



Weak solutions for potential mean field games of controls

P. Jameson Graber, Alan Mullenix  and Laurent Pfeiffer

Abstract. We analyze a system of partial differential equations that model a potential mean field game of controls, briefly MFGC. Such a game describes the interaction of infinitely many negligible players competing to optimize a personal value function that depends in aggregate on the state and, most notably, control choice of all other players. A solution of the system corresponds to a Nash Equilibrium, a strategy for which no one player can improve by altering only their own action. We investigate the second order, possibly degenerate, case with non-strictly elliptic diffusion operator and local coupling function. The main result exploits potentiality to employ variational techniques to provide a unique weak solution to the system, with additional space and time regularity results under additional assumptions. New analytical subtleties occur in obtaining a priori estimates with the introduction of an additional coupling that depends on the state distribution as well as feedback.

Mathematics Subject Classification. 35Q91, 35F61, 49J20.

Keywords. Mean field games, Mean field games of controls, Calculus of variations, Optimal control, Weak solutions.

1. Introduction

Mean Field Games (MFG), introduced simultaneously in 2006–7 by Lasry and Lions [27] and Huang et al. [22], have seen swift development into a vibrant and substantial subfield of partial differential equations. See, for instance, the monographs [2, 9, 10]. Considered are high population games of homogeneous, negligibly powerful players all attempting to optimize a cost while contending with the effects of the choices of all other players.

The term Mean Field, inspired by physics, relates to each player viewing the remaining players as one large entity. The cost functional that has to be optimized by each player typically incorporates an interaction term $f(m)$,

where m denotes the distribution of player states. Mean Field Games of Controls (briefly, MFGC), also called Extended Mean Field Games, introduce a control element into the Mean Field, so that not only can players “detect” (via the Mean Field) the positions of others, but also their control choices. Such an extension naturally arises in many applications, for example in economics [11, 12, 16, 19–21, 24]. MFGC have been studied by D. Gomes and V. Voskanyan, who have results on classical solutions with S. Patrizi in the stationary (time independent) second order case where the diffusion is explicitly the Laplacian [14], and also in the time dependent first order case [15]. In the second order uniformly parabolic time dependent case, Z. Kobeissi has proved the existence of classical solutions under sufficient structural and smoothness assumptions, with uniqueness under additional assumptions, as well as results on approximate solutions [25, 26]. P. Cardaliaguet and C.-A. Lehalle have provided a theorem giving the existence of weak solutions to a general system of MFGC, under the assumption that the Lagrangian is monotone with respect to the measure variable and that the Hamiltonian is sufficiently smooth; in particular it must depend on the density of players nonlocally [8]. See also Carmona and Delarue [9] for a probabilistic approach to Extended Mean Field Games.

In this article, we investigate the second order degenerate case (which can, in particular, be first order) featuring a non-strictly elliptic diffusion operator with space dependence. The MFGC system to be studied is

$$\begin{cases} (i) & -\partial_t u - A_{ij} \partial_{ij} u + H(x, Du(x, t) + G(x)^\top P(t)) = f(x, m(x, t)) & (x, t) \in Q, \\ (ii) & \partial_t m - \partial_{ij} (A_{ij} m) + \nabla \cdot (vm) = 0 & (x, t) \in Q, \\ (iii) & P(t) = \Psi\left(\int_{\mathbb{T}^d} G(x)v(x, t)m(x, t) dx\right) & t \in [0, T], \\ (iv) & v(x, t) = -D_\xi H(x, Du(x, t) + G(x)^\top P(t)) & (x, t) \in Q, \\ (v) & m(x, 0) = m_0(x), \quad u(x, T) = u_T(x), & x \in \mathbb{T}^d \end{cases} \quad (1.1)$$

where u, m are scalar functions, v is a vector field in $\mathbb{R}^d$, $P = P(t) \in \mathbb{R}^k$, $Q := \mathbb{T}^d \times [0, T]$, and $A(x) = [A_{ij}(x)]_{1 \leq i, j \leq d}$ is a given matrix-valued function on $\mathbb{T}^d$ whose values are symmetric and non-negative.

The heuristic interpretation of the above system is the following. The variable m describes the distribution of the state of the players and P is a time-dependent price variable. These two variables are interaction terms; they describe the collective behavior of the agents. Isolated agents have no impact of them. Each agent controls the following dynamical system in $\mathbb{T}^d$:

$$dX_t = \alpha_t dt + \sqrt{2}\Sigma(X_t) dB_t$$

where $(B_t)_{t \in [0, T]}$ is a standard Brownian motion in $\mathbb{R}^D$, α is an adapted process in $\mathbb{R}^d$, and $\Sigma: \mathbb{T}^d \rightarrow \mathbb{R}^{d \times D}$ is such that $A(x) = \Sigma(x)\Sigma(x)^\top$, for all $x \in \mathbb{T}^d$. Given the interaction terms m and P , the associated cost (to be minimized) is given by

$$\mathbb{E} \left[\int_0^T \left(H^*(X_t, -\alpha_t) + \langle P(t), G(X_t)\alpha_t \rangle + f(X_t, m(X_t, t)) \right) dt + u_T(X_T) \right].$$

Equation (1.1)(i) is a Hamilton–Jacobi–Bellman equation and characterizes the value function u associated with the optimal control problem to be solved by each agent, given m and P . Equation (1.1)(iv) gives the corresponding solution in feedback form:

$$\alpha_t = v(X_t, t) = -D_p H(X_t, Du(X_t, t) + G(X_t)^\top P(t)).$$

Assuming now that all agents employ the feedback v , the evolution of their distribution is given by the Fokker–Planck equation (1.1)(ii) (the initial distribution m_0 is fixed). Then equation (1.1)(iii) gives the price in function of m and v . When $G(x)$ is the identity matrix, P is simply a function Ψ of the average control of the agents. Let us emphasize that P induces an interaction of the agents not only through their states but also via their controls. In summary, a (mean field) Nash equilibrium is attained when v is a best response with respect to the interaction terms m and P (equations (i) and (iv)), and when m and P can be deduced from v via equations (ii) and (iii).

A natural economic interpretation of P is as negative market price (the negative sign is introduced so that we have a minimization problem). In this interpretation, P increases (i.e. demand decreases) in each dimension along which the total market supply of a particular good increases. See [3] for more details. Alternatively, one could interpret System (1.1) as a model of congestion penalization with an additional dispersive forcing term given by P . For example, P may be proportional to average velocity. In this case, whereas f imposes a cost corresponding to population density, P pushes agents to move in a direction opposite to that general motion of the crowd, thus encouraging the crowd to disperse.

The basic structural assumptions are

- (1) $\mathbb{T}^d \times \mathbb{R}^d \ni (x, \xi) \mapsto H(x, \xi) \in \mathbb{R}$ is convex in ξ
- (2) $\mathbb{T}^d \times [0, \infty) \ni (x, m) \mapsto f(x, m) \in \mathbb{R}$ is monotone increasing in m
- (3) $\mathbb{R}^k \ni z \mapsto \Psi(z) \in \mathbb{R}^k$ is monotone in z , i.e. $\langle \Psi(t, z_1) - \Psi(t, z_2), z_1 - z_2 \rangle \geq 0$ for all $z_1, z_2 \in \mathbb{R}^k$.

See Sect. 1.2 for more detailed assumptions on the data.

We will focus in the article on the MFG system obtained after performing the Benamou–Brenier change of variables $w = mv$ [1]:

$$\begin{cases} (i) & -\partial_t u - A_{ij} \partial_{ij} u + H(x, Du(x, t) + G(x)^\top P(t)) = f(x, m(x, t)) & (x, t) \in Q, \\ (ii) & \partial_t m - \partial_{ij} (A_{ij} m) + \nabla \cdot w = 0 & (x, t) \in Q, \\ (iii) & P(t) = \Psi\left(\int_{\mathbb{T}^d} G(x) w(x, t) dx\right) & t \in [0, T], \\ (iv) & w(x, t) = -D_\xi H(x, Du(x, t) + G(x)^\top P(t)) m(x, t) & (x, t) \in Q, \\ (v) & m(x, 0) = m_0(x), \quad u(x, T) = u_T(x), & x \in \mathbb{T}^d. \end{cases} \quad (1.2)$$

It is worth mentioning that although the function w as defined above is determined by u , m , and P , labeling greatly reduces clutter in the statements and calculations to follow. For the same reason, we keep P , although it is determined by w .

In [4] the authors prove the existence of classical solutions to (1.1) when A is the identity matrix and the congestion term f is nonlocal. They also showed that the game is “potential,” which means that a Nash equilibrium

can be interpreted as a critical point of a suitably chosen functional, which we may call the potential. Cf. [27, Section 6.2]. When the potential is strictly convex, we formally have that the Nash equilibrium is unique, and under suitable assumptions one can show that the PDE system is well-posed. In what follows we will provide the existence and uniqueness of a suitably defined “weak solution” to the MFGC system with local coupling and provide additional regularity results involving the solution u and the distribution evolution m . The method follows that of Cardaliaguet, Graber, Porretta, and Tonon in [7]—see also [5, 6]—in the study of the case of first and second order “classical” MFG systems, which are also potential. The nonlocal interaction term $P(t)$ introduces new subtleties into the analysis, especially as it does not introduce any a priori gain of regularity. On the contrary, a priori estimates on solutions to the Hamilton–Jacobi Equation (1.2)(i) are highly sensitive to the L^p norms of $P(t)$. See Sect. 3.

We first lay out the required assumptions on the data (Sect. 1). We then view the MFGC system as a system of optimality for the minimization of a suitably defined convex potential. Next, we consider the dual problem and show that the correct relaxation of it provides existence and a.e. uniqueness of an adjoint state (Sect. 3). The solutions to these minimization problems are then shown to be proper candidates for the weak solution to the MFGC, whose existence is then proved (Sect. 4). Finally, with some additional assumptions on the data, we include some space and time regularity results for the weak solution based on previous techniques of Graber and Meřzáros [17] (Sect. 5).

We now lay out the notation and assumptions to hold throughout the paper.

1.1. Notation

We denote by $\langle x, y \rangle$ the Euclidean scalar product of two vectors $x, y \in \mathbb{R}^d$ and by $|x|$ the Euclidean norm of x . We use conventions on repeated indices: for instance, if $a, b \in \mathbb{R}^d$, we often write $a_i b_i$ for the scalar product $\langle a, b \rangle$. More generally, if A and B are two square symmetric matrices of size $d \times d$, we write $A_{ij} B_{ij}$ for $\text{Tr}(AB)$.

To avoid further difficulties arising from boundary issues, we work in the flat d –dimensional torus $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ with $Q := \mathbb{T}^d \times [0, T]$ for some fixed time $T > 0$. Our methods can be applied in a more or less straightforward way when the boundary conditions are of Neumann type; with some further technical assumptions they may be applied on the whole space. Other boundary conditions, which may be more appropriate for economics applications, tend to introduce greater technicalities; this is a subject for future research. We denote by $P(\mathbb{T}^d)$ the set of Borel probability measures over $\mathbb{T}^d$. It is endowed with the weak convergence. For $k, n \in \mathbb{N}$ and $T > 0$, we denote by $\mathcal{C}^k(Q, \mathbb{R}^n)$ the space of maps $G = G(t, x)$ of class $\mathcal{C}^k$ in time and space with values in $\mathbb{R}^n$. For $p \in [1, \infty]$ and $T > 0$, we denote by $L^p(\mathbb{T}^d)$ and $L^p(Q)$ the set of p –integrable maps over $\mathbb{T}^d$ and Q respectively. We often abbreviate $L^p(\mathbb{T}^d)$ and $L^p(Q)$ into L^p . We denote by $\|f\|_p$ the L^p –norm of a map $f \in L^p$. The conjugate of a real $p > 1$ is denoted by p' , i.e. $1/p + 1/p' = 1$.

1.2. Assumptions

We now collect the assumptions on the “congestion coupling” f , the “aggregate control coupling” Ψ , the Hamiltonian H , and the initial and terminal conditions m_0 and u_T .

Along the article, we assume that there exist some constants $C_1 > 0$, $C_2 > 0$, $C_3 > 0$, $C_4 > 0$, $q > 1$, $r > 1$, and $s > 1$ such that the following hypotheses hold true. We denote

$$p = q'.$$

(H1) (Conditions on the coupling)

- The map $f: \mathbb{T}^d \times [0, +\infty) \rightarrow \mathbb{R}$ is continuous in both variables, increasing with respect to the second variable m , and satisfies

$$\frac{1}{C_1}|m|^{q-1} - C_1 \leq f(x, m) \leq C_1|m|^{q-1} + C_1 \quad \forall m \geq 0. \quad (1.3)$$

Moreover $f(x, 0) = 0$ for all $x \in \mathbb{T}^d$.

- The map $\Psi: \mathbb{R}^k \rightarrow \mathbb{R}^k$ is the continuous gradient of some convex function $\Phi: \mathbb{R}^k \rightarrow \mathbb{R}$. Without loss of generality, we assume that $\Phi(0) = 0$. Moreover,

$$\Phi(z) \leq C_2|z|^s + C_2 \quad \forall z \in \mathbb{R}^k. \quad (1.4)$$

Changing C_2 if necessary, we have

$$\Phi^*(z) \geq \frac{1}{C_2}|z|^{s'} - C_2 \quad \forall z \in \mathbb{R}^k.$$

If $\frac{1}{s} + \frac{1}{pr} < 1$, we assume that

$$\frac{1}{C_2}|z|^s - C_2 \leq \Phi(z) \quad \forall z \in \mathbb{R}^k. \quad (1.5)$$

- The map $G: \mathbb{T}^d \rightarrow \mathcal{L}(\mathbb{R}^d; \mathbb{R}^k)$ is continuously differentiable. If $\frac{1}{s} + \frac{1}{pr} < 1$, we assume that it is constant.

