d}t + g_0(\gamma(T), m_T^{\eta_1})$$
$$u^2(\gamma(0)) \le \int_0^T (L(\gamma(t), v(t)) + \langle P^{\eta_2}(t), v(t)\rangle + f(\gamma(t), m_t^{\eta_2}))\,\mathrm{d}t + g_0(\gamma(T), m_T^{\eta_2}).$$

Integrating w.r.t. η_1 we obtain

$$\begin{aligned}\int_{\Gamma\times\mathcal{U}} (u^1(\gamma(0)) - u^2(\gamma(0)))\mathrm{d}\eta_1(\gamma, v) + \int_{\Gamma\times\mathcal{U}}\int_0^T \langle P^{\eta_2}(t) - P^{\eta_1}(t), v(t)\rangle \mathrm{d}t\mathrm{d}\eta_1(\gamma, v)\\ \ge \int_{\Gamma\times\mathcal{U}} (g_0(\gamma(T), m_T^{\eta_1}) - g_0(\gamma(T), m_T^{\eta_2}))\,\mathrm{d}\eta_1(\gamma, v)\\ + \int_{\Gamma\times\mathcal{U}}\int_0^T (f(\gamma(t), m_t^{\eta_1}) - f(\gamma(t), m_t^{\eta_2}))\,\mathrm{d}t\mathrm{d}\eta_1(\gamma, v).\end{aligned}$$

By the definition of m^{η_1} we obtain

$$\begin{aligned}\int_{\mathbb{R}^n} (u^1(x) - u^2(x))\mathrm{d}m_0(x) + \int_{\Gamma\times\mathcal{U}}\int_0^T \langle P^{\eta_2}(t) - P^{\eta_1}(t), v(t)\rangle \mathrm{d}t\mathrm{d}\eta_1(\gamma, v)\\ \ge \int_{\mathbb{R}^n} (g_0(x, m_T^{\eta_1}) - g_0(x, m_T^{\eta_2}))\,\mathrm{d}m_T^{\eta_1}(x)\\ + \int_0^T\int_{\mathbb{R}^n} (f(x, m_t^{\eta_1}) - f(x, m_t^{\eta_2}))\,\mathrm{d}m_t^{\eta_1}(x)\mathrm{d}t.\end{aligned}$$

Arguing in a similar way for η_2, we deduce

$$\begin{aligned}\int_0^T\int_{\Gamma\times\mathcal{U}} \langle P^{\eta_2}(t) - P^{\eta_1}(t), v(t)\rangle \mathrm{d}(\eta_1 - \eta_2)(\gamma, v)\mathrm{d}t\\ \ge \int_{\mathbb{R}^n} (g_0(x, m_T^{\eta_1}) - g_0(x, m_T^{\eta_2}))\,\mathrm{d}\left(m_T^{\eta_1} - m_T^{\eta_2}(x)\right)(x)\\ + \int_0^T\int_{\mathbb{R}^n} (f(x, m_t^{\eta_1}) - f(x, m_t^{\eta_2}))\,\mathrm{d}\left(m_t^{\eta_1} - m_t^{\eta_2}\right)(x)\mathrm{d}t. \qquad (5.1)\end{aligned}$$

By the definition of P^{η_i} we deduce

$$\begin{aligned}\int_0^T\int_{\Gamma\times\mathcal{U}} \langle P^{\eta_2}(t) - P^{\eta_1}(t), v(t)\rangle \mathrm{d}(\eta_1 - \eta_2)(\gamma, v)\mathrm{d}t\\ = \int_0^T \Big\langle \nabla\phi\left(\textstyle\int v\mathrm{d}\eta_2(\gamma, v)\right) - \nabla\phi\left(\textstyle\int v\mathrm{d}\eta_1(\gamma, v)\right), \textstyle\int v\mathrm{d}\eta_1(\gamma, v) - \int v\mathrm{d}\eta_2(\gamma, v)\Big\rangle \mathrm{d}t\end{aligned}$$

and the r.h.s. is non-positive, since ϕ is convex. In addition, since f and g_0 are monotone, we deduce that the three terms in (5.1) vanish. Since f and g_0 are strictly monotone we obtain for all $x \in \mathbb{R}^n$ and a.e. $t \in (0, T)$,

$$f(x, m_t^{\eta_1}) = f(x, m_t^{\eta_2}) \quad \text{and} \quad g_0(x, m_T^{\eta_1}) = g_0(x, m_T^{\eta_2}).$$

