HAL
open science

First order Mean Field Games with density constraints: Pressure equals Price

Pierre Cardaliaguet, Alpár Richárd Mészáros, Filippo Santambrogio

To cite this version:

Pierre Cardaliaguet, Alpár Richárd Mészáros, Filippo Santambrogio. First order Mean Field Games with density constraints: Pressure equals Price. 2015. hal-01173947v1

HAL Id: hal-01173947

https://hal.science/hal-01173947v1

Preprint submitted on 7 Jul 2015 (v1), last revised 6 Dec 2016 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FIRST ORDER MEAN FIELD GAMES WITH DENSITY CONSTRAINTS: PRESSURE EQUALS PRICE

PIERRE CARDALIAGUET, ALPÁR R. MÉSZÁROS, AND FILIPPO SANTAMBROGIO

Abstract. In this paper we study Mean Field Game systems under density constraints as optimality conditions of two optimization problems in duality. A weak solution of the system contains an extra term, an additional price imposed on the saturated zones. We show that this price corresponds to the pressure field from the models of incompressible Euler's equations *à la Brenier*. By this observation we manage to obtain a minimal regularity, which allows to write optimality conditions at the level of single agent trajectories and to define a weak notion of Nash equilibrium for our model.

1. Introduction

1.1. **The MFG system.** Introduced by Lasry-Lions [22, 23, 24] (see also Huang-Malhamé-Caines [21]) the mean field game system (in short, MFG system) describes a differential game with infinitely many identical players who interact through their repartition density. The first order MFG system with a local coupling takes the form

$$\begin{cases} \text{(i)} & -\partial_t u + H(x, Du) = f(x,m) & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(ii)} & \partial_t m - \operatorname{div}(mD_pH(x,Du)) = 0 & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(iii)} & u(T,x) = g(x), \quad m(0,x) = m_0(x) & \text{in } \mathbb{T}^d. \end{cases} \tag{1.1}$$

Here, to avoid the discussion of the boundary data, we work for simplicity with periodic boundary conditions, i.e., in the torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. The Hamiltonian $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is typically convex with respect to the last variable and the coupling cost $f : \mathbb{T}^d \times [0, +\infty)$ is nondecreasing with respect to the last variable.

Let us briefly describe the interpretation of (1.1). In the above backward-forward system, $u = u(t,x)$ is the value function associated to any tiny player while $m = m(t,x)$ is the density of players at time t and at position x. The value function $u(t,x)$ is formally given by

$$u(t,x) = \inf_\gamma \int_t^T L(\gamma(s), \dot\gamma(s)) + f(\gamma(s), m(s,\gamma(s))\,\mathrm{d}s + g(\gamma(T))$$

where the player minimizes over the paths $\gamma : [t,T] \to \mathbb{T}^d$ with $\gamma(t) = x$, $f = f(x, m(t,x))$ is the running cost, L is obtained from the Fenchel conjugate of H with respect to the last variable, and $g : \mathbb{T}^d \to \mathbb{R}$ is the terminal cost at the terminal time $t = T$. The running cost f couples the two equations and acts as a penalization for those regions where the density m is too high.

At the initial time $t = 0$, the initial distribution is m_0 (a probability measure on $\mathbb{T}^d$). Then the density evolvs according to the motion of the players. Since – by standard argument in optimal control – it is optimal for the players to play $\dot\gamma(s) = -D_pH(\gamma(s), Du(s,\gamma(s))$, the evolution of the density is given by the continuity equation (1.1)-(ii).

Note that each tiny players acts as if he knew the evolution of the players' density $m = m(t,x)$ (he somehow "forecasts" it, as usually in a "rational expectations" framework). The mean field game system corresponds to an equilibrium situation in where the "forecast" of the players is correct: the solution of the continuity equation is indeed $m = m(t,x)$. In terms of game theory, this corresponds to a Nash equilibrium.

Date: July 8, 2015.

Existence and uniqueness of solutions for the above problem are discussed by P.-L. Lions in [25] (through a reduction to an elliptic equation in time-space when the coefficients are smooth, under the additional assumption $\lim_{m\to 0} f(x,m) = -\infty$, which is satisfied for instance for log-like couplings, which guarantees $m > 0$ and hence ellipticity) and in Cardaliaguet [11], Graber [20], Cardaliaguet-Graber [13], Cardaliaguet-Porretta-Tonon [15] (following an approach by variational methods suggested in [24] and also inspired by Benamou-Brenier [4]). Recently, in [5] Benamou and Carlier used similar variational techniques to study an augmented Lagrangian scheme for MFG problems and obtain efficient numerical simulations.

1.2. **The problem with a density constraint.** In this paper we study the behavior of the MFG system when there is a *density constraint*, i.e., when the density m cannot exceed some given value $\overline{m} > 1/|\mathbb{T}^d| = 1$. Namely: $0 \le m(t,x) \le \overline{m}$ at any point (t,x). In other words, the players pay an infinite price when the density goes above $\overline{m}$: $f(x,m) = +\infty$ if $m > \overline{m}$. The question of how to model this situation was first introduced by Santambrogio [28] and then investigated Mészáros and Silva in [27] in the framework of stationary second order models. We emphasize the fact that imposing a density constraint will result in a so-called "hard congestion" effect in the model. Models of MFGs where so-called "soft congestion" (meaning that agents slow down when they reach zones with high density) effects have been studied recently by Gomes and Mitake in [18], by Gomes and Voskanyan in [19] and by Burger, Di Francesco, Markowich and Wolfram in [10].

Coming back to our model, there are several issues in the interpretation of system (1.1) when there is a density constraints. Indeed, the above interpretation does not make sense anymore for the following reason: if, on the one hand, the constraint $m \le \overline{m}$ is fulfilled, then the minimization problem of the agents (due to the fact that they are considered negligible against the others) does not see this constraint and the pair (u,m) is the solution of a standard MFG system; but this solution has no reason to satisfy the constraint, and there is a contradiction. On the other hand, if there are places where $m(t,x) > \overline{m}$, then the players do not go through these places because their cost is infinite there: but then the density at such places is zero, and there is again a contradiction. So, in order to understand the MFG system with a density constraint, one has to change the point of view. We shall see that there are several ways to understand more deeply the phenomena behind this question. We warn the reader that he model that we will obtain significantly differ from that in [28].

Perhaps the simplest approach is to go through an approximation argument: let us consider the solution $(u^\varepsilon, m^\varepsilon)$ corresponding to a running cost f^ε which is finite everywhere, but tends to infinity as ε tends to 0 when $m > \overline{m}$. In orther words, $f^\varepsilon(x,m) \to f(x,m)$ if $m \le \overline{m}$ and $f(x,m) \to +\infty$ if $m > \overline{m}$, as $\varepsilon \to 0$. In this case the MFG system with a density constraint should simply be the limit configuration (a limit which should be proven to be well-defined).

We indeed show that the pair $(u^\varepsilon, m^\varepsilon)$ has (up to subsequences) a limit (u,m) which satisfies (in a weak sense) the following system:

$$\begin{cases} \text{(i)} & -\partial_t u + H(x,Du) = f(x,m) + \beta & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(ii)} & \partial_t m - \operatorname{div}(mD_pH(x,Du)) = 0 & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(iii)} & u(T,x) = g(x) + \beta_T, \quad m(0,x) = m_0(x) & \text{in } \mathbb{T}^d \\ \text{(iv)} & 0 \le m \le \overline{m} & \text{in } [0,T]\times\mathbb{T}^d \end{cases} \tag{1.2}$$

Besides the expected density constraint (iv), two extra terms appear: β in (i) and β_T in (iii). These two quantities turn out to be nonnegative and concentrated on the set $\{m = \overline{m}\}$. They formally correspond to an extra price payed by the players to go through zones where the concentration is saturated, i.e., where $m = \overline{m}$ (in traffic language, this would be a *toll*). In other words, the new optimal control problem for the players is now (formally)

$$u(t,x) = \inf_{\substack{\gamma \\ \gamma(t)=x}} \int_t^T L(\gamma(s),\dot\gamma(s)) + f(\gamma(s),m(s,\gamma(s)) + \beta(s,\gamma(s))\,\mathrm{d}s + g(\gamma(T)) + \beta_T(\gamma(T)), \tag{1.3}$$

and thus (still formally) satisfies the dynamic programming principle: for any $0 \le t_1 \le t_2 < T$,

$$u(t_1, x) = \inf_{\substack{\gamma \\ \gamma(t_1)=x}} \int_{t_1}^{t_2} L(\gamma(s), \dot\gamma(s)) + f(\gamma(s), m(s, \gamma(s)) + \beta(s, \gamma(s))\, \mathrm{d}s + u(t_2, \gamma(t_2)). \tag{1.4}$$

The "extra prices" β and β_T discourage too many players to be attracted by the area where the constraint is saturated, thus ensuring the density constraints (iv) to be fulfilled.This is exactly the typical role of prices in theoretical economics: a price is a quantity determined by a global configuration, which replaces, in the individual choices of the agents, the presence of the constraint.

1.3. **The variational method.** Another way to see the problem is the following: it is known (see [24]) that the solution (u, m) to (1.1) can be obtained by variational methods at least when f is finite everywhere. More precisely, the value function u is (formally) given as a minimizer of the functional

$$\mathcal{A}(u) := \int_0^T \int_{\mathbb{T}^d} F^*(x, -\partial_t u + H(x, Du))\, \mathrm{d}x\, \mathrm{d}t - \int_{\mathbb{T}^d} u(0, x)\, \mathrm{d}m_0(x),$$

subject to the constraint that $u(T, x) = g(x)$, where $F = F(x, m)$ is an antiderivative of $f = f(x, m)$ with respect to m and F^* is its Legendre-Fenchel conjugate w.r.t. the second variable. In the same way m is (formally) given as a minimizer of the problem

$$\mathcal{B}(m, w) := \int_{\mathbb{T}^d} g(x) m(T, x)\, \mathrm{d}x + \int_0^T \int_{\mathbb{T}^d} m(t, x) H^* \left(x, -\frac{w}{m}\right) + F(x, m(t, x))\ \mathrm{d}x\, \mathrm{d}t$$

subject to the constraint

$$\partial_t m + \operatorname{div}(w) = 0 \quad \text{in } (0, T) \times \mathbb{T}^d, \qquad m(0) = m_0,$$

where H^* is the convex conjugate of H with respect to the last variable. It turns out that both problems make perfectly sense, even when $f(x, m) = +\infty$ if $m > \overline{m}$. In fact, if f^ε is a finite approximation of f as before, one can expect the minimizers of $\mathcal{A}^\varepsilon$ and $\mathcal{B}^\varepsilon$ (corresponding to f^ε) to converge to the minimizers of $\mathcal{A}$ and $\mathcal{B}$ as $\varepsilon \to 0$ (as a simple consequence of $\Gamma-$convergence). This is precisely what happens. Note that, as $f(x, m) = +\infty$ for $m > \overline{m}$, $F(x, m)$ has the same property, so that $F^*(x, m)$ is linear on $[\overline{m}, +\infty)$. This linear behavior explains the appearance of the terms β and β_T described above.

1.4. **Connections between MFGs with density constraints and the incompressible Euler's equations à la Brenier.** It is not surprising, due to the constraint $m \le \overline{m}$, that some strong connections between our model and the variational models for the incompressible Euler's equations studied by Brenier (see [7]) and also by Ambrosio and Figalli (see [2]) arise. What was unexpected at the beginning of our study is the role that this connection would play in reguarity. In order to understand the analogy, notice that the incompressibility constraint in the model of Brenier to study perfect fluids is what introduces the pressure field. Morally the same effect happens imposing density constraint for MFG. Using the common variational structure, similar also to the one introduced by Benamou and Brenier in [4], shared by the incompressible Euler equation and by our model, we can easily interpret the terms β and β_T, that we call "additional prices/costs" for the agents (appearing only if they pass through saturated zones) in (1.2) as a sort of pressure field from fluid mechanics. This observation motivates the title of our work as well.

Using similar techniques as in [7] and [2, 3] we show that β is an

$$L^2_{\mathrm{loc}}((0, T); BV(\mathbb{T}^d)) \hookrightarrow L^{d/(d-1)}_{\mathrm{loc}}((0, T) \times \mathbb{T}^d)$$

function (while a priori it was only supposed to be a measure) and β_T is $L^1(\mathbb{T}^d)$. With the help of an example we show that this local integrability cannot be extended so as to include the final time $t = T$, which shows that the result is somehow sharp. This regularity property will allow us to give a clearer (weak) meaning to the control problem (1.4), obtaining optimality conditions along single agent trajectories. Our techniques to proceed with the analysis rely on the properties of measures

defined on paths, that we shall call *density-constrained flows* in our context, and we are exploiting some properties of a Hardy-Littlewood type maximal functional as well (this is very much inspired by [2]).

After this analysis we deduce the existence of a *local weak Nash equilibrium* for our model.

The paper is organized in the following way. We first introduce our main notation and assumptions (Section 2). Then we discuss the two optimization problems for $\mathcal{A}$ and $\mathcal{B}$ described above (Section 3). We introduce the definition of the MFG system with a density constraint, present our main existence result as well as the approximation by standard MFG systems in Section 4. In Section 5, by means of an example, we study some finer properties of a solution (m, u, β, β_T) of the MFG system with density constraints. Section 6 is devoted to the proof of the $L^{d/(d-1)}_{\text{loc}}$ integrability of the additional price β under some additional assumptions on the Hamiltonian and the coupling. Finally, having in hand this integrability property, we introduce in Section 7 the optimal density-constrained flows and derive optimality conditions along single agent paths, which allow in particular to study the existence of the local weak Nash equilibrium.

Acknowledgement: The first author was partially supported by the ANR (Agence Nationale de la Recherche) projects ANR-10-BLAN 0112, ANR-12-BS01-0008-01 and ANR-14-ACHN-0030-01. The second and third authors were partly supported by the ANR project ANR-12-MONU-0013 and by the iCODE project "strategic crowds" of the IDEX *Université Paris-Saclay.*

2. Notations, assumptions and preliminaries

We consider the MFG system with a density constraint (1.2) under the assumption that all the maps are periodic in space. Typical conditions are

(H1) The density constraint $\overline{m}$ is larger than $1 = 1/|\mathbb{T}^d|$.

(H2) (Conditions on the initial and final conditions) m_0 is a probability measure on $\mathbb{T}^d$ which is absolutely continuous with respect to Lebesgue measure and there exists $\overline{c} > 0$ such that $0 \le m_0 < \overline{m} - \overline{c}$ a.e. on $\mathbb{T}^d$. We assume that $g : \mathbb{T}^d \to \mathbb{R}$ is a C^1 function on $\mathbb{T}^d$.

(H3) (Conditions on the Hamiltonian) $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is continuous in both variables, convex and differentiable in the second variable, with D_pH continuous in both variables. Moreover, H has superlinear growth in the gradient variable: there exist $r > 1$ and $C > 0$ such that

$$\frac{1}{rC}|p|^r - C \le H(x,p) \le \frac{C}{r}|p|^r + C. \tag{2.1}$$

We denote by $H^*(x,\cdot)$ the Fenchel conjugate of $H(x,\cdot)$, which, due to the above assumptions, satisfies

$$\frac{1}{r'C}|q|^{r'} - C \le H^*(x,q) \le \frac{C}{r'}|q|^{r'} + C, \tag{2.2}$$

where r' is the conjugate of r. We will also denote by L the Lagrangian given by $L(x,q) = H^*(x,-q)$, which thus satisfies the same bounds as H^*.

(H4) (Conditions on the coupling) Let f be continuous on $\mathbb{T}^d \times [0,\overline{m}]$, non-decreasing in the second variable with $f(x,0) = 0$.

We define F so that $F(x,\cdot)$ is an antiderivative of $f(x,\cdot)$ on $[0,\overline{m}]$, that is,

$$F(x,m) = \int_0^m f(x,s)\,\mathrm{d}s, \quad \forall\, m \in [0,\overline{m}], \tag{2.3}$$

and extend F to $+\infty$ on $(-\infty,0) \times (\overline{m},+\infty)$. It follows that F is continuous on $\mathbb{T}^d \times [0,\overline{m}]$, is convex and differentiable in the second variable. We also define $F^*(x,\cdot)$ to be the Fenchel conjugate of $F(x,\cdot)$ for each x. Note that

$$F^*(x,\alpha) \ge \alpha\overline{m} - F(x,\overline{m}) \tag{2.4}$$

and $F^*(\cdot,\alpha)=0$ for all $\alpha\le 0$. Following the approach of Cardaliaguet-Carlier-Nazaret [12] (see also Cardaliaguet [11], Graber [20] or Cardaliaguet-Graber [13]) it seems that the solution to (1.2) can be obtained as the system of optimality conditions for optimal control problems.

2.1. **Optimal transport toolbox.** In this subsection we collect some basic notions and results from the theory of optimal transportation which we will need in the sequel. We refer to [29, 30] general references to this theory. Let $\Omega\subset\mathbb{R}^d$ be a compact subset (or any compact subset of a Polish space). Even if in the whole paper we will restrict ourselves to the case of $\Omega=\mathbb{T}^d$, we state the following results in the general case. Given two probability measures $\mu,\nu\in\mathcal{P}(\Omega)$ and for $p\ge 1$ we define the usual Wasserstein metric by means of the Monge-Kantorovich optimal transportation problem

$$W_p(\mu,\nu):=\inf\left\{\int_{\Omega\times\Omega}|x-y|^p\,\mathrm{d}\gamma(x,y)\ :\ \gamma\in\Pi(\mu,\nu)\right\}^{\frac{1}{p}},$$

where $\Pi(\mu,\nu):=\{\gamma\in\mathcal{P}(\Omega\times\Omega):\quad (\pi^x)_\#\gamma=\mu,\ (\pi^y)_\#\gamma=\nu\}$ and π^x and π^y denote the canonical projections from $\Omega\times\Omega$ onto Ω (in a more general setting, Ω being any compact subset of a Polish space, in the definition of W_p one has to replace the Euclidean distance $|x-y|$ by the distance induced by the underlining metric d). This quantity happens to be a distance on $\mathcal{P}(\Omega)$ which metrizes the weak-$\star$ convergence of probability measures; we denote by $\mathbb{W}_p(\Omega)$ the space of probabilities on Ω endowed with this distance.

Historically, the quadratic case $p=2$ had been understood first. So, let us state the most fundamental results in this case. Under the additional assumption $\mu\ll\mathcal{L}^d\llcorner\Omega$ (the $d-$dimensional Lebesgue measure on Ω) Brenier showed (see [8, 9]) that the optimal γ in the above problem is actually induced by a map, which turns out to be the gradient of a convex function, i.e. there exists $T:\Omega\to\Omega$ and $\psi:\Omega\to\mathbb{R}$ convex such that $T=\nabla\psi$ and $\overline{\gamma}:=(\mathrm{id},T)_\#\mu$. The function ψ is obtained as $\psi(x)=\frac{1}{2}|x|^2-\varphi(x)$, where φ is the so-called Kantorovich potential for the transport from μ to ν, and is characterized as the solution of a dual problem. In this way, the optimal transport map T can also be written as $T(x)=x-\nabla\varphi(x)$. Later, McCann (see [26]) introduced a useful notion of interpolation between probability measures: the curve $\mu_t:=((1-t)x+ty)_\#\,\overline{\gamma}$, for $t\in[0,1]$, gives a constant speed geodesic in the Wasserstein space connecting $\mu_0:=\mu$ and $\mu_1:=\nu$.