(H2) (Conditions on the Hamiltonian) The Hamiltonian $H: \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is continuous in both variables, convex and differentiable in the second variable, with $D_\xi H$ continuous in both variables, and has a superlinear growth in the gradient variable:

$$\frac{1}{C_3}|\xi|^r - C_3 \leq H(x, \xi) \leq C_3|\xi|^r + C_3 \quad \forall (x, \xi) \in \mathbb{T}^d \times \mathbb{R}^d. \quad (1.6)$$

We note for later use that the Fenchel conjugate H^* of H with respect to the second variable, given by

$$H^*(x, \zeta) = \sup_{\gamma} [\langle \zeta, \gamma \rangle - H(x, \gamma)],$$

is continuous and satisfies similar inequalities to H (changing C_3 if necessary):

$$\frac{1}{C_3}|\xi|^{r'} - C_3 \leq H^*(x, \xi) \leq C_3|\xi|^{r'} + C_3 \quad \forall (x, \xi) \in \mathbb{T}^d \times \mathbb{R}^d. \quad (1.7)$$

- (H3) (Conditions on A) There exists a Lipschitz continuous map $\Sigma : \mathbb{T}^d \rightarrow \mathbb{R}^{d \times D}$ such that $\Sigma \Sigma^T = A$ and such that

$$|\Sigma(x) - \Sigma(y)| \leq C_4 |x - y| \quad \forall x, y \in \mathbb{T}^d. \quad (1.8)$$

- (H4) (Conditions on the initial and terminal conditions) $u_T : \mathbb{T}^d \rightarrow \mathbb{R}$ is of class $\mathcal{C}^2$, while $m_0 : \mathbb{T}^d \rightarrow \mathbb{R}$ is a C^1 positive density (namely $m_0(x) > 0$ and $\int_{\mathbb{T}^d} m_0 dx = 1$).
- (H5) (Restrictions on the exponents). We consider 4 cases, depending on whether $s' < r$ or $s' \geq r$ and whether A is constant or not.

	Case 1: $s' < r$	Case 2: $s' \geq r$
Case A: A is not constant	$s' \geq p$	$r \geq p$
Case B: A is constant	$\frac{s'(d+1)}{d} \geq p$	$[s' \geq 1+d] \text{ or } [s' < 1+d \text{ and } \frac{s'(1+d)}{d-s'+1} > p]$

Remark 1.1. (1) The condition $f(x, 0) = 0$ is just a normalization condition, which we may assume without loss of generality, as explained in [7, Section 2].

- (2) Let us compare the different cases of Assumption (H5).
- (a) Assumption (H5) is stronger in cases 1A and 2A than in cases 1B and 2B, respectively, that is, Assumption (H5) is stronger in the case of a non-constant A than in the constant case.
 - (b) If A is not constant (cases 1A and 2A), then Assumption (H5) can be summarized by $\min(s', r) \geq p$.
 - (c) If A is constant, it is easy to verify that Assumption (H5) is stronger in the case 1B ($s' < r$) than in case 2B ($s' \geq r$).
- (3) If $\Psi = 0$, then we are back to the framework of [7] and our assumptions coincide. Indeed, (1.4) is then satisfied with any $s > 1$. Taking s sufficiently close to 1, we have $1/s + 1/(rp) \geq 1$, so that (1.5) is not necessary, and we have $s' \geq r$, so that we are either in case 2A or 2B in hypothesis (H5). If A is constant, we must choose s close enough to 1, so that $s' \geq 1 + d$.

Let us set

$$F(x, m) = \begin{cases} \int_0^m f(x, \tau) d\tau & \text{if } m \geq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Then F is continuous on $\mathbb{T}^d \times (0, +\infty)$, differentiable and strictly convex in m and satisfies

$$\frac{1}{C_1} |m|^q - C_1 \leq F(x, m) \leq C_1 |m|^q + C_1 \quad \forall m \geq 0, \quad (1.9)$$

changing C_1 if necessary. Let F^* be the Fenchel conjugate of F with respect to the second variable. Note that $F^*(x, a) = 0$ for $a \leq 0$ because $F(x, m)$ is nonnegative, equal to $+\infty$ for $m < 0$, and null at zero. Moreover,

$$\frac{1}{C_1}|a|^p - C_1 \leq F^*(x, a) \leq C_1|a|^p + C_1 \quad \forall a \geq 0, \quad (1.10)$$

changing again C_1 if necessary.

Remark 1.2. Most of the results in this paper hold also for time-dependent data, in particular when f and H depend on t . It suffices to have the estimates in this subsection hold uniformly with respect to t .

2. Two problems in duality

The approach that we follow consists in viewing the MFG system as an optimality condition for two convex problems, which we introduce now.

Let $\mathcal{K}_0$ be the set of all triples $(u, P, \gamma) \in \mathcal{C}^2(Q) \times \mathcal{C}^0([0, T]; \mathbb{R}^k) \times \mathcal{C}^0(Q)$ satisfying

$$\begin{cases} -\partial_t u - A_{ij} \partial_{ij} u + H(x, Du(x, t) + G(x)^\top P(t)) = \gamma, \\ u(x, T) = u_T(x). \end{cases} \quad (2.1)$$

The associated cost is given by

$$D(u, P, \gamma) = - \int_{\mathbb{T}^d} u(x, 0) m_0(x) dx + \int_0^T \Phi^*(P(t)) dt + \iint_Q F^*(x, \gamma(x, t)) dx dt. \quad (2.2)$$

The first problem is

$$\inf_{(u, P, \gamma) \in \mathcal{K}_0} D(u, P, \gamma). \quad (2.3)$$

We consider now the set $\mathcal{K}_1$ of all pairs $(m, w) \in L^1(Q) \times L^1(Q; \mathbb{R}^d)$ such that $m \geq 0$ a.e., $\int_{\mathbb{T}^d} m(x, t) dx = 1$ for a.e. $t \in (0, T)$, and such that the continuity equation

$$\begin{cases} \partial_t m - \partial_{ij} (A_{ij} m) + \nabla \cdot w = 0, \\ m(x, 0) = m_0(x) \end{cases} \quad (2.4)$$

holds in the sense of distributions. For $(m, w) \in \mathcal{K}_1$, consider the following cost functional:

$$\begin{aligned} B(m, w) = & \iint_Q \left(H^* \left(x, -\frac{w(x, t)}{m(x, t)} \right) m(x, t) + F(x, m(x, t)) \right) dx dt \\ & + \int_0^T \Phi \left(\int_{\mathbb{T}^d} G(x) w(x, t) dx \right) dt + \int_{\mathbb{T}^d} u_T(x) m(x, T) dx, \end{aligned} \quad (2.5)$$

where for $m(t, x) = 0$, we impose that

$$m(t, x) H^* \left(x, -\frac{w(t, x)}{m(t, x)} \right) = \begin{cases} +\infty & \text{if } w(t, x) \neq 0 \\ 0 & \text{if } w(t, x) = 0. \end{cases}$$

Since H^* and F are bounded from below and $m \geq 0$, the first integral in B is well defined in $\mathbb{R} \cup \{+\infty\}$. Concerning the second term in B , we simply need

to observe that since Φ is convex, there exists a constant $C > 0$ such that $\Phi(z) \geq -C(1 + |z|)$, for all $z \in \mathbb{R}^k$. For $w \in L^1(Q; \mathbb{R}^d)$, the term

$$\int_{\mathbb{T}^d} G(x)w(x, \cdot) dx$$

is integrable in time and therefore

$$\int_0^T \Phi \left(\int_{\mathbb{T}^d} G(x)w(x, t) dx \right) dt$$

is well-defined in $\mathbb{R} \cup \{+\infty\}$. For the third term, we refer the reader to [7, Lemma 4.1], where it is proved that for $(m, w) \in \mathcal{K}_1$, m can be seen as a continuous map from $[0, T]$ to $P(\mathbb{T}^d)$ for the Rubinstein-Kantorovich distance $\mathbf{d}_1$. Finally, the second optimization problem is the following:

$$\inf_{(m, w) \in \mathcal{K}_1} B(m, w). \quad (2.6)$$

Lemma 2.1. *We have*

$$\inf_{(u, P, \gamma) \in \mathcal{K}_0} D(u, P, \gamma) = - \min_{(m, w) \in \mathcal{K}_1} B(m, w).$$

Moreover, the minimum in the right-hand side is achieved by a unique pair $(m, w) \in \mathcal{K}_1$ satisfying

$$m \in L^q(Q), \quad w \in L^{\frac{r'q}{r'+q-1}}(Q; \mathbb{R}^d), \quad \text{and} \quad \int_{\mathbb{T}^d} G(x)w(x, \cdot) dx \in L^s((0, T); \mathbb{R}^k). \quad (2.7)$$

Proof. Following previous papers [5–7], we look to apply the Fenchel–Rockafellar duality theorem. In order to do so, we reformulate the first optimization problem into a more suitable form.

Let $E_0 = \mathcal{C}^2(Q) \times \mathcal{C}^0([0, T], \mathbb{R}^k)$ and $E_1 = \mathcal{C}^0(Q) \times \mathcal{C}^0(Q; \mathbb{R}^d)$. Define on E_0 the functional

$$\mathcal{F}(u, P) = \int_0^T \Phi^*(P(t)) dt - \int_{\mathbb{T}^d} u(0, x)m_0(x) dx + \chi_S(u),$$

where χ_S is the convex characteristic function of the set $S = \{u \in E_0, u(T, \cdot) = u_T\}$, i.e., $\chi_S(u) = 0$ if $u \in S$ and $+\infty$ otherwise. For $(a, b) \in E_1$, we define

$$\mathcal{G}(a, b) = \iint_Q F^*(x, -a + H(x, b)) dx dt.$$

The functional $\mathcal{F}$ is convex and lower semi-continuous on E_0 while $\mathcal{G}$ is convex and continuous on E_1 . Let $\Lambda : E_0 \rightarrow E_1$ be the bounded linear operator defined by $\Lambda(u, P) = (\partial_t u + A_{ij} \partial_{ij} u, Du + G^\top P)$. We can observe that

$$\inf_{(u, P, \gamma) \in \mathcal{K}_0} D(u, P, \gamma) = \inf_{(u, P) \in E_0} \{\mathcal{F}(u, P) + \mathcal{G}(\Lambda(u, P))\}.$$

In the interest of employing the Fenchel–Rockafellar duality theorem, note that $\mathcal{F}(u, P) < +\infty$ for $(u_T, 0)$ and $\mathcal{G}(\Lambda(u, P))$ is continuous at $(u_T, 0)$. This satisfies the duality theorem, and so

$$\inf_{(u, P) \in E_0} \{\mathcal{F}(u, P) + \mathcal{G}(\Lambda(u, P))\} = \max_{(m, w) \in E'_1} \{-\mathcal{F}^*(\Lambda^*(m, w)) - \mathcal{G}^*(-(m, w))\},$$

where E'_1 is the dual space of E_1 , i.e., the set of vector valued Radon measures (m, w) over Q with values in $\mathbb{R} \times \mathbb{R}^d$, E'_0 is the dual space of E_0 , $\Lambda^* : E'_1 \rightarrow E'_0$ is the dual operator of Λ and $\mathcal{F}^*$ and $\mathcal{G}^*$ are the convex conjugates of $\mathcal{F}$ and $\mathcal{G}$ respectively. We now compute the relevant conjugate transforms.

$$\begin{aligned}
\mathcal{F}^*(\Lambda^*(m, w)) &= \sup_{(u, P)} \left\{ \langle (u, P), \Lambda^*(m, w) \rangle - \mathcal{F}(u, P) \right\} \\
&= \sup_{\substack{(u, P) \\ u \in S}} \left\{ \langle \Lambda(u, P), (m, w) \rangle - \int_0^T \Phi^*(P(t)) dt + \int_{\mathbb{T}^d} u(0, x) m_0(x) dx \right\} \\
&= \sup_{\substack{(u, P) \\ u \in S}} \left\{ \langle \partial_t u + A_{ij} \partial_{ij} u, m \rangle + \langle Du, w \rangle + \langle G^\top P, w \rangle \right. \\
&\quad \left. - \int_0^T \Phi^*(P(t)) dt + \int_{\mathbb{T}^d} u(0, x) m_0(x) dx \right\} \\
&= \sup_{\substack{(u, P) \\ u \in S}} \left\{ \langle -\partial_t m + \partial_{ij} (A_{ij} m) - \nabla \cdot w, u \rangle + \langle G^\top P, w \rangle \right. \\
&\quad \left. - \int_0^T \Phi^*(P(t)) dt + \int_{\mathbb{T}^d} u_T(x) m(T, x) dx \right\}.
\end{aligned}$$

It is evident here that if $-\partial_t m + \partial_{ij} (A_{ij} m) - \nabla \cdot w \neq 0$ in the sense of distributions, this supremum is infinite. If the condition does hold however, the supremum no longer depends on u , and so the calculation is reduced to

$$\begin{aligned}
&\sup_P \left\{ \int_0^T \left(\int_{\mathbb{T}^d} \langle G(x)^\top P(t), w(x, t) \rangle dx \right) - \Phi^*(P(t)) dt \right\} \\
&= \sup_P \left\{ \int_0^T \left\langle P(t), \int_{\mathbb{T}^d} G(x) w(x, t) dx \right\rangle - \Phi^*(P(t)) dt \right\} \\
&= \sup_P \left\{ \int_0^T \Phi \left(\int_{\mathbb{T}^d} G(x) w(x, t) dx \right) dt \right\}.
\end{aligned}$$

Combined with the conditions above, we have

$$\begin{aligned}
&\mathcal{F}^*(\Lambda^*(m, w)) \\
&= \begin{cases} \int_0^T \Phi \left(\int_{\mathbb{T}^d} G w \right) + \int_{\mathbb{T}^d} u_T(x) m(T, x) dx, & \text{if } -\partial_t m + A_{ij} \partial_{ij} m - \nabla \cdot w = 0, \\ +\infty & \text{otherwise.} \end{cases}
\end{aligned}$$

Following [5], we have that $\mathcal{G}^*(m, w) = +\infty$ if $(m, w) \notin L^1(Q) \times L^1(Q; \mathbb{R}^d)$ and

$$\mathcal{G}^*(m, w) = \iint_Q K^*(x, m(t, x), w(t, x)) dx dt$$

otherwise, where K^* is given by

$$K^*(x, m, w) = \begin{cases} F(x, -m) - mH^*\left(x, -\frac{w}{m}\right) & \text{if } m < 0 \\ 0 & \text{if } m = 0 \text{ and } w = 0 \\ +\infty & \text{otherwise,} \end{cases}$$

it is the convex conjugate of

$$K(x, a, b) = F^*(x, -a + H(x, b)) \quad \forall (x, a, b) \in \mathbb{T}^d \times \mathbb{R} \times \mathbb{R}^d.$$

Therefore

$$\begin{aligned} & \max_{(m, w) \in E'_1} \left\{ -\mathcal{F}^*(\Lambda^*(m, w)) - \mathcal{G}^*(-(m, w)) \right\} \\ &= \max \left\{ \iint_Q -F(x, m) - mH^*\left(x, -\frac{w}{m}\right) dx dt \right. \\ & \quad \left. - \int_0^T \Phi\left(\int_{\mathbb{T}^d} Gw\right) dt - \int_{\mathbb{T}^d} u_T(x)m(T, x) dx \right\} \end{aligned}$$

with the last maximum taken over the L^1 maps (m, w) such that $m \geq 0$ a.e. and

$$-\partial_t m + \partial_{ij}(A_{ij}m) - \nabla \cdot w = 0, \quad m(0) = m_0$$

holds in the sense of distributions. Since $\int_{\mathbb{T}^d} m_0 = 1$, it follows that $\int_{\mathbb{T}^d} m(t) = 1$ for any $t \in [0, T]$. Thus the pair (m, w) belongs to the set $\mathcal{K}_1$ and the first part of the statement is proved.