By the strict convexity of ϕ we have

$$\int v \mathrm{d}\eta_1(\gamma, v) = \int v \mathrm{d}\eta_2(\gamma, v), \quad \text{a.e. } t \in (0,T),$$

which in particular implies $P^{\eta_1} = P^{\eta_2}$. The result follows. ∎

Remark 5.3. As noted in [8], if we assume that ϕ is strictly convex, g_0 is monotone and f satisfies

$$\int_{\mathbb{R}^n} \left(f(x, m_1) - f(x, m_2)\right) \mathrm{d}\left(m_1 - m_2\right)(x) \leq 0 \ \Rightarrow m_1 = m_2,$$

then, following the ideas of the above proof, we obtain $P^{\eta_1} = P^{\eta_2}$ and $m^{\eta_1} = m^{\eta_2}$.

6. Conclusion. We have proved the existence of a Lagrangian equilibrium for an MFG of controls with final state and mixed state-control constraints, and a class of nonlinear dynamics. Using auxiliary mappings and a priori estimates on optimal state-costate trajectories, we have reformulated the problem as a fixed point problem on a compact set of probability measures on state-costate trajectories. As explained in Remark 4.9, this reformulation was necessary, in the absence of smoothing properties of the price interaction.

We want to point out that for the problem studied here, it is not possible to think of an equivalent notion of equilibrium defined as a solution to a system of coupled partial differential equations (HJB and continuity equation), as it is usual for the unconstrained case, because of the final-state constraints. It should be possible to obtained a well-posed HJB equation in the presence of state-control constraints only. Yet it seems quite difficult to construct a pointwise solution to some appropriate PDE system, as it is done in [9]. In this reference, a feedback control is constructed thanks to the differentiability of the value function, itself obtained with the strict convexity of the Hamiltonian. This last property is however lost (in general) in the presence of mixed state-control constraints.

In some future work, one could address the extension of our aggregative MFG model to the case of pure state constraints, as those considered in [8]. As we already pointed out, our analysis relies in a crucial way on some a priori estimates on the costate, whose evolution is not impacted by the price variable. Proving the regularity of the costate, in the presence of pure state constraints and a merely measurable price function, seems however to be a great challenge.

Concerning the numerical resolution of the problem, we mention that the mean field game problem investigated here has a variational structure if the congestion term is the derivative (in some sense) of a convex potential. Therefore, the ideas for the discretization of variational MFGs in Lagrangian form investigated in [27] could certainly be applied to our setting.

Appendix A. Proof of optimality conditions.

We provide in this section a proof of the optimality conditions stated in Proposition 3.5. An important difficulty is the fact that optimal controls are not a priori known to be bounded (we are not able to prove the boundedness of optimal controls without having the optimality conditions at hand). It is therefore not possible to formulate the optimal control problem as an abstract problem satisfying a qualification condition in L^∞ and to derive easily optimality conditions, as it is done in [5] for example. It turns out that the optimal control problem can be naturally formulated as an optimal control problem for which the dynamic constraint takes the form of a differential inclusion. This enables us to use the associated optimality conditions, referred to as *extended Euler-Lagrange* conditions in the literature. More precisely, our analysis is based on [28, Theorem 7.5.1], which covers the case of unbounded controls and requires few regularity assumptions.

