HAL open science

Time-dependent focusing Mean-Field Games: the sub-critical case

Marco Cirant, Daniela Tonon

▶ To cite this version:

Marco Cirant, Daniela Tonon. Time-dependent focusing Mean-Field Games: the sub-critical case. Journal of Dynamics and Differential Equations, 2019, 31 (1), pp.49-79. hal-01508506v2

HAL Id: hal-01508506
https://hal.science/hal-01508506v2

Submitted on 30 Aug 2018

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

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

Time-dependent focusing Mean-Field Games: the sub-critical case

Marco Cirant and Daniela Tonon

April 19, 2018

Abstract

We consider time-dependent viscous Mean-Field Games systems in the case of local, decreasing and unbounded couplings. These systems arise in mean-field game theory, and describe Nash equilibria of games with a large number of agents aiming at aggregation. We prove the existence of weak solutions that are minimizers of an associated non-convex functional, by rephrasing the problem in a convex framework. Under additional assumptions involving the growth at infinity of the coupling, the Hamiltonian, and the space dimension, we show that such minimizers are indeed classical solutions by a blow-up argument and additional Sobolev regularity for the Fokker-Planck equation. We exhibit an example of non-uniqueness of solutions. Finally, by means of a contraction principle, we observe that classical solutions exist just by local regularity of the coupling if the time horizon is short.

AMS-Subject Classification. 35K55, 49N70.

Keywords. Variational formulation of Mean Field Games, local decreasing coupling, non-uniqueness.

1 Introduction

Mean Field Games (MFG) theory models the behavior of an infinite number of indistinguishable rational agents aiming at minimizing a common cost. The theory was introduced in the seminal papers by Lasry and Lions [19, 20, 21] and by Huang, Caines and Malhamé [16] to describe Nash equilibria in differential games with infinitely many players. A large part of MFG literature is devoted to the study of MFG systems with increasing coupling. Heuristically, this assumption means that agents prefer sparsely populated areas (indeed concentration costs), and it is well-suited to model competitive cases. The increasing monotonicity of the coupling ensures existence and regularity of solutions in many circumstances (see, e.g., [13] and references therein, or the courses at Collège de France by P.-L. Lions [22]); it is also a key assumption if one looks for uniqueness of equilibria. Although those models have a wide range of applications, they rule out the possibility to apply the MFG theory to analyse aggregation phenomena, that is when agents aim at converging to a common state. To cite an example, in [14], Guéant considered simple population models where individuals have preferences about resembling to each other. Very few results exist in this direction and they only deal with very particular cases. See [12, 15] and [5], for the quadratic and linear-quadratic case. Our goal is to better understand this class of "focusing" MFG systems, where the coupling is monotone decreasing and it is a local function of the distribution, so that no regularising effect can be expected. Actually, non-existence, non-uniqueness of solutions, non-smoothness, and concentration are likely to arise, as shown by the first author in [10], where the stationary focusing case is considered. In that paper, it is proven that there exists a threshold for the growth of the coupling, after which solutions to the MFG system may not even exist. Indeed, the focusing character of the MFG induces solutions to concentrate and develop singularities.

Let us enter into the details of the kind of MFG systems we deal with in this paper. In order to avoid boundary issues and exploit the compactness of the state space, we set our problem on the N-dimensional flat torus $\mathbb{T}^N$. Let $Q = Q_T = \mathbb{T}^N \times (0,T)$. We consider MFG systems of the form

$$\begin{cases} -u_t - \Delta u + H(\nabla u) = -f(x, m(x,t)), & \text{in } Q, \\ m_t - \Delta m - \operatorname{div}(\nabla H(\nabla u)\, m) = 0 & \text{in } Q, \\ m(x,0) = m_0(x), \quad u(x,T) = u_T(x) & \text{on } \mathbb{T}^N. \end{cases} \tag{1}$$

where $\int_{\mathbb{T}^N} m_0\, dx = 1$, $m_0 > 0$, $m_0 \in C^1(\mathbb{T}^N)$ and $u_T \in C^2(\mathbb{T}^N)$. In the system above, the first is the Hamilton-Jacobi-Bellman equation for the value function u of a single agent, the second is the Kolmogorov-Fokker-Planck equation that governs the evolution of the distribution of the population m.

Even more than in the competitive case, the assumptions on the Hamiltonian H, the growth of the coupling f and the dimension of the state space will affect the qualitative behavior of the system. Let us clearly state the assumptions we make throughout the article.

The Hamiltonian $H : \mathbb{R}^N \to \mathbb{R}$ is convex, $C^1(\mathbb{R}^N)$, and has super-linear growth: there exist $\gamma > 1$, $C_H > 0$ such that

$$C_H^{-1}|p|^\gamma \le H(p) \le C_H(|p|^\gamma + 1), \tag{2}$$

for all $p \in \mathbb{R}^N$. Its Legendre transform, $L(q) := \sup_{p\in\mathbb{R}^N}\{p \cdot q - H(p)\}$ satisfies for some $C_L > 0$,

$$C_L^{-1}|q|^{\gamma'} - C_L \le L(q) \le C_L(|q|^{\gamma'} + 1) \tag{3}$$

for all $q \in \mathbb{R}^N$, where γ' is the conjugate exponent of γ, i.e. $\frac{1}{\gamma} + \frac{1}{\gamma'} = 1$. Note that L is strictly convex.

The local coupling $f : \mathbb{T}^N \times [0, +\infty) \to \mathbb{R}$, $f \ge 0$ is continuous in both variables, differentiable w.r.t. the second variable and satisfies, for $\alpha > 0$,

$$|\partial_m f(x,m)| \le c_f(m+1)^{\alpha-1} \tag{4}$$

for all $(x,m) \in \mathbb{T}^N \times [0, +\infty)$. Note that (4) implies

$$0 \le f(x,m) \le \frac{c_f}{\alpha}(m+1)^\alpha - \frac{c_f}{\alpha} + f(x,0) \le C_f(m^\alpha + 1),$$

for all $(x,m) \in \mathbb{T}^N \times [0, +\infty)$, but in general f is not bounded by above. Actually, what is important here is to have a control on the behaviour of the coupling at infinity, rather than requiring a restriction on the monotonicity of f, that need not be increasing with respect to m. Note that f could also depend explicitly on time, without giving any additional difficulty. We also mention that an additional regular dependence of H with respect to x could be easily added, provided it preserves the growth requirement (2) and convexity (see in particular (H2) in [8]). Let

$$F(x,m) = \int_0^m f(x,\sigma)d\sigma \quad \forall (x,m) \in \mathbb{T}^N \times [0,+\infty), \quad F(x,m) = 0 \quad \text{otherwise},$$

we then have, for all $m \ge 0$ and a $C_F > 0$,

$$0 \le F(x,m) \le C_F(m^{\alpha+1} + 1). \tag{5}$$

Before stating our results, let us have a look at the ones obtained in [10] for the stationary case. In this setting, scaling properties of the system and regularity of the Kolmogorov equation can be exploited to prove that: if $\alpha < \gamma'/N$, there exists a classical solution basically by means of a control of the "energy" associated to the system. In other words, the decaying of the coupling is well compensated by the regularising properties of the diffusion. If $\gamma'/N \le \alpha < \gamma'/(N-\gamma')$, such a control turns out to be more delicate, as aggregation may become the leading effect. Therefore, existence of classical solutions can be obtained only under additional assumptions on the coupling. In both cases, existence is shown by a blow-up method, together with Schauder's fixed point theorem. If $\alpha > \gamma'/(N-\gamma')$, it is shown that classical solutions may not exist, i.e. concentration due to the fast decay of the coupling cannot be compensated by the diffusion.

In the evolutionary case, we expect a similar behaviour. Indeed, one of the tools that turns out to be fundamental for understanding a MFG system of the form (1) is its variational formulation: let $\mathcal{K}_{\gamma',\alpha} \subset L^{\alpha+1}(Q) \times L^1(Q)$ be the pairs (m, w) satisfying $|w|^{\gamma'} m^{1-\gamma'} \in L^1(Q)$, $m \ge 0$ and

$$\int_Q (m\varphi_t + w \cdot \nabla\varphi - \nabla m \cdot \nabla\varphi)dxdt + \int_{\mathbb{T}^N} m_0(x)\varphi(x,0)dx = 0 \quad \forall\varphi \in C_0^\infty(\mathbb{T}^N \times [0,T)). \tag{6}$$

then the energy functional $\mathcal{E}$ associated with (1) is defined on $\mathcal{K}_{\gamma',\alpha}$ as[1]

$$\mathcal{E}(m,w) := \int_Q mL\left(-\frac{w}{m}\right) - F(x,m)\,dxdt + \int_{\mathbb{T}^N} u_T(x)m(x,T)\,dx. \tag{7}$$

The functional $\mathcal{E}$ shares many features with its stationary analogue, associated with the stationary problems considered in [10]. On the other hand, some useful scaling properties are missing in the parabolic case by the presence of time. Note that, due to the behaviour of f at infinity, $\mathcal{E}$ is not convex and it may not be bounded from below in general. Hence, the usual variational methods that link minimizers of this functional to solutions of the MFG system cannot be used in general for this case. However, if we restrict to the regime $\alpha < \gamma'/N$, the energy $\mathcal{E}$ becomes bounded from below; this key fact is proven in Lemma 3.2. Let us comment upon this last assumption. In the stationary case, using the rescaling properties of the Kolmogorov equation (that acts as a constraint), it can be shown that it is a necessary and sufficient condition for the stationary version of the energy functional $\mathcal{E}$ to be bounded from below. In particular, if $\alpha \ge \gamma'/N$, by scaling a test competitor one observes that the term $-\int F(x,m)$ prevails on $\int mL(-w/m)$, making the energy unbounded. In the evolutionary case, however, such a procedure does not apply directly, because any rescaling in the space variable changes the initial datum and jeopardises the constraint $\mathcal{K}$. Nevertheless, this hypothesis is crucial in Proposition 2.5, where we prove estimates for a superlinear power of the term

$$\int_Q m^{\alpha+1}\,dxdt.$$

in terms of $\int mL(-w/m)$. Such an estimate is based on the Gagliardo-Nirenberg inequality and parabolic regularity applied to the Fokker-Planck equation.

In what follows we will always suppose

$$\alpha < \frac{\gamma'}{N},$$

so that $\mathcal{E}$ is bounded by below on $\mathcal{K}$: for this reason, we will call this regime *sub-critical*. Under this assumption, it can be shown through a convexification procedure that the energy functional $\mathcal{E}$ possess a minimum that can be linked to a weak solution (in the sense of Definition 3.1) of the MFG system. The idea is that variational techniques similar to the ones presented by the second

[1] The term $mL(-w/m)$ has to be intended equal to zero if $(m,w) = (0,0)$ and $+\infty$ if $w \ne 0$ and $m \le 0$.

author et al. in [8], can be applied to prove the existence of a weak solution to the convexified MFG system. Then it is enough to prove that the solution of the convexified problem provides a solution of the original one. Moreover, to have convergence of minimizing sequences in $L^{\alpha+1}$ we also need (see Lemma 3.3)

$$\alpha < \min\left\{\frac{\gamma'}{N}, \frac{\gamma'}{N+2-\gamma'}\right\} = \begin{cases} \frac{\gamma'}{N} & \text{when } \gamma' \geq 2 \\ \frac{\gamma'}{N+2-\gamma'} & \text{when } 1 < \gamma' \leq 2 \end{cases}. \tag{8}$$

Note that existence of minimizers in the superquadratic case, i.e. $\gamma > 2$, is more delicate, and requires additional restrictions on α.

When $\gamma' > N+2$, thanks to Corollary 2.7, more regularity can be proven for weak solutions, that are indeed classical ones.

When $2 < \gamma' \leq N+2$, we are able to prove the existence of classical solutions only under the additional hypothesis

$$\alpha < \min\left\{\frac{\gamma'}{N}, \frac{\gamma'-2}{N+2-\gamma'}\right\}.$$

In this case a penalisation argument allows to obtain a sequence of solutions applying Theorem 1.1. Then a series of a priori estimates will show the convergence of the sequence to a classical solution of the original problem. These estimates rely on some blow-up techniques for which the hypothesis on α and the requirement $\gamma' > 2$ (i.e. H is subquadratic) are necessary.

The main theorems are then the following. See Remark 2.1 and Definition 3.1 for the expected regularity of (u, m).

Theorem 1.1. *Suppose that* (2), (4) *hold,* $\gamma' > N+2$ *and*

$$\alpha < \frac{\gamma'}{N}.$$

Then, there exists a classical solution (u, m) *of* (1) *such that* $(m, -m\nabla H(\nabla u))$ *is a minimizer of* $\mathcal{E}$ *in* $\mathcal{K}_{\gamma',\alpha}$.

Theorem 1.2. *Suppose that* (2), (4) *hold,* $1 < \gamma' \leq N+2$ *and*

$$\alpha < \min\left\{\frac{\gamma'}{N}, \frac{\gamma'}{N+2-\gamma'}\right\}.$$

Then, there exists a weak solution (u, m) *of* (1) *such that* $(m, -m\nabla H(\nabla u))$ *is a minimizer of* $\mathcal{E}$ *in* $\mathcal{K}_{\gamma',\alpha}$.

Theorem 1.3. *Suppose that* (2), (4) *hold,* $2 < \gamma' \leq N+2$ *and*

$$\alpha < \min\left\{\frac{\gamma'}{N}, \frac{\gamma'-2}{N+2-\gamma'}\right\}.$$

Then, there exists a classical solution (u, m) *of* (1) *such that* $(m, -m\nabla H(\nabla u))$ *is a minimizer of* $\mathcal{E}$ *in* $\mathcal{K}_{N+3,\alpha}$.

A convexification argument has already been used by Briani and Cardaliaguet in [7] to prove the existence of classical solutions for the case of energy functionals where the coupling is a regularising function of m. In their setting, the energy is not necessarily convex, but it is bounded by below. Their result is part of a more general analysis on stability of solutions in MFG having multiple equilibria.

Up to our knowledge, the only examples of non-uniqueness of solutions available in the literature for the evolutionary problem are the ones presented in Briani and Cardaliaguet [7], Bardi and Fischer [4], by the first author in [11], and in the recorded courses at Collège de France by P.-L. Lions [22]. Here, we present a new example (see in particular Section 4) that involves a class of MFG systems where $f(x, m) = f(m)$, $u_T = 0$ and $m_0 = 1$. We show that the solution corresponding to a minimum of $\mathcal{E}$ is not equal to the trivial solution $(\bar{u}, \bar{m}) = ((t-T)(f(1)+H(0)), 1)$. This follows by construction of a suitable competitor (m, w) for which $\mathcal{E}(m, w) < \mathcal{E}(1, 0)$. Note that non-uniqueness is obtained in [7, 11] in cases where $H \in C^2$ and T large enough (if $H \in C^2$ uniqueness for small time horizon T holds, see [3, 4]), while in [4] an example with Lipschitz Hamiltonian and $T > 0$ arbitrary is discussed. In our examples, Hamiltonians can be $C^{1,\beta}$, with $0 < \beta < 1$, and non-uniqueness of solutions shows up for all $T > 0$.

