THE PLANNING PROBLEM IN MEAN FIELD GAMES AS REGULARIZED MASS TRANSPORT

P. JAMESON GRABER, ALPÁR R. MÉSZÁROS, FRANCISCO J. SILVA, AND DANIELA TONON

Version: November 17, 2018

Abstract. In this paper, using variational approaches, we investigate the first order planning problem arising in the theory of mean field games. We show the existence and uniqueness of weak solutions of the problem in the case of a large class of Hamiltonians with arbitrary superlinear order of growth at infinity and local coupling functions. We require the initial and final measures to be merely summable. As an alternative way, we show that solutions of the planning problem can be approximated, via a Γ-convergence procedure, by solutions of standard mean field games with suitable penalized final couplings. In the same time (relying on the techniques developed recently in [GM18]), under stronger monotonicity and convexity conditions on the data, we obtain Sobolev estimates on the solutions of mean field games with general final couplings and the planning problem as well, both for space and time derivatives.

Contents

Keywords and phrases: planning problem of mean field games; regularized mass transport; Hamilton-Jacobi equations; Sobolev regularity
2010 AMS Subject Classification: 49K20; 35Q91; 49N60; 49N15; 49N70.

1. Introduction

The purpose of this article is to study the first order planning problem in mean field game theory, which can be formulated as a system of nonlinear partial differential equations:

$$
(1.1) \qquad \begin{cases} -\partial_t u + H(x, \nabla u) = f(x, m), & \text{in } (0,T)\times\mathbb{T}^d, \\ \partial_t m - \nabla\cdot(D_\xi H(x,\nabla u)m) = 0, & \text{in } (0,T)\times\mathbb{T}^d, \\ m(0,\cdot) = m_0, \quad m(T,\cdot) = m_T, & \text{in } \mathbb{T}^d. \end{cases}
$$

The data consist of probability measures $m_0, m_T \in \mathscr{P}(\mathbb{T}^d)$, a fixed time horizon $T > 0$, a coupling function $f : \mathbb{T}^d \times [0,+\infty) \to \mathbb{R}$ and a Hamiltonian $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$. Our aim is to find conditions on the data for which weak solutions to (1.1) can be shown to exist and are unique.

The theory of Mean Field Games (briefly MFG in what follows) was thrust into the limelight by the works of J.-M. Lasry and P.-L. Lions on the one hand (see [LL06a, LL06b, LL07]) and M. Huang, R. Malhamé and P. Caines on the other (see [HMC06]). Their main motivation was to study limits of Nash equilibria of (stochastic or deterministic) differential games when the number of players tends to infinity. Since then, it has become a very lively and active branch of the theory nonlinear partial differential equations. In addition to studying Nash equilibria, Lions [Lio] proposed a corresponding *planning problem*, in which a central planner would like to steer a population to a predetermined final configuration while still allowing individuals to choose their own strategies.

Let us give a simple, brief interpretation of System (1.1) in terms of large numbers of interacting agents. The solution of the Hamilton-Jacobi Equation (1.1)(i) is supposed to be the value function for an optimal control problem of the form

$$
(1.2) \qquad \inf_{\alpha}\left\{\int_t^T \left[L(x(s),\alpha(s)) + f(x(s), m(s,x(s)))\right]\mathrm{d}s + u(T,x(T))\right\} =: u(t,x)
$$

subject to

$$
\begin{cases} x'(s) = \alpha(s), & s \in (t,T] \\ x(t) = x \in \mathbb{T}^d. \end{cases}
$$

Here the Lagrangian $L : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is the Legendre-Fenchel transform of H w.r.t. the second variable. Formally the optimal strategy is given in feedback form, hence for the agent it is optimal to play $-D_\xi H(x(s), \nabla u(s, x(s)))$. Having this velocity field as a drift, the evolution of the agents' density is given by the solution of the second equation in (1.1). Then the coupling of the equations in (1.1) implies that every player is acting optimally with respect to the competing choices, i.e. the game is in equilibrium. We underline the fact that when considering 'standard' mean field games, typically, the final cost $u(T,\cdot) = u_T$ is treated as given, along with an initial population density $m(0,\cdot) = m_0$. For the planning problem, however, we fix a target population density $m(T,\cdot) = m_T$, leaving $u(T,\cdot)$ as an adjustable variable by which a central planner may determine the final outcome. Thus in the above control problem in particular $u(T,\cdot)$ is part of the problem itself.

As in many studies of mean field games, we restrict our attention to the case where $f(x,m)$ is an increasing function in the m variable, namely $\partial_m f(x,\cdot) > 0$ for all $x \in \mathbb{T}^d$. We interpret this to mean that agents have a preference for low-density regions, i.e. they want to avoid congestion.

In spite of the large number of studies available on MFG, the literature on System (1.1) is sparse. The cases when f is increasing and $D^2_{\xi\xi}H(x,\cdot) > 0$ (and both are smooth) are well-understood in the literature for both first order and second order non-degenerate models. In his lectures P.-L. Lions showed how to transform (1.1) into a uniformly elliptic system (thanks to smoothness assumptions on f and H) on space-time, and he showed the existence of classical solutions. One can summarize these results as follows.

Theorem 1.1. [Lio] *Let m_0, m_T be strictly positive probability densities of class $C^{1,\alpha}(\mathbb{T}^d)$ ($0 < \alpha < 1$), let moreover f and H be smooth such that $\partial_m f(x,\cdot) > 0$ for all $x \in \mathbb{T}^d$ and $D^2_{\xi\xi}H(x,\cdot) > 0$.*

Then, there exists a unique solution $(u,m) \in C^{2,\alpha}([0,T]\times\mathbb{T}^d)\times C^{1,\alpha}([0,T]\times\mathbb{T}^d)$ *to* (1.1) *(here the uniqueness of* u *has to be understood modulo constants). Moreover,* $(u,m)\in C^\infty((0,T)\times\mathbb{T}^d)\times C^\infty((0,T)\times\mathbb{T}^d)$.

Similar results can be achieved for the case of nondegenerate second order model as well, for purely quadratic Hamiltonians, or which are close at infinity to purely quadratic ones. The techniques used in this case are slightly different and they rely on the Hopf-Lax transformation, which is possible because of the quadratic Hamiltonian structure.

Existence of weak solutions to (1.1) in this latter case of nondegenerate second order models were obtained by A. Porretta in [Por13, Por14] using energy methods. These results can be summarized as follows.

Theorem 1.2. *(*[Por14, Theorem 1.3], [Por13, Theorem 2]*) Let us consider the nondegenerate diffusive model and let* $f(x,m)$ *be continuous, and nondecreasing w.r.t. the* m *variable. Let* $H(x,\xi)$ *be* C^1 *and convex in* ξ *with quadratic growth. Let* $m_0, m_T \in C^1(T^d)$ *be strictly positive probability densities. Then there exists a weak solution* $(u,m)\in L^2([0,T];H^1(\mathbb{T}^d))\times C^0([0,T];L^1(\mathbb{T}^d))$ *to* (1.1)*. Moreover, if* H *is strictly convex in the* p *variable, the solution is unique (modulo constants in the case of* u*).*

In the recent paper [OPS] C. Orrieri, A. Porretta and G. Savaré study weak solutions of System (1.1) set on the whole space $\mathbb{R}^d$ with Hamiltonians of quadratic growth and couplings with general growth. Their analysis relies on the variational structure of the problem and on a suitable weak theory – which they develop in the paper – of distributional sub-solutions and their traces of Hamilton-Jacobi equations with summable right hand sides.

In [GS], D.A. Gomes and T. Seneci explore displacement convexity properties, introduced by the Benamou-Brenier formulation for optimal transport, in order to obtain a priori estimates for solutions of (1.1). Finally, some numerical aspects of the mean field planning problem were investigated by Y. Achdou, F. Camilli and I. Capuzzo-Dolcetta in [ACCD12].

Our goal in this paper is to prove the existence and uniqueness of solutions of (1.1) for general Hamiltonians H and with as few assumptions as possible on the initial/final conditions. In order to accomplish this, we will make use of variational methods, which have proved to be very useful in mean field games, both to prove existence of weak solutions [Car15, CG15, CGPT15, CMS16] and also to establish additional regularity [GM18, PS17]. The main idea behind this point of view (which is also exploited in [OPS]) is that System (1.1) can be seen formally as first order necessary optimality conditions of two convex optimization problems in duality, cf. [LL07, Section 2.6].

The first problem is a control problem associated to the continuity equation, i.e.

$$\inf_{(m,w)}\int_0^T\int_{\mathbb{T}^d}\left[mH^*(x,-w(t,x)/m(t,x))+F(x,m(t,x))\right]\mathrm{d}x\,\mathrm{d}t$$

subject to $\partial_t m+\nabla\cdot w=0$ and $m(0,\cdot)=m_0$, $m(T,\cdot)=m_T$, where $H^*(x,\cdot)$ is the Fenchel conjugate of $H(x,\cdot)$ and $F(x,\cdot)$ is the antiderivative of $f(x,\cdot)$ w.r.t. the second variable.

The formal dual of this problem can be seen as a control problem associated to the Hamilton-Jacobi equation

$$\inf_u\int_0^T\int_{\mathbb{T}^d}F^*(x,-\partial_t u+H(x,\nabla u))\,\mathrm{d}x\,\mathrm{d}t+\int_{\mathbb{T}^d}u(T,x)\,\mathrm{d}m_T(x)-\int_{\mathbb{T}^d}u(0,x)\,\mathrm{d}m_0(x),$$

where $F^*(x,\cdot)$ is the Fenchel conjugate of $F(x,\cdot)$.

As the first results of our paper, in Section 2 we show the well-posedness of System (1.1) relying on the duality between the previous two convex optimization problems. To show the existence of a solution to the dual problem, we relax it in a suitable way and use a sort of 'renormalization trick' that was first used in [CCN13].

At this point, let us remark that in our existence proof for the dual problem, we require a joint condition on the order of growth of H in the momentum variable and the order of growth of f in the second variable (similarly as in [Car15, CG15]). This is mainly due to the lack of enough summability on m. It is worth mentioning that L^∞ estimates on m would allow us to drop this joint condition on H and f. In this context, for the planning problem, in the case of purely quadratic Hamiltonians and $m_0, m_T \in L^\infty(\mathbb{T}^d)$ such L^∞ estimates were obtained by Lions in his lectures (see [Lio]). Using completely different techniques, but still in the quadratic Hamiltonian case, such L^∞ estimates on m were obtained recently for mean field games by Lavenant and Santambrogio in [LS17]. Since in this paper our aim is to consider as general Hamiltonians and initial and final measures as possible, we are not pursuing the higher order summability estimates on m. Such questions would deserve a completely independent study, by themselves.

Coming back to our results, as a second approach, we use one similar in spirit to that of Porretta in [Por14]. We first prove existence of solutions for a mean field game in which the final cost is also given by a coupling function. Thus, we show the well-posedness of mean field games of the form

$$\left\{\begin{array}{ll} -\partial_t u + H(x,\nabla u) = f(x,m), & \text{in } (0,T)\times\mathbb{T}^d, \\ \partial_t m - \nabla\cdot(D_\xi H(x,\nabla u)m) = 0, & \text{in } (0,T)\times\mathbb{T}^d, \\ m(0,\cdot) = m_0, \quad u(T,\cdot) = g(\cdot, m(T,\cdot)), & \text{in } \mathbb{T}^d. \end{array}\right. \tag{1.3}$$

Then, we choose a sequence g_ε such that as $\varepsilon \to 0$, the corresponding solutions m_ε and a renormalized version of u_ε converge to a solution of the planning problem (1.1). Unlike Porretta, however, we use variational methods; in particular we show a sort of Γ-convergence of certain auxiliary problems to the optimal control problem and its dual given above. All these are achieved in Section 3. In a sense, then, our paper acts also as a companion to [Car15, CG15] and [CGPT15], generalizing the approaches developed there. Because of the lack of such general results in the literature, we prove all the results on System (1.3). To keep the current paper at a reasonable length, we study the MFG system (1.3) tailored to the particular class of penalizations g_ε that we choose. However, we emphasize the fact that all those results remain valid in more general scenarios, and their proofs would not require additional effort. In order to apply this second approach we need to require higher L^p summability (for some $p > 1$) for the final measure m_T (instead of mere integrability), however the interest of this approach resides in giving an approximation result useful for the numerical analysis of the Problem (1.1), see [ACCD12].

Moreover, in contrast to [Car15, CG15] and [CGPT15], we also address questions of regularity of weak solutions, based on techniques developed in the recent work [GM18] by the first two authors. The inspiration for these results comes from the alternative interpretation of the planning problem in terms of optimal mass transport. Indeed, the variational formulation of both the planning problem and mean field games has its roots in the dynamic formulation of the Monge-Kantorovich optimal transport problem (see [BB00]). In the same way, such convex variational problems are underneath other models studying weak solutions to the incompressible Euler equations (see for instance [Bre99, AF08]). The strong convexity present in these problems led Y. Brenier to develop a regularity theory for the pressure field in his model. Inspired by these techniques, the very same ideas were used later successfully to obtain Sobolev regularity for weak solutions of mean field games (see [PS17, San18, GM18]). After this series of results it is not unexpected that such results should be obtained for the solutions of the planning problem. This fact also motivates our title, i.e. the planning problem can be seen as a 'regularized' optimal transport problem, where the presence of the coupling and convexity of the Hamiltonian imply immediate Sobolev estimates on the distributional weak solutions. We think that these regularity results in particular could have further impacts on other problems arising in optimal transportation. In this context, let us also mention the very recent paper [LL] where the authors observe a different (but similar in spirit) regularization effect for very similar optimal transport type problems, in the presence of strong 'mean field interaction effects'.

The organization of the paper is as follows.

Next, in Section 2, we show the well-posedness of System (1.1) via the 'direct' variational approach, relying on the two convex optimization problems in duality. While the existence and uniqueness of a solution to the control problem associated to the continuity equation is a simple consequence of the well-known Fenchel-Rockafellar duality theorem, showing the existence of a solution to the other problem is more challenging. Inspired from some previous works in the literature (see [CCN13]) we find an interesting class of convex Hamiltonians, allowing general order of growth at infinity, for which we are able achieve this task. Then, solutions to the two optimization problems are shown to be equivalent to weak solutions of the planning problem. We emphasize the fact that our well-posedness results stand for initial and final distributions chosen merely in $L^1(\mathbb{T}^d)$. This generality, however, requires some constraints on the growth of the coupling term f.

In Section 3 we consider the existence and uniqueness of solutions to mean field games with final costs given by coupling functions. The arguments more or less follow those developed in [CGPT15], cf. [Car15, CCN13, CG15, Gra14, CMS16]. However, there are a few technical assumptions which we were able to relax, and this turns out to be crucial for our application to the planning problem. In particular, we were able to improve the uniqueness proof in [CGPT15], which is valid now for System (1.3) with final coupling. Also, another significant improvement consists in relaxing all the previous technical assumptions on the initial density m_0, which is assumed to be any general L^1 probability measure on $\mathbb{T}^d$. In the same time, we rely on all the assumptions on H and f as in Section 2. We end this section by showing that solutions to the planning problem can be obtained (via a Γ-convergence type procedure) as limits of solutions of mean field games with final coupling functions chosen as L^p-type penalizations.

We also consider additional regularity of weak solutions of both Systems (1.3) and (1.1) in Section 4. These estimates, which are interesting independently of existence and uniqueness for the planning problem, are based on the recent work in [GM18]. Here, unlike in [GM18] we present also general local in time Sobolev estimates for time derivatives, which complete in some sense the results from [GM18]. Similar results on time derivatives in the particular cases of quadratic Hamiltonians without space dependence were obtained for the m variable in [PS17]. In contrast to these, a nontrivial adaptation of the techniques from [GM18] allow us to obtain estimates on the u variable as well, under fairly general conditions on the Hamiltonian and the coupling function. We also observe that the regularity estimates obtained on the weak solutions of System (1.3) are in some sense inherited by weak solutions to the planning problem.

Finally, in Section 5 we conclude with some open questions and ideas for further research. Let us finish this introduction by summarizing our main results.

Our results on the planning problem are given in Section 2, Section 3 and Section 4.3, and can be informally summarized as follows.

Theorem 1.3 (Theorem 2.6, Proposition 2.8). *Let $m_0, m_T \in L^1(\mathbb{T}^d)$ be probability densities. Then under suitable growth and regularity assumptions on H and f (we refer to Section 2 for the precise hypotheses) we have that System* (1.1) *has a weak solution (in the sense of Definition 2.5). Moreover, if f is strictly increasing in its second variable, then m is unique and if H is strictly convex in the momentum variable, ∇u is unique on* $\mathrm{spt}(m)$.

In the same way, the well-posedness result on (1.3) and the convergence result to the solutions of (1.1) can be summarized as follows.

Theorem 1.4 (Theorem 3.7, Theorem 3.11). *Let $m_0 \in L^1(\mathbb{T}^d)$ be a probability density. Under standard growth and regularity assumption on H, f and g (we refer to Section 2 for the precise conditions), System* (1.3) *has a unique weak solution (understood in the sense of Definition 3.4).*

Then, choosing a suitable class of L^p-type penalizations g_ε and $m_T \in L^p(\mathbb{T}^d)$ (for arbitrary $p>1$), solutions of (3.1) *converge as $\varepsilon \downarrow 0$ to solutions of* (1.1)*.*

The Sobolev regularity estimates on Systems (1.3) and (1.1) can be informally summarized as follows.

Theorem 1.5 (Propositions 4.1-4.2, Propositions 4.3-4.4)**.** *(i) Suppose that $m_0 \in W^{2,1}(\mathbb{T}^d)$ and suppose further strong convexity and monotonicity conditions on H, f and g respectively (we refer to Section 4 for the details). Then the solution (u,m) of* (1.3) *satisfies*

$$\|m^{\frac{q}{2}-1}\nabla m\|_{L^2([0,T]\times\mathbb{T}^d)} \le C, \quad \|m^{1/2}D(j_1(\nabla u))\|_{L^2([0,T]\times\mathbb{T}^d)} \le C,$$

and

$$\|m(T,\cdot)^{\frac{p}{2}-1}\nabla m(T,\cdot)\|_{L^2(\mathbb{T}^d)} \le \ C,$$

where $(q-1)$ and $(p-1)$ are the order of growth of f and g $(p,q>1)$, respectively, and if $H(x,\cdot) \sim |\cdot|^r$ $(r>1)$, then j_1 is a function growing like $|\cdot|^{r/2}$.

(ii) Under similar assumptions on the data, but no assumptions on m_0, using the same notations as previously, we have

$$m^{1/2}\partial_t(j_1(\nabla u)) \in L^2_{\rm loc}((0,T);L^2(\mathbb{T}^d))$$

and

$$\partial_t(m^{q/2}) \in L^2_{\rm loc}((0,T);L^2(\mathbb{T}^d)).$$

(iii) Supposing $m_0, m_T \in W^{2,1}(\mathbb{T}^d)$, under the same strong convexity and monotonicity on H and f as before, we have the very same Sobolev estimates on the solutions of (1.1) *as for the solutions of* (1.3)*.*

Acknowledgements. The first author was supported by the National Science Foundation through Grant DMS-1612880. The second author was partially supported by the Air Force under the grant AFOSR MURI FA9550-18-1-0502. The last two authors were partially supported by the ANR project MFG ANR-16-CE40-0015-01 and the PEPS-INSMI Jeunes project "Some open problems in Mean Field Games" for the years 2016, 2017 and 2018. This article also benefited from the support of the FMJH Program PGMO and from the support of EDF, Thales, Orange et Criteo. Reference ANR-11-LABX-0056-LMH, LabEx LMH, Project PGMO VarPDEMFG. The second author would like to thank the hospitality of the Baylor University during his visit in April 2018.

2. Well-posedness of the planning problem via two convex optimization problems in duality

The general outline for the planning problem is largely the same as for variational mean field games (see e.g. [CCN13, Car15, CG15, CGPT15]): present two optimization problems in duality, prove that their minimizers are equivalent to solutions to (1.1), and show the existence of minimizers. We perform this last step in two ways: 'directly', working at the level of the optimization problems and 'indirectly', meaning that minimizers will be obtained as a limit of solutions to auxiliary problems of the form (1.3). The former approach will be done in this section and the latter one will be done in Section 3.

2.1. **Assumptions.** We assume the following.

(H1) (Conditions on the Hamiltonian I) $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is continuous in both variables, convex and differentiable in the second variable ξ, with $D_\xi H$ continuous in both variables. Moreover, H has superlinear growth in the gradient variable: there exist $r > 1$ and $C > 0$ such that

$$\frac{1}{rC}|\xi|^r - C \leq H(x,\xi) \leq \frac{C}{r}|\xi|^r + C, \quad \forall\, (x,\xi) \in \mathbb{T}^d \times \mathbb{R}^d. \tag{2.1}$$

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

$$\frac{1}{r'C}|\zeta|^{r'} - C \leq H^*(x,\zeta) \leq \frac{C}{r'}|\zeta|^{r'} + C, \quad \forall\, (x,\zeta) \in \mathbb{T}^d \times \mathbb{R}^d, \tag{2.2}$$

where we always use the notation $s' := s/(s-1)$ to denote the conjugate exponent of a number $s \in (1,\infty)$.

(H2) (Conditions on the Hamiltonian II) The Hamiltonian satisfies

$$H(x, a\xi) \leq aH(x,\xi), \quad \forall\, x \in \mathbb{T}^d,\ \forall\, \xi \in \mathbb{R}^d,\ \forall a \in [0,1].$$

A typical Hamiltonian that satisfies (H1) and (H2) is $H(x,\xi) = b(x)|\xi|^r - c(x)$ for some $b : \mathbb{T}^d \to \mathbb{R}$ continuous positive function and $c : \mathbb{T}^d \to \mathbb{R}$ continuous nonnegative function.

(H3) (Conditions on the coupling) The function f is continuous on $\mathbb{T}^d \times (0,\infty)$, strictly increasing in the second variable. Assume there exist $C > 0$ and $q > 1$ such that $r > \max\{d(q-1), 1\}$ and

$$\frac{1}{C}|m|^{q-1} - C \leq f(x,m) \leq C|m|^{q-1} + C, \quad \forall\, m \geq 0,\ \forall\, x \in \mathbb{T}^d. \tag{2.3}$$

With no real loss of generality, we ask for the following normalization:

$$f(x,0) = 0 \quad \forall\, x \in \mathbb{T}^d. \tag{2.4}$$

(H4) (Conditions on the initial and final measures) The probability measures m_0 and m_T are absolutely continuous with respect to $\mathscr{L}^d \llcorner \mathbb{T}^d$, with densities still denoted by m_0 and m_T, respectively.