We first introduce two definitions of cones, used for the expression of the optimality conditions for problems with differential inclusions. Given a closed subset K of $\mathbb{R}^\ell$ and $x \in K$, we call proximal normal cone of K at x the set $N_K^P(x)$ defined by

$$N_K^P(x) = \big\{ p \in \mathbb{R}^\ell : \exists C > 0,\, \forall y \in K,\, \langle p, y - x \rangle \leq C|y-x|^2 \big\}.$$

That is, $p \in N_K^P(x)$ if and only if, for some $C > 0$,

$$x \in \operatorname{argmin} \left\{ \langle -p, y \rangle + C|y-x|^2 \ : \ y \in K \right\}. \tag{A.1}$$

The limiting normal cone $N_K(x)$ is defined by

$$N_K(x) = \left\{ p \in \mathbb{R}^\ell \,\middle|\, \exists (x_k, p_k)_{k \in \mathbb{N}} \text{ such that: } \begin{array}{l} (x_k, p_k) \to (x, p), \text{ as } k \to \infty \\ x_k \in K,\ p_k \in N_K^P(x_k) \ \ \forall k \in \mathbb{N} \end{array} \right\}.$$

Proof of Proposition 3.5. Step 1: reformulation of the optimal control problem. Let us fix a solution $(\bar{\gamma}, \bar{v}) \in H^1(0,T;\mathbb{R}^n) \times L^2(0,T;\mathbb{R}^m)$ to (2.1). In order to alleviate the notation, we first define

$$\tilde{L}(t,x,v) = L(x,v) + \langle P(t), v \rangle + f(x, m(t)),$$

for all $(x,v) \in \mathbb{R}^{n+m}$ and for a.e. $t \in (0,T)$. We work with an augmented state variable $x = (x^{(1)}, x^{(2)}, x^{(3)}) \in \mathbb{R}^{n+1+m}$. We define the set-valued map $F\colon [0,T] \times \mathbb{R}^{n+1+m} \rightrightarrows \mathbb{R}^{n+1+m}$ by $F(t,x) = \big\{ \xi(t,x,v,z) :\ c(x,v) \leq 0,\ z \geq 0 \big\}$, where

$$\xi^{(1)}(t,x,v,z) = a(x^{(1)}) + b(x^{(1)})v, \ \ \xi^{(2)}(t,x,v,z) = \tilde{L}(t,x^{(1)},v) + z, \ \ \xi^{(3)}(t,x,v,z) = v.$$

The component $\xi^{(1)}$ coincides with the dynamics of the original state variable. The second component allows to put the problem in Mayer form. The third component has a technical purpose, it allows in particular to prove easily that $F(t,x)$ is closed (which would be delicate otherwise, since the controls are not necessarily bounded). The initial condition associated with the new state variable is defined by $\bar{x}_0 = (x_0, 0, 0) \in \mathbb{R}^{n+1+m}$. Let $K \subseteq \mathbb{R}^{2(n+1+m)}$ be given by

$$K = \big\{ (x_i, x_f) \in \mathbb{R}^{2(n+1+m)} : x_i = \bar{x}_0,\ g_1\big(x_f^{(1)}\big) = 0,\ g_2\big(x_f^{(1)}\big) \leq 0 \big\}.$$

We define $\Phi\colon \mathbb{R}^{2(n+1+m)} \to \mathbb{R}$ by $\Phi(x_i, x_f) = g_0(x_f^{(1)}) + x_f^{(2)}$. The optimal control problem (2.1) can finally be reformulated as follows:

$$\inf_{x \in H^1(0,T;\mathbb{R}^{n+1+m})} \Phi(x(0), x(T)), \ \ \text{subject to: } \left\{ \begin{array}{l} \dot{x}(t) \in F(t, x(t)), \text{ for a.e. } t \in (0,T), \\ (x(0), x(T)) \in K. \end{array} \right. \tag{A.2}$$

More precisely, the trajectory $\bar{x}$, defined by

$$\begin{cases} \bar{x}^{(1)}(t) = \bar{\gamma}(t) \\ \bar{x}^{(2)}(t) = \int_0^t \tilde{L}(s, \bar{\gamma}(s), \bar{v}(s)) \mathrm{d}s \\ \bar{x}^{(3)}(t) = \int_0^t \bar{v}(s) \mathrm{d}s \end{cases}$$

is a solution to (A.2). Denoting $\bar{\xi} = \dot{\bar{x}}$, we note that $\bar{\xi}(t) = \xi(t, \bar{x}(t), \bar{v}(t), 0)$.