Finally, one should expect similar features for the stationary and non-stationary problems when the time horizon T is large. Conversely, if T is small, the two settings may exhibit different behaviours. As mentioned before, existence of solutions might be false in the stationary case when the coupling is very strong. On the other hand, in the parabolic case a standard contraction argument applies: we prove the existence of a classical solution of (1) for small T, without requiring *any* assumption on the growth at infinity of f.

Theorem 1.4. *Suppose that $f, H \in C^3$, $u_T \in C^{2,\beta}$ and $m_0 \in C^{2,\nu}$ for some $0 < \nu < \beta < 1$. Then, there exists $T^* > 0$ such that for all $T \in (0, T^*]$, (1) has a classical solution.*

The idea is to exploit the local regularity of f and H (note that in Theorem 1.4 f need not be with controlled growth nor H to be convex); even if the coupling is very strong, during a small time interval the distribution m remains close to the initial datum m_0, without developing singularities. Note that the contraction mapping principle ensures that solutions are unique in a neighborhood of (u_T, m_0), see in particular Remark A.1.

We mention that the contraction theorem has already been used by Ambrose [1, 2] in the MFG setting, to prove the existence of small, locally unique, strong solutions over any finite time interval in the case of general local couplings, if m_0 is chosen sufficiently close to a uniform density or if other "smallness" conditions are satisfied.

The paper is organized as follows: we first collect embedding theorems and estimates for the Hamilton-Jacobi equation and Fokker-Planck equation. Section 3 is devoted to the proof of Theorems 1.1, 1.2 and 1.3. A non-uniqueness example is shown in Section 4. Finally, the appendix contains the proof of the existence of classical solutions for small T.

Acknowledgements. The first author is partially supported by the Fondazione CaRiPaRo Project "Nonlinear Partial Differential Equations: Asymptotic Problems and Mean-Field Games" and the INdAM-GNAMPA project "Fenomeni di segregazione in sistemi stazionari di tipo Mean Field Games a più popolazioni". The second author is partially supported by the ANR project MFG ANR-16-CE40-0015-01, the PEPS-INSMI Jeunes project "SOME OPEN PROBLEMS IN MEAN FIELD GAMES" for the years 2016 and 2017. This article benefited from the support of the FMJH Program Gaspard Monge for optimization and operations research and their interactions with data science and by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH, PGMO project VarPDEMFG. We warmly thank the anonymous referee for his/her careful reading and valuable comments.

2 Notations and preliminaries

In the sequel, we will use the notation $\mathcal{K} = \mathcal{K}_{\gamma',\alpha}$, and write explicitly the subscript if does not coincide with γ', α. $\mathcal{P}(\mathbb{T}^N)$ will denote the space of probability measures, endowed with the standard weak* topology.

We begin with the definition of some particular Banach spaces involving time and space weak derivatives. Let $Q = Q_T = \mathbb{T}^N \times (0,T)$ and $r > 1$. We denote by $W_r^{2,1}(Q)$ the space of functions $u \in L^r(Q)$ having weak space derivatives $D_x^\beta u \in L^r(Q)$ for any integer multi-index β such that $|\beta| \le 2$, and weak time derivative $u_t \in L^r(Q)$, equipped with the norm

$$\|u\|_{W_r^{2,1}(Q)} = \|u\|_{L^r(Q)} + \|u_t\|_{L^r(Q)} + \sum_{1\le|\beta|\le 2} \|D_x^\beta u\|_{L^r(Q)}.$$

Similarly, $W_r^{1,0}(Q)$ is endowed with the norm

$$\|u\|_{W_r^{1,0}(Q)} = \|u\|_{L^r(Q)} + \sum_{|\beta|=1} \|D_x^\beta u\|_{L^r(Q)}.$$

Finally, we denote by $\mathcal{H}^{r,1}(Q)$ the space of functions $u \in W_r^{1,0}(Q)$ with $u_t \in (W_{r'}^{1,0}(Q))'$, equipped with the norm

$$\|u\|_{\mathcal{H}^{r,1}(Q)} = \|u\|_{W_r^{1,0}(Q)} + \|u_t\|_{(W_{r'}^{1,0}(Q))'}.$$

For a given Banach space X, $L^p((0,T), X)$ and $C^{0,\theta}([0,T], X)$ will denote the usual Lebesgue and Hölder parabolic spaces respectively.

Remark 2.1. Under the standing assumptions on the Hamiltonian (2), in particular if $H \in C^1(\mathbb{R}^N)$, we will say that (u,m) is classical solution to (1) in the following sense (with a slight abuse of terminology): first, u is (at least) C^2 in the x-variable and C^1 in the t-variable, and solves the HJB equation in the classical sense. Second, m is a weak solution of the Fokker-Planck equation. Since $\nabla H(\nabla u)$ is just bounded a priori, one expects m to be in $L^2((0,T), H^1(\mathbb{T}^N)) \cap L^\infty((0,T)\times\mathbb{T}^N)$ and Hölder continuous, while $m_t \in L^2((0,T), H^{-1}(\mathbb{T}^N))$ (see, e.g. [17, Chapter 3]). Note that scarce regularity of ∇H is the only obstruction for m to be a classical solution: if $H \in C^{2,a}$, $a \in (0,1)$, by Schauder estimates m is indeed classical (for example, as in Theorem 1.4). In this way, we can deal with model power-like Hamiltonians of the form

$$H(p) = C_H |p|^\gamma, \qquad \gamma > 1.$$

We recall here some embedding properties enjoyed by $\mathcal{H}^{r,1}(Q)$.

Proposition 2.2. *The following embeddings hold.*

(i) If $1 < r < N+2$, then $\mathcal{H}^{r,1}(Q)$ is continuously embedded in $L^\ell(Q)$, for all $1 \le \ell \le \frac{(N+2)r}{N+2-r}$.

(ii) If $r \ge N+2$, then $\mathcal{H}^{r,1}(Q)$ is continuously embedded in $L^\ell(Q)$, for all $1 \le \ell < \infty$.

(iii) If $r > N+2$, then there exist $\nu, \theta \in (0,1)$ such that $\mathcal{H}^{r,1}(Q)$ is continuously embedded in $C^{0,\nu}([0,T], C^{0,\theta}(\mathbb{T}^N))$.

(iv) If $1 < r < N+2$, then $\mathcal{H}^{r,1}(Q)$ is compactly embedded in $L^\ell(Q)$, for all $1 \le \ell < \frac{(N+2)r}{N+2-r}$.

Proof. A proof of the continuous embeddings *(i)* and *(ii)* can be found in [24, Theorem 7.1] (see also [6, Theorem 6.2.2 (i), and refs. therein]), in the case $Q = \mathbb{R}^N \times (0,T)$, that is, for all $\tilde{u} \in \mathcal{H}^{r,1}(\mathbb{R}^N \times (0,T))$ it holds

$$\|\tilde{u}\|_{L^\ell(\mathbb{R}^N\times(0,T))} \le C\|\tilde{u}\|_{\mathcal{H}^{r,1}(\mathbb{R}^N\times(0,T))} \tag{9}$$

for some $C > 0$ not depending on $\tilde{u}$. To derive the same embeddings for $Q = \mathbb{T}^N \times (0,T)$, it is sufficient to construct by standard methods a linear operator extending continuously $\mathcal{H}^{r,1}(\mathbb{T}^N \times (0,T))$ to $\mathcal{H}^{r,1}(\mathbb{R}^N \times (0,T))$. In particular, fix any $\chi \in C_0^\infty(\mathbb{R}^N)$ such that $\chi \equiv 1$ on $[0,1]^N$. Then, any $u \in \mathcal{H}^{r,1}(\mathbb{T}^N \times (0,T))$ can be extended naturally to a function on $\mathbb{R}^N \times (0,T)$ that is periodic in the x variable, i.e. $u(x,t) = u(x+k,t)$ for all $k \in \mathbb{Z}^N$ and a.e. $(x,t) \in \mathbb{R}^N \times (0,T)$. Let $\tilde{u}(x,t) = \chi(x)u(x,t)$, then one has

$$\|\tilde{u}\|_{\mathcal{H}^{r,1}(\mathbb{R}^N \times (0,T))} \le C_1 \|u\|_{\mathcal{H}^{r,1}(\mathbb{T}^N \times (0,T))} \tag{10}$$

for some positive constant C_1 that depends only on the choice of χ. It is then sufficient to combine (9), (10) and the fact that $u \equiv \tilde{u}$ on $[0,1]^N$ to conclude *(i)* and *(ii)*.

As for *(iii)*, see [6, Theorem 6.2.2 (ii)].

Finally, the proof of *(iv)* of relies on the so-called Aubin-Lions-Simon Lemma. Let $1 < r < N+2$. Note first that $W^{1,r'}(\mathbb{T}^N)$ is reflexive and separable. Therefore, $L^r((0,T),(W^{1,r'}(\mathbb{T}^N)')$ is isomorphic to $(L^{r'}((0,T),W^{1,r'}(\mathbb{T}^N)))'$, and the latter space coincides with $(W^{1,0}_{r'}(Q))'$. Since $W^{1,0}_r(Q)$ coincides with $L^r((0,T),W^{1,r}(\mathbb{T}^N))$, we have that the space $\mathcal{H}^{r,1}(Q)$ is isomorphic to

$$E = \{u \in L^r((0,T),W^{1,r}(\mathbb{T}^N)),\, u_t \in L^r((0,T),(W^{1,r'}(\mathbb{T}^N)')\}.$$

As $W^{1,r}(\mathbb{T}^N)$ is compactly embedded in $L^r(\mathbb{T}^N)$, and $L^r(\mathbb{T}^N)$ is continuously embedded in $(W^{1,r'}(\mathbb{T}^N))'$, the Aubin-Lions-Simon lemma (see in particular [27, Corollary 5]) states that E is compactly embedded in $L^r(Q)$. Hence, $\mathcal{H}^{r,1}(Q)$ is compactly embedded in $L^r(Q)$ and the result follows for $1 \le \ell \le r$.

Let now u_n be a bounded sequence in $\mathcal{H}^{r,1}(Q)$; we may extract a subsequence u_{n_k} that converges to u strongly in $L^r(Q)$. For any $r < \ell < \frac{(N+2)r}{N+2-r}$, by interpolation, there exists $0 < \theta < 1$ such that

$$\|u_{n_k} - u_{n_j}\|_{L^\ell(Q)} \le \|u_{n_k} - u_{n_j}\|^\theta_{L^r(Q)} \|u_{n_k} - u_{n_j}\|^{1-\theta}_{L^{\frac{(N+2)r}{N+2-r}}(Q)} \to 0,$$

as $j,k \to \infty$, being u_{n_k} bounded in $\mathcal{H}^{r,1}(Q)$ and hence in $L^{\frac{(N+2)r}{N+2-r}}(Q)$ by *(i)*. Therefore, u_{n_k} converges strongly in $L^\ell(Q)$ also. □

The following is a classical Hölder regularity result for Hamilton-Jacobi equations with quadratic or sub-quadratic Hamiltonians.

Proposition 2.3. *Let Ω be a bounded domain of $\mathbb{R}^N$ and $\widetilde{H} \in C(\mathbb{R}^N)$. Suppose that v is a classical solution to*

$$-v_t - \Delta v + \widetilde{H}(\nabla v) = W(x,t) \quad \text{in } \widetilde{Q} = \Omega \times (0,\tau),\ \tau > 0$$

and that for some $K > 0$,

$$\begin{aligned} &i) \quad \|v\|_{L^\infty(\widetilde{Q})}, \|W\|_{L^\infty(\widetilde{Q})}, \|v(\cdot,\tau)\|_{C^2(\Omega)} \le K, \\ &ii) \quad |\widetilde{H}(p)| \le K(|p|^2+1) \text{ for all } p \in \mathbb{R}^N. \end{aligned}$$

Then, for all $\Omega' \subset\subset \Omega$, there exists $C > 0$, $\beta > 0$ (depending on K, Ω', but not on τ), such that

$$\|\nabla v\|_{C^{0,\beta}(\Omega' \times (0,\tau))} \le C.$$

Proof. See [17, Theorem V.3.1]. □

2.1 Regularity of the Fokker-Planck equation

In what follows, let $m \in L^1(Q)$ and A be a measurable vector field such that m is a weak solution of the Fokker-Planck equation

$$\begin{cases} m_t - \Delta m + \mathrm{div}(A\, m) = 0 & \text{in } Q, \\ m(x,0) = m_0(x) & \text{on } \mathbb{T}^N. \end{cases} \tag{11}$$

Suppose $|A|^{\gamma'} m \in L^1(Q)$, for a $\gamma' > 1$, and set E to be the quantity

$$E := \int_Q |A|^{\gamma'} m \, dxdt.$$

In this section we will state some regularity results and a priori estimates for m, depending on E, that will be used in the sequel. We stress that our aim here is to obtain regularity of m in terms of $|A|m^{1/\gamma'}$, rather than in terms of $|A|$ itself. The former quantity is indeed associated to the energy of the system. The exponent γ' and α used in this section are not necessarily linked with the ones given by the hypotheses on H and f (unless otherwise specified).

Note that, by setting $w := Am$ on the set $\{m > 0\}$, and $w \equiv 0$ where m vanishes, the couple (m, w) solves

$$\begin{cases} m_t - \Delta m + \mathrm{div}(w) = 0 & \text{in } Q, \\ m(x,0) = m_0(x) & \text{on } \mathbb{T}^N, \end{cases}$$

and E can be rewritten as $E = \int_Q |w/m|^{\gamma'} m \, dxdt$. Recall that $E < +\infty$, since we assumed $|A|^{\gamma'} m \in L^1(Q)$. The proof of the following result follows the lines of [24, Proposition 3.2].

Proposition 2.4. *Let $m \in L^\ell(Q)$, for some $\ell > 1$, be a weak solution of (11), and r be such that*

$$\frac{1}{r} := \frac{1}{\gamma'} + \left(1 - \frac{1}{\gamma'}\right) \frac{1}{\ell}. \tag{12}$$

Then, there exists $C > 0$, depending on T, ℓ, N, γ' and $\|\nabla m_0\|_{L^r(\mathbb{T}^N)}$, such that

$$\|m\|_{\mathcal{H}^{r,1}(Q)} \leq C(E^{1/\gamma'} \|m\|_{L^\ell(Q)}^{1/\gamma} + 1), \tag{13}$$

where γ is the conjugate exponent of γ'.