Based on this notion of interpolation, Benamou and Brenier using also some ideas from fluid mechanics, gave a dynamical formulation to the Monge-Kantorovich problem (see [4]). They showed that

$$\frac{1}{p}W_p^p(\mu,\nu)=\inf\left\{\mathcal{B}_p(E_t,\mu_t)\ :\ \partial_t\mu_t+\mathrm{div}(E_t)=0,\ \mu_0=\mu,\ \mu_1=\nu\right\},$$

where $\mathcal{B}_p:\mathfrak{M}([0,1]\times\Omega)^d\times L^\infty([0,1];\mathbb{W}_p(\Omega))\to\mathbb{R}\cup\{+\infty\}$[1] is given by

$$\mathcal{B}_p(E,\mu):=\begin{cases}\displaystyle\int_0^1\int_\Omega\frac{1}{p}\left|\frac{\mathrm{d}E}{\mathrm{d}\mu}\right|^p\,\mathrm{d}\mu_t(x)\,\mathrm{d}t, & \text{if } E\ll\mu,\\ +\infty, & \text{otherwise.}\end{cases}$$

It is well-known that $\mathcal{B}_p$ is jointly convex and l.s.c. w.r.t the weak-$\star$ convergence of measures (see Section 5.3.1 in [29]) and that, if $\partial_t\mu_t+\mathrm{div}(E_t)=0$, then $\mathcal{B}_p(E,\mu)<+\infty$ implies that $t\mapsto\mu_t$ is a continuous curve, belonging to $W^{1,p}([0,1];\mathbb{W}_p(\Omega))$. In particular μ_t is well defined for all $t\in[0,1]$.

We shall use the notion of *narrow convergence* in $\mathcal{P}(\Omega)$, which is the weak-$\star$ convergence in duality with continuous and bounded functions on Ω. Since in our case Ω is compact, then $\mathcal{P}(\Omega)$ is also compact for this convergence.

[1]We denote by $\mathfrak{M}(X)$ the signed Radon measures on X. Observe that $\mu\in L^\infty([0,1];\mathbb{W}_p(\Omega))$ only means that $\mu=(\mu_t)_t$ is a time-dependent family of probability measures.

2.2. **Measures on curves and a superposition principle.** Let us denote by Γ the set of absolutely continuous curves $\gamma : [0,T] \to \mathbb{T}^d$. We denote by $\mathcal{P}(\Gamma)$ the set of Borel probability measures defined on Γ. Let us set $\mathcal{P}_r(\Gamma)$ $(r \geq 1)$ to be the subset of $\mathcal{P}(\Gamma)$ such that

$$\int_\Gamma \int_0^T |\dot\gamma(s)|^r \,\mathrm{d}s \,\mathrm{d}\boldsymbol{\eta}(\gamma) < +\infty.$$

Note that the space Γ, which is naturally endowed with the uniform convergence topology, is not compact, and hence $\mathcal{P}(\Gamma)$ is not compact for the narrow convergence. Yet, Prokhorov Theorem guarantees that any family of probability measures on a Polish space Ω is relatively sequentially compact w.r.t. the narrow convergence provided it is *tight*. Tight means that for any $\varepsilon > 0$ there exists a compact set $K \subseteq \Omega$ such that for any μ from this family $\mu(\Omega \setminus K) < \varepsilon$. In the case of $\mathcal{P}(\Gamma)$, it is easy to see that a uniform bound on $\displaystyle\int_\Gamma \int_0^T |\dot\gamma(s)|^r \,\mathrm{d}s \,\mathrm{d}\boldsymbol{\eta}(\gamma)$ is enough to provide tightness. This will be useful later in the paper.

We also define the evaluation maps $e_t : \Gamma \to \mathbb{T}^d$, given by $e_t(\gamma) := \gamma(t)$ for all $t \in [0,T]$. This allows to state a well-known result, a connection between the solutions of the continuity equation and the measures on paths, called *superposition principle*, which can be considered as a weaker version of the DiPerna-Lions-Ambrosio theory (see for instance Theorem 8.2.1. from [1]).

Theorem 2.1. *Let $\mu : [0,T] \to \mathcal{P}(\mathbb{T}^d)$ be a narrowly continuous solution of the continuity equation $\partial_t\mu + \mathrm{div}(v\mu) = 0$, $\mu_0 \in \mathcal{P}_2(\mathbb{T}^d)$ for a velocity field $v : (0,T)\times\mathbb{T}^d \to \mathbb{R}^d$ satisfying $\displaystyle\int_0^T \int_{\mathbb{T}^d} |v_t|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t < +\infty$. Then there exists $\boldsymbol{\eta} \in \mathcal{P}(\Gamma)$ such that*

(i) *$\mu_t = (e_t)_\#\boldsymbol{\eta}$ for all $t \in [0,T]$;*
(ii) *we have the energy inequality*

$$\int_\Gamma \int_0^T |\dot\gamma(t)|^2 \,\mathrm{d}t \,\mathrm{d}\boldsymbol{\eta}(\gamma) \leq \int_0^T \int_{\mathbb{T}^d} |v|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t;$$

(iii) *$\dot\gamma(t) = v_t(\gamma(t))$, for $\boldsymbol{\eta}-$a.e. γ and a.e. $t \in [0,T]$.*

3. Optimal control problems

We look in this section at two PDE control problems that will lead to our Mean Field Game model.

The first one is an optimal control problem of Hamilton-Jacobi equations: denote by $\mathcal{K}_P$ the set of functions $u \in C^1([0,T]\times\mathbb{T}^d)$ such that $u(T,x) = g(x)$ (the subscript P stands for "primal"). Let us define on $\mathcal{K}_P$ the functional

$$\mathcal{A}(u) = \int_0^T \int_{\mathbb{T}^d} F^*(x, -\partial_t u + H(x,Du)) \,\mathrm{d}x \,\mathrm{d}t - \int_{\mathbb{T}^d} u(0,x) \,\mathrm{d}m_0(x). \tag{3.1}$$

Then we have our first optimal control problem.

Problem 3.1 (Optimal control of HJ)**.** *Find* $\inf_{u\in\mathcal{K}_P} \mathcal{A}(u)$.

It is easy to check that one can restrict the optimization to the class of minimizers such that $-\partial_t u + H(x,Du) \geq 0$, because $F^*(x,\alpha) = 0$ for $\alpha \leq 0$ (see Lemma 3.2 in [11]).

The second problem is an optimal control problem for the continuity equation: define $\mathcal{K}_D$ to be the set of all pairs $(m,w) \in L^1([0,T]\times\mathbb{T}^d) \times L^1([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$ such that $m \geq 0$ almost everywhere, $\displaystyle\int_{\mathbb{T}^d} m(t,x)\,\mathrm{d}x = 1$ for a.e. $t \in [0,T]$ (the subscript D stands for "dual"), and

$$\begin{cases} \partial_t m + \mathrm{div}(w) &= 0 \quad \text{in } (0,T)\times\mathbb{T}^d \\ m(0,\cdot) &= m_0 \quad \text{in } \mathbb{T}^d. \end{cases}$$

in the sense of distributions. Because of the integrability assumption on w, it follows that $t \mapsto m(t)$ has a unique narrowly continuous representative (cf. [1]). It is to this representative that we refer when we write $m(t)$, and thus $m(t)$ is well-defined as a probability density for all $t \in [0,T]$.

Define the functional

$$\mathcal{B}(m,w) = \int_{\mathbb{T}^d} g(x)m(T,x)\,\mathrm{d}x + \int_0^T \int_{\mathbb{T}^d} m(t,x)L\left(x, \frac{w(t,x)}{m(t,x)}\right) + F(x, m(t,x))\,\mathrm{d}x\,\mathrm{d}t \tag{3.2}$$

on $\mathcal{K}_D$. Recall that L is defined just after (2.2). We follow the convention that

$$mL\left(x, \frac{w}{m}\right) = \begin{cases} +\infty, & \text{if} \quad m = 0 \text{ and } w \neq 0, \\ 0, & \text{if} \quad m = 0 \text{ and } w = 0. \end{cases} \tag{3.3}$$

Since $m \geq 0$, the second integral in (3.2) is well-defined in $(-\infty, \infty]$ by the assumptions on F and L. The first integral is well-defined and necessarily finite by the continuity of g and the fact that $m(T,x)\,\mathrm{d}x$ is a probability measure.

We next state the "dual problem" as

Problem 3.2 (Dual Problem)**.** *Find* $\inf_{(m,w)\in\mathcal{K}_D} \mathcal{B}(m,w)$.

Proposition 3.3. *Problems 3.1 and 3.2 are in duality, i.e.*

$$\inf_{u\in\mathcal{K}_P} \mathcal{A}(u) = -\min_{(m,w)\in\mathcal{K}_D} \mathcal{B}(m,w) \tag{3.4}$$

Moreover, the minimum on the right-hand side is achieved by a pair $(m,w) \in \mathcal{K}_D$ *with* $m \in L^\infty([0,T]\times\mathbb{T}^d)$ *and* $w \in L^{r'}([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$.

Proof. The proof relies on the Fenchel-Rockafellar duality theorem (see for example [17]) and basically follows the lines of the proof of Lemma 2.1 from [11], hence we omit it. The integrability of (m,w) is just coming from the density constraint and from the growth condition of H^*. □

Remark 3.4. *If f is strictly increasing with respect to the second variable in $\mathbb{T}^d \times (0,\overline{m})$, then the minimizer (m,w) is unique.*

In general one cannot expect Problem 3.1 to have a solution. This motivates us to relax it and search for solutions in a larger class. For this let us first state the following observation.

Lemma 3.5. *Let (u_n) be a minimizing sequence for Problem 3.1 and set $\alpha_n = -\partial_t u_n + H(x, Du_n)$. Then (u_n) is bounded in $BV([0,T]\times\mathbb{T}^d)\cap L^r([0,T]\times\mathbb{T}^d)$, the sequence (α_n) is bounded in $L^1([0,T]\times\mathbb{T}^d)$, with $\alpha_n \geq 0$ a.e., while (Du_n) is bounded in $L^r([0,T]\times\mathbb{T}^d)$. Finally, there exists a Lipschitz continuous function $\psi : [0,T]\times\mathbb{T}^d \to \mathbb{R}$ such that $\psi(T,\cdot) = g$ and $u_n \geq \psi$ for any n.*

Proof. As $F^*(\cdot,\alpha) = 0$ for $\alpha \leq 0$, we can assume without loss of generality that $\alpha_n \geq 0$ (indeed, if we replace α_n with its positive part, the part with F^* does not increase, and the value of $u_n(0,x)$ does not decrease, by the maximum principle applied to the HJ equation connecting u_n to α_n). By comparison, $u_n \geq \psi$ where ψ is the unique Lipschitz continuous viscosity solution to

$$\begin{cases} -\partial_t\psi + H(x, D\psi) &= 0 & \text{in } (0,T)\times\mathbb{T}^d \\ \psi(T,x) &= g(x) & \text{in } \mathbb{T}^d. \end{cases}$$

So (u_n) is uniformly bounded from below. Integrating the equation for (u_n) on $[0,T]\times\mathbb{T}^d$ and using the fact that $H \geq -C$ and the fact that g is bounded, we get (up to redefining the constant $C > 0$)

$$\int_{\mathbb{T}^d} u_n(0,x)\,\mathrm{d}x \leq \int_0^T \int_{\mathbb{T}^d} \alpha_n\,\mathrm{d}x\,\mathrm{d}t + C.$$

So, by (2.4) and for n large enough,

$$\begin{aligned}\inf_{u\in\mathcal{K}_P}\mathcal{A}(u)+1 &\geq \int_0^T\int_{\mathbb{T}^d} F^*(x,\alpha_n)\,\mathrm{d}x\,\mathrm{d}t-\int_{\mathbb{T}^d}u_n(0,x)m_0\,\mathrm{d}x\\ &\geq \int_0^T\int_{\mathbb{T}^d}\overline{m}\alpha_n\,\mathrm{d}x\,\mathrm{d}t-\int_{\mathbb{T}^d}u_n(0,x)m_0\,\mathrm{d}x-C\\ &\geq \int_{\mathbb{T}^d}u_n(0,x)(\overline{m}-m_0)\,\mathrm{d}x-C.\end{aligned}$$

By **(H2)** $\overline{m}-m_0>\overline{c}$: as we know that $u_n(0,\cdot)$ is bounded from below, we get that $(u_n(0,\cdot))$ is bounded in $L^1(\mathbb{T}^d)$. Thus, as $\alpha_n\geq 0$, we also have that (α_n) is bounded in $L^1([0,T]\times\mathbb{T}^d)$. Then integrating the equation $\alpha_n=-\partial_t u_n+H(x,Du_n)$ over $[t,T]\times\mathbb{T}^d$ and using the lower bound on H, we get on the one hand

$$\int_{\mathbb{T}^d}u_n(t,x)\,\mathrm{d}x\leq\int_t^T\int_{\mathbb{T}^d}\alpha_n\,\mathrm{d}x\,\mathrm{d}t+C,$$

which, in view of the lower bound on u_n, gives an $L^\infty([0,T])$ bound on $\langle u_n(t,\cdot)\rangle=\int_{\mathbb{T}^d}u_n(t,x)\,\mathrm{d}x$. We integrate again the equation $\alpha_n=-\partial_t u_n+H(x,Du_n)$ over $[0,T]\times\mathbb{T}^d$ and use the coercivity of H and Poincaré's inequality to get

$$\begin{aligned}C\ &\geq\int_0^T\int_{\mathbb{T}^d}\alpha_n\,\mathrm{d}x\,\mathrm{d}t+C\ \geq\ \int_0^T\int_{\mathbb{T}^d}H(x,Du_n)\,\mathrm{d}x\,\mathrm{d}t\ \geq\ (1/C)\int_0^T\int_{\mathbb{T}^d}|Du_n|^r\,\mathrm{d}x\,\mathrm{d}t-C\\ &\geq(1/C)\int_0^T\int_{\mathbb{T}^d}|u_n-\langle u_n(t,\cdot)\rangle|^r\,\mathrm{d}x\,\mathrm{d}t-C\ \geq\ (1/C)\int_0^T\int_{\mathbb{T}^d}|u_n|^r\,\mathrm{d}x\,\mathrm{d}t-C.\end{aligned}$$

In particular (Du_n) and (u_n) are bounded in $L^r([0,T]\times\mathbb{T}^d)$. Thus $\partial_t u_n=-\alpha_n+H(x,Du_n)$ is bounded in $L^1([0,T]\times\mathbb{T}^d)$. The result follows. □

By the results of Lemma 3.5 we introduce a relaxation of the Problem 3.1. Let us denote by $\mathcal{K}_R$ (here the subscript R stands for "relaxed") the set of pairs (u,α) such that $u\in BV([0,T]\times\mathbb{T}^d)$ with $Du\in L^r([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$ and $u(T^-,\cdot)\geq g$ a.e., α is a nonnegative measure on $[0,T]\times\mathbb{T}^d$, and, if we extend (u,α) by setting $u=g$ and $\alpha:=H(\cdot,Dg)\,\mathrm{d}x\,\mathrm{d}t$ on $(T,T+1)\times\mathbb{T}^d$, then the pair (u,α) satisfies

$$-\partial_t u+H(x,Du)\leq\alpha$$

in the sense of distribution in $(0,T+1)\times\mathbb{T}^d$. Note that the extension of (u,α) to $[0,T+1]\times\mathbb{T}^d$ just expresses the fact that $u(T^+)=g$ and that α compensates the possible jump from $u(T^-)$ to g. We set

$$\mathcal{A}(u,\alpha)=\int_0^T\int_{\mathbb{T}^d}F^*(x,\alpha^{ac}(t,x))\,\mathrm{d}x\,\mathrm{d}t+\overline{m}\alpha^s([0,T]\times\mathbb{T}^d)-\int_{\mathbb{T}^d}u(0^+,x)\ \mathrm{d}m_0(x).$$

where α^{ac} and α^s are respectively the absolutely continuous part and the singular part of the measure α.

Problem 3.6 (Relaxed Problem)**.** *Find* $\inf\limits_{(u,\alpha)\in\mathcal{K}_R}\mathcal{A}(u,\alpha)$.

Let us consider the following result as a counterpart of Lemma 2.7 from [13] in our case.

Lemma 3.7. *Let $(m,w)\in\mathcal{K}_D$ such that $m\in L^\infty([0,T]\times\mathbb{T}^d)$ and $(u,\alpha)\in\mathcal{K}_R$ an arbitrary competitor for Problem 3.6. Then, for every $t\in[0,T]$, we have*

$$\begin{aligned}\int_0^t\int_{\mathbb{T}^d}-mH^*\left(x,-\frac{w}{m}\right)\,\mathrm{d}x\,\mathrm{d}t\leq&\int_{\mathbb{T}^d}m(t,x)u(t^-,x)\,\mathrm{d}x-\int_{\mathbb{T}^d}m_0(x)u(0^+,x)\,\mathrm{d}x\\ &+\int_0^t\int_{\mathbb{T}^d}\alpha^{ac}m\,\mathrm{d}x\,\mathrm{d}t+\overline{m}\alpha^s([0,t]\times\mathbb{T}^d)\end{aligned}$$

and

$$\int_t^T \int_{\mathbb{T}^d} -mH^*\left(x, -\frac{w}{m}\right) \,\mathrm{d}x\,\mathrm{d}t \le \int_{\mathbb{T}^d} m(T,x)g(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} m(t,x)u(t^+,x)\,\mathrm{d}x + \int_t^T \int_{\mathbb{T}^d} \alpha^{ac} m\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([t,T]\times\mathbb{T}^d).$$

Moreover we can take $t=0$ in the above inequalities. If, finally, equality holds in the second inequality when $t=0$, then $w = -mD_pH(\cdot, Du)$ a.e. and

$$\limsup_{\varepsilon\to 0} m_\varepsilon(t,x) = \overline{m} \qquad for\ \alpha^s - a.e.\ (t,x)\in[0,T]\times\mathbb{T}^d,$$

where m_ε is any standard mollification of m.