It remains to show (2.7). Taking an optimal $(m, w) \in \mathcal{K}_1$ in the above problem, we have that $w(t, x) = 0$ for all $(t, x) \in [0, T] \times \mathbb{T}^d$ whenever $m(t, x) = 0$. By convexity of Φ , we have

$$\int_0^T \Phi\left(\int_{\mathbb{T}^d} Gw dx\right) dt \geq \int_0^T \Phi(0) + \langle \Psi(0), \int_{\mathbb{T}^d} Gw dx \rangle dt \geq C - C\|w\|_1. \quad (2.8)$$

Moreover, by Hölder's inequality,

$$\|w\|_1 = \iint_Q \frac{|w|}{m} m dx dt \leq \left(\iint_Q \left(\frac{|w|}{m} \right)^{r'} m \right)^{1/r'}.$$

It follows with Young's inequality that

$$\|w\|_1 \leq \frac{\varepsilon}{r'} \iint_Q |w|^{r'} m^{1-r'} dx dt + \varepsilon^{-(r-1)},$$

for any $\varepsilon > 0$. Therefore,

$$\int_0^T \Phi\left(\int_{\mathbb{T}^d} Gw dx\right) dt \geq C - \frac{\varepsilon C}{r'} \iint_Q |w|^{r'} m^{1-r'} dx dt - \varepsilon^{-(r-1)}.$$

Using successively the optimality of (m, w) , the growth conditions on F and H^* and the above inequality, we obtain

$$\begin{aligned} C &\geq \iint_Q F(x, m) + mH^*\left(x, -\frac{w}{m}\right) dx dt + \int_0^T \Phi(\int_{\mathbb{T}^d} Gw) dt + \int_{\mathbb{T}^d} u_T(x)m(T, x) dx \\ &\geq \frac{1}{C} \|m\|_q^q + \left(\frac{1}{C} - \varepsilon C\right) \iint_Q |w|^{r'} m^{1-r'} dx dt - \varepsilon^{-(r-1)} - C, \end{aligned}$$

for some constant $C > 0$ independent of ε . Choosing $\varepsilon > 0$ sufficiently small, we deduce that $m \in L^q(Q)$ and that $|w|^{r'} m^{1-r'} \in L^1(Q)$. To investigate the claim on w given in the statement, write, for some ρ to be determined,

$$\|w\|_{\rho r'}^{r'} = \left\| m^{r'-1} \frac{|w|^{r'}}{m^{r'-1}} \right\|_{\rho} \leq \|m^{r'-1}\|_{\frac{q}{r'-1}} \left\| \frac{|w|^{r'}}{m^{r'-1}} \right\|_1 = \|m\|_q^{r'-1} \left\| \frac{|w|^{r'}}{m^{r'-1}} \right\|_1 < \infty.$$

For this interpolative Hölder inequality to hold, we must have $\rho = \frac{q}{r'+q-1}$. Thus $w \in L^\sigma(Q; \mathbb{R}^d)$, with

$$\sigma = \frac{r'q}{r' + q - 1}, \quad \text{i.e. } \sigma' = rp. \quad (2.9)$$

Two cases must be considered. If $\frac{1}{s} \geq 1 - \frac{1}{rp} = \frac{1}{\sigma}$, then we have $\sigma \geq s$, thus $w \in L^s(Q; \mathbb{R}^d)$ and (2.7) follows. In the other case, we have by Hypothesis (H1) the growth assumption $\Phi(z) \geq \frac{1}{C}|z|^s - C$. It can be employed to get a better bound from below in (2.8). We obtain (2.7) with a straightforward adaptation of the above proof. $\square$

3. Optimal control problem of the HJ equation

In general we do not expect problem (2.3) to have a solution. In this section we exhibit a relaxation for (2.3) (Proposition 3.8) and show that the obtained relaxed problem has at least one solution (Proposition 3.10).

3.1. Estimates on subsolutions to HJ equations

In this subsection we prove estimates in Lebesgue spaces for subsolutions of Hamilton–Jacobi equations of the form

$$\begin{cases} (i) & -\partial_t u - A_{ij} \partial_{ij} u + H(Du + G^\top P) \leq \gamma \\ (ii) & u(x, T) \leq u_T(x) \end{cases} \quad (3.1)$$

in terms of Lebesgue norms of γ , u_T , and P . Equation (3.1) is understood in the sense of distributions. This means that $Du + G^\top P \in L^r(Q)$ and for any nonnegative test function $\zeta \in C_c^\infty((0, T] \times \mathbb{T}^d)$,

$$-\int_{\mathbb{T}^d} \zeta(T) u_T + \iint_Q u \partial_t \zeta + \langle D\zeta, ADu \rangle + \zeta (\partial_i A_{ij} \partial_j u + H(Du + G^\top P)) \leq \iint_Q \gamma \zeta. \quad (3.2)$$

Let us introduce some notation. For all $\tilde{r} > 1$ and for all $\tilde{p} \geq 1$, let us define $\bar{\kappa}(\tilde{r}, \tilde{p})$ and $\bar{\eta}(\tilde{r}, \tilde{p})$ by

$$\bar{\eta}(\tilde{r}, \tilde{p}) = \frac{d(\tilde{r}(\tilde{p} - 1) + 1)}{d - \tilde{r}(\tilde{p} - 1)} \quad \text{and} \quad \bar{\kappa}(\tilde{r}, \tilde{p}) = \frac{\tilde{r}\tilde{p}(1 + d)}{d - \tilde{r}(\tilde{p} - 1)}$$

if $\tilde{p} < 1 + \frac{d}{\tilde{r}}$ and

$$\bar{\eta}(\tilde{r}, \tilde{p}) = \infty \quad \text{and} \quad \bar{\kappa}(\tilde{r}, \tilde{p}) = \infty$$

if $\tilde{p} > 1 + \frac{d}{\tilde{r}}$. In the border line case $\tilde{p} = 1 + \frac{d}{\tilde{r}}$, $\bar{\eta}(\tilde{r}, \tilde{p})$ and $\bar{\kappa}(\tilde{r}, \tilde{p})$ can be fixed to arbitrarily large values. We let the reader verify that

$$\bar{\kappa}(\tilde{r}, \tilde{p}) \geq \tilde{r} \quad \text{and} \quad \bar{\kappa}(\tilde{r}, \tilde{p}) \geq \tilde{p}, \quad (3.3)$$

assuming that the assigned value to $\bar{\kappa}(\tilde{r}, \tilde{p})$ is large enough in the border line case.

We now restate [7, Theorem 3.3], since it will prove useful below.

Theorem 3.1. *Let u satisfy*

$$\begin{cases} (i) & -\partial_t u - A_{ij} \partial_{ij} u + \frac{1}{K} |Du|^{\tilde{r}} \leq \gamma \\ (ii) & u(x, T) \leq u_T(x) \end{cases} \quad (3.4)$$

in the sense of distributions, with $\gamma \in L^{\tilde{p}}(Q)$ for some $\tilde{p} \geq 1$ and $\tilde{r} > 1$. Then

$$\|u_+\|_{L^\infty((0,T),L^\eta(\mathbb{T}))} + \|u_+\|_{L^\kappa(Q)} \leq C,$$

where $u_+ := \max\{u, 0\}$, $\kappa = \bar{\kappa}(\tilde{r}, \tilde{p})$, $\eta = \bar{\eta}(\tilde{r}, \tilde{p})$. The constant C depends only on $T, d, \tilde{r}, \tilde{p}$, K (appearing in (3.4)), C_4 (appearing in Hypothesis (H3)), $\|\gamma\|_{\tilde{p}}$, and $\|u_T\|_{\eta}$.

Remark 3.2. Although the case $\tilde{p} = 1$ is not explicitly mentioned in [7, Theorem 3.3], it is not hard to check that the theorem also applies in that case.

Corollary 3.3. *Let $P \in C^0([0, T]; \mathbb{R}^k)$ and $\gamma \in C^0(Q; \mathbb{R})$. Let u be the viscosity solution to the HJ equation*

$$\begin{cases} -\partial_t u - A_{ij} \partial_{ij} u + H(x, Du(x, t) + G(x, t)^\top P(t)) = \gamma, \\ u(x, T) = u_T(x). \end{cases} \quad (3.5)$$

Let $\tilde{r} > 1$. Define

$$\kappa = \bar{\kappa}(\tilde{r}, 1) = \frac{\tilde{r}(1 + d)}{d}. \quad (3.6)$$

Then

$$\|u\|_{L^\infty((0,T),L^1(\mathbb{T}))} + \|u\|_{L^\kappa(Q)} \leq C,$$

where the constant C depends only on $T, d, \tilde{r}, \tilde{p}$, C_4 (appearing in Hypothesis (H3)), $\|\gamma\|_1$, $\|Du\|_{\tilde{r}}$, $\|H(Du + G^\top P)\|_1$, and $\|u_T\|_1$.

Proof. By [23], u also satisfies the HJ equation in the sense of distributions. Observe that (3.5) can be rewritten

$$-\partial_t u - A_{ij} \partial_{ij} u + |Du|^{\tilde{r}} = \gamma - H(Du + G^\top P) + |Du|^{\tilde{r}}. \quad (3.7)$$

The L^1 norm of the right-hand side depends on $\|\gamma\|_1$, $\|H(Du + G^\top P)\|_1$, and $\|Du\|_{\tilde{r}}$. Similarly, $-u$ is a weak subsolution of a HJ equation with right-hand side

$$-\gamma + H(Du + G^\top P) + |Du|^{\tilde{r}}.$$

By applying Theorem 3.1 to both u and $-u$, we deduce the desired estimate. $\square$

When $s' \geq r$, the growth assumption on the Hamiltonian (Hypothesis (H2)) can be exploited to derive a more precise estimate on the solution to (3.5).

Corollary 3.4. *Let P , u , and γ be as in Corollary 3.3. Assume moreover that $\gamma \geq 0$ and $s' \geq r$. Take $\tilde{r} \in (1, r]$ and define*

$$\tilde{p} = \min\left(p, \frac{s'}{\tilde{r}}\right), \quad \kappa = \bar{\kappa}(\tilde{r}, \tilde{p}), \quad \text{and} \quad \eta = \bar{\eta}(\tilde{r}, \tilde{p}).$$

Then

$$\|u\|_{L^\infty((0,T), L^\eta(\mathbb{T}^d))} + \|u\|_{L^\kappa(Q)} \leq C.$$

The constant C depends only on $T, d, \tilde{r}, \tilde{p}$, C_3 (appearing in Hypothesis (H2)), C_4 (appearing in Hypothesis (H3)), $\|\gamma\|_p$, $\|P\|_{s'}$, and $\|u_T\|_\eta$.

Proof. We have $\gamma \geq 0$ and the upper bound

$$H(Du + G^\top P) \leq C|Du|^r + C|P|^r,$$

therefore,

$$-\partial_t u - A_{ij}\partial_{ij}u + C|Du|^r \geq -C|P|^r - C, \quad u(T, x) \geq \left(\min_{x' \in \mathbb{T}^d} u_T(x')\right). \quad (3.8)$$

Let $\hat{u}$ be defined by

$$\hat{u}(x, t) = \left(\min_{x' \in \mathbb{T}^d} u_T(x')\right) - \int_t^T |P(t)|^r dt - C(T - t),$$

which solves (3.8) with inequality replaced by equality. By the comparison principle, $u \geq \hat{u} \geq -C$, where C depends only on T , the growth of H , $\min u_T$, and $\|P\|_r$. Note that $\|P\|_r$ depends only on $\|P\|_{s'}$ and T because $s' \geq r$.

Next, by the growth condition of H we have

$$-\partial_t u - A_{ij}\partial_{ij}u + \frac{1}{C_3}|Du + G^\top P|^r - C_3 \leq \gamma.$$

Observe that by Young's inequality,

$$|Du|^{\tilde{r}} \leq 2^{\tilde{r}-1}|Du + G^\top P|^{\tilde{r}} + 2^{\tilde{r}-1}|G^\top P|^{\tilde{r}} \leq 2^{\tilde{r}-1}\left(\frac{\tilde{r}}{r}|Du + G^\top P|^r + 1\right) + C|P|^{\tilde{r}},$$

since $r \geq \tilde{r}$. It follows that

$$|Du + G^\top P|^r \geq \frac{1}{C}|Du|^{\tilde{r}} - C|P|^{\tilde{r}} - C,$$

therefore,

$$-\partial_t u - A_{ij}\partial_{ij}u + \frac{1}{C}|Du|^{\tilde{r}} \leq \gamma + C|P|^{\tilde{r}} + C. \quad (3.9)$$

Since $|P|^{\tilde{r}}$ lies in $L^{s'/\tilde{r}}(Q)$, we have that the right hand side of (3.9) is bounded in $L^{\tilde{p}}$. Combining this with the lower bound on u , the conclusion follows from Theorem 3.1. $\square$

We can now fix the values of the coefficients $\tilde{r} \in (1, r]$, $\kappa > 1$, and $\eta > 1$ to be employed in the sequel, consistently with Corollary 3.3 (if $s' < r$) and Corollary 3.4 (if $s' \geq r$). As will appear later in the proofs of Lemma 3.9 and Proposition 3.10, these coefficients must satisfy the following:

$$[s' \geq \tilde{r}], \quad [\kappa \geq p], \quad \text{and} \quad [A \text{ is not constant} \implies \tilde{r} \geq p].$$

This is the reason why four subcases have been introduced in Hypothesis (H5) and why we have a specific definition of the coefficients for each of the subcases. In order to deal with the case 2B, we need the following lemma.

Lemma 3.5. *Assume that $s' \geq r$ and A is constant, that is, consider the case 2B of Assumption (H5). Then the corresponding condition:*

$$\left[s' \geq 1 + d \right] \quad \text{or} \quad \left[s' < 1 + d \quad \text{and} \quad \frac{s'(1+d)}{d-s'+1} > p \right] \quad (3.10)$$

is satisfied if and only if there exists $\tilde{r} \in (1, r]$ such that

$$\bar{\kappa}\left(\tilde{r}, \min\left(p, \frac{s'}{\tilde{r}}\right)\right) \geq p. \quad (3.11)$$

Proof. Several cases must be distinguished.