Proof. We first assume that m is a smooth function; we will remove this requirement at the end of the proof. Let $\varphi \in C^{2,1}(\overline{Q})$ be a test function such that $\varphi(\cdot, T) = 0$. Then, being m a weak solution,

$$\int_Q m(-\varphi_t - \Delta\varphi - A \cdot \nabla\varphi) \, dxdt = \int_{\mathbb{T}^N} m_0(x)\varphi(x,0)dx.$$

Hence,

$$\left| \int_Q m(-\varphi_t - \Delta\varphi) \, dxdt \right| \leq \left| \int_{\mathbb{T}^N} m_0(x)\varphi(x,0)dx \right| + \int_Q |A| m^{1/\gamma'} m^{1-1/\gamma'} |\nabla\varphi| \, dxdt,$$

and, applying twice Hölder inequality in the last term, we have

$$\left| \int_Q m(-\varphi_t - \Delta\varphi) \, dxdt \right| \leq \left| \int_{\mathbb{T}^N} m_0(x)\varphi(x,0)dx \right| + E^{1/\gamma'} \|m\|_{L^\ell(Q)}^{1/\gamma} \|\nabla\varphi\|_{L^{r'}(Q)}, \tag{14}$$

where, r is as in (12).

Let now $i = 1, \dots, N$ be fixed, and $\psi \in C^{2,1}(Q)$ be the solution of

$$\begin{cases} -\psi_t - \Delta\psi = |\partial_{x_i} m|^{r-2}\partial_{x_i} m & \text{in } Q, \\ \psi(x,T) = 0 & \text{on } \mathbb{T}^N. \end{cases} \tag{15}$$

Note that, by standard parabolic regularity (see, e. g., [17, Theorem IV.9.1]),

$$\|\psi\|_{W^{2,1}_{r'}(Q)} \le C\| \, |\partial_{x_i} m|^{r-1} \, \|_{L^{r'}(Q)} = C\|\partial_{x_i} m\|^{r-1}_{L^r(Q)}, \tag{16}$$

where r' is the conjugate exponent of r. Moreover, $\psi(x,0) = -\int_0^T \psi_t(x,s)\, ds$ on $\mathbb{T}^N$, therefore, using Hölder inequality,

$$\begin{aligned} \left| \int_{\mathbb{T}^N} \partial_{x_i} m_0(x)\psi(x,0)\, dx \right| &\le \int_Q |\partial_{x_i} m_0(x)\psi_t(x,s)|\, dxds \le T^{1/r}\|\partial_{x_i} m_0\|_{L^r(\mathbb{T}^N)}\|\psi_t\|_{L^{r'}(Q)} \\ &\le C\|\psi\|_{W^{2,1}_{r'}(Q)}. \end{aligned} \tag{17}$$

If we now let $\varphi = \partial_{x_i}\psi$ in (14), integrating by parts, we obtain

$$\left| \int_Q \partial_{x_i} m(-\psi_t - \Delta\psi)\, dxdt \right| \le \left| \int_{\mathbb{T}^N} \partial_{x_i} m_0(x)\psi(x,0) dx \right| + E^{1/\gamma'}\|m\|^{1/\gamma}_{L^\ell(Q)}\|\nabla(\partial_{x_i}\psi)\|_{L^{r'}(Q)},$$

and by (17) and the fact that ψ solves (15),

$$\int_Q |\partial_{x_i} m|^r\, dxdt = \left| \int_Q \partial_{x_i} m(-\psi_t - \Delta\psi)\, dxdt \right| \le C\|\psi\|_{W^{2,1}_{r'}(Q)}(1 + E^{1/\gamma'}\|m\|^{1/\gamma}_{L^\ell(Q)}).$$

Hence, using again (16),

$$\|\nabla m\|^r_{L^r(Q)} = \int_Q |\nabla m|^r\, dxdt \le C(E^{1/\gamma'}\|m\|^{1/\gamma}_{L^\ell(Q)} + 1)^r. \tag{18}$$

By Poincaré inequality and (16), since $\frac{1}{|Q|}\int_Q m\, dxdt = 1$, we infer

$$\|m\|_{L^r(Q)} \le \|m - 1\|_{L^r(Q)} + T \le C(E^{1/\gamma'}\|m\|^{1/\gamma}_{L^\ell(Q)} + 1). \tag{19}$$

Finally, we multiply the Fokker-Planck equation by any test function (that may not vanish at time T) $\varphi \in C^{2,1}(\overline{Q})$ to get

$$\begin{aligned} \left| \int_Q m_t\varphi\, dxdt \right| &\le \left| \int_Q \nabla m \cdot \nabla\varphi\, dxdt \right| + E^{1/\gamma'}\|m\|^{1/\gamma}_{L^\ell(Q)}\|\nabla\varphi\|_{L^{r'}(Q)} \\ &\le (\|\nabla m\|_{L^r(Q)} + E^{1/\gamma'}\|m\|^{1/\gamma}_{L^\ell(Q)})\|\nabla\varphi\|_{L^{r'}(Q)}, \end{aligned}$$

where Hölder inequality has been used for the integral of $A \cdot \nabla\varphi m$ as in (14). Hence

$$\|m_t\|_{(W^{1,r'}(Q))'} \le C(E^{1/\gamma'}\|m\|^{1/\gamma}_{L^\ell(Q)} + 1),$$

and by (18) and (19) we conclude.

If m is not smooth, consider a regularised sequence $(m_n, w_n) := (m \star \xi_n, w \star \xi_n)$, where ξ_n is a (space-time) smoothing kernel. Then, $(m_n)_t - \Delta m_n + \mathrm{div}(w_n) = 0$ holds, and (13) is verified with $E_n = \int_Q |w_n/m_n|^{\gamma'} m_n\, dxdt$. Since $E_n \to E$ as $n \to \infty$, see [9], eq. (2.19) in Lemma 2.7, we obtain (13) in the general case (actually, it is enough to invoke weak lower semicontinuity of the functional $\int_Q |w/m|^{\gamma'} m\, dx$, ensured by convexity, to conclude).

□

The following is a crucial estimate that links the energy term $\left(\int_Q m^{\alpha+1}\,dxdt\right)^\delta$, for a $\delta > 1$, with the quantity E. The assumption $\alpha < \gamma'/N$ basically allows to apply the Gagliardo-Nirenberg and interpolation inequalities.

Proposition 2.5. *Let $m \in L^{\alpha+1}(Q)$ for some $0 \le \alpha < \gamma'/N$, be a weak solution of* (11)*. Then, there exist $C > 0$ and $\delta > 1$, depending on T, α, N, γ and $\|\nabla m_0\|_{L^{\frac{\gamma'(\alpha+1)}{\gamma'+\alpha}}(\mathbb{T}^N)}$, such that*

$$\left(\int_Q m^{\alpha+1}\,dxdt\right)^\delta \le C\,(E+1)\,. \tag{20}$$

Proof. Since $\|m\|_{L^1(Q)} = T$, if $\alpha = 0$, we are done. Suppose then $\alpha > 0$. Proposition 2.4 applies with $\ell = \alpha + 1$, so $m \in \mathcal{H}^{r,1}(Q)$, where $r = \gamma'(\alpha+1)/(\gamma'+\alpha)$. Hence, $m(\cdot,t) \in W^{1,r}(\mathbb{T}^N)$ for a. e. $t \in (0,T)$. Let $\eta := r(N+1)/N$. By the Gagliardo-Nirenberg inequality (see in particular [25, p. 125-126]), there exists $C > 0$ such that for a.e. t,

$$\|m(\cdot,t)\|_{L^\eta(\mathbb{T}^N)} \le C\left(\|\nabla m(\cdot,t)\|_{L^r(\mathbb{T}^N)}^{\frac{r}{\eta}}\|m(\cdot,t)\|_{L^1(\mathbb{T}^N)}^{\frac{1}{N+1}} + \|m(\cdot,t)\|_{L^1(\mathbb{T}^N)}\right),$$

thus, by the fact that $\int_{\mathbb{T}^N} m(\cdot,t)dx = 1$ for a. e. $t \in (0,T)$,

$$\int_{\mathbb{T}^N} m^\eta(x,t)\,dx \le C\left(\int_{\mathbb{T}^N} |\nabla m(x,t)|^r\,dx + 1\right).$$

If we now integrate the previous equality over $(0,T)$, in view of (13), it follows that

$$\|m\|_{L^\eta(Q)}^{(N+1)/N} \le C(\|m\|_{L^{\alpha+1}(Q)}^{1/\gamma} E^{1/\gamma'} + 1).$$

It is now crucial to observe that $\alpha < \gamma'/N$ implies $1 < \alpha+1 < \eta$. Indeed,

$$\eta = r\frac{N+1}{N} = (\alpha+1)\frac{\gamma' N + \gamma'}{\gamma' N + \alpha N} > \alpha + 1.$$

Arguing by interpolation, there exists some $0 < \theta < 1$ such that $\|m\|_{L^{\alpha+1}(Q)} \le C\|m\|_{L^\eta(Q)}^\theta$, so

$$\|m\|_{L^{\alpha+1}(Q)}^{(N+1)/(\theta N)} \le C(\|m\|_{L^{\alpha+1}(Q)}^{1/\gamma} E^{1/\gamma'} + 1),$$

which implies

$$\|m\|_{L^{\alpha+1}(Q)}^{1+\gamma'(N+1)/(\theta N)-\gamma'} \le C(E+1).$$

The inequality (20) then follows setting $\delta := \frac{1+\gamma'(N+1)/(\theta N)-\gamma'}{\alpha+1}$. Note that $\delta > 1$ since $\gamma'(N+1)/(\theta N) - \gamma' \ge \gamma'/N > \alpha$.

□

We next state a proposition that gives a bound on the $\mathcal{H}^{r,1}(Q)$ norm of the solution of the Fokker-Planck equation in terms of E, for certain values of r that depends on γ' and N.

Proposition 2.6. *Suppose that $E \le K$ and $\|m\|_{L^{\ell_0}(Q)} \le K$, for some $K > 0$,*

$$\begin{cases} \ell_0 \in \left(1, \frac{N+2}{N+2-\gamma'}\right) & \textit{if } \gamma' < N+2, \\ \ell_0 \in (1, +\infty) & \textit{if } \gamma' \ge N+2. \end{cases}$$

Let r be such that

$$\begin{cases} r \in \left(1, \frac{N+2}{N+3-\gamma'}\right) & \text{if } \gamma' < N+2, \\ r \in (1, \gamma') & \text{if } \gamma' \geq N+2. \end{cases} \tag{21}$$

Then, there exists $C > 0$ depending on K, r, T, N, γ' such that

$$\|m\|_{\mathcal{H}^{r,1}(Q)} \leq C.$$

Proof. Let the sequences ℓ_n, r_n be defined by induction as follows: for a given ℓ_n, let r_n be such that

$$\frac{1}{r_n} := \frac{1}{\gamma'} + \left(1 - \frac{1}{\gamma'}\right) \frac{1}{\ell_n} > \frac{1}{\gamma'}.$$

Moreover, $\ell_{n+1} := \frac{(N+2)r_n}{N+2-r_n}$, i.e.

$$\frac{1}{\ell_{n+1}} = \frac{1}{r_n} - \frac{1}{N+2} = \frac{1}{\gamma'} - \frac{1}{N+2} + \left(1 - \frac{1}{\gamma'}\right) \frac{1}{\ell_n}.$$

We use here the convention $\frac{N+2-\gamma'}{(N+2)} = 0$, when $\gamma' \geq N+2$. Since $1 < \ell_0 < \frac{N+2}{N+2-\gamma'}$, r_n, ℓ_n are increasing sequences. Indeed, we have

$$\frac{\ell_n}{\ell_{n+1}} = \frac{N+2-\gamma'}{\gamma'(N+2)} \ell_n - \frac{1}{\gamma'} + 1 < 1$$

as soon as $\gamma' \geq N+2$, or as $\ell_n < \frac{N+2}{N+2-\gamma'}$.

In the case $\gamma' < N+2$, ℓ_n converges to $\frac{N+2}{N+2-\gamma'}$, while r_n converges to $\frac{N+2}{N+3-\gamma'}$. By Proposition 2.2 (i), (ii) and Proposition 2.4, we have that

$$\|m\|_{L^{\ell_{n+1}}(Q)} \leq C \|m\|_{\mathcal{H}^{r_n,1}(Q)} \leq C_1 (E^{1/\gamma'} \|m\|^{1/\gamma}_{L^{\ell_n}(Q)} + 1),$$

so we obtain the assertion by iterating the last inequality a finite number of times.

As for the case $\gamma' \geq N+2$, one argues in a similar way, with the difference that $\ell_n \to +\infty$ and $r_n \to \gamma'$.

□

Together with the embedding results of Proposition 2.2, Proposition 2.6 allows to prove the strong convergence of a (sub)sequence of weak solutions m_n, as shown in the following corollary.

Corollary 2.7. *Let (m_n, A_n) be a sequence solving the Fokker-Planck equation* (11) *in the weak sense. Suppose that*

$$E_n := \int_Q |A_n|^{\gamma'} m_n \, dxdt \leq K$$

and $\|m_n\|_{L^{\ell_0}(Q)} \leq K$ for some $K > 0$,

$$\begin{cases} \ell_0 \in \left(1, \frac{N+2}{N+2-\gamma'}\right) & \text{if } \gamma' < N+2, \\ \ell_0 \in (1, +\infty) & \text{if } \gamma' \geq N+2. \end{cases}$$

Then, up to subsequences, m_n converges:

- *strongly in $L^\ell(Q)$ for any $\ell \in [1, \frac{N+2}{N+2-\gamma'})$, if $\gamma' < N+2$,*

- *strongly in $L^\ell(Q)$ for any $\ell \in [1,+\infty)$, if $\gamma' \geq N+2$,*
- *in $C^{0,\theta}(Q)$ for some $\theta > 0$, if $\gamma' > N+2$.*

Proof. By Proposition 2.6, m_n is bounded in $\mathcal{H}^{r,1}(Q)$, for all r defined as in (21).

If $\gamma' < N+2$, for any $\ell \in [1, \frac{N+2}{N+2-\gamma'})$, we can find $\bar r \in \left(1, \frac{N+2}{N+3-\gamma'}\right)$, $\bar r < \frac{N+2}{N+3-\gamma'} < N+2$, such that $\ell \in [1, \frac{(N+2)\bar r}{N+2-\bar r})$ and m_n is bounded in $\mathcal{H}^{\bar r,1}(Q)$. Hence it is sufficient to apply Proposition 2.2 (iv) to obtain that m_n converges strongly (up to subsequences) in $L^\ell(Q)$.

If $\gamma' \geq N+2$ then for any $\ell \in [1,+\infty)$ we can find $\bar r \in (1,\gamma')$, $\bar r < N+2$, such that $\ell \in [1, \frac{(N+2)\bar r}{N+2-\bar r})$ and m_n is bounded in $\mathcal{H}^{\bar r,1}(Q)$. Therefore we can apply Proposition 2.2 (iv) to obtain that m_n converges strongly (up to subsequences) in $L^\ell(Q)$.