Proof. We prove the result only for $t=0$, the general case follows by a similar (and simpler) argument. We first extend the pairs (u,α) and (m,w) to $(0,T+1)\times\mathbb{T}^d$ by setting $u=g$ and $\alpha := H(\cdot, Dg)\,\mathrm{d}x\,\mathrm{d}t$, $m(s,x)=m(T,x)$, $w(s,x)=0$ on $(T,T+1)\times\mathbb{T}^d$. Note that

$$\partial_t m + \mathrm{div}(w) = 0 \quad \text{and} \quad -\partial_t u + H(x,Du) \le \alpha \quad \text{on} \quad (0,T+1)\times\mathbb{T}^d.$$

We smoothen the pair (m,w) in a standard way into $(m_\varepsilon, w_\varepsilon)$: $m_\varepsilon := m\star\rho_\varepsilon$ and $w_\varepsilon := w\star\rho_\varepsilon$, where the mollifier ρ has a support in the unit ball of $\mathbb{R}^{d+1}$ and $\rho_\varepsilon := \varepsilon^{-d-1}\rho(\cdot/\varepsilon)$. Then, for any $\eta>\varepsilon$, we have, since $m_\varepsilon \le \overline{m}$,

$$\begin{aligned}\int_\eta^{T+\eta}\int_{\mathbb{T}^d} u\partial_t m_\varepsilon + m_\varepsilon H(x,Du)\,\mathrm{d}x\,\mathrm{d}t - \left[\int_{\mathbb{T}^d} m_\varepsilon u\,\mathrm{d}x\right]_{\eta^+}^{T+\eta} \\ \le \int_\eta^{T+\eta}\int_{\mathbb{T}^d} m_\varepsilon\,\mathrm{d}\alpha \le \int_\eta^{T+\eta}\int_{\mathbb{T}^d}\alpha^{ac}m_\varepsilon\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([\eta,T+\eta]\times\mathbb{T}^d)\end{aligned} \tag{3.5}$$

where, as $\partial_t m_\varepsilon + \mathrm{div}(w_\varepsilon) = 0$,

$$\int_\eta^{T+\eta}\int_{\mathbb{T}^d} u\partial_t m_\varepsilon\,\mathrm{d}x\,\mathrm{d}t = \int_\eta^{T+\eta}\int_{\mathbb{T}^d} Du\cdot w_\varepsilon\,\mathrm{d}x\,\mathrm{d}t.$$

In the followings, we shall consider only such η's for which no jump of u occurs, in particular $\int_{\mathbb{T}^d} u(\eta^-,x)\,\mathrm{d}x = \int_{\mathbb{T}^d} u(\eta^+,x)\,\mathrm{d}x = \int_{\mathbb{T}^d} u(\eta,x)\,\mathrm{d}x$. So, by convexity of H, (3.5) and the above equality,

$$\begin{aligned}\int_\eta^{T+\eta}\int_{\mathbb{T}^d} -m_\varepsilon H^*\left(x, -\frac{w_\varepsilon}{m_\varepsilon}\right)\,\mathrm{d}x\,\mathrm{d}t \ &\le \ \int_\eta^{T+\eta}\int_{\mathbb{T}^d} w_\varepsilon\cdot Du + m_\varepsilon H(x,Du)\,\mathrm{d}x\,\mathrm{d}t \\ &\le \ \left[\int_{\mathbb{T}^d} m_\varepsilon u\,\mathrm{d}x\right]_\eta^{T+\eta} + \int_\eta^{T+\eta}\int_{\mathbb{T}^d}\alpha^{ac}m_\varepsilon\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([\eta,T+\eta]\times\mathbb{T}^d).\end{aligned}$$

We multiply the inequality $-\partial_t u + H(x,Du)\le\alpha$ by $m_\varepsilon(\eta,\cdot)$ and integrate on $(0,\eta)\times\mathbb{T}^d$ to get, as m_ε is bounded by $\overline{m}$ and H is bounded from below,

$$\int_{\mathbb{T}^d} u(0^+,x)m_\varepsilon(\eta,x)\,\mathrm{d}x \le \int_{\mathbb{T}^d} u(\eta,x)m_\varepsilon(\eta,x)\,\mathrm{d}x + C\eta + \overline{m}\alpha((0,\eta)\times\mathbb{T}^d).$$

Note that $\overline{m}\alpha((0,\eta)\times\mathbb{T}^d) = \overline{m}\alpha^s([0,\eta)\times\mathbb{T}^d) + o(1)$, where $o(1)\to 0$ as $\eta\to 0$. So

$$\begin{aligned}&\int_\eta^{T+\eta}\int_{\mathbb{T}^d} -m_\varepsilon H^*\left(x, -\frac{w_\varepsilon}{m_\varepsilon}\right)\,\mathrm{d}x\,\mathrm{d}t \le \\ &\le \int_{\mathbb{T}^d}[m_\varepsilon(T+\eta,x)u(T+\eta,x) - u(0^+,x)m_\varepsilon(\eta,x)]\,\mathrm{d}x \\ &\qquad + \int_\eta^{T+\eta}\int_{\mathbb{T}^d}\alpha^{ac}m_\varepsilon\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0,T+\eta]\times\mathbb{T}^d) + o(1).\end{aligned}$$

We now let $\varepsilon \to 0$. By convergence of $(m_\varepsilon, w_\varepsilon)$ to (m, w) in $L^q \times L^{r'}$, $\forall q \geq 1$, and by the fact that

$$\lim_{\varepsilon\to 0} \int_\eta^{T+\eta} \int_{\mathbb{T}^d} m_\varepsilon H^*(x, -w_\varepsilon/m_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t = \int_\eta^{T+\eta} \int_{\mathbb{T}^d} mH^*(x, -w/m)\,\mathrm{d}x\,\mathrm{d}t,$$

(see the proof of Lemma 2.7. from [13]) we obtain for all $\eta > 0$, chosen above,

$$\begin{aligned}
&\int_\eta^{T+\eta} \int_{\mathbb{T}^d} -mH^*\left(x, -\frac{w}{m}\right)\,\mathrm{d}x\,\mathrm{d}t \leq \\
&\leq \int_{\mathbb{T}^d} [m(T+\eta, x)u(T+\eta, x) - u(0^+, x)m(\eta, x)]\,\mathrm{d}x \\
&\qquad + \int_\eta^{T+\eta} \int_{\mathbb{T}^d} \alpha^{ac} m\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0, T+\eta] \times \mathbb{T}^d) + o(1).
\end{aligned}$$

By definition of the extension of the maps u and m,

$$\begin{aligned}
&\int_{\mathbb{T}^d} [m(T+\eta)u(T+\eta) - u(0^+, x)m(\eta, x)]\,\mathrm{d}x + \int_\eta^{T+\eta} \int_{\mathbb{T}^d} \alpha^{ac} m\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0, T+\eta] \times \mathbb{T}^d) \\
&\quad = \int_{\mathbb{T}^d} [m(T, x)g(x) - u(0^+, x)m(\eta, x)]\,\mathrm{d}x + \int_\eta^{T+\eta} \int_{\mathbb{T}^d} \alpha^{ac} m\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0, T] \times \mathbb{T}^d).
\end{aligned}$$

We finally let $\eta \to 0$ and get

$$\begin{aligned}
&\int_0^T \int_{\mathbb{T}^d} -mH^*\left(x, -\frac{w}{m}\right)\,\mathrm{d}x\,\mathrm{d}t \\
&\quad \leq \int_{\mathbb{T}^d} [m(T, x)g(x) - u(0^+, x)m_0(x)]\,\mathrm{d}x + \int_0^T \int_{\mathbb{T}^d} \alpha^{ac} m\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0, T] \times \mathbb{T}^d)
\end{aligned}$$

thanks to the $L^\infty-$weak-$\star$ continuity of $t \mapsto m(t)$ and the L^1 integrability of $u(0^+, \cdot)$.

The proof of the equality $w = -mD_pH(\cdot, Du)$ when equality holds in the above inequality follows exactly the proof of the corresponding statement in [13], so we omit it. Note that, if equality holds, then all the above inequalities must become equalities as ε and then η tend to 0. In particular, from inequality (3.5), we must have

$$\lim_{\eta\to 0} \limsup_{\varepsilon\to 0} \int_\eta^{T+\eta} \int_{\mathbb{T}^d} m_\varepsilon(t, x)\,\mathrm{d}\alpha^s(t, x) = \overline{m}\alpha^s([0, T] \times \mathbb{T}^d).$$

By Fatou's lemma, this implies that

$$\overline{m}\alpha^s([0, T] \times \mathbb{T}^d) \leq \int_0^T \int_{\mathbb{T}^d} \limsup_{\varepsilon\to 0} m_\varepsilon(t, x)\,\mathrm{d}\alpha^s(t, x),$$

where the right-hand side is also bounded above by the left-hand side since $m_\varepsilon \leq \overline{m}$. So $\limsup_{\varepsilon\to 0} m_\varepsilon = \overline{m}$ α^s-a.e. □

Proposition 3.8. *We have*

$$\inf_{u\in\mathcal{K}_P} \mathcal{A}(u) = \min_{(u,\alpha)\in\mathcal{K}_R} \mathcal{A}(u, \alpha). \tag{3.6}$$

Moreover, if (u, α) is a minimum of $\mathcal{A}$, then $\alpha = \alpha \llcorner [0, T) \times \mathbb{T}^d + (u(T^-, \cdot) - g)\,\mathrm{d}\left(\delta_T \otimes \mathcal{H}^d \llcorner \mathbb{T}^d\right)$.

Proof. We follow here [20]. Inequality $\geq$ is obvious and we now prove the reverse one. Let us fix $(u,\alpha)\in\mathcal{K}_R$ and let (m,w) be an optimal solution for the dual problem. Then

$$\begin{aligned}\mathcal{A}(u,\alpha) &= \int_0^T\int_{\mathbb{T}^d} F^*(x,\alpha^{ac})\,\mathrm{d}x\,\mathrm{d}t+\overline{m}\alpha^s([0,T]\times\mathbb{T}^d)-\int_{\mathbb{T}^d} m_0(x)u(0^+,x)\,\mathrm{d}x\\ &\geq \int_0^T\int_{\mathbb{T}^d}(m\alpha^{ac}-F(x,m))\,\mathrm{d}x\,\mathrm{d}t+\overline{m}\alpha^s([0,T]\times\mathbb{T}^d)-\int_{\mathbb{T}^d} m_0(x)u(0^+,x)\,\mathrm{d}x\\ &\geq \int_0^T\int_{\mathbb{T}^d}(-mH^*(x,-w/m)-F(x,m))\,\mathrm{d}x\,\mathrm{d}t-\int_{\mathbb{T}^d} m(T,x)g(x)\,\mathrm{d}x \;=\; -\mathcal{B}(m,w)\end{aligned}$$

where the last inequality comes from Lemma 3.7. By optimality of (m,w) and (3.4) we obtain therefore

$$\mathcal{A}(u,\alpha)\geq\inf_{u\in\mathcal{K}_P}\mathcal{A}(u),$$

which shows equality (3.6).

To prove that the problem in the right-hand side of (3.6) has a minimum we consider a minimizing sequence (u_n) for Problem 3.1. We extend $u_n=g$ on $(T,T+1]\times\mathbb{T}^d$ and set $\alpha_n:=-\partial_t u_n+H(x,Du_n)$ on $[0,T+1]\times\mathbb{T}^d$ and note that, in view of Lemma 3.5, there is a subsequence, again denoted by (u_n,α_n), such that (u_n) converges in L^1 to a BV map u, (Du_n) converges weakly in L^r, and (α_n) converges in sense of measures to α on $[0,T+1]\times\mathbb{T}^d$. As $u_n\geq\psi$ on $[0,T]\times\mathbb{T}^d$, we also have $u\geq\psi$ $(0,T)\times\mathbb{T}^d$, so that $u(T^-,\cdot)\geq\psi(T,\cdot)=g$. By convexity of H with respect to p, the pair (u,α) belongs to $\mathcal{K}_R$. One easily shows by standard relaxation that

$$\mathcal{A}(u,\alpha)\leq\liminf_{n\to\infty}\mathcal{A}(u_n).$$

Hence (u,α) is a minimum.

Let us finally check that $\alpha=\alpha\llcorner(0,T)\times\mathbb{T}^d+(u(T^-,\cdot)-g)\,\mathrm{d}\left(\delta_T\otimes\mathcal{H}^d\llcorner\mathbb{T}^d\right)$. Indeed, by definition of $\mathcal{K}_R$, we can extend (u,α) by setting $(u,\alpha):=(g,H(\cdot,Dg))$ on $(T,T+1)\times\mathbb{T}^d$ and the following inequality holds in the sense of measure in $(0,T+1)\times\mathbb{T}^d$:

$$-\partial_t u+H(x,Du)\leq\alpha$$

Let $\phi\in C^\infty(\mathbb{T}^d)$, with $\phi\geq 0$, be a test function. We multiply the above inequality by ϕ and integrate on $(T-\eta,T+\eta)\times\mathbb{T}^d$ to get

$$\int_{\mathbb{T}^d}\phi(x)u((T-\eta)^+,x)\,\mathrm{d}x+\int_{T-\eta}^{T+\eta}\int_{\mathbb{T}^d}\phi H(x,Du)\,\mathrm{d}x\,\mathrm{d}t\leq\int_{\mathbb{T}^d}g\phi\,\mathrm{d}x+\int_{T-\eta}^{T+\eta}\int_{\mathbb{T}^d}\phi\,\mathrm{d}\alpha.$$

Letting $\eta\to 0$ along a suitable sequence such that $u((T-\eta)^+,\cdot)\to u(T^-,\cdot)$ in L^1, we obtain:

$$\int_{\mathbb{T}^d}\phi u(T^-,x)\,\mathrm{d}x\leq\int_{\mathbb{T}^d}g\phi\,\mathrm{d}x+\int_{\mathbb{T}^d}\phi\,\mathrm{d}(\alpha\llcorner\{T\}\times\mathbb{T}^d).$$

This means that $u(T^-,\cdot)\leq g+\alpha\llcorner\{T\}\times\mathbb{T}^d$. Let us now replace α by

$$\tilde{\alpha}:=\alpha\llcorner(0,T)\times\mathbb{T}^d+(u(T^-,\cdot)-g)\,\mathrm{d}\left(\delta_T\otimes\mathcal{H}^d\llcorner\mathbb{T}^d\right).$$

We claim that the pair $(u,\tilde{\alpha})$ still belongs to $\mathcal{K}_R$. For this we just have to check that, if we extend $(u,\tilde{\alpha})$ to $(0,T+1)\times\mathbb{T}^d$ as before, then

$$-\partial_t u+H(x,Du)\leq\tilde{\alpha}\text{ on }(0,T+1)\times\mathbb{T}^d$$

holds in the sense of distributions. Let $\phi \in C_c^\infty((0,T+1)\times\mathbb{T}^d)$ with $\phi \geq 0$. Then

$$\begin{aligned}\int_0^{T+1}\int_{\mathbb{T}^d} u\partial_t\phi + \phi H(x,Du)\,\mathrm{d}x\,\mathrm{d}t =& \\ &= \int_0^T\int_{\mathbb{T}^d} u\partial_t\phi + \phi H(x,Du)\,\mathrm{d}x\,\mathrm{d}t + \int_T^{T+1}\int_{\mathbb{T}^d} g\partial_t\phi + \phi H(x,Dg)\,\mathrm{d}x\,\mathrm{d}t \\ &\leq \int_0^T\int_{\mathbb{T}^d} \phi\,\mathrm{d}(\alpha \llcorner (0,T)\times\mathbb{T}^d) + \int_{\mathbb{T}^d}(u(T^-,x)-g)\phi(T,x)\,\mathrm{d}x + \int_T^{T+1}\int_{\mathbb{T}^d}\phi H(x,Dg)\,\mathrm{d}x\,\mathrm{d}t \\ &\leq \int_0^{T+1}\int_{\mathbb{T}^d}\phi\,\mathrm{d}\tilde{\alpha}.\end{aligned}$$

This proves that the pair $(u,\tilde{\alpha})$ belongs to $\mathcal{K}_R$. In particular, $\mathcal{A}(u,\alpha)\leq\mathcal{A}(u,\tilde{\alpha})$, so that

$$\overline{m}\alpha^s(\{T\}\times\mathbb{T}^d) \leq \overline{m}\int_{\mathbb{T}^d} u(T^-,x)-g(x)\,\mathrm{d}x.$$

Since we have proved that $u(T^-,\cdot)\leq g+\alpha\llcorner\{T\}\times\mathbb{T}^d$, we have therefore an equality in the above inequality, which means that $\alpha\llcorner\{T\}\times\mathbb{T}^d = (u(T^-,\cdot)-g)\,\mathrm{d}x$. □

4. The MFG system with density constraints

In this section, we study the existence of solutions for the MFG system with density constraints:

$$\begin{cases} \text{(i)} & -\partial_t u + H(x,Du) = f(x,m)+\beta & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(ii)} & \partial_t m - \operatorname{div}(mD_pH(x,Du)) = 0 & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(iii)} & u(T,x)=g(x)+\beta_T, \quad m(0,x)=m_0(x) & \text{in } \mathbb{T}^d \\ \text{(iv)} & 0\leq m\leq\overline{m} & \text{in } (0,T)\times\mathbb{T}^d \end{cases} \tag{4.1}$$

under the assumptions on H, f, g and m_0 stated in Section 2. We also study the approximation of the solution of this system by the solution of the classical MFG system.

4.1. Solutions of the MFG system with density constraints.

Definition 4.1. *We say that (u,m,β,β_T) is a solution to the MFG system (4.1) if*

(1) *Integrability conditions: β is a nonnegative Radon measure on $(0,T)\times\mathbb{T}^d$, $\beta_T\in L^1(\mathbb{T}^d)$ is nonnegative, $u\in BV([0,T]\times\mathbb{T}^d)\cap L^r([0,T]\times\mathbb{T}^d)$, $Du\in L^r([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$, $m\in L^1([0,T]\times\mathbb{T}^d)$ and $0\leq m\leq\overline{m}$ a.e.,*

(2) *The following inequality*

$$-\partial_t u + H(x,Du(t,x)) \leq f(x,m)+\beta \tag{4.2}$$

holds in $(0,T)\times\mathbb{T}^d$ in the sense of measures, with the boundary condition on $\mathbb{T}^d$

$$g\leq u(T^-,\cdot) = g+\beta_T \text{ a.e.}$$

Moreover, $\beta^{ac}=0$ a.e. in $\{m<\overline{m}\}$ and

$$\limsup_{\varepsilon\to 0} m_\varepsilon(t,x) = \overline{m} \qquad \beta^s - a.e. \text{ if } t<T \text{ and a.e. in } \{\beta_T>0\} \text{ if } t=T, \tag{4.3}$$

where m_ε is any standard mollification of m.

(3) *Equality*

$$\partial_t m - \operatorname{div}(mD_pH(x,Du(t,x)) = 0, \qquad m(0)=m_0$$

holds in the sense of distribution,

(4) *Equality*

$$\int_0^T \int_{\mathbb{T}^d} m(-H(x,Du) + Du \cdot D_pH(x,Du) + f(x,m) + \beta^{ac})\,\mathrm{d}x\,\mathrm{d}t$$
$$+\overline{m}\beta^s((0,T)\times\mathbb{T}^d) + \overline{m}\int_{\mathbb{T}^d}\beta_T\,\mathrm{d}x = \int_{\mathbb{T}^d} m_0(x)u(0^+,x) - m(T,x)g(x)\,\mathrm{d}x$$

holds.

Let us recall that, in the above definition, β^{ac} and β^s denote the absolutely continuous part and the singular part of the measure β.

Some comments on the definition are now necessary. Equality (4.3) is a weak way of stating that $m = \overline{m}$ in the support of β and β_T respectively, while the last requirement formally says that equality $-\partial_t u + H(x, Du(t,x)) = f(x,m) + \beta$ holds.

We can state the main result of this section.

Theorem 4.2. *Let $(u,\alpha) \in \mathcal{K}_R$ be a solution of the relaxed Problem 3.6 and $(m,w) \in \mathcal{K}_D$ be a solution of the dual Problem 3.2. Then $\alpha \geq f(\cdot,m)$ as measures, and, if we set*

$$\beta := \alpha \llcorner [0,T)\times\mathbb{T}^d - f(\cdot,m)\,\mathrm{d}x\,\mathrm{d}t$$

and $\beta_T := \alpha \llcorner \{T\}\times\mathbb{T}^d$, the quadruplet (u,m,β,β_T) is a solution of the MFG system (4.1).

Conversely, let (u,m,β,β_T) be a solution of the MFG system (4.1). Let us set

$$\alpha := f(\cdot,m)\,\mathrm{d}x\,\mathrm{d}t + \beta + \beta_T\,\mathrm{d}(\delta_T \otimes \mathcal{H}^d \llcorner \mathbb{T}^d) \tag{4.4}$$

and $w = -mD_pH(x,Du)$. Then the pair (u,α) is a solution of the relaxed problem while the pair (m,w) is a solution of the dual problem.

The proof of this results goes along the same lines as in [13] (Theorem 3.5). However for the sake of completeness (and because of some differences) we sketch it here.

Proof. Let $(u,\alpha) \in \mathcal{K}_R$ be a solution of the Problem 3.6 and $(m,w) \in \mathcal{K}_D$ be the solution of Problem 3.2. First, by the definition of Legendre-Fenchel transform we have for a.e. $(t,x) \in [0,T]\times\mathbb{T}^d$

$$F^*(x,\alpha^{ac}(t,x)) + F(x,m(t,x)) - \alpha^{ac}(t,x)m(t,x) \geq 0. \tag{4.5}$$

On the other hand by optimality we have that

$$\begin{aligned}
0 &= \mathcal{A}(u,\alpha) + \mathcal{B}(m,w)\\
&= \int_0^T\int_{\mathbb{T}^d} F^*(x,\alpha^{ac}(t,x))\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0,T]\times\mathbb{T}^d) - \int_{\mathbb{T}^d} u(0^+,x)\,m_0(x)\,\mathrm{d}x\\
&+ \int_{\mathbb{T}^d} g(x)m(T,x)\,\mathrm{d}x + \int_0^T\int_{\mathbb{T}^d} m(t,x)H^*\left(x,-\frac{w(t,x)}{m(t,x)}\right) + F(x,m(t,x))\,\mathrm{d}x\,\mathrm{d}t\\
&\geq \int_0^T\int_{\mathbb{T}^d} \alpha^{ac}m\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0,T]\times\mathbb{T}^d) + \int_{\mathbb{T}^d} g(x)m(T,x) - u(0^+,x)\,m_0(x)\,\mathrm{d}x\\
&+ \int_0^T\int_{\mathbb{T}^d} m(t,x)H^*\left(x,-\frac{w(t,x)}{m(t,x)}\right)\,\mathrm{d}x\,\mathrm{d}t \;\geq\; 0,
\end{aligned}$$

where we used the Lemma 3.7 for the last inequality. This means that all the inequalities in the previous lines are equalities. In particular we have an equality in (4.5), which implies

$$\alpha^{ac}(t,x) \in \partial_m F(x,m(t,x)) \qquad \text{a.e.}$$

As $\partial_m F(x,m(t,x)) = \{f(x,m(t,x))\}$ for $0 < m(t,x) < \overline{m}$ a.e., we have $\alpha^{ac}(t,x) = f(x,m(t,x))$ a.e. in $\{0 < m < \overline{m}\}$. Moreover, as $\partial_m F(x,0) = (-\infty,0]$ and $\alpha^{ac} \geq 0$, we also have $\alpha^{ac} = 0 = f(\cdot,0)$ a.e. in $\{m = 0\}$. Finally, since $\partial_m F(x,\overline{m}) = [f(x,\overline{m}),+\infty)$, $\alpha^{ac} \geq f(x,m(t,x))$ a.e. on $\{m = \overline{m}\}$.