- Case (i): $s' > p$. In that case, either $s' \geq 1 + d$ or $s' < 1 + d$ and then

$$\frac{s'(1+d)}{d-s'+1} > s' > p.$$

Thus, if $s' > p$, then (3.10) holds true. Then we can set $\tilde{r} = \frac{s'}{p} > 1$. We have $\min\left(p, \frac{s'}{\tilde{r}}\right) = p$ and therefore $\kappa = \bar{\kappa}(\tilde{r}, p) \geq p$, by inequality (3.3).

- Case (ii): $s' \leq p$. Then whatever the choice of $\tilde{r} \in (1, r]$, we have

$$\tilde{p}(\tilde{r}) := \min\left(p, \frac{s'}{\tilde{r}}\right) = \frac{s'}{\tilde{r}}.$$

- Case (iia): $s' > 1 + d$. Then we can chose $\tilde{r}$ sufficiently close to 1, so that

$$\frac{s' - d}{\tilde{r}} > 1.$$

Then we have $\tilde{p}(\tilde{r}) > 1 + \frac{d}{\tilde{r}}$, thus $\kappa = \bar{\kappa}(\tilde{r}, \tilde{p}(\tilde{r})) = \infty$.

- Case (iib): $s' \leq 1 + d$. Then whatever the choice of $\tilde{r} \in (1, r]$, we have $\tilde{p}(\tilde{r}) < 1 + \frac{d}{\tilde{r}}$ and therefore, condition (3.11) is equivalent to:

$$\exists \tilde{r} \in (1, r], \quad \frac{s'(d+1)}{d+\tilde{r}-s'} \geq p.$$

The above condition is clearly satisfied if and only if either $s' = 1 + d$ or $s' < 1 + d$ and $\frac{s'(d+1)}{d+1-s'} > p$.

□

We can finally fix $\tilde{r}$, κ , and η .

- In cases 1A and 1B (i.e. $s' < r$), we set

$$\tilde{r} = s', \quad \kappa = \bar{\kappa}(\tilde{r}, 1), \quad \eta = \bar{\eta}(\tilde{r}, 1).$$

Then we have $\kappa \geq \tilde{r} \geq p$. In case 1A, $\tilde{r} \geq p$.

- In case 2A (i.e. $s' \geq r$ and A is not constant), we set $\tilde{r} = p$. In case 2B (i.e. $s' \geq r$ and A is constant), we assign a value to $\tilde{r}$ so that (3.11) holds true. In both cases 2A and 2B, we set

$$\kappa = \bar{\kappa}\left(\tilde{r}, \min\left(p, \frac{s'}{\tilde{r}}\right)\right) \quad \text{and} \quad \eta = \bar{\eta}\left(\tilde{r}, \min\left(p, \frac{s'}{\tilde{r}}\right)\right).$$

In case 2A, we have $\kappa \geq \tilde{r} = p$ by inequality (3.3). In case 2B, we have $\kappa \geq p$ by definition.

Remark 3.6. In case 2B, it is easy to deduce from the proof of Lemma 3.5 an explicit $\tilde{r} \in (1, r]$ such that (3.11) holds. Note that the obtained $\tilde{r}$ may not be the best one (i.e. the largest one). For example, if $s' \geq pr$, then one can take $\tilde{r} = r$. Then $\frac{s'}{\tilde{r}} \geq p$ and therefore $\kappa = \bar{\kappa}(\tilde{r}, p) \geq p$, by inequality (3.3).

3.2. The relaxed problem

We propose in this subsection an appropriate relaxation of problem (2.3). Let $\mathcal{K}$ denote the set of triplets $(u, P, \gamma) \in L^\kappa(Q) \times L^{s'}(0, T) \times L^p(Q)$ such that $Du + G^\top P \in L^r(Q; \mathbb{R}^d)$, $Du \in L^{\tilde{r}}(Q; \mathbb{R}^d)$, and such that (3.1) holds in the sense of distributions.

The following statement explains that u has a “trace” in a weak sense.

Lemma 3.7. *Let $f \in L^1(Q)$ and let $u \in L^1(Q)$ satisfy $Du \in L^{\tilde{r}}(Q; \mathbb{R}^d)$ and*

$$-\partial_t u - A_{ij}\partial_{ij}u \leq f, \quad u(T) \leq u_T \tag{3.12}$$

in the sense of distributions, i.e. for every non-negative function $\vartheta \in C_c^\infty(\mathbb{T}^d \times (0, T])$ we have

$$\iint_Q (u\partial_t \vartheta + \partial_j(A_{ij}\vartheta)\partial_i u - f\vartheta) \, dx \, dt \leq \int_{\mathbb{T}^d} \vartheta(x, T)u_T(x) \, dx.$$

Then, for any C^1 map $\vartheta: [0, T] \times \mathbb{T}^d \rightarrow \mathbb{R}$, the function

$$t \mapsto \int_{\mathbb{T}} \vartheta(x, t)u(x, t)dx$$

has a BV representative on $[0, T]$. In particular, for any nonnegative C^1 map $\vartheta: [0, T] \times \mathbb{T}^d \rightarrow \mathbb{R}$, one has the integration by parts formula: for any $0 \leq t_1 \leq t_2 \leq T$,

$$-\left[\int_{\mathbb{T}^d} \vartheta u\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} u\partial_t \vartheta + \langle D\vartheta, ADu \rangle + \vartheta\partial_i A_{ij}\partial_j u \leq \int_{t_1}^{t_2} \int_{\mathbb{T}^d} f\vartheta. \tag{3.13}$$

Proof. First, observe that $x \mapsto u(x, t)$ is a well-defined $L^1(\mathbb{T}^d)$ function for a.e. $t \in (0, T)$. Then by standard convolution smoothing arguments, one can check that (3.13) holds for a.e. $t_1, t_2 \in [0, T]$ with $t_1 \leq t_2$. Indeed, if ξ_ε is a convolution kernel, then $u_\varepsilon = \xi_\varepsilon * u$, $f_\varepsilon = \xi_\varepsilon * f$ can be shown to satisfy

$$-\partial_t u_\varepsilon - A_{ij} \partial_{ij} u_\varepsilon \leq f_\varepsilon + R_\varepsilon \quad (3.14)$$

where $R_\varepsilon \rightarrow 0$ in L^1 as $\varepsilon \rightarrow 0$. Then integration by parts implies, for $0 < t_1 \leq t_2 < T$ and ε small enough, that

$$-\left[\int_{\mathbb{T}^d} \vartheta u_\varepsilon\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} u_\varepsilon \partial_t \vartheta + \langle D\vartheta, ADu_\varepsilon \rangle + \vartheta \partial_i A_{ij} \partial_j u_\varepsilon \leq \int_{t_1}^{t_2} \int_{\mathbb{T}^d} (f_\varepsilon + R_\varepsilon) \vartheta. \quad (3.15)$$

Since $u_\varepsilon(\cdot, t) \rightarrow u(\cdot, t)$ in $L^1(\mathbb{T}^d)$ for a.e. t , and likewise $u_\varepsilon \rightarrow u$, $Du_\varepsilon \rightarrow Du$, and $f_\varepsilon \rightarrow f$ in L^1 , so by letting $\varepsilon \rightarrow 0$ we deduce the (3.13) for a.e. $t_1, t_2 \in [0, T]$ with $t_1 \leq t_2$.

Now define, for a.e. $t \in [0, T]$, the functions

$$G(t) = \int_{\mathbb{T}^d} \vartheta(x, t) u(x, t) dx + F(t),$$

$$F(t) = \int_t^T \int_{\mathbb{T}^d} u \partial_t \vartheta + \langle D\vartheta, ADu \rangle + \vartheta \partial_i A_{ij} \partial_j u - f \vartheta.$$

Now F is absolutely continuous, being the integral of an $L^1(0, T)$ function. By what we have shown $G(t)$ is increasing on its domain, and moreover $G(T) \leq \int_{\mathbb{T}^d} \vartheta(x, T) u_T(x) dx$ by hypothesis. Thus $I := G - F$ is BV, and (3.13) indeed continues to hold for all $0 \leq t_1 \leq t_2 \leq T$, even if we replace $\int_{\mathbb{T}^d} \vartheta(x, t) u(x, t) dx$ by any value between $I(t+)$ and $I(t-)$. $\square$

We extend the functional D to triplets $(u, P, \gamma) \in \mathcal{K}$:

$$D(u, P, \gamma) = - \int_{\mathbb{T}^d} u(x, 0^+) m_0(x) dx + \int_0^T \Phi^*(P(t)) dt + \iint_Q F^*(x, \gamma(x, t)) dx dt.$$

We consider the following relaxation of problem (2.3):

$$\inf_{(u, P, \gamma) \in \mathcal{K}} D(u, P, \gamma). \quad (3.16)$$

Proposition 3.8. *We have*

$$\inf_{(u, P, \gamma) \in \mathcal{K}_0} D(u, P, \gamma) = \inf_{(u, P, \gamma) \in \mathcal{K}} D(u, P, \gamma).$$

The proof requires an integration by parts formula, established in the following lemma.

Lemma 3.9. *Let $(u, P, \gamma) \in \mathcal{K}$ and $(m, w) \in \mathcal{K}_1$ satisfy (2.7). Assume that $mH^*(\cdot, -w/m) \in L^1(Q)$. Then*

$$\gamma m \in L^1((0, T) \times \mathbb{T}^d), \quad \langle P(\cdot), \int_{\mathbb{T}^d} G(x) w(x, \cdot) dx \rangle \in L^1(0, T)$$

and for almost all $t \in (0, T)$ we have

$$\begin{aligned} & \int_{\mathbb{T}^d} (u(T)m_T - u(t)m(t)) \, dx + \int_t^T \int_{\mathbb{T}^d} (m\gamma + mH^*(x, -\frac{w}{m}) \\ & \quad + \langle P(t), G(x)w(x, t) \rangle) \, dx \, dt \geq 0, \end{aligned} \quad (3.17)$$

and

$$\begin{aligned} & \int_{\mathbb{T}^d} (u(t)m(t) - u(0)m_0) \, dx + \int_0^t \int_{\mathbb{T}^d} (m\gamma + mH^*(x, -\frac{w}{m}) \\ & \quad + \langle P(t), G(x, t)w(x, t) \rangle) \, dx \, dt \geq 0. \end{aligned} \quad (3.18)$$

Moreover, if equality holds in the inequality (3.17) for $t = 0$, then $w = -mD_\xi H(\cdot, Du + G^\top P)$ a.e.

Proof. In the interest of smoothing (m, w) by convolution, extend the pair to $[-1, T+1] \times \mathbb{T}^d$ by defining $m = m_0$ on $[-1, 0]$, $m = m(T)$ on $[T, T+1]$, and $w(s, x) = 0$ for $(s, x) \in (-1, 0) \cup (T, T+1) \times \mathbb{T}^d$. Let $\tilde{A}_{ij}$ be an extension of A_{ij} with $\tilde{A}_{ij} = A_{ij}$ if $t \in (0, T)$ and zero otherwise. Note that with these described extensions, (m, w) solves

$$\partial_t m - \partial_{ij}(\tilde{A}_{ij}(t, x)m) + \nabla \cdot w = 0 \quad \text{on } (-1, T+1) \times \mathbb{T}^d.$$

Let $\xi^\epsilon = \xi^\epsilon(t, x) = \xi_1^\epsilon(t)\xi_2^\epsilon(x)$ be a smooth convolution kernel with support in a ball of radius ϵ . We smoothen the pair (m, w) with this kernel in a standard way into $(m_\epsilon, w_\epsilon) = (\xi^\epsilon * m, \xi^\epsilon * w)$. Then (m_ϵ, w_ϵ) solves

$$\partial_t m_\epsilon - \partial_{ij}(\tilde{A}_{ij}m_\epsilon) + \nabla \cdot w_\epsilon = \partial_j R_\epsilon \quad \text{in } \left(-\frac{1}{2}, T + \frac{1}{2}\right) \quad (3.19)$$

in the sense of distributions, where

$$R_\epsilon := [\xi^\epsilon, \partial_j \tilde{A}_{ij}](m) + [\xi^\epsilon, \tilde{A}_{ij} \partial_j](m). \quad (3.20)$$

Here we use again the commutator notation [13]

$$[\xi^\epsilon, c](f) := \xi^\epsilon \star (cf) - c(\xi^\epsilon \star f). \quad (3.21)$$

By [13, Lemma II.1], we have that $R_\epsilon \rightarrow 0$ in L^q , since $m \in L^q$ and $\tilde{A}_{ij} \in W^{1, \infty}$. Fix time $t \in (0, T)$ at which $u(t^+) = u(t^-) = u(t)$ in $L^\kappa(\mathbb{T}^d)$ and $m_\epsilon(t)$ converges to $m(t)$. We have the following inequality based on the equality in (2.1),

$$-\partial_t u - A_{ij} \partial_{ij} u + H(x, Du + G^\top P) \leq \gamma. \quad (3.22)$$

Integrating this inequality against m_ϵ yields

$$\begin{aligned} & \int_t^T \int_{\mathbb{T}^d} u \partial_t m_\epsilon + \partial_i u \partial_j (\tilde{A}_{ij} m_\epsilon) + m_\epsilon H(x, Du + G^\top P) + \int_{\mathbb{T}^d} m_\epsilon(t) u(t) - m_\epsilon(T) u_T \\ & \leq \int_t^T \int_{\mathbb{T}^d} \gamma m_\epsilon. \end{aligned} \quad (3.23)$$

By (3.19), we have

$$\int_t^T \int_{\mathbb{T}^d} u \partial_t m_\epsilon + \partial_i u \partial_j (\tilde{A}_{ij} m_\epsilon) = \int_t^T \int_{\mathbb{T}^d} -\partial_i u R_\epsilon + \langle Du, w_\epsilon \rangle,$$

while the convexity of H in the last variable gives

$$\int_t^T \int_{\mathbb{T}^d} -m_\epsilon H^* \left(x, -\frac{w_\epsilon}{m_\epsilon} \right) \leq \int_t^T \int_{\mathbb{T}^d} \langle w_\epsilon, Du + G^\top P \rangle + m_\epsilon H(x, Du + G^\top P). \quad (3.24)$$