If $\gamma' > N+2$, we can find $\bar r \in (1,\gamma')$, $\bar r > N+2$, such that m_n is bounded in $\mathcal{H}^{\bar r,1}(Q)$. Hence by Proposition 2.2 (iii), m_n is bounded in some $C^{0,\nu}([0,T], C^{0,\theta'}(\mathbb{T}^N))$ (compactness follows by choosing $\theta < \min\{\theta', \nu\}$).

□

2.2 Known results on convex MFG

For the reader's convenience, we rephrase here some of the results from [8] that will be used in Section 3, in the setting of this paper. We consider a MFG system of the form

$$\begin{cases} -u_t - \Delta u + H(\nabla u) = \rho(x,t,m(x,t)), & \text{in } Q, \\ m_t - \Delta m - \operatorname{div}(\nabla H(\nabla u)\, m) = 0 & \text{in } Q, \\ m(x,0) = m_0(x), \quad u(x,T) = u_T(x) & \text{on } \mathbb{T}^N, \end{cases} \tag{22}$$

where the Hamiltonian, $m_0(\cdot)$ and $u_T(\cdot)$ satisfy the same hypotheses as before and the coupling $\rho : Q \times [0,+\infty) \to \mathbb{R}$ is continuous in all variables, strictly increasing with respect to the third variable m, and there exist $\alpha > 0$ and C_1 such that

$$\frac{1}{C_1} m^\alpha - C_1 \leq \rho(x,t,m) \leq C_1 m^\alpha + C_1 \qquad \forall m \geq 0\,. \tag{23}$$

As noted in [8], ρ can depend explicitly on time. Moreover, we ask the following normalization condition to hold:

$$\rho(x,t,0) = 0 \qquad \forall (x,t) \in Q\,. \tag{24}$$

Let α, γ be the exponents defined by the hypotheses on H and ρ, and define

$$\vartheta := \begin{cases} \frac{\gamma(\alpha+1)(1+N)}{\alpha N - \gamma} & \text{if } \alpha > \frac{\gamma}{N} \\ +\infty & \text{if } \alpha \leq \frac{\gamma}{N} \end{cases}, \quad \bar\vartheta := \begin{cases} \frac{N(\gamma+\alpha)}{\alpha N - \gamma} & \text{if } \alpha > \frac{\gamma}{N} \\ +\infty & \text{if } \alpha \leq \frac{\gamma}{N} \end{cases}.$$

The following theorem corresponds to Theorem 3.3 in [8].

Theorem 2.8. *Assume that* (2) *holds true and let u satisfy in the weak sense*

$$\begin{cases} (i) & -u_t - \Delta u + H(\nabla u) \leq \beta(x,t) \\ (ii) & u(x,T) \leq u_T(x), \end{cases}$$

with $\beta \in L^{(\alpha+1)'}(Q)$, $u_T \in L^\infty(\mathbb{T}^N)$. (Here $(\alpha+1)'$ is the conjugate exponent of $\alpha+1$). Then, if u is bounded below, we have

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

with a constant C depending on $T, \alpha, N, \gamma, C_H$ (appearing in (2)*) and on $\|\beta\|_{L^{(\alpha+1)'}(Q)}$, $\|u_T\|_{L^{\bar\vartheta}(\mathbb{T}^N)}$.*

This Lebesgue estimate holds typically for sub-solutions of Hamilton-Jacobi equations and is a consequence of the divergence structure of the second order term.

Thanks to the above regularity for u and estimate (4) on f, we look for weak solutions $(u,m) \in L^{\vartheta}(Q) \times L^{\alpha+1}(Q)$.

Definition 2.9. We say that a pair $(u,m) \in L^{\vartheta}(Q) \times L^{\alpha+1}(Q)$ is a weak solution to (22), if

(i) the following integrability conditions hold:

$$\nabla u \in L^{\gamma}(Q),\ mL(\nabla H(\nabla u)) \in L^1(Q) \quad \text{and} \quad m\nabla H(\nabla u) \in L^1(Q).$$

(ii) Equation (1)-(i) holds in the following sense: inequality

$$-u_t - \Delta u + H(\nabla u) \le \rho(x,t,m) \quad \text{in } Q, \tag{25}$$

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

(iii) Equation (1)-(ii) holds:

$$m_t - \Delta m - \operatorname{div}(m\nabla H(\nabla u))) = 0 \text{ in } Q, \quad m(0) = m_0 \tag{26}$$

in the sense of distributions,

(iv) The following equality holds:

$$\int_Q m(x,t)(\rho(x,t,m(x,t)) + L(\nabla H(\nabla u)(x,t)))dxdt \tag{27}$$
$$+ \int_{\mathbb{T}^N} (m(x,T)u_T(x) - m_0(x)u(x,0))dx = 0.$$

The following theorems correspond to Theorem 6.2 and to a part of Theorem 6.4 in [8].

Theorem 2.10. *There exists a unique weak solution (u,m) to the MFG system (22). By uniqueness, we mean that m is indeed unique and u is uniquely defined in $\{m > 0\}$. Finally, there exists a solution which is bounded below by a constant depending on $\|u_T\|_{C^2(\mathbb{T}^N)}$.*

Theorem 2.11. *Let (u,m) be the unique solution to the MFG system (22) and define $w := -m\nabla H(\nabla u)$. Then the couple (m,w) is a minimizer for*

$$\int_Q mL\left(-\frac{w}{m}\right) + \Psi(x,t,m)\,dxdt + \int_{\mathbb{T}^N} u_T(x)m(x,T)\,dx,$$

over the set $\mathcal{K}$, where

$$\Psi(x,t,m) = \int_0^m \rho(x,t,\sigma)d\sigma \quad \forall (x,m) \in \mathbb{T}^N \times [0,+\infty), \quad \Psi(x,t,m) = +\infty \quad \textit{otherwise.}$$

3 Existence of solutions

In this section we discuss the existence of solution for the MFG system (1). We begin giving the definition of weak solution.

Let α, γ be the exponents defined by the hypotheses on H and f, and $\alpha < \min\left\{\frac{\gamma'}{N}, \frac{\gamma'}{N+2-\gamma'}\right\}$. Let ϑ be defined as in Section 2.2:

$$\vartheta := \begin{cases} \frac{\gamma(\alpha+1)(1+N)}{\alpha N - \gamma} & \text{if } \alpha > \frac{\gamma}{N} \\ +\infty & \text{if } \alpha \le \frac{\gamma}{N} \end{cases}.$$

Definition 3.1. We say that a pair $(u, m) \in L^{\vartheta}(Q) \times L^{\alpha+1}(Q)$ is a weak solution to (1), if it satisfies Definition 2.9 with ρ replaced by the coupling $-f$.

Using the estimates obtained for the solution of the Fokker-Planck equation, we are able to prove that the energy functional $\mathcal{E}$, defined as in (7), is bounded from below over the set $\mathcal{K}$.

Lemma 3.2. *There exists $c \in \mathbb{R}$ such that*

$$c = \inf_{(m,w)\in\mathcal{K}} \mathcal{E}(m, w).$$

Moreover, suppose that for a sequence $(m^n, w^n) \in \mathcal{K}$, there exists $e \in \mathbb{R}$ such that $\mathcal{E}(m^n, w^n) \leq e$ for all $n \in \mathbb{N}$. Then, for some c_1 (depending on e), and for all $n \in \mathbb{N}$

$$\int_Q (m^n)^{\alpha+1} dxdt + \int_Q \frac{|w^n|^{\gamma'}}{(m^n)^{\gamma'-1}} dxdt \leq c_1.$$

Proof. Let $(m, w) \in \mathcal{K}$, hence $m \in L^{\alpha+1}(Q)$, $w \in L^1(Q)$. Since m is a weak solution of the Fokker-Planck equation with drift $A = w/m$, we may apply Proposition 2.5 to infer the existence of $C > 0$ and $\delta > 1$ (depending on the data), such that

$$\left(\int_Q m^{\alpha+1} dxdt\right)^{\delta} \leq C\left(\int_Q \frac{|w|^{\gamma'}}{m^{\gamma'-1}} dxdt + 1\right). \tag{28}$$

Moreover, by (3) and (5), we have that

$$\begin{aligned} \mathcal{E}(m, w) &\geq C_L^{-1} \int_Q \frac{|w|^{\gamma'}}{m^{\gamma'-1}} dxdt - C_F \int_Q m^{\alpha+1} dxdt - \|u_T\|_{L^\infty(\mathbb{T}^N)} - C_F T \\ &\geq C\left(\int_Q m^{\alpha+1} dxdt\right)^{\delta} - C_F \int_Q m^{\alpha+1} dxdt - \|u_T\|_{L^\infty(\mathbb{T}^N)} - C, \end{aligned} \tag{29}$$

which has a finite infimum, since $\delta > 1$.

In order to prove the second assertion, it suffices to use (28) for the sequence (m^n, w^n), rewriting (29) as

$$\begin{aligned} C_L^{-1} \int_Q \frac{|w^n|^{\gamma'}}{(m^n)^{\gamma'-1}} dxdt &\leq \mathcal{E}(m^n, w^n) + C_F \int_Q (m^n)^{\alpha+1} dxdt + \|u_T\|_{L^\infty(\mathbb{T}^N)} + C_F T \\ &\leq C + C\left(\int_Q \frac{|w^n|^{\gamma'}}{(m^n)^{\gamma'-1}} dxdt + 1\right)^{1/\delta}, \end{aligned}$$

and by (28) we conclude. ☐

We prove now that, up to subsequences, a minimizing sequence of $\mathcal{E}$ converges (in different spaces, according to the value of γ') to a minimizer.
A key point here is that, thanks to the convexity of $mL(w/m)$, lower semicontinuity of $\int_Q mL(w/m)$ in a strong topology will immediately imply weak lower semicontinuity w.r.t the corresponding weak topology. In turn, (strong) lower semicontinuity can be easily verified in $L^1(Q) \times L^1(Q)$. This result will be used several times in the following. We also refer to [26, Chapter 5] for a proof of lower semicontinuity in a more general setting.

Lemma 3.3. *Let $(m^n, w^n) \in \mathcal{K}$ be a minimizing sequence, that is*

$$\mathcal{E}(m^n, w^n) \xrightarrow{n\to\infty} c := \inf_{(m,w)\in\mathcal{K}} \mathcal{E}(m, w). \tag{30}$$

Then, up to subsequences, $m^n \to \bar{m}$ strongly in $L^\ell(Q)$ and $w^n \rightharpoonup \bar{w}$ weakly in $L^{\frac{\gamma'\ell}{\gamma'+\ell-1}}(Q)$, where:

- $\ell \in [1, \frac{N+2}{N+2-\gamma'})$, *if* $1 < \gamma' \le N + 2$;
- $\ell \in [1, +\infty)$, *if* $\gamma' \ge N + 2$.

Moreover, if $\gamma' > N + 2$, then $m^n \to \bar{m}$ in $C^{0,\theta}(Q)$ for some $\theta > 0$.

In particular, we can always take $\ell = \alpha + 1$ in the above statement, and, up to subsequences, $m^n \to \bar{m}$ a.e. in Q. Then, the couple $(\bar{m}, \bar{w})$ is a minimizer of $\mathcal{E}$ in $\mathcal{K}$.

Proof. Let $(m^n, w^n) \in \mathcal{K}$ be a minimizing sequence. By choosing n large enough, $\mathcal{E}(m^n, w^n) \le c + 1$, and Lemma 3.2 implies that

$$\int_Q \frac{|w^n|^{\gamma'}}{(m^n)^{\gamma'-1}} dxdt \le c_1$$

for some $c_1 > 0$. Note that since m^n is a weak solution to the Fokker-Planck equation with drift $A^n = w^n/m^n$,

$$\int_Q |A^n|^{\gamma'} m^n \, dxdt = \int_Q \frac{|w^n|^{\gamma'}}{(m^n)^{\gamma'-1}} \, dxdt \le c_1. \tag{31}$$

The thesis for m^n follows applying Corollary 2.7. In all cases, $\bar{m} \in L^{\alpha+1}(Q)$ (since $\alpha+1 < \frac{N+2}{N+2-\gamma'}$ by (8)) and, up to subsequences, $m^n \to \bar{m}$ a.e. in Q.

Concerning w^n, for all ℓ such that $m_n \to \bar{m}$ strongly in $L^\ell(Q)$, by Hölder inequality

$$\int_Q |w^n|^{\frac{\gamma'\ell}{\gamma'+\ell-1}} = \int_{\{m^n>0\}} |w^n|^{\frac{\gamma'\ell}{\gamma'+\ell-1}} \le \|m^n\|_{L^\ell(Q)}^{\frac{\gamma'-1}{\gamma'+\ell-1}} \left(\int_Q \frac{|w^n|^{\gamma'}}{(m^n)^{\gamma'-1}} \, dxdt \right)^{\frac{\ell}{\gamma'+\ell-1}} \le C.$$

Hence, w^n converges weakly to $\bar{w}$ in $L^{\frac{\gamma'\ell}{\gamma'+\ell-1}}(Q)$ and we can take in particular $\ell = \alpha + 1$. The fact that the limit $(\bar{m}, \bar{w})$ is a minimizer readily follows by weak lower semicontinuity of $\mathcal{E}$, that is ensured by convexity, and strong convergence of m_n. □

3.1 A convex problem

In order to find a link between the minimizer of $\mathcal{E}$ and the solution of (1), being the energy $\mathcal{E}$ not convex in (m, w) due to the presence of the term $-\int_Q F(x, m)dxdt$, we convexify it by adding a term that vanishes in $\bar{m}$, the limit of the minimizing sequence m^n. Therefore, let us define

$$\overline{\mathcal{E}}(m, w) = \mathcal{E}(m, w) + \int_Q G(x, t, m)dxdt,$$

where for $(x, t, m) \in Q \times [0, +\infty)$

$$G(x, t, m) := \frac{c_f + 1}{\alpha(\alpha + 1)} \left[(m + 1)^{\alpha+1} - (\bar{m}(x, t) + 1)^{\alpha+1}\right] - \frac{c_f + 1}{\alpha} (\bar{m}(x, t) + 1)^\alpha (m - \bar{m}(x, t))$$

and $G(x,t,m) = +\infty$ otherwise. Then, for $(x,t,m) \in Q \times [0,+\infty)$

$$\begin{aligned} g(x,t,m) &:= \partial_m G(x,t,m) = \frac{c_f+1}{\alpha}(m+1)^\alpha - \frac{c_f+1}{\alpha}(\bar m(x,t)+1)^\alpha, \\ \partial_m g(x,t,m) &= \partial_{mm} G(x,t,m) = (c_f+1)(m+1)^{\alpha-1}. \end{aligned}$$

Note that $G(x,t,\bar m(t,x)) = \partial_m G(x,t,\bar m(t,x)) = 0$ for all $(x,t) \in Q$, and $\partial_{mm}G(x,t,m) \geq 0$ for all $m \geq 0$, so that $G(x,m) \geq 0$ everywhere. Moreover,

$$\begin{aligned} \frac{c_f}{\alpha} - f(x,0) + \frac{1}{\alpha}(m+1)^\alpha - \frac{c_f+1}{\alpha}(\bar m(x,t)+1)^\alpha \leq \\ (-f+g)(x,t,m) \leq \frac{c_f+1}{\alpha}(m+1)^\alpha - \frac{c_f+1}{\alpha}(\bar m(x,t)+1)^\alpha. \end{aligned} \tag{32}$$

Lemma 3.4. *$\overline{\mathcal{E}}$ is convex on $\mathcal{K}$, and it is strictly convex with respect to m, that is*

$$\overline{\mathcal{E}}(\tau m + (1-\tau)\mu, \tau w + (1-\tau)v) \leq \tau \overline{\mathcal{E}}(m,w) + (1-\tau)\overline{\mathcal{E}}(\mu,v) - \frac{1}{2}\tau(1-\tau)\int_Q \psi(m,\mu)\, dxdt$$

for all $\tau \in [0,1]$, $(m,w),(\mu,v) \in \mathcal{K}$, where $\psi(x,y) = \min\{(x+1)^{\alpha-1},(y+1)^{\alpha-1}\}(x-y)^2$.

