Minimax Theory and its Applications
Volume 3 (2018), No. 2, 261–284

Displacement Convexity for First-Order Mean-Field Games*

Diogo A. Gomes
King Abdullah University of Science and Technology (KAUST), CEMSE Division, Thuwal 23955-6900, Saudi Arabia
diogo.gomes@kaust.edu.sa

Tommaso Seneci
King Abdullah University of Science and Technology (KAUST), CEMSE Division, Thuwal 23955-6900, Saudi Arabia
tommaso.seneci@kaust.edu.sa

Received: July 20, 2018
Accepted: September 26, 2018

Here, we consider the planning problem for first-order mean-field games (MFG). When there is no coupling between players, MFG degenerate into optimal transport problems. Displacement convexity is a fundamental tool in optimal transport that often reveals hidden convexity of functionals and, thus, has numerous applications in the calculus of variations. We explore the similarities between the Benamou-Brenier formulation of optimal transport and MFG to extend displacement convexity methods to MFG. In particular, we identify a class of functions, that depend on solutions of MFG, that are convex in time and, thus, obtain new a priori bounds for solutions of MFG. A remarkable consequence is the log-convexity of L^q norms. This convexity gives bounds for the density of solutions of the planning problem and extends displacement convexity of L^q norms from optimal transport. Additionally, we prove the convexity of L^q norms for MFG with congestion.

Keywords: Mean field game, congestion, optimal transport, displacement convexity.

2010 Mathematics Subject Classification: 91A13, 35Q91, 26B25

1. Introduction

Displacement convexity is an alternative concept of convexity used often in minimization problems in spaces of measures. Displacement convexity was introduced in [34] to study a non-convex variation problem where it revealed a hidden convexity that gives existence and uniqueness of minimizers.

Given two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, we say that a map $T\colon \mathbb{R}^d \to \mathbb{R}^d$ *transports* μ into ν if $\nu = T_\# \mu$, where

$$\int_{\mathbb{R}^d} f(x)\, dT_\#\mu(x) = \int_{\mathbb{R}^d} f(T(x))\, d\nu(x)$$

*The authors were partially supported by KAUST baseline funds and KAUST OSR-CRG2017-3452.

ISSN 2199-1413 (printed), ISSN 2199-1421 (electronic) / \$ 2.50

for all bounded continuous $f\colon \mathbb{R}^d \to \mathbb{R}$. In optimal transport, we are given two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, and we seek to transport μ into ν in the most efficient way according to a given transport cost, see for example the surveys [39], [41], and [42]. While this problem is discrete in nature, a remarkable alternative formulation due to Benamou and Brenier [3], looks instead at paths in $\mathcal{P}(\mathbb{R}^d)$ that connect μ to ν. The Benamou-Brenier formulation of optimal transport consists of minimizing the energy functional

$$\int_{\mathbb{R}^d}\int_0^1 \rho^t(x)|v(x,t)|^2 dxdt,$$

over all smooth velocity fields $v(x,t)$, with trajectories $T_x(t)$, and densities $\rho^t = T_{(\cdot)}(t)_{\#}\mu$, such that $\rho^0 = \mu$ and $\rho^1 = \nu$. Under suitable regularity conditions, the optimality conditions of this variational problem are

$$\begin{cases} -u_t + \frac{|Du|^2}{2} = \bar{H} \\ (\rho^t)_t - \operatorname{div}(\rho^t Du) = 0 \\ \rho^t \in \mathcal{P}_{ac}(\mathbb{R}^d) \quad \forall\, t \in [0,1] \\ \rho^0 = \mu,\ \rho^1 = \nu, \end{cases} \tag{1}$$

where $v(x,t) = -D_x u(x,t)$ and $\bar{H} \in \mathbb{R}$. The value of the constant $\bar{H}$ is irrelevant and we can take it to be 0 by adding a linear function of time to u. The *displacement interpolant* between μ and ν is the minimizer of the Benamou-Brenier problem. A functional, $\mathcal{F}\colon \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$, is *displacement convex* if $t \mapsto \mathcal{F}(\rho^t)$ is convex for all displacement interpolants ρ^t. By using (1), we can differentiate twice $\mathcal{F}(\rho^t)$ to study displacement convexity. This methodology was used in [40], where the author identifies a new class of displacement convex functionals that depend on spatial derivatives of the density.