Combining these results yields

$$\int_{\mathbb{T}^d} m_\epsilon(t)u(t) \leq \int_{\mathbb{T}^d} m_\epsilon(T)u_T + \int_t^T \int_{\mathbb{T}^d} m_\epsilon \left(\gamma + H^* \left(x, -\frac{w_\epsilon}{m_\epsilon} \right) \right) + \langle Gw_\epsilon, P \rangle + \partial_j u R_\epsilon.$$

Following now [7], we have that as $Du \in L^{\tilde{r}}$ (where we recall that $\tilde{r} \geq p$ or A is a constant matrix), and as m is continuous in time with respect to the topology on $P(\mathbb{T}^d)$, we have, as $\epsilon \rightarrow 0$,

$$\begin{aligned} \int_t^T \int_{\mathbb{T}^d} \partial_j u R_\epsilon &\rightarrow 0, \\ \int_t^T \int_{\mathbb{T}^d} -m_\epsilon H^* \left(x, -\frac{w_\epsilon}{m_\epsilon} \right) &\rightarrow \int_t^T \int_{\mathbb{T}^d} -m H^* \left(x, -\frac{w}{m} \right), \\ \int_{\mathbb{T}^d} m_\epsilon(T)u_T &\rightarrow \int_{\mathbb{T}^d} m(T)u_T. \end{aligned}$$

(For the second limit, cf. [6].) We also claim that

$$\int_t^T \int_{\mathbb{T}^d} \langle Gw_\epsilon, P \rangle \rightarrow \int_t^T \int_{\mathbb{T}^d} \langle Gw, P \rangle.$$

To see this, recall Equation (2.7) from Lemma 2.1. If $\frac{1}{s} + \frac{1}{pr} \geq 1$, we deduce that $w \in L^s$ and therefore $w_\epsilon \rightarrow w$ in L^s ; from this the claim follows immediately. Otherwise, if $\frac{1}{s} + \frac{1}{pr} < 1$, then by Hypothesis (H1) we assume that G is constant. Therefore, we have

$$\int_t^T \int_{\mathbb{T}^d} \langle Gw_\epsilon, P \rangle = \int_t^T \left\langle \int_{\mathbb{T}^d} Gw_\epsilon, P \right\rangle = \int_t^T \left\langle \xi_1^\epsilon * \int_{\mathbb{T}^d} Gw, P \right\rangle \rightarrow \int_t^T \left\langle \int_{\mathbb{T}^d} Gw, P \right\rangle$$

because $t \mapsto \int_{\mathbb{T}^d} Gw$ is in L^s . Now since $u \in L^\kappa(Q)$, $m \in L^q(Q)$, and $\kappa \geq p$, $m_\epsilon u$ strongly converges to mu in $L^1(Q)$ and thus up to a subsequence, $\int_{\mathbb{T}^d} m_\epsilon(t)u(t) \rightarrow \int_{\mathbb{T}^d} m(t)u(t)$ a.e. We now have that

$$\int_{\mathbb{T}^d} m(t)u(t) \, dx \leq \int_{\mathbb{T}^d} m(T)u_T \, dx + \int_t^T \int_{\mathbb{T}^d} m \left(\gamma + H^* \left(x, -\frac{w}{m} \right) \right) + \langle P, Gw \rangle \, dx \, dt.$$

An analogous argument produces the other desired inequality, so now assume that equality holds in inequality (3.18) with $t = 0$. Then there is $t^* \in (0, T)$ where equality holds with $t = t^*$. Let

$$\begin{aligned} E_\sigma(t) &:= \left\{ (s, y) : s \in [t, T], \, m \left(H^* \left(y, -\frac{w}{m} \right) \right. \right. \\ &\quad \left. \left. + H(x, Du + G^\top P) \right) \geq -\langle w, Du + G^\top P \rangle + \sigma \right\}. \end{aligned}$$

If $|E_\sigma(t)| > 0$, then for $\epsilon > 0$ small enough, the set of s, y satisfying

$$m_\epsilon \left(H^* \left(y, -\frac{w_\epsilon}{m_\epsilon} \right) + H(x, Du + G^\top P) \right) \geq -\langle w_\epsilon, Du + G^\top P \rangle + \frac{\sigma}{2}$$

has measure larger than $\frac{|E_\sigma(t)|}{2}$. Then by (3.24), for the fixed choice of ϵ ,

$$\begin{aligned} \int_{t^*}^T \int_{\mathbb{T}^d} -m_\epsilon H^* \left(x, -\frac{w_\epsilon}{m_\epsilon} \right) &\leq \int_{t^*}^T \int_{\mathbb{T}^d} \langle w_\epsilon, Du + G^\top P \rangle \\ &\quad + m_\epsilon H(x, Du + G^\top P) - |E_\sigma(t)|\sigma/4, \end{aligned}$$

whereby we obtain strict inequality in (3.18) with $t = t^*$, a contradiction. Thus $|E_\sigma(t)| = 0$ for any σ and a.e. t ,

$$\langle -w, Du + G^\top P \rangle = m \left(H(x, Du + G^\top P) + H^*(y, -\frac{w}{m}) \right),$$

and hence

$$w = -mD_\xi H(\cdot, Du + G^\top P) \quad \text{a.e. in } (0, T) \times \mathbb{T}^d.$$

□

Proof of Proposition 3.8. It is clear that the value of the relaxed problem is smaller than the value of problem (2.3). It remains to show the other inequality. For any $(m, w) \in \mathcal{K}_1$ with $mH^*(-w/m) \in L^1(Q)$, we have, by Fenchel–Young inequality and Lemma 3.9,

$$\begin{aligned} D(u, P, \gamma) &\geq - \int_{\mathbb{T}^d} u(0)m_0 + \int_0^T \left(\langle P(t), \int_{\mathbb{T}^d} Gw \rangle - \Phi(\int_{\mathbb{T}^d} Gw) \right) + \iint_Q (\gamma m - F(m)) \\ &\geq - \int_{\mathbb{T}^d} u_T m(T) - \iint_Q mH^*\left(-\frac{w}{m}\right) - \int_0^T \Phi(\int_{\mathbb{T}^d} Gw) - \iint_Q F(m) \\ &= -B(m, w). \end{aligned}$$

Maximizing the right-hand side with respect to (m, w) , we obtain with Lemma 2.1 that

$$D(u, P, \gamma) \geq - \inf_{(m, w) \in \mathcal{K}_1} B(m, w) = \inf_{(u, P, \gamma) \in \mathcal{K}_0} D(u, P, \gamma),$$

which concludes the proof. □

3.3. Existence of a relaxed solution

We establish now the existence of a relaxed solution.

Proposition 3.10. *The relaxed problem (3.16) has at least one solution $(u, P, \gamma) \in \mathcal{K}$.*

Proof. Let (u_n, P_n, γ_n) be a minimizing sequence for problem (2.3). By Proposition (3.8), it is also a minimizing sequence for the relaxed problem (3.16). We can, without loss of generality, assume that $\gamma_n \geq 0$, so long as we only require u_n to be a viscosity solution to the Hamilton–Jacobi equation. Let us replace γ_n with its positive part, i.e. $(\gamma_n)_+ := \max\{\gamma_n, 0\}$. Then we replace u_n with $\tilde{u}_n$, the continuous viscosity solution of

$$-\partial_t \tilde{u}_n - A_{ij} \partial_{ij} \tilde{u}_n + H(D\tilde{u}_n + G^\top P_n) = (\gamma_n)_+, \quad \tilde{u}_n(x, T) = u_T(x).$$

By [23], $\tilde{u}_n$ also satisfies this equation in the sense of distributions, and thus the new triple $(\tilde{u}_n, P_n, (\gamma_n)_+)$ is also a member of $\mathcal{K}$. We have $\tilde{u}_n \geq u_n$ and $F^*(\gamma_n) = F^*((\gamma_n)_+)$ for all $(x, t) \in Q$. Therefore, $D(\tilde{u}_n, P_n, (\gamma_n)_+) \leq D(u_n, P_n, \gamma_n)$, and thus the new sequence also minimizes D . The arguments below will then apply to $(\tilde{u}_n, P_n, (\gamma_n)_+)$ in place of (u_n, P_n, γ_n) .

Step 1: [Bounds for (γ_n) , (P_n) , and (Du_n)]:

All constants C used in this part of the proof are independent of n . We integrate (3.1) against m_0 on Q and obtain

$$\begin{aligned} & \int_{\mathbb{T}^d} u_n(0)m_0 + \iint_Q \partial_j u_n \partial_i (A_{ij} m_0) + \iint_Q H(Du_n + G^\top P_n) m_0 \\ & \leq \iint_Q \gamma_n m_0 + \int_{\mathbb{T}^d} u_T m_0. \end{aligned} \quad (3.25)$$

Let us recall that $m_0 \geq \frac{1}{C}$. The Hamiltonian term can be bounded from below as:

$$\iint_Q H(Du_n + G^\top P_n) m_0 \geq \frac{1}{C} \|Du_n + G^\top P_n\|_r^r - C. \quad (3.26)$$

In light of the regularity assumptions on A and m_0 , we also have that

$$\left| \iint_Q \partial_j u_n \partial_i (A_{ij} m_0) \right| \leq C \|Du_n\|_1 \leq C \|Du_n\|_{\tilde{r}}. \quad (3.27)$$

Finally, the right-hand side of (3.25) is bounded by $C\|\gamma_n\|_p + C$. Combining this estimate with (3.26) and (3.27), we obtain that

$$\int_{\mathbb{T}^d} u_n(0)m_0 + \frac{1}{CB} \|Du_n + G^\top P_n\|_r^r - C \|Du_n\|_{\tilde{r}} \leq C \|\gamma_n\|_p + C \quad (3.28)$$

for any choice of $B \geq 1$. The constants C used are also independent of B . Now we use the fact that (u_n, P_n, γ_n) is a minimizing sequence and the growth assumptions on F^* and Φ^* to derive

$$- \int_{\mathbb{T}^d} u_n(0)m_0 + \frac{1}{C} \|P_n\|_{s'}^{s'} + \frac{1}{C} \|\gamma_n\|_p^p - C \leq D(u_n, P_n, \gamma_n) \leq C. \quad (3.29)$$

Summing up (3.28) and (3.29), we obtain

$$\frac{1}{CB} \|Du_n + G^\top P_n\|_r^r - C \|Du_n\|_{\tilde{r}} + \frac{1}{C} \|P_n\|_{s'}^{s'} + \frac{1}{C} \|\gamma_n\|_p^p \leq C \|\gamma_n\|_p + C. \quad (3.30)$$

Now by Hölder's inequality we have

$$\|Du\|_{\tilde{r}}^{\tilde{r}} \leq C (\|Du + G^\top P\|_{\tilde{r}}^{\tilde{r}} + \|G^\top P\|_{\tilde{r}}^{\tilde{r}}) \leq C (\|Du + G^\top P\|_r^r + \|P\|_{s'}^{s'} + 1)$$

and so

$$\frac{1}{CB} \|Du_n\|_{\tilde{r}}^{\tilde{r}} + \left[\frac{1}{C} - \frac{C}{B} \right] \|P_n\|_{s'}^{s'} - C \|Du_n\|_{\tilde{r}} + \frac{1}{C} \|\gamma_n\|_p^p \leq C \|\gamma_n\|_p + C. \quad (3.31)$$

We fix now $B = 2C^2$. The terms $\|Du_n\|_{\tilde{r}}$, $\|\gamma_n\|_p$ can be absorbed. For instance, the former can be absorbed by $\|Du\|_{\tilde{r}}^{\tilde{r}}$ insofar as for an arbitrarily small $\varepsilon > 0$, there exists $C > 0$ (depending on ε) such that

$$\|Du_n\|_{\tilde{r}} \leq \varepsilon \|Du_n\|_{\tilde{r}}^{\tilde{r}} + C. \quad (3.32)$$

Taking ε small enough, we finally deduce from (3.31) the estimate

$$\|Du_n\|_{\tilde{r}}^{\tilde{r}} + \|P_n\|_{s'}^{s'} + \|\gamma_n\|_p^p \leq C, \quad (3.33)$$

so that $(\gamma_n)_{n \in \mathbb{N}}$ is bounded in $L^p(Q)$, $(P_n)_{n \in \mathbb{N}}$ is bounded in $L^{s'}((0, T); \mathbb{R}^k)$ and $(Du_n)_{n \in \mathbb{N}}$ is bounded in $L^{\tilde{r}}(Q)$. Inequality (3.30) further shows that $Du_n + G^\top P_n$ is bounded in $L^r(Q; \mathbb{R}^d)$. This implies that

$$\|H(Du_n + G^\top P_n)\|_1 \leq C.$$

Step 2 [Bound of u_n in $L^\kappa(Q)$]:

Now that we have estimates on P_n in $L^{s'}$, γ_n in L^p , Du_n in $L^{\tilde{r}}$, and $H(Du_n + G^\top P_n)$ in L^1 , we can apply Corollary 3.3 in case $s' < r$ or Corollary 3.4 in case $s' \geq r$ and obtain $\|u_n\|_\kappa \leq C$, where κ is defined at the end of Sect. 3.1.

Step 3 [Conclusion]:

The rest of the proof is very similar to the proof of [7, Proposition 5.4], we only give the main lines. By passing to a subsequence, we assume without loss of generality that

$$\begin{aligned} u_n &\rightharpoonup \bar{u} \text{ in } L^\kappa(Q), \quad Du_n \rightharpoonup D\bar{u} \text{ in } L^{\tilde{r}}(Q), \quad Du_n + G^\top P_n \rightharpoonup D\bar{u} + G^\top \bar{P} \text{ in } L^r(Q; \mathbb{R}^d), \\ \gamma_n &\rightharpoonup \bar{\gamma} \text{ in } L^p(Q), \quad P_n \rightharpoonup \bar{P} \text{ in } L^{s'}(0, T). \end{aligned}$$

Since H is convex, $(\bar{u}, \bar{P}, \bar{\gamma}) \in \mathcal{K}$. By weak lower semicontinuity arguments, we have

$$\liminf_{n \rightarrow \infty} \iint_Q F^*(\gamma_n) + \int_0^T \Phi^*(P_n) \geq \iint_Q F^*(\bar{\gamma}) + \int_0^T \Phi^*(\bar{P}).$$

Using exactly the same arguments as in [7, Proposition 5.4 (Step 3)], one can prove that

$$\limsup_{n \rightarrow \infty} \int_{\mathbb{T}^d} u_n(0) m_0 \leq \int_{\mathbb{T}^d} \bar{u}(0) m_0,$$

which proves the optimality of $(\bar{u}, \bar{P}, \bar{\gamma})$. $\square$

4. Existence and uniqueness of a solution for the MFG system

We prove in this section the existence and uniqueness of a *weak solution* to the MFG system (1.2).