Step 2: verification of the technical conditions of [28, Theorem 7.5.1]. It is easily verified that for a.e. $t \in (0,T)$, $F(t,x)$ is non-empty and convex, as a consequence of Assumptions (H1)-(i), (H1)-(ii), and (H4)-(i). It is also easily verified that F is measurable and has a closed graph. It remains to show that there exist $\eta > 0$ and $k \in L^1(0,T)$ such that

$$F(t,\tilde{x}) \cap \big(\dot{\bar{x}}(t) + \eta k(t)\bar{B}(1)\big) \subseteq F(t,x) + k(t)|\tilde{x} - x|\bar{B}(1), \tag{A.3}$$

for a.e. $t \in (0,T)$ and for all x and $\tilde{x}$ such that $|x - \bar{x}(t)| \leq \eta$ and $|\tilde{x} - \bar{x}(t)| \leq \eta$. Let $t \in (0,T)$, let x and $\tilde{x}$ be such that $|x - \bar{x}| \leq \delta/2$ and $|\tilde{x} - \bar{x}| \leq \delta/2$, where δ is given by Lemma 3.1, with $R = \|\bar{\gamma}\|_{L^\infty(0,T;\mathbb{R}^n)}$. Let $\bar{k}(t) = 1 + |\bar{v}(t)|^2$. Let $\tilde{\xi} \in F(t,\tilde{x}) \cap \big(\dot{\bar{x}}(t) + \bar{k}(t)\bar{B}(1)\big)$. Let $\tilde{v} \in \mathbb{R}^m$ and $\tilde{z} \in \mathbb{R}$ be such that $\tilde{\xi} = \xi(t,\tilde{x},\tilde{v},\tilde{z})$, $c(\tilde{x},\tilde{v}) \leq 0$ and $\tilde{z} \geq 0$. Since $|\tilde{\xi} - \bar{\xi}(t)| \leq \bar{k}(t)$, we deduce that

$$|\tilde{\xi}^{(3)} - \bar{\xi}^{(3)}(t)| = |\tilde{v} - \bar{v}(t)| \leq \bar{k}(t).$$

Therefore

$$|\tilde{v}| \leq |\tilde{v} - \bar{v}(t)| + |\bar{v}(t)| \leq \bar{k}(t) + \frac{1}{2} + \frac{1}{2}|\bar{v}(t)|^2 \leq \frac{3}{2}\bar{k}(t).$$

We also have $|x - \tilde{x}| \leq |x - \bar{x}(t)| + |\tilde{x} - \bar{x}(t)| \leq \delta$. Thus by Lemma 3.1, there exists $v \in \mathbb{R}^m$ such that $c(x,v) \leq 0$ and $|v - \tilde{v}| \leq C|x - \tilde{x}|$ (note that all constants C involved for the verification of (A.3) are independent of $(t,\tilde{x},\tilde{v},x,v)$). Let $\xi = \xi(t,x,v,\tilde{z})$. We have $\xi \in F(t,x)$. It remains to bound $|\xi - \tilde{\xi}|$. We first have

$$\begin{aligned}
|\xi^{(1)} - \tilde{\xi}^{(1)}| &\leq |a(x^{(1)}) - a(\tilde{x}^{(1)})| + |b(x^{(1)})| \cdot |v - \tilde{v}| + |b(x^{(1)}) - b(\tilde{x}^{(1)})| \cdot |\tilde{v}| \\
&\leq C\big(|x - \tilde{x}| + |v - \tilde{v}|\big)(1 + |\bar{v}(t)|^2) \\
&\leq C|x - \tilde{x}|\bar{k}(t),
\end{aligned}$$

by (H2)-(ii). The same estimate can be established for $|\xi^{(2)} - \tilde{\xi}^{(2)}|$ (with the help of Assumption (H3)-(i)) and for $|\xi^{(3)} - \tilde{\xi}^{(3)}|$, thus

$$|\xi - \tilde{\xi}| \leq C|x - \tilde{x}|\bar{k}(t). \tag{A.4}$$

The inclusion (A.3) follows, taking $k(t) = C\bar{k}(t)$ and $\eta = \min\big(\delta/2, 1/C\big)$, where C is the constant appearing in the right-hand side of (A.4).