Note that if $m,\mu \in L^{\alpha+1}(\mathbb{T}^N)$ satisfy $\int_Q \psi(m,\mu)\,dxdt \leq 0$, then $m = \mu$ a. e. in Q.

Proof. It is standard to show that $(m,w) \mapsto \int_Q mL\left(-\frac{w}{m}\right) dxdt + \int_{\mathbb{T}^N} u_T(x)m(x,T)\,dx$ is convex. It is then sufficient to note that $\partial_{mm}(-F+G)(x,m) = -\partial_m f(x,m) + (c_f+1)(m+1)^{\alpha-1} \geq (m+1)^{\alpha-1}$, because of (4). ☐

We consider now the MFG system associated to $\overline{\mathcal{E}}$:

$$\begin{cases} -u_t - \Delta u + H(\nabla u) = -f(x,m) + g(x,t,m), \\ m_t - \Delta m - \operatorname{div}(\nabla H(\nabla u)\, m) = 0 & \text{in } Q, \\ m(x,0) = m_0(x), \quad u(x,T) = u_T(x) & \text{on } \mathbb{T}^N. \end{cases} \tag{33}$$

Proposition 3.5. *If $\gamma' > 1$, there exists a weak solution $(\tilde u, \tilde m)$ to the MFG system* (33)*, such that the couple $(\tilde m, \tilde m \nabla H(\nabla \tilde u))$ is a minimizer of the (convex) energy functional $\overline{\mathcal{E}}$ in $\mathcal{K}$. Moreover, if $\gamma' > N+2$, such a solution is classical.*

Proof. **1.** We consider first the case $1 < \gamma' \leq N+2$.

In order to apply the standard variational theory for convex systems, see [8] and Section 2.2, we need to truncate $\bar m$ because estimate (32) has left and right bounds which depend on (x,t), while this is not the case for (23).

1.1 Let us, therefore, consider the truncated function $\bar m_M$ defined for $M > 0$, as

$$\bar m_M(x,t) := \begin{cases} \bar m(x,t) & \forall (x,t) \in Q \text{ s.t } 0 \leq \bar m(x,t) \leq M, \\ M & \text{otherwise} \end{cases}$$

and the truncated energy functional

$$\overline{\mathcal{E}}_M(m,w) := \mathcal{E}(m,w) + \int_Q G_M(x,t,m)dxdt,$$

where for $(x,t,m) \in Q \times [0,+\infty)$

$$G_M(x,t,m) := \frac{c_f+1}{\alpha(\alpha+1)}\left[(m+1)^{\alpha+1} - (\bar m_M(x,t)+1)^{\alpha+1}\right] - \frac{c_f+1}{\alpha}(\bar m_M(x,t)+1)^\alpha(m - \bar m_M(x,t)).$$

and $G_M(x,t,m) = +\infty$ otherwise. Then, defining for $(x,t,m) \in Q \times [0,+\infty)$

$$g_M(x,t,m) := \partial_m G_M(x,t,m) = \frac{c_f+1}{\alpha}(m(x,t)+1)^\alpha - \frac{c_f+1}{\alpha}(\bar{m}_M(x,t)+1)^\alpha,$$

we have, in particular, that for $(x,t,m) \in Q \times [0,+\infty)$

$$\frac{c_f}{\alpha} - f(x,0) + \frac{1}{\alpha}(m+1)^\alpha - \frac{c_f+1}{\alpha}(M+1)^\alpha \le (-f+g_M)(x,t,m) \le \frac{c_f+1}{\alpha}(m+1)^\alpha - \frac{c_f+1}{\alpha}(M+1)^\alpha.$$

Thus $\overline{\mathcal{E}}_M(m,w)$ is strictly convex w.r.t. m and it is precisely a functional of the type studied in [8]. Indeed the coupling $-f+g_M$ is strictly increasing and continuous w.r.t. m and satisfies (23). Note that it should also verify the normalization condition (24). This condition is satisfied modifying G_M with $G_M + (f(x,0) - g_M(x,t,0))m$. Since

$$0 \le f(x,0) - g_M(x,t,0) \le C_f - \frac{c_f+1}{\alpha} + \frac{c_f+1}{\alpha}(M+1)^\alpha,$$

this will only modify the constants in the right hand side of the previous inequality for $-f+g_M$.

Therefore, for all $M>0$, Theorem 2.10 gives us a weak solution (u_M, m_M) such that u_M is bounded below by a constant depending on $\|u_T\|_{C^2}$ and on $\|H(\nabla u_T)\|_\infty$.

1.2. We have now to show the stability of solutions with respect to this approximation. We follow Section 6.4 in [8]. Note that it is enough to set $A = Id$ in their second order MFG system in order to obtain our truncated convex problem. On the one side, we are in a simpler case than the one in [8], since we are only approximating the coupling function $-f+g$ with the sequence $-f+g_M$, while all the other data are not approximated. On the other, even if $-f+g_M$ converges locally uniformly to $-f+g$, the limit $-f+g$ satisfies the more general inequality (32), where the right and left bounds depend also on (x,t) (not only in m). However, we have the additional information that they are bounded in $L^{1+\frac{1}{\alpha}}(Q)$, since $\bar{m} \in L^{\alpha+1}(Q)$. Note that the fact that the sequence $-f+g_M$ depends also on time does not add any difficulty as stated in the introduction of [8].

Let $w_M := -m_M \nabla H(\nabla u_M)$. By Theorem 2.11, the couple (m_M, w_M) is a minimizer for $\overline{\mathcal{E}}_M$. Proceeding as in [8] (or as we did for the minimizing sequence (m^n, w^n) in Lemma 3.3), we can prove that

$$\|m_M\|_{L^{\alpha+1}(Q)} + \|w_M\|_{L^{\frac{\gamma'(\alpha+1)}{\gamma'+\alpha}}(Q)} + \left\| \frac{|w_M|^{\gamma'}}{(m_M)^{\gamma'-1}} \right\|_{L^1(Q)} \le C$$

and for all $M>0$, $t \to m_M(t)$ are uniformly Hölder continuous with respect to the standard weak* topology in $\mathcal{P}(\mathbb{T}^N)$, see Lemma 4.1 in [8]. Hence, up to a subsequence, (m_M, w_M) converges weakly in $L^{\alpha+1}(Q) \times L^{\frac{\gamma'(\alpha+1)}{\gamma'+\alpha}}(Q)$ to some $(\tilde{m}, \tilde{w})$ and $m_M(t)$ converges to $\tilde{m}(t)$ in $C^0([0,T], \mathcal{P}(\mathbb{T}^N))$ as $M \to +\infty$. It follows that $(\tilde{m}, \tilde{w})$ satisfy $\tilde{w}^{\gamma'} \tilde{m}^{1-\gamma'} \in L^1(Q)$ and for all $\varphi \in C_0^\infty(\mathbb{T}^N \times [0,T))$

$$\begin{aligned} 0 &= \lim_{M\to+\infty} \int_Q (m_M \varphi_t + w_M \cdot \nabla\varphi + m_M \Delta\varphi) dxdt + \int_{\mathbb{T}^N} m_0(x)\varphi(x,0)dx \\ &= \int_Q (\tilde{m}\varphi_t + \tilde{w}\cdot\nabla\varphi + \tilde{m}\Delta\varphi) dxdt + \int_{\mathbb{T}^N} m_0(x)\varphi(x,0)dx. \end{aligned}$$

Hence $(\tilde{m}, \tilde{w})$ satisfies (6) and belongs to $\mathcal{K}$.

We claim that

$$\limsup_{M\to+\infty} \inf_{(m,w)\in\mathcal{K}} \overline{\mathcal{E}}_M(m,w) \le \inf_{(m,w)\in\mathcal{K}} \overline{\mathcal{E}}(m,w). \tag{34}$$

Indeed,

$$\limsup_{M\to+\infty} \inf_{(m,w)\in\mathcal{K}} \overline{\mathcal{E}}_M(m,w)$$
$$\leq \limsup_{M\to+\infty} \int_Q mL\left(-\frac{w}{m}\right) - F(x,m) + G_M(x,t,m)\,dxdt + \int_{\mathbb{T}^N} u_T(x)m(x,T)\,dx,$$

for all $(m,w)\in\mathcal{K}$. Now, the locally uniform convergence of G_M to G gives, by dominated convergence theorem,

$$\limsup_{M\to+\infty} \int_Q G_M(x,t,m)\,dxdt = \lim_{M\to+\infty} \int_Q G_M(x,t,m)\,dxdt = \int_Q G(x,t,m)\,dxdt.$$

Hence,

$$\limsup_{M\to+\infty} \int_Q mL\left(-\frac{w}{m}\right) - F(x,m) + G_M(x,t,m)\,dxdt + \int_{\mathbb{T}^N} u_T(x)m(x,T)\,dx = \overline{\mathcal{E}}(m,w)$$

for all $(m,w)\in\mathcal{K}$. Thus,

$$\limsup_{M\to+\infty} \inf_{(m,w)\in\mathcal{K}} \overline{\mathcal{E}}_M(m,w) \leq \overline{\mathcal{E}}(m,w)$$

for all $(m,w)\in\mathcal{K}$ and (34) holds.

Let us now show that

$$\liminf_{M\to+\infty} \overline{\mathcal{E}}_M(m_M,w_M) \geq \overline{\mathcal{E}}(\tilde{m},\tilde{w}).$$

Recall that, by convexity, we have that $(m,w)\mapsto mL(-w/m)$ is weakly lower semicontinuous, hence

$$\liminf_{M\to+\infty} \int_Q m_M L\left(-\frac{w_M}{m_M}\right) dxdt \geq \int_Q \tilde{m}L\left(-\frac{\tilde{w}}{\tilde{m}}\right) dxdt,$$

moreover by the definition of F and G_M we have that

$$-F(x,m) + G_M(x,t,m) = \int_0^m -f(x,\sigma) + g_M(x,t,\sigma)d\sigma \quad \forall (x,t,m)\in Q\times[0,+\infty),$$
$$-F(x,m) + G_M(x,t,m) = +\infty \quad \text{otherwise.}$$

Hence $-F+G_M$ is lower semicontinuous and by convexity of $-F+G_M$, weak lower semicontinuity follows. Thus, by strong convergence of $\bar{m}_M$ to $\bar{m}$ in $L^{\alpha+1}(Q)$, we have the convergence of $\int_Q G_M\,dxdt$ to $\int_Q G\,dxdt$, hence

$$\liminf_{M\to+\infty} \int_Q (-F(x,m_M) + G_M(x,t,m_M))\,dxdt \geq \int_Q (-F(x,\tilde{m}) + G(x,t,\tilde{m}))\,dxdt.$$

Finally, by the weak* convergence of $m_M(T)$ in $\mathcal{P}(\mathbb{T}^N)$ we have

$$\lim_{M\to+\infty} \int_{\mathbb{T}^N} u_T(x)m_M(x,T)\,dx = \int_{\mathbb{T}^N} u_T(x)\tilde{m}(x,T)\,dx.$$

Hence,

$$\liminf_{M\to+\infty} \overline{\mathcal{E}}_M(m_M,w_M) =$$
$$\liminf_{M\to+\infty} \int_Q m_M L\left(-\frac{w_M}{m_M}\right) - F(x,m_M) + G_M(x,t,m_M)\,dxdt + \int_{\mathbb{T}^N} u_T(x)m_M(x,T)\,dx$$
$$\geq \int_Q \tilde{m}L\left(-\frac{\tilde{w}}{\tilde{m}}\right) - F(x,\tilde{m}) + G(x,t,\tilde{m})\,dxdt + \int_{\mathbb{T}^N} u_T(x)\tilde{m}(x,T)\,dx$$
$$= \overline{\mathcal{E}}(\tilde{m},\tilde{w}).$$

Therefore, thanks to (34),

$$\overline{\mathcal{E}}(\tilde m,\tilde w)\le \liminf_{M\to+\infty}\overline{\mathcal{E}}_M(m_M,w_M)\le \limsup_{M\to+\infty}\inf_{(m,w)\in\mathcal{K}}\overline{\mathcal{E}}_M(m,w)\le \inf_{(m,w)\in\mathcal{K}}\overline{\mathcal{E}}(m,w).$$

thus

$$\overline{\mathcal{E}}(\tilde m,\tilde w)=\lim_{M\to+\infty}\overline{\mathcal{E}}_M(m_M,w_M)=\lim_{M\to+\infty}\min_{(m,w)\in\mathcal{K}}\overline{\mathcal{E}}_M(m,w)=\min_{(m,w)\in\mathcal{K}}\overline{\mathcal{E}}(m,w).$$

In particular $(\tilde m,\tilde w)$ minimizes $\overline{\mathcal{E}}$, and

$$\lim_{M\to+\infty}\int_Q m_M L\left(-\frac{w_M}{m_M}\right)dxdt=\int_Q \tilde m L\left(-\frac{\tilde w}{\tilde m}\right)dxdt$$

and

$$\lim_{M\to+\infty}\int_Q(-F(x,m_M)+G_M(x,t,m_M))\,dxdt=\int_Q(-F(x,\tilde m)+G(x,t,\tilde m))\,dxdt.$$

Since $-F+G_M$ are bounded below and $-F+G$ is strictly convex, an argument using Young measures as the one in [8], gives us the strong convergence of every subsequence of (m_M,w_M) to a minimizer $(\tilde m,\tilde w)$ of $\overline{\mathcal{E}}$, in $L^{\alpha+1}(Q)\times L^{\frac{\gamma'(\alpha+1)}{\gamma'+\alpha}}(Q)$. Due to strict convexity of $\bar{\mathcal{E}}$ w.r.t. m, the minimizer $\tilde m$ is unique, and by strict convexity of L, we have uniqueness of $\tilde w/\tilde m$ where $\tilde m>0$. By optimality, $\tilde w=0$ when $\tilde m=0$ (otherwise $\tilde m L(-\frac{\tilde w}{\tilde m})=+\infty$), hence uniqueness follows also for $\tilde w$. Being the minimizer unique, the full sequence (m_M,w_M) strongly converges to $(\tilde m,\tilde w)$.