Definition 4.1. We say that a quadruplet $(u, P, m, w) \in L^\kappa(Q) \times L^{s'}(0, T) \times L^q(Q) \times L^{\frac{r'q}{r'+q-1}}(Q)$ is a weak solution if

- (i) The following integrability conditions hold: $Du \in L^{\tilde{r}}(Q)$ and $mH^*(\cdot, -m/w) \in L^1(Q)$.
- (ii) Equation (1.2)-(i) holds in the sense of distributions,

$$-\partial_t u - A_{ij} \partial_{ij} u + H(Du + G^\top P) \leq f(m), \quad u(T) \leq u_T$$

- (iii) Equation (1.2)-(ii) holds in the sense of distributions,

$$\partial_t m - \partial_{ij} (A_{ij} m) - \nabla \cdot w = 0, \quad m(0) = m_0,$$

- (iv) Equation (1.2)-(iii)-(iv) hold almost everywhere,

(v) The following equality holds:

$$\iint_Q \left(mf(m) + mH^*(-w/m) + \langle P, Gw \rangle \right) + \int_{\mathbb{T}^d} m(T)u_T - \int_{\mathbb{T}^d} m_0 u(0) = 0. \quad (4.1)$$

Theorem 4.2. *There exists a weak solution (u, P, m, w) to the MFG system (1.2). It is unique in the following sense: if (u, P, m, w) and (u', P', m', w') are two solutions, then $m = m'$, $w = w'$, $P = P'$ a.e. and $u = u'$ in $\{m > 0\}$.*

Theorem 4.3. *Let $(\bar{m}, \bar{w}) \in \mathcal{K}_1$ be a minimizer of (2.6) and $(\bar{u}, \bar{P}, \bar{\gamma})$ be a minimizer of (3.16). Then, $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ is a weak solution of the MFG system and $\bar{\gamma} = f(\bar{m})$.*

Conversely, any weak solution $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ of the MFG system is such that $(\bar{m}, \bar{w})$ is the solution to (2.6) and $(\bar{u}, \bar{P}, f(\bar{m}))$ is a solution to (3.16).

Proof. Part 1. Let $(\bar{m}, \bar{w}) \in \mathcal{K}_1$ be the solution to (2.6) and $(\bar{u}, \bar{P}, \bar{\gamma}) \in \mathcal{K}$ be a solution to (3.16). Condition (iii) of Definition (4.1) is verified by the definition of $\mathcal{K}_1$. By Lemma 2.1 and Proposition 3.8, these two problems have the same value, thus

$$\begin{aligned} 0 &= D(\bar{u}, \bar{P}, \bar{\gamma}) + B(\bar{m}, \bar{w}) \\ &= \iint_Q (F^*(\bar{\gamma}) + F(\bar{m})) + \int_0^T \left(\Phi^*(\bar{P}) + \Phi(\int_{\mathbb{T}^d} G\bar{w}) \right) \\ &\quad + \iint_Q \bar{m}H^*(-\bar{w}/\bar{m}) + \int_{\mathbb{T}^d} u_T \bar{m}(T) - \int_{\mathbb{T}^d} \bar{u}(0)m_0. \end{aligned}$$

By the Fenchel–Young inequality, we have

$$F^*(\bar{\gamma}) + F(\bar{m}) \geq \bar{\gamma}\bar{m} \quad \text{for a.e. } (x, t) \in Q, \quad (4.2)$$

$$\Phi^*(\bar{P}) + \Phi(\int_{\mathbb{T}^d} G\bar{w}) \geq \langle \bar{P}, \int_{\mathbb{T}^d} G\bar{w} \rangle \quad \text{for a.e. } t \in (0, T) \quad (4.3)$$

thus

$$0 \geq \iint_Q \left(\bar{m}\bar{\gamma} + \bar{m}H^*(-\bar{w}/\bar{m}) + \langle P, Gw \rangle \right) + \int_{\mathbb{T}^d} \bar{m}(T)u_T - \int_{\mathbb{T}^d} m_0 \bar{u}(0). \quad (4.4)$$

This implies first that $\bar{m}H^*(-\bar{w}/\bar{m}) \in L^1(Q)$. Moreover, by Lemma 3.9, inequality (4.4) is in fact an equality and $\bar{w} = -\bar{m}D_\xi H(D\bar{u} + G^\top \bar{P})$. Moreover, the equality holds a.e. in (4.2) and (4.3) therefore,

$$\begin{aligned} \bar{\gamma} &= F'(\bar{m}) = f(\bar{m}) \quad \text{for a.e. } (x, t) \in Q, \\ \bar{P} &= D\Phi(\int_{\mathbb{T}^d} G\bar{w}) = \Psi(\int_{\mathbb{T}^d} G\bar{w}) \quad \text{for a.e. } t \in (0, T). \end{aligned}$$

Since (4.4) is an equality and $\bar{\gamma} = f(\bar{m})$, (4.1) (condition (v)) holds true. Further, by the definition of $\mathcal{K}$ and $\bar{\gamma} = f(\bar{m})$, condition (ii) holds. We conclude then that $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ is a weak solution to the MFG system.

Part 2. Let $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ be a weak solution to (1.2). Let $\bar{\gamma} = f(\bar{m})$. The growth condition on f implies that $\bar{\gamma} \in L^p(Q)$. Therefore, $(\bar{m}, \bar{w}) \in \mathcal{K}_1$ and

$(\bar{u}, \bar{P}, \bar{\gamma}) \in \mathcal{K}$. It remains to show that $(\bar{u}, \bar{P}, \bar{\gamma})$ solves (3.16) and that $(\bar{m}, \bar{w})$ solves (3.16).

The argument is very similar to the one used in Proposition 3.8. It mainly consists in showing that $D(\bar{u}, \bar{P}, \bar{\gamma}) + B(\bar{m}, \bar{w}) = 0$. Since $\bar{\gamma} = f(\bar{m}) = F'(\bar{m})$ a.e., we have by convexity of F that

$$F(\bar{m}) + F^*(\bar{\gamma}) = \bar{\gamma}\bar{m}, \quad \text{for a.e. } (x, t) \in Q.$$

Similarly, since $\bar{P} = \Psi(\int_{\mathbb{T}^d} Gw) = D\Phi(\int_{\mathbb{T}^d} Gw)$, we have

$$\Phi(\int_{\mathbb{T}^d} Gw) + \Phi^*(\bar{P}) = \langle P, \int_{\mathbb{T}^d} Gw \rangle, \quad \text{for a.e. } t \in (0, T).$$

These two equalities and (4.1) yield:

$$\begin{aligned} D(\bar{u}, \bar{P}, \bar{\gamma}) + B(\bar{m}, \bar{w}) &= \iint_Q (F^*(\bar{\gamma}) + F(\bar{m})) + \int_0^T \left(\Phi(\int_{\mathbb{T}^d} G\bar{w}) + \Phi^*(\bar{P}) \right) \\ &\quad + \iint_Q \bar{m}H^*(-\bar{w}/\bar{m}) + \int_{\mathbb{T}^d} (u_T\bar{m}(T) - \bar{u}(0)m_0) \\ &= \iint_Q \bar{m}\bar{\gamma} + \iint_Q \langle \bar{P}, G\bar{w} \rangle \\ &\quad + \int_{\mathbb{T}^d} u_T\bar{m}(T) - \bar{u}(0)m_0 + \iint_Q \bar{m}H^*(-\bar{w}/\bar{m}) \\ &= 0. \end{aligned}$$

As a consequence, we obtain

$$\inf_{(u, P, \gamma) \in \mathcal{K}} D(u, P, \gamma) \leq D(\bar{u}, \bar{P}, \bar{\gamma}) = -B(\bar{m}, \bar{w}) \leq - \min_{(m, w) \in \mathcal{K}_1} B(m, w).$$

The first and the last term being equal, the two above inequalities are equalities, which shows the optimality of optimality of $(\bar{u}, \bar{P}, \bar{\gamma})$ and $(\bar{m}, \bar{w})$, respectively. $\square$

Proof of Theorem 4.2. By Lemma 2.1, problem (2.6) has a solution $(\bar{m}, \bar{w})$ and by Proposition 3.10, problem (3.16) has a solution $(\bar{u}, \bar{P}, \bar{\gamma})$. By Theorem 4.3, $(\bar{u}, \bar{P}, \bar{m}, \bar{w})$ is a weak solution to the MFG system.

Now, let (u_1, P_1, m_1, w_1) and (u_2, P_2, m_2, w_2) be two weak solutions. By Theorem 4.3, (m_1, w_1) and (m_2, w_2) are solutions to problem (2.6), they are therefore equal. Relation (1.2)-(iii) implies that $P_1 = P_2$. Let $(\bar{m}, \bar{w}, \bar{P}) = (m_1, w_1, P_1)$ denote the common values. Let $\bar{\gamma} = f(\bar{m})$. Then $(u_1, \bar{P}, \bar{\gamma})$ and $(u_2, \bar{P}, \bar{\gamma})$ lie in $\mathcal{K}$ (by definition of weak solutions).

For $t \in (0, T)$, $(m, w) \in \mathcal{K}_1$, and $(u, P, \gamma) \in \mathcal{K}$, we introduce

$$\begin{aligned} B_t(m, w) &= \int_t^T \int_{\mathbb{T}^d} \left(mH^*(-w/m) + F(m) \right) + \int_t^T \Phi(\int_{\mathbb{T}^d} Gw) + \int_{\mathbb{T}^d} u_T m(T) \\ D_t(u, P, \gamma) &= - \int_{\mathbb{T}^d} u(0)m_0 + \int_t^T \Phi^*(P) + \int_t^T \int_{\mathbb{T}^d} F^*(\gamma). \end{aligned}$$

Proceeding as in the proof of Proposition 3.8, we obtain that

$$\inf_{(u, P, \gamma) \in \mathcal{K}} D_t(u, P, \gamma) \geq -B_t(\bar{m}, \bar{w}).$$

By Lemma 3.9 and relation (4.1),

$$\int_t^T \int_{\mathbb{T}^d} \left(\bar{m} f(\bar{m}) + \bar{m} H^*(-\bar{w}/\bar{m}) + \langle \bar{P}, G\bar{w} \rangle \right) + \int_{\mathbb{T}^d} \bar{m}(T) u_T - \int_{\mathbb{T}^d} \bar{m}(t) u_i(t) = 0$$

for a.e. $t \in (0, T)$ and for $i = 1, 2$. Proceeding as in the proof of Theorem 4.3, we obtain that $-B_t(\bar{m}, \bar{w}) = D_t(u_i, \bar{P}, \bar{\gamma})$. Thus $(u_1, \bar{P}, \bar{\gamma})$ and $(u_2, \bar{P}, \bar{\gamma})$ minimize D_t over $\mathcal{K}$.

Let $\bar{u} = u_1 \vee u_2$. Adapting the proof in [7, Theorem 6.2], we obtain that $(\bar{u}, \bar{P}, \bar{\gamma}) \in \mathcal{K}$. Since $D_t(\bar{u}) \leq D_t(u_i)$, we deduce that $(\bar{u}, \bar{P}, \bar{\gamma})$ also minimize D_t . It follows that

$$\int_{\mathbb{T}^d} u_1(t) \bar{m}(t) = \int_{\mathbb{T}^d} u_2(t) \bar{m}(t) = \int_{\mathbb{T}^d} \bar{u}(t) \bar{m}(t)$$

As $u_1 \leq \bar{u}$ and $u_2 \leq \bar{u}$, this implies that $u_1 = u_2 = \bar{u}$ a.e. in $\{\bar{m} > 0\}$ and concludes the proof. $\square$

5. Regularity estimates

In this section we adapt the methods used in [17, 18] to show that weak solutions of (1.1) possess extra regularity–Sobolev estimates in both space and time—not required by Definition 4.1. These estimates hold under general *strong monotonicity* assumptions on the coupling $f(x, m)$ and *coercivity* on the Hamiltonian. We divide our results into “space regularity,” i.e. estimates on derivatives with respect to x , and “time regularity,” estimates on derivatives with respect to t .

5.1. Space regularity

Before stating the result, let us enumerate a few additional assumptions.

Assumption 5.1. A_{ij} is constant for every i, j .

Assumption 5.2. (*Strong monotonicity*) We have a Lipschitz estimate on f of the form

$$|f(x, m) - f(y, m)| \leq C(m^{q-1} + 1)|x - y| \quad \forall x, y \in \mathbb{T}^d, \quad m \geq 0. \quad (5.1)$$

We also assume that $f(x, m)$ is strongly monotone in m , that is, there exists $c_f > 0$ such that

$$(f(x, \tilde{m}) - f(x, m))(\tilde{m} - m) \geq c_f \min\{\tilde{m}^{q-2}, m^{q-2}\} |\tilde{m} - m|^2 \quad \forall \tilde{m}, m \geq 0, \quad \tilde{m} \neq m. \quad (5.2)$$

If $q < 2$ one should interpret 0^{q-2} as $+\infty$ in (5.2). In this way, when $\tilde{m} = 0$, for instance, (5.2) reduces to $f(x, m)m \geq c_f m^q$, as in the more regular case $q \geq 2$.

Assumption 5.3. (*Coercivity*) There exist $j_1, j_2 : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $c_H > 0$ such that

$$H(x, \xi) + H^*(x, \zeta) - \xi \cdot \zeta \geq c_H |j_1(\xi) - j_2(\zeta)|^2. \quad (5.3)$$

In particular, and in light of our restriction on the growth of H , we specify that $j_1(\xi) \sim |\xi|^{r/2-1}\xi$ and $j_2(\zeta) \sim |\zeta|^{r'/2-1}\zeta$.

Assumption 5.4. $m_0 \in W^{2,\infty}(\mathbb{T}^d)$, $u_T \in W^{2,\infty}(\mathbb{T}^d)$, $G \in W^{2,\infty}(\mathbb{T}^d; \mathcal{L}(\mathbb{R}^k, \mathbb{R}^d))$, and H^* is twice continuously differentiable in x with

$$|D_{xx}^2 H^*(x, \zeta)| \leq C_H \left(|\zeta|^{r'} + 1 \right). \quad (5.4)$$

Notice that Assumption 5.2 holds for the canonical case $f(x, m) = m^{q-1}$ or even if $\tilde{f}(x, m) = \tilde{f}(x)m^{q-1}$ for some strictly positive, Lipschitz continuous function $\tilde{f}$ on $\mathbb{T}^d$. Assumption 5.3 likewise holds for a canonical structure $H(x, \xi) = c(x)|\xi|^r$ for some strictly positive, $\mathcal{C}^2$ smooth function $c(x)$ on $\mathbb{T}^d$.