Mean-field games (MFG) model the interaction between identical rational agents, see the original papers in [27, 28] and [29, 30, 31], or the surveys [4, 6, 22, 23]. In these games, each agent minimizes a value function, which is the same for every agent. In classical MFG, agents choose their trajectories given an initial configuration and a terminal cost. In the MFG planning problem [1, 33, 37], the initial and terminal distribution of the agents are prescribed while the terminal cost is unknown. Here, we focus on the planning problem for first-order MFG. These games are given by a Hamilton-Jacobi equation coupled with a continuity equation

$$\begin{cases} -u_t + H(Du) = g(m) \\ m_t - \operatorname{div}(mD_pH(Du)) = 0 \qquad \forall\ (x,t) \in \mathbb{T}^d \times (0,T) \\ m(\cdot,0) = m^0(\cdot),\ m(\cdot,T) = m^T(\cdot). \end{cases} \tag{2}$$

Here, we use periodic boundary conditions so that the spatial domain for (2) is $\mathbb{T}^d$, the d-dimensional torus. A classical solution of (2) is a pair $(u(x,t), m(x,t)) \in$

$C^\infty(\mathbb{T}^d \times [0,T])$, such that $m(x,t) \geq 0$ and m and u are periodic in x for all time $t \in [0,T]$. The function m represents the statistical distribution of the agents in space, whereas u represents their value function. The Hamiltonian, $H(Du)$, takes into accounts the movement cost of the agents and their preferred direction of motion, and $g(m)$ determines the interactions between agents. As can be seen by comparing (1) with (2), the optimal transport problem is a special case of a first-order MFGs where the interaction between the agents does not exist; that is, $g = 0$. In the *initial-terminal value* problem, (2) is endowed with initial, $m(\cdot,0) = m^0(\cdot)$, and terminal, $u(\cdot,T) = u^T(\cdot)$, conditions; that is, agents are given a terminal cost, and their initial distribution is specified. In contrast, in the *planning problem*, m^0 and m^T, the initial and terminal distributions, are specified. Thus, our goal is to find a cost, u, that steers agents from an initial distribution, m^0, to a desired terminal distribution, m^T.

The initial-terminal value problem for second-order MFGs is now well understood. The existence and uniqueness of smooth solutions of the time dependent problem were first studied in [30, 31], and examined in detail in [33]. Subsequently, several authors considered classical [18, 19, 20, 21] and weak solutions [10, 38]. For first-order MFGs, several Sobolev regularity results were obtained in [8], [9], [11], and [26]. The planning problem was addressed from a variational numerical perspective in [1] and for second-order MFGs in [37, 38].

Here, we explore displacement convexity properties to construct a new class of estimates for first-order MFGs. In particular, the primary goal of this paper is to identify functions $U\colon \mathbb{R}_0^+ \to \mathbb{R}$ (recall that $\mathbb{R}_0^+ = [0,+\infty)$) such that

$$t \mapsto \int U(m(x,t))\,dx \quad \text{is convex,} \tag{3}$$

where $m(x,t)$ solves (2). In the case of optimal transport, (3) is displacement convex if

$$\begin{cases} P(z) = U'(z)z - U(z), \\ P \in C^1(\mathbb{R}_0^+), \; P(z) \geq 0, \\ \frac{P(z)}{z^{1-\frac{1}{d}}} \quad \text{non-decreasing.} \end{cases} \tag{4}$$

The convexity of the preceding functional gives the following a priori bound:

$$\int U(m(x,t))dx \leq \frac{t}{T}\int U(m(x,T))dx + \left(1-\frac{t}{T}\right)\int U(m(x,0))dx,$$

which are particularly interesting in the case of the planning problem because $m(x,0)$ and $m(x,T)$ are known.