Step 3: abstract optimality conditions and interpretation. Applying [28, Theorem 7.5.1], we obtain the existence of $\bar{p} \in W^{1,1}(0,T;\mathbb{R}^{n+1+m})$ and $\lambda_0 \geq 0$ such that:

(i) $(\bar{p}, \lambda_0) \neq (0,0)$,

(ii) $-\dot{\bar{p}}(t) \in \operatorname{conv}\big\{q : (q, -\bar{p}(t)) \in N_{\operatorname{Gr}(F(t,\cdot))}(\bar{x}(t), \bar{\xi}(t))\big\}$,

(iii) $(-\bar{p}(0), \bar{p}(T)) \in \lambda_0 \nabla\Phi(\bar{x}(0), \bar{x}(T)) + N_K(\bar{x}(0), \bar{x}(T))$,

where $\operatorname{Gr}(F(t,\cdot)) = \{(x,\xi) : \xi \in F(t,x)\}$. We let the reader verify that the condition (iii) (together with Assumptions (H5)-(i) and (H5)-(ii)) implies the existence of $\lambda_1 \in \mathbb{R}^{n_{g_1}}$ and $\lambda_2 \in \mathbb{R}^{n_{g_2}}$, $\lambda_2 \geq 0$, such that

$$\begin{aligned}
\bar{p}^{(1)}(T)^\top &= \lambda_0 Dg_0(\bar{x}^{(1)}(T)) + \lambda_1^\top Dg_1(\bar{x}^{(1)}(T)) + \lambda_2^\top Dg_2(\bar{x}^{(1)}(T)), \\
\bar{p}^{(2)}(T)^\top &= \lambda_0, \\
\bar{p}^{(3)}(T)^\top &= 0.
\end{aligned}$$

and such that $\langle g_2(\bar{x}^{(1)}(T)), \lambda_2 \rangle = 0$. For the interpretation of the adjoint equation (condition (ii)), we need to examine the limiting normal cone of the graph of $F(t,\cdot)$.

Let $x \in \mathbb{R}^{n+1+m}$, let $\xi \in F(t,x)$, and let $(q,-p) \in N_{\mathrm{Gr}(F(t,\cdot))}(x,\xi)$. Let $x_k \to x$, $\xi_k = \xi(t,x_k,v_k,z_k) \to \xi$, $\xi_k \in F(t,x_k)$, $(q_k,p_k) \to (q,p)$ be such that $(q_k,-p_k) \in N^P_{\mathrm{Gr}(F(t,\cdot))}(x_k,\xi_k)$. By definition of the proximal normal cone, see (A.1), (x_k,ξ_k) is for some $C>0$ (depending on k) solution of the minimization problem

$$\operatorname*{Min}_{(x,\xi)\in \mathrm{Gr}(F)} \sum_{i=1}^{3} \left(\langle -q_k^{(i)}, x^{(i)} \rangle + \langle p_k^{(i)}, \xi^{(i)} \rangle \right) + C(|x-x_k|^2 + |\xi-\xi_k|^2).$$