We define $F : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}$ so that $F(x,\cdot)$ is an antiderivative of $f(x,\cdot)$ on $(0,\infty)$, that is,

$$F(x,m) = \int_0^m f(x,s)\,\mathrm{d}s,\ \forall\, m \geq 0. \tag{2.5}$$

For $m < 0$ we set $F(x,m) = +\infty$. Note that $F(x,m) \geq 0$ thanks to hypothesis (2.4). Moreover, it follows from (H3) that F is continuous on $\mathbb{T}^d \times [0,\infty)$, for each $x \in \mathbb{T}^d$ the function $F(x,\cdot)$ is strictly convex and differentiable in $(0,+\infty)$, and satisfies the growth condition

$$\frac{1}{qC}|m|^q - C \leq F(x,m) \leq \frac{C}{q}|m|^q + C, \quad \forall\, m \geq 0,\ \forall\, x \in \mathbb{T}^d. \tag{2.6}$$

We define $F^*(x,\cdot) : \mathbb{R} \to \mathbb{R}$ to be the Fenchel conjugate of $F(x,\cdot)$, i.e.

$$F^*(x,a) = \sup_{m \geq 0} \{am - F(x,m)\}.$$

Note that $F^*(x,\cdot)$ is continuous, increasing and $F^*(x,a) = 0$ for all $a \leq 0$. We also have

$$\frac{1}{q'C}|a|^{q'} - C \leq F^*(x,a) \leq \frac{C}{q'}|a|^{q'} + C, \quad \forall\, a \geq 0,\ \forall\, x \in \mathbb{T}^d. \tag{2.7}$$

2.2. **Two optimization problems in duality.** The planning problem has a variational formulation analogous to what is introduced in [Car15] in the context of MFGs.

The first optimization problem is described as follows: let us denote $\mathcal{K}_0 = \mathcal{C}^1([0,T]\times\mathbb{T}^d)$ and define, on $\mathcal{K}_0$, the functional

$$\mathcal{A}(u) = \int_0^T \int_{\mathbb{T}^d} F^*\left(x, -\partial_t u(t,x) + H(x,\nabla u(t,x))\right) \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} u(T,x)m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} u(0,x)m_0(x)\,\mathrm{d}x. \tag{2.8}$$

Notice that, $F^*(x,\cdot)$ being increasing and convex, for every $x\in\mathbb{T}^d$ the function $\mathbb{R}\times\mathbb{R}^d \ni (a,b)\mapsto E(x,a,b) := F^*\left(x,-a+H(x,b)\right)\in\mathbb{R}$ is convex and, hence, $\mathcal{A}$ is a convex function.

The first optimization problem is given by

$$\inf_{u\in\mathcal{K}_0} \mathcal{A}(u). \tag{2.9}$$

Now, suppose that $(m,w)\in L^1((0,T)\times\mathbb{T}^d)\times L^1((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$ are such that $m(t,x)\geq 0$ for a.e. $(t,x)\in[0,T]\times\mathbb{T}^d$ and the continuity equation

$$\partial_t m + \operatorname{div}(w) = 0 \text{ in } (0,T)\times\mathbb{T}^d, \tag{2.10}$$

is satisfied in the distributional sense, i.e. for all $\varphi\in C_c^1((0,T)\times\mathbb{T}^d)$ we have

$$\int_0^T\int_{\mathbb{T}^d} \left[\partial_t\varphi m + \nabla\varphi\cdot w\right]\mathrm{d}x\,\mathrm{d}t = 0. \tag{2.11}$$

Let us denote by $\mathscr{M}(\mathbb{T}^d)$ the space of Radon measures over $\mathbb{T}^d$ and by $\mathscr{M}_+(\mathbb{T}^d)$ the subset of $\mathscr{M}(\mathbb{T}^d)$ given by the non-negative Radon measures over $\mathbb{T}^d$. By [DNS09, Lemma 4.1] (see also the discussion in [CCN13]), if (2.11) holds, then there exists a unique weakly-* continuous curve $[0,T]\ni t\mapsto\tilde{m}(t)\in\mathscr{M}_+(\mathbb{T}^d)$ such that for a.e. $t\in[0,T]$ the measure $\tilde{m}(t)$ is absolutely continuous w.r.t. the Lebesgue measure, with density given by $m(t,\cdot)$, and for all $\varphi\in C^1([0,T]\times\mathbb{T}^d)$ the following equality holds

$$\int_{t_1}^{t_2}\int_{\mathbb{T}^d}\left[\partial_t\varphi m + \nabla\varphi\cdot w\right]\mathrm{d}x\,\mathrm{d}t = \int_{\mathbb{T}^d}\varphi(t_2,x)\,\mathrm{d}\tilde{m}(t_2)(x)\,\mathrm{d}x - \int_{\mathbb{T}^d}\varphi(t_1,x)\,\mathrm{d}\tilde{m}(t_1)(x), \tag{2.12}$$

for all $0\leq t_1<t_2\leq T$.

Define $\mathcal{K}_1$ as the set of 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 $m(t,x)\geq 0$ for a.e. $(t,x)\in[0,T]\times\mathbb{T}^d$, equation (2.11) is satisfied in the distributional sense and $\tilde{m}(0)$ and $\tilde{m}(T)$ are absolutely continuous with respect to the Lebesgue measure with densities given by m_0 and m_T, respectively. Note that (2.12) implies that the last two requirements are equivalent to the fact that (m,w) satisfies

$$\int_0^T\int_{\mathbb{T}^d}\left[\partial_t\varphi m + \nabla\varphi\cdot w\right]\mathrm{d}x\,\mathrm{d}t = \int_{\mathbb{T}^d}\varphi(T,x)m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d}\varphi(0,x)m_0(x)\,\mathrm{d}x, \tag{2.13}$$

for all $\varphi\in C^1([0,T]\times\mathbb{T}^d)$. Notice that if $(m,w)\in\mathcal{K}_1$, then $\int_{\mathbb{T}^d} m_0(x)\,\mathrm{d}x = 1$ and (2.12) imply that $\int_{\mathbb{T}^d} m(t,x)\,\mathrm{d}x = 1$ for a.e. $t\in[0,T]$. On $\mathcal{K}_1$, let us define

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

where, for $a=0$ and $b\in\mathbb{R}^d$, we impose that

$$aH^*\left(x,-\frac{b}{a}\right) = \begin{cases} +\infty & \text{if } b\neq 0,\\ 0 & \text{if } b=0.\end{cases}$$

Under this definition, it is easy to check that for every $x \in \mathbb{T}^d$, the function $\mathbb{R}_+ \times \mathbb{R}^d \ni (a,b) \mapsto aH^*\left(x, -\frac{b}{a}\right) \in \mathbb{R} \cup \{+\infty\}$ is proper, convex and lower-semicontinuous and, hence, $\mathcal{B}$ is convex.

The second optimization problem is the following:

$$\inf_{(m,w)\in\mathcal{K}_1} \mathcal{B}(m,w)\,. \tag{2.14}$$

The following lemma is proved using a similar argumentation as in [CCN13, Proposition 2.1] and [Car15, Lemma 2]. For the sake of completeness, we provide the details of the proof.

Lemma 2.1. *We have*

$$\inf_{u\in\mathcal{K}_0} \mathcal{A}(u) = -\min_{(m,w)\in\mathcal{K}_1} \mathcal{B}(m,w). \tag{2.15}$$

Moreover, there exists a unique $(\bar{m},\bar{w}) \in \mathcal{K}_1$ such that $\mathcal{B}(\bar{m},\bar{w}) = \min_{(m,w)\in\mathcal{K}_1} \mathcal{B}(m,w)$. Setting $\ell := \frac{r'q}{r'+q-1} > 1$, this minimizer also satisfies $(\bar{m},\bar{w}) \in L^q((0,T)\times\mathbb{T}^d) \times L^\ell((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$ and

$$\|\bar{m}\|_{L^q} + \|\bar{w}\|_{L^\ell} \leq C, \tag{2.16}$$

where $C>0$ is a constant independent of m_0 and m_T.

Proof. Let $\mathcal{H}_0 := \mathcal{C}^0([0,T]\times\mathbb{T}^d) \times \mathcal{C}^0([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$ and define the bounded linear operator $\Lambda : \mathcal{K}_0 \to \mathcal{H}_0$ by $\Lambda u = (\partial_t u, \nabla u)$, and the functionals $J_1 : \mathcal{H}_0 \to \mathbb{R}$, $J_2 : \mathcal{K}_0 \to \mathbb{R}$, respectively, by

$$\begin{aligned} J_1(\phi_1,\phi_2) &= \textstyle\int_0^T \int_{\mathbb{T}^d} E(x,\phi_1(x),\phi_2(x))\,\mathrm{d}x\,\mathrm{d}t, \\ J_2(u) &= \textstyle\int_{\mathbb{T}^d} u(T,x)m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} u(0,x)m_0(x)\,\mathrm{d}x, \end{aligned}$$

where we recall that $E(x,a,b) := F^*(x,-a+H(x,b))$ for all $x\in\mathbb{T}^d$, $a\in\mathbb{R}$ and $b\in\mathbb{R}^d$. Thus, problem (2.9) can be rewritten as

$$\inf_{u\in\mathcal{K}_0} \{J_1(\Lambda u) + J_2(u)\}\,. \tag{2.17}$$

Since

$$J_1\left(\Lambda\left(u - \min_{x\in\mathbb{T}^d} u(T,x)\right)\right) + J_2\left(u - \min_{x\in\mathbb{T}^d} u(T,x)\right) = J_1(\Lambda u) + J_2(u), \quad \forall\, u\in\mathcal{K}_0,$$

we can assume that the infimum in (2.17) is taken over $u\in\mathcal{K}_0$ such that $\inf_{x\in\mathbb{T}^d} u(T,x) = 0$. Using this fact, (2.7), estimate (2.37) (proved after Lemma 2.7 below), and setting $\bar{a} := \|-\partial_t u + H(\cdot,\nabla u)\|_{L^{q'}}$, we get the existence of $c_1>0$, $c_2\in\mathbb{R}$ and $c_3\in\mathbb{R}$ (independent of m_0 and m_T) such that

$$J_1(\Lambda u) + J_2(u) \geq c_1\bar{a}^{q'} + c_2\bar{a} + c_3 \geq \underline{c} := \inf_{\tau\in\mathbb{R}_+}\left\{c_1\tau^{q'} + c_2\tau\right\} + c_3 > -\infty. \tag{2.18}$$

This proves that the infimum in (2.17) is finite. Using that J_1 and J_2 are continuous, by the Fenchel-Rockafellar theorem (see e.g. [ET76, Chapter 3, Theorem 4.1]) we have that

$$\inf_{u\in\mathcal{K}_0} \mathcal{A}(u) = -\min\left\{J_1^*(-(m,w)) + J_2^*(\Lambda^*(m,w)) \,|\, (m,w)\in\mathcal{H}_0^*\right\}, \tag{2.19}$$

where $\mathcal{H}_0^* = \mathscr{M}((0,T)\times\mathbb{T}^d)) \times \mathscr{M}((0,T)\times\mathbb{T}^d)^d$. Let us provide a more explicit expression of the right hand side above. By [Roc71, Theorem 5], we have that

$$J_1^*(m,w) = \int_0^T\int_{\mathbb{T}^d} E^*(x,m^{ac},w^{ac})\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_{\mathbb{T}^d} E^*_\infty\left(x,\frac{\mathrm{d}m^s}{\mathrm{d}\theta},\frac{\mathrm{d}w^s}{\mathrm{d}\theta}\right)\mathrm{d}\theta(t,x),$$

where (m^{ac},w^{ac}) and (m^s,w^s) denote, respectively, the absolutely continuous and singular parts of (m,w) w.r.t. the Lebesgue measure, $\theta\in\mathscr{M}((0,T)\times\mathbb{T}^d)$ is any Radon measure such that (m^s,w^s)

is absolutely continuous w.r.t. θ, and $E^*_\infty(x,\cdot,\cdot)$ is the recession function of $E^*(x,\cdot,\cdot)$. We easily check that

$$E^*(x,m,w) = \begin{cases} -mH^*\left(x, -\frac{w}{m}\right) + F(-m), & \text{if } m<0, \\ 0, & \text{if } (m,w)=(0,0), \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.20}$$

Since $E^*(x,0,0) = 0 < +\infty$, the recession function can be computed as follows

$$E^*_\infty(x,h_m,h_w) = \lim_{\lambda\to+\infty} \frac{E^*(x,\lambda h_m,\lambda h_w)}{\lambda} = \begin{cases} 0, & \text{if } (h_m,h_w)=(0,0), \\ +\infty, & \text{otherwise.} \end{cases}$$

In the second equality above, we have used (2.7). We deduce that if $(m,w)\notin L^1((0,T)\times\mathbb{T}^d)\times L^1((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$, then $J_1^*(m,w)=+\infty$. If $(m,w)\in L^1((0,T)\times\mathbb{T}^d)\times L^1((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$, then

$$J_2^*(\Lambda^*(m,w)) = \sup_{u\in\mathcal{K}_0}\left\{\langle\Lambda^*(m,w),u\rangle_{\mathcal{K}_0^*,\mathcal{K}_0} - \int_{\mathbb{T}^d} u(T,x)m_T(x)\,\mathrm{d}x + \int_{\mathbb{T}^d} u(0,x)m_0(x)\,\mathrm{d}x\right\},$$

where $\langle\cdot,\cdot\rangle_{\mathcal{K}_0^*,\mathcal{K}_0}$ denotes the duality product between $\mathcal{K}_0$ and $\mathcal{K}_0^*$. Using that

$$\langle\Lambda^*(m,w),u\rangle_{\mathcal{K}_0^*,\mathcal{K}_0} = \langle(m,w),\Lambda u\rangle_{\mathcal{H}_0^*,\mathcal{H}_0} = \int_0^T\int_{\mathbb{T}^d}[m\partial_t u + w\cdot\nabla u]\,\mathrm{d}x\,\mathrm{d}t,$$

we get that $J_2^*(\Lambda^*(m,w))<+\infty$ if and only if $J_2^*(\Lambda^*(m,w))=0$, which is equivalent to the fact that (m,w) satisfies (2.13). We conclude that the optimization problem in the r.h.s. of (2.19) admits a solution $(\bar m,\bar w)$, is equivalent to problem (2.14) and, hence, (2.15) holds true.

By (2.18), we have $\mathcal{B}(\bar m,\bar w)\le -\underline{c}$ with $\underline{c}$ independent of m_0 and m_T. Using this estimate and arguing as in the proof of [Car15, Lemma 2], we easily obtain (2.16). Finally, the uniqueness of the solution $(\bar m,\bar w)$ to (2.14) follows exactly as in the proof of [Car15, Lemma 2]. □

Remark 2.2. *The previous proof shows that the results in Lemma 2.1 are valid also when m_0 and m_T belong to $\mathscr{P}(\mathbb{T}^d)$, without any summability assumptions.*

Now, we consider a relaxation of Problem (2.9). Let $\mathcal{K}$ be the set of pairs $(u,\alpha)\in BV((0,T)\times\mathbb{T}^d)\times L^{q'}((0,T)\times\mathbb{T}^d)$ such that $\nabla u\in L^r((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$, $u\in L^\infty((0,T)\times\mathbb{T}^d)$, the traces $u(0,\cdot)$, $u(T,\cdot)$ of u on $\{0\}\times\mathbb{T}^d$ and $\{T\}\times\mathbb{T}^d$ (see e.g. [AFP00, Section 3.8]), respectively, belong to $L^\infty(\mathbb{T}^d)$, and

$$-\partial_t u + H(x,\nabla u)\le\alpha$$

holds in the sense of distributions on $(0,T)\times\mathbb{T}^d$.

We extend the functional $\mathcal{A}$ to $\mathcal{K}$ by setting

$$\mathcal{A}(u,\alpha) = \int_0^T\int_{\mathbb{T}^d} F^*(x,\alpha(t,x))\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} u(T,x)m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} u(0,x)m_0(x)\,\mathrm{d}x \quad \forall\,(u,\alpha)\in\mathcal{K}.$$

We consider the following relaxation of Problem (2.9):

$$\inf_{(u,\alpha)\in\mathcal{K}}\mathcal{A}(u,\alpha) \tag{2.21}$$

Proposition 2.3. *We have*

$$\inf_{u\in\mathcal{K}_0}\mathcal{A}(u) = \inf_{(u,\alpha)\in\mathcal{K}}\mathcal{A}(u,\alpha) = -\min_{(m,w)\in\mathcal{K}_1}\mathcal{B}(m,w). \tag{2.22}$$

The proof of Proposition 2.3 follows easily once we have the following technical lemma.

Lemma 2.4. *Let* $(u,\alpha)\in\mathcal{K}$ *and* $(m,w)\in\mathcal{K}_1$. *Assume that* $mH^*(\cdot,-w/m)\in L^1((0,T)\times\mathbb{T}^d)$, $m\in L^q((0,T)\times\mathbb{T}^d)$, *and* m_0, $m_T\in L^1(\mathbb{T}^d)$. *Then* $\alpha m\in L^1((0,T)\times\mathbb{T}^d)$, *and for almost all* $t\in(0,T)$ *we have*

$$\int_{\mathbb{T}^d}(u(T)m_T-u(t)m(t))\,\mathrm{d}x+\int_t^T\int_{\mathbb{T}^d}m\left[\alpha+H^*\left(x,-\frac{w}{m}\right)\right]\mathrm{d}x\,\mathrm{d}t\ \geq\ 0,\tag{2.23}$$

and

$$\int_{\mathbb{T}^d}(u(t)m(t)-u(0)m_0)\,\mathrm{d}x+\int_0^t\int_{\mathbb{T}^d}m\left[\alpha+H^*\left(x,-\frac{w}{m}\right)\right]\mathrm{d}x\,\mathrm{d}t\ \geq\ 0.\tag{2.24}$$

Moreover, if equality holds in the inequality (2.23) *for* $t=0$, *then* $w=-mD_\xi H(\cdot,\nabla u)$ *a.e. and* $-\partial_t u^{ac}(t,x)+H(x,\nabla u(t,x))=\alpha(t,x)$ *for* m*-a.e.* $(t,x)\in(0,T)\times\mathbb{T}^d$, *where* $\partial_t u^{ac}$ *is the absolutely continuous part of the measure* $\partial_t u$.

Proof. The proof is an adaptation of the argument seen in [CG15, Lemma 2.4]. We will prove (2.23); the proof of (2.24) is analogous.

We first extend (m,w) to $\mathbb{R}\times\mathbb{T}^d$ by setting $(m,w)=(m_0,0)$ on $(-\infty,0)\times\mathbb{T}^d$ and $(m,w)=(m(T),0)$ on $(T,+\infty)\times\mathbb{T}^d$. Note that we still have $\partial_t m+\operatorname{div}\,w=0$ on $\mathbb{R}\times\mathbb{T}^d$. For $\varepsilon>0$, let $\xi_\varepsilon(t,x)$ be a sequence of smooth convolution kernels that we will define below. Define $m_\varepsilon:=\xi_\varepsilon*m$ and $w_\varepsilon:=\xi_\varepsilon*w$. Then m_ε and w_ε are C^∞ smooth, $m_\varepsilon>0$, and

$$\partial_t m_\varepsilon+\operatorname{div}\,w_\varepsilon=0.\tag{2.25}$$

Recalling that $-\partial_t u+H(x,\nabla u)\leq\alpha$ in the sense of distributions, we deduce

$$\begin{aligned}\int_{\mathbb{T}^d}u(t)m_\varepsilon(t)\,\mathrm{d}x-\int_{\mathbb{T}^d}u(T)m_\varepsilon(T)\,\mathrm{d}x&\leq\int_t^T\int_{\mathbb{T}^d}[-w_\varepsilon\cdot\nabla u-m_\varepsilon H(x,\nabla u)+m_\varepsilon\alpha]\,\mathrm{d}x\,\mathrm{d}t\\&\leq\int_t^T\int_{\mathbb{T}^d}\left[m_\varepsilon H^*\left(x,-\frac{w_\varepsilon}{m_\varepsilon}\right)+m_\varepsilon\alpha\right]\mathrm{d}x\,\mathrm{d}t\end{aligned}\tag{2.26}$$

for any $t\in(0,T)$. As $\varepsilon\to0$, we have that $m_\varepsilon\to m$ in $L^q((0,T)\times\mathbb{T}^d)$, and in particular $m_\varepsilon(t)\to m(t)$ in $L^q(\mathbb{T}^d)$ for almost every $t\in(0,T)$, while $m_\varepsilon\alpha\to m\alpha$ in $L^1((0,T)\times\mathbb{T}^d)$ since $\alpha\in L^{q'}((0,T)\times\mathbb{T}^d)$. Thus as $u\in L^\infty((0,T)\times\mathbb{T}^d)$, we get $\int_{\mathbb{T}^d}u(t)m_\varepsilon(t)\,\mathrm{d}x\to\int_{\mathbb{T}^d}u(t)m(t)\,\mathrm{d}x$ for almost every $t\in(0,T)$. On the other hand, by the argument given in [CG15, Lemma 2.4], we have

$$\lim_{\varepsilon\to0}\int_t^T\int_{\mathbb{T}^d}m_\varepsilon H^*\left(x,-\frac{w_\varepsilon}{m_\varepsilon}\right)\mathrm{d}x\,\mathrm{d}t=\int_t^T\int_{\mathbb{T}^d}mH^*\left(x,-\frac{w}{m}\right)\mathrm{d}x\,\mathrm{d}t.\tag{2.27}$$

Then

$$\int_{\mathbb{T}^d}u(t)m(t)\,\mathrm{d}x\leq\limsup_{\varepsilon\to0}\int_{\mathbb{T}^d}u(T)m_\varepsilon(T)\,\mathrm{d}x+\int_t^T\int_{\mathbb{T}^d}\left[mH^*\left(x,-\frac{w}{m}\right)+m\alpha\right]\mathrm{d}x\,\mathrm{d}t.\tag{2.28}$$

To conclude, we just need to show that

$$\int_{\mathbb{T}^d}u(T)m_\varepsilon(T)\,\mathrm{d}x\to\int_{\mathbb{T}^d}u(T)m_T\,\mathrm{d}x,\ \varepsilon\downarrow0.\tag{2.29}$$

Since $u(T)\in L^\infty(\mathbb{T}^d)$, it is enough to show $m_\varepsilon(T)\to m_T$ in $L^1(\mathbb{T}^d)$ as $\varepsilon\downarrow0$. For this we choose a particular construction of the convolution kernel ξ_ε.