In Section 3, we prove the following result on the convexity of functionals that depend on the density of solutions of first-order MFGs, as in (3).

Theorem 1.1. *Let $m, u \in C^\infty(\mathbb{T}^d \times [0,T])$, $m \geq 0$, be periodic solutions of the first-order MFG*

$$\begin{cases} -u_t + H(Du) = g(m) \\ m_t - \operatorname{div}(mD_pH(Du)) = 0 \end{cases} \tag{5}$$

with $g\colon \mathbb{R}_0^+ \to \mathbb{R}$, $H\colon \mathbb{R}^d \to \mathbb{R}$ smooth, g non-decreasing, and H convex. If $U\colon \mathbb{R}_0^+ \to \mathbb{R}$ is such that (4) holds, then

$$t \mapsto \int_{\mathbb{T}^d} U(m(x,t))\, dx \quad \text{is convex.}$$

Functionals of the form $t \mapsto \int_{\mathbb{T}^d} m(x,t)^q dx$ satisfy the conditions of the preceding theorem. Moreover, a careful computation reveals the convexity of $t \to \ln(\|m(\cdot,t)\|_{L^q(\mathbb{T}^d)})$ for all $1 \leq q \leq \infty$, see Proposition 3. Furthermore, this log-convexity generalizes the result in [34] concerning the displacement convexity of $\rho \mapsto \int \rho(x)^q dx$. Here, we should also mention the recent work [32] where, using a discretization method and ideas that are reminiscent from displacement convexity, the authors prove additional regularity for mean-field games.

MFGs with congestion model the case where the agents' displacement cost increases in high-density regions. These games correspond to the system

$$\begin{cases} -u_t + m^\alpha H\left(\frac{Du}{m^\alpha}\right) = g(m) \\ m_t - \operatorname{div}\left(mD_pH\left(\frac{Du}{m^\alpha}\right)\right) = 0 & \forall\ (x,t) \in \mathbb{T}^d \times (0,T) \\ m(\cdot,t) \in \mathcal{P}_{ac}(\mathbb{R}^d) & \forall\ t \in (0,T) \\ m(\cdot,0) = m^0(\cdot),\ m(\cdot,T) = m^T(\cdot) \end{cases}$$

for $\alpha > 0$. The existence and uniqueness of solutions of second-order classical MFG with congestion were studied in [16, 18] in the stationary case and in [2, 25] in the non-stationary case. First-order MFG with congestion were studied in the stationary case in [15] , [36], and [17] and in the time-dependent case, for the forward-forward model, in [24]. In particular, as far as the authors are aware the planning problem was not studied previously. Here, in Section 4, we examine the case where $H(p) = |p|^\beta/\beta$ and, in Theorem 4.1, prove the convexity of $t \mapsto \int_{\mathbb{T}^d} m(x,t)^p dx$, p depending on α and β. As an application, we obtain L^∞ bounds for the density in Corollary 4.2.

2. Preliminaries

Here, we briefly review the optimal transport problem and the Benamou-Brenier formulation. Subsequently, we recall displacement convexity and discuss elementary examples.

2.1. Optimal transport

Let $\mathcal{P}(\mathbb{R}^d)$ be the set of probability measures in $\mathbb{R}^d$, and $\mathcal{P}_{ac}(\mathbb{R}^d)$ the subset of those probabilities that are absolutely continuous with respect to the Lebesgue measure.

The *optimal transport problem*, also known as the *Monge-Kantorovich problem*, studies the optimal way of moving mass between two different locations. We are given an initial distribution of mass determined by a probability measure, $\mu \in \mathcal{P}(\mathbb{R}^d)$, and a target distribution given by another probability measure, $\nu \in \mathcal{P}(\mathbb{R}^d)$. To each unit of mass moved from $x \in \mathbb{R}^d$ to $y \in \mathbb{R}^d$, we associate a cost, $c(x, y)$. The Monge-Kantorovich problem consists of minimizing the total cost,