In view of the expression of the multimapping F, this holds if and only if, for some (v_k,z_k), (x_k,v_k,z_k) is solution of

$$\begin{aligned} &\mathrm{Min}_{(x,v,z)\in \mathrm{Gr}(F)} \textstyle\sum_{i=1}^3 \langle -q_k^{(i)}, x^{(i)}\rangle + \langle p_k^{(1)}, a(x^{(1)}) + b(x^{(1)})v\rangle \\ &\quad + \langle p_k^{(2)}, \tilde{L}(t,x^{(1)},v) + z\rangle + \langle p_k^{(3)}, v\rangle + C(|x-x_k|^2 + |\xi-\xi_k|^2), \\ &\text{s.t. } c(x,v) \le 0 \text{ and } z \ge 0. \end{aligned}$$

Since this problem is qualified, we obtain the existence of $\nu_k \in \mathbb{R}^{n_c}$, $\nu_k \ge 0$, such that the following stationarity and complementarity conditions hold:

- Stationarity with respect to z: $p_k^{(2)} \ge 0$.
- Stationarity with respect to v:

$$(p_k^{(1)})^\top b(x_k^{(1)}) + p_k^{(2)} D_v \tilde{L}(x_k^{(1)}, v_k) + (p_k^{(3)})^\top + \nu_k^\top D_v c(x_k^{(1)}, v_k) = 0. \tag{A.5}$$

- Stationarity with respect to $x^{(1)}$:

$$\begin{aligned} -(q_k^{(1)})^\top + (p_k^{(1)})^\top \left(Da(x_k^{(1)}) + \textstyle\sum_{i=1}^m Db(x_k^{(1)}) v_{k,i} \right) \\ + p_k^{(2)} D_x \tilde{L}(x_k^{(1)}, v_k^{(1)}) + \nu_k^\top D_x c(x_k^{(1)}, v_k) = 0. \end{aligned}$$

- Stationarity with respect to $x^{(2)}$: $q_k^{(2)} = 0$.
- Stationarity with respect to $x^{(3)}$: $q_k^{(3)} = 0$.
- Complementarity: $\langle c(x_k^{(1)}, v_k), \nu_k \rangle = 0$.

The inward pointing condition, Assumption (H5)-(iv), yields a uniform bound on ν_k (with respect to k). This allows to pass to the limit in the above relations, using the continuity assumptions on $\tilde{L}$, a, b, and c (note that $\xi_k^{(3)} \to \xi^{(3)}$ implies that $v_k \to v$). We deduce that $\dot{\bar{p}}^{(2)} = 0$ and $\dot{\bar{p}}^{(3)} = 0$, thus $\bar{p}^{(2)}(t) = \lambda_0$ and $\bar{p}^{(3)}(t) = 0$. Since $\bar{p} \in W^{1,1}(0,T;\mathbb{R}^{n+1+m})$, it belongs to $L^\infty(0,T;\mathbb{R}^{n+1+m})$. Passing to the limit in (A.5) we obtain

$$\lambda_0 D_v \tilde{L}(\bar{x}^{(1)}(t), \bar{v}(t)) + (\bar{p}^{(1)}(t))^\top b(\bar{x}^{(1)}(t)) + \bar{\nu}(t)^\top D_v c(\bar{x}^{(1)}(t), \bar{v}(t)) = 0. \tag{A.6}$$

If $\lambda_0 \neq 0$, $\bar{\nu}(t) = \nu\left[\bar{x}^{(1)}(t), \lambda_0 P(t) + (\bar{p}^{(1)}(t))^\top b(\bar{x}^{(1)}(t))\right]$, by Lemma 3.2. Since $\nu[\cdot,\cdot]$ is Lipschitz continuous on bounded sets, we obtain $\bar{\nu} \in L^\infty(0,T;\mathbb{R}^{n_c})$. Analogously, $\bar{v} \in L^\infty(0,T;\mathbb{R}^m)$, therefore we deduce that $\bar{p} \in W^{1,\infty}(0,T;\mathbb{R}^{n+1+m})$.