1.3 We are left to prove that we can find a weak solution for (33) from the minimizer $(\tilde m,\tilde w)$.

Let $\beta_M(x,t):=(-f+g_M)(x,t,m_M(x,t))$ on Q. Thanks to the growth condition on $-f+g_M$ and the uniform bound of m_M in $L^{\alpha+1}(Q)$, the sequence β_M weakly converges in $L^{1+\frac{1}{\alpha}}(Q)$ to $\tilde\beta$.

Being u_M uniformly bounded by below, Theorem 2.8 gives

$$\|u_M\|_{L^\infty((0,T),L^{\bar\vartheta}(\mathbb{T}^N))}+\|u_M\|_{L^{\vartheta}(Q)}\le C,$$

where

$$\bar\vartheta:=\begin{cases}\frac{N(\gamma+\alpha)}{\alpha N-\gamma} & \text{if } \alpha>\frac{\gamma}{N}\\ +\infty & \text{if } \alpha\le\frac{\gamma}{N}\end{cases}.$$

Hence, up to a subsequence, u_M weakly converges to $\tilde u$ in $L^{\vartheta}(Q)$. Moreover, proceeding as in [8] we can prove that ∇u_M converges weakly to $\nabla\tilde u$ in $L^{\gamma}(Q)$. Hence by convexity of H we have that $(\tilde u,\tilde\beta)$ satisfies

$$-\tilde u_t-\Delta\tilde u+H(\nabla\tilde u)\le\tilde\beta\quad\text{in } Q,$$

in the sense of distributions.

By Lemma 5.3 in [8] (which does not involve the coupling function hence it holds even for our limit functions), we have

$$\left[\int_{\mathbb{T}^N}\tilde m\tilde u\right]_0^T+\int_Q\tilde m\left(\tilde\beta+L\left(-\frac{\tilde w}{\tilde m}\right)\right)\ \ge\ 0. \tag{35}$$

Moreover, for all $M>0$, being (u_M,m_M) a weak solution of (1) with coupling function $-f+g_M$, we have that (27) is satisfied. For a.e. $(x,t)\in Q$,

$$(-F+G_M)^*(x,t,\beta_M)+(-F+G_M)(x,t,m_M)=\beta_M(x,t)m_M(x,t)=(-f+g_M)(x,t,m_M)m_M(x,t),$$

(here, $(-F+G_M)^*$ is the Legendre transform of $-F+G_M$ with respect to m), hence, by the definition of w_M and β_M,

$$\int_Q (-F+G_M)^*(x,t,\beta_M) + (-F+G_M)(x,t,m_M) + m_M L\left(-\frac{w_M}{m_M}\right) dxdt$$
$$+ \int_{\mathbb{T}^N} u_T m_M(T) - u_M(0)m_0 dx = 0.$$

Following Step 3 of the proof of Proposition 5.4 in [8] we can prove that

$$\limsup_{M\to+\infty} \int_{\mathbb{T}^N} u_M(0)m_0 dx \le \int_{\mathbb{T}^N} \tilde{u}(0)m_0.$$

Hence passing to the limit in the previous identity, due to the convexity of the functionals involved, which implies weak lower semicontinuity, we have

$$\int_Q (-F+G)^*(x,t,\tilde{\beta}) + (-F+G)(x,t,\tilde{m}) + \tilde{m}L\left(-\frac{\tilde{w}}{\tilde{m}}\right) dxdt + \int_{\mathbb{T}^N} u_T\tilde{m}(T) - \tilde{u}(0)m_0 dx \le 0.$$

Using the convexity of $-F+G$, one obtains

$$(-F+G)^*(x,t,\tilde{\beta}) + (-F+G)(x,t,\tilde{m}) - \tilde{\beta}(x,t)\tilde{m}(x,t) \ge 0. \tag{36}$$

Therefore (35) is an equality, hence by Lemma 5.3 in [8], we have $\tilde{w} = -\tilde{m}\nabla H(\nabla\tilde{u})$. Moreover (36) holds a.e., hence $\tilde{\beta} = (-f+g)(\cdot,\cdot,\tilde{m})$. This proves that $(\tilde{u},\tilde{m})$ is a weak solution for (33). Note that this also implies that

$$\int_Q (-F+G)^*(x,t,\tilde{\beta}) + (-F+G)(x,t,\tilde{m}) + \tilde{m}L\left(-\frac{\tilde{w}}{\tilde{m}}\right) dxdt + \int_{\mathbb{T}^N} u_T\tilde{m}(T) - \tilde{u}(0)m_0 dx = 0.$$

2. Suppose now that $\gamma' > N+2$, so that, by Lemma 3.3, $\bar{m} \in C^{0,\theta}(Q)$, for a $\theta > 0$. Then, Theorem 6.4 in [8] applies directly giving a weak solution $(\tilde{u},\tilde{m})$ to our convex MFG problem (33), with no need to truncate $\bar{m}$, since there exists M such that $0 \le \bar{m}(x,t) \le M$ for all $(x,t) \in Q$.

We just have to show that the weak solution $(\tilde{u},\tilde{m})$ enjoys more regularity and it is indeed a classical solution of (33). Indeed, by Corollary 2.7, also $\tilde{m} \in C^{0,\theta}(Q)$, hence $\tilde{\beta} := (-f+g)(\cdot,\cdot,\tilde{m})$ is Hölder continuous on Q; moreover, we know by [8] that the pair $(\tilde{u},\tilde{\beta})$ is a minimizer of

$$\inf_{(u,\beta)\in\overline{\mathcal{K}}} \int_Q (-F+G)^*(x,\beta(x,t))\, dxdt - \int_{\mathbb{T}^N} u(0,x)m_0(x)\, dx, \tag{37}$$

where $\overline{\mathcal{K}}$ is the set of pairs (u,β) satisfying $(u,\beta) \in L^{\vartheta}(Q) \times L^{(\alpha+1)'}(Q)$, (where $(\alpha+1)'$ is the conjugate exponent of $\alpha+1$ and ϑ is as in Definition 2.9), and

$$-u_t - \Delta u + H(\nabla u) \le \beta \quad \text{in } Q,$$

and $u(T,\cdot) \le u_T(\cdot)$ in the sense of distributions. We may then consider the classical solution u_1 to

$$-(u_1)_t - \Delta u_1 + H(\nabla u_1) = \tilde{\beta}(x,t) \quad \text{in } Q,$$

$u(T,\cdot) = u_T(\cdot)$. Then, by comparison, the couple $(\tilde{u}_1,\tilde{\beta})$ is still a minimizer of (37), so $(u_1,\tilde{m})$ is also a solution to (33) (again by Theorem 6.4 in [8]). Since $u_1 \in C^2(Q)$, $\tilde{m}$ is a weak solution of the Fokker-Planck equation in the sense of Remark 2.1.

□

Since the existence of a solution of the convexified problem has been established, we are now ready to conclude the proofs of Theorems 1.1 and 1.2.

Proof of Theorem 1.1 and Theorem 1.2. We just need to show that the solution $(\tilde{u}, \tilde{m})$ of (33) given by Proposition 3.5 is such that $\tilde{m} = \bar{m}$ and $\tilde{w} = \bar{w}$, where $(\bar{m}, \bar{w})$ is given by Lemma 3.3. This will immediately imply that $(\tilde{u}, \tilde{m}) = (\tilde{u}, \bar{m})$ is not only a solution of (33), but also a solution of (1), and the couple $(\bar{m}, -\bar{m}\nabla H(\nabla\tilde{u}))$ is a minimizer of $\mathcal{E}$ in $\mathcal{K}$.

We observe that $\overline{\mathcal{E}}(m, w) \geq \mathcal{E}(m, w)$ (since $G \geq 0$), so

$$\overline{\mathcal{E}}(\tilde{m}, \tilde{w}) = \inf_{(m,w)\in\mathcal{K}} \overline{\mathcal{E}}(m, w) \geq \inf_{(m,w)\in\mathcal{K}} \mathcal{E}(m, w) = \mathcal{E}(\bar{m}, \bar{w}) = \overline{\mathcal{E}}(\bar{m}, \bar{w}).$$

Hence $(\bar{m}, \bar{w})$ is also a minimizer of $\overline{\mathcal{E}}$. By strict convexity of $\overline{\mathcal{E}}$ with respect to m, $\bar{m} = \tilde{m}$, and by strict convexity of L we have that $\frac{\bar{w}}{\bar{m}} = \frac{\tilde{w}}{\tilde{m}}$ on the set where $\tilde{m} = \bar{m} > 0$. Being $\bar{w} = 0$ when $\bar{m} = 0$, we also have $\bar{w} = \tilde{w}$, as wanted.

□

3.2 $2 < \gamma' \leq N + 2$: smooth solutions

We show in this section that, when $2 < \gamma' \leq N + 2$, we can find smooth solutions through a penalisation argument under additional hypothesis on α.

We consider the approximated (or penalised) Lagrangian

$$L_\eta(q) := L(q) + \frac{\eta}{N+3}|q|^{N+3}, \quad \forall q \in \mathbb{R}^N, \eta > 0,$$

and the associated functional

$$\mathcal{E}_\eta(m, w) = \int_Q mL_\eta\left(-\frac{w}{m}\right) - F(x, m)\, dxdt + \int_{\mathbb{T}^N} u_T(x)m(x, T)\, dx.$$

Note that the minimization of $\mathcal{E}_\eta$ has to be performed on $\mathcal{K}_{N+3,\alpha} \subset \mathcal{K}_{\gamma',\alpha}$. We are basically increasing the growth of L in order to gain regularity for minimizers of the energy. In particular, if L grows faster that $|q|^{N+2}$, a solution of the Fokker-Planck equation, that enters in the constraint $\mathcal{K}$, enjoys automatically Hölder regularity (see Corollary 2.7). Note that a similar penalisation argument has been implemented in [23] in the stationary setting. For any fixed $\eta > 0$, $L_\eta(q)$ behaves like $|q|^{N+3}$ as $|q| \to +\infty$, namely

$$c_\eta^{-1}|q|^{N+3} - c_\eta \leq L_\eta(q) \leq c_\eta(|q|^{N+3} + 1),$$

for all $q \in \mathbb{R}^N$ and some positive c_η (depending on η). The corresponding family of Hamiltonians H_η satisfy

$$h_\eta^{-1}|p|^{\frac{N+3}{N+2}} - h_\eta \leq H_\eta(p) \leq h_\eta(|p|^{\frac{N+3}{N+2}} + 1),$$

together with additional bounds independent of η. In particular, we have by (3) that

$$-C_L \leq -L(0) \leq H_\eta(p) = \sup_{q\in\mathbb{R}^N}[p \cdot q - L_\eta(q)] \leq \sup_{q\in\mathbb{R}^N}[p \cdot q - L(q)] = H(p) \leq C_H(|p|^\gamma + 1), \quad (38)$$

where C does not depend on η. Moreover,

$$|\nabla H_\eta(p)| \leq C(|p|^{\gamma-1} + 1). \quad (39)$$

Indeed, by the definition of the Legendre transform,

$$L_\eta(\nabla H_\eta(p)) = \nabla H_\eta(p)\cdot p - H_\eta(p) \quad \forall p \in \mathbb{R}^N.$$

Therefore, by (38),

$$C_L^{-1}|\nabla H_\eta(p))|^{\gamma'} - C_L \le L_\eta(\nabla H_\eta(p)) \le |\nabla H_\eta(p)|\cdot|p| + C_L \le \frac{C_L^{-1}}{2}|\nabla H_\eta(p))|^{\gamma'} + C|p|^\gamma + C_L,$$

which implies (39), as $\gamma/\gamma' = \gamma - 1$.

When $\eta > 0$ is fixed, H_η satisfies (2) with $\gamma = \frac{N+3}{N+2}$, hence its conjugate $\gamma' = N+3 > N+2$, so Theorem 1.1 applies. In particular, there exists a classical solution (u_η, m_η) of

$$\begin{cases} -u_t - \Delta u + H_\eta(\nabla u) = -f(x, m(x,t)), & \\ m_t - \Delta m - \operatorname{div}(\nabla H_\eta(\nabla u)\, m) = 0 & \text{in } Q, \\ m(x,0) = m_0(x), \quad u(x,T) = u_T(x) & \text{on } \mathbb{T}^N. \end{cases} \tag{40}$$

such that, setting $w_\eta := -m_\eta \nabla H(\nabla u_\eta)$, then (m_η, w_η) is a minimizer of $\mathcal{E}_\eta$ in $\mathcal{K}_{N+3,\alpha}$. We will show that (u_η, m_η) converges as $\eta \to 0$ to a solution of the original problem. Before proving Theorem 1.3 we state some a priori estimates that will be crucial to pass to the limit.

Lemma 3.6. *For all $\ell \in [1, \frac{N+2}{N+2-\gamma'})$, there exists $C_\ell > 0$ such that*

$$\|m_\eta\|_{L^\ell(Q)} \le C_\ell. \tag{41}$$

Proof. We observe that $\mathcal{E}_\eta(m,w) \le \mathcal{E}_1(m,w)$ for all $\eta \le 1$ and $(m,w) \in \mathcal{K}_{N+3,\alpha}$. Hence, $\mathcal{E}_\eta(m_\eta, w_\eta) = \min \mathcal{E}_\eta \le \min \mathcal{E}_1$ for all $\eta \le 1$. Since

$$L_\eta(q) = L(q) + \frac{\eta}{N+3}|q|^{N+3} \ge C_L^{-1}|q|^{\gamma'} - C_L,$$

arguing as in Lemma 3.2 and recalling that m_η solves the Fokker-Planck equation with drift $A_\eta = \nabla H_\eta(\nabla u_\eta)$, we get

$$\int_Q m_\eta^{\alpha+1}\, dxdt + \int_Q |\nabla H_\eta(\nabla u_\eta))|^{\gamma'} m_\eta\, dxdt \le C.$$

Then we can apply Corollary 2.7 to conclude. □

Lemma 3.7. *Suppose that*

$$\alpha < \min\left\{\frac{\gamma'}{N}, \frac{\gamma'-2}{N+2-\gamma'}\right\}.$$

Then, there exists $C > 0$ such that

$$\|m_\eta\|_{L^\infty(Q)} \le C \tag{42}$$

for all $\eta > 0$.