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y)\, d\pi(x, y), \tag{6}$$

over the set of plans $\Pi[\mu, \nu] = \{\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \text{ with marginals } \mu, \nu\}$. If $c(x, y)$ is positive and lower semi-continuous, there exist a minimizer of (6), see [41]. For a quadratic cost, $c(x, y) = |x - y|^2$, a duality formulation due to Kantorovich uncovers remarkable properties of the optimal plan. If $\mu, \nu \in \mathcal{P}_{ac}(\mathbb{R}^d)$ and have finite second-order moments, the minimizing plan is unique and can be written in the form

$$\pi = (Id \times D\phi)_{\#}\mu,$$

where $D\phi$ is the unique gradient of a convex function such that $\nu = D\phi_{\#}\mu$; that is, for every $E \subset \mathbb{R}^d$ measurable,

$$\nu(E) = (D\phi_{\#}\mu)(E) = \mu((D\phi)^{-1}(E)).$$

Thus, the minimum of (6) equals to

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - D\phi(x)|^2\, d\mu(x). \tag{7}$$

In the literature, $D\phi$ is called the Brenier's map transporting μ into ν, see [5].

2.2. The Benamou-Brenier formulation

In the time-dependent optimal transport problem, each particle moves from μ to ν according to a piecewise C^1 trajectory $T_x(t)\colon [0, 1] \to \mathbb{R}^d$. At $t = 0$, $T_x(0) = x$, and at time $t = 1$ particles reach their destination in $\operatorname{supp}(\nu)$. Accordingly, we require $\nu = T_{(\cdot)}(1)_{\#}\mu$. The time-dependent optimal transport problem consists of minimizing a displacement cost $C = C(T_x(\cdot))$ over all trajectories $\{T_x(\cdot)\}_{x \in \mathbb{R}^d}$ transporting μ into ν, i.e.

$$\inf \left\{ \int_X C(\{T_x(t)\}_{t \in (0,1)})\, d\mu(x) : T_x(0) = x,\ T_{(\cdot)}(1)_{\#}\mu = \nu \right\}.$$

An important case is given by differential cost function

$$C(\{T_x(t)\}_{t \in (0,1)}) = \int_0^1 c(\dot{T}_x(t))\, dt,$$

where c is a convex function. Thanks to Jensen's inequality, we find

$$\int_0^1 c(\dot{T}_x(t))\,dx \geq c\left(\int_0^1 \dot{T}_x(t)\,dx\right) = c(y-x). \tag{8}$$

For $c(x) = |x|^2$, by comparing (8) with (7), we see that straight lines are admissible trajectories. Hence, they are minimizers. Thus, the optimal velocities are $x - D\phi(x)$, $D\phi(x)$ being the Brenier's map transporting μ into ν. This means that the minimizing straight lines are

$$T_x(t) = (1-t)x + tD\phi(x). \tag{9}$$

At each time $t \in [0,1]$, μ is transported into

$$\rho^t = ((1-t)x + tD\phi(x))_{\#}\mu.$$

The previous discussion suggests we move our perspective to the Eulerian point of view. For that, we fix $x \in \mathbb{R}^d$ and consider a smooth trajectory $T_x(t)$ determined by a Lipschitz velocity field $v(x,t)$; that is,

$$\begin{cases} \dot{T}_x(t) = v(T_x(t),(t)) \\ T_x(0) = x. \end{cases}$$

If $\{T_{(\cdot)}(t)\}_{0\leq t\leq T}$ is a Lipschitz family of diffeomorphisms, the pushforward $\rho^t = T_{(\cdot)}(t)_{\#}\mu$ is the unique solution of the *continuity equation*

$$\frac{\partial \rho^t}{\partial t} + \operatorname{div}(\rho^t v) = 0 \tag{10}$$

in the weak sense. We look for a path ρ^t that minimizes the total action

$$A[\rho,v] = \int_0^1 E(t)dt = \int_0^1 \int_{\mathbb{R}^d} \rho^t(x)\frac{|v(x,t)|^2}{2}\,dx\,dt. \tag{11}$$