If $\lambda_0 = 0$, we denote by $\bar{\nu}_I(t)$ the components of $\bar{\nu}(t)$ whose indices belong to the set $I(\bar{x}^{(1)}(t), \bar{v}(t))$. Skipping the time arguments, from (A.6) we deduce

$$\bar{\nu}_I = -\left[D_v c_I(\bar{x}^{(1)}, \bar{v}) D_v c_I(\bar{x}^{(1)}, \bar{v})^\top \right]^{-1} D_v c_I(\bar{x}^{(1)}, \bar{v}) b(\bar{x}^{(1)})^\top \bar{p}^{(1)}.$$

The matrix $\left[D_v c_I(\bar{x}^{(1)}, \bar{v}) D_v c_I(\bar{x}^{(1)}, \bar{v})^\top\right]$ is invertible by (H5)-(iii). Since $\bar{x}^{(1)}, \bar{p}^{(1)} \in L^\infty(0,T;\mathbb{R}^n)$, by (H2)-(i), (H3)-(ii) and (H3)-(iii) we deduce that $\bar{\nu} \in L^\infty(0,T;\mathbb{R}^{n_c})$. In this case, we only have $\bar{v} \in L^2(0,T;\mathbb{R}^m)$, so we obtain $\bar{p} \in W^{1,2}(0,T;\mathbb{R}^{n+1+m})$. ∎

Appendix B. Application to a gas storage problem. Consider the case when the scalar state $\gamma(t)$ represents a scaled energy storage, with value in $[0,1]$ and integrator dynamics

$$\dot{\gamma}(t) = v(t).$$

In addition we have control constraints

$$v_m \le v(t) \le v_M,$$

with $v_m < 0 < v_M$. Finally we have limitations on the efficiency of pumping depending on the storage level, namely

$$\varphi_1(\gamma(t)) \le v(t) \le \varphi_2(\gamma(t)),$$

with φ_1 (resp. φ_2) having negative (resp. positive) values except for $\varphi_1(0) = \varphi_2(1) = 0$, and for some $c_1 > 0$ and $c_2 > 0$:

$$-c_1 x \le \varphi_1(x); \quad \varphi_2(x) \le c_2(1-x).$$

For example, we could take $\varphi_1(x) = -c_1 x$ and $\varphi_2(x) = c_2(1-x)$. Observe that these constraints imply that the state remains between 0 and 1 (assuming of course that $\gamma(0) \in [0,1]$). Therefore we can discard the pure state constraint $\gamma(t) \in [0,1]$.

Overall we can write the constraints to be taken into account in the model as

$$\max(v_m, \varphi_1(\gamma(t))) \le v(t) \le \min(v_M, \varphi_2(\gamma(t))). \tag{B.1}$$

The two constraints in (B.1) cannot be active simultaneously, since the constraint on the l.h.s. (resp. r.h.s) has negative (resp. positive) values when $x \in (0,1)$, and they cannot be equal for $x \in \{0,1\}$. It follows that

$$\delta := \min_{x\in[0,1]} [\min(v_M, \varphi_2(x)) - \max(v_m, \varphi_1(x))]$$

is positive. While this is a model with mixed control-state constraints, our theory does not apply directly since these constraints are not smooth. Instead, consider the constraints

$$\varphi_1'(\gamma(t)) \le v(t) \le \varphi_2'(\gamma(t)),$$

with

$$\varphi_1'(x) := M_\varepsilon(v_m, \varphi_1(x)); \quad \varphi_2'(x) := -M_\varepsilon(-\varphi_2(x), -v_M),$$

where $\varepsilon > 0$ and M_ε is the smoothed maximum function given by

$$M_\varepsilon(a,b) := \tfrac{1}{2}(a+b) + \tfrac{1}{2}\sqrt{(a-b)^2 + 4\varepsilon^2}.$$

We have the error bound

$$\max(a,b) \le M_\varepsilon(a,b) \le \max(a,b) + \varepsilon.$$

So, choosing $\varepsilon \in (0, \delta/2)$, we obtain that the two bounds cannot be simultaneously active. It can also be checked that hypotheses (H3)-(i), (H3)-(ii), (H5)-(iii), (H5)-(iv) hold. So, our theory applies to this setting.