Let $\eta:\mathbb{R}\to(0,\infty)$ and $\psi:\mathbb{R}^d\to(0,\infty)$ be even convolution kernels, each with compact support in the unit ball, $\delta>0$ and set $\eta_\delta(t)=\delta^{-1}\eta(t/\delta)$ and $\psi_\varepsilon(x)=\varepsilon^{-d}\psi(x/\varepsilon)$. We will choose $\xi_\varepsilon(t,x)=\eta_\delta(t)\psi_\varepsilon(x)$ where $\delta=\delta(\varepsilon)$ will be determined by the following calculations. Set $m_{T,\varepsilon}=\xi_\varepsilon * m_T$. Our first observation is that

$$
\begin{aligned}
(2.30)\quad \int_{\mathbb{T}^d}|m_\varepsilon(T,x)-m_T(x)|\,\mathrm{d}x &\le \int_{\mathbb{T}^d}\left|\int_{T-\delta}^{T+\delta}\int_{\mathbb{T}^d}\eta_\delta(T-s)\psi_\varepsilon(x-y)(m(s,y)-m_T(y))\,\mathrm{d}y\,\mathrm{d}s\right|\mathrm{d}x\\
&\quad+\int_{\mathbb{T}^d}\left|\int_{T-\delta}^{T+\delta}\int_{\mathbb{T}^d}\eta_\delta(T-s)\psi_\varepsilon(x-y)(m_T(y)-m_T(x))\,\mathrm{d}y\,\mathrm{d}s\right|\mathrm{d}x\\
&=\int_{\mathbb{T}^d}\left|\int_{T-\delta}^{T+\delta}\int_{\mathbb{T}^d}\int_s^T\eta_\delta(T-s)\nabla\psi_\varepsilon(x-y)\cdot w(\tau,y)\,\mathrm{d}y\,\mathrm{d}\tau\,\mathrm{d}s\right|\mathrm{d}x+\int_{\mathbb{T}^d}|m_{T,\varepsilon}(x)-m_T(x)|\,\mathrm{d}x\\
&\le\frac{C}{\varepsilon^{d+1}}\int_{T-\delta}^T\int_{\mathbb{T}^d}|w(\tau,y)|\,\mathrm{d}\tau\,\mathrm{d}y+\int_{\mathbb{T}^d}|m_{T,\varepsilon}(x)-m_T(x)|\,\mathrm{d}x.
\end{aligned}
$$

We set $\delta=\delta(\varepsilon)$ small enough such that $\frac{C}{\varepsilon^{d+1}}\int_{T-\delta}^T\int_{\mathbb{T}^d}|w(\tau,y)|\,\mathrm{d}\tau\,\mathrm{d}y\le\varepsilon$. Then (2.30) proves that $m_\varepsilon(T,\cdot)\to m_T$ in L^1 as $\varepsilon\to0$. The proof of (2.23) is complete.

□

Proof of Proposition 2.3 . Fixing $t\in(0,T)$ such that (2.23) and (2.24) hold, by adding both inequalities we get that

$$
(2.31)\qquad \int_{\mathbb{T}^d}(u(T)m_T-u(0)m(0))\,\mathrm{d}x+\int_0^T\int_{\mathbb{T}^d}m\left[\alpha+H^*\left(x,-\frac{w}{m}\right)\right]\mathrm{d}x\,\mathrm{d}t\ \ge\ 0,
$$

for every $(u,\alpha)\in\mathcal{K}$ and $(m,w)\in\mathcal{K}_1$ satisfying the assumptions of Lemma 2.4. Thus,

$$
\begin{aligned}
\mathcal{A}(u,\alpha)\ &\ge\ -\int_0^T\int_{\mathbb{T}^d}\left[m\left(\alpha+H^*\left(x,-\frac{w}{m}\right)\right)-F^*(x,\alpha)\right]\mathrm{d}x\,\mathrm{d}t\\
&\ge\ -\int_0^T\int_{\mathbb{T}^d}\left[H^*\left(x,-\frac{w}{m}\right)m+F(x,m)\right]\mathrm{d}x\,\mathrm{d}t,
\end{aligned}
$$

from which we deduce that $\inf_{(u,\alpha)\in\mathcal{K}}\mathcal{A}(u,\alpha)\ge-\min_{(m,w)\in\mathcal{K}_1}\mathcal{B}(m,w)$. Therefore, (2.22) follows from the inequalities

$$
-\min_{(m,w)\in\mathcal{K}_1}\mathcal{B}(m,w)=\inf_{u\in\mathcal{K}_0}\mathcal{A}(u,\alpha)\ge\inf_{(u,\alpha)\in\mathcal{K}}\mathcal{A}(u,\alpha)\ge-\min_{(m,w)\in\mathcal{K}_1}\mathcal{B}(m,w).
$$

□

2.3. **Weak solutions and minimizers.** The definition of weak solutions for the planning problem is analogous to that of the mean field game system (see [Car15, CG15]).

Definition 2.5. *Let $(u,m)\in BV((0,T)\times\mathbb{T}^d)\times L^q((0,T)\times\mathbb{T}^d)$. We say that (u,m) is a weak solution to (1.1) if*

(i) *the following integrability conditions hold:*

$$\nabla u\in L^r((0,T)\times\mathbb{T}^d;\mathbb{R}^d),\quad u\in L^\infty((0,T)\times\mathbb{T}^d),$$

the traces $u(0,\cdot)$, $u(T,\cdot)$ belong to $L^\infty(\mathbb{T}^d)$,

$$mH^*(\cdot,D_\xi H(\cdot,\nabla u))\in L^1((0,T)\times\mathbb{T}^d),\quad mD_\xi H(\cdot,\nabla u))\in L^1((0,T)\times\mathbb{T}^d;\mathbb{R}^d).$$

(ii) *Equation (1.1)-(i) holds in the following sense: inequality*

$$(2.32)\qquad -\partial_t u+H(x,\nabla u)\le f(x,m)\quad\text{in }(0,T)\times\mathbb{T}^d$$

holds in the sense of distributions.

(iii) *Equation* (1.1)-(ii) *holds:*

$$\partial_t m - \operatorname{div}(m D_\xi H(x, \nabla u))) = 0 \text{ in } (0,T)\times\mathbb{T}^d, \quad m(0) = m_0,\ m(T) = m_T \tag{2.33}$$

in the weak sense (2.13)*; and, finally,*

(iv) *the following equality holds:*

$$\int_0^T \int_{\mathbb{T}^d} m(t,x)\left[f(x,m(t,x)) + H^*(x, D_\xi H(x,\nabla u)(t,x))\right] \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} [m_T(x)u(T,x) - m_0(x)u(0,x)]\,\mathrm{d}x = 0. \tag{2.34}$$

To prove the existence of weak solutions, we will use the fact that they are equivalent to minimizers of the two optimization problems presented in Section 2.2. In Section 2.4 below we will show the existence of minimizers for problem (2.21), and, hence, the existence of solutions to (1.1).

Theorem 2.6. *Let* $(\bar m, \bar w) \in \mathcal{K}_1$ *be a minimizer of* (2.14) *and* $(\bar u, \bar\alpha) \in \mathcal{K}$ *be a minimizer of* (2.21)*. Then* $(\bar u, \bar m)$ *is a weak solution of the planning problem* (1.1) *and* $\bar w = -\bar m D_\xi H(\cdot, \nabla \bar u)$*, while* $\bar\alpha = f(\cdot, \bar m)$ *a.e.*

Conversely, any weak solution $(\bar u, \bar m)$ *of* (1.1) *is such that the pair* $(\bar m, -\bar m D_\xi H(\cdot,\nabla\bar u))$ *is the minimizer of* (2.14) *while* $(\bar u, f(\cdot,\bar m))$ *is a minimizer of* (2.21)*.*

Proof. Let $(\bar m, \bar w) \in \mathcal{K}_1$ be a minimizer of Problem (2.14) and $(\bar u, \bar\alpha,) \in \mathcal{K}$ be a minimizer of Problem (2.21). Due to Proposition 2.3, we have

$$\int_0^T \int_{\mathbb{T}^d} \left[F^*(x,\bar\alpha) + F(x,\bar m) + \bar m H^*\left(x, -\frac{\bar w}{\bar m}\right)\right] \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} [\bar u(T)m_T - \bar u(0)m_0]\,\mathrm{d}x = 0.$$

We show that $\bar\alpha = f(x,\bar m)$. Indeed, by the definition of Legendre transform,

$$F^*(x,\bar\alpha(t,x)) + F(x,\bar m(t,x)) \geq \bar\alpha(t,x)\bar m(t,x), \tag{2.35}$$

hence

$$\int_0^T \int_{\mathbb{T}^d} \left[\bar\alpha(t,x)\bar m(t,x) + \bar m H^*\left(x, -\frac{\bar w}{\bar m}\right)\right] \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} [\bar u(T)m_T - \bar u(0)m_0]\,\mathrm{d}x \leq 0.$$

Thanks to Lemma 2.4, the above inequality is in fact an equality, $\bar w = -\bar m D_\xi H(\cdot,\nabla\bar u)$ a.e. and Equation (2.35) becomes equality a.e. Therefore, by the convexity and differentiability of F,

$$\bar\alpha(t,x) = f(x,\bar m(t,x)) \tag{2.36}$$

almost everywhere and (2.34) holds for $(\bar u, \bar m)$. In particular, $\bar m H^*(\cdot, D_\xi H(\cdot,\nabla\bar u)) \in L^1((0,T)\times\mathbb{T}^d)$. Moreover, since $(\bar u,\bar\alpha) \in \mathcal{K}$ and Equation (2.36) holds, we have $-\partial_t \bar u + H(x,\nabla\bar u) \leq f(x,\bar m)$ in the sense of distributions. Furthermore, since $(\bar u,\bar\alpha) \in \mathcal{K}$ and $\bar w = -\bar m D_\xi H(\cdot,\nabla\bar u)$, we have that $\bar m D_\xi H(\cdot,\nabla\bar u) \in L^1((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$ and (2.33) holds in the sense of distributions. Therefore $(\bar u,\bar m)$ is a solution in the sense of Definition 2.5.

Suppose now that $(\bar u,\bar m)$ is a weak solution of (1.1) as in Definition 2.5. Set $\bar w = -\bar m D_\xi H(\cdot,\nabla\bar u)$, $\bar\alpha(t,x) = f(x,\bar m(t,x))$. By definition of weak solution $(\bar w,\bar\alpha) \in L^1((0,T)\times\mathbb{T}^d;\mathbb{R}^d)\times L^1((0,T)\times\mathbb{T}^d)$, $\bar m \in L^q((0,T)\times\mathbb{T}^d)$, and $\bar u \in L^\infty((0,T)\times\mathbb{T}^d)$. Moreover, since $\bar m \in L^q((0,T)\times\mathbb{T}^d)$, the growth condition (2.3) implies that $\bar\alpha \in L^{q'}((0,T)\times\mathbb{T}^d)$. Therefore $(\bar m,\bar w) \in \mathcal{K}_1$ and $(\bar u,\bar\alpha) \in \mathcal{K}$.

It remains to show that $(\bar{u},\bar{\alpha})$ minimizes $\mathcal{A}$ and $(\bar{m},\bar{w})$ minimizes $\mathcal{B}$. Let $(\bar{u}',\bar{\alpha}') \in \mathcal{K}$. By the convexity and differentiability of F^* in the second variable, we have

$$\begin{aligned}
\mathcal{A}(\bar{u}',\bar{\alpha}') &= \int_0^T \int_{\mathbb{T}^d} F^*(x,\bar{\alpha}'(t,x))\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} \left[\bar{u}'(T,x)m_T(x) - \bar{u}'(0,x)m_0(x)\right]\mathrm{d}x \\
&\geq \int_0^T \int_{\mathbb{T}^d} \left[F^*(x,\bar{\alpha}(t,x)) + \partial_\alpha F^*(x,\bar{\alpha}(t,x))(\bar{\alpha}'(t,x)-\bar{\alpha}(t,x))\right]\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_{\mathbb{T}^d} \left[\bar{u}'(T,x)m_T(x) - \bar{u}'(0,x)m_0(x)\right]\mathrm{d}x \\
&= \int_0^T \int_{\mathbb{T}^d} \left[F^*(x,\bar{\alpha}(t,x)) + \bar{m}(t,x)(\bar{\alpha}'(t,x)-\bar{\alpha}(t,x))\right]\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_{\mathbb{T}^d} \left[\bar{u}'(T,x)m_T(x) - \bar{u}'(0,x)m_0(x)\right]\mathrm{d}x \\
&= \mathcal{A}(\bar{u},\bar{\alpha}) + \int_0^T \int_{\mathbb{T}^d} \bar{m}(t,x)(\bar{\alpha}'(t,x)-\bar{\alpha}(t,x))\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_{\mathbb{T}^d} (\bar{u}'(T,x)-\bar{u}(T,x))m_T(x)\,\mathrm{d}x + \int_{\mathbb{T}^d} (\bar{u}(0,x)-\bar{u}'(0,x))m_0(x)\,\mathrm{d}x \\
&= \mathcal{A}(\bar{u},\bar{\alpha}) + \int_0^T \int_{\mathbb{T}^d} \left[\bar{m}(t,x)\bar{\alpha}'(t,x) + \bar{m}(t,x)H^*\left(x,-\frac{\bar{w}(t,x)}{\bar{m}(t,x)}\right)\right]\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_{\mathbb{T}^d} \bar{u}'(T,x)m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} \bar{u}'(0,x)m_0(x)\,\mathrm{d}x
\end{aligned}$$

where the last equality follows from Equation (2.34). Applying Lemma 2.4 applied to $(\bar{u}',\bar{\alpha}')$ and $(\bar{m},\bar{w})$, we deduce

$$\mathcal{A}(\bar{u}',\bar{\alpha}') \geq \mathcal{A}(\bar{u},\bar{\alpha}),$$

and so $(\bar{u},\bar{\alpha})$ is a minimizer of $\mathcal{A}$.

The argument for $(\bar{m},\bar{w})$ is similar. Let $(\bar{m}',\bar{w}')$ be a competitor for $\mathcal{B}$. Then because F is convex and differentiable in the second variable, we have, using Equation (2.34),

$$\begin{aligned}
\mathcal{B}(\bar{m}',\bar{w}') &= \int_0^T \int_{\mathbb{T}^d} \left[\bar{m}'H^*\left(x,-\frac{\bar{w}'}{\bar{m}'}\right) + F(x,\bar{m}')\right]\mathrm{d}x\,\mathrm{d}t \\
&\geq \int_0^T \int_{\mathbb{T}^d} \left[\bar{m}'H^*\left(x,-\frac{\bar{w}'}{\bar{m}'}\right) + F(x,\bar{m}) + f(x,\bar{m})(\bar{m}'-\bar{m})\right]\mathrm{d}x\,\mathrm{d}t \\
&= \int_{\mathbb{T}^d} \left[\bar{u}(T)m_T - \bar{u}(0)m_0\right]\mathrm{d}x \\
&\qquad + \int_0^T \int_{\mathbb{T}^d} \left[\bar{m}'H^*\left(x,-\frac{\bar{w}'}{\bar{m}'}\right) + \bar{m}H^*\left(x,-\frac{\bar{w}}{\bar{m}}\right) + F(x,\bar{m}) + \bar{\alpha}\bar{m}'\right]\mathrm{d}x\,\mathrm{d}t \\
&= \mathcal{B}(\bar{m},\bar{w}) + \int_{\mathbb{T}^d} \left[\bar{u}(T)m_T - \bar{u}(0)m_0\right]\mathrm{d}x + \int_0^T \int_{\mathbb{T}^d} \left[\bar{m}'H^*\left(x,-\frac{\bar{w}'}{\bar{m}'}\right) + \bar{\alpha}\bar{m}'\right]\mathrm{d}x\,\mathrm{d}t \\
&\geq \mathcal{B}(\bar{m},\bar{w}).
\end{aligned}$$

Here we applied Lemma 2.4 to $(\bar{u},\bar{\alpha})$ and $(\bar{m}',\bar{w}')$ in the last line. Therefore $(\bar{m},\bar{w})$ is a minimizer of $\mathcal{B}$.

□

2.4. **Existence of solutions by direct method.** We will now show that Problem (2.21) admits at least one solution. We will need the following preliminary result proved in [CG15, Lemma 2.7].

Lemma 2.7. *Let α be a continuous function, and set*

$$\nu = \frac{r - d(q-1)}{d(q-1)(r-1) + rq},$$

which by Hypothesis (H3) is positive. Then there exists $C > 0$ such that for any smooth subsolution of $-\partial_t u + H(x, \nabla u) \leq \alpha$,

$$u(t_1, x) \leq u(t_2, y) + C\left[|x-y|^{r'}(t_2 - t_1)^{1-r'} + \left((t_2 - t_1)^{\nu} \wedge 1 + T^{1/q}\right)\left(\|\alpha_+\|_{q'} + 1\right)\right]$$

for all $0 \leq t_1 < t_2 \leq T$ and x, $y \in \mathbb{T}^d$.

As a consequence of the previous lemma, we have that, for any $x \in \mathbb{T}^d$,

$$u(0, x) \leq u(T, y) + C\left[|x-y|^{r'} T^{1-r'} + \left(T^{\nu} \wedge 1 + T^{1/q}\right)\left(\|\alpha_+\|_{q'} + 1\right)\right] \quad \forall\, y \in \mathbb{T}^d,$$

and since x and y belong to a bounded set, up to redefining C, we get

$$u(0, x) \leq \inf_{y \in \mathbb{T}^d} u(T, y) + C\left[T^{1-r'} + \left(T^{\nu} \wedge 1 + T^{1/q}\right)\left(\|\alpha_+\|_{q'} + 1\right)\right] \quad \forall\, x \in \mathbb{T}^d. \tag{2.37}$$

Proposition 2.8. *Problem (2.21) admits at least one solution (u, α). The function u is Hölder continuous in $[0, T) \times \mathbb{T}^d$, $\alpha \geq 0$ a.e. and there exists $C > 0$, independent of m_0 and m_T, such that*

$$\sup_{(t,x) \in [0,T] \times \mathbb{T}^d} |u(t, x)| + \|\nabla u\|_{L^r} + \|\partial_t u\|_{\mathscr{M}} + \|\alpha\|_{L^{q'}} \leq C, \tag{2.38}$$

where $\|\cdot\|_{\mathscr{M}}$ denotes the usual norm of $\mathscr{M}$ as dual space of $\mathcal{C}^0([0, T] \times \mathbb{T}^d)$.

Proof. Consider a smooth minimizing sequence $(u_n)_n$ for Problem (2.9). Using that $\mathcal{A}(u_n) = \mathcal{A}(u_n + c)$ for all $c \in \mathbb{R}$, by subtracting $\min_{x \in \mathbb{T}^d} u_n(T, x)$ we can suppose that $\min_{x \in \mathbb{T}^d} u_n(T, x) = 0$. For all $x \in \mathbb{T}^d, t \in [0, T]$, let

$$\alpha_n(t, x) := -\partial_t u_n(t, x) + H(x, \nabla u_n(t, x)).$$

Then, inequality (2.37) applies to all u_n giving

$$u_n(0, x) \leq C\left[T^{1-r'} + \left(T^{\nu} \wedge 1 + T^{1/q}\right)\left(\|(\alpha_n)_+\|_{L^{q'}} + 1\right)\right] \quad \forall\, x \in \mathbb{T}^d. \tag{2.39}$$

Moreover, Proposition 2.3 implies that $(u_n, \alpha_n)_n$ is a minimizing sequence for Problem (2.21). Hence, since Hypothesis (H2) and $F^*(x, a) = 0$ for all $a \leq 0$, imply $\mathcal{A}(0) = 0$, we have

$$\begin{aligned} 0 &\geq \int_0^T \int_{\mathbb{T}^d} F^*(x, \alpha_n)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} u_n(T, x)\,\mathrm{d}m_T(x) - \int_{\mathbb{T}^d} u_n(0, x)\,\mathrm{d}m_0(x) \\ &\geq \frac{1}{q'C}\|(\alpha_n)_+\|_{L^{q'}}^{q'} - C\left(T^{1-r'} + \left(T^{\nu} \wedge 1 + T^{1/q}\right)\left(\|(\alpha_n)_+\|_{L^{q'}} + 1\right)\right), \end{aligned} \tag{2.40}$$

where we used the growth condition (2.7) on F^*, inequality (2.39), $u_n(T, \cdot) \geq 0$ and $\int_{\mathbb{T}^d} m_0(x)\,\mathrm{d}x = 1$. We deduce that $((\alpha_n)_+)_n$ is a bounded sequence in $L^{q'}([0, T] \times \mathbb{T}^d)$ uniformly with respect to m_0 and m_T. Moreover, there exists a constant $C_1 > 0$ such that $u_n(0, x) \leq C_1$ for all $x \in \mathbb{T}^d$.

In order to obtain uniform bounds on $(u_n)_n$, we need to modify the sequence. Let $\eta \in C^1(\mathbb{R})$ such that $0 \leq \eta' \leq 1$, $|\eta| \leq 2C_1$ and $\eta(s) = s$ if $|s| \leq C_1$, and set $\tilde{u}_n := \eta \circ u_n$. Since u_n is Lipschitz continuous, we have that $\tilde{u}_n$ is Lipschitz continuous and, therefore,

$$\begin{aligned} -\partial_t \tilde{u}_n + H(x, \nabla \tilde{u}_n) &\leq \eta'(\tilde{u}_n)\left(-\partial_t u_n + H(x, \nabla u_n)\right) \\ &\leq \eta'(\tilde{u}_n)(\alpha_n)_+ \\ &\leq (\alpha_n)_+, \end{aligned}$$

where we have used that $0 \leq \eta' \leq 1$ and assumption (H2).