Therefore $\alpha^{ac} \geq f(\cdot, m)$ a.e. Let us set $\beta := \alpha \llcorner [0,T) \times \mathbb{T}^d - f(\cdot, m)\,\mathrm{d}x\,\mathrm{d}t$ and $\beta_T := \alpha \llcorner \{T\} \times \mathbb{T}^d$. From Proposition 3.8 we know that $\beta_T = u(T^-, \cdot) - g$.

Since equality holds in the above inequalities, there is an equality in the inequality of Lemma 3.7: thus point 4 holds in Definition 4.1. Moreover, by Lemma 3.7, we have that $w = -mD_pH(\cdot, Du)$ a.e. and (4.3) holds. In conclusion, since $(u, \alpha) \in \mathcal{K}_R$ and $(m, w) \in \mathcal{K}_D$, the quadruplet (u, m, β, β_T) is a solution to the MFG system (4.1).

Now let us prove the converse statement. For this let us take a solution (m, u, β, β_T) of the MFG system (4.1) in the sense of the Definition 4.1. Let us define α as in (4.4) and $w := -mD_pH(\cdot, Du)$. We shall prove that (u, α) is a solution for the Problem 3.6 and (m, w) is a solution of Problem 3.2. For the first one, one easily checks, following the argument of Proposition 3.8, that $(u, \alpha) \in \mathcal{K}_R$. Let us now consider a competitor $(\tilde{u}, \tilde{\alpha}) \in \mathcal{K}_R$. Using the equality in Lemma 3.7 for $(u, \alpha, m, -mD_pH(\cdot, Du))$ and the inequality for $(\tilde{u}, \tilde{\alpha}, m, -mD_pH(\cdot, Du))$ we have

$$\begin{aligned}
\mathcal{A}(\tilde{u}, \tilde{\alpha}) &= \int_0^T \int_{\mathbb{T}^d} F^*(x, \tilde{\alpha}^{ac}(x))\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\tilde{\alpha}^s([0,T] \times \mathbb{T}^d) - \int_{\mathbb{T}^d} \tilde{u}(0^+, x) m_0(x)\,\mathrm{d}x \\
&\geq \int_0^T \int_{\mathbb{T}^d} F^*(x, \alpha^{ac}(x)) + m(\tilde{\alpha}^{ac} - \alpha^{ac})\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\tilde{\alpha}^s([0,T] \times \mathbb{T}^d) - \int_{\mathbb{T}^d} \tilde{u}(0^+, x) m_0(x)\,\mathrm{d}x \\
&\geq \int_0^T \int_{\mathbb{T}^d} F^*(x, \alpha^{ac}(x))\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([0,T] \times \mathbb{T}^d) - \int_{\mathbb{T}^d} u(0^+, x) m_0(x)\,\mathrm{d}x,
\end{aligned}$$

thus (u, α) is a minimizer for the Problem 3.6.

In a similar manner we can show that (m, w) is a solution for Problem 3.2. Hence the statement of the theorem follows. □

We now briefly discuss the issue of the approximation of any solution to the MFG system with density constraints. If $F = F(x, m)$ is strictly convex on $[0, \overline{m}]$ with respect to the m variable, then, as $H^* = H^*(x, q)$ is strictly convex with respect to q (because $H = H(x, p)$ is C^1 in p), we can conclude that the dual Problem 3.2 has a unique minimizer. In particular, in this case, the m component of the MFG system (4.1) is unique. We do not expect uniqueness of the u component: this is not the case in the "classical setting", i.e., without density constraint (see, however, the discussion in [11]). For this reason, the fact that one can approximate *any* solution of the MFG system (4.1) by regular maps with suitable property is not straightforward. This is the aim of the next lemma, needed in the sequel, where we explain that the β component of any solution can be approached by a minimizing sequence of Lipschitz maps with some optimality property.

Lemma 4.3. *Let (u, m, β, β_T) be a solution to the MFG system (4.1). Then there exists Lipschitz continuous maps (u_n, α_n) such that*

(i) *u_n satisfies a.e. and in the viscosity sense,*

$$-\partial_t u_n + H(x, Du_n) = \alpha_n \text{ in } (0,T) \times \mathbb{T}^d,$$

(ii) *the pair (u_n, α_n) is a minimizing sequence for Problem 3.1 and Problem 3.6,*

(iii) *(u_n) is bounded from below, is bounded in $BV([0,T] \times \mathbb{T}^d) \cap L^r([0,T] \times \mathbb{T}^d)$ and (Du_n) is bounded in $L^r([0,T] \times \mathbb{T}^d)$,*

(iv) *(u_n) converges to some $\tilde{u} \geq u$ in $L^1((0,T) \times \mathbb{T}^d)$ with $\tilde{u} = u$ $m-$a.e.,*

(v) *(α_n) is bounded in $L^1((0,T) \times \mathbb{T}^d)$ and converges in measure on $[0,T] \times \mathbb{T}^d$ to α defined from (β, β_T) by (4.4),*

(vi) *$(\tilde{u}, m, \beta, \beta_T)$ is a solution to the MFG system (4.1),*

Proof. Let us define α as in (4.4) and recall that, by Theorem 4.2, (u, α) is a minimum in the relaxed Problem 3.6. In particular, (u, α) belongs to $\mathcal{K}_R$, which means that, if we extend (u, α) by $(g, H(\cdot, Dg(\cdot)))$ in $[T, T+1] \times \mathbb{T}^d$, then

$$-\partial_t u + H(x, Du) \leq \alpha \text{ in } (0, T+1) \times \mathbb{T}^d.$$

For $\eta \in (0,1)$ we set $u^\eta(t,x) := u(t+\eta,x)$ and $\alpha^\eta := \tau_\eta \sharp \alpha$, where $\tau_\eta : [0,T+1)\times\mathbb{T}^d \to [-\eta, T+1-\eta)\times\mathbb{T}^d$ is the time shift $\tau_\eta(t,x) = (t-\eta,x)$. We then smoothen u^η into $u^\eta \star \rho_\varepsilon$ where $\varepsilon \in (0,\eta/2)$, ρ is a standard even mollifier supported in the unit ball of $\mathbb{R}^{d+1}$ and $\rho_\varepsilon(\cdot) := \varepsilon^{-d-1}\rho(\cdot/\varepsilon)$. We note that $u^\eta \star \rho_\varepsilon(t,x)$ for $t \in [T-\varepsilon, T]$ is a mollified version of g. We finally slightly modify $u^\eta \star \rho_\varepsilon$ so that it satisfies the boundary condition: let $\zeta : \mathbb{R} \to [0,1]$ be smooth, increasing, with $\zeta(s) = 0$ for $s \le -1$ and $\zeta(s) = 1$ for $s \ge 0$. Set $\zeta_\varepsilon(s) = \zeta(\varepsilon^{-1}s)$, $u^{\eta,\varepsilon}(t,x) := (1-\zeta_\varepsilon(t-T))u^\eta \star \rho_\varepsilon(t,x) + \zeta_\varepsilon(t-T)g(x)$. Then $u^{\eta,\varepsilon}(T,x) = g(x)$ and

$$-\partial_t u^{\eta,\varepsilon} + H(x, Du^{\eta,\varepsilon}) \le \alpha^{\eta,\varepsilon} \text{ in } (0,T+1)\times\mathbb{T}^d,$$

where

$$\alpha^{\eta,\varepsilon} := \left[(1-\zeta_\varepsilon(t-T))\left(\alpha^\eta - H(\cdot, Du^\eta)\right) \star \rho_\varepsilon + H(x, Du^{\eta,\varepsilon}) - \zeta_\varepsilon'(t-T)(g(x) - g\star\rho_\varepsilon(t,x))\right]_+$$

(observe that the last convolution is done in $(d+1)$ variables, even if g is a function only depeding on x). As $\varepsilon \to 0$, $u^{\eta,\varepsilon}$ is bounded in BV and converges to u^η in L^1 while $\alpha^{\eta,\varepsilon}$ is nonnegative, bounded in L^1 and converges to α^η as a measure. We have

$$\mathcal{A}(u^{\eta,\varepsilon}, \alpha^{\eta,\varepsilon}) \;=\; \int_0^T \int_{\mathbb{T}^d} F^*\left(x, \alpha^{\eta,\varepsilon}(t,x)\right) \,\mathrm{d}x\,\mathrm{d}t - \int_{\mathbb{T}^d} u\star\rho_\varepsilon(\eta,x) m_0(x)\,\mathrm{d}x.$$

As $\varepsilon \to 0$, the first integral in the right-hand side converges to

$$\int_0^T \int_{\mathbb{T}^d} F^*(x, (\alpha^\eta)^{ac}(t,x))\,\mathrm{d}x\,\mathrm{d}t + \overline{m}(\alpha^\eta)^s((0,T]\times\mathbb{T}^d)$$
$$= \int_0^T \int_{\mathbb{T}^d} F^*(x, \alpha^{ac}(t+\eta,x))\,\mathrm{d}x\,\mathrm{d}t + \overline{m}\alpha^s([\eta, T+\eta]\times\mathbb{T}^d).$$

This convergence is technical, but not difficult: without explicit dependence on x this is just the combination of the l.s.c. behavior of this integral functional with the fact that convex functionals which are invariant by translation decrease by convolution; in the x-dependent case, one just needs to estimate the error using the regularity in x.

Pick now a sequence (η_n) tending to 0, such that $u\star\rho_\varepsilon(\eta_n,\cdot)$ converges in L^1 to $u(\eta_n,\cdot)$ as $\varepsilon \to 0$ (this is the case for a.e. η) and $(u(\eta_n,\cdot))$ tends in L^1 to $u(0^+,\cdot)$ as $n \to +\infty$: then

$$\limsup_n \limsup_{\varepsilon\to 0} \mathcal{A}(u^{\eta_n,\varepsilon}) \le \limsup_n \mathcal{A}(u^{\eta_n}, \alpha^{\eta_n}) = \mathcal{A}(u,\alpha).$$

As (u,α) is a minimum in the relaxed Problem 3.6, we can find $\varepsilon_n \to 0$ such that $(u^{\eta_n,\varepsilon_n}, \alpha^{\eta_n,\varepsilon_n})$ is a minimizing sequence for Problem 3.6 thanks to Proposition 3.8.

Let now $\tilde{u}_n$ be the viscosity solution to

$$\begin{cases} -\partial_t u + H(x, Du) = \alpha^{\varepsilon_n,\eta_n} \text{ in } (0,T)\times\mathbb{T}^d \\ u(T,x) = g(x) \text{ in } \mathbb{T}^d. \end{cases}$$

Standard results on viscosity solutions imply that $\tilde{u}_n$ is Lipschitz continuous (because so are $\alpha^{\varepsilon_n,\eta_n}$ and g), satisfies the equation a.e. and, by comparison, is such that $\tilde{u}_n \ge u^{\eta_n,\varepsilon_n}$. Therefore

$$\mathcal{A}(\tilde{u}_n, \alpha^{\varepsilon_n,\eta_n}) \le \mathcal{A}(u^{\eta_n,\varepsilon_n}, \alpha^{\eta_n,\varepsilon_n}),$$

so that $(\tilde{u}_n, \alpha^{\varepsilon_n,\eta_n})$ is also a minimizing sequence for Problem 3.6. By Lemma 3.5, $(\tilde{u}_n)$ is bounded from below, is bounded in $BV([0,T]\times\mathbb{T}^d) \cap L^r([0,T]\times\mathbb{T}^d)$ and (Du_n) is bounded in $L^r([0,T]\times\mathbb{T}^d)$. Up to a subsequence, $(\tilde{u}_n)$ converges to a BV map $\tilde{u}$ in L^1 such that $\tilde{u} \ge u$. Note that, as in the proof of Proposition 3.8, $(\tilde{u},\alpha)$ is also a minimizer of Problem 3.6, so that, by Theorem 4.2 , $(\tilde{u}, m, \beta, \beta_T)$ is also a solution to the MFG system (4.1). In particular, by (4) in the definition of solution, the inequalities of Lemma 3.7 must be equalities for $(\tilde{u},\alpha)$ and (u,α) so that,

$$\int_{\mathbb{T}^d} m(t,x)u(t,x)\,\mathrm{d}x = \int_{\mathbb{T}^d} m(t,x)\tilde{u}(t,x)\,\mathrm{d}x \qquad \text{for a.e. } t \in [0,T].$$

As $\tilde{u} \geq u$, this implies that $\tilde{u} = u$ $m-$a.e. In conclusion the pair $(\tilde{u}_n, \alpha^{\eta_n,\varepsilon_n})$ satisfies our requirements. □

4.2. **Approximation by classical MFG systems.** We now study to what extent the solution of the MFG system with density constraint introduced above can be obtained as the limit of the solutions of classical MFG systems. For this, we assume that $f^\varepsilon : \mathbb{T}^d \times [0,+\infty) \to \mathbb{R}$ is a continuous function for each $\varepsilon > 0$, increasing with respect to m, with $f^\varepsilon(\cdot, 0) = 0$, and which fulfills the growth condition: there exists $\theta > 1 + d/r$ and $C,\ C_\varepsilon > 0$ such that

$$C^{-1}m^{\theta-1} - C \leq f^\varepsilon(x,m) \leq C_\varepsilon m^{\theta-1} + C_\varepsilon.$$

We consider $(u^\varepsilon, m^\varepsilon)$ the solution to the classical MFG system

$$\begin{cases} \text{(i)} & -\partial_t u^\varepsilon + H(x, Du^\varepsilon) = f^\varepsilon(x, m^\varepsilon) & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(ii)} & \partial_t m^\varepsilon - \operatorname{div}(m^\varepsilon D_p H(x, Du^\varepsilon)) = 0 & \text{in } (0,T)\times\mathbb{T}^d \\ \text{(iii)} & u^\varepsilon(T,x) = g(x), \quad m^\varepsilon(0,x) = m_0(x) & \text{in } \mathbb{T}^d \end{cases} \tag{4.6}$$

Following Cardaliaguet [11], Cardaliaguet-Graber [13], we know that the MFG system (4.6) has a unique (weak) solution $(u^\varepsilon, m^\varepsilon)$: namely, $(u^\varepsilon, m^\varepsilon) \in C^0([0,T]\times\mathbb{T}^d) \times L^\theta([0,T]\times\mathbb{T}^d)$ and the following hold:

(i) the following integrability conditions hold:

$$Du^\varepsilon \in L^r,\ m^\varepsilon H^*(\cdot, D_pH(\cdot, Du^\varepsilon)) \in L^1 \quad \text{and} \quad m^\varepsilon D_pH(\cdot, Du^\varepsilon)) \in L^1.$$

(ii) Equation (4.6)-(i) holds in the following sense: the inequality

$$-\partial_t u^\varepsilon + H(x, Du^\varepsilon) \leq f(x, m^\varepsilon) \quad \text{in } (0,T)\times\mathbb{T}^d, \tag{4.7}$$

holds in the sense of distributions, with $u^\varepsilon(T,\cdot) = g$,

(iii) Equation (4.6)-(ii) holds:

$$\partial_t m^\varepsilon - \operatorname{div}(m^\varepsilon D_pH(x, Du^\varepsilon)) = 0 \text{ in } (0,T)\times\mathbb{T}^d, \quad m^\varepsilon(0) = m_0 \tag{4.8}$$

in the sense of distributions,

(iv) The following equality holds:

$$\begin{aligned} \int_0^T \int_{\mathbb{T}^d} m^\varepsilon(t,x)\left(f(x, m^\varepsilon(t,x)) + H^*(x, D_pH(x, Du^\varepsilon)(t,x))\right) \mathrm{d}x\,\mathrm{d}t \\ + \int_{\mathbb{T}^d} m^\varepsilon(T,x)g(x) - m_0(x)u^\varepsilon(0,x)\,\mathrm{d}x = 0. \end{aligned} \tag{4.9}$$

In addition, u^ε is Hölder continuous and in $W^{1,1}$, and equality $-\partial_t u^\varepsilon + H(x, Du^\varepsilon) = f(x, m^\varepsilon)$ holds a.e. in $(0,T)\times\mathbb{T}^d$ (see Cardaliaguet-Porretta-Tonon [15]).

Let us now suppose that $f^\varepsilon(x,m) \to f(x,m)$ uniformly with respect to x for any $m \leq \overline{m}$ and $f^\varepsilon(x,m) \to +\infty$ uniformly with respect to x for any $m > \overline{m}$ as $\varepsilon \to 0^+$.

Proposition 4.4. *Under the above assumptions,*

(1) *the family (u^ε) is bounded in $BV([0,T]\times\mathbb{T}^d) \cap L^r([0,T]\times\mathbb{T}^d)$ while (Du^ε) is bounded in $L^r([0,T]\times\mathbb{T}^d)$, the family $(\alpha^\varepsilon := -\partial_t u^\varepsilon + H(\cdot, Du^\varepsilon))$ is bounded in $L^1([0,T]\times\mathbb{T}^d)$, with $\alpha^\varepsilon \geq 0$ a.e., (m^ε) is bounded in $L^\theta([0,T]\times\mathbb{T}^d)$ while (w^ε) is bounded in $L^{r'}([0,T]\times\mathbb{T}^d)$.*

(2) *If (u,m,α) is any cluster point for the weak convergence of $(u^\varepsilon, m^\varepsilon, \alpha^\varepsilon)$, then $\alpha \geq f(\cdot, m)$ and, if we set $\beta := \alpha \llcorner (0,T)\times\mathbb{T}^d$ and $\beta_T := u(T^-,\cdot) - g$, then the quadruplet (u,m,β,β_T) is a solution of the MFG system with density constraint (4.1).*

Proof. The proof is a straightforward adaptation of our previous constructions. According to [14], we know that $(u^\varepsilon, \alpha^\varepsilon)$ is minimizer over $\mathcal{K}_R$ of the functional

$$\mathcal{A}^\varepsilon(u,\alpha) = \int_0^T \int_{\mathbb{T}^d} (F^\varepsilon)^*(x,\alpha)\,\mathrm{d}x\,\mathrm{d}t - \int_{\mathbb{T}^d} u(0^+,x)\,\mathrm{d}m_0(x),$$

where $F^\varepsilon(x,m) := \int_0^m f^\varepsilon(x,s)\,\mathrm{d}s$ and $(F^\varepsilon)^*$ is the Fenchel conjugate of F^ε with respect to the last variable. Then, by convexity,

$$(F^\varepsilon)^*(x,\alpha) \geq \alpha\overline{m} - F^\varepsilon(x,\overline{m}),$$

where, by our assumptions, $F^\varepsilon(x,\overline{m})$ converges uniformly with respect to x to $F(x,\overline{m})$. Let ψ be the Lipschitz continuous viscosity solution to

$$\begin{cases} -\partial_t\psi + H(x,D\psi) &= 0 \\ \psi(T,x) &= g(x). \end{cases}$$

It is also an a.e. solution, so that $(\psi,0)$ belongs to $\mathcal{K}_R$. Then

$$\mathcal{A}^\varepsilon(u^\varepsilon,\alpha^\varepsilon) \leq \mathcal{A}^\varepsilon(\psi,0) \leq -\int_{\mathbb{T}^d}\psi(0,x)\,\mathrm{d}m_0(x) \leq C.$$

So $(\mathcal{A}^\varepsilon(u^\varepsilon,\alpha^\varepsilon))$ is bounded from above and one can then argue exactly as in the proof of Lemma 3.5 to obtain the bounds on (u^ε) and (α^ε) as well as a bound for $(\mathcal{A}^\varepsilon(u^\varepsilon,\alpha^\varepsilon))$.