REFERENCES

[1] C. Alasseur, I. Ben Taher, and A. Matoussi. An extended mean field game for storage in smart grids. *Journal of Optimization Theory and Applications*, 184(2):644–670, 2020.

[2] L. Ambrosio, N. Gigli, and G. Savare. *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics. ETH Zürich. Birkhäuser Basel, 2008.

[3] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. In *Active Particles, Volume 1*, pages 141–171. Springer, 2017.

[4] C. Bertucci, J.-M. Lasry, and P.-L. Lions. Some remarks on mean field games. *Communications in Partial Differential Equations*, 44(3):205–227, 2019.

[5] J. F. Bonnans, X. Dupuis, and L. Pfeiffer. Second-order necessary conditions in Pontryagin form for optimal control problems. *SIAM Journal on Control and Optimization*, 52(6):3887–3916, 2014.

[6] J. F. Bonnans, S. Hadikhanloo, and L. Pfeiffer. Schauder estimates for a class of potential mean field games of controls. *Applied Mathematics & Optimization*, Jul 2019.

[7] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2000.

[8] P. Cannarsa and R. Capuani. Existence and uniqueness for mean field games with state constraints. In *PDE models for multi-agent phenomena*, pages 49–71. Springer, 2018.

[9] P. Cannarsa, R. Capuani, and P. Cardaliaguet. Mean field games with state constraints: from mild to pointwise solutions of the PDE system. *arXiv preprint arXiv:1812.11374*, 2018.

[10] P. Cannarsa, R. Capuani, and P. Cardaliaguet. $C^{1,1}$ smoothness of constrained solutions in the calculus of variations with application to mean field games. *Mathematics in Engineering*, 1(1):174–203, 2019.

[11] P. Cardaliaguet and C.-A. Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018.

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

[13] A. De Paola, D. Angeli, and G. Strbac. Distributed control of micro-storage devices with mean field games. *IEEE Transactions on Smart Grid*, 7(2):1119–1127, 2015.

[14] R. M. Dudley. *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2002.

[15] D. A. Gomes, S. Patrizi, and V. Voskanyan. On the existence of classical solutions for stationary extended mean field games. *Nonlinear Anal.: Theory, Methods and Appl.*, 99:49–79, 2014.

[16] D. A. Gomes and V. K. Voskanyan. Extended deterministic mean-field games. *SIAM Journal on Control and Optimization*, 54(2):1030–1055, 2016.

[17] J. Graber, A. Mullenix, and L. Pfeiffer. Weak solutions for potential mean field games of controls, 2020. arXiv preprint.

[18] 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(3):221–252, 2006.

[19] Z. Kobeissi. Mean field games with monotonous interactions through the law of states and controls of the agents, 2020. arXiv preprint.

[20] Z. Kobeissi. On classical solutions to the mean field game system of controls, 2020. arXiv preprint.

[21] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. i – le cas stationnaire. *Comptes Rendus Mathematique*, 343(9):619 – 625, 2006.

[22] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. ii – horizon fini et contrôle optimal. *Comptes Rendus Mathematique*, 343(10):679 – 684, 2006.

[23] J.-M. Lasry and P.-L. Lions. Mean field games. *Japanese J. of Math.*, 2(1):229–260, Mar 2007.

[24] K. Liu, N. Oudjane, and C. Wan. Approximate Nash equilibria in large nonconvex aggregative games. *arXiv preprint arXiv:2011.12604*, 2020.

[25] G. Mazanti and F. Santambrogio. Minimal-time mean field games. *Mathematical Models and Methods in Applied Sciences*, 29(08):1413–1464, 2019.

[26] D. Paccagnan, M. Kamgarpour, and J. Lygeros. On aggregative and mean field games with applications to electricity markets. In *2016 European Control Conference (ECC)*, pages 196–201. IEEE, 2016.

[27] C. Sarrazin. Lagrangian discretization of variational mean field games. *arXiv preprint arXiv:2010.11519*, 2020.

[28] R. Vinter. *Optimal control*. Springer Science & Business Media, 2010.