Proof. By contradiction, let $M_\eta > 0, x_\eta \in \mathbb{T}^N$ be such that

$$0 < M_\eta := m_\eta(x_\eta, t_\eta) = \max_{(x,t)\in\overline{Q}} m_\eta(x,t) \to \infty, \quad \text{as } \eta \to 0.$$

1. Let us define the following blow-up sequences

$$v_\eta(x,t) := a_\eta^{\gamma'-2} u_\eta(x_\eta + a_\eta x, t_\eta + a_\eta^2 t), \quad \mu_\eta(x,t) := \frac{1}{M_\eta} m_\eta(x_\eta + a_\eta x, t_\eta + a_\eta^2 t), \quad a_\eta = M_\eta^{-\alpha/(\gamma'-2)}, \tag{43}$$

for all $(x,t) \in Q_\eta := \{(x,t) \in \mathbb{R}^{N+1} : (x_\eta + a_\eta x, t_\eta + a_\eta^2 t) \in Q\}$. Then, (v_η, μ_η) solves

$$\begin{cases} -(v_\eta)_t - \Delta v_\eta + \widehat{H}_\eta(\nabla v_\eta) = -f_\eta(x, \mu_\eta) \\ (\mu_\eta)_t - \Delta \mu_\eta(x) - \operatorname{div}(\nabla \widehat{H}_\eta(\nabla v_\eta(x))\, \mu_\eta(x)) = 0 \quad \text{in } Q_\eta, \\ \mu_\eta\left(x, -t_\eta/a_\eta^2\right) = \frac{1}{M_\eta} m_0(x_\eta + a_\eta x), \\ v_\eta\left(x, (T - t_\eta)/a_\eta^2\right) = a_\eta^{\gamma'-2} u_T(x_\eta + a_\eta x) \qquad \text{on } T_\eta\ , \end{cases} \tag{44}$$

where $T_\eta = \{x : x_\eta + a_\eta x \in \mathbb{T}^N\}$, $\widehat{H}_\eta(p) = a_\eta^{\gamma'} H_\eta(a_\eta^{1-\gamma'} p)$, $f_\eta(x,\mu) = a_\eta^{\gamma'} f(x_\eta + a_\eta x, M_\eta \mu_\eta)$. Note that

$$a_\eta \to 0 \quad \text{as } \eta \to 0,$$

and $\widehat{H}_\eta$ satisfies for some $C_1 > 0$

$$-a_\eta^{\gamma'} C_1 \le \widehat{H}_\eta(p) \le C_H(|p|^\gamma + a_\eta^{\gamma'}) \le C_1(|p|^2 + 1) \tag{45}$$

for all $\eta > 0$ and $p \in \mathbb{R}^N$, by (38). Similarly,

$$|\nabla \widehat{H}_\eta(p)| \le C(|p|^{\gamma-1} + 1), \tag{46}$$

by (39).

Moreover, $\mu_\eta \le 1$ on Q_η and

$$0 \le f_\eta(x, \mu_\eta) \le C_f a_\eta^{\gamma'}(M_\eta^\alpha + 1) \le C_f(a_\eta^2 + a_\eta^{\gamma'}) \tag{47}$$

for all η, by (4). Moreover, since $\gamma' \ge 2$, μ_η and v_η are bounded (uniformly in η) in $W^{2,\infty}(T_\eta)$ at initial and final time respectively.

2. We show that v_η and its gradient are bounded on Q_η. It suffices to observe that

$$\bar{v}_\eta(x,t) := a_\eta^{\gamma'-2} \sup_{x \in \mathbb{T}^N} u_T(x) + C_1 a_\eta^{\gamma'} \left(\frac{T - t_\eta}{a_\eta^2} - t \right),$$
$$\underline{v}_\eta(x,t) := a_\eta^{\gamma'-2} \inf_{x \in \mathbb{T}^N} u_T(x) - (C_H a_\eta^{\gamma'} + C_f a_\eta^{\gamma'} + C_f a_\eta^2) \left(\frac{T - t_\eta}{a_\eta^2} - t \right)$$

are supersolutions and subsolutions respectively of the (backward) Cauchy problem for the HJB equation in (44). Hence,

$$-C \le \underline{v}_\eta(x,t) \le v_\eta(x,t) \le \bar{v}_\eta(x,t) \le C \quad \text{on } Q_\eta \tag{48}$$

by the Comparison Principle. Proposition 2.3 applies with an arbitrary choice of a ball $\Omega' = B_1(x_0)$, as (45), (47), (48) hold with constants that do not depend on x_0, thus

$$\|\nabla v_\eta\|_{L^\infty(B_1(x_0) \times (-t_\eta/a_\eta^2, (T - t_\eta)/a_\eta^2))} \le C.$$

Therefore, $\|\nabla v_\eta\|_{L^\infty(Q_\eta)} \le C$.

3. The rescaled distribution μ_η is a solution of the following linear equation

$$(\mu_\eta)_t - \Delta \mu_\eta = \operatorname{div}(\Phi_\eta(x,t)) \quad \text{in } Q_\eta,$$

where $\Phi_\eta(x,t) = \nabla \widehat{H}_\eta(\nabla v_\eta(x,t))\mu_\eta(x,t)$ is bounded in $L^\infty(Q_\eta)$ uniformly with respect to η, by the previous step and (46). Note that $|\mu_\eta|$ itself is bounded by one. Thus, by classical parabolic regularity (see, for example, [17, Theorem V.1.1]), we conclude that $\|\mu_\eta\|_{C^{0,\theta/2}([0,T],C^{0,\theta}(\mathbb{T}^N))} \le C$ for some $\theta > 0$. It follows that μ_η is bounded away from zero in a neighbourhood of zero, as $\mu_\eta(0,0) = 1$. Therefore, for any fixed $\ell > 0$, there exists some neighbourhood U of $(x,t) = (0,0)$ and $\delta > 0$, depending on ℓ but not on η, such that

$$\int_U (\mu_\eta)^\ell \, dxdt \ge \delta. \tag{49}$$

We may choose ℓ so that

$$\alpha \frac{N+2}{\gamma' - 2} < \ell < \frac{N+2}{N+2-\gamma'},$$

by the assumptions on α. Thus, in view of (41),

$$\int_{Q_\eta} (\mu_\eta)^\ell \, dxdt = \frac{1}{M_\eta^\ell} \int_{Q_\eta} (m_\eta(x_\eta + a_\eta x, t_\eta + a_\eta^2 t))^\ell \, dxdt = M_\eta^{\alpha \frac{(N+2)}{\gamma'-2} - \ell} \|m_\eta\|^\ell_{L^\ell(Q)} \to 0$$

as $\eta \to 0$, by the assumptions on α and the fact that $M_\eta \to \infty$, but this contradicts (49).

□

Proof of Theorem 1.3. Once $L^\infty(Q)$ estimates on m_η are in force, we just have to improve the bounds on (u_η, m_η) to pass to (classical) limits in (40). The Maximum Principle and (42) guarantee that u_η is bounded in $L^\infty(Q)$. Then, by Proposition 2.3, (38) and (42), ∇u_η is bounded in some Hölder space, independently on η. Note that such a bound can be extended to $(0,T]$ by the regularity of the final datum u_T. Standard parabolic estimates then provide classical regularity of u_η: indeed, one may apply first [17, Theorem V.1.1 or Theorem III.10.1] to the Fokker-Planck equation to obtain Hölder continuity of m_η and use Schauder estimates [17, Section IV.5] to the HJB equation to conclude. Finally, note that also $-w_\eta/m_\eta = \nabla H(\nabla u_\eta)$ is bounded in $L^\infty(Q)$ independently on η. Therefore, the penalisation term $\frac{\eta}{N+3}\int_Q m_\eta |w_\eta/m_\eta|^{N+3}$ in $\mathcal{E}_\eta$ vanishes, implying

$$\min_{\mathcal{K}_{N+3,\alpha}} \mathcal{E}_\eta \to \min_{\mathcal{K}_{N+3,\alpha}} \mathcal{E} \quad \text{as } \eta \to 0.$$

□

Remark 3.8. Note that the penalization procedure leads to a minimizer of $\mathcal{E}$ in $\mathcal{K}_{N+3,\alpha}$, rather than in the natural and larger constraint set $\mathcal{K}_{\gamma',\alpha}$.

Remark 3.9. In Lemma 3.7, the rescaling is designed so that (v_η, μ_η) solves the Hamilton-Jacobi-Bellman equation

$$-(v_\eta)_t - \Delta v_\eta + \widehat{H}_\eta(\nabla v_\eta) = -f_\eta(x, \mu_\eta),$$

where $|f_\eta| \le 1$ on Q for all η. Then, gradient bounds $\|\nabla v_\eta\|_{L^\infty(Q_\eta)} \le C$ are used to run an argument by contradiction; those bounds are obtained by first estimating $\|v_\eta\|_{L^\infty(Q_\eta)} \le C$ through a comparison principle. Lipschitz estimates then follow.

Note that gradient estimates available in the literature usually depend on bounds on the solution itself. On the other hand, for *stationary* HJB equations, namely

$$\lambda - \Delta v(x) + H(\nabla v(x)) = F(x),$$

it is possible to prove Lipschitz estimates that *do not* depend a priori on $\|v\|_{L^\infty(Q)}$, for example by means of the Bernstein method (see [18]). This key fact has been used in [10] to prove existence of classical solutions to (1) in the stationary case, for a wider range of couplings f. If similar estimates were available also in the time dependent case, i.e. $\|\nabla v_\eta\|_{L^\infty(Q_\eta)} \le C$, with C depending on $\|f_\eta\|_{L^\infty(Q_\eta)}$, $\|v_\eta(\cdot,T)\|_{L^\infty(Q_\eta)}$ but not on $\|v\|_{L^\infty(Q)}$ and the time horizon T, then it would have been possible to have Lemma 3.7 in the whole range $\alpha \in (0, \gamma'/N)$, and hence the existence of classical solutions. Note that our v_η and f_η are space-periodic, but the period $a_\eta^{-1} \to \infty$ as $\eta \to 0$, so periodicity disappears in the blow-up limit.

4 A non-uniqueness example

The aim of this section is to prove that, under additional assumptions, (1) has multiple solutions. Consider $f(x,m) = f(m)$, $u_T \equiv 0$ and $m_0 \equiv 1$. Note that the corresponding system

$$\begin{cases} -u_t - \Delta u + H(\nabla u) = -f(m(x,t)) & \text{in } Q, \\ m_t - \Delta m - \operatorname{div}(\nabla H(\nabla u)\, m) = 0 & \text{in } Q, \\ m(x,0) = 1, \quad u(x,T) = 0 & \text{on } \mathbb{T}^N. \end{cases} \tag{50}$$

has always a trivial solution $(\bar u, \bar m) = ((t-T)(f(1)+H(0)), 1)$. We look for a solution to (50) that is not the trivial one.

We will require that for some $b > 2$, $c > 0$,

$$\begin{aligned} &\bullet \quad 0 \le L(q) \le c|q|^b \quad \text{for all } |q| \le 1, \\ &\bullet \quad f'(m) \ge c \quad \text{for all } |m| \le 2. \end{aligned} \tag{51}$$

Note that (51) implies that H cannot be of class C^2 in a neighborhood of $p = 0$. As an example, if $L(q) = \frac{1}{b}|q|^b$ with $b > 2$, then $H(p) = \frac{b-1}{b}|p|^{\frac{b}{b-1}} \in C^{1,\frac{1}{b-1}}(\mathbb{R}^N)$.

Proposition 4.1. *Under the assumptions of Theorem 1.1 and* (51)*, for any $T > 0$ there exist at least two (different) classical solutions to* (50)*.*

Proof. In view of Theorem 1.1, there exists a classical solution $(\tilde u, \tilde m)$ of (1) such that the couple $(\tilde m, \tilde w) = (\tilde m, -\tilde m \nabla H(\nabla \tilde u))$ is a minimizer of $\mathcal{E}$. By adding a constant to $\mathcal{E}$, we may assume that $F(1) = 0$, hence $\mathcal{E}(1,0) = 0$. Our aim is to show that $(\tilde u, \tilde m)$ cannot be the trivial solution; this will be achieved by proving that $\mathcal{E}(\tilde m, -\tilde m \nabla H(\nabla \tilde u)) = \min_{\mathcal{K}} \mathcal{E}(m,w) < \mathcal{E}(1,0) = 0$.

In order to build a suitable competitor (m,w), let us consider $\mu(x) := 1 + \epsilon\varphi(x)$, where $\epsilon > 0$ will be chosen later and $\varphi \in C^2(\mathbb{T}^N)$ is a non-trivial eigenfunction of the Laplacian, i.e.

$$-\Delta\varphi = \lambda\varphi \quad \text{on } \mathbb{T}^N$$

for some $\lambda > 0$. Note that $\int_{\mathbb{T}^N} \varphi dx = 0$. Set $v(x) := \nabla\mu(x) = \epsilon\nabla\varphi(x)$. In order to connect the initial datum $m_0 \equiv 1$ to μ at $t = T/2$, we define on Q

$$\begin{aligned} m(x,t) &:= 1 + \zeta(t)(\mu(x) - 1) = 1 + \epsilon\zeta(t)\varphi(x), \\ w(x,t) &:= \zeta(t)v(x) + \frac{\epsilon}{\lambda}\zeta'(t)\nabla\varphi(x) = \epsilon\nabla\varphi(x)[\zeta(t) + \lambda^{-1}\zeta'(t)], \end{aligned}$$

where $\zeta : [0,T] \to \mathbb{R}$ is a smooth function which is zero at $t = 0$ and equal to one in the interval $[T/2, T]$. One easily verifies that $(u,m) \in \mathcal{K}$. Moreover, there exist constants C_1, C_2 depending on T, but not on $\epsilon \le \epsilon_0$, such that by (51)

$$\int_0^T \int_{\mathbb{T}^N} mL\left(-\frac{w}{m}\right) dxdt \le c \int_0^T \int_{\mathbb{T}^N} \epsilon^b |1 + \epsilon\zeta\varphi|^{1-b} |\nabla\varphi(\zeta + \lambda^{-1}\zeta')|^b \, dxdt \le C_1 T\epsilon^b, \tag{52}$$

$$\int_0^T \int_{\mathbb{T}^N} F(m) dxdt \ge \int_0^T \int_{\mathbb{T}^N} F(1) + \epsilon f(1)\zeta\varphi + \frac{c}{2}\epsilon^2\zeta^2\varphi^2 dxdt \ge C_2 T\epsilon^2. \tag{53}$$

Therefore, by choosing ϵ small enough,

$$\mathcal{E}(m,w) = \int_0^T \int_{\mathbb{T}^N} mL\left(-\frac{w}{m}\right) - F(m)\,dxdt \le T(C_1\epsilon^b - C_2\epsilon^2) < 0 = \mathcal{E}(1,0)$$

Hence, the minimum of $\mathcal{E}$ is not achieved by $(1,0)$, and $(\tilde{u},\tilde{m})$ cannot be the trivial solution. □

Remark 4.2. The conclusion of Proposition 4.1 holds if one replaces the assumptions of Theorem 1.1 with the assumptions of Theorem 1.3 or Theorem 1.2. In the latter case, the non-trivial solution minimizing the energy $\mathcal{E}$ has to be intended in the weak sense.