Following [14], we also know that the pair $(m^\varepsilon,w^\varepsilon) := (m^\varepsilon, -m^\varepsilon D_pH(\cdot,Du^\varepsilon))$ is a minimizer over $\mathcal{K}_D$ of

$$\mathcal{B}^\varepsilon(m,w) = \int_{\mathbb{T}^d} g(x)m(T,x)\,\mathrm{d}x + \int_0^T\int_{\mathbb{T}^d} m(t,x)L\left(x,\frac{w(t,x)}{m(t,x)}\right) + F^\varepsilon(x,m(t,x))\,\mathrm{d}x\,\mathrm{d}t$$

Since, by [14], $\mathcal{A}^\varepsilon(u^\varepsilon,\alpha^\varepsilon) = -\mathcal{B}^\varepsilon(m^\varepsilon,w^\varepsilon)$, $(\mathcal{B}^\varepsilon(m^\varepsilon,w^\varepsilon))$ is bounded. From our assumption on f^ε we have therefore that (m^ε) is bounded in $L^\theta([0,T]\times\mathbb{T}^d)$ while (w^ε) is bounded in $L^{r'}([0,T]\times\mathbb{T}^d)$.

Let finally (u,α) be a cluster point of $(u^\varepsilon,\alpha^\varepsilon)$ and (m,w) be a cluster point of $(m^\varepsilon,w^\varepsilon)$ for the weak convergence. Then, standard arguments from the theory of Γ-convergence show that (u,α) minimizes $\mathcal{A}$ while (m,w) minimizes $\mathcal{B}$, so that, if we set $\beta := \alpha \llcorner (0,T)\times\mathbb{T}^d$ and $\beta_T := u(T^-,\cdot) - g$, the quadruplet (u,m,β,β_T) is a solution of the MFG system (4.1) according to Theorem 4.2. □

5. No coupling, space homogeneity, power-like Hamiltonians and $m_0 < \overline{m}$

In this section, we study through an example some finer properties of the solutions of (4.1). Let us consider $f(x,m)\equiv 0$, $H(x,p) = \frac{1}{s}|p|^s$ $(s>1)$ and $T=1$. The terminal cost g is a given smooth function. As usual, we assume that the initial density of the population satisfies $m_0 < \overline{m} - c$ a.e. in $\mathbb{T}^d$ for a given constant $0<c<\overline{m}$ (here this assumption will be essential, while it is not clear whether for the considerations of the previous sections it is purely technical or not). For simplicity, let us set $s=2$. In this case the functional $\mathcal{B}$ for the Problem 3.2 has the form

$$\mathcal{B}(m,w) = \int_0^1\int_{\mathbb{T}^d}\frac{1}{2}\frac{|w|^2}{m} + F(x,m)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} g(x)m(1,x)\,\mathrm{d}x,$$

where we use the convention (3.3). Let us also chose $F(x,m)\equiv 0$ for $m\in[0,\overline{m}]$ and $F\equiv+\infty$ otherwise. This functional recalls the one introduced by Bemamou and Brenier to give a dynamical formulation for the Monge-Kantorovich's optimal transportation problem (see [4]). Only a constraint on the density m and a penalization on the final value have been added.

Indeed, forgetting for a while the density constraint, Problem 3.2 can be reformulated as

$$\min\left\{\frac{1}{2}W_2^2(m_0,m_1) + \int_{\mathbb{T}^d} gm_1\,\mathrm{d}x \;:\; m_1\in\mathcal{P}(\mathbb{T}^d),\; m_1\leq\overline{m}\right\}. \tag{5.1}$$

We remark that the above formulation gives always a geodesic curve connecting m_0 and m_1 (thus m_t is defined for all $t\in[0,1]$). Since the admissible set in the above problem is geodesically convex (and $\mathbb{T}^d$ is a convex set), the density constraint is satisfied as soon as it is satisfied at the terminal time. Hence the problem in (5.1) is completely equivalent to Problem 3.2. Actually we can prove something more: if the initial density satisfies strictly the constraint, then saturation may happen

only at the final time. This result is not a straightforward consequence of geodesic convexity, and we give a complete proof of it here below.

Lemma 5.1. *Let $m_0 < \overline{m} - c$ (for a given constant $0 < c < \overline{m}$) a.e. in $\mathbb{T}^d$ and m_1 be the solution of Problem 5.1. Let (m_t) be the geodesic connecting m_0 to m_1. Then, for any $t \in (0,1)$ we have $\|m_t\|_{L^\infty} \leq \overline{m}\frac{\lambda}{(1-t)+t\lambda^{1/d})^d}$, where $\lambda := (\overline{m} - c)/\overline{m} < 1$ (note $\frac{\lambda}{(1-t)+t\lambda^{1/d})^d} < 1$ for $t < 1$).*

Proof. As we mentioned before, since the admissible set in Problem 5.1 is geodesically convex, we get $m_t \leq \overline{m}$ a.e. in $\mathbb{T}^d$ for all $t \in [0,1]$. On the other hand, since m_t is absolutely continuous for all $t \in [0,1]$ we know that there exist optimal transport maps $T^t, S^t : \mathbb{T}^d \to \mathbb{T}^d$ such that $(T^t)_\# m_0 = m_t$ and $(S^t)_\# m_t = m_0$ with $T^t \circ S^t = \mathrm{id}$. The maps (S_t) and (T_t) are given by McCann's interpolation in terms of S^1 and T^1 respectively, which is $T_t := (1-t)\mathrm{id} + tT^1$ and $S_t = t\mathrm{id} + (1-t)S^1$. Moreover T_t and S_t are countably Lipschitz (i.e. the domain can be decomposed, up to negligible sets, into a countable union of sets where these maps are Lipschitz continuous), hence we can write the Jacobian equation

$$\det(DT_t) = \frac{m_0}{m_t \circ T_t}.$$

Hence, the density m_t is given by

$$m_t = \frac{m_0}{\det(DT_t)} \circ S_t. \tag{5.2}$$

Using the concavity of $\det^{1/d}$ (for positive definite matrices) we obtain that

$$\det(DT_t) = \det\left((1-t)I_d + tDT^1\right) \geq \left((1-t) + t\det(DT^1)^{1/d}\right)^d = \left((1-t) + t\left(\frac{m_0}{m_1 \circ T^1}\right)^{1/d}\right)^d$$
$$\geq \left((1-t) + t\left(\frac{m_0}{\overline{m}}\right)^{1/d}\right)^d.$$

Hence by (5.2) we have that

$$m_t \leq \frac{m_0 \circ S_t}{\left((1-t) + t\left(\frac{m_0 \circ S_t}{\overline{m}}\right)^{1/d}\right)^d}.$$

Let us set $\lambda := (\overline{m} - c)/\overline{m} < 1$. Then, for any $t \in (0,1)$ we have

$$m_t \leq \overline{m}\frac{\lambda}{(1-t) + t\lambda^{1/d})^d},$$

and the coefficient $\frac{\lambda}{(1-t)+t\lambda^{1/d})^d}$ is strictly less than 1 for every $t, \lambda < 1$. □

5.1. **Some properties of β, β_1 and u.** Let us discuss now some further properties of β, β_1 and u.

Proposition 5.2. *Let (u, m, β, β_1) be a solution of the MFG system with density constraints and let us assume that we are in the setting of this section. Then $\beta \equiv 0$ and u and β_1 are bounded.*

Proof. From Theorem 4.2, we know that the pair $(m, -mD_pH(\cdot, Du))$ is a minimizer of $\mathcal{B}$. In view of Lemma 5.1, we have therefore $m(t,x) < \overline{m}$ for a.e. $(t,x) \in (0,T) \times \mathbb{T}^d$. By Definition 4.1, this implies that $\beta^{ac} = 0$. Recall on the other hand that

$$\limsup_{\varepsilon \to 0} m_\varepsilon(t,x) = \overline{m} \qquad \beta^s - \text{a.e. if } t < 1,$$

where m_ε is any standard mollification of m. But, still by Lemma 5.1, for any $t \in (0,1)$, we get an upper bound on $m_\varepsilon(t,x)$ which is strictly less than $\overline{m}$. Hence, $\beta^s = 0$ on $[0,1) \times \mathbb{T}^d$.

Let us now check that u is bounded. We note that, as $\beta = 0$ and H satisfies the growth condition (2.1), u satisfies a.e. the inequality $-\partial_t u + |Du|^2/2 \leq 0$ in $(0,1) \times \mathbb{T}^d$. Thus, if we mollify u in the

usual way, u_ε is a classical sub solution to $-\partial_t u_\varepsilon + |Du_\varepsilon|^2/2 \le 0$ on $(\varepsilon, 1-\varepsilon)\times\mathbb{T}^d$. By Hopf's formula we get therefore

$$u_\varepsilon(t,x) \le \inf_{y\in\mathbb{T}^d}\left\{u_\varepsilon(1-\varepsilon,y) + C\frac{|x-y|^2}{(1-\varepsilon-t)} + C(1-t)\right\} \qquad \forall (t,x)\in(\varepsilon,1-\varepsilon)\times\mathbb{T}^d.$$

Hence

$$u_\varepsilon(t,x) \le \inf_{y\in\mathbb{T}^d}\{u_\varepsilon(1-\varepsilon,y)+C\} \qquad \forall (t,x)\in(\varepsilon,1/2)\times\mathbb{T}^d.$$

Recalling that $\int_{\mathbb{T}^d} u(t,x)\,\mathrm{d}x$ is bounded for a.e. t (see the proof of Lemma 3.5), we also have that $\int_{\mathbb{T}^d} u_\varepsilon(t,x)\,\mathrm{d}x$ is bounded as well for all t and therefore $\inf_{y\in\mathbb{T}^d} u_\varepsilon(1,y)$ is bounded from above. So we have proved that u_ε is bounded from above by a constant C_0 on $(\varepsilon,1/2)\times\mathbb{T}^d$, where C_0 is independent of ε. This shows that u is bounded from above by C_0 on $(0,1/2)\times\mathbb{T}^d$.

Let us set $z(t,x) := (C_0 + \|H(\cdot,0)\|_{L^\infty}) \vee \|g\|_{L^\infty} - \|H(\cdot,0)\|_{L^\infty}(1-t)$. Then z is a subsolution to $-\partial_t z + H(x,Dz) \le 0$ which satisfies $z(1,\cdot)\ge g$ and $z(0,\cdot)\ge C_0 \ge u(0,\cdot)$. Therefore the map $\tilde u(t,x) := u(t,x)\wedge z(t,x)$ is still a subsolution (because $H = H(x,p)$ is convex with respect to p), which satisfies $\tilde u(0,\cdot) = u(0,\cdot)$ a.e. and $g(x)\le \tilde u(1^-,x)\le u(1^-,x)$. Let us set $\tilde\alpha := (\tilde u(1^-,\cdot)-g)\,\mathrm{d}(\delta_1\otimes\mathcal{H}^d \llcorner \mathbb{T}^d)$. Then the pair $(\tilde u,\tilde\alpha)$ belongs to $\mathcal{K}_R$ and by optimality of (u,α) we have

$$\begin{aligned}\mathcal{A}(u,\alpha) &= \int_{\mathbb{T}^d}\big(u(1^-,x)-g(x)\big)\,\mathrm{d}x - \int_{\mathbb{T}^d} m_0(x)u(0,x)\,\mathrm{d}x\\ &\le \mathcal{A}(\tilde u,\tilde\alpha) = \int_{\mathbb{T}^d}\big(\tilde u(1^-,x)-g(x)\big)\,\mathrm{d}x - \int_{\mathbb{T}^d} m_0(x)\tilde u(0,x)\,\mathrm{d}x.\end{aligned}$$

As $\tilde u(0,\cdot) = u(0,\cdot)$ and $\tilde u(1^-,x)\le u(1^-,x)$, this proves that $\tilde u(1^-,x) = u(1^-,x)$ a.e., which means that $u(1^-,\cdot)$ is bounded from above. Since we already know that u is bounded from below (see the proof of Lemma 3.5), we have established that $u(1^-,\cdot)$ is bounded. By Hopf's formula, this entails the boundedness of u on $(0,1)\times\mathbb{T}^d$, from where the boundedness of β_1 follows as well. □

Remark 5.3 (Nash-type equilibrium)**.** *For this example a notion of Nash equilibrium can be formulated by the means of (m,β_1), i.e. by the means of the "additional price" β_1 to be payed by the agents at the final time. This price, whose value is precisely $\beta_1 = (u(1^-,\cdot)-g)$, clearly has to be payed only if agents arrive to the saturated zone at the final time. Let us postpone the precise definition and the details on the question of the Nash equilibrium, which will be established for more general cases in Section 7 (see Definition 7.9).*

6. Regularity of the "additional price" β

In this section we show, under some additional reguarity assumption on the data, that the measure β is absolutely continuous and belongs to $L^{d/(d-1)}_{\mathrm{loc}}((0,T)\times\mathbb{T}^d)$. In this respect, our model recalls those studied by Brenier (see [7]) and later by Ambrosio-Figalli (see [2, 3]), where they analyzed the motion of incompressible perfect fluids driven by the Euler's equations.

We will see in the next section that this regularity is essential in order to define Nash equilibria in our context. For this, we assume in addition to the previous hypotheses the following conditions: there exists $\lambda>0$ such that

(HP1) (Assumption for H): H and H^* are of class C^2 with

$$\lambda I_d \le D^2_{pp}H \le \lambda^{-1}I_d \ \text{ and } \ \lambda I_d \le D^2_{qq}H^* \le \lambda^{-1}I_d, \tag{6.1}$$

$$|D^2_{xx}H^*(x,p)| \le C(1+|p|^2), \qquad |D^2_{xp}H^*(x,p)| \le C(1+|p|). \tag{6.2}$$

(HP2) (Assumption on F): f is of class C^2 on $\mathbb{T}^d \times [0, \overline{m}]$ and, for any $m \in [0, \overline{m}]$ and $\alpha \geq 0$,

$$F(x,m) + F^*(x,\alpha) - \alpha m \geq \frac{\lambda}{2}|\alpha_1 - f(x,m)|^2 + p(\overline{m} - m), \tag{6.3}$$

where $p = (\alpha - f(x,\overline{m}))_+$ and $\alpha_1 = \alpha - p$.

(HP3) (Assumption on g): g is of class C^2.

Remark 6.1. *Observe that the assumption* **(HP2)** *is fulfilled if $\partial^2_{\alpha\alpha} F^*(x,\alpha) \geq \lambda$ on $(0, f(x,\overline{m}))$ for some $\lambda > 0$. This assumption of course holds if $f(x,\overline{m}) = 0$. If $f(x,\overline{m}) > 0$, since $\partial_\alpha F^*(x, f(x,m)) = m$ on $(0,\overline{m})$, the Implicit Function Theorem implies that $\partial^2_{\alpha\alpha} F^*(x,\alpha) = 1/\partial_m f(x,m)$, which means that the assumption is indeed satisfied as soon as $\partial_m f(x,m)$ is bounded from above on $(0,\overline{m})$, which makes it a very natural assumption. Among the examples that we have in mind, we underline the case where $H(x,p) = |p|^2/2 - f(x)$ and $F(x,m) = 0$ if $m \in [0,\overline{m}]$ and $+\infty$ otherwise. Notice that the same example could be written with $F(x,m) = f(x)m$ if $m \in [0,\overline{m}]$ and $+\infty$ otherwise and $H(x,p) = |p|^2/2$ (these different expression give rise to the same global and individual problems), but in this case* **(HP2)** *would not be satisfied (and also* **(H4)** *would be violated).*

Theorem 6.2. *Let (u, m, β, β_T) be a solution of the MFG system (4.1). Under the above assumptions, $f(\cdot, m(\cdot,\cdot)) \in H^1_{\mathrm{loc}}((0,T) \times \mathbb{T}^d)$ and β is absolutely continuous in $(0,T) \times \mathbb{T}^d$ with*

$$\beta \in L^2_{\mathrm{loc}}((0,T); BV(\mathbb{T}^d)) \hookrightarrow L^{d/(d-1)}_{\mathrm{loc}}((0,T) \times \mathbb{T}^d).$$

As we said, the proof is largely inspired by the works of Brenier (see [7]) and Ambrosio-Figalli (see [2]) on the incompressible Euler's equations.

Proof of the Theorem 6.2. By abuse of notion, we use $\mathcal{B}(m', v')$ meaning $\mathcal{B}(m', m'v')$ for any admissible pair $(m', m'v')$ in the dual problem (v' denoting the velocity field).

Throughout the proof, (u, m, β, β_T) is a fixed solution of the MFG system (4.1) and we define α by (4.4) and set $w = -mD_pH(x, Du)$. Recall that (m, w) is a minimizer for $\mathcal{B}$. We also set $v := w/m$ and construct competitors $(m^{\delta,\eta}, m^{\delta,\eta}v^{\delta,\eta})$ in the following way: let us fix $0 < t_1 < t_2 < T$ and let $\zeta \in C^\infty_c((0,T);[0,1])$ be a smooth cut-off such that $\zeta \equiv 1$ on $[t_1, t_2]$; for $\eta > 0$ small and $\delta \in \mathbb{R}^d$ small (such that $t + \zeta(t)\eta \in (0,T)$ for all $t \in [0,T]$), we denote

$$m^{\delta,\eta}(t,x) := m(t + \zeta(t)\eta, x + \zeta(t)\delta)$$

the time-space translation of the density and let

$$v^{\delta,\eta}(t,x) := v(t + \zeta(t)\eta, x + \zeta(t)\delta)(1 + \eta\zeta'(t)) - \zeta'(t)\delta$$

the velocity field associated to $m^{\delta,\eta}$. Indeed, by construction $(m^{\delta,\eta}, m^{\delta,\eta}v^{\delta,\eta})$ solves the continuity equation, and satisfies the other constraints.

Step 0. Let us collect some tools now.

First, we have

$$\mathcal{B}(m^{\delta,\eta}, v^{\delta,\eta}) \leq \mathcal{B}(m,v) + C(\eta^2 + |\delta|^2). \tag{6.4}$$

Indeed, let us denote by ξ_η the inverse map of $t \mapsto t + \zeta(t)\eta$. Then, after changing variables,

$$\begin{aligned}
&\mathcal{B}(m^{\delta,\eta}, v^{\delta,\eta}) \\
&\quad = \int_0^T \int_{\mathbb{T}^d} \big[m(s,y) H^* \left(y - \zeta(\xi_\eta(s))\delta, -v(s,y)(1 + \eta\zeta'(\xi_\eta(s))) + \zeta'(\xi_\eta(s))\delta \right) \\
&\qquad\qquad + F(y - \zeta(\xi_\eta(s))\delta, m(s,y)) \big] \, \xi_\eta'(s) \, \mathrm{d}y \, \mathrm{d}s + \int_{\mathbb{T}^d} g(y) m(T,y) \, \mathrm{d}y.
\end{aligned}$$

In view of our C^2 regularity assumptions on H^*, F and g, the map $(\eta, \delta) \mapsto \mathcal{B}(m^{\delta,\eta}, v^{\delta,\eta})$ is C^2. We obtain (6.4) by optimality of (m, v).