As in [3], if $\mu,\nu \in P_{ac}(\mathbb{R}^d)$ are compactly supported and satisfy suitable conditions [41], then

$$\inf_{\pi\in\Pi[\mu,\nu]} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 d\pi(x,y) = \inf_{\rho,v} A[\rho,v],$$

where the infimum on the r.h.s is taken over all smooth (ρ,v) solving (10) with $\rho^0 = \mu$ and $\rho^1 = \nu$. The optimality conditions of this problem correspond to (1), as we describe next. We already have a partial differential equation solved by the density ρ^t. Moreover, if $\{T_x(\cdot)\}_{x\in\mathbb{R}^d}$ are constant speed trajectories such that $x \mapsto T_x(t)$ are diffeomorphisms for all t, then $v(x,t)$ solves

$$\frac{\partial v}{\partial t} + v \cdot Dv = 0.$$

Because of (9), it turns out that v is a gradient. Thus, $v = -D_x u$. Consequently, u solves a Hamilton-Jacobi equation

$$-\frac{\partial u}{\partial t} + \frac{|Du|^2}{2} = \bar{H}, \;\; \bar{H} \in \mathbb{R}. \tag{12}$$

If we combine (12) with (10), we get the following system

$$\begin{cases} -\frac{\partial u}{\partial t} + \frac{|Du|^2}{2} & = \bar{H} \\ \frac{\partial \rho^t}{\partial t} - \operatorname{div}(\rho^t Du) & = 0. \end{cases} \tag{13}$$

Because displacement interpolants are constant speed trajectories, (13) are the corresponding optimality conditions.

The system (13) has a triangular structure. The first equation does not depend on m, while the second one depends on Du. First-order MFGs are recovered by adding a coupling $g = g(m)$ to the Hamilton-Jacobi equation. Due to this coupling, MFGs no longer have this triangular structure, and, thus, their study becomes substantially more challenging.

2.3. Displacement convexity

Displacement convexity was introduced by McCann in [34] to study a non-convex variational problem from the theory of interacting gases. In that problem, the gas density is determined by a probability, $\rho \in P_{ac}(\mathbb{R}^d)$. Each particle is subject to two forces: one given by an interaction potential $W(x-y)$ that increases with the distance between particles, and the other given by the internal energy, $U(z)$. The potential is

$$\mathcal{W}(\rho) = \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} W(x-y)\, d\rho(x)\, d\rho(y),$$

and the internal energy is

$$\mathcal{U}(\rho) = \int_{\mathbb{R}^d} U(\rho(x))dx.$$

In that model, the configuration of the gas minimizes the energy

$$E(\rho) = \mathcal{U}(\rho) + \mathcal{W}(\rho).$$

Given the variational nature of the problem, the convexity of E is of paramount importance. If U is convex, then $\mathcal{U}$ is also convex. However, convexity of W does not imply the convexity of $\mathcal{W}$.

A fundamental contribution in [34] is a new way of interpolating two probabilities densities, $\mu, \nu \in \mathcal{P}_{ac}(\mathbb{R}^d)$, that reveals a hidden convexity in $\mathcal{U}$ and $\mathcal{W}$. For a given family of trajectories $\{T_{(\cdot)}(t)\}_{t \in (0,1)}$, $T_{(\cdot)}(t) \colon \mathbb{R}^d \to \mathbb{R}^d$, we set $\rho^t = T_{(\cdot)}(t)_{\#}\rho$. Thus,

$$\begin{aligned}\mathcal{W}(\rho^t) &= \iint W(x-y)\, d\rho^t(x)\, d\rho^t(y)\\ &= \iint W(x-y)\, d(T_{(\cdot)}(t)_{\#}\rho)(y)\, d(T_{(\cdot)}(t)_{\#}\rho)(x)\\ &= \iint W(T_x(t)-T_y(t))\, d\rho(y)\, d\rho(x).\end{aligned}$$

Therefore, if $T_x(t)$ is linear in time, the map $t \mapsto \mathcal{W}(\rho^t)$ is convex.