Thus, $(\tilde{u}_n, (\alpha_n)_+)_n \in \mathcal{K}$, $\|\tilde{u}_n\|_{L^\infty} \leq 2C_1$, i.e. $(\tilde{u}_n)_n$ is uniformly bounded, and $(\tilde{u}_n, (\alpha_n)_+)_n$ is a minimizing sequence. In fact, we have that $\tilde{u}_n(0,\cdot) \geq u_n(0,\cdot)$ and $\tilde{u}_n(T,\cdot) \leq u_n(T,\cdot)$ because $\eta(a) \geq a$ for all $a < 0$, $\eta(0) = 0$ and $\eta(a) \leq a$ for all $a > 0$ (recall $u_n(0,x) \leq C_1$ and $u_n(T,x) \geq 0$ for all $x \in \mathbb{T}^d$). Moreover we can prove that $(\partial_t \tilde{u}_n)_n$ is bounded in $L^1([0,T]\times\mathbb{T}^d)$ and $(\nabla \tilde{u}_n)$ is bounded in $L^r([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$ uniformly w.r.t. m_0 and m_T. Indeed, by the growth condition (2.1) on H, for a.e. (t,x), we have

$$\partial_t \tilde{u}_n(t,x) + (\alpha_n)_+(t,x) + C \geq \frac{1}{Cr}|\nabla \tilde{u}_n|^r \geq 0.$$

Therefore, since $|\partial_t \tilde{u}_n| - |(\alpha_n)_+ + C| \leq |\partial_t \tilde{u}_n + (\alpha_n)_+ + C|$, we have

$$\begin{aligned}\int_0^T \int_{\mathbb{T}^d} |\partial_t \tilde{u}_n| \,\mathrm{d}x\,\mathrm{d}t &\leq \int_0^T \int_{\mathbb{T}^d} (\alpha_n)_+ \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_{\mathbb{T}^d} \left(\partial_t \tilde{u}_n(t,x) + (\alpha_n)_+ + C\right) \mathrm{d}x\,\mathrm{d}t \\ &\leq C + \int_{\mathbb{T}^d} \left(\tilde{u}_n(T,x) - \tilde{u}_n(0,x)\right) \mathrm{d}x \\ &\leq C,\end{aligned}$$

where we used the fact that $((\alpha_n)_+)_n$ is bounded in $L^{q'}([0,T]\times\mathbb{T}^d)$, hence in $L^1([0,T]\times\mathbb{T}^d)$, and that $(\tilde{u}_n)_n$ is uniformly bounded. Moreover,

$$\int_0^T \int_{\mathbb{T}^d} |\nabla \tilde{u}_n(t,x)|^r \,\mathrm{d}x\,\mathrm{d}t \leq Cr \int_0^T \int_{\mathbb{T}^d} \left(\partial_t \tilde{u}_n(t,x) + (\alpha_n)_+(t,x) + C\right) \mathrm{d}x\,\mathrm{d}t \leq C.$$

Summarizing all the estimates, we have

$$\|\tilde{u}_n\|_{L^\infty} + \|\nabla \tilde{u}_n\|_{L^r} + \|\partial_t \tilde{u}_n\|_{L^1} + \|(\alpha_n)_+\|_{L^{q'}} \leq C, \tag{2.41}$$

with $C > 0$ independent of m_0 and m_T. From this estimate we immediately deduce that, up to some subsequence, $(\nabla \tilde{u}_n)_n$ weakly converges in $L^r([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$, $(\partial_t \tilde{u}_n)_n$ weakly-* converges to a measure and $((\alpha_n)_+)_n$ weakly converges in $L^{q'}([0,T]\times\mathbb{T}^d)$.

Thanks to [Car15, Lemma 1] (see also [CS12, Theorem 1.3]), we have that $(\tilde{u}_n)_n$ is a sequence of locally uniformly Hölder continuous functions on $[0,T)\times\mathbb{T}^d$. Therefore, by the Arzelà-Ascoli theorem, we have that $(\tilde{u}_n)_n$ uniformly converges to $u \in \mathcal{C}^0([0,T)\times\mathbb{T}^d)$ on any compact set of $[0,T)\times\mathbb{T}^d$. From (2.41) we get that $u \in BV((0,T)\times\mathbb{T}^d)$ and $(\partial_t u, \nabla u)$ is the weak-* limit of $(\partial_t \tilde{u}_n, \nabla \tilde{u}_n)_n$.

Let $\alpha \in L^{q'}([0,T]\times\mathbb{T}^d)$ be a weak limit of $((\alpha_n)_+)_n$ in $L^{q'}([0,T]\times\mathbb{T}^d)$. Note that $\alpha \geq 0$ a.e. and, since $q' > 1$, α is also a weak-* limit of $((\alpha_n)_+)_n$ in $L^1([0,T]\times\mathbb{T}^d)$. As a consequence of the last assertion, the pair (u,α) satisfies (2.38) for some $C > 0$. Now, take φ a nonnegative test function in $\mathcal{C}_c^\infty([0,T]\times\mathbb{T}^d)$, then for all n, we have

$$\int_0^T \int_{\mathbb{T}^d} -\partial_t \tilde{u}_n(t,x)\varphi(t,x) \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_{\mathbb{T}^d} \varphi(t,x) H(x,\nabla \tilde{u}_n) \,\mathrm{d}x\,\mathrm{d}t \leq \int_0^T \int_{\mathbb{T}^d} \varphi(t,x)(\alpha_n)_+(t,x) \,\mathrm{d}x\,\mathrm{d}t.$$

The first integral on the left hand side converge by the weak* convergence of $(\partial_t \tilde{u}_n)_n$ and the integral on the right hand side converge due to the weak convergence of $((\alpha_n)_+)_n$ in $L^{q'}([0,T]\times\mathbb{T}^d)$, while thanks to the convexity of H in the gradient variable, we have

$$\int_0^T \int_{\mathbb{T}^d} \varphi(t,x) H(x,\nabla u) \,\mathrm{d}x\,\mathrm{d}t \leq \liminf_{n\to\infty} \int_0^T \int_{\mathbb{T}^d} \varphi(t,x) H(x,\nabla \tilde{u}_n) \,\mathrm{d}x\,\mathrm{d}t.$$

Therefore, (u,α) satisfies

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

in the sense of distributions and in particular, $(u,\alpha) \in \mathcal{K}$.

Let us now prove that (u,α) is a minimizer. Thanks to the convexity of F^*, we have the lower semicontinuity

$$\int_0^T\int_{\mathbb{T}^d} F^*(x,\alpha(t,x))\,\mathrm{d}x\,\mathrm{d}t \leq \liminf_{n\to\infty}\int_0^T\int_{\mathbb{T}^d} F^*(x,(\alpha_n)_+(t,x))\,\mathrm{d}x\,\mathrm{d}t. \tag{2.42}$$

The uniform convergence of $(\tilde{u}_n)_n$ on any compact set of $[0,T)\times\mathbb{T}^d$ implies that $(\tilde{u}_n(0,\cdot))_n$ converges uniformly to $u(0,\cdot)$, thus $\|u(0,\cdot)\|_{L^\infty}\leq C$ and

$$\int_{\mathbb{T}^d} u(0,x)\,\mathrm{d}m_0(x) = \lim_{n\to\infty}\int_{\mathbb{T}^d}\tilde{u}_n(0,x)\,\mathrm{d}m_0(x). \tag{2.43}$$

The pointwise convergence of $(\tilde{u}_n(T,\cdot))_n$ is not ensured. However, since $(\|\tilde{u}_n(T,\cdot)\|_\infty)_n$ is uniformly bounded by C, there exists $g\in L^\infty(\mathbb{T}^d)$ such that $\|g\|_\infty\leq C$ and, up to some subsequence, $u_n(T,\cdot)$ converges to g in the weak-* topology $\sigma(L^\infty,L^1)$. Therefore,

$$\int_{\mathbb{T}^d}\phi(x)g(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_{\mathbb{T}^d}\phi(x)u_n(T,x)\,\mathrm{d}x \quad \forall\,\phi\in L^1(\mathbb{T}^d). \tag{2.44}$$

Now, let $\phi\in\mathcal{C}^0(\mathbb{T}^d)$. We have that

$$\begin{aligned}\int_{\mathbb{T}^d}\phi(x)u_n(T,x)\,\mathrm{d}x &= \int_{\mathbb{T}^d}\int_0^T\phi(x)\partial_t u_n(t,x)\,\mathrm{d}t\,\mathrm{d}x + \int_{\mathbb{T}^d}\phi(x)u_n(0,x)\,\mathrm{d}x\\ &\to \int_{\mathbb{T}^d}\int_0^T\phi(x)\partial_t u(\mathrm{d}t,\mathrm{d}x) + \int_{\mathbb{T}^d}\phi(x)u(0,x)\,\mathrm{d}x.\end{aligned} \tag{2.45}$$

Using that the trace $u(T,\cdot)\in L^1(\mathbb{T}^d)$ of u at $\{T\}\times\mathbb{T}^d$ satisfies

$$\int_{\mathbb{T}^d}\int_0^T\phi(x)\partial_t u(\mathrm{d}t,\mathrm{d}x) = \int_{\mathbb{T}^d}\phi(x)u(T,x)\,\mathrm{d}x - \int_{\mathbb{T}^d}\phi(x)u(0,x)\,\mathrm{d}x,$$

relations (2.44) and (2.45) imply that $g=u(T,\cdot)$. Combining this result with (2.42) and (2.43), we deduce that (u,α) solves Problem (2.21). The result follows. □

Now, for $\varepsilon>0$ let us consider two probability densities m_0^ε and $m_T^\varepsilon\in L^1(\mathbb{T}^d)$ and denote by $(m_\varepsilon,w_\varepsilon)$ the unique solution to problem (2.14) with m_0 and m_T replaced by m_0^ε and $m_T^\varepsilon\in L^1(\mathbb{T}^d)$, respectively. Likewise, we denote by $(u_\varepsilon,\alpha_\varepsilon)\in\mathcal{K}$ a solution to the corresponding problem (2.21) such that (2.38) holds true.

The following stability result is a consequence of Γ-convergence and it follows easily from the statement and the proof of Proposition 2.8.

Corollary 2.9. *Suppose that, as $\varepsilon\to 0^+$, $(m_0^\varepsilon)_\varepsilon$ and $(m_T^\varepsilon)_\varepsilon$ converge in $L^1(\mathbb{T}^d)$ to m_0 and m_T, respectively. Then, the following assertions hold true:*

- (i) *$(m_\varepsilon,w_\varepsilon)$ converges weakly in $L^q((0,T)\times\mathbb{T}^d)\times L^{\frac{r'q}{r'+q-1}}((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$ to (m,w), the unique solution to (2.14).*
- (ii) *Up to some subsequence, $u_\varepsilon\to u$ uniformly on every compact subset of $[0,T)\times\mathbb{T}^d$, $u_\varepsilon(T,\cdot)\to u(T,\cdot)$ weakly-* in $L^\infty(\mathbb{T}^d)$, and $(\partial_t u_\varepsilon,\nabla u_\varepsilon,\alpha_\varepsilon)\to(\partial_t u,\nabla u,\alpha)$ weakly* in $\mathscr{M}((0,T)\times\mathbb{T}^d)\times L^r((0,T)\times\mathbb{T}^d;\mathbb{R}^d)\times L^{q'}((0,T)\times\mathbb{T}^d)$, where (u,α) is a solution to (2.21) satisfying (2.21).*

Proof. For $(u,\alpha)\in\mathcal{K}$ let us define

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

Define also $\mathcal{K}_1^\varepsilon$ as $\mathcal{K}_1$ with m_0 and m_T replaced by m_0^ε and m_T^ε, respectively.

Notice that Proposition 2.8 implies that for all $\varepsilon > 0$ we have

$$\inf_{(u,\alpha)\in\mathcal{K}} \mathcal{A}_\varepsilon(u,\alpha) = \inf\left\{\mathcal{A}_\varepsilon(u,\alpha) \mid (u,\alpha)\in\mathcal{K},\ \ \|u(0,\cdot)\|_{L^\infty}\leq C,\ \text{and}\ \|u(T,\cdot)\|_{L^\infty}\leq C\right\}. \tag{2.46}$$

Using that

$$|\mathcal{A}_\varepsilon(u,\alpha)-\mathcal{A}(u,\alpha)| \leq C\left(\|m_0^\varepsilon - m_0\|_{L^1} + \|m_T^\varepsilon - m_T\|_{L^1}\right),$$

for all $(u,\alpha)\in\mathcal{K}$ such that $\|u(0,\cdot)\|_{L^\infty}\leq C$ and $\|u(T,\cdot)\|_{L^\infty}\leq C$, relation (2.46) implies that

$$\lim_{\varepsilon\to 0^+} -\min_{(m,w)\in\mathcal{K}_1^\varepsilon}\mathcal{B}(m,w) = \lim_{\varepsilon\to0^+}\inf_{(u,\alpha)\in\mathcal{K}}\mathcal{A}_\varepsilon(u,\alpha) = \inf_{(u,\alpha)\in\mathcal{K}}\mathcal{A}(u,\alpha) = -\min_{(m,w)\in\mathcal{K}_1}\mathcal{B}(m,w).$$

Arguing as in the proof of Proposition 2.8, we have the existence of $(u,\alpha)\in\mathcal{K}$ such that, up to some subsequence, $(u_\varepsilon,\alpha_\varepsilon)_\varepsilon$ converges to (u,α) in the sense of (ii), and

$$\mathcal{A}(u,\alpha)\leq \lim_{\varepsilon\to0^+}\mathcal{A}_\varepsilon(u,\alpha) = \inf_{(u,\alpha)\in\mathcal{K}}\mathcal{A}(u,\alpha),$$

which implies (ii). In addition, Lemma 2.1 yields that $(m_\varepsilon,w_\varepsilon)$ is uniformly bounded in $L^q((0,T)\times\mathbb{T}^d)\times L^\ell((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$, where $\ell := \frac{r'q}{r'+q-1}$. Then, the lower semicontinuity of the convex functional $\mathcal{B}$ implies that any weak limit point (m,w) of $((m_\varepsilon,w_\varepsilon))_\varepsilon$ satisfies

$$\mathcal{B}(m,w)\leq \lim_{\varepsilon\to0}\mathcal{B}(m_\varepsilon,w_\varepsilon) = \min_{(m',w')\in\mathcal{K}_1}\mathcal{B}(m',w').$$

Since $(m_\varepsilon,w_\varepsilon)$ satisfies (2.13) with initial and final conditions given by m_0^ε and m_T^ε, respectively, we can pass to the limit in that equation to obtain that (m,w) also satisfies (2.13) with initial and final conditions given by m_0 and m_T, respectively. Finally, since $m_\varepsilon\geq 0$ a.e. we also get that $m\geq 0$ a.e., which implies that $(m,w)\in\mathcal{K}_1$. Therefore, (m,w) is the unique solution to (2.14) and the whole sequence $(m_\varepsilon,w_\varepsilon)_\varepsilon$ converges to (m,w) weakly in $L^q((0,T)\times\mathbb{T}^d)\times L^\ell((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$. The result follows. □

2.5. **Uniqueness.** In this subsection address uniqueness of solutions to the planning problem. Let $(\bar u,\bar m)$ be a weak solution to (1.1). In light of Theorem 2.6, the pair $(\bar m,\bar w) = (\bar m, -\bar m D_\xi H(\cdot,\nabla\bar u))$ is the minimizer of (2.14) while $(\bar u, f(\cdot,\bar m))$ is a solution of (2.21). In particular, $\bar m$ and $\bar w$ are unique because of the uniqueness of the solution of (2.14).

On the other hand, if H is strictly convex in the second variable, then uniqueness of $\bar w$ implies that $\nabla\bar u$ is unique on the set $\{\bar m > 0\}$ (Cf. the statement of Theorem 6.15 in [OPS]).

3. The planning problem as limit of penalized MFG

Consider the following system of equations, corresponding to a mean field game:

$$\begin{cases} -\partial_t u + H(x,\nabla u) = f(x,m), & \text{in } (0,T)\times\mathbb{T}^d,\\ \partial_t m - \nabla\cdot(D_\xi H(x,\nabla u)m) = 0, & \text{in } (0,T)\times\mathbb{T}^d,\\ m(0,\cdot)=m_0,\quad u(T,\cdot) = g_\varepsilon(\cdot,m(T,\cdot)), & \text{in } \mathbb{T}^d,\end{cases} \tag{3.1}$$

with the final coupling given by

$$g_\varepsilon(x,m(T)) = \frac{|m(T)-m_T(x)|^{p-2}(m(T)-m_T(x))}{\varepsilon}, \tag{3.2}$$

where $p>1$ is fixed and $\varepsilon>0$ is a fixed small parameter that we will take $\varepsilon\downarrow 0$ later. This corresponds to a final cost of

$$G_\varepsilon(x,m(T)) = \frac{|m(T)-m_T(x)|^p}{p\varepsilon}. \tag{3.3}$$

Its Fenchel conjugate is

$$G_\varepsilon^*(x,b) = bm_T(x) + \varepsilon^{\frac{p'}{p}} \frac{1}{p'}|b|^{p'}.$$

The goal of this section is twofold. First we prove existence and uniqueness of solutions for System (3.1). Then we prove that the solution $u^\varepsilon, m^\varepsilon$ converges in a suitable sense to a solution of the planning problem.

Our method of proving existence and uniqueness derives mainly from [Car15, CG15]. There are serious technical difficulties in establishing existence under the same weak hypotheses as in the previous section, namely m_0, m_T only in $L^1(\mathbb{T}^d)$. Although the convex minimization problem is "relaxed" in the sense that the constraint $m(T) = m_T$ is weakened, we actually impose an extra integrability criterion $m(T) \in L^p(\mathbb{T}^d)$ for some arbitrary $p > 1$.

Based on the results of this section and similarly to [GM18], we provide in Section 4 some Sobolev regularity of solutions of (3.1). Thus the present section has some intrinsic interest besides its application to the planning problem. Nevertheless, its primary purpose for this article will be to serve as a way to generate approximating solutions to System (1.1).

Throughout this section we assume that all the hypotheses from Section 2 are in force, with the following additional assumption:

(H5) $m_T \in L^p(\mathbb{T}^d)$, for a given $p > 1$.

3.1. **Two optimization problems in duality.** In this subsection too, we rely on the fact that our problem can be studied as an optimality condition between two problems in duality.

The first optimization problem is described as follows: let us denote $\mathcal{K}_0^\varepsilon = \mathcal{C}^1([0,T] \times \mathbb{T}^d) = \mathcal{K}_0$ and define on it the functional

$$\mathcal{A}_\varepsilon(u) = \int_0^T \int_{\mathbb{T}^d} F^*\left(x, -\partial_t u(t,x) + H(x, \nabla u(t,x))\right) \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} G_\varepsilon^*(x, u(T,x))\,\mathrm{d}x - \int_{\mathbb{T}^d} u(0,x) m_0(x)\,\mathrm{d}x. \tag{3.4}$$

Notice that

$$\mathcal{A}_\varepsilon(u) := \mathcal{A}(u) + \varepsilon^{\frac{p'}{p}} \frac{1}{p'} \int_{\mathbb{T}^d} |u(T,x)|^{p'}\,\mathrm{d}x.$$

The problem consists in optimizing

$$\inf_{u \in \mathcal{K}_0^\varepsilon} \mathcal{A}_\varepsilon(u)\ . \tag{3.5}$$

For the second optimization problem, let $\mathcal{K}_1^\varepsilon$ be the set of 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 $m(t,x) \geq 0$ a.e., with $\int_{\mathbb{T}^d} m(t,x)\,\mathrm{d}x = 1$ for a.e. $t \in (0,T)$, and which satisfy in the sense of distributions the continuity equation

$$\partial_t m + \operatorname{div}(w) = 0 \text{ in } (0,T) \times \mathbb{T}^d, \qquad m(0) = m_0. \tag{3.6}$$

On the set $\mathcal{K}_1^\varepsilon$, let us define the following functional

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

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

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

Since H^* and F are bounded from below and $m \geq 0$ a.e., the first integral in $\mathcal{B}(m,w)$ is well defined in $\mathbb{R}\cup\{+\infty\}$. Notice that

$$\mathcal{B}_\varepsilon(m,w) := \mathcal{B}(m,w) + \int_{\mathbb{T}^d} G_\varepsilon(x, m(T))\,\mathrm{d}x.$$

The second optimal control problem is the following:

$$\inf_{(m,w)\in\mathcal{K}_1^\varepsilon} \mathcal{B}_\varepsilon(m,w)\ . \tag{3.7}$$

Lemma 3.1. *We have*

$$\inf_{u\in\mathcal{K}_0^\varepsilon} \mathcal{A}_\varepsilon(u) = -\min_{(m,w)\in\mathcal{K}_1^\varepsilon} \mathcal{B}_\varepsilon(m,w).$$

Moreover, the minimum in the right-hand side is achieved by a unique pair $(m,w)\in\mathcal{K}_1^\varepsilon$ *satisfying* $(m,w)\in L^q((0,T)\times\mathbb{T}^d)\times L^{\frac{r'q}{r'+q-1}}((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$ *and* $m(T)\in L^p(\mathbb{T}^d)$.

The proof closely resembles that of Lemma 2.1, so we omit it.

3.2. **Analysis of the optimal control of the HJ equation.** Let $\mathcal{K}^\varepsilon$ be the set of triples $(u,\alpha,\beta)\in BV((0,T)\times\mathbb{T}^d)\times L^{q'}((0,T)\times\mathbb{T}^d)\times L^{p'}(\mathbb{T}^d)$ such that $u\in L^\infty((0,T)\times\mathbb{T}^d)$, $\nabla u\in L^r((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$, the traces $u(0,\cdot), u(T,\cdot)$ belong to $L^\infty(\mathbb{T}^d)$, and

$$-\partial_t u + H(x,\nabla u) \leq \alpha, \qquad u(T,x)\leq \beta(x) \qquad (t,x)\in(0,T)\times\mathbb{T}^d,$$

holds in the sense of distributions.

We extend the functional $\mathcal{A}_\varepsilon$ to $\mathcal{K}^\varepsilon$ by setting

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

The next proposition explains that the problem

$$\inf_{(u,\alpha,\beta)\in\mathcal{K}^\varepsilon} \mathcal{A}_\varepsilon(u,\alpha,\beta) \tag{3.8}$$

is the relaxed problem of (3.5).