Second, by stationarity of the problem for $\mathcal{B}$ (it is enough to consider perturbations of form $(m^{0,\eta}, v^{0,\eta})$ for ζ with compact support, not necessarily 1 on $[t_1, t_2]$), we have

$$\int_{\mathbb{T}^d} \{m(H^*(x, -v(t,x)) + D_q H^*(x, -v(t,x)) \cdot v(t,x)) + F(x, m(t,x))\} \, \mathrm{d}x = \textit{constant}.$$

From our assumption on H^*, we have

$$H^*(x, -v) + D_q H^*(x, -v) \cdot v \leq H^*(x, 0) - \frac{\lambda}{2}|v|^2.$$

Thus

$$\mathrm{ess} - \sup_{t\in[0,T]} \int_{\mathbb{T}^d} m(t,x)|v(t,x)|^2 \, \mathrm{d}x \leq C. \tag{6.5}$$

By (6.1), we have $D^2_{qq} H^* \leq (1/\lambda) I_d$ and therefore (6.5) implies

$$\mathrm{ess} - \sup_{t\in[0,T]} \int_{\mathbb{T}^d} m(t,x)|D_q H^*(x, -v(t,x))|^2 \, \mathrm{d}x \leq C. \tag{6.6}$$

Third, for any smooth map (u', α'), with $\alpha' \geq 0$, and $(m', w') \in \mathcal{K}_D$ (where $v' = w'/m'$) competitor for the primal and the dual problems respectively, we have

$$\begin{aligned}\mathcal{A}(u', \alpha') + \mathcal{B}(m', v') \geq & \int_0^T \int_{\mathbb{T}^d} \left\{ m'(H(x, Du') + H^*(x, -v') + v' \cdot Du') \right\} \, \mathrm{d}x \, \mathrm{d}t \\ & + \int_0^T \int_{\mathbb{T}^d} \left\{ F(x, m') + F^*(x, \alpha') - \alpha' m' \right\} \, \mathrm{d}x \, \mathrm{d}t.\end{aligned}$$

In view of our assumptions on **(HP1)** and **(HP2)**, we have the key inequality

$$\begin{aligned}\mathcal{A}(u', \alpha') + \mathcal{B}(m', v') \geq \; & \frac{\lambda}{4} \int_0^T \int_{\mathbb{T}^d} m'(t,x)|Du' - D_q H^*(x, -v')|^2 \, \mathrm{d}x \, \mathrm{d}t \\ & + \frac{\lambda}{4} \int_0^T \int_{\mathbb{T}^d} m'(t,x)|v' + D_p H(x, Du')|^2 \, \mathrm{d}x \, \mathrm{d}t \\ & + \int_0^T \int_{\mathbb{T}^d} \left\{ \frac{\lambda}{2} |\alpha'_1 - f(x, m')|^2 + p'(\overline{m} - m') \right\} \, \mathrm{d}x \, \mathrm{d}t\end{aligned} \tag{6.7}$$

where $p' = (\alpha' - f(x, \overline{m}))_+$ and $\alpha'_1 = \alpha' - p'$.

With the help of these tools let us show now the statements of the theorem.

Step 1. We first check that $f(\cdot, m) \in H^1_{\mathrm{loc}}((0,T) \times \mathbb{T}^d)$. Let us fix (m', v') to be a smooth competitor for $\mathcal{B}$ and let (u_n, α_n) be the minimizing sequence for Problem 3.6 defined in Lemma 4.3: we know that (α_n) is bounded in L^1 and converges to the (nonnegative) measure α defined from (β, β_T) by (4.4). Then, passing to the limit in the inequality

$$\mathcal{A}(u_n, \alpha_n) + \mathcal{B}(m', v') \geq \int_0^T \int_{\mathbb{T}^d} \left\{ F(x, m') + F^*(x, \alpha_n) - \alpha_n m' \right\} \, \mathrm{d}x \, \mathrm{d}t,$$

we get

$$\inf_{\mathcal{K}_R} \mathcal{A} + \mathcal{B}(m', v') \geq \int_0^T \int_{\mathbb{T}^d} \left\{ F(x, m') + F^*(x, \alpha^{ac}) - \alpha^{ac} m' \right\} \, \mathrm{d}x \, \mathrm{d}t + \int_0^T \int_{\mathbb{T}^d} (\overline{m} - m') \, \mathrm{d}\alpha^s.$$

In view of the proof of Theorem 4.2, we have $\alpha^{ac} \geq f(\cdot, m)$, with an equality in $\{m < \overline{m}\}$. So, if we set as above $p = (\alpha^{ac} - f(x, \overline{m}))_+$ and $\alpha^{ac}_1 = \alpha^{ac} - p$, then $\alpha^{ac}_1 = f(\cdot, m)$. By (6.3), this implies that

$$\inf_{\mathcal{K}_R} \mathcal{A} + \mathcal{B}(m', v') \geq \int_0^T \int_{\mathbb{T}^d} \frac{\lambda}{2} |f(x, m) - f(x, m')|^2 \, \mathrm{d}x \, \mathrm{d}t,$$

an inequality which remains true for any $(m', v') \in \mathcal{K}_D$ (not necessarily smooth ones). Adding $\inf_{\mathcal{K}_R} \mathcal{A}$ to inequality (6.4) and using the duality $\inf_{\mathcal{K}_R} \mathcal{A} + \min_{\mathcal{K}_D} \mathcal{B} = 0$ we have

$$\inf_{\mathcal{K}_R} \mathcal{A} + \mathcal{B}(m^{\delta,\eta}, v^{\delta,\eta}) \leq C(\eta^2 + |\delta|^2), \tag{6.8}$$

which implies

$$\int_{t_1}^{t_2} \int_{\mathbb{T}^d} |f(x,m) - f(x, m^{\delta,\eta})|^2 \,\mathrm{d}x \,\mathrm{d}t \leq C(\eta^2 + |\delta|^2)$$

and the regularity of f in x allows to conclude $f(\cdot, m) \in H^1_{\mathrm{loc}}((0,T) \times \mathbb{T}^d)$.

Step 2. Let (u_n, α_n) be the minimizing sequence defined by Lemma 4.3. Without loss of generality, we can assume that

$$\mathcal{A}(u_n, \alpha_n) - \inf_{\mathcal{K}_R} \mathcal{A} \leq 1/n. \tag{6.9}$$

We set $p_n := (\alpha_n - f(\cdot, \overline{m}))_+$ and $\alpha_{1,n} := \alpha_n - p_n$. For $\varphi : [0,T] \times \mathbb{T}^d \to \mathbb{R}$ and for $\eta > 0$ small, let us define the average of φ on the $[t, t+\eta]$ by

$$\varphi^\eta(t,x) := \int_0^1 \varphi(t + \theta\eta, x) \,\mathrm{d}\theta,$$

which is well-defined on $[t_1, t_2] \times \mathbb{T}^d$. With this procedure, we similarly define the functions p_n^η, α_n^η, etc. Let us take moreover $\sigma \in C^\infty([t_1, t_2]; [0, +\infty))$.

The aim of this step consists in estimating the quantity

$$I := \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma(t) \overline{m} \, |p_n^\eta(t, x+\delta) - p_n^\eta(t,x)| \,\mathrm{d}x \,\mathrm{d}t.$$

Namely, we prove that

$$\begin{aligned} I \quad &\leq C\|\sigma\|_{L^2} \left(|\delta| + \left(1 + \frac{|\delta|}{\eta}\right) (1/n + |\delta|^2 + \eta^2)^{1/2} \right) \\ &+ C\|\sigma\|_{L^\infty} (1/n + \eta^2 + |\delta|^2)^{1/2} \left[(1/n + \eta^2 + |\delta|^2)^{1/2} + |\delta| \left(1 + (1/n + \eta^2 + |\delta|^2)^{1/2}\right) \right] \\ &+ C \left\{ \|\sigma\|_{L^2}^2 + \|\sigma\|_{L^\infty} \left[(1/n + |\delta|^2 + \eta^2) + (1/n + |\delta|^2 + \eta^2)^{1/2} \right] \right\}^{1/2} (1/n + |\delta|^2 + \eta^2)^{1/2} \\ &=: X(\sigma, 1/n, \delta, \eta). \end{aligned} \tag{6.10}$$

We will show in the last two steps that this inequality easily entails the desired estimates on p and β.

The proof of (6.10) is quite long and relies on the combination of (6.7), (6.8) and (6.9) which imply that

$$\begin{aligned} 1/n + C(\eta^2 + |\delta|^2) \geq \quad & \frac{\lambda}{4} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} m^{\delta,\eta}(t,x) |Du_n - D_q H^*(x, -v^{\delta,\eta})|^2 \,\mathrm{d}x \,\mathrm{d}t \\ & + \frac{\lambda}{4} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} m^{\delta,\eta}(t,x) |v^{\delta,\eta} + D_p H(x, Du_n)|^2 \,\mathrm{d}x \,\mathrm{d}t \\ & + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \left\{ \frac{\lambda}{2} |\alpha_{1,n} - f(x, m^{\delta,\eta})|^2 + p_n(\overline{m} - m^{\delta,\eta}) \right\} \,\mathrm{d}x \,\mathrm{d}t \end{aligned} \tag{6.11}$$

We have

$$\begin{aligned} I \leq & \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma(t)(\overline{m} - m(t,x)) |p_n^\eta(t, x+\delta) - p_n^\eta(t,x)| \,\mathrm{d}x \,\mathrm{d}t \\ & + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma(t) m(t,x) |p_n^\eta(t, x+\delta) - p_n^\eta(t,x)| \,\mathrm{d}x \,\mathrm{d}t \\ & =: I_{01} + I_{02} \end{aligned}$$

where the the first term can be estimated as follows:

$$\begin{aligned}
I_{01} &\leq \|\sigma\|_{L^\infty} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} (\overline{m} - m(t,x)) \left\{ |p_n^\eta(t, x+\delta)| + |p_n^\eta(t,x)| \right\} \, \mathrm{d}x \, \mathrm{d}t \\
&= \|\sigma\|_{L^\infty} \int_0^1 \mathrm{d}\theta \int_{t_1}^{t_2} \int_{\mathbb{T}^d} (\overline{m} - m(t,x)) \left\{ p_n(t+\theta\eta, x+\delta) + p_n(t+\theta\eta, x) \right\} \, \mathrm{d}x \, \mathrm{d}t \\
&\leq \|\sigma\|_{L^\infty} \int_0^1 \mathrm{d}\theta \int_0^T \int_{\mathbb{T}^d} (\overline{m} - m^{-\delta,-\theta\eta}) p_n \, \mathrm{d}x \, \mathrm{d}t \\
&+ \|\sigma\|_{L^\infty} \int_0^1 \mathrm{d}\theta \int_0^T \int_{\mathbb{T}^d} (\overline{m} - m^{0,-\theta\eta}) p_n \, \mathrm{d}x \, \mathrm{d}t.
\end{aligned}$$

Now by (6.11) we obtain that

$$I_{01} \leq C\|\sigma\|_{L^\infty}(1/n + |\delta|^2 + \eta^2).$$

For the second term we have

$$\begin{aligned}
I_{02} \leq &\int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma(t) m(t,x) |\alpha_n^\eta(t, x+\delta) - \alpha_n^\eta(t,x)| \, \mathrm{d}x \, \mathrm{d}t \\
&+ \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma(t) \overline{m} |\alpha_{1,n}^\eta(t, x+\delta) - \alpha_{1,n}^\eta(t,x)| \, \mathrm{d}x \, \mathrm{d}t \\
&:= I_1 + I_2.
\end{aligned}$$

To estimate the term I_1, let us compute

$$\begin{aligned}
\alpha_n^\eta(t, x+\delta) - \alpha_n^\eta(t,x) = &\int_0^1 -\partial_t u_n(t+\theta\eta, x+\delta) + H(x+\delta, Du_n(t+\theta\eta, x+\delta)) \, \mathrm{d}\theta \\
&- \int_0^1 -\partial_t u_n(t+\theta\eta, x) + H(x, Du_n(t+\theta\eta, x)) \, \mathrm{d}\theta \\
&= -\eta^{-1} \int_0^1 [Du_n(t+\eta, x+s\delta) - Du_n(t, x+s\delta)] \cdot \delta \, \mathrm{d}s \\
&+ \int_0^1 \int_0^1 D_x H(x+s\delta, Du_n(t+\theta\eta, x+s\delta)) \cdot \delta \; ds d\theta \\
&+ \int_0^1 \mathrm{d}\theta \int_0^1 D_p H(x+s\delta, \xi_s) \cdot [Du_n(t+\theta\eta, x+\delta) - Du_n(t+\theta\eta, x)] \, \mathrm{d}s \, \mathrm{d}\theta
\end{aligned}$$

where $\xi_s := (1-s)Du_n(t+\theta\eta, x) + sDu_n(t+\theta\eta, x+\delta)$. Thus,

$$\begin{aligned}
|\alpha_n^\eta(t, x+\delta) - \alpha_n^\eta(t,x)| \leq &|\delta|\eta^{-1} \int_0^1 |Du_n(t+\eta, x+s\delta) - Du_n(t, x+s\delta)| \, \mathrm{d}s \\
&+ |\delta| \int_0^1 \int_0^1 |D_x H(x+s\delta, Du_n(t+\theta\eta, x+s\delta))| \, \mathrm{d}s \, \mathrm{d}\theta \\
&+ \int_0^1 \int_0^1 |D_p H(x+s\delta, \xi_s)| \; |Du_n(t+\theta\eta, x+\delta) - Du_n(t+\theta\eta, x)| \, \mathrm{d}s \, \mathrm{d}\theta.
\end{aligned}$$

In view of our assumption (6.2) on D_xH and D_pH:

$$\begin{aligned}
I_1 &= \int_{t_1}^{t_2}\int_{\mathbb{T}^d} \sigma(t)m(t,x)|\alpha_n^\eta(t,x+\delta)-\alpha_n^\eta(t,x)|\,\mathrm{d}x\,\mathrm{d}t\\
&\le |\delta|\eta^{-1}\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1 \sigma(t)m(t,x)|Du_n(t+\eta,x+s\delta)-Du_n(t,x+s\delta)|\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\\
&+ C|\delta|\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1\int_0^1 \sigma(t)m(t,x)\left\{1+|Du_n(t+\theta\eta,x+s\delta)|^2\right\}\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t\\
&+ C\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1 \sigma(t)m(t,x)\left\{1+|Du_n(t+\theta\eta,x)|+|Du_n(t+\theta\eta,x+\delta)|\right\}\\
&\qquad\qquad\times|Du_n(t+\theta\eta,x+\delta)-Du_n(t+\theta\eta,x)|\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t\\
&:= I_{11}+I_{12}+I_{13}.
\end{aligned}$$

For I_{11}, we have

$$\begin{aligned}
I_{11} &\le |\delta|\eta^{-1}\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1 \Big\{\sigma(t)m(t,x)\big(|Du_n(t+\eta,x+s\delta)-D_qH^*(x,-v(t,x))|\\
&\qquad\qquad + |D_qH^*(x,-v(t,x))-Du_n(t,x+s\delta)|\big)\Big\}\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\\
&\le |\delta|\eta^{-1}\int_{t_1+\eta}^{t_2+\eta}\int_{\mathbb{T}^d}\int_0^1 \sigma^{0,-\eta}(t)m^{-s\delta,-\eta}(t,x)|Du_n(t,x)-D_qH^*(x-s\delta,-v^{-s\delta,-\eta}(t,x))|\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\\
&+ |\delta|\eta^{-1}\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1 \sigma(t)m^{-s\delta,0}(t,x)|Du_n(t,x)-D_qH^*(x-s\delta,-v^{-s\delta,0}(t,x))|\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}$$

By Cauchy-Schwarz and (6.11) we obtain:

$$I_{11}\le C|\delta|\eta^{-1}\|\sigma\|_{L^2}(1/n+\eta^2+|\delta|^2)^{1/2}.$$

We now estimate I_{12}, that we bound from above as follows:

$$\begin{aligned}
I_{12} &\le C\|\sigma\|_{L^2}|\delta| + C|\delta|\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1\int_0^1 \sigma(t)m(t,x)|D_qH^*(x,-v(t,x))|^2\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t\\
&+ C|\delta|\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1\int_0^1 \sigma(t)m(t,x)\left\{|Du_n(t+\theta\eta,x+s\delta)|^2-|D_qH^*(x,-v(t,x))|^2\right\}\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}$$

The second term can be estimated by (6.6), while, for the third one, we use the inequality $|a|^2-|b|^2\le |a-b|^2+|a-b||b|$ to get:

$$\begin{aligned}
I_{12} &\le C\|\sigma\|_{L^2}|\delta| +\\
&+ C\|\sigma\|_{L^\infty}|\delta|\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1\int_0^1 m(t,x)|Du_n(t+\theta\eta,x+s\delta)-D_qH^*(x,-v(t,x))|^2\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t\\
&+ 2C\|\sigma\|_{L^\infty}|\delta|\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1\int_0^1 \Big\{m(t,x)|Du_n(t+\theta\eta,x+s\delta)-D_qH^*(x,-v(t,x))|\\
&\qquad\qquad\times|D_qH^*(x,-v(t,x))|\Big\}\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&\leq C\|\sigma\|_{L^2}|\delta|+\\
&+C\|\sigma\|_{L^\infty}|\delta|\int_{t_1+\theta\eta}^{t_2+\theta\eta}\int_{\mathbb{T}^d}\int_0^1\int_0^1 m^{-s\delta,-\theta\eta}(t,x)|Du_n(t,x)-D_qH^*(x-s\delta,-v^{-s\delta,-\theta\eta}(t,x))|^2\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t\\
&+2C\|\sigma\|_{L^\infty}|\delta|\int_{t_1+\theta\eta}^{t_2+\theta\eta}\int_{\mathbb{T}^d}\int_0^1\int_0^1\Big\{m^{-s\delta,-\theta\eta}(t,x)|Du_n(t,x)-D_qH^*(x-s\delta,-v^{-s\delta,-\theta\eta}(t,x))|\\
&\qquad\qquad\times|D_qH^*(x-s\delta,-v^{-s\delta,-\theta\eta}(t,x))|\Big\}\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

As before, using the energy estimates (6.6) and (6.11) together with a Cauchy-Schwarz inequality in the last integral we obtain

$$
I_{12}\leq C\|\sigma\|_{L^2}|\delta|+C\|\sigma\|_{L^\infty}|\delta|\left\{(1/n+\eta^2+|\delta|^2)+C(1/n+\eta^2+|\delta|^2)^{1/2}\right\}.
$$

It is easy to see that with the help of the estimations for I_{11} and I_{12} we can estimate I_{13} as well. Hence we obtain

$$
\begin{aligned}
I_{13}&\leq C\|\sigma\|_{L^2}(1/n+|\delta|^2+\eta^2)^{1/2}\\
&\quad+C\left\{C\|\sigma\|_{L^2}^2+C\|\sigma\|_{L^\infty}\left[(1/n+|\delta|^2+\eta^2)+(1/n+|\delta|^2+\eta^2)^{1/2}\right]\right\}^{1/2}(1/n+|\delta|^2+\eta^2)^{1/2}.
\end{aligned}
$$

Let us now take care of I_2. Setting $f^\eta(t,x):=\displaystyle\int_0^1 f(x,m(t+\theta\eta,x))\,\mathrm{d}\theta$, we have

$$
\begin{aligned}
I_2&=\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\sigma(t)\overline{m}|\alpha_{1,n}^\eta(t,x+\delta)-\alpha_{1,n}^\eta(t,x)|\,\mathrm{d}x\,\mathrm{d}t\\
&\leq\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\sigma(t)\overline{m}|\alpha_{1,n}^\eta(t,x+\delta)-f^\eta(t,x+\delta)|\,\mathrm{d}x\,\mathrm{d}t+\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\sigma(t)\overline{m}|f^\eta(t,x+\delta)-f^\eta(t,x)|\,\mathrm{d}x\,\mathrm{d}t\\
&+\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\sigma(t)\overline{m}|f^\eta(t,x)-\alpha_{1,n}^\eta(t,x)|\,\mathrm{d}x\,\mathrm{d}t\\
&:=\ I_{21}+I_{22}+I_{23}.
\end{aligned}
$$

Since

$$
\begin{aligned}
I_{21}&\leq C\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\int_0^1\sigma(t)|\alpha_{1,n}(t+\theta\eta,x+\delta)-f(x+\delta,m(t+\theta\eta,x+\delta))|\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}t\\
&\leq C\int_0^1\int_{t_1-\theta\eta}^{t_2-\theta\eta}\int_{\mathbb{T}^d}\sigma^{-\theta\eta}(t)|\alpha_{1,n}(t,x)-f(x,m(t,x))|\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\theta
\end{aligned}
$$

we obtain by Cauchy-Schwarz and (6.11):

$$
I_{21}\ \leq C\|\sigma\|_{L^2}(1/n+|\delta|^2+\eta^2)^{1/2}.
$$

The term I_{23} can be treated in the same way. For I_{22}, we have

$$
\begin{aligned}
I_{22}&\leq C\|\sigma\|_{L^2}\int_0^1\left(\int_{t_1}^{t_2}\int_{\mathbb{T}^d}|f(x+\delta,m(t+\theta\eta,x+\delta))-f(x,m(t+\theta\eta,x))|^2\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\mathrm{d}\theta\\
&\leq C\|\sigma\|_{L^2}|\delta|
\end{aligned}
$$

because $f(\cdot,m(\cdot,\cdot))$ is in $H^1_{\mathrm{loc}}((0,T)\times\mathbb{T}^d)$.