Definition 2.1. (1) Let $\mu, \nu \in \mathcal{P}_{ac}(\mathbb{R}^d)$ and $D\phi\colon \mathbb{R}^d \to \mathbb{R}^d$ the unique gradient of a convex function such that $\nu = D\phi_{\#}\mu$. The *displacement interpolant* between μ and ν is

$$\rho^t = ((1-t)(\cdot) + tD\phi(\cdot))_{\#}\mu.$$

(2) A function $\mathcal{F}\colon \mathcal{P}_{ac}(\mathbb{R}^d) \to \mathbb{R}$ is *displacement convex* if it is convex along displacement interpolants; that is,

$$t \mapsto \mathcal{F}(\rho^t) \text{ is convex for all } \rho^t \text{ displacement interpolants.}$$

As we have seen, the map $t \mapsto \mathcal{W}(\rho^t)$ is convex; that is, $\rho \mapsto \mathcal{W}(\rho)$ is displacement convex. In general, even if U is convex, $\mathcal{U}$ may not be displacement convex. However, the following condition states the required convexity, as proven in [34]: if

$$z \mapsto z^d U(z^{-d}), z \in \mathbb{R}^+ \quad \text{is convex, non-increasing and } \; U(0) = 0, \tag{14}$$

then
$$t \mapsto \mathcal{U}(\rho^t) = \int U(\rho^t(x))\, dx \quad \text{is convex.}$$

In [34], the author derives conditions equivalent to (14) for U sufficiently smooth in terms of the pressure

$$P(z) = U'(z)z - U(z).$$

By differentiating twice $z \mapsto z^d U(z^{-d})$ and using the preceding identity into the resulting expression, we conclude that for $U \in C^1(\mathbb{R}_0^+)$ and if P satisfies (4), then $\rho \mapsto \int U(\rho)$ is displacement convex. Notice that P' is non-negative, as we recover by differentiating $\frac{P(z)}{z^{1-\frac{1}{d}}}$,

$$zP'(z) \geq \left(1 - \frac{1}{d}\right) P(z) \geq 0.$$

Consequently, we can differentiate P to show that the above condition implies the convexity of U:

$$P'(z) = U''(z)z + U'(z) - U'(z) = U''(z)z \geq 0.$$

Here, we use an alternative approach explored in [40] to study displacement convexity. Formally, because displacement interpolants are solutions of the

Benamou-Brenier problem, (13), to check displacement convexity, it is enough to prove that

$$\frac{d^2}{dt^2}\int U(\rho^t(x))\,dx \geq 0 \tag{15}$$

for all (ρ^t, u) smooth solutions of (13). Because first-order MFGs are recovered by coupling the Hamilton-Jacobi equation in (13), differentiating (15) for $\rho^t \equiv m$ solving (13) may lead to similar displacement convexity inequalities. In Section 3, we prove that this holds provided that the coupling $g(m)$ is non-decreasing and $H(p)$ is convex.

2.4. Extensions of displacement convexity

Several authors have investigated extensions of displacement convexity to other settings. For example, in [14], the Benamou-Brenier functional, (11), is replaced by

$$\tilde{A}[\rho, v] = \int_0^1 E(t)dt = \int_0^1 \int_{\mathbb{R}^d} m(\rho^t(x))\frac{|v(x,t)|^2}{2}dxdt. \tag{16}$$

The preceding functional gives rise to a class of dynamic distances where the mobility depends on the density. These models are natural in chemotaxis, where they prevent overcrowding, and also appear in phase segregation and thin-film models. The optimality conditions corresponding to (16) become

$$\begin{cases} -\frac{\partial u}{\partial t} + \frac{m'(\rho)}{2}|Du|^2 = 0 \\ \frac{\partial \rho}{\partial t} - \operatorname{div}(m(\rho)Du) = 0. \end{cases}$$

In [12], this formulation is used to characterize the displacement convexity of various functionals. As a by-product of that displacement convexity, they obtain existence and stability results for a non-linear diffusion equation. The results in [14] were further refined in [7], where the connection between this problem and mean-field games was also observed.

Until recently, displacement convexity was only established for functionals that depend on the density and not on their derivatives. The first example of a higher-order displacement convexity functional was presented in [13]. These results were further extended in [40].