Also condition (51) can be weakened: the construction of a minimum of $\mathcal{E}$ that is not the trivial solution can be achieved by means of a couple (μ, v) satisfying the constraint

$$\Delta\mu + \operatorname{div}(v) = 0 \text{ on } \mathbb{T}^N, \quad \int_{\mathbb{T}^N} \mu\,dx = 1, \quad \mu \ge 0,$$

such that

$$\mathcal{E}_S(\mu, v) := \int_{\mathbb{T}^N} \mu L\left(-\frac{v}{\mu}\right) - F(\mu)\,dx < L(0) - F(1).$$

In other words, this can be seen as requiring that the energy functional $\mathcal{E}_S$ associated to the *stationary* version of (50) is not minimized by the couple $(1,0)$. Such a μ has to be "connected" to the initial datum in order to serve as a competitor for the time-dependent problem (as in the proof of Proposition 4.1). If $H \in C^2$, one will be able to produce (m,w) so that $\mathcal{E}(m,w) < \mathcal{E}(1,0)$ only if T is large enough by the aforementioned uniqueness results.

A Existence for small T

To prove existence in the small time-horizon regime, we implement a standard contraction mapping principle (see, for example, [28, Chapter 15]). This tool has already been used in [1, 2] in the MFG framework to prove existence of solutions to (1), but with a different spirit: existence for arbitrarily large T but initial data close to $\bar{m} \equiv 1$, or other "smallness" conditions (in the mentioned works, the functional space setting is indeed different).

Let us rewrite (1) in integral form; set $v(\cdot,t) := u(\cdot, T-t)$ for all $t \in [0,T]$, then, by the classical Duhamel formula

$$\begin{cases} v(x,t) = e^{t\Delta}u_T(x) - \int_0^t e^{(t-s)\Delta}\Phi^v[v,m](s)(x)ds, \\ m(x,t) = e^{t\Delta}m_0(x) + \int_0^t e^{(t-s)\Delta}\Phi^m[v,m](s)(x)ds, \end{cases} \tag{54}$$

where

$$\Phi^v[v,m](s)(\cdot) := f(m(\cdot, T-s)) - H(\nabla v(\cdot,s)),$$
$$\Phi^m[v,m](s)(\cdot) := \operatorname{div}(\nabla H(\nabla v(\cdot, T-s))m(s)) \quad \forall s \in [0,T],$$

and $e^{t\Delta}$ is the (strongly continuous) semigroup associated with the parabolic equation $\varphi_t = \Delta\varphi$, defined on suitable Hölder spaces (see *iii)* below). Note that (54) has the form of a forward-forward system for v, m. Here, the *local* regularity of H, f plays a role, rather than the time direction in the two equations or the behaviour of f at infinity: we just assume $f, H \in C^3$ and do not require (2) and (4) to hold. We stress that this argument could be adapted to more general MFG systems (congestion problems, u_T depending on m, ...).

Let us define $X^{k,\nu} := C([0,T], C^{k,\nu}(\mathbb{T}^N))$ for all integers k and $\nu \in (0,1)$. We will need the following facts: let $0 < \nu < \beta < 1$, then *i)* If $h \in C^2(\mathbb{R}^N)$ and $\|g_1\|_{C^{1,\nu}(\mathbb{T}^N)}, \|g_2\|_{C^{1,\nu}(\mathbb{T}^N)} \le K$, then $\|h(g_1) - h(g_2)\|_{C^{1,\nu}(\mathbb{T}^N)} \le C\|g_1 - g_2\|_{C^{1,\nu}(\mathbb{T}^N)}$, for some $C = C(h, K, \nu) > 0$.

ii) $\|g_1 g_2\|_{C^\nu(\mathbb{T}^N)} \le C\|g_1\|_{C^\nu(\mathbb{T}^N)}\|g_2\|_{C^\beta(\mathbb{T}^N)}$ for all $g_1 \in C^\nu(\mathbb{T}^N)$, $g_2 \in C^\beta(\mathbb{T}^N)$.

iii) $\|e^{t\Delta}u\|_{C^{2,\nu}(\mathbb{T}^N)} \le Ct^{-\frac{2+\nu-\beta}{2}}\|u\|_{C^{0,\beta}(\mathbb{T}^N)}$, $\|e^{t\Delta}u\|_{C^{2,\beta}(\mathbb{T}^N)} \le Ct^{-\frac{1}{2}}\|u\|_{C^{1,\beta}(\mathbb{T}^N)}$ for all $t \in (0, 1]$.

Items *i)* and *ii)* follows by computation; as for *iii)* see, for example, [28, p. 274]. Recall also that $C^{2,\nu}(\mathbb{T}^N)$ is continuously embedded into $C^{1,\beta}(\mathbb{T}^N)$.

Remark A.1. Note that the contraction mapping principle implies also uniqueness of solutions in the set $\mathcal{Z}_a$ (see (55) below), that is, there is only one equilibrium (u, m) that is close to the final-initial data (u_T, m_0).

Proof of Theorem 1.4. Fix $0 < \nu < \beta < 1$ so that $u_T \in C^{2,\beta}(T)$ and $m_0 \in C^{2,\nu}(T)$. Then, $\Phi^v : X^{2,\beta} \times X^{2,\nu} \to X^{1,\beta}$ and $\Phi^m : X^{2,\beta} \times X^{2,\nu} \to X^{0,\beta}$. Let $a > 0$ and

$$\begin{aligned}\mathcal{Z}_a := \{(v, m) \in X^{2,\beta} \times X^{2,\nu} : v(0) = u_T,\ m(0) = m_0,\\ \|v(t) - u_T\|_{C^{2,\beta}(\mathbb{T}^N)} \le a,\ \|m(t) - m_0\|_{C^{2,\nu}(\mathbb{T}^N)} \le a \text{ for all } t \in [0, T]\}.\end{aligned} \tag{55}$$

Let $(v, m) \mapsto (\hat{v}, \hat{m}) := \Psi(v, m)$, where

$$\begin{cases} \hat{v}(t) = e^{t\Delta}u_T - \int_0^t e^{(t-s)\Delta}\Phi^v[v, m](s)ds, \\ \hat{m}(t) = e^{t\Delta}m_0 + \int_0^t e^{(t-s)\Delta}\Phi^m[v, m](s)ds. \end{cases}$$

Our aim is to prove that Ψ has a fixed point, by means of the contraction mapping theorem. First, we claim that Ψ maps $\mathcal{Z}_a$ into itself when $T = T(a)$ is small. Indeed,

$$\|e^{t\Delta}u_T - u_T\|_{C^{2,\beta}(\mathbb{T}^N)} \le a/2, \quad \|e^{t\Delta}m_0 - m_0\|_{C^{2,\nu}(\mathbb{T}^N)} \le a/2$$

if t is small, by continuity of the semigroup $e^{t\Delta}$. Moreover,

$$\begin{aligned}\left\|\int_0^t e^{(t-s)\Delta}\Phi^v[v, m](s)ds\right\|_{C^{2,\beta}(\mathbb{T}^N)} &\le \int_0^t \|e^{(t-s)\Delta}\Phi^v[v, m](s)\|_{C^{2,\beta}(\mathbb{T}^N)}ds \\ &\le C\int_0^t (t-s)^{-1/2}\|\Phi^v[v, m](s)\|_{C^{1,\beta}(\mathbb{T}^N)}ds \le C_1T^{1/2} \le a/2\end{aligned}$$

whenever T is small (here, C, C_1, ..., are positive constants depending on a, but not on T). Similarly,

$$\left\|\int_0^t e^{(t-s)\Delta}\Phi^m[v, m](s)ds\right\|_{C^{2,\nu}(\mathbb{T}^N)} \le CT^{(\beta-\nu)/2} \le a/2,$$

so $\Psi : \mathcal{Z}_a \to \mathcal{Z}_a$.

To show that Ψ is a contraction, note that for all $s \in [0, T]$,

$$\begin{aligned}&\|\Phi^m[v_1, m_1](s) - \Phi^m[v_2, m_2](s)\|_{C^{0,\beta}(\mathbb{T}^N)} \le \\ &\le \|\nabla H(\nabla v_1(T-s))(m_1 - m_2)\|_{C^{1,\beta}(\mathbb{T}^N)} + \|m_2\,[\nabla H(\nabla v_1(T-s)) - \nabla H(\nabla v_2(T-s))]\|_{C^{1,\beta}(\mathbb{T}^N)} \\ &\le C\|\nabla H(\nabla v_1(T-s))\|_{C^{1,\beta}(\mathbb{T}^N)}\|m_1 - m_2\|_{C^{1,\beta}(\mathbb{T}^N)} + \|m_2\|_{C^{1,\beta}(\mathbb{T}^N)}\|(\nabla v_1 - \nabla v_2)(T-s)\|_{C^{1,\beta}(\mathbb{T}^N)} \\ &\le C_1(\|m_1 - m_2\|_{C^{2,\nu}(\mathbb{T}^N)} + \|v_1(T-s) - v_2(T-s)\|_{C^{2,\beta}(\mathbb{T}^N)}).\end{aligned}$$

Therefore,

$$\begin{aligned}
&\left\| \int_0^t e^{(t-s)\Delta} \left(\Phi^m[v_1, m_1](s) - \Phi^m[v_2, m_2](s) \right) ds \right\|_{C^{2,\nu}(\mathbb{T}^N)} \leq \\
&\leq C \int_0^t (t-s)^{-\frac{2+\nu-\beta}{2}} (\|m_1 - m_2\|_{C^{2,\nu}(\mathbb{T}^N)} + \|v_1(T-s) - v_2(T-s)\|_{C^{2,\beta}(\mathbb{T}^N)}) ds \\
&\leq CT^{(\beta-\nu)/2} \sup_{s\in[0,T]} (\|m_1 - m_2\|_{C^{2,\nu}(\mathbb{T}^N)} + \|v_1(T-s) - v_2(T-s)\|_{C^{2,\beta}(\mathbb{T}^N)}) \\
&\leq \|m_1 - m_2\|_{X^{2,\nu}} + \|v_1 - v_2\|_{X^{2,\beta}}.
\end{aligned}$$

by eventually reducing T. In a similar way, one shows that

$$\left\| \int_0^t e^{(t-s)\Delta} \Phi^v[v_1, m_1](s) - \Phi^v[v_2, m_2](s) ds \right\|_{C^{2,\beta}(\mathbb{T}^N)} \leq \|m_1 - m_2\|_{X^{2,\nu}} + \|v_1 - v_2\|_{X^{2,\beta}},$$

hence Ψ is a contraction; its fixed point in $\mathcal{Z}_a$ is a solution to (54), and hence a classical solution to (1) by classical Schauder regularity results.

□

References

[1] D. M. Ambrose. Small strong solutions for time-dependent mean field games with local coupling. *C. R. Math. Acad. Sci. Paris*, 354(6):589–594, 2016.

[2] D. M. Ambrose. Strong solutions for time-dependent mean field games with non-separable Hamiltonians. *J. Math. Pures Appl. (9)*, 113:141–154, 2018.

[3] M. Bardi and M. Cirant. Uniqueness of solutions in mean field games with several populations and neumann conditions. *To appear in "PDE models for multi-agent phenomena" (P. Cardaliaguet, A. Porretta, F. Salvarani eds.), Springer INdAM Series*, 2018.

[4] M. Bardi and M. Fischer. On non-uniqueness and uniqueness of solutions in finite-horizon mean field games. *To appear in ESAIM Control Optim. Calc. Var.*, 2018.

[5] M. Bardi and F. S. Priuli. Linear-quadratic N-person and mean-field games with ergodic cost. *SIAM J. Control Optim.*, 52(5):3022–3052, 2014.

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

[7] A. Briani and P. Cardaliaguet. Stable solutions in potential mean field game systems. *NoDEA Nonlinear Differential Equations Appl.*, 25(1):Art. 1, 26, 2018.

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

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

[10] M. Cirant. Stationary focusing mean-field games. *Comm. Partial Differential Equations*, 41(8):1324–1346, 2016.

[11] M. Cirant. On the existence of oscillating solutions in non-monotone mean-field games. arXiv preprint, https://arxiv.org/abs/1711.08047, 2017.

[12] D. A. Gomes, L. Nurbekyan, and M. Prazeres. One-Dimensional Stationary Mean-Field Games with Local Coupling. *Dyn. Games Appl.*, 8(2):315–351, 2018.

[13] D. A. Gomes, E. A. Pimentel, and V. Voskanyan. *Regularity theory for mean-field game systems.* SpringerBriefs in Mathematics. Springer, 2016.

[14] O. Guéant. A reference case for mean field games models. *J. Math. Pures Appl. (9)*, 92(3):276–294, 2009.

[15] O. Guéant. Mean field games with a quadratic Hamiltonian: a constructive scheme. In *Advances in dynamic games*, volume 12 of *Ann. Internat. Soc. Dynam. Games*, pages 229–241. Birkhäuser/Springer, New York, 2012.

[16] 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.

[17] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Uralceva. *Linear and quasilinear equations of parabolic type.* Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968.

[18] J.-M. Lasry and P.-L. Lions. Nonlinear elliptic equations with singular boundary conditions and stochastic control with state constraints. I. The model problem. *Math. Ann.*, 283(4):583–630, 1989.

[19] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.

[20] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.

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

[22] P.-L. Lions. In cours au collége de france. *www.college-de-france.fr.*

[23] A. R. Mészáros and F. J. Silva. A variational approach to second order mean field games with density constraints: the stationary case. *J. Math. Pures Appl. (9)*, 104(6):1135–1159, 2015.

[24] G. Metafune, D. Pallara, and A. Rhandi. Global properties of transition probabilities of singular diffusions. *Theory Probab. Appl.*, 54(1):68–96, 2010.

[25] L. Nirenberg. On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 13:115–162, 1959.

[26] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications.* Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

[27] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl.*, 146(1):65–96, 1986.

[28] M. E. Taylor. *Partial differential equations III. Nonlinear equations*, volume 117 of *Applied Mathematical Sciences.* Springer, New York, second edition, 2011.

cirant@math.unipd.it
Dipartimento di Matematica “Tullio Levi-Civita”
Università di Padova
via Trieste 63, 35121 Padova (Italy)

tonon@ceremade.dauphine.fr
Université Paris-Dauphine
PSL Research University
CNRS, Ceremade
Place du Maréchal de Lattre de Tassigny
75775 Paris cedex 16 (France)