Proposition 3.2. *We have*

$$\inf_{u\in\mathcal{K}_0^\varepsilon} \mathcal{A}_\varepsilon(u) = \inf_{(u,\alpha,\beta)\in\mathcal{K}^\varepsilon} \mathcal{A}_\varepsilon(u,\alpha,\beta).$$

The proof requires the following inequality:

Lemma 3.3. *Let* $(u,\alpha,\beta)\in\mathcal{K}^\varepsilon$ *and* $(m,w)\in\mathcal{K}_1^\varepsilon$. *Assume that* $mH^*(\cdot,-w/m)\in L^1((0,T)\times\mathbb{T}^d)$, $m\in L^q((0,T)\times\mathbb{T}^d)$, *and* $m(T)\in L^p(\mathbb{T}^d)$. *Then for a.e.* $t\in[0,T]$,

$$\int_{\mathbb{T}^d} (\beta m(T) - u(t)m(t))\,\mathrm{d}x + \int_t^T\int_{\mathbb{T}^d} m\left[\alpha + H^*\left(x, -\frac{w}{m}\right)\right]\mathrm{d}x\,\mathrm{d}t \ \geq\ 0 \tag{3.9}$$

and

$$\int_{\mathbb{T}^d} (u(t)m(t) - u(0)m_0)\,\mathrm{d}x + \int_0^t\int_{\mathbb{T}^d} m\left[\alpha + H^*\left(x, -\frac{w}{m}\right)\right]\mathrm{d}x\,\mathrm{d}t \ \geq\ 0.$$

Moreover, if equality holds in the inequality (3.9) *for* $t=0$, *then* $w = -mD_\xi H(x,\nabla u)$ *a.e.*

Proof. We may use the same argument as in the proof of Lemma 2.4. The only difference is that, when we introduced the regularization m_ε, we need to show $m_\varepsilon(T) \to m(T)$ in $L^p(\mathbb{T}^d)$. This is almost the same as showing the L^1 convergence as before:

(3.10)
$$\begin{aligned}\int_{\mathbb{T}^d} |m_\varepsilon(T,x) - m(T,x)|^p \,\mathrm{d}x &\le 2^{p-1}\int_{\mathbb{T}^d}\left|\int_{T-\delta}^{T+\delta}\int_{\mathbb{T}^d}\eta_\delta(T-s)\psi_\varepsilon(x-y)(m(s,y)-m_T(y))\,\mathrm{d}y\,\mathrm{d}s\right|^p \mathrm{d}x \\ &\quad + 2^{p-1}\int_{\mathbb{T}^d}\left|\int_{T-\delta}^{T+\delta}\int_{\mathbb{T}^d}\eta_\delta(T-s)\psi_\varepsilon(x-y)(m_T(y)-m(T,x))\,\mathrm{d}y\,\mathrm{d}s\right|^p \mathrm{d}x \\ = 2^{p-1}\int_{\mathbb{T}^d}\left|\int_{T-\delta}^{T+\delta}\int_{\mathbb{T}^d}\int_s^T \eta_\delta(T-s)\nabla\psi_\varepsilon(x-y)\cdot w(\tau,y)\,\mathrm{d}y\,\mathrm{d}\tau\,\mathrm{d}s\right|^p &\mathrm{d}x + 2^{p-1}\int_{\mathbb{T}^d}|m_{T,\varepsilon}(x)-m(T,x)|^p\,\mathrm{d}x \\ &\le \frac{C}{\varepsilon^{d+1}}\left(\int_{T-\delta}^T\int_{\mathbb{T}^d}|w(\tau,y)|\,\mathrm{d}\tau\,\mathrm{d}y\right)^p + 2^{p-1}\int_{\mathbb{T}^d}|m_{T,\varepsilon}(x)-m(T,x)|^p\,\mathrm{d}x.\end{aligned}$$

Letting $\delta(\varepsilon)$ be small enough, we see that the right-hand side goes to zero as $\varepsilon \to 0$. In a similar manner, $m_\varepsilon(0) \to m(0)$ in $L^1(\mathbb{T}^d)$, which is sufficient because $u(0) \in L^\infty(\mathbb{T}^d)$. The other estimates are exactly as in the proof of Lemma 2.4, so we omit them. □

Proof of Proposition 3.2. The argument is almost exactly the same as the proof of Proposition 2.3, so we omit it. □

3.3. **Existence of a solution for the relaxed problem.**

The next proposition explains the interest of considering the relaxed problem (3.8) instead of the original one (3.5).

Proposition 3.4. *The relaxed problem (3.8) has at least one solution (u, α, β). The function u is Hölder continuous in $[0,T) \times \mathbb{T}^d$, $\alpha \ge 0$ a.e., and there exists $C > 0$, independent of m_0, m_T and ε, and $C_\varepsilon > 0$, independent of m_0 and m_T, such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{T}^d}|u(t,x)| + \|\nabla u\|_{L^r} + \|\partial_t u\|_{\mathscr{M}} \le C_\varepsilon \quad \text{and} \quad \|\alpha\|_{L^{q'}} + \varepsilon^{\frac{p'}{p}}\|\beta\|_{L^{p'}} \le C, \tag{3.11}$$

where $\|\cdot\|_{\mathscr{M}}$ denotes the usual norm of $\mathscr{M}$ as dual space of $\mathcal{C}^0([0,T]\times\mathbb{T}^d)$.

Proof. Consider a minimizing sequence $(u_n)_n$ for Problem (3.5). Observe that if

$$c_n := \min_{x\in\mathbb{T}^d} u_n(T,x) > 0,$$

then

$$\mathcal{A}_\varepsilon(u_n - c_n) = \mathcal{A}(u_n - c_n) + \varepsilon^{\frac{p'}{p}}\frac{1}{p'}\int_{\mathbb{T}^d}(u_n(T,x)-c_n)^{p'}\,\mathrm{d}x \le \mathcal{A}(u_n) + \varepsilon^{\frac{p'}{p}}\frac{1}{p'}\int_{\mathbb{T}^d}u_n(T,x)^{p'}\,\mathrm{d}x = \mathcal{A}_\varepsilon(u_n).$$

where we have used the fact that $\mathcal{A}(u_n - c_n) = \mathcal{A}(u_n)$ and $u_n(T,x) - c_n \le u_n(T,x)$ being $c_n > 0$. Therefore, by replacing u_n with $u_n - c_n$ whenever $c_n > 0$, without loss of generality we may assume $c_n = \min_{x\in\mathbb{T}^d} u_n(T,x) \le 0$.

We set

$$\alpha_n(t,x) = -\partial_t u_n(t,x) + H(x, \nabla u_n(t,x)), \quad \beta_n(x) = u_n(T,x). \tag{3.12}$$

Proposition 3.2 implies that $(u_n, \alpha_n, \beta_n)_n$ is a minimizing sequence for Problem (3.8). Moreover, inequality (2.37) applies to all u_n giving

$$u_n(0,x) \le c_n + C\left[T^{1-r'} + \left(T^\nu \wedge 1 + T^{1/q}\right)\left(\|(\alpha_n)_+\|_{L^{q'}} + 1\right)\right] \quad \forall\, x \in \mathbb{T}^d. \tag{3.13}$$

Now since $\mathcal{A}_\varepsilon(0)=0$, we can assume without loss of generality that

(3.14)
$$
\begin{aligned}
0 &\geq \int_0^T \int_{\mathbb{T}^d} F^*(x,\alpha_n)\,\mathrm{d}x\,\mathrm{d}t + \frac{\varepsilon^{\frac{p'}{p}}}{p'}\int_{\mathbb{T}^d}|\beta_n(x)|^{p'}\,\mathrm{d}x + \int_{\mathbb{T}^d}\beta_n(x)\,\mathrm{d}m_T(x) - \int_{\mathbb{T}^d} u_n(0,x)\,\mathrm{d}m_0(x)\\
&\geq \frac{1}{q'C}\|(\alpha_n)_+\|_{L^{q'}}^{q'} + \frac{\varepsilon^{\frac{p'}{p}}}{p'}\|\beta_n\|_{L^{p'}}^{p'} - C\left(T^{1-r'} + \left(T^\nu \wedge 1 + T^{1/q}\right)\left(\|(\alpha_n)_+\|_{L^{q'}}+1\right)\right),
\end{aligned}
$$

where we used the growth condition (2.7) on F^*, inequality (3.13), $u_n(T,\cdot)=\beta_n(x)\geq c_n$ and $\int_{\mathbb{T}^d} m_0(x)\,\mathrm{d}x = \int_{\mathbb{T}^d} m_T(x)\,\mathrm{d}x = 1$. We deduce that $((\alpha_n)_+)_n$ is a bounded sequence in $L^{q'}([0,T]\times\mathbb{T}^d)$, and $(\beta_n)_n$ is a bounded sequence in $L^{p'}(\mathbb{T}^d)$ uniformly with respect to m_0 and ε. Moreover, there exists a constant $C_1>0$, independent of m_0 and ε, such that $u_n(0,x)\leq C_1$ for all $x\in\mathbb{T}^d$.

By Young's inequality and the lower bound on F^*, we deduce from (3.14) that

$$
-\int_{\mathbb{T}^d} u_n(0,x)m_0(x)\,\mathrm{d}x \leq C + C_\varepsilon\|m_T\|_{L^p}^p. \tag{3.15}
$$

Integrating (3.13) against $m_0(x)$ over $\mathbb{T}^d$ we obtain

$$
c_n \geq -C_\varepsilon\|m_T\|_{L^p}^p - C\left[1 + T^{1-r'} + \left(T^\nu\wedge 1 + T^{1/q}\right)\left(\|(\alpha_n)_+\|_{L^{q'}}+1\right)\right] \geq -\tilde{C}_\varepsilon. \tag{3.16}
$$

By increasing $\tilde{C}_\varepsilon$ if necessary, we will assume it is larger than C_1. In analogy with the proof of Proposition 2.8, let $\eta_\varepsilon\in C^1(\mathbb{R})$ such that $0\leq\eta_\varepsilon'\leq 1$, $|\eta_\varepsilon|\leq 2\tilde{C}_\varepsilon$ and $\eta_\varepsilon(s)=s$ if $-\tilde{C}_\varepsilon\leq s\leq C_1$, and set $\tilde{u}_n := \eta_\varepsilon\circ u_n$, then $\tilde{u}_n$ is Lipschitz continuous. We again deduce that

$$
-\partial_t\tilde{u}_n + H(x,\nabla\tilde{u}_n) \leq (\alpha_n)_+. \tag{3.17}
$$

On the other hand, we have that $\tilde{u}_n(0,\cdot)\geq u_n(0,\cdot)$ and $\tilde{u}_n(T,\cdot)\leq u_n(T,\cdot)$ because $\eta_\varepsilon(a)\geq a$ for all $a<0$, $\eta_\varepsilon(0)=0$ and $\eta_\varepsilon(a)\leq a$ for all $a>0$ (recall $u_n(0,x)\leq C_1$ and $u_n(T,x)\geq c_n\geq -\tilde{C}_\varepsilon$ for all $x\in\mathbb{T}^d$), in particular $\tilde{u}_n(T,\cdot)\leq\beta_n$. It follows that $\mathcal{A}_\varepsilon(\tilde{u}_n,(\alpha_n)_+,\beta_n)\leq\mathcal{A}_\varepsilon(u_n,(\alpha_n)_+,\beta_n)$ and $(\tilde{u}_n,(\alpha_n)_+,\beta_n)\in\mathcal{K}_\varepsilon$ is a new minimizing sequence. We note that $(\tilde{u}_n)_n$ is uniformly bounded in L^∞ by $2\tilde{C}_\varepsilon$.

Proceeding as in the proof of Proposition 2.8, we obtain

$$
\|\tilde{u}_n\|_{L^\infty} + \|\nabla\tilde{u}_n\|_{L^r} + \|\partial_t\tilde{u}_n\|_{L^1} \leq C_\varepsilon \tag{3.18}
$$

for some constant C_ε depending on ε. It remains to show that $(\tilde{u}_n,(\alpha_n)_+,\beta_n)$ has an accumulation point $(u,\alpha,\beta)\in\mathcal{K}_\varepsilon$ that minimizes $\mathcal{A}_\varepsilon$. The argument is exactly as in Proposition 2.8, and we omit the details. □

Corollary 3.5. *Let (u,α,β) be the minimizer obtained in Proposition 3.4. Then there exists $(v,\alpha)\in\mathcal{K}$ such that $\mathcal{A}(v,\alpha)\leq\mathcal{A}(u,\alpha)$ and the following estimates hold uniformly with respect to ε (and m_0 and m_T):*

$$
\|v\|_{L^\infty} + \|\nabla v\|_{L^r} + \|\partial_t v\|_{L^1} \leq C. \tag{3.19}
$$

Furthermore, v is locally Hölder continuous on $[0,T)\times\mathbb{T}^d$, and there exists a constant $c\leq 0$ possibly depending on ε such that $c\leq\inf_x u(T,x)$ and such that v satisfies

$$
v = \eta(u-c)
$$

for some smooth function η as in the proof of Proposition 2.8, satisfying $0\leq\eta'\leq 1$ and $|\eta|\leq 2C_1$.

We stress that (v,α,β) is not necessarily a minimizer of $\mathcal{A}_\varepsilon$, and it need not even be an element of $\mathcal{K}^\varepsilon$. We can think of v as a "regularized modification" of u.

Proof. Let $(u_n)_n$ be the same minimizing sequence as in the proof of Proposition 3.4 and again set

$$\alpha_n(t,x) = -\partial_t u_n(t,x) + H(x, \nabla u_n(t,x)), \quad c_n = \min_{x\in\mathbb{T}^d} u_n(T,x) \leq 0. \tag{3.20}$$

Define $\tilde{u}_n$ as in the previous proof so that (3.17), (3.21) hold as well as $\|(\alpha_n)_+\|_{L^{q'}} \leq C$ for some C independent of ε. Note also that $c_n = \min_{x\in\mathbb{T}^d} \tilde{u}_n(T,x)$. Let $v_n = \tilde{u}_n - c_n$. Then $\mathcal{A}(\tilde{u}_n, (\alpha_n)_+) = \mathcal{A}(v_n, (\alpha_n)_+)$ (we do not claim anything about $\mathcal{A}_\varepsilon(v_n)$!). As in the proof of Proposition 2.8, let $\eta \in C^1(\mathbb{R})$ such that $0 \leq \eta' \leq 1$, $|\eta| \leq 2C_1$ and $\eta(s) = s$ if $|s| \leq C_1$ (where now $C_1 > 0$ is such that $v_n(0,\cdot) \leq C_1$), and set $\tilde{v}_n := \eta \circ v_n$. By the same argument as in that proof, we obtain

$$-\partial_t \tilde{v}_n + H(x, \nabla \tilde{v}_n) \leq (\alpha_n)_+ \text{ a.e.},$$

$\mathcal{A}(\tilde{v}_n, (\alpha_n)_+) \leq \mathcal{A}(v_n, (\alpha_n)_+) = \mathcal{A}(\tilde{u}_n, (\alpha_n)_+)$, (since, again $\tilde{v}_n(T,\cdot) \leq v_n(T,\cdot)$, $\tilde{v}_n(0,\cdot) \geq v_n(0,\cdot)$) and

$$\|\tilde{v}_n\|_{L^\infty} + \|\nabla \tilde{v}_n\|_{L^r} + \|\partial_t \tilde{v}_n\|_{L^1} + \|(\alpha_n)_+\|_{L^{q'}} \leq C, \tag{3.21}$$

where C is uniform in n, m_0, m_T, and ε. Here we have used the uniform estimate on $\|(\alpha_n)_+\|_{L^{q'}}$ obtained in the proof of Proposition 3.4.

From this estimate we deduce that, up to some subsequence, $(\nabla \tilde{v}_n)_n$ weakly converges in $L^r([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$, $(\partial_t \tilde{v}_n)_n$ weakly-* converges to a measure, and $((\alpha_n)_+)_n$ converges weakly to α in $L^{q'}([0,T]\times\mathbb{T}^d)$ (the same as in the proof of Proposition 3.4). Again, thanks to [Car15, Lemma 1], we have that $(\tilde{v}_n)_n$ is a sequence of locally uniformly Hölder continuous functions on $[0,T)\times\mathbb{T}^d$, and so $(\tilde{v}_n)_n$ uniformly converges to $v \in \mathcal{C}^0([0,T)\times\mathbb{T}^d)$ on any compact subset of $[0,T)\times\mathbb{T}^d$. By (3.21) we have $v \in BV((0,T)\times\mathbb{T}^d)$ and $(\partial_t v, \nabla v)$ is the weak-* limit of $(\partial_t \tilde{v}_n, \nabla \tilde{v}_n)_n$. As before we deduce that (v,α) satisfies

$$-\partial_t v + H(x, \nabla v) \leq \alpha$$

in the sense of distributions. As in the proof of Proposition 2.8, we also have $\tilde{v}_n(T,\cdot) \rightharpoonup v(T,\cdot)$, as $n \to +\infty$ in the weak-* topology $\sigma(L^\infty, L^1)$. In particular, $(v,\alpha) \in \mathcal{K}$.

Finally, by lower semicontinuity arguments as before, we get $\mathcal{A}(v,\alpha) \leq \mathcal{A}(u,\alpha)$. To see this, note that

$$\int_{\mathbb{T}^d} v(0,x) m_0(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{\mathbb{T}^d} \tilde{v}_n(0,x) m_0(x)\,\mathrm{d}x$$

and

$$\int_{\mathbb{T}^d} v(T,x) m_T(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{\mathbb{T}^d} \tilde{v}_n(T,x) m_T(x)\,\mathrm{d}x$$

by the same proof as in Proposition 2.8, and so

$$\begin{aligned}
\mathcal{A}(v,\alpha) &= \int_0^T \int_{\mathbb{T}^d} F^*(x,\alpha)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} v(T,x) m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} v(0,x) m_0(x)\,\mathrm{d}x \\
&= \liminf_{n\to\infty} \int_0^T \int_{\mathbb{T}^d} F^*(x,(\alpha_n)_+)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} \tilde{v}_n(T,x) m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} \tilde{v}_n(0,x) m_0(x)\,\mathrm{d}x \\
&\leq \liminf_{n\to\infty} \int_0^T \int_{\mathbb{T}^d} F^*(x,(\alpha_n)_+)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} v_n(T,x) m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} v_n(0,x) m_0(x)\,\mathrm{d}x \\
&= \liminf_{n\to\infty} \int_0^T \int_{\mathbb{T}^d} F^*(x,(\alpha_n)_+)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} \tilde{u}_n(T,x) m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} \tilde{u}_n(0,x) m_0(x)\,\mathrm{d}x \\
&= \int_0^T \int_{\mathbb{T}^d} F^*(x,\alpha)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{T}^d} u(T,x) m_T(x)\,\mathrm{d}x - \int_{\mathbb{T}^d} u(0,x) m_0(x)\,\mathrm{d}x = \mathcal{A}(u,\alpha).
\end{aligned}$$

Also from the uniform estimates on c_n derived in Proposition 3.4, we can assume (again, up to a subsequence) that $c_n \to c \in [-\tilde{C}_\varepsilon, 0]$. We can see that $c \le \inf_x u(T,x)$, since

$$\int_{\mathbb{T}^d} (u(T,x) - c)\phi(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{\mathbb{T}^d} (u_n(T,x) - c_n)\phi(x)\,\mathrm{d}x \ge 0$$

for all integrable functions $\phi \ge 0$, so $u(T,x) \ge c$ a.e. It follows that

$$v(t,x) = \lim_{n\to\infty} \eta(v_n(t,x)) = \lim_{n\to\infty} \eta(\tilde{u}_n(t,x) - c_n) = \eta(\max\{u(t,x) - c; 0\}) \quad \forall t \in [0,T), \forall x \in \mathbb{T}^d,$$

using the local uniform convergence $\tilde{v}_n \to v$ and $\tilde{u}_n \to u$ to in $[0,T) \times \mathbb{T}^d$.

□

3.4. **Definition of weak solutions.** The variational method described above provides weak solutions for the MFG system. By a weak solution, we mean the following:

Definition 3.6. *Let $(u,m) \in BV((0,T)\times\mathbb{T}^d) \times L^q((0,T)\times\mathbb{T}^d)$ with $u \in L^\infty((0,T)\times\mathbb{T}^d)$ and $u(0,\cdot), u(T,\cdot) \in L^\infty(\mathbb{T}^d)$. We say that (u,m) is a weak solution to (3.1) if*

(i) *the following integrability conditions hold:*

$$\nabla u \in L^r((0,T)\times\mathbb{T}^d;\mathbb{R}^d),\ mH^*(\cdot, D_\xi H(\cdot,\nabla u)) \in L^1((0,T)\times\mathbb{T}^d),\ m(T) \in L^p(\mathbb{T}^d),$$

and

$$mD_\xi H(\cdot,\nabla u)) \in L^1((0,T)\times\mathbb{T}^d;\mathbb{R}^d).$$

(ii) *Equation (3.1)-(i) holds in the following sense: inequality*

$$-\partial_t u + H(x,\nabla u) \le f(x,m) \quad \text{in } (0,T)\times\mathbb{T}^d, \tag{3.22}$$

with $u(T,\cdot) \le g_\varepsilon(\cdot, m(T,\cdot))$, holds in the sense of distributions,

(iii) *Equation (3.1)-(ii) holds:*

$$\partial_t m - \mathrm{div}(mD_\xi H(x,\nabla u)) = 0 \text{ in } (0,T)\times\mathbb{T}^d, \quad m(0) = m_0 \tag{3.23}$$

in the sense of distributions,

(iv) *The following equality holds:*

$$\begin{aligned}\int_0^T \int_{\mathbb{T}^d} m(t,x)\,[f(x,m(t,x)) + H^*(x, D_\xi H(x,\nabla u)(t,x))]\,\mathrm{d}x\,\mathrm{d}t \\ + \int_{\mathbb{T}^d} [m(T,x)g_\varepsilon(x,m(T,x)) - m_0(x)u(0,x)]\,\mathrm{d}x = 0.\end{aligned} \tag{3.24}$$

Our main result for this section is the following existence and uniqueness theorem:

Theorem 3.7. *There exists a weak solution (u,m) to the MFG system (3.1). Moreover this solution is unique in the following sense: if (u,m) and (u',m') are two solutions, then $m = m'$ a.e. and $u = u'$ in $\{m > 0\}$.*

Remark 3.8. *Let us underline the fact that unlike the planning problem, we can expect m-a.e. uniqueness of the adjoint state u.*

3.5. **Existence of a weak solution.** The first step towards the proof of Theorem 3.7 consists in showing a one-to-one equivalence between solutions of the MFG system and the two optimization problems (3.7) and (3.8).

Theorem 3.9. *Let $(\bar{m},\bar{w}) \in \mathcal{K}_1^\varepsilon$ be a minimizer of (3.7) and $(\bar{u},\bar{\alpha},\bar{\beta}) \in \mathcal{K}^\varepsilon$ be a minimizer of (3.8). Then $(\bar{u},\bar{m})$ is a weak solution of the mean field games system (3.1) and $\bar{w} = -\bar{m}D_\xi H(\cdot,\nabla\bar{u})$, while $\bar{\alpha} = f(\cdot,\bar{m})$ and $\bar{\beta} = g_\varepsilon(\cdot,\bar{m}(T,\cdot))$ a.e.*

Conversely, any weak solution $(\bar{u},\bar{m})$ of (3.1) is such that the pair $(\bar{m}, -\bar{m}D_\xi H(\cdot,\nabla\bar{u}))$ is the minimizer of (3.7) while $(\bar{u}, f(\cdot,\bar{m}), g_\varepsilon(\cdot,\bar{m}(T,\cdot))$ is a minimizer of (3.8).