3. Displacement convexity in first-order mean-field games

Here, we prove that, if U satisfies (4), then $t \mapsto \int_{\mathbb{T}^d} U(m(x,t))dx$ is convex, where (u, m) solves (5). We end this section by examining the one-dimensional case, where more precise results can be proven.

3.1. Convex functionals for first-order mean-field games

Here, we prove our main result, Theorem 1.1, that extends displacement convexity to MFG. However, we first need a small lemma.

Lemma 3.1. *Let $A, B \in \mathbb{R}^{d\times d}$ be symmetric matrices with A positive semidefinite, then*

$$\operatorname{tr}\left((AB)^2\right) \geq \frac{1}{d}\operatorname{tr}(AB)^2.$$

Proof. Notice that, if A is symmetric positive semidefinite, then $A_\varepsilon = A + \varepsilon\mathcal{I}$ is symmetric positive definite and converges to A as $\varepsilon \to 0$. By approximation, it is then enough to prove the lemma for A positive definite.

Since A is symmetric positive definite, it admits an invertible square root $A^{\frac{1}{2}}$. Multiplying AB by $A^{-\frac{1}{2}}$ on the left and by $A^{\frac{1}{2}}$ on the right, we conclude that

$$A^{-\frac{1}{2}}(AB)A^{\frac{1}{2}} = A^{\frac{1}{2}}BA^{\frac{1}{2}};$$

that is, AB is similar to a symmetric matrix and, thus, it is diagonalizable. Accordingly, we have $AB = S^{-1}\Lambda S$, where $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_d)$. Then,

$$\operatorname{tr}(AB)^2 = (\sum_i \lambda_i)^2 = \sum_{i,j}\lambda_i\lambda_j \leq \sum_{i,j}\frac{\lambda_i^2}{2} + \frac{\lambda_j^2}{2} = d\sum_i \lambda_i^2 = \operatorname{tr}\left((AB)^2\right). \quad \square$$

Proof of Theorem 1.1. We begin by the following computation

$$\begin{aligned}
\frac{d}{dt}\int U(m) &= \int U'(m)m_t = \int U'(m)\operatorname{div}(mD_pH) \\
&= \int U'(m)m\operatorname{div}(D_pH) + U'(m)DmD_pH \\
&= \int U'(m)m\operatorname{div}(D_pH) + D(U(m))D_pH \\
&= \int U'(m)m\operatorname{div}(D_pH) - U(m)\operatorname{div}(D_pH) \\
&= \int P(m)\operatorname{div}(D_pH),
\end{aligned}$$

where $P(m)$ is given by (4). Differentiating again, we obtain

$$\begin{aligned}
\frac{d^2}{dt^2}\int U(m) &= \int P'(m)m_t\operatorname{div}(D_pH) + P(m)\operatorname{div}(D_pH_t) \\
&= \int P'(m)\operatorname{div}(mD_pH)\operatorname{div}(D_pH) + P(m)\operatorname{div}(D^2_{pp}HDu_t) \\
&= \int \overbrace{P'(m)m\operatorname{div}(D_pH)^2}^{A} + \overbrace{P'(m)DmD_pH\operatorname{div}(D_pH)}^{B} \\
&\quad + \overbrace{P(m)\operatorname{div}(D^2_{pp}HD(H))}^{C} - \overbrace{P(m)\operatorname{div}(g'(m)D^2_{pp}HDm)}^{D}.
\end{aligned}$$

We want to generalize Lemma 5.43 in [41]; thus, we expect to get an inequality of the form

$$\frac{d^2}{dt^2}\int U(m) \geq \int \overbrace{mP'(m)\operatorname{div}(D_pH)^2}^{A'} - \overbrace{P(m)\operatorname{div}(D_pH)^2}^{B'} + \overbrace{\frac{P(m)}{d}\operatorname{div}(D_pH)^2}^{C'}\,dx + \text{ non-negative terms.}$$