Proposition 5.5. *Let Assumptions 5.1, 5.2, 5.3, 5.4 hold. Then, if (u, m) is a weak solution of (1.1),*

$$\|m^{\frac{q}{2}-1} Dm\|_{L^2(Q)} \leq C \quad \text{and} \quad \|m^{1/2} D(j_1(Du))\|_{L^2(Q)} \leq C,$$

where C is a constant depending only on the data.

Proof. Throughout we use the notation $g^\delta(x) = g(x + \delta)$ for any function depending on $x \in \mathbb{T}^d$.

Take a smooth minimizing sequence $(u_n, P_n, \gamma_n) \in \mathcal{K}_0$ for the dual problem. Integrate (2.4) by parts against u_n and rearrange to get

$$\begin{aligned} \iint_Q m H(x, Du_n + G^\top P_n) \, dx \, dt &= \int_{\mathbb{T}^d} (u_T m(T) - u_n(0) m_0) \, dx \\ &+ \iint_Q \gamma_n m - \langle Du_n, w \rangle \, dx \, dt. \end{aligned} \quad (5.5)$$

Step 1. The following estimates show that (up to a subsequence) $Du_n \rightharpoonup Du$ in $L^{\tilde{r}}_{m'}([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$ (see Sect. 3 for definition of $\tilde{r}$, and NB $\tilde{r} \leq \min\{r, s'\}$):

Using Young's inequality and Assumption (H2) we get

$$\begin{aligned} \frac{1}{C} \iint_Q |Du_n + G^\top P_n|^{\tilde{r}} m \, dx \, dt &\leq \frac{1}{C} \iint_Q |Du_n + G^\top P_n|^r m \, dx \, dt + C \\ &\leq \|u_T\|_\infty + \int_{\mathbb{T}^d} |u_n(0)| m_0 + \iint_Q \left((\gamma_n)_+ m + C m \left| \frac{w}{m} \right|^{r'} \right) \, dx \, dt + C. \end{aligned} \quad (5.6)$$

By possibly increasing C we get

$$\begin{aligned}
& \frac{1}{C} \iint_Q |Du_n|^{\tilde{r}} m \, dx \, dt \\
& \leq \|u_T\|_\infty + \int_{\mathbb{T}^d} |u_n(0)| m_0 \\
& \quad + \iint_Q \left((\gamma_n)_+ m + Cm \left| \frac{w}{m} \right|^{r'} + C |G^\top P_n|^{\tilde{r}} m \right) \, dx \, dt + C \\
& \leq \|u_T\|_\infty + \int_{\mathbb{T}^d} |u_n(0)| m_0 \\
& \quad + \iint_Q \left((\gamma_n)_+ m + Cm \left| \frac{w}{m} \right|^{r'} \right) \, dx \, dt + C \int_0^T |P_n|^{s'} \, dt + C. \quad (5.7)
\end{aligned}$$

Since P_n is bounded in $L^{s'}$, we have that Du_n is bounded in $L_m^{\tilde{r}}$ where we recall that $\tilde{r} = \min(r, s')$. Thus, up to a subsequence, $Du_n \rightharpoonup \zeta$ for some $\zeta \in L_m^{\tilde{r}}$. The argument that $\zeta = Du$ m -a.e. follows as in [17]. We also have, up to a subsequence, that $P_n \rightharpoonup P$ in $L^{s'}(0, T)$, and thus also that $Du_n + G^\top P_n \rightharpoonup Du + G^\top P$ in $L_m^r([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$. Indeed, the upper bound given by (5.6) shows that $Du_n + G^\top P_n$ converges weakly in L_m^r , and its limit must be equal to $Du + G^\top P$ a.e. by taking the limit of each summand.

Step 2. Now use u_n^δ and $u_n^{-\delta}$ as test functions in (2.4) to get

$$\begin{aligned}
& \int_{\mathbb{T}^d} (u_T^\delta m(T) - u_n^\delta(0) m_0) \, dx \\
& = \iint_Q (H(x + \delta, Du_n^\delta + (G^\delta)^\top P_n) m - \gamma_n^\delta m + Du_n^\delta \cdot w) \, dx \, dt \quad (5.8)
\end{aligned}$$

and

$$\begin{aligned}
& \int_{\mathbb{T}^d} (u_T^{-\delta} m(T) - u_n^{-\delta}(0) m_0) \, dx \\
& = \iint_Q (H(x - \delta, Du_n^{-\delta} + (G^{-\delta})^\top P_n) m - \gamma_n^{-\delta} m + Du_n^{-\delta} \cdot w) \, dx \, dt \\
& \quad (5.9)
\end{aligned}$$

We have the optimality condition

$$\int_{\mathbb{T}^d} (u_T m(T) - u(0) m_0) \, dx = - \iint_Q \left(H^* \left(x, -\frac{w}{m} \right) m + P \cdot (Gw) + f(x, m) m \right) \, dx \, dt. \quad (5.10)$$

Take (5.8) + (5.9) -2(5.10) to get

$$\begin{aligned}
& \int_{\mathbb{T}^d} \left((u_T^\delta + u_T^{-\delta} - 2u_T) m(T) \right. \\
& \quad \left. - u_n(0) (m_0^\delta + m_0^{-\delta} - 2m_0) \right) dx \\
&= \iint_Q \left(H(x + \delta, Du_n^\delta + (G^\delta)^T P_n) m + H^* \left(x + \delta, -\frac{w}{m} \right) m \right. \\
& \quad \left. + (Du_n^\delta + (G^\delta)^T P_n) \cdot w \right) dx dt \\
& \quad + \iint_Q \left(H(x - \delta, Du_n^{-\delta} + (G^{-\delta})^T P_n) m + H^* \left(x - \delta, -\frac{w}{m} \right) m \right. \\
& \quad \left. + (Du_n^{-\delta} + (G^{-\delta})^T P_n) \cdot w \right) dx dt \\
& \quad + \iint_Q \left((2f(x, m) - \gamma_n^\delta - \gamma_n^{-\delta}) m + 2P \cdot (Gw) - P_n \cdot (G^\delta w + G^{-\delta} w) \right) dx dt - I
\end{aligned} \tag{5.11}$$

where

$$I := \iint_Q \left(H^* \left(x + \delta, -\frac{w}{m} \right) + H^* \left(x - \delta, -\frac{w}{m} \right) - 2H^* \left(x, -\frac{w}{m} \right) \right) m \tag{5.12}$$

and where we have used

$$\int_{\mathbb{T}^d} (u_n^\delta(0) + u_n^{-\delta}(0) - 2u(0)) m_0 dx = \int_{\mathbb{T}^d} u_n(0) (m_0^\delta + m_0^{-\delta} - 2m_0) dx. \tag{5.13}$$

Since H is convex in the third argument, by the result of Step 1 and weak lower semicontinuity we have

$$\begin{aligned}
& \iint_Q H(x \pm \delta, Du^{\pm\delta} + (G^{\pm\delta})^T P) m dx dt \\
& \leq \liminf \iint_Q H(x \pm \delta, Du_n^{\pm\delta} + (G^{\pm\delta})^T P_n) m dx dt.
\end{aligned}$$

Letting $n \rightarrow \infty$ in (5.11) we obtain

$$\begin{aligned}
& \iint_Q (H(x + \delta, Du^\delta + (G^\delta)^T P) m \\
& \quad + H^* \left(x + \delta, -\frac{w}{m} \right) m + (Du^\delta + (G^\delta)^T P) \cdot w) dx dt \\
& \quad + \iint_Q (H(x - \delta, Du^{-\delta} + (G^{-\delta})^T P) m + H^* \left(x - \delta, -\frac{w}{m} \right) m + (Du^{-\delta} \\
& \quad + (G^{-\delta})^T P) \cdot w) dx dt \\
& \leq \int_{\mathbb{T}^d} ((u_T^\delta + u_T^{-\delta} - 2u_T) m(T) - u(0) (m_0^\delta + m_0^{-\delta} - 2m_0)) dx + I \\
& \quad + \iint_Q ((f(x + \delta, m^\delta) + f(x - \delta, m^{-\delta}) - 2f(x, m)) m \\
& \quad + P \cdot ((G^\delta + G^{-\delta} - 2G) w)) dx dt.
\end{aligned} \tag{5.14}$$

By [17, computation (4.25)] we have

$$\begin{aligned} & \int_{\mathbb{T}^d} \left(f(x + \delta, m^\delta) + f(x - \delta, m^{-\delta}) - 2f(x, m) \right) m \, dx \\ & \leq C|\delta|^2 \left(1 + \int_{\mathbb{T}^d} \min\{m^\delta, m\}^q \, dx \right) - \frac{cf}{2} \int_{\mathbb{T}^d} \min\{(m^\delta)^{q-2}, m^{q-2}\} |m^\delta - m|^2 \, dx. \end{aligned} \quad (5.15)$$

Using estimate (5.3) on the left-hand side of (5.14), then using $|a + b|^2 \leq 2|a|^2 + 2|b|^2$, and combining with (5.15), then using Assumption 5.4 we deduce

$$\begin{aligned} & \frac{c_H}{2} \iint_Q \left| \left(Du^\delta + (G^\delta)^T P \right)^{r/2} - \left(Du^{-\delta} + (G^{-\delta})^T P \right)^{r/2} \right| m \, dx \, dt \\ & \quad + \frac{cf}{2} \int_{\mathbb{T}^d} \min\{(m^\delta)^{q-2}, m^{q-2}\} |m^\delta - m|^2 \, dx \\ & \leq |\delta|^2 \left(\|u_T\|_{W^{2,\infty}} + \|m_0\|_{W^{2,\infty}} \int_{\mathbb{T}^d} |u(0)| \, dx + C_H \left(\left\| \frac{w}{m} \right\|_{L_m^r} + 1 \right) \right) \\ & \quad |\delta|^2 \left(C \left(1 + \int_{\mathbb{T}^d} \min\{m^\delta, m\}^q \, dx \right) + \|G\|_{W^{2,\infty}} \iint_Q |P \cdot w| \, dx \, dt \right). \end{aligned} \quad (5.16)$$

□

5.2. Time regularity

As in the previous subsection, we enumerate our assumptions before stating the main result.

Assumption 5.6. *We assume that $A_{ij} = 0$.*

We remark that Assumption 5.6 is much stronger than Assumption 5.1 but appears to be necessary, for technical reasons that appear in the estimates below.

Assumption 5.7. *(Strong monotonicity in time) We assume that (5.2) holds.*

We assume that Ψ is invertible, with inverse denoted by $(\Psi^{-1})(t, \cdot)$ (for instance, it suffices to assume that its primitive Φ is strictly convex). We assume that, for some constant $c_\Psi > 0$,

$$\begin{aligned} & \left(\Psi^{-1}(\tilde{P}) - \Psi^{-1}(P) \right) (\tilde{P} - P) \\ & \geq c_\Psi \min \left\{ \left| \tilde{P} \right|^{s'-2}, |P|^{s'-2} \right\} |\tilde{P} - P|^2 \quad \forall t, \tau \in [0, T], \, P \in \mathbb{R}^k \end{aligned} \quad (5.17)$$

Proposition 5.8. *Under Assumptions 5.6, 5.7, and (5.4), for every $\varepsilon > 0$, there exists a constant $C(\varepsilon)$, depending only on ε and the data, such that*

$$\left\| \partial_t \left(m^{q/2} \right) \right\|_{L^2(Q_\varepsilon)} + \left\| \frac{d}{dt} \left(P^{s'/2} \right) \right\|_{L^2(\varepsilon, T-\varepsilon)} \leq C(\varepsilon) \quad (5.18)$$

where $Q_\varepsilon := \mathbb{T}^d \times (\varepsilon, T - \varepsilon)$.

Remark 5.9. The proposition could also be proved for data depending on time, in particular with $f(x, m)$ and $H(x, \xi)$ replaced by $f(t, x, m)$ and $H(t, x, \xi)$, respectively. The only additional assumption needed would be a Lipschitz estimate in t , where the Lipschitz constant can depend on x (but not on m or ξ).

Proof. Let $\varepsilon \in \mathbb{R}$ be small and $\eta : [0, T] \rightarrow [0, 1]$ be smooth and compactly supported (with the support of η denoted $\text{spt}(\eta)$) in $(0, T)$ such that $|\varepsilon| < \min\{\text{dist}(0, \text{spt}(\eta)); \text{dist}(T, \text{spt}(\eta))\}$ and $\max_t |\varepsilon \eta'(t)| < 1$. If $\varepsilon > 0$ we set $\eta_\varepsilon(t) = t + \varepsilon \eta(t)$, which is a strictly increasing bijection from $[0, T]$ to itself. Then we set $\eta_{-\varepsilon} = \eta_\varepsilon^{-1}$, which is also smooth by the inverse function theorem. For competitors (u, P, γ) of the minimization problem for $\mathcal{A}$, let us define

$$u^\varepsilon(x, t) := u(x, \eta_\varepsilon(t)); \quad P^\varepsilon(t) = P(\eta_\varepsilon(t)); \quad \gamma^\varepsilon(x, t) := \eta'_\varepsilon(t) \gamma(x, \eta_\varepsilon(t)).$$

Notice that by construction, if $t \in \{0, T\}$ then $u(x, t) = u^\varepsilon(x, t)$ and $\gamma(x, t) = \gamma^\varepsilon(x, t)$, provided that $\gamma(x, t)$ is well-defined.

Similarly, for competitors (m, w) of minimization problem for $\mathcal{B}$, we define

$$m^\varepsilon(x, t) := m(x, \eta_\varepsilon(t)); \quad w^\varepsilon(x, t) := \eta'_\varepsilon(t) w(x, \eta_\varepsilon(t))$$

and here as well if $t \in \{0, T\}$ then $m(x, t) = m^\varepsilon(x, t)$ and $w(x, t) = w^\varepsilon(x, t)$.