Proof. The argument is very similar to the proof of Theorem 2.6, so we omit it. □

3.6. **Uniqueness.** In this subsection we prove the uniqueness part of Theorem 3.7. Let $(\bar{u}, \bar{m})$ be a weak solution to (3.1). In light of Theorem 3.7, the pair $(\bar{m}, -\bar{m} D_\xi H(\cdot, \nabla \bar{u}))$ is the minimizer of (3.7) while $(\bar{u}, f(\cdot, \bar{m}))$ is a solution of (3.8). In particular, $\bar{m}$ is unique because of the uniqueness of the solution of (3.7).

Thus we assume that $(u_1, \bar{m}), (u_2, \bar{m})$ are two weak solutions of (3.1). Set $\bar{\alpha} = f(\cdot, \bar{m})$ and $\bar{\beta} = g_\varepsilon(\cdot, \bar{m}(T, \cdot))$. To show that $u_1 = u_2$ on the set $\{\bar{m} > 0\}$, it is sufficient to show that $\bar{u} := \max\{u_1, u_2\}$ satisfies

$$-\partial_t \bar{u} + H(x, \nabla \bar{u}) \leq \bar{\alpha} \tag{3.25}$$

and

$$\bar{u}(T, \cdot) \leq \bar{\beta} \tag{3.26}$$

in the sense of distributions. This is because $(\bar{u}, \bar{\alpha}, \bar{\beta})$ can be seen to minimize the cost

$$\int_t^T \int_{\mathbb{T}^d} F^*(x, \alpha) \,\mathrm{d}x \,\mathrm{d}t + \int_{\mathbb{T}^d} G_\varepsilon^*(x, \beta) - \bar{m}(t) u(t) \,\mathrm{d}x$$

for any t, and so $\int_{\mathbb{T}^d} \bar{m}(t) \bar{u}(t) \,\mathrm{d}x = \int_{\mathbb{T}^d} \bar{m}(t) u_1(t) \,\mathrm{d}x$, which implies $u_1 = \bar{u}$ a.e. in $\{\bar{m} > 0\}$ (cf. [CGPT15, Section 5.3]). The same argument applies with u_2 in place of u_1.

Our goal is to show that $\bar{u}$ satisfies (3.25) and (3.26), in the sense of distributions. Let us notice immediately that (3.26) is automatically satisfied. Let $\varepsilon > 0$. Introduce the following translation and extension of $(u_k, \bar{\alpha}), k = 1, 2$:

$$\tilde{u}_k(t, x) = \begin{cases} u_k(t + 2\varepsilon, x) & \text{if } t \in [-2\varepsilon, T - 2\varepsilon), \\ 0 & \text{if } t \in [T - 2\varepsilon, T + 2\varepsilon] \end{cases} \tag{3.27}$$

and

$$\tilde{\alpha}(t, x) = \begin{cases} \bar{\alpha}(t + 2\varepsilon, x) & \text{if } t \in [-2\varepsilon, T - 2\varepsilon), \\ H(x, 0) & \text{if } t \in [T - 2\varepsilon, T + 2\varepsilon]. \end{cases} \tag{3.28}$$

Then we have that

$$-\partial_t \tilde{u}_k + H(x, \nabla \tilde{u}_k) \leq \tilde{\alpha} + \bar{\beta}(x) \partial_t \mathbf{1}_{[T-2\varepsilon, T+2\varepsilon]}(t) \tag{3.29}$$

in the sense of distributions on $(-2\varepsilon, T + 2\varepsilon) \times \mathbb{T}^d$.

Fix a smooth vector field ψ on $[0, T] \times \mathbb{T}^d$. Notice that

$$-\partial_t \tilde{u}_k + \psi \cdot \nabla \tilde{u}_k \leq \tilde{\alpha} + H^*(x, \psi) + \bar{\beta}(x) \partial_t \mathbf{1}_{[T-2\varepsilon, T+2\varepsilon]}(t) \tag{3.30}$$

in the sense of distributions on $(-2\varepsilon, T + 2\varepsilon) \times \mathbb{T}^d$. Let $\eta^\varepsilon(t, x) = \eta_x^\varepsilon(x) \eta_t^\varepsilon(t)$ be the product of standard convolution kernels, one in the x variable and the other in the t variable. Then set $\alpha^\varepsilon = \eta^\varepsilon * \tilde{\alpha}$, $u_k^\varepsilon = \eta^\varepsilon * \tilde{u}_k$, $u^\varepsilon = \max\{u_1^\varepsilon, u_2^\varepsilon\}$, and $\beta^\varepsilon = \eta_x^\varepsilon * \bar{\beta}$. Then for each $k = 1, 2$,

$$-\partial_t u_k^\varepsilon + \psi \cdot \nabla u_k^\varepsilon \leq \alpha^\varepsilon + \eta^\varepsilon * H^*(\cdot, \psi) + \beta^\varepsilon(x) \partial_t (\eta_t^\varepsilon * \mathbf{1}_{[T-2\varepsilon, T+2\varepsilon]})(t) + R^\varepsilon \tag{3.31}$$

in a pointwise sense on $[0, T] \times \mathbb{T}^d$, where

$$R^\varepsilon := [\psi, \eta^\varepsilon](\nabla \tilde{u}_k), k = 1, 2,$$

where we have used the commutator notation from DiPerna-Lions [DL89] (cf. LeBris-Lions [LBL08] where this is applied to the Fokker-Planck equation). By [DL89, Lemma II.1], we have $R^\varepsilon \to 0$ in $L^r((0, T) \times \mathbb{T}^d)$. Now, since the maximum of two subsolutions is itself a subsolution, we have

$$-\partial_t u^\varepsilon + \psi \cdot \nabla u^\varepsilon \leq \alpha^\varepsilon + \eta^\varepsilon * H^*(\cdot, \psi) + \beta^\varepsilon(x) \partial_t (\eta_t^\varepsilon * \mathbf{1}_{[T-2\varepsilon, T+2\varepsilon]})(t) + R^\varepsilon \tag{3.32}$$

in a viscosity sense on $[0,T]\times\mathbb{T}^d$, hence also in the sense of distributions by the result of Ishii ([Ish95], see also [CGPT15, Section 6.3]). Here we also use the fact that $u_1^\varepsilon(T)=u_2^\varepsilon(T)=0$. In other words,

$$
\begin{aligned}
(3.33)\quad &\int_0^T\int_{\mathbb{T}^d} u^\varepsilon\partial_t\zeta+\zeta\psi\cdot\nabla u^\varepsilon\,\mathrm{d}x\,\mathrm{d}t\\
&\le \int_0^T\int_{\mathbb{T}^d}\zeta(\alpha^\varepsilon+\eta^\varepsilon*H^*(\cdot,\psi))\,\mathrm{d}x\,\mathrm{d}t+\int_{\mathbb{T}^d}\zeta(T)\beta^\varepsilon\,\mathrm{d}x-\int_0^T\int_{\mathbb{T}^d}\partial_t\zeta\beta^\varepsilon\eta_t^\varepsilon*\mathbf{1}_{[T-2\varepsilon,T+2\varepsilon]}\,\mathrm{d}x\,\mathrm{d}t+\int_0^T\int_{\mathbb{T}^d}R^\varepsilon\zeta\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
$$

for any non-negative test function $\zeta\in C_c^1((0,T]\times\mathbb{T}^d)$. As $\varepsilon\to0$, a routine check shows that, at least up to a subsequence, $u^\varepsilon\to\bar u$ pointwise and in $L^\gamma([0,T]\times\mathbb{T}^d)$ for any $\gamma\ge1$, $\nabla u^\varepsilon\rightharpoonup\nabla\bar u$ weakly in $L^r([0,T]\times\mathbb{T}^d)$, $\eta^\varepsilon*H^*(\cdot,\psi)\to H^*(\cdot,\psi)$ uniformly, $\beta^\varepsilon\to\beta$ in $L^p(\mathbb{T}^d)$ and $\alpha^\varepsilon\to\alpha$ in $L^{q'}([0,T]\times\mathbb{T}^d)$. Finally, to see that the penultimate integral vanishes, we observe that

$$
\begin{aligned}
(3.34)\quad &\left|\int_0^T\int_{\mathbb{T}^d}\partial_t\zeta\beta^\varepsilon\eta_t^\varepsilon*\mathbf{1}_{[T-2\varepsilon,T+2\varepsilon]}\,\mathrm{d}x\,\mathrm{d}t\right|=\left|\int_0^T\int_{\mathbb{T}^d}\int_{-\varepsilon}^{\varepsilon}\partial_t\zeta\beta^\varepsilon\eta_t^\varepsilon(s)\mathbf{1}_{[T-2\varepsilon,T+2\varepsilon]}(t-s)\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\right|\\
&\le\int_0^T\int_{\mathbb{T}^d}\int_{-\varepsilon}^{\varepsilon}|\partial_t\zeta||\beta^\varepsilon|\eta_t^\varepsilon(s)\mathbf{1}_{[T-3\varepsilon,T+3\varepsilon]}(t)\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t=\int_{T-3\varepsilon}^T\int_{\mathbb{T}^d}|\partial_t\zeta||\beta^\varepsilon|\,\mathrm{d}x\,\mathrm{d}t\to0.
\end{aligned}
$$

Thus, letting $\varepsilon\to0$ in (3.33) we get

$$
(3.35)\qquad -\int_{\mathbb{T}^d}\zeta(T)\beta\,\mathrm{d}x+\int_0^T\int_{\mathbb{T}^d}\bar u\partial_t\zeta+\zeta\psi\cdot\nabla\bar u\,\mathrm{d}x\,\mathrm{d}t\le\int_0^T\int_{\mathbb{T}^d}\zeta(\alpha+H^*(x,\psi))\,\mathrm{d}x\,\mathrm{d}t.
$$

Finally, taking a sequence of such ψ converging to $D_\xi H(x,\nabla\bar u)$ in $L^{r'}((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$, we get the desired inequality. The proof is complete.

3.7. **The limit as $\varepsilon\to0$.**

Lemma 3.10 (Uniform bounds in $\varepsilon>0$)**.** *Let $\varepsilon>0$ and $(u_\varepsilon,m_\varepsilon)$ be the solution of System (3.1) with $\alpha_\varepsilon=f(\cdot,m_\varepsilon),\beta_\varepsilon=g_\varepsilon(\cdot,m_\varepsilon(T))$ as coupling, where $(u_\varepsilon,\alpha_\varepsilon,\beta_\varepsilon)$ and $(m_\varepsilon,w_\varepsilon)$ are minimizers of the variational problems (3.8) and (3.7), respectively. Then*

- (i) $\left(\int_{\mathbb{T}^d}G_\varepsilon(x,m_\varepsilon(T))\,\mathrm{d}x\right)_{\varepsilon>0}$ *is uniformly bounded and* $(\varepsilon^{1/p}\beta_\varepsilon)_{\varepsilon>0}$ *is uniformly bounded in* $L^{p'}(\mathbb{T}^d)$*, and in particular* $m_\varepsilon(T)\to m_T$ *strongly in* $L^p(\mathbb{T}^d)$*;*
- (ii) $(m_\varepsilon,w_\varepsilon)_{\varepsilon>0}$ *is uniformly bounded in* $L^q([0,T]\times\mathbb{T}^d)\times L^{\frac{r'q}{r'+q-1}}([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$*;*
- (iii) $(\alpha_\varepsilon)_{\varepsilon>0}=(f(\cdot,m_\varepsilon))_{\varepsilon>0}$ *is uniformly bounded below by 0 and in* $L^{q'}([0,T]\times\mathbb{T}^d)$*.*

Proof. Using the assumptions on H^* and F, one sees that

$$
(3.36)\qquad 0\le\int_{\mathbb{T}^d}G_\varepsilon(x,m_\varepsilon(T))\,\mathrm{d}x\le C+\mathcal{B}_\varepsilon(m_\varepsilon,w_\varepsilon)=C+\inf_{\mathcal{K}_1^\varepsilon}\mathcal{B}_\varepsilon.
$$

Taking any fixed pair $(\tilde m,\tilde w)\in\mathcal{K}_1$ for the functional $\mathcal{B}$, we can use this pair as a competitor in $\mathcal{B}_\varepsilon$ to reveal

$$
(3.37)\qquad \inf_{\mathcal{K}_1^\varepsilon}\mathcal{B}_\varepsilon\le\int_0^T\int_{\mathbb{T}^d}\left[\tilde mH^*\left(x,-\frac{\tilde w}{\tilde m}\right)+F(x,\tilde m)\right]\mathrm{d}x\,\mathrm{d}t,
$$

where we use the fact that $\tilde m(T)=m_T$. Thus $\int_{\mathbb{T}^d}G_\varepsilon(x,m_\varepsilon(T))\,\mathrm{d}x$ is bounded uniformly in ε. From the definition of G_ε, it follows that $\int_{\mathbb{T}^d}|m_\varepsilon(T)-m_T|^p\,\mathrm{d}x\le p\varepsilon C\to0$, i.e. $m_\varepsilon(T)\to m_T$ in $L^p(\mathbb{T}^d)$. Furthermore, by the definition of g_ε, $\varepsilon^{1/p}g_\varepsilon(\cdot,m_\varepsilon(T))$ is uniformly bounded in $L^{p'}(\mathbb{T}^d)$. Thus (i) follows.

The previous arguments and (ii) imply also that $\int_0^T\int_{\mathbb{T}^d}F(\cdot,m_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t$ is uniformly bounded, which by the growth condition on F implies that $(m_\varepsilon)_\varepsilon$ is uniformly bounded in $L^q([0,T]\times\mathbb{T}^d)$.

From Lemma 3.1 one deduces also that $(w_\varepsilon)_\varepsilon$ is uniformly bounded in $L^{\frac{r'q}{r'+q-1}}([0,T]\times\mathbb{T}^d;\mathbb{R}^d)$. Thus (ii) follows.

Part (iii) follows from Proposition 3.4. □

Theorem 3.11 (Approximation of minimizers in the optimization problems (2.14) and (2.21))**.** *For $\varepsilon>0$ let $(u_\varepsilon,m_\varepsilon)$ be the solution of System (3.1) with $g_\varepsilon(\cdot,m_\varepsilon(T))$ as final coupling, where $(u_\varepsilon,\alpha_\varepsilon,\beta_\varepsilon)$ and $(m_\varepsilon,w_\varepsilon)$ are the minimizers of the variational problems (3.8) and (3.7), respectively. Let moreover v_ε be constructed as in Corollary 3.5, satisfying the uniform estimate (3.19).*

Then there exist $(m,w)\in L^q([0,T]\times\mathbb{T}^d)\times L^{\frac{r'q}{r'+q-1}}([0,T]\times\mathbb{T}^d)$ and $(u,\alpha)\in(L^\infty((0,T)\times\mathbb{T}^d)\cap BV((0,T)\times\mathbb{T}^d))\times L^{q'}([0,T]\times\mathbb{T}^d)$ such that choosing a suitable positive vanishing sequence that we denote only by ε, $(m_\varepsilon,w_\varepsilon)\rightharpoonup(m,w)$ and $(v_\varepsilon,\alpha_\varepsilon)\rightharpoonup(u,\alpha)$ as $\varepsilon\downarrow 0$ in the corresponding spaces. Moreover, (u,α) and (m,w) are optimizers in (2.21) and (2.14) respectively.

Proof. Let us once more recall that there exists at least one competitor $(\hat m,\hat w)\in\mathcal{K}_1$ in (2.14) and this infimum is finite.

The existence of (m,w), (u,α) and the previously mentioned weak convergences are consequences of Lemma 3.10 and Corollary 3.5. In fact, let us notice that by Corollary 3.5 (since the estimates on $(v_\varepsilon)_\varepsilon$ are uniform in ε), we have that there exists $u\in BV((0,T)\times\mathbb{T}^d)\cap L^\infty((0,T)\times\mathbb{T}^d)$ with $\nabla u\in L^r((0,T)\times\mathbb{T}^d;\mathbb{R}^d)$ such that possibly passing to a subsequence again with ε, we have that $v_\varepsilon\rightharpoonup u$ as $\varepsilon\downarrow 0$ and the convergence is also locally uniform on $[0,T)\times\mathbb{T}^d$. Then we use once more the same argument as in Proposition 2.8 to see that $v_\varepsilon(T,\cdot)$ converges to $u(T,\cdot)$ in the weak* topology on $\sigma(L^\infty,L^1)$. The convergence of $(m_\varepsilon,w_\varepsilon)$ is straightforward.

Let us show that (m,w) is an optimizer in (2.14). For this, take $(\hat m,\hat w)\in\mathcal{K}_1$ any competitor. This will be a competitor also in the problem for $\mathcal{B}_\varepsilon$ and in particular $G_\varepsilon(\cdot,\hat m(T))=0$. By the definition of G_ε, we have $G_\varepsilon\geq 0$, thus one obtains

$$\mathcal{B}(\hat m,\hat w)=\mathcal{B}_\varepsilon(\hat m,\hat w)\geq\mathcal{B}_\varepsilon(m_\varepsilon,w_\varepsilon)\geq\mathcal{B}(m_\varepsilon,w_\varepsilon)\geq\liminf_{\varepsilon\downarrow 0}\mathcal{B}(m_\varepsilon,w_\varepsilon)\geq\mathcal{B}(m,w),$$

where we have used the optimality of $(m_\varepsilon,w_\varepsilon)$ in (3.7), the fact that $\mathcal{B}_\varepsilon(\bar m,\bar w)\geq\mathcal{B}(\bar m,\bar w)$ for each admissible $(\bar m,\bar w)$ and the weak l.s.c. of $\mathcal{B}$. Notice that by Lemma 3.10(1), we have that $m_\varepsilon(T)\to m_T$ strongly in $L^p(\mathbb{T}^d)$ as $\varepsilon\downarrow 0$, which together with the previous chain of inequalities imply that (m,w) is a solution to (2.14).

Let us show now that (u,α) is a minimizer in (2.21). Take a competitor $(\hat u,\hat\alpha)$. By duality and by the fact that (m,w) is a minimizer in (2.14), one has

$$\begin{aligned}\mathcal{A}(\hat u,\hat\alpha)&\geq-\mathcal{B}(m,w)\geq-\liminf_{\varepsilon\downarrow 0}\mathcal{B}(m_\varepsilon,w_\varepsilon)=\limsup_{\varepsilon\downarrow 0}(-\mathcal{B}(m_\varepsilon,w_\varepsilon))\geq\limsup_{\varepsilon\downarrow 0}(-\mathcal{B}_\varepsilon(m_\varepsilon,w_\varepsilon))\\
&=\limsup_{\varepsilon\downarrow 0}\mathcal{A}_\varepsilon(u_\varepsilon,\alpha_\varepsilon,\beta_\varepsilon)\geq\limsup_{\varepsilon\downarrow 0}\mathcal{A}(u_\varepsilon,\alpha_\varepsilon)\geq\limsup_{\varepsilon\downarrow 0}\mathcal{A}(v_\varepsilon,\alpha_\varepsilon)\\
&\geq\limsup_{\varepsilon\downarrow 0}\left\{\int_0^T\int_{\mathbb{T}^d}F^*(x,\alpha_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t+\int_{\mathbb{T}^d}[v_\varepsilon(T,x)m_T(x)-v_\varepsilon(0,x)m_0(x)]\,\mathrm{d}x\right\}\\
&\geq\int_0^T\int_{\mathbb{T}^d}F^*(x,\alpha)\,\mathrm{d}x\,\mathrm{d}t+\int_{\mathbb{T}^d}[u(T,x)m_T(x)-u(0,x)m_0(x)]\,\mathrm{d}x\\
&=\mathcal{A}(u,\alpha),\end{aligned}\tag{3.38}$$

using the convergence outlined above.

It is routine to verify that (u,α) is actually an element of $\mathcal{K}$. Indeed, since the estimates in Corollary 3.5 are uniform in ε, using the convexity of H in its second variable, passing to the limit as $\varepsilon\downarrow 0$ in the constraint inequality

$$-\partial_t v_\varepsilon+H(x,\nabla v_\varepsilon)\leq\alpha_\varepsilon,$$

one obtains also that (u, α) fulfills the constraint inequality in $\mathcal{K}$, thus (u, α) is a minimizer for the problem (2.21). $\square$

4. Sobolev regularity of weak solutions

In this section, by applying the techniques used in [GM18], we prove some additional a priori regularity for the weak solutions for both the planning problem (1.1) and the mean field game (3.1). (The definition of weak solution is given in Sections 2.3 and 3.4.) We will start with the mean field game; our arguments are easily adapted then to the planning problem. Let us underline the fact that all the estimates below apply to general mean field games with final couplings. However, we mainly emphasize the case of our penalized problem. To simplify the notation, we drop the subscript ε in the definition of g. For general mean field games, we need to assume the following hypotheses.

Additional assumptions

(H6) (Conditions on the coupling) Let f, g be continuous on $\mathbb{T}^d \times (0, \infty)$, strictly increasing in the second variable, satisfying (2.3), (3.2). Moreover, we will assume that $f(x, m), g(x, m)$ are Lipschitz with respect to x, specifically

$$|f(x,m) - f(y,m)| \le C(m^{q-1} + 1)|x - y| \ \ \forall x, y \in \mathbb{T}^d, \ m \ge 0 \tag{4.1}$$

and

$$|g(x,m) - g(y,m)| \le C(m^{p-1} + 1)|x - y| \ \ \forall x, y \in \mathbb{T}^d, \ m \ge 0. \tag{4.2}$$

To ensure that (4.2) holds, in the particular case of g_ε, it suffices to assume that m_T is Lipschitz on $\mathbb{T}^d$ and $p \ge 2$.

We also assume that $f(x, m), g(x, m)$ are strongly monotone in m. That is, we assume there exist $c_f, c_g > 0$ such that

$$\left(f(x,\tilde m) - f(x,m)\right)(\tilde m - m) \ge c_f \min\{\tilde m^{q-2}, m^{q-2}\}|\tilde m - m|^2 \ \forall \tilde m, m \ge 0, \ \tilde m \ne m \tag{4.3}$$

and

$$\begin{aligned}\left(g(x,\tilde m) - g(x,m)\right)(\tilde m - m)& \\ \ge c_g \min\{|\tilde m - m_T(x)|^{p-2}, &|m - m_T(x)|^{p-2}\}|\tilde m - m|^2 \ \forall \tilde m, m \ge 0, \ \tilde m \ne m.\end{aligned} \tag{4.4}$$

Equation (4.4) actually follows from (3.2), where $c_g \sim \varepsilon^{-1}$. If $q < 2$ one should interpret 0^{q-2} as $+\infty$ in (4.3). In this way, when $\tilde m = 0$, for instance, (4.3) reduces to $f(x, m)m \ge c_f m^q$, as in the more regular case $q \ge 2$. An analogous remark holds for $p < 2$ in (4.4). This has to be imposed only when we consider general final coupling functions for which we can allow a growth order of $1 < p < 2$. For our particular choice of penalization g_ε, this remark is irrelevant.