Putting the above inequalities together gives (6.10).

Step 3. We now show that $p_n := (\alpha_n - f(\cdot, \overline{m}))_+$ belongs to the space $L^2([t_1, t_2]; BV(\mathbb{T}^d))$. Let us take a test function $\psi \in C_c^\infty((0,T) \times \mathbb{T}^d)$, $e \in \mathbb{R}^n$ with $|e| = 1$, $\eta > 0$ small and let us set $\delta := \eta e$. Let us estimate

$$\begin{aligned}
&\int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma^{-\eta}(t) \frac{\psi^{-\eta}(t,x) - \psi^{-\eta}(t, x-\delta)}{\eta} p_n(t,x) \,\mathrm{d}x \,\mathrm{d}t \\
&\qquad = \int_{t_1-\eta}^{t_2-\eta} \int_{\mathbb{T}^d} \sigma(t) \frac{\psi(t,x) - \psi(t, x-\delta)}{\eta} p_n^\eta(t,x) \,\mathrm{d}x \,\mathrm{d}t \\
&\qquad = \int_{t_1-\eta}^{t_2-\eta} \int_{\mathbb{T}^d} \sigma(t)\psi(x) \frac{p_n^\eta(t,x) - p_n^\eta(t, x+\delta)}{\eta} \,\mathrm{d}x \,\mathrm{d}t \\
&\qquad \le \|\psi\|_{L^\infty} \frac{1}{\eta} \int_{t_1-\eta}^{t_2-\eta} \int_{\mathbb{T}^d} \sigma(t) |p_n^\eta(t,x) - p_n^\eta(t, x+\delta)| \,\mathrm{d}x \,\mathrm{d}t \\
&\qquad \le \|\psi\|_{L^\infty} \frac{1}{\eta} X(\sigma, 1/n, \eta e, \eta).
\end{aligned}$$

First let us recall that $p_n \overset{*}{\rightharpoonup} p$ as $n \to +\infty$ in $\mathfrak{M}([0,T] \times \mathbb{T}^d)$, which allows us to pass to the limit in the above inequality as $n \to +\infty$ and obtain

$$\int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma^{-\eta}(t) \frac{\psi^{-\eta}(t,x) - \psi^{-\eta}(t, x-\delta)}{\eta} \,\mathrm{d}p(t,x) \le \|\psi\|_{L^\infty} \frac{1}{\eta} X(\sigma, 0, \eta e, \eta).$$

Now sending $\eta \to 0$ and recalling (6.10) we have

$$\int_{t_1}^{t_2} \int_{\mathbb{T}^d} \sigma(t) \nabla\psi(t,x) \cdot e \,\mathrm{d}p(t,x) \le C \|\psi\|_{L^\infty} \|\sigma\|_{L^2}.$$

Therefore we obtain that $p \in L^2([t_1,t_2]; BV(\mathbb{T}^d)) \hookrightarrow L^2([t_1,t_2]; L^{d/(d-1)}(\mathbb{T}^d))$ and in particular, by the arbitrariness of t_1 and t_2 and by an injection we have $p \in L^{d/(d-1)}_{\mathrm{loc}}((0,T) \times \mathbb{T}^d)$.

Step 4. Conclusion: as $0 \le \alpha_n \le p_n + f(\cdot, \overline{m})$ and (α_n) converges to α defined by (4.4), we have $0 \le \beta \le p + f(\cdot, \overline{m})$ in $(0,T) \times \mathbb{T}^d$. This proves that β is absolutely continuous and belongs to $L^{d/(d-1)}_{\mathrm{loc}}((0,T) \times \mathbb{T}^d)$. □

Remark 6.3. *Note that by the example provided in the Section 5 we have the sharpness of the above integrability result in the following sense: we cannot expect a bound for p_n in $L^{d/(d-1)}_{\mathrm{loc}}((0,T] \times \mathbb{T}^d)$, i.e., up to the final time, because of the occurrence of a possible jump at $t = T$.*

7. Nash equilibria for MFG with density constraints

Let us suppose in this section the additional assumptions **(HP1)**, **(HP2)** and **(HP3)** as in Section 6. To define a proper notion of Nash equilibrium, we use the techniques for measures on paths, corresponding the trajectories of single agents. This will also allow us to clarify the meaning of the control problem (1.4). The used machinery is inspired by [3] (Section 6) and also by [11] (Section 4.3) and [12] (Section 4). We remark also some similarities of this approach with the works modeling traffic congestion and Wardrop equilibria (see [6, 16]).

7.1. **Density-constrained flows and a first optimality condition.** Let us recall that Γ denotes the set of absolutely continuous curves $\gamma : [0,T] \to \mathbb{T}^d$ and $\mathcal{P}_2(\Gamma)$ the set of Borel probability measures $\tilde{\boldsymbol{\eta}}$ defined on Γ such that

$$\int_\Gamma \int_0^T |\dot\gamma(s)|^2 \,\mathrm{d}s \,\mathrm{d}\tilde{\boldsymbol{\eta}}(\gamma) < +\infty.$$

We call $\tilde{\boldsymbol{\eta}}$ an *almost density-constrained flow* if there exists $C = C(\tilde{\boldsymbol{\eta}}) > 0$ such that $0 \le \tilde{m}_t \le C(\tilde{\boldsymbol{\eta}})$ a.e. in $\mathbb{T}^d$ for all $t \in [0,T]$, where $\tilde{m}_t := (e_t)_\# \tilde{\boldsymbol{\eta}}$. If $C(\tilde{\boldsymbol{\eta}}) \le \overline{m}$ (the density constraint, given by our

model) then we call $\tilde{\boldsymbol{\eta}}$ a *density-constrained flow*. Let us recall moreover that we use the definition of the Lagrangian as $L(x,v) = H^*(x,-v)$.

In the whole section we consider a solution (u, m, β, β_T) of the MFG system (4.1). By Theorem 4.2 and Theorem 6.2 this corresponds to (u,α) and (m,w) solutions of Problem 3.6 and Problem 3.2 respectively, where

$$\alpha = f(\cdot, m)\,\mathrm{d}x\,\mathrm{d}t + \beta\,\mathrm{d}x\,\mathrm{d}t + \beta_T\,\mathrm{d}(\delta_T \otimes \mathcal{H}^d \llcorner \mathbb{T}^d) \quad \text{and} \quad w = -mD_pH(x, Du).$$

Let us state the following results (in the spirit of Lemma 4.6.-4.8. from [11]) which characterize the density-constrained flows.

Lemma 7.1. *Let $\tilde{\boldsymbol{\eta}} \in \mathcal{P}_2(\Gamma)$ be an almost density-constrained flow and set $\tilde{m}_t := (e_t)_\#\tilde{\boldsymbol{\eta}}$. Then*

(i) *for all $0 \le t_1 < t_2 \le T$ we have*

$$\int_{\mathbb{T}^d} u(t_1^+, x)\tilde{m}(t_1, x)\,\mathrm{d}x \le \int_{\mathbb{T}^d} u(t_2^-, x)\tilde{m}(t_2, x)\,\mathrm{d}x + \int_\Gamma \int_{t_1}^{t_2} L(\gamma(t), \dot\gamma(t))\,\mathrm{d}t\,\mathrm{d}\tilde{\boldsymbol{\eta}}(\gamma) + \int_{t_1}^{t_2}\int_{\mathbb{T}^d} \alpha(t,x)\tilde{m}(t,x)\,\mathrm{d}x\,\mathrm{d}t.$$

(we recall $\alpha(t,x) := f(x, m(t,x)) + \beta(t,x)$)

(ii) *In particular, for all $0 \le t_1 < T$*

$$\int_{\mathbb{T}^d} u(t_1^+, x)\tilde{m}(t_1, x)\,\mathrm{d}x \le \int_{\mathbb{T}^d} (g(x) + \beta_T(x))\tilde{m}(T, x)\,\mathrm{d}x + \int_\Gamma \int_0^T L(\gamma(t), \dot\gamma(t))\,\mathrm{d}t\,\mathrm{d}\tilde{\boldsymbol{\eta}}(\gamma) + \int_0^T\int_{\mathbb{T}^d} \alpha(t,x)\tilde{m}(t,x)\,\mathrm{d}x\,\mathrm{d}t.$$

Proof. Let us recall that u satisfies, in the sense of measures,

$$-\partial_t u + H(x, Du) \le \alpha \qquad \text{in } (0,T)\times\mathbb{T}^d,$$

where α belongs to $L^{d/(d-1)}_{loc}((0,T)\times\mathbb{T}^d)$ thanks to Theorem 6.2. If we regularize u into u_n and α into α_n by convolution (with a compact support in $B_{1/n}(0)$), we obtain

$$-\partial_t u_n + H(x, Du_n) \le \alpha_n + r_n \qquad \text{in } (1/n, T-1/n)\times\mathbb{T}^d,$$

where

$$r_n(t,x) = H(x, Du_n(t,x)) - H(\cdot, Du) \star \rho_n(t,x).$$

Note that (r_n) tends to 0 in $L^1((0,T)\times\mathbb{T}^d)$. By the way, if H is independent of x, one also has $r_n \le 0$. Let us fix $0 < t_1 < t_2 < T$ and n large. Now for any $\gamma \in H^1([0,T])$ we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(u_n(t,\gamma(t)) - \int_t^T L(\gamma(s), \dot\gamma(s))\,\mathrm{d}s\right) &\ge \partial_t u_n(t,\gamma(t)) - H(\gamma(t), Du_n(t,\gamma(t)))\\ &\ge -\alpha_n(t,\gamma(t)) - r_n(t,\gamma(t)).\end{aligned} \tag{7.1}$$

Integrating this inequality on $[t_1,t_2]$, then over Γ w.r.t. $\tilde{\boldsymbol{\eta}}$, we obtain

$$\int_{\mathbb{T}^d} u_n(t_1,x)\tilde{m}(t_1,x)\,\mathrm{d}x \le \int_{\mathbb{T}^d} u_n(t_2,x)\tilde{m}(t_2,x)\,\mathrm{d}x + \int_\Gamma\int_{t_1}^{t_2} L(\gamma(t),\dot\gamma(t))\,\mathrm{d}t\,\mathrm{d}\tilde{\boldsymbol{\eta}}(\gamma) + \int_{\mathbb{T}^d}\int_{t_1}^{t_2} [\alpha_n(t,x) + r_n(t,x)]\tilde{m}(t,x)\,\mathrm{d}t\,\mathrm{d}x.$$

We recall the fact that $\tilde{m} \in L^\infty([0,T]\times\mathbb{T}^d)$. Since (u_n) strongly converges in L^1 to $u \in BV([0,T]\times\mathbb{T}^d)$, we have the existence of $J \subset (0,T)$ of full measure such that for every $t_1, t_2 \in J, t_1 < t_2$, the first two integrals pass to the limit as $n \to +\infty$. By the strong convergence in $L^{d/(d-1)}([t_1,t_2]\times\mathbb{T}^d)$ of (α_n) to

α and in $L^1([t_1,t_2]\times\mathbb{T}^d)$ of (r_n) to 0, we can pass to the limit as $n\to+\infty$ is the last integral as well. So, for a.e. $0<t_1<t_2<T$, we have

$$\int_{\mathbb{T}^d} u(t_1,x)\tilde{m}(t_1,x)\,\mathrm{d}x \le \int_{\mathbb{T}^d} u(t_2,x)\tilde{m}(t_2,x)\,\mathrm{d}x + \int_\Gamma\int_{t_1}^{t_2} L(\gamma(t),\dot\gamma(t))\,\mathrm{d}t\,\mathrm{d}\tilde{\boldsymbol{\eta}}(\gamma) + \int_{\mathbb{T}^d}\int_{t_1}^{t_2}\alpha(t,x)\tilde{m}(t,x)\,\mathrm{d}t\,\mathrm{d}x.$$

In order to show that the inequality holds for any $t_1<t_2$, let us now check that

$$\lim_{t'\to t^\pm}\int_{\mathbb{T}^d} u((t')^+,x)\tilde{m}(t,x)\,\mathrm{d}x = \int_{\mathbb{T}^d} u(t^\pm,x)\tilde{m}(t,x)\,\mathrm{d}x,$$

where $u((t')^\pm,\cdot)$ is understood in the sense of trace and $\tilde{m}(t,\cdot)$ is the (bounded) density of the continuous representative of the map $t\mapsto\tilde{m}(t,\cdot)\,\mathrm{d}x$ (for the L^∞ weak$-\star$ convergence). The above limit basically follows from the trace properties of BV functions, but for the sake of completeness let us sketch it below. Let u_n be a standard mollification in space of u. As u is in BV, $u((t')^+,\cdot)$ converges in L^1 to $u(t^\pm,\cdot)$ as $t'\to t^\pm$, so that $u_n((t')^+,\cdot)$ uniformly converges to $u_n(t^\pm,\cdot)$. Let us write $\int_{\mathbb{T}^d} u((t')^+,x)\tilde{m}(t',x)\,\mathrm{d}x$ as

$$\int_{\mathbb{T}^d} u_n((t')^+,x)\tilde{m}(t',x)\,\mathrm{d}x + \int_{\mathbb{T}^d}(u((t')^+,x)-u_n((t')^+,x))\tilde{m}(t',x))\,\mathrm{d}x. \tag{7.2}$$

By uniform convergence of $u_n((t')^+,\cdot)$, the first term in (7.2) converges to $\int_{\mathbb{T}^d} u_n(t^\pm,x)\tilde{m}(t,x)\,\mathrm{d}x$, which is arbitrary close to $\int_{\mathbb{T}^d} u(t^\pm,x)\tilde{m}(t,x)\,\mathrm{d}x$ for n large. As for the second term in (7.2), it is bounded by $\|u((t')^+,\cdot)-u_n((t')^+,\cdot)\|_{L^1}\|m\|_{L^\infty}$, which, by L^1 convergence of $u((t')^+,\cdot)$ to $u(t^\pm,\cdot)$, tends to 0 uniformly in t'. This proves (i).

For (ii), we just apply (i) for $t_2=T$, since $u(T^-,\cdot)=g+\beta_T$. □

Definition 7.2. *We say that an $\boldsymbol{\eta}\in\mathcal{P}_2(\Gamma)$ is an* optimal density-constrained flow *associated with the solution (u,m,β,β_T) if $m(t,\cdot)=(e_t)_\#\boldsymbol{\eta}$, for all $t\in[0,T]$ and the following energy equality holds*

$$\int_{\mathbb{T}^d} u(0^+,x)m_0(x)\,\mathrm{d}x = \int_{\mathbb{T}^d} g(x)m(T,x)\,\mathrm{d}x + \overline{m}\int_{\mathbb{T}^d}\beta_T\,\mathrm{d}x + \int_\Gamma\int_0^T L(\gamma(t),\dot\gamma(t))\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\eta}(\gamma) + \int_0^T\int_{\mathbb{T}^d}(f(x,m(t,x))+\beta(t,x))\,m(t,x)\,\mathrm{d}x\,\mathrm{d}t.$$

Note that the above definition is the reformulation in terms of density-constrained flows of the energy equality from Definition 4.1 point (4).

Remark 7.3. *Let us observe that for an optimal density-constrained flow $\boldsymbol{\eta}$, the energy equality in Definition 7.2 holds for any $0\le t_1<t_2\le T$ as well, i.e.*

$$\int_{\mathbb{T}^d} u(t_1^+,x)m(t_1,x)\,\mathrm{d}x = \int_{\mathbb{T}^d} u(t_2^-,x)m(t_2,x)\,\mathrm{d}x + \int_\Gamma\int_{t_1}^{t_2} L(\gamma(t),\dot\gamma(t))\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\eta}(\gamma) + \int_{t_1}^{t_2}\int_{\mathbb{T}^d}(f(x,m(t,x))+\beta(t,x))\,m(t,x)\,\mathrm{d}x\,\mathrm{d}t. \tag{7.3}$$

This can be easily deduced using the inequalities from Lemma 7.1 three times on the intervals $[0,t_1],[t_1,t_2]$ and $[t_2,T]$ together with the global equality from Definition 7.2 and the fact that we have $(\partial_t u)^s$ is a non-negative measure (by the fact that $-\partial_t u+H(\cdot,Du)\le\alpha$ and α does not have singular part in $(0,T)\times\mathbb{T}^d$), i.e. one has always $u(t^-,\cdot)\le u(t^+,\cdot)$ a.e. for all $t\in(0,T)$.

Identity (7.3) *implies also that* $(\partial_t u)^s = 0$ *on the support of* m*, more precisely*

$$\int_{\mathbb{T}^d} u(t^+, x) m(t, x)\, \mathrm{d}x = \int_{\mathbb{T}^d} u(t^-, x) m(t, x)\, \mathrm{d}x, \tag{7.4}$$

for all $t \in (0, T)$.

The following proposition gives the existence result for an optimal density-constrained flow $\boldsymbol{\eta}$.

Proposition 7.4. *There exists at least one optimal density-constrained flow* $\boldsymbol{\eta} \in \mathcal{P}_2(\Gamma)$ *in the sense of the Definition 7.2.*

Proof. The proof uses the same construction and goes along the same lines as in [11]. Nevertheless, we discuss the main steps here.

We construct a family $(\boldsymbol{\eta}_\varepsilon)_{\varepsilon>0}$ of density-constrained flows by

$$\int_\Gamma \Psi(\gamma)\, \mathrm{d}\boldsymbol{\eta}_\varepsilon(\gamma) := \int_{\mathbb{T}^d} \Psi(X_\varepsilon^x) m_0(x)\, \mathrm{d}x,$$

for any bounded and continuous map $\Psi : \Gamma \to \mathbb{R}$, where X_ε^x is the solution of the Cauchy problem

$$\begin{cases} \dot{x}(t) = \dfrac{w_\varepsilon(t, x(t))}{m_\varepsilon(t, x(t))}, & \text{a.e in } [0, T], \\ x(0) = x, \end{cases}$$

$(m_\varepsilon, w_\varepsilon)$ being a standard mollification of (m, w) such that $0 < m_\varepsilon \leq \overline{m}$. One easily checks that $m_\varepsilon(t, \cdot) = (e_t)_\# \boldsymbol{\eta}_\varepsilon$.

Using Lemma 4.7. from [11] we obtain that the family $(\boldsymbol{\eta}_\varepsilon)_{\varepsilon>0}$ is tight. Denoting by $\boldsymbol{\eta}$ the limit of a suitable subsequence of such a family, this is an optimal density-constrained flow in the sense of Definition 7.2. The proof of this statement goes exactly as for Lemma 4.8. in [11], using the equality (4) from Definition 4.1 and the inequality (ii) from Lemma 7.1. □

7.2. **Optimality conditions on the level of single agent trajectories.** In this subsection our aim is to show that the optimal density-constrained flows are actually concentrated on paths which are optimal (in some weak sense) for the control problem (1.4) (see Definition 7.5). We will show that they satisfy a weak dynamic programming principle.

Let us recall that $\beta \in L^2_{\mathrm{loc}}((0, T); BV(\mathbb{T}^d))$ and $\beta_T \in L^1(\mathbb{T}^d)$. In order to handle the evaluation of β along single agent paths we shall work with specific representative of it (which is defined everywhere in $\mathbb{T}^d$).

For an L^1_{loc} function $h : \mathbb{T}^d \to \mathbb{R}$ we define the specific representative of h by

$$\hat{h}(x) := \limsup_{\varepsilon \downarrow 0} h_\varepsilon(x), \tag{7.5}$$

where

$$h_\varepsilon(x) := \int_{\mathbb{R}^d} h(x + \varepsilon y) \rho(y)\, \mathrm{d}y$$

and ρ being the heat kernel

$$\rho(y) := (2\pi)^{-d/2} e^{-|y|^2/2}. \tag{7.6}$$

We use this specific regularization via the heat kernel because of the semigroup property $(h_\varepsilon)_{\varepsilon'} = h_{\varepsilon+\varepsilon'}$ we shall profit on later.