We define moreover perturbations on the data as

$$\begin{aligned} \Phi^\varepsilon(t, v) &:= \eta'_\varepsilon(t) \Phi(v / \eta'_\varepsilon(t)), \\ f^\varepsilon(t, x, m) &:= \eta'_\varepsilon(t) f(x, m); \quad F^\varepsilon(t, x, m) := \eta'_\varepsilon(t) F(x, m), \end{aligned}$$

from which the Legendre transforms w.r.t. the last variable satisfy

$$(\Phi^\varepsilon)^*(t, P) = \eta'_\varepsilon(t) \Phi(P), \quad (F^\varepsilon)^*(t, x, \gamma) := \eta'_\varepsilon(t) F^*(x, \gamma / \eta'_\varepsilon(t)).$$

Finally, we define

$$H^\varepsilon(t, x, \xi) := \eta'_\varepsilon(t) H(x, \xi), \quad \text{thus } (H^\varepsilon)^*(x, \zeta) := \eta'_\varepsilon(t) H^*(x, \zeta / \eta'_\varepsilon(t)).$$

Step 1. Take a smooth minimizing sequence (u_n, P_n, γ_n) in $\mathcal{K}_0$. Use $u_n^{\pm\varepsilon}$ as a test function in $\partial_t m + \nabla \cdot w = 0$ to get

$$\begin{aligned} & \int_{\mathbb{T}^d} (u_T m(T) - u_n(0) m_0) \, dx \\ & \geq \iint_Q (H^\varepsilon(t, x, Du_n^\varepsilon + G^\top P^\varepsilon(t)) m - \gamma_n^\varepsilon m + Du_n^\varepsilon \cdot w) \, dx \, dt \quad (5.19) \end{aligned}$$

and

$$\begin{aligned} & \int_{\mathbb{T}^d} (u_T m(T) - u_n(0) m_0) \, dx \\ & \geq \iint_Q (H^{-\varepsilon}(t, x, Du_n^{-\varepsilon} + G^\top P^{-\varepsilon}(t)) m - \gamma_n^{-\varepsilon} m + Du_n^{-\varepsilon} \cdot w) \, dx \, dt. \end{aligned} \quad (5.20)$$

Take (5.19)+(5.20)-2(5.10) to get

$$\begin{aligned}
& \int_{\mathbb{T}^d} 2(u(0) - u_n(0)) m_0 \, dx \\
& \geq \iint_Q \left(H^\varepsilon(x, Du_n^\varepsilon + G^\top P^\varepsilon) + H^* \left(x, -\frac{w}{m} \right) + (Du_n^\varepsilon + G^\top P^\varepsilon) \cdot \frac{w}{m} \right) m \, dx \, dt \\
& \quad + \iint_Q \left(H^{-\varepsilon}(x, Du_n^{-\varepsilon} + G^\top P^{-\varepsilon}) + H^* \left(x, -\frac{w}{m} \right) \right. \\
& \quad \left. + (Du_n^{-\varepsilon} + G^\top P^{-\varepsilon}) \cdot \frac{w}{m} \right) m \, dx \, dt \\
& \quad + \iint_Q (2f(x, t, m) - \gamma_n^\varepsilon - \gamma_n^{-\varepsilon}) m \, dx \, dt \\
& \quad + \iint_Q (2P \cdot (Gw) - P^\varepsilon \cdot (Gw) - P^{-\varepsilon} \cdot (Gw)) \, dx \, dt. \tag{5.21}
\end{aligned}$$

Letting $n \rightarrow \infty$ we get

$$\begin{aligned}
& \iint_Q \left(H^\varepsilon(x, Du^\varepsilon + G^\top P^\varepsilon) + (H^*)^\varepsilon \left(x, -\frac{w}{m} \right) + (Du^\varepsilon + G^\top P^\varepsilon) \cdot \frac{w}{m} \right) m \, dx \, dt \\
& \quad + \iint_Q \left(H^{-\varepsilon}(x, Du^{-\varepsilon} + G^\top P^{-\varepsilon}) + (H^*)^{-\varepsilon} \left(x, -\frac{w}{m} \right) + (Du^{-\varepsilon} + G^\top P^{-\varepsilon}) \cdot \frac{w}{m} \right) m \, dx \, dt \\
& \quad + \iint_Q (2f(x, m) - f^\varepsilon(t, x, m^\varepsilon) - f^{-\varepsilon}(t, x, m^{-\varepsilon})) m \, dx \, dt \\
& \quad + \iint_Q (2P \cdot (Gw) - P^\varepsilon \cdot (Gw) - P^{-\varepsilon} \cdot (Gw)) \, dx \, dt \leq R(\varepsilon) \tag{5.22}
\end{aligned}$$

where

$$R(\varepsilon) := \iint_Q \left((H^*)^\varepsilon \left(x, -\frac{w}{m} \right) + (H^*)^{-\varepsilon} \left(x, -\frac{w}{m} \right) - 2H^* \left(x, -\frac{w}{m} \right) \right) \, dx \, dt. \tag{5.23}$$

Arguing as in [18, Proposition 3.3, Step 1] and using the estimate on $D_{xx}^2 H^*$, we have $R(\varepsilon) = O(\varepsilon^2)$.

Next we perform changes of variables and the relation $P = D_v \Phi(t, \int Gw)$ to rewrite

$$\begin{aligned}
& \iint_Q (2P \cdot (Gw) - P^\varepsilon \cdot (Gw) - P^{-\varepsilon} \cdot (Gw)) \, dx \, dt = \iint_Q (P^\varepsilon - P) \cdot (Gw^\varepsilon - Gw) \, dx \, dt \\
& = \int_0^T (P^\varepsilon - P) \cdot ((\Psi^{-1})^\varepsilon(t, P^\varepsilon) - \Psi^{-1}(t, P)) \, dt. \tag{5.24}
\end{aligned}$$

Using the same argument as in [18, Proposition 3.3, Step 4], Assumption 5.7 implies

$$\begin{aligned}
& \int_0^T (P^\varepsilon - P) \cdot \left((\Psi^{-1})^\varepsilon(t, P^\varepsilon) - (\Psi^{-1})(t, P) \right) dt \\
& \geq \frac{c_\Psi}{2} \int_0^T \min \{ |P^\varepsilon(t)|, |P(t)| \}^{s'-2} |P^\varepsilon(t) - P(t)|^2 dt - C |\varepsilon|^2 \int_0^T |P(t)|^{s'} dt.
\end{aligned} \tag{5.25}$$

We use an analogous argument (or see [18, Proposition 3.3, Step 4]) we deduce

$$\begin{aligned}
& \iint_Q (2f(x, m) - f^\varepsilon(t, x, m^\varepsilon) - f^{-\varepsilon}(t, x, m^{-\varepsilon})) m dx dt \\
& = \iint_Q (f^\varepsilon(t, x, m^\varepsilon) - f(x, m)) (m^\varepsilon - m) dx dt \\
& \geq \frac{c_f}{2} \iint_Q \min \{ m^\varepsilon, m \}^{q-2} |m^\varepsilon - m|^2 dx dt - C |\varepsilon|^2 \iint_Q m^q dx dt.
\end{aligned} \tag{5.26}$$

Finally, using Assumption 5.3 we get

$$\begin{aligned}
& \iint_Q \left(H^\varepsilon(x, Du^\varepsilon + G^\top P^\varepsilon) + (H^*)^\varepsilon \left(x, -\frac{w}{m} \right) \right. \\
& \quad \left. + (Du^\varepsilon + G^\top P^\varepsilon) \cdot \frac{w}{m} \right) m dx dt \\
& + \iint_Q \left(H^{-\varepsilon}(x, Du^{-\varepsilon} + G^\top P^{-\varepsilon}) + (H^*)^{-\varepsilon} \left(x, -\frac{w}{m} \right) \right. \\
& \quad \left. + (Du^{-\varepsilon} + G^\top P^{-\varepsilon}) \cdot \frac{w}{m} \right) m dx dt \\
& \geq c_H \iint_Q \left| j_1(Du^\varepsilon + G^\top P^\varepsilon) - j_2 \left(\frac{w}{m} \right) \right|^2 m dx dt \\
& + c_H \iint_Q \left| j_1(Du^{-\varepsilon} + G^\top P^{-\varepsilon}) - j_2 \left(\frac{w}{m} \right) \right|^2 m dx dt \\
& \geq \frac{c_H}{2} \iint_Q |j_1(Du^\varepsilon + G^\top P^\varepsilon) - j_1(Du^{-\varepsilon} + G^\top P^{-\varepsilon})|^2 m dx dt
\end{aligned} \tag{5.27}$$

Combining (5.23), (5.25), (5.26), and (5.27) with (5.22), we get

$$\begin{aligned}
& \frac{c_H}{2} \iint_Q |j_1(Du^\varepsilon + G^\top P^\varepsilon) - j_1(Du^{-\varepsilon} + G^\top P^{-\varepsilon})|^2 m dx dt \\
& + \frac{c_f}{2} \iint_Q \min \{ m^\varepsilon, m \}^{q-2} |m^\varepsilon - m|^2 dx dt \\
& + \frac{c_\Psi}{2} \int_0^T \min \{ |P^\varepsilon(t)|, |P(t)| \}^{s'-2} |P^\varepsilon(t) - P(t)|^2 dt \leq C |\varepsilon|^2, \tag{5.28}
\end{aligned}$$

where we have used the estimates on $\iint_Q m^q dx dt$ and $\int_0^T |P(t)|^{s'} dt$. The conclusion follows. $\square$

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- [1] Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84**(3), 375–393 (2000)
- [2] Bensoussan, A., Frehse, J., Yam, P.: *Mean Field Games and Mean Field Type Control Theory*. Springer, Berlin (2013)
- [3] Bonnans, J.F., Gianatti, J., Pfeiffer, L.: A Lagrangian approach for aggregative mean field games of controls with mixed and final constraints (2021)
- [4] Bonnans, J.F., Hadikhanloo, S., Pfeiffer, L.: Schauder estimates for a class of potential mean field games of controls. *Appl. Math. Optim.* **83**, 1–34 (2019)
- [5] Cardaliaguet, P.: Weak solutions for first order mean field games with local coupling. In: Bettiol, P., Cannarsa, P., Colombo, G., Motta, M., Rampazzo, F. (eds.) *Analysis and Geometry in Control Theory and its Applications*. Springer INdAM Series, vol 11. Springer, Cham. https://doi.org/10.1007/978-3-319-06917-3_5
- [6] Cardaliaguet, P., Graber, P.J.: Mean field games systems of first order. *ESAIM COCV* **21**(3), 690–722 (2015)
- [7] Cardaliaguet, P., Graber, P.J., Porretta, A., Tonon, D.: Second order mean field games with degenerate diffusion and local coupling. *Nonlinear Differ. Equ. Appl. NoDEA* **22**(5), 1287–1317 (2015)
- [8] Cardaliaguet, P., Lehalle, C.-A.: Mean field game of controls and an application to trade crowding. *Math. Financ. Econ.* **12**, 335–363 (2018)
- [9] Carmona, R., Delarue, F. (2017) Probabilistic theory of mean field games: vol. I, mean field FBSDEs, control, and games. In: *Stochastic Analysis and Applications*. Springer, Berlin
- [10] Carmona, R., Delarue, F.: Probabilistic theory of mean field games: vol. II, mean field games with common noise and master equations. In: *Stochastic Analysis and Applications*. Springer (2017)
- [11] Chan, P., Sircar, R.: Bertrand and Cournot mean field games. *Appl. Math. Optim.* **71**(3), 533–569 (2015)
- [12] Chan, P., Sircar, R.: Fracking, renewables, and mean field games. *SIAM Rev.* **59**(3), 588–615 (2017)
- [13] DiPerna, R.J., Lions, P.-L.: Ordinary differential equations, transport theory and sobolev spaces. *Invent. Math.* **98**(3), 511–547 (1989)

- [14] Gomes, D.A., Patrizi, S., Voskanyan, V.: On the existence of classical solutions for stationary extended mean field games. *Nonlinear Anal. Theory Methods Appl.* **99**, 49–79 (2014)
- [15] Gomes, D.A., Voskanyan, V.K.: Extended deterministic mean-field games. *SIAM J. Control Optim.* **54**(2), 1030–1055 (2016)
- [16] Graber, P.J., Bensoussan, A.: Existence and uniqueness of solutions for Bertrand and Cournot mean field games. *Appl. Math. Optim.* **77**(1), 47–71 (2018)
- [17] Graber, P.J., Mészáros, A.R.: Sobolev regularity for first order mean field games. *Annales de l’Institut Henri Poincaré C, Analyse non linéaire* **35**(6), 1557–1576 (2018)
- [18] Graber, P.J., Mészáros, A.R., Silva, F.J., Tonon, D.: The planning problem in mean field games as regularized mass transport. *Calc. Var. Partial Differ. Equ.* **58**(3), 115 (2019)
- [19] Graber, P.J., Mouzouni, C.: Variational mean field games for market competition. In: Cardaliaguet, P., Porretta, A., Salvarani, F. (eds.) *PDE Models for Multi-Agent Phenomena*. Springer INdAM Series, vol. 28. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01947-1_5
- [20] Graber, P.J., Mouzouni, C.: On mean field games models for exhaustible commodities trade. *ESAIM COCV* **26**, 11 (2020)
- [21] Guéant, O., Lasry, J.-M., Lions, P.-L.: Mean field games and applications. In: *Paris-Princeton lectures on mathematical finance 2010*, pp. 205–266. Springer, Berlin, Heidelberg (2011)
- [22] Huang, M., Malhamé, R.P., Caines, P.E.: Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the nash certainty equivalence principle. *Commun. Inf. Syst.* **6**(3), 221–252 (2006)
- [23] Ishii, H.: On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions. *Funkc. Ekvac* **38**(1), 101–120 (1995)
- [24] Jameson Graber, P., Ignazio, V., Neufeld, A.: Nonlocal Bertrand and Cournot mean field games with general nonlinear demand schedule. *J. Math. Pures Appl.* **148**, 150–198 (2021). <https://doi.org/10.1016/j.matpur.2021.02.002>
- [25] Kobeissi, Z.: On classical solutions to the mean field game system of controls (2019)
- [26] Kobeissi, Z.: Mean field games with monotonous interactions through the law of states and controls of the agents (2020)
- [27] Lasry, J.-M., Lions, P.-L.: Mean field games. *Jpn. J. Math.* **2**(1), 229–260 (2007)

P. Jameson Graber and Alan Mullenix
Department of Mathematics
Baylor University
One Bear Place #97328
Waco TX 76798-7328
USA
e-mail: Alan.Mullenix@baylor.edu

P. Jameson Graber
e-mail: Jameson_Graber@baylor.edu

Laurent Pfeiffer
Inria and CMAP (UMR 7641), CNRS, Ecole Polytechnique
Institut Polytechnique de Paris
Route de Saclay
91128 Palaiseau
France
e-mail: laurent.pfeiffer@inria.fr

Received: 2 November 2020.

Accepted: 30 June 2021.