(H7) (Coercivity assumptions.) There exist $j_1, j_2 : \mathbb{R}^d \to \mathbb{R}^d$ and $c_H > 0$ such that

$$H(x,\xi) + H^*(x,\zeta) - \xi\cdot\zeta \ge c_H|j_1(\xi) - j_2(\zeta)|^2. \tag{4.5}$$

In particular, and in light of our restriction (2.1), we will have that $j_1(\xi) \sim |\xi|^{r/2-1}\xi$ and $j_2(\zeta) \sim |\zeta|^{r'/2-1}\zeta$.

4.1. **Global space regularity.** By using arguments analogous to those in [GM18, Proposition 4.3], we get

Proposition 4.1. *In addition to all the previous assumption, we assume that* $m_0 \in W^{2,1}(\mathbb{T}^d)$*, and assume that* H^* *is twice continuously differentiable in* x *with*

$$|D^2_{xx}H^*(x,\zeta)| \leq C|\zeta|^{r'} + C. \tag{H8}$$

Then $\|m^{\frac{q}{2}-1}\nabla m\|_{L^2([0,T]\times\mathbb{T}^d)} \leq C$, $\|m^{1/2}D(j_1(\nabla u))\|_{L^2([0,T]\times\mathbb{T}^d)} \leq C$ *and*

$$\|\tilde{m}(T,\cdot)^{\frac{p}{2}-1}\nabla\tilde{m}(T,\cdot)\|_{L^2(\mathbb{T}^d)} \leq \ C,$$

where $\tilde{m} := m - m_T$.

Proof. We give only a sketch, leaving the reader to find the remaining details in [GM18, Proposition 4.3]. Let (u_n, α_n, β_n) be an approximating sequence of the minimizer $(u, f(\cdot, m), g(\cdot, m(T))$, cf. the proof of Proposition 3.4. Fix $\delta \in \mathbb{T}^d$. Denoting $u_n^\delta(t,x) = u_n(t,x+\delta)$, etc. we deduce

$$\begin{aligned}
&\int_0^T\int_{\mathbb{T}^d}(H(x+\delta,\nabla u_n^\delta) + H^*(x+\delta,-w/m) + \nabla u_n^\delta\cdot w/m)m\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_0^T\int_{\mathbb{T}^d}(H(x-\delta,\nabla u_n^{-\delta}) + H^*(x-\delta,-w/m) + \nabla u_n^{-\delta}\cdot w/m)m\,\mathrm{d}x\,\mathrm{d}t \\
&= \int_{\mathbb{T}^d}\left(\beta_n^\delta + \beta_n^{-\delta} - 2\beta_n\right)m(T)\,\mathrm{d}x - \int_{\mathbb{T}^d}(u_n(0)(m_0^\delta + m_0^{-\delta}) - 2u(0)m_0)\,\mathrm{d}x \\
&+ \int_0^T\int_{\mathbb{T}^d}\left(\alpha_n^\delta + \alpha_n^{-\delta} - 2f(m)\right)m\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_{\mathbb{T}^d}\int_0^1\int_s^{-s}\langle D^2_{xx}H^*(x+\tau\delta,-w/m)\delta,\delta\rangle m\,\mathrm{d}\tau\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{4.6}$$

Equation (4.6) can be obtained by using u_n^δ as a test function in (3.1)(ii) and u_n as a test function in the same equation with m replaced by m^δ, then using the optimality condition (3.24). Notice that in the proof of [GM18, Proposition 4.3] we replace

$$\int_{\mathbb{T}^d}\left(u_T^\delta + u_T^{-\delta} - 2u_T\right)m(T)\,\mathrm{d}x$$

with

$$\int_{\mathbb{T}^d}\left(\beta_n^\delta + \beta_n^{-\delta} - 2\beta_n\right)m(T)\,\mathrm{d}x.$$

Letting $n \to \infty$ in (4.6), following the arguments given in [GM18, Proposition 4.3], we get

$$\begin{aligned}
&\int_0^T\int_{\mathbb{T}^d}(H(x+\delta,\nabla u^\delta) + H^*(x+\delta,-w/m) + \nabla u^\delta\cdot w/m)m\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_0^T\int_{\mathbb{T}^d}(H(x-\delta,\nabla u^{-\delta}) + H^*(x-\delta,-w/m) + \nabla u^{-\delta}\cdot w/m)m\,\mathrm{d}x\,\mathrm{d}t \\
&\leq \int_{\mathbb{T}^d}\left(g^\delta(m^\delta(T)) + g^{-\delta}(m^{-\delta}(T)) - 2g(m(T))\right)m(T)\,\mathrm{d}x - \int_{\mathbb{T}^d}u(0)(m_0^\delta + m_0^{-\delta} - 2m_0)\,\mathrm{d}x \\
&\quad + \int_0^T\int_{\mathbb{T}^d}\left(f^\delta(m^\delta) + f^{-\delta}(m^{-\delta}) - 2f(m)\right)m\,\mathrm{d}x\,\mathrm{d}t + C|\delta|^2\int_0^T\int_{\mathbb{T}^d}(|w/m|^{r'} + 1)m\,\mathrm{d}x\,\mathrm{d}t,
\end{aligned} \tag{4.7}$$

where we have used Hypothesis (H8). We have (using the same arguments as in [GM18])

$$
\int_{\mathbb{T}^d} \left(g^\delta(m^\delta(T)) + g^{-\delta}(m^{-\delta}(T)) - 2g(m(T)) \right) m(T)\,\mathrm{d}x \tag{4.8}
$$

$$
\leq C|\delta|^2 \left(1 + \int_{\mathbb{T}^d} \min\{|\tilde{m}^\delta(T)|, |\tilde{m}(T)|\}^p\,\mathrm{d}x \right) - \frac{c_g}{2} \int_{\mathbb{T}^d} \min\{|\tilde{m}^\delta(T)|^{p-2}, |\tilde{m}(T)|^{p-2}\}|\tilde{m}^\delta(T) - \tilde{m}(T)|^2\,\mathrm{d}x
$$

and similarly for f. Combining these estimates with assumption (4.5), we get

(4.9)

$$
\begin{aligned}
&\frac{c_H}{2}\int_0^T\int_{\mathbb{T}^d}\left(|j_1(\nabla u^\delta) - j_1(\nabla u^{-\delta})|^2\right) m\,\mathrm{d}x\,\mathrm{d}t + \frac{c_f}{2}\int_0^T\int_{\mathbb{T}^d}\min\{(m^\delta)^{q-2}, m^{q-2}\}|m^\delta - m|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \frac{c_g}{2}\int_{\mathbb{T}^d}\min\{|\tilde{m}^\delta(T)|^{p-2}, |\tilde{m}(T)|^{p-2}\}|\tilde{m}^\delta(T) - \tilde{m}(T)|^2\,\mathrm{d}x \\
&\qquad \leq C|\delta|^2\left(1 + \int_{\mathbb{T}^d}\min\{|\tilde{m}^\delta(T)|, |\tilde{m}(T)|\}^p\,\mathrm{d}x + \int_0^T\int_{\mathbb{T}^d}(|w/m|^{r'} + 1)m\,\mathrm{d}x\,\mathrm{d}t\right) \\
&\qquad\qquad + C|\delta|^2\left(T + \int_0^T\int_{\mathbb{T}^d}\min\{m^\delta, m\}^q\,\mathrm{d}x\,\mathrm{d}t\right) + C|\delta|^2\|m_0\|_{W^{2,1}}\|u(0)\|_\infty,
\end{aligned}
$$

where we have used the L^∞ estimate on $u(0)$ from Proposition 3.4. Since also

$$
\lim_{\delta\to 0}\int_{\mathbb{T}^d}\min\{|\tilde{m}^\delta(T)|, |\tilde{m}(T)|\}^p\,\mathrm{d}x + \int_0^T\int_{\mathbb{T}^d}(|w/m|^{r'} + 1)m\,\mathrm{d}x\,\mathrm{d}t \leq C\mathcal{B}(m,w) + C < \infty
$$

we conclude that there exists some C such that

(4.10)

$$
\begin{aligned}
&\frac{c_H}{2}\int_0^T\int_{\mathbb{T}^d}\left(|j_1(\nabla u^\delta) - j_1(\nabla u^{-\delta})|^2\right) m\,\mathrm{d}x\,\mathrm{d}t + \frac{c_f}{2}\int_0^T\int_{\mathbb{T}^d}\min\{(m^\delta)^{q-2}, m^{q-2}\}|m^\delta - m|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \frac{c_g}{2}\int_{\mathbb{T}^d}\min\{|\tilde{m}^\delta(T)|^{p-2}, |\tilde{m}(T)|^{p-2}\}|\tilde{m}^\delta(T) - \tilde{m}(T)|^2\,\mathrm{d}x \leq C|\delta|^2.
\end{aligned}
$$

Dividing by $|\delta|^2$ and letting $\delta \to 0$, we easily obtain the result. □

4.2. **Local time regularity.** We rely on very similar arguments as those found in [GM18, Proposition 4.3], but applied to time rather than space. Our translations in time will be localized so as to avoid conflict with the initial-final conditions.

Let $\varepsilon \in \mathbb{R}$ be small and $\eta : [0,T] \to [0,1]$ be smooth and compactly supported on $(0,T)$ such that $|\varepsilon| < \min\{\mathrm{dist}(0, \mathrm{spt}(\eta)); \mathrm{dist}(T, \mathrm{spt}(\eta))\}$. For competitors (u, α, β) of the minimization problem for $\mathcal{A}$, let us define

$$
u^\varepsilon(t,x) := u(t + \varepsilon\eta(t), x); \quad \alpha^\varepsilon(t,x) := (1 + \varepsilon\eta'(t))\alpha(t + \varepsilon\eta(t), x).
$$

Notice that by construction, if $t \in \{0, T\}$ then $u(t,x) = u^\varepsilon(t,x)$ and $\alpha(t,x) = \alpha^\varepsilon(t,x)$.

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

$$
m^\varepsilon(t,x) := m(t + \varepsilon\eta(t), x); \quad w^\varepsilon(t,x) := (1 + \varepsilon\eta'(t))w(t + \varepsilon\eta(t), x)
$$

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

We define moreover perturbations on the data as

$$
f^\varepsilon(t,x,m) := (1 + \varepsilon\eta'(t))f(x,m); \quad F^\varepsilon(t,x,m) := (1 + \varepsilon\eta'(t))F(x,m),
$$

from which the Legendre transform w.r.t. the last variable satisfies

$$
(F^\varepsilon)^*(t,x,\alpha) := (1 + \varepsilon\eta'(t))F^*(x, \alpha/(1 + \varepsilon\eta'(t))).
$$

Finally, we define

$$H^\varepsilon(t,x,\xi) := (1+\varepsilon\eta'(t))H(x,\xi), \quad \text{thus } (H^\varepsilon)^*(t,x,\zeta) := (1+\varepsilon\eta'(t))H^*(x,\zeta/(1+\varepsilon\eta'(t))).$$

We refer to $\mathcal{A}^\varepsilon$ as the functional $\mathcal{A}$ in Section 3.1 with the data $H^\varepsilon, (F^\varepsilon)^*$ and to $\mathcal{B}^\varepsilon$ as the functional $\mathcal{B}$ in Section 3.1 with data $(H^\varepsilon)^*, F^\varepsilon$.

A very important remark is that by construction (u,α,β) is a minimizer of the problem for $\mathcal{A}$ if and only if $(u^\varepsilon,\alpha^\varepsilon,\beta)$ is a minimizer of the problem for $\mathcal{A}^\varepsilon$. Similarly, (m,w) is a minimizer of the problem for $\mathcal{B}$ if and only if $(m^\varepsilon,w^\varepsilon)$ is a minimizer of the problem for $\mathcal{B}^\varepsilon$. Cf. [GM18, Section 4.1].

In the same spirit as Proposition 4.1, we can formulate

Proposition 4.2. *Let the same monotonicity/coercivity assumptions as for the space regularity be fulfilled and let us suppose moreover that*

$$\text{(H9)} \qquad |D_\zeta H^*(x,\zeta)\cdot\zeta| \le C|\zeta|^{r'} + C, \quad |D^2_{\zeta\zeta}H^*(x,\zeta)| \le C|\zeta|^{r'-2} \quad \text{a.e. } \zeta\in\mathbb{R}^d.$$

If in addition $j_2:\mathbb{R}^d\to\mathbb{R}^d$ satisfies

$$\text{(H10)} \qquad |D_\zeta j_2(\zeta)\cdot\zeta|^2 \le C|\zeta|^{r'} + C, \quad \text{a.e. } \zeta\in\mathbb{R}^d,$$

then

$$m^{1/2}\partial_t(j_1(\nabla u)) \in L^2_{\rm loc}((0,T);L^2(\mathbb{T}^d)).$$

We also have (only assuming the monotonicity condition on f and g and not (4.1) and (4.2))

$$\partial_t(m^{q/2}) \in L^2_{\rm loc}((0,T);L^2(\mathbb{T}^d)),$$

and the two previous bounds depend only on the data.

Proof of Proposition 4.2. The proof closely follows the steps of [GM18, Proposition 4.3], but with translations in time rather than space.

Step 0. Preparatory step. Take a smooth minimizing sequence $(u_n)_{n\ge0}$ in the problem for $\mathcal{A}$, defining

$$\alpha_n = -\partial_t u_n + H(x,\nabla u_n).$$

Now use u_n as test function for $\partial_t m^\varepsilon + \nabla\cdot w^\varepsilon = 0$. In the same way, use u_n^ε (defined as u^ε) as test function for $\partial_t m + \nabla\cdot w = 0$.

Then on the one hand one has

$$\text{(4.11)} \qquad \int_{\mathbb{T}^d}[u_n(T)m(T) - u_n(0)m_0]\,\mathrm{d}x = \int_0^T\int_{\mathbb{T}^d}[w^\varepsilon\cdot\nabla u_n + (H(x,\nabla u_n)-\alpha_n)m^\varepsilon]\,\mathrm{d}x\,\mathrm{d}t.$$

On the other hand, one gets

$$\text{(4.12)} \qquad \int_{\mathbb{T}^d}[u_n(T)m(T) - u_n(0)m_0]\,\mathrm{d}x = \int_0^T\int_{\mathbb{T}^d}[w\cdot\nabla u_n^\varepsilon + (H^\varepsilon(t,x,\nabla u_n^\varepsilon)-\alpha_n^\varepsilon)m]\,\mathrm{d}x\,\mathrm{d}t.$$

Combining (4.12) with (3.24) one obtains

$$\begin{aligned}\text{(4.13)} \qquad &\int_{\mathbb{T}^d}[u_n(T) - g(x,m(T))]m(T) - [u_n(0)-u(0)]m_0\,\mathrm{d}x \\ &= \int_0^T\int_{\mathbb{T}^d}[H^\varepsilon(t,x,\nabla u_n^\varepsilon) + H^*(x,-w/m) + \nabla u_n^\varepsilon\cdot w/m + f(x,m) - \alpha_n^\varepsilon]\,m\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_0^T\int_{\mathbb{T}^d}[H^\varepsilon(t,x,\nabla u_n^\varepsilon) + (H^\varepsilon)^*(t,x,-w/m) + \nabla u_n^\varepsilon\cdot w/m + f(x,m) - \alpha_n^\varepsilon]\,m\,\mathrm{d}x\,\mathrm{d}t \\ &+ \int_0^T\int_{\mathbb{T}^d}[H^*(x,-w/m) - (H^\varepsilon)^*(t,x,-w/m)]\,m\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

Similarly, combining (4.11) with (3.24) (for the ε-translates) one obtains

(4.14)
$$
\begin{aligned}
&\int_{\mathbb{T}^d}[u_n(T)-g(x,m(T))]m(T)-[u_n(0)-u(0)]m_0\,\mathrm{d}x\\
&=\int_0^T\int_{\mathbb{T}^d}[H(x,\nabla u_n)+(H^\varepsilon)^*(t,x,-w^\varepsilon/m^\varepsilon)+\nabla u_n\cdot w^\varepsilon/m^\varepsilon+f^\varepsilon(t,x,m^\varepsilon)-\alpha_n]\,m^\varepsilon\,\mathrm{d}x\,\mathrm{d}t\\
&=\int_0^T\int_{\mathbb{T}^d}\left[H^{-\varepsilon}(s,x,\nabla u_n^{-\varepsilon})+H^*(x,-w/m)+\nabla u_n^{-\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}s\\
&\quad+\int_0^T\int_{\mathbb{T}^d}O(\varepsilon^2)H(x,\nabla u_n^{-\varepsilon})m\,\mathrm{d}x\,\mathrm{d}s+\int_0^T\int_{\mathbb{T}^d}[f^\varepsilon(t,x,m^\varepsilon)-\alpha_n]\,m^\varepsilon\,\mathrm{d}x\,\mathrm{d}t\\
&=\int_0^T\int_{\mathbb{T}^d}\left[H^{-\varepsilon}(s,x,\nabla u_n^{-\varepsilon})+(H^{-\varepsilon})^*(s,x,-w/m)+\nabla u_n^{-\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}s\\
&\quad+\int_0^T\int_{\mathbb{T}^d}\left[H^*(x,-w/m)-(H^{-\varepsilon})^*(s,x,-w/m)\right]m\,\mathrm{d}x\,\mathrm{d}s+\int_0^T\int_{\mathbb{T}^d}O(\varepsilon^2)H(x,\nabla u_n^{-\varepsilon})m\,\mathrm{d}x\,\mathrm{d}s\\
&\quad+\int_0^T\int_{\mathbb{T}^d}[f^\varepsilon(t,x,m^\varepsilon)-\alpha_n]\,m^\varepsilon\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
$$

where in the penultimate equation we used the change of variable $s=t+\varepsilon\eta(t)$ (which means in particular that $t=s-\varepsilon\eta(s)+O(\varepsilon^2)$ and $\frac{1}{1+\varepsilon\eta'(t)}=1-\varepsilon\eta'(s)+O(\varepsilon^2)$). By slight abuse of notation we denoted

$$u_n^{-\varepsilon}(s,x):=u_n(s-\varepsilon\eta(s)+O(\varepsilon^2),x),$$

and we use the original notation for $H^{-\varepsilon}$ and $(H^{-\varepsilon})^*$. Adding (4.13) to (4.14) we get

$$
\begin{aligned}
(4.15)\quad&\int_0^T\int_{\mathbb{T}^d}\left[H^{-\varepsilon}(t,x,\nabla u_n^{-\varepsilon})+(H^{-\varepsilon})^*(t,x,-w/m)+\nabla u_n^{-\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}t\\
&\quad+\int_0^T\int_{\mathbb{T}^d}[H^\varepsilon(t,x,\nabla u_n^\varepsilon)+(H^\varepsilon)^*(t,x,-w/m)+\nabla u_n^\varepsilon\cdot w/m]\,m\,\mathrm{d}x\,\mathrm{d}t\\
&=\int_0^T\int_{\mathbb{T}^d}[\alpha_n^\varepsilon-f(x,m)]\,m\,\mathrm{d}x\,\mathrm{d}t+\int_0^T\int_{\mathbb{T}^d}[\alpha_n-f^\varepsilon(t,x,m^\varepsilon)]\,m^\varepsilon\,\mathrm{d}x\,\mathrm{d}t\\
&\quad+2\int_{\mathbb{T}^d}[u_n(T)-g(x,m(T))]m(T)-[u_n(0)-u(0)]m_0\,\mathrm{d}x+R_n(\varepsilon)
\end{aligned}
$$

where the remainder term satisfies

$$
\begin{aligned}
R_n(\varepsilon)=&\int_0^T\int_{\mathbb{T}^d}\left[(H^\varepsilon)^*(t,x,-w/m)+(H^{-\varepsilon})^*(t,x,-w/m)-2H^*(x,-w/m)\right]m\,\mathrm{d}x\,\mathrm{d}t\\
&+O(\varepsilon^2)\int_0^T\int_{\mathbb{T}^d}H(x,\nabla u_n^{-\varepsilon})m\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

Step 1. Error term. Before letting $n\to\infty$ let us first show that $R_n(\varepsilon)=O(\varepsilon^2)$ (uniformly in n). To that end we estimate the terms $H^*-(H^\varepsilon)^*$ and $H^*-(H^{-\varepsilon})^*$. By a Taylor expansion, we have

$$
\begin{aligned}
(H^\varepsilon)^*(t,x,\zeta)&=(1+\varepsilon\eta'(t))H^*(x,\zeta/(1+\varepsilon\eta'(t)))=(1+\varepsilon\eta'(t))H^*(x,(1-\varepsilon\eta'(t)+O(\varepsilon^2))\zeta)\\
&=(1+\varepsilon\eta'(t))\left[H^*(x,\zeta)-\varepsilon\eta'(t)D_\zeta H^*(x,\zeta)\cdot\zeta+O(\varepsilon^2)D_\zeta H^*(x,\zeta)\cdot\zeta\right]\\
&\quad+(1+\varepsilon\eta'(t))\left[\varepsilon\eta'(t)+O(\varepsilon^2)\right]^2\frac{1}{2}D^2_{\zeta\zeta}H^*(x,\zeta^*_\varepsilon)\zeta\cdot\zeta
\end{aligned}
$$

where ζ_ε^* is a point on the segment between ζ and $(1-\varepsilon\eta'(t)+O(\varepsilon^2))\zeta$. Let us notice that due to the growth condition (H9) we have that

$$|D^2_{\zeta\zeta}H^*(x,\zeta_\varepsilon^*)\zeta\cdot\zeta| \leq C|\zeta_\varepsilon^*|^{r'-2}|\zeta|^2,$$

where, by the comparison $(1-|\varepsilon|)|\zeta| \leq |\zeta_\varepsilon^*| \leq (1+|\varepsilon|)|\zeta|$, the right-hand side is finite even when $\zeta_\varepsilon^* = 0$, and in particular

$$|D^2_{\zeta\zeta}H^*(x,\zeta_\varepsilon^*)\zeta\cdot\zeta| \leq C\max\left\{(1-|\varepsilon|)^{r'-2},(1+|\varepsilon|)^{r'-2}\right\}|\zeta|^{r'}.$$

Therefore, by (H9) we have

$$(4.16)\quad (H^\varepsilon)^*(t,x,\zeta) - (1+\varepsilon\eta'(t))H^*(x,\zeta) + \varepsilon\eta'(t)D_\zeta H^*(x,\zeta)\cdot\zeta \\ = O(\varepsilon^2)\left[D_\zeta H^*(x,\zeta)\cdot\zeta + \frac{1}{2}D^2_{\zeta\zeta}H^*(x,\zeta_\varepsilon^*)\zeta\cdot\zeta\right] = O(\varepsilon^2)(|\zeta|^{r'}+1).$$

Using a similar argument, we deduce

$$(4.17)\quad (H^{-\varepsilon})^*(t,x,\zeta) - (1-\varepsilon\eta'(t))H^*(x,\zeta) - \varepsilon\eta'(t)D_\zeta H^*(x,\zeta)\cdot\zeta = O(\varepsilon^2)(|\zeta|^{r'}+1).$$

Adding together (4.16) and (4.17), setting $\zeta = -w/m$ and then integrating against m, we get

$$(4.18)\quad \int_0^T\int_{\mathbb{T}^d}\left[(H^\varepsilon)^*(t,x,-w/m) + (H^{-\varepsilon})^*(t,x,-w/m) - 2H^*(x,-w/m)\right]m\,\mathrm{d}x\,\mathrm{d}t \\ = \int_0^T\int_{\mathbb{T}^d} O(\varepsilon^2)\left|\frac{w}{m}\right|^{r'} m\,\mathrm{d}x\,\mathrm{d}t = O(\varepsilon^2),$$

where in the last equation we used the assumption (2.2) and the fact that $\mathcal{B}(m,w)$ is finite.