To treat passages to limit (in the regularization, as $\varepsilon \downarrow 0$, similarly as in Section 6 from [3]) we will need some uniform point-wise bounds on β_ε, hence we shall use the properties of the Hardy-Littlewood-type maximal function defined with the help of the heat kernel (7.6). Thus for any $h \in L^1(\mathbb{T}^d)$ we set

$$(Mh)(x) := \sup_{\varepsilon>0} \int_{\mathbb{R}^d} |h(x + \varepsilon y)| \rho(y)\, \mathrm{d}y.$$

Let us state some basic properties of the maximal functional M that we will use in our setting. First because of the semigroup property we have

$$Mh_\varepsilon = \sup_{\varepsilon'>0} |h_\varepsilon|_{\varepsilon'} \leq \sup_{\tilde\varepsilon>0} |h|_{\tilde\varepsilon} = Mh.$$

Secondly it is well-known that M leaves invariant any L^p space with $1 < p \leq +\infty$ and there exists $C_p > 0$ such that

$$\|Mh\|_{L^p(\mathbb{T}^d)} \leq C_p \|h\|_{L^p(\mathbb{T}^d)}.$$

Let us recall that by Theorem 6.2 we have that $M\beta \in L^{d/(d-1)}_{\mathrm{loc}}((0,T)\times\mathbb{T}^d) \hookrightarrow L^1_{\mathrm{loc}}((0,T)\times\mathbb{T}^d)$. The integrability property we need is only $M\beta \in L^1_{\mathrm{loc}}((0,T)\times\mathbb{T}^d)$, but to guarantee this, $\beta \in L^1_{\mathrm{loc}}((0,T)\times\mathbb{T}^d)$ is not enough.

As usual we set $\alpha(t,x) := f(x, m(t,x)) + \beta(t,x)$ and we use its representative $\hat\alpha$ (obtained as in (7.5)).

Definition 7.5. *Given $0 < t_1 < t_2 < T$, we say that a path $\gamma \in H^1([0,T];\mathbb{T}^d)$ with $M\hat\alpha(\cdot,\gamma) \in L^1_{\mathrm{loc}}((0,T))$ is* minimizing *on the time interval $[t_1,t_2]$ in the problem (1.4) if we have*

$$\hat u(t_2,\gamma(t_2)) + \int_{t_1}^{t_2} L(\gamma(t),\dot\gamma(t)) + \hat\alpha(t,\gamma(t))\,\mathrm{d}t \leq \hat u(t_2,\gamma(t_2)+\omega(t_2)) + \tag{7.7}$$

$$+ \int_{t_1}^{t_2} L(\gamma(t)+\omega(t),\dot\gamma(t)+\dot\omega(t)) + \hat\alpha(t,\gamma(t)+\omega(t))\,\mathrm{d}t, \tag{7.8}$$

for all $\omega \in H^1([t_1,t_2];\mathbb{T}^d)$ such that $\omega(t_1) = 0$ and $M\hat\alpha(\cdot,\gamma+\omega) \in L^1([t_1,t_2])$.

Remark 7.6. *Let us notice that for any density-constrained flow $\tilde{\boldsymbol{\eta}}$ the integrability property $M\hat\alpha(\cdot,\gamma) \in L^1_{\mathrm{loc}}((0,T))$ is natural, since it is satisfied $\tilde{\boldsymbol{\eta}}$-a.e., if $M\hat\alpha \in L^1_{\mathrm{loc}}((0,T)\times\mathbb{T}^d)$. Indeed, we have*

$$\int_\Gamma \int_{t_1}^{t_2} M\hat\alpha(t,\gamma(t))\,\mathrm{d}t\,\mathrm{d}\tilde{\boldsymbol{\eta}}(\gamma) = \int_{t_1}^{t_2}\int_{\mathbb{T}^d} M\hat\alpha(t,x)\tilde m(t,x)\,\mathrm{d}x\,\mathrm{d}t < +\infty,$$

for all $0 < t_1 < t_2 < T$, where $\tilde m(t,\cdot)\,\mathrm{d}x = (e_t)_\#\tilde{\boldsymbol{\eta}}$.

Theorem 7.7. *For any $0 < t_1 < t_2 < T$, any optimal density-constrained flow $\boldsymbol{\eta}$ is concentrated on minimizing paths on the time interval $[t_1,t_2]$ for the problem (1.4) in the sense of the Definition 7.5.*

Proof. We follow here Ambrosio-Figalli [3]. Let us take an optimal density-constrained flow $\boldsymbol{\eta}$ given by Proposition 7.4, fix $0 < t_1 < t_2 < T$ and $y \in \mathbb{T}^d$, take $\omega \in H^1([t_1,t_2];\mathbb{T}^d)$ with $\omega(t_1) = 0$ and $\chi \in C^1_c((0,T);[0,1])$ with $\chi > 0$ on $(t_1,t_2]$ and $\chi(t_1) = 0$ a smooth cut-off function. Let us take a Borel subset $E \subset \Gamma$ such that $\boldsymbol{\eta}(E)$ is positive. For $\varepsilon > 0$ and $y \in \mathbb{T}^d$ we introduce the map $T_{\varepsilon,y} : \Gamma \to \Gamma$ by

$$T_{\varepsilon,y}(\gamma) := \begin{cases} \gamma, & \text{if } \gamma \notin E, \\ \gamma + \omega + \varepsilon\chi y, & \text{if } \gamma \in E. \end{cases}$$

Now let us define $\boldsymbol{\eta}_{\varepsilon,y} := (T_{\varepsilon,y})_\#\boldsymbol{\eta}$, which in particular is an admissible density-constrained flow satisfying the inequalities from Lemma 7.1. In addition let us remark that $(e_{t_1})_\#\boldsymbol{\eta}_{\varepsilon,y} = (e_{t_1})_\#\boldsymbol{\eta} = m(t_1,\cdot)\,\mathrm{d}x$.

Using the inequality (i) from Lemma 7.1 for $\boldsymbol{\eta}_{\varepsilon,y}$ (on the time interval $[t_1,t_2]$) and the equality (7.3) for $\boldsymbol{\eta}$ (on the same interval $[t_1,t_2]$) we obtain

$$\int_E \Big[\hat u(t_2^+,\gamma(t_2)) + \int_{t_1}^{t_2} L(\gamma(t),\dot\gamma(t)) + \hat\alpha(t,\gamma(t))\,\mathrm{d}t\Big]\,\mathrm{d}\boldsymbol{\eta}(\gamma) \leq$$

$$\int_E \Big[\hat u(t_2^-,\gamma(t_2)+\omega(t_2)+\varepsilon\chi(t_2)y) + \int_{t_1}^{t_2} L(\gamma(t)+\omega(t)+\varepsilon\chi(t)y,\dot\gamma(t)+\dot\omega(t)+\varepsilon\dot\chi(t)y)\,\mathrm{d}t\Big]\,\mathrm{d}\boldsymbol{\eta}(\gamma)$$

$$+ \int_E\int_{t_1}^{t_2} \hat\alpha(t,\gamma(t)+\omega(t)+\varepsilon\chi(t)y)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\eta}(\gamma).$$

where we are allowed to use any representative of u and α, thus we use the specially constructed ones $\hat{u}$ and $\hat{\alpha}$. Let us average this last inequality w.r.t. the variable y using the kernel ρ introduced in (7.6). We obtain

$$\int_E \Big[\hat{u}(t_2^+,\gamma(t_2)) + \int_{t_1}^{t_2} L(\gamma(t),\dot\gamma(t)) + \hat\alpha(t,\gamma(t))\,\mathrm{d}t\Big]\,\mathrm{d}\boldsymbol{\eta}(\gamma) \le$$
$$\int_E\int_{\mathbb{R}^d}\Big[\hat{u}(t_2^-,\gamma(t_2)+\omega(t_2)+\varepsilon\chi(t_2)y)+\int_{t_1}^{t_2} L(\gamma(t)+\omega(t)+\varepsilon\chi(t)y,\dot\gamma(t)+\dot\omega(t)+\varepsilon\dot\chi(t)y)\,\mathrm{d}t\Big]\rho(y)\,\mathrm{d}y\,\mathrm{d}\boldsymbol{\eta}(\gamma)$$
$$+\int_E\int_{t_1}^{t_2}\hat\alpha_{\varepsilon\chi(t)}(t,\gamma(t)+\omega(t))\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\eta}(\gamma).$$

Now choosing $\mathcal{D}\subset H^1([t_1,t_2];\mathbb{T}^d)$ a dense subset with $\omega(t_1)=0$ for all $\omega\in\mathcal{D}$, by the arbitrariness of E for $\boldsymbol{\eta}$-almost every curve $\gamma\in\Gamma$ we deduce that

$$\hat{u}(t_2^+,\gamma(t_2)) + \int_{t_1}^{t_2} L(\gamma(t),\dot\gamma(t)) + \hat\alpha(t,\gamma(t))\,\mathrm{d}t$$
$$\le \int_{\mathbb{R}^d}\int_{t_1}^{t_2} L(\gamma(t)+\omega(t)+\varepsilon\chi(t)y,\dot\gamma(t)+\dot\omega(t)+\varepsilon\dot\chi(t)y)\rho(y)\,\mathrm{d}t\,\mathrm{d}y$$
$$+\hat{u}_{\varepsilon\chi(t_2)}(t_2^-,\gamma(t_2)+\omega(t_2)) + \int_{t_1}^{t_2}\hat\alpha_{\varepsilon\chi(t)}(t,\gamma(t)+\omega(t))\,\mathrm{d}t,$$

for all $\omega\in\mathcal{D}$ and $\varepsilon=1/n$. By a density argument the above inequality holds for any $\omega\in H^1([t_1,t_2];\mathbb{T}^d)$ with $\omega(t_1)=0$. We finally let $\varepsilon\downarrow 0$. As $M\hat\alpha(t,\gamma+\omega)\in L^1([t_1,t_2])$ and using the domination $|\alpha_\varepsilon|\le M\hat\alpha$, we can pass to the limit in the last term of the above inequality. By the dominate convergence theorem we can also pass to the limit in the first term thanks to the growth property and the continuity of L. In this way we obtain the inequality (7.7) with $\hat{u}(t_2^+,\gamma(t_2))$ instead of $\hat{u}(t_2,\gamma(t_2))$ and $\hat{u}(t_2^-,\gamma(t_2)+\omega(t_2))$. To conclude, it is sufficient to use $u(t_2^-,\cdot)\le u(t_2,\cdot)$ (a consequence of $\partial_t^s u\ge 0$) and $u(t_2^+,\cdot)=u(t_2,\cdot)$ m_{t_2}-a.e. (given by (7.4)). □

Remark 7.8. *The global version of Theorem 7.7 (to arrive up to the initial time* 0 *and the final time* T*) remains an open question. This is mainly due to the local integrability property for the additional price* $\beta\in L^2_{\mathrm{loc}}((0,T);BV(\mathbb{T}^d))$ *we are aware of for the moment. Let us remark that an integrability property* $\beta\in L^1([0,T];L^{1+\varepsilon}(\mathbb{T}^d))$ *for some* $\varepsilon>0$ *would be enough to conclude in the global version.*

The *notion of Nash equilibria* has now a clearer formulation. Since we are able to give a weak meaning for the optimization problem along single agent trajectories, a solution (u,m,β,β_T) of the MFG system with density constraints gives the following notion of equilibrium.

Definition 7.9 (Local weak Nash equilibria)**.** *Let* (u,m,β,β_T) *be a solution of the MFG system with density constraints in the sense of Definition 4.1 on* $[0,T]\times\mathbb{T}^d$*. We say that* (m,β,β_T) *is a* local weak Nash equilibrium*, if there exists an optimal density-constrained flow* $\boldsymbol{\eta}\in\mathcal{P}_2(\Gamma)$ *in the sense of Definition 7.2 (constructed with the help of* (m,β,β_T)*) which is concentrated on locally minimizing paths for Problem* (1.4) *in the sense of Definition 7.5. In particular one has that* $m_t=(e_t)_\#\boldsymbol{\eta}$ *and* $0\le m_t\le\overline{m}$ *a.e. in* $\mathbb{T}^d$ *for all* $t\in[0,T]$*.*

Remark 7.10. *Let us remark that by Proposition 7.4 and Theorem 7.7 for any solution* (u,m,β,β_T) *for the MFG system with density constraints obtained with the additional assumptions* **(HP1)**, **(HP2)** *and* **(HP3)** *the triple* (m,β,β_T) *is always a local weak Nash equilibrium in the sense of the above definition.*

7.3. **The case without density constraint.** Let us have a few words on the Nash equilibrium and on the optimality condition on the level of single agent trajectories in the case when we do not impose density constraints. More precisely, our aim is to clarify Remark 4.9. from [11].

Let us recall that in Section 4.3. from [11] it was considered a class of flows $\tilde{\boldsymbol{\eta}} \in \mathcal{P}_{r'}(\mathbb{T}^d)$ such that $\tilde{m} \in L^q([0,T] \times \mathbb{T}^d)$ where $\tilde{m}_t := (e_t)_\# \tilde{\boldsymbol{\eta}}$, where $r' > 1$ is the growth of the Lagrangian L in the velocity variable, while $q-1$ (where $q > 1$) is the growth of the continuous coupling f in the second variable. Because of this growth condition and since $m \in L^q([0,T] \times \mathbb{T}^d)$ we have first that $\alpha(t,x) := f(x, m(t,x)) \in L^{q'}([0,T] \times \mathbb{T}^d)$. Moreover Lemma 7.1 and Proposition 7.4 hold with $\beta \equiv 0$ and $\beta_T \equiv 0$, since we did not impose any density constraint (see the corresponding Lemma 4.6-4.8 from [11]).

The difference, compared to our analysis in the previous section, is that we can consider globally minimizing paths in Definition 7.5. More precisely, by the global integrability property of $\hat{\alpha}$, and hence $M\hat{\alpha} \in L^{q'}([0,T] \times \mathbb{T}^d)$ we allow curves $\gamma \in W^{1,r'}([0,T])$ (and their variations) such that $M\hat{\alpha}(\cdot, \gamma) \in L^{q'}([0,T])$. This is once again a natural class, since for any flow $\tilde{\boldsymbol{\eta}}$, with the above described properties, satisfies that

$$\int_\Gamma \int_{t_1}^{t_2} M\hat{\alpha}(t, \gamma(t)) \, \mathrm{d}t \, \mathrm{d}\tilde{\boldsymbol{\eta}}(\gamma) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} M\hat{\alpha}(t,x) \tilde{m}(t,x) \, \mathrm{d}x \, \mathrm{d}t < +\infty,$$

for all $0 \le t_1 < t_2 \le T$, since $M\hat{\alpha} \in L^{q'}([0,T] \times \mathbb{T}^d)$ and $\tilde{m} \in L^q([0,T] \times \mathbb{T}^d)$ where $\tilde{m}_t = (e_t)_\# \tilde{\boldsymbol{\eta}}$.

By these observations in the statement of Theorem 7.7 one can change now the word "locally" to "globally" and the proof goes along the same lines.

References

[1] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probability measure*, Birkhäuser, (2008).

[2] L. Ambrosio, A. Figalli, On the regularity of the pressure field of Brenier's weak solutions to incompressible Euler equations, *Calc. Var. PDE*, 31 (2008) No. 4, 497-509.

[3] L. Ambrosio, A. Figalli, Geodesics in the space of measure-preserving maps and plans, *Arch. Rational Mech. Anal.*, 194 (2009), 421-462.

[4] J.-D. Benamou, Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84 (2000), 375-393.

[5] J.-D. Benamou, G. Carlier, Augmented Lagrangian Methods for Transport Optimization, Mean Field Games and Degenerate Elliptic Equations, *J. Opt. Theor. Appl.*, (2015), *to appear.*

[6] L. Brasco, G. Carlier, F. Santambrogio, Congested traffic dynamics, weak flows and very degenerate elliptic equations, *J. Math. Pures Appl.*, (9) 93 (2010), No. 6, 652-671.

[7] Y. Brenier, Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations, *Comm. Pure Appl. Math.*, 52 (1999) 4, 411-452.

[8] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, *C.R.A.S. Paris, Série I*, 305 (1987), 805-808.

[9] Y. Brenier, Polar factorization and monotone rearrangement of vector valued functions, *Comm. Pure Appl. Math.*, 44 (1991), No. 4, 375-417.

[10] M. Burger, M. Di Francesco, P. Markowich, M.-T. Wolfram, Mean field games with nonlinear mobilities in pedestrian dynamics, *Discrete Contin. Dyn. Syst. Ser. B*, 19 (2014), No. 5, 1311-1333.

[11] P. Cardaliaguet, Weak solutions for first order mean field games with local coupling, (2013), *preprint available at* http://arxiv.org/abs/1305.7015.

[12] P. Cardaliaguet, G. Carlier, B. Nazaret, Geodesics for a class of distances in the space of probability measures, *Calc. Var. PDE*, 48 (2013), no. 2-3, 395-420.

[13] P. Cardaliaguet, P.J. Graber, Mean field games systems of first order, *ESAIM: Contr. Opt. and Calc. Var.*, (2015), *to appear.*

[14] P. Cardaliaguet, P.J. Graber, A. Porretta, D. Tonon, Second order mean field games with degenerate diffusion and local coupling, *NoDEA Nonlinear Differential Equations Appl.*, (2015), to appear.

[15] P. Cardaliaguet, A. Porretta, D. Tonon, Sobolev regularity for the first order Hamilton-Jacobi equation, (2014), *preprint available at* http://arxiv.org/abs/1411.0227.

[16] G. Carlier, C. Jimenez, F. Santambrogio, Optimal transportation with traffic congestion and Wardrop equilibria, *SIAM J. Control Optim.*, 47 (2008), No. 3, 1330-1350.

[17] I. Ekeland, R. Témam, *Convex analysis and variational problems*, Translated from the French. Studies in Mathematics and its Applications, Vol. 1. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, (1976).

[18] D.A. Gomes, H.Mitake, Existence for stationary mean field games with quadratic Hamiltonians with congestion, (2014), *preprint available at* http://arxiv.org/abs/1407.8267.

[19] D.A. Gomes, V. Voskanyan, Short-time existence of solutions for mean-field games with congestion, (2015), *preprint available at* http://arxiv.org/abs/1503.06442.

[20] P.J. Graber, Optimal control of first-order Hamilton-Jacobi equations with linearly bounded Hamiltonian, *Appl. Math. Optim.*, 70 (2014), no. 2, 185-224.

[21] M. Huang, R.P. Malhamé, P.E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, *Communication in information and systems*, 6 (2006), No. 3, 221-252.

[22] J.-M. Lasry, P.-L. Lions, Jeux à champ moyen. I. Le cas stationnaire, *C. R. Math. Acad. Sci. Paris*, 343 (2006), No. 9, 619-625.

[23] J.-M. Lasry, P.-L. Lions, Jeux à champ moyen. II. Horizon fini et contrôle optimal, *C. R. Math. Acad. Sci. Paris*, 343 (2006), No. 10, 679-684.

[24] J.-M. Lasry, P.-L. Lions, Mean field games, *Jpn. J. Math.*, 2 (2007), no. 1, 229-260.

[25] P.-L. Lions, *Cours au Collège de France*, www.college-de-france.fr.

[26] R. J. McCann, A convexity principle for interacting gases, *Adv. Math.* 128 (1997), No. 1, 153-179.

[27] A.R. Mészáros, F.J. Silva, A variational approach to second order Mean Field Games with density constraints: the stationary case, *J. Math. Pures et Appl.*, to appear, *preprint available at* http://arxiv.org/abs/1502.06026.

[28] F. Santambrogio, A modest proposal for MFG with density constraints, *Netw. Heterog. Media*, 7 (2012) No. 2, 337-347.

[29] F. Santambrogio, *Optimal Transport for Applied Mathematicians.* Birkäuser, NY. To be published in Sept. 2015.

[30] C. Villani *Topics in Optimal Transportation.* Graduate Studies in Mathematics, AMS, (2003).

Ceremade, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16 - France
E-mail address, P. Cardaliaguet: cardaliaguet@ceremade.dauphine.fr

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, France
E-mail address, A.R. Mészáros: alpar.meszaros@math.u-psud.fr

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, France
E-mail address, F. Santambrogio: filippo.santambrogio@math.u-psud.fr