As for what remains of $R_n(\varepsilon)$, we use $u_n^{-\varepsilon}$ as a test function in $\partial_t m + \nabla\cdot w = 0$; with the appropriate change of variable we get

$$(4.19)\quad \int_0^T\int_{\mathbb{T}^d} H(x,\nabla u_n^{-\varepsilon})m\,\mathrm{d}x\,\mathrm{d}t \\ \leq (1+O(\varepsilon))\left[\int_{\mathbb{T}^d} u_n(T)_+ m(T) - u_n(0)m_0\,\mathrm{d}x + \int_0^T\int_{\mathbb{T}^d}(\alpha_n)_+(s-\eta(s)+O(\varepsilon^2),x)m(s,x)\,\mathrm{d}x\,\mathrm{d}s\right].$$

Recalling from the definition of weak solution that $m \in L^q((0,T)\times\mathbb{T}^d), m(T) \in L^p(\mathbb{T}^d)$ while from the proof of Proposition 3.4 we have that $(\alpha_n)_+$ is bounded in $L^{q'}$ and $u_n(T)$ is bounded in L^∞, it follows that $\int_0^T\int_{\mathbb{T}^d} H(x,\nabla u_n^{-\varepsilon})m\,\mathrm{d}x\,\mathrm{d}s$ is bounded. We can now rewrite (4.15) as

$$(4.20)\quad \int_0^T\int_{\mathbb{T}^d}\left[H^{-\varepsilon}(t,x,\nabla u_n^{-\varepsilon}) + (H^{-\varepsilon})^*(t,x,-w/m) + \nabla u_n^{-\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}t \\ + \int_0^T\int_{\mathbb{T}^d}\left[H^\varepsilon(t,x,\nabla u_n^\varepsilon) + (H^\varepsilon)^*(x,-w/m) + \nabla u_n^\varepsilon\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}t \\ = \int_0^T\int_{\mathbb{T}^d}\left[\alpha_n^\varepsilon - f(x,m)\right]m\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_{\mathbb{T}^d}\left[\alpha_n - f^\varepsilon(t,x,m^\varepsilon)\right]m^\varepsilon\,\mathrm{d}x\,\mathrm{d}t \\ + 2\int_{\mathbb{T}^d}[u_n(T) - g(x,m(T))]m(T) - [u_n(0)-u(0)]m_0\,\mathrm{d}x + O(\varepsilon^2).$$

Step 2. Taking $n\to\infty$. We can now proceed exactly as in the proof of [GM18, Proposition 4.3.] when taking limits as $n\to+\infty$ in (4.20). First notice that we have the weak convergence (up to a subsequence) of $\nabla u_n^{-\varepsilon} \rightharpoonup \nabla u^{-\varepsilon}, \nabla u_n^\varepsilon \rightharpoonup \nabla u^\varepsilon$ in $L^r((0,T)\times\mathbb{T}^d)$. Second recall that α_n converges weakly in $L^{q'}((0,T)\times\mathbb{T}^d)$ to $f(x,m)$ on the set where it is bounded below by an arbitrary constant, and likewise α_n^ε converges weakly to $f^\varepsilon(t,x,m^\varepsilon)$. Third recall that $u_n(T)$ converges to $g(x,m(T))$ in the weak* topology $\sigma(L^\infty,L^1)$, and $u_n(0)\to u(0)$ uniformly (see the proof of Proposition 2.8). Arguing as in [GM18], we deduce that

$$\begin{aligned}(4.21)\quad &\int_0^T\int_{\mathbb{T}^d}\left[H^{-\varepsilon}(t,x,\nabla u^{-\varepsilon})+(H^{-\varepsilon})^*(t,x,-w/m)+\nabla u^{-\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}t\\ &\quad+\int_0^T\int_{\mathbb{T}^d}\left[H^{\varepsilon}(t,x,\nabla u^{\varepsilon})+(H^{\varepsilon})^*(t,x,-w/m)+\nabla u^{\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}t\\ &\qquad\le-\int_0^T\int_{\mathbb{T}^d}\left(f^\varepsilon(t,x,m^\varepsilon)-f(x,m)\right)(m^\varepsilon-m)\,\mathrm{d}x\,\mathrm{d}t+O(\varepsilon^2).\end{aligned}$$

Step 3. Time regularity for u.

By the coercivity condition on H and H^* for any $\gamma>0$ we have that

$$\gamma H(x,\xi)+\gamma H^*(x,\zeta)-\gamma\xi\cdot\zeta\ge\gamma c_H|j_1(\xi)-j_2(\zeta)|^2,\quad\forall x\in\mathbb{T}^d,\xi,\zeta\in\mathbb{R}^d.$$

In particular, setting $\tilde\zeta:=\gamma\zeta$, this implies

$$\gamma H(x,\xi)+\gamma H^*(x,\tilde\zeta/\gamma)-\xi\cdot\tilde\zeta\ge\gamma c_H|j_1(\xi)-j_2(\tilde\zeta/\gamma)|^2,\quad\forall x\in\mathbb{T}^d,\xi,\tilde\zeta\in\mathbb{R}^d.$$

Therefore, we have

$$\begin{aligned}(4.22)\quad &\int_0^T\int_{\mathbb{T}^d}c_H(1+\varepsilon\eta'(t))\left|j_1(\nabla u^\varepsilon)-j_2\left(-\frac{w}{(1+\varepsilon\eta'(t))m}\right)\right|^2m\,\mathrm{d}x\,\mathrm{d}t\\ &\qquad\le\int_0^T\int_{\mathbb{T}^d}\left[H^{\varepsilon}(t,x,\nabla u^{\varepsilon})+(H^{\varepsilon})^*(t,x,-w/m)+\nabla u^{\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}t\end{aligned}$$

and similarly

(4.23)

$$\begin{aligned}&\int_0^T\int_{\mathbb{T}^d}c_H(1-\varepsilon\eta'(t))\left|j_1(\nabla u^{-\varepsilon})-j_2\left(-\frac{w}{(1-\varepsilon\eta'(t))m}\right)\right|^2m\,\mathrm{d}x\,\mathrm{d}t\\ &\qquad\le\int_0^T\int_{\mathbb{T}^d}\left[H^{-\varepsilon}(t,x,\nabla u^{-\varepsilon})+(H^{-\varepsilon})^*(t,x,-w/m)+\nabla u^{\varepsilon}\cdot w/m\right]m\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

By the triangle inequality,

$$\begin{aligned}(4.24)\quad &\int_0^T\int_{\mathbb{T}^d}\frac{c_H}{3}\min\{1+\varepsilon\eta'(t);1-\varepsilon\eta'(t)\}\left|j_1(\nabla u^\varepsilon)-j_1(\nabla u^{-\varepsilon})\right|^2m\,\mathrm{d}x\,\mathrm{d}t\\ &\le(1+\varepsilon)c_H\int_0^T\int_{\mathbb{T}^d}\left|j_1(\nabla u^\varepsilon)-j_2\left(-\frac{w}{(1+\varepsilon\eta'(t))m}\right)\right|^2m\,\mathrm{d}x\,\mathrm{d}t\\ &+(1+\varepsilon)c_H\int_0^T\int_{\mathbb{T}^d}\left|j_2\left(-\frac{w}{(1+\varepsilon\eta'(t))m}\right)-j_2\left(-\frac{w}{(1-\varepsilon\eta'(t))m}\right)\right|^2m\,\mathrm{d}x\,\mathrm{d}t\\ &+(1+\varepsilon)c_H\int_0^T\int_{\mathbb{T}^d}\left|j_1(\nabla u^{-\varepsilon})-j_2\left(-\frac{w}{(1-\varepsilon\eta'(t))m}\right)\right|^2m\,\mathrm{d}x\,\mathrm{d}t,\end{aligned}$$

where it remains to estimate the second term on the right-hand side. For this we note that

$$\begin{aligned}&\left|j_2(\zeta/(1+\varepsilon\eta'(t)))-j_2(\zeta/(1-\varepsilon\eta'(t)))\right|^2=\left|j_2(\zeta(1-\varepsilon\eta'(t)+O(\varepsilon^2)))-j_2(\zeta(1+\varepsilon\eta'(t)+O(\varepsilon^2)))\right|^2\\ &=|D_\zeta j_2(\zeta(1-\varepsilon\eta'(t)+O(\varepsilon^2)))\cdot\zeta\eta'(t)|^2\varepsilon^2\le C|\zeta|^{r'}\varepsilon^2,\end{aligned}$$

where the last constant depends only on $\eta'(t)$ and the constant in the hypothesis (H10). Setting $\zeta := -w/m$ in the previous inequality, we find that the second term on the right-hand side of (4.24) is $O(\varepsilon^2)$ since $\int_0^T \int_{\mathbb{T}^d} |w/m|^{r'} m \,\mathrm{d}x\,\mathrm{d}t$ is finite. Equation (4.21) now becomes

$$\text{(4.25)} \quad \frac{c_H}{6} \int_0^T \int_{\mathbb{T}^d} \left|j_1(\nabla u^\varepsilon) - j_1(\nabla u^{-\varepsilon})\right|^2 m \,\mathrm{d}x\,\mathrm{d}t \\ \leq - \int_0^T \int_{\mathbb{T}^d} \left(f^\varepsilon(t,x,m^\varepsilon) - f(x,m)\right)(m^\varepsilon - m)\,\mathrm{d}x\,\mathrm{d}t + O(\varepsilon^2)$$

for ε small enough.

Step 4. Time regularity for m.
We have

$$\begin{aligned}
&- \int_0^T \int_{\mathbb{T}^d} \left(f^\varepsilon(t,x,m^\varepsilon) - f(x,m)\right)(m^\varepsilon - m)\,\mathrm{d}x\,\mathrm{d}t \\
&= - \iint_{\{m^\varepsilon \leq m\}} (f^\varepsilon(t,x,m^\varepsilon) - f(x,m^\varepsilon))(m^\varepsilon - m)\,\mathrm{d}x\,\mathrm{d}t - \iint_{\{m^\varepsilon \leq m\}} (f(x,m^\varepsilon) - f(x,m))(m^\varepsilon - m)\,\mathrm{d}x\,\mathrm{d}t \\
&- \iint_{\{m < m^\varepsilon\}} (f^\varepsilon(t,x,m^\varepsilon) - f^\varepsilon(t,x,m))(m^\varepsilon - m)\,\mathrm{d}x\,\mathrm{d}t - \iint_{\{m < m^\varepsilon\}} (f^\varepsilon(t,x,m) - f(x,m))(m^\varepsilon - m)\,\mathrm{d}x\,\mathrm{d}t \\
&\leq C \int_0^T \int_{\mathbb{T}^d} |\varepsilon| \min\{(m^\varepsilon)^{q-1}, m^{q-1}\}|m^\varepsilon - m|\,\mathrm{d}x\,\mathrm{d}t - c_0 \int_0^T \int_{\mathbb{T}^d} \min\{(m^\varepsilon)^{q-2}, m^{q-2}\}|m^\varepsilon - m|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\leq C|\varepsilon|^2 \int_0^T \int_{\mathbb{T}^d} \min\{m^\varepsilon, m\}^q\,\mathrm{d}x\,\mathrm{d}t - \frac{c_0}{2} \int_0^T \int_{\mathbb{T}^d} \min\{(m^\varepsilon)^{q-2}, m^{q-2}\}|m^\varepsilon - m|^2\,\mathrm{d}x\,\mathrm{d}t,
\end{aligned}$$

where, we used Young's inequality in the last inequality, and the expression $\min\{(m^\varepsilon)^{q-2}, m^{q-2}\}|m^\varepsilon - m|^2$ is treated as zero whenever $m^\varepsilon = m$ (even in the case $q < 2$). Since

$$\int_0^T \int_{\mathbb{T}^d} \min\{m^\varepsilon, m\}^q\,\mathrm{d}x\,\mathrm{d}t \leq \int_0^T \int_{\mathbb{T}^d} m^q\,\mathrm{d}x\,\mathrm{d}t \leq C,$$

Equation (4.25) now becomes

$$\frac{c_H}{6} \int_0^T \int_{\mathbb{T}^d} \left|j_1(\nabla u^\varepsilon) - j_1(\nabla u^{-\varepsilon})\right|^2 m\,\mathrm{d}x\,\mathrm{d}t + \frac{c_0}{2} \int_0^T \int_{\mathbb{T}^d} \min\{(m^\varepsilon)^{q-2}, m^{q-2}\}|m^\varepsilon - m|^2\,\mathrm{d}x\,\mathrm{d}t = O(\varepsilon^2).$$

Dividing the previous identity by ε^2 and letting $\varepsilon \to 0$, we conclude that $m^{1/2}\partial_t j_1(\nabla u),\ \partial_t(m^{q/2}) \in L^2_{\mathrm{loc}}((0,T); L^2(\mathbb{T}^d))$, with norms estimated by a constant depending only on the data. The proof is complete. □

4.3. **Sobolev estimates for the solution of the planning problem.** One can rely on the same arguments as in Sections 4.1 4.2 to obtain Sobolev estimates for the solutions on the planning problem. Thus, we are in position to formulate the following results.

Proposition 4.3 (Global in time space regularity)**.** *Let $m_0, m_T \in W^{2,1}(\mathbb{T}^d)$ and let us assume that the hypotheses of Proposition 4.1 are fulfilled. Then*

$$m^{\frac{q}{2}-1}\nabla m \in L^2([0,T] \times \mathbb{T}^d; \mathbb{R}^d) \quad \text{and} \quad m^{1/2} D(j_1(\nabla u)) \in L^2([0,T] \times \mathbb{T}^d; \mathbb{R}^{d\times d}),$$

where the bounds depend only on the data.

Proof. The proof follows the same lines as the one of Proposition 4.1. We observe that using the argument of Proposition 2.8 we can construct suitable minimizing sequences (u_n, α_n) for the problem (2.21) which satisfy (2.38) uniformly in n. The only difference in the proof is to treat differently one of the time-boundary terms. To estimate the term

$$\int_{\mathbb{T}^d} \left(u_n^{\delta}(T) + u_n^{-\delta}(T) - 2u_n(T)\right) m(T)\, \mathrm{d}x,$$

in the proof of Proposition 4.1, we pass the translates to m_T. It becomes now

$$\int_{\mathbb{T}^d} u_n(T) \left(m_T^{\delta} + m_T^{-\delta} - 2m_T\right) \mathrm{d}x,$$

which, by the assumption on m_T and the fact that $u_n(T)$ is uniformly bounded, is $O(|\delta|^2)$.

We can deal with the boundary term at time zero in the exact same way as before, and in particular no change in the proof is needed there. □

Proposition 4.4 (Local time regularity)**.** *Let us assume that the hypotheses of Proposition 4.2 are fulfilled. Then*

$$m^{1/2}\partial_t(j_1(\nabla u)) \in L^2_{\mathrm{loc}}((0,T); L^2(\mathbb{T}^d)),$$

and

$$\partial_t(m^{q/2}) \in L^2_{\mathrm{loc}}((0,T); L^2(\mathbb{T}^d)),$$

and the two previous bounds depend only on the data.

Proof. We use the same proof as in Proposition 4.2, again using an approximating sequence derived from Proposition 2.8. We only remark that nowhere in the proof of Proposition 4.2 do we need $m(T) \in L^p(\mathbb{T}^d)$; here it suffices to have $m_T \in L^1(\mathbb{T}^d)$. □

5. Open questions and further directions

• An interesting direction of study would be the relaxation of the joint assumption on q and r (the growth exponent of F and H, respectively) in the case of the planning problem. This should somehow imply also a more precise link between our work and the results of Orrieri-Porretta-Savaré in [OPS]. This direction would be strongly related to the search for higher order summability estimates on the m variable, also in the spirit of [LS17].

• The well-posedness of the second order (degenerate) planning problem is largely open (except the non-degenerate case with essentially quadratic Hamiltonians in [Lio, Por13, Por14]). Our hope is that the variational approach developed in the present paper and exploited also in [OPS] will provide hints to attack the general second order problem.

In this direction the exact controllability problem of the Fokker-Planck equation with general initial and final conditions (pointed out also by Lions in his lectures) seems to be an interesting open question, by its own.

• We are aiming to pursue the global in time a priori Sobolev estimates in the time variable for the solutions of both mean field games and the planning problem, which so far seem to be inaccessible by relying only on our current techniques.

References

[ACCD12] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: numerical methods for the planning problem. SIAM J. Control Optim., 50(1):77–109, (2012).

[AF08] L. Ambrosio and A. Figalli. On the regularity of the pressure field of Brenier's weak solutions to incompressible Euler equations. Calc. Var. Partial Differential Equations, 31(4):497–509, (2008).

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara. Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, (2000).

[BB00] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numer. Math., 84(3):375–393, (2000).

[Bre99] Y. Brenier. Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. Comm. Pure Appl. Math., 52(4):411–452, (1999).

[Car15] P. Cardaliaguet. Weak solutions for first order mean field games with local coupling. In Analysis and geometry in control theory and its applications, volume 11 of Springer INdAM Ser., pages 111–158. Springer, Cham, (2015).

[CCN13] P. Cardaliaguet, G. Carlier, and B. Nazaret. Geodesics for a class of distances in the space of probability measures. Calc. Var. Partial Differential Equations, 48(3-4):395–420, (2013).

[CG15] P. Cardaliaguet and P. J. Graber. Mean field games systems of first order. ESAIM Control Optim. Calc. Var., 21(3):690–722, (2015).

[CGPT15] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon. Second order mean field games with degenerate diffusion and local coupling. NoDEA Nonlinear Differential Equations Appl., 22(5):1287–1317, (2015).

[CMS16] P. Cardaliaguet, A.R. Mészáros, and F. Santambrogio. First order Mean Field Games with density constraints: pressure equals price. SIAM J. Control Optim., 54(5):2672–2709, (2016).

[CS12] P. Cardaliaguet and L. Silvestre. Hölder continuity to Hamilton-Jacobi equations with superquadratic growth in the gradient and unbounded right-hand side. Communications in Partial Differential Equations, 37(9):1668–1688, (2012).

[DL89] R.J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and sobolev spaces. Inventiones mathematicae, 98(3):511–547, (1989).

[DNS09] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. Calc. Var. Partial Differential Equations, 34:193–231, (2009).

[ET76] I. Ekeland and 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).

[GM18] P. J. Graber and A. R. Mészáros. Sobolev regularity for first order mean field games. Ann. Inst. H. Poincaré Anal. Non Linéaire, to appear, (2018).

[Gra14] P. J. Graber. Optimal control of first-order Hamilton-Jacobi equations with linearly bounded Hamiltonian. Appl. Math. Optim., 70(2):185–224, (2014).

[GS] D. A. Gomes and T. Seneci. Displacement convexity for first-order mean-field games. arXiv:1807.07090.

[HMC06] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. Commun. Inf. Syst., 6(3):221–251, (2006).

[Ish95] H. Ishii. On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions. Funkcial. Ekvac, 38(1):101–120, (1995).

[LBL08] C. Le Bris and P-L. Lions. Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. Communications in Partial Differential Equations, 33(7):1272–1317, (2008).

[Lio] P.-L. Lions. *Cours au Collège de France*. www.college-de-france.fr, (lectures on November 27th, December 4th-11th, 2009).

[LL] J. Liu and G. Loeper. Optimal transport with discrete long range mean field interactions. arXiv:1809.07432.

[LL06a] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen I. Le cas stationnaire. C. R. Math. Acad. Sci. Paris, 343:619–625, (2006).

[LL06b] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen II. Horizon fini et contrôle optimal. C. R. Math. Acad. Sci. Paris, 343:679–684, (2006).

[LL07] J.-M. Lasry and P.-L. Lions. Mean field games. Jpn. J. Math., 2:229–260, (2007).

[LS17] H. Lavenant and F. Santambrogio. Optimal density evolution with congestion: L^∞ bounds via flow interchange techniques and applications to variational mean field games. preprint, (2017).

[OPS] C. Orrieri, A. Porretta, and G. Savaré. A variational approach to the mean field planning problem. arXiv:1807.09874.

[Por13] A. Porretta. On the planning problem for a class of mean field games. C. R. Math. Acad. Sci. Paris, 351(11-12):457–462, (2013).

[Por14] A. Porretta. On the planning problem for the mean field games system. Dyn. Games Appl., 4(2):231–256, (2014).

[PS17] A. Prosinski and F. Santambrogio. Global-in-time regularity via duality for congestion-penalized mean field games. Stochastics, 89(6-7):923–942, (2017).

[Roc71] R T Rockafellar. Integrals which are convex functionals. II. Pacific Journal of Mathematics, 39(2):439–469, 1971.

[San18] F. Santambrogio. Regularity via duality in calculus of variations and degenerate elliptic PDEs. J. Math. Anal. Appl., 457(2):1649–1674, (2018).

Department of Mathematics, Baylor University, One Bear Place, Waco, TX 97328, USA
E-mail address: jameson_graber@baylor.edu

Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, CA 90095, USA
E-mail address: alpar@math.ucla.edu

Institut de recherche XLIM-DMI, UMR-CNRS 7252 Faculté des sciences et techniques Université de Limoges, 87060 Limoges, France
E-mail address: francisco.silva@unilim.fr

CEREMADE, CNRS, [UMR 7534], Université Paris-Dauphine, PSL Research University, 75016 Paris, France
E-mail address: tonon@ceremade.dauphine.fr