The first three terms, A', B' and C', correspond to the optimal transport case; that is, $g = 0$. Due to conditions (4), $A' + B' + C' \geq 0$. We note that $A = A'$. Next, we integrate by parts B to get

$$\begin{aligned} B &= \int P'(m)DmD_pH\operatorname{div}(D_pH) = \int D(P(m))D_pH\operatorname{div}(D_pH) \\ &= -\int P(m)\operatorname{div}(D_pH\operatorname{div}(D_pH)) \\ &= \int -\overbrace{P(m)\operatorname{div}(D_pH)^2}^{B'} - \overbrace{P(m)D_pHD(\operatorname{div}(D_pH))}^{Q}. \end{aligned} \tag{17}$$

To simplify C, we compute

$$\begin{aligned} \operatorname{div}(D^2_{pp}HD(H)) &= \operatorname{div}(D^2_{pp}HD^2uD_pH) = \operatorname{div}(D(D_pH)D_pH) \\ &= ((H_{p_i})_{x_j}H_{p_j})_{x_i} = (H_{p_i})_{x_i,x_j}H_{p_j} + (H_{p_i})_{x_j}(H_{p_j})_{x_i} \\ &= D(\operatorname{div}(D_pH))D_pH + \operatorname{tr}((D(D_pH))^2). \end{aligned} \tag{18}$$

Then, we expand C as follows

$$\begin{aligned} C &= \int P(m)\operatorname{div}(D^2_{pp}HD(H)) \\ &= \int \overbrace{P(m)D(\operatorname{div}(D_pH))D_pH}^{Q} + P(m)\operatorname{tr}((D(D_pH))^2) \end{aligned}$$

and notice that Q cancels the corresponding term in (17). Finally, D is

$$D = \int(-P(m)\operatorname{div}(D^2_{pp}HD(g(m)))) = \int P'(m)g'(m)DmD^2_{pp}HDm.$$

According to the preceding identities, we get

$$\begin{aligned} \frac{d^2}{dt^2}\int U(m) = \int & \overbrace{P'(m)m\operatorname{div}(D_pH)^2}^{A'} - \overbrace{P(m)\operatorname{div}(D_pH)^2}^{B'} \\ & + P(m)\operatorname{tr}((D(D_pH))^2) + P'(m)g'(m)DmD^2_{pp}HDm. \end{aligned}$$

Because $D(D_pH) = D_{pp}HD^2u$ is the product of a positive semidefinite matrix and a symmetric matrix, Lemma 3.1 implies

$$\operatorname{tr}((D(D_pH))^2) \geq \frac{1}{d}\operatorname{tr}(D(D_pH))^2 = \frac{1}{d}\operatorname{div}(D_pH)^2.$$

Since P is non-negative, we obtain

$$\int P(m)\operatorname{tr}((D(D_pH))^2) \geq \int \overbrace{\frac{1}{d}P(m)\operatorname{div}(D_pH)^2}^{C'}.$$

Finally, in view of the preceding identities, the last expression becomes

$$\begin{aligned}\frac{d^2}{dt^2}\int U(m) \geq \int \left(P'(m)m - P(m) + \frac{1}{d}P(m)\right)\operatorname{div}(D_pH)^2 \\ + P'(m)g'(m)DmD^2_{pp}HDm \geq 0,\end{aligned} \tag{19}$$

which is convex because (4) holds, because $g'(m) \geq 0$, and because $H(p)$ is convex. □

3.2. L^q Estimates

In the previous section, we identified conditions on U such that $t \mapsto \int U(m(x,t))dx$ is convex when (u,m) solves a first-order MFG. The function $U(z) = z^q$ satisfies (4) for all $1 \leq q < \infty$. Here, we refine this result and prove the log-convexity of the L^q norms of the density.

Proposition 3.2. *Let $u, m \in C^\infty(\mathbb{T}^d \times [0,T])$ be periodic solutions of* (5) *with g, H smooth, g non-decreasing, and H convex. Then, for all $1 \leq q \leq \infty$,*

$$\|m(\cdot,t)\|_{L^q(\mathbb{T}^d)} \leq \|m^0(\cdot)\|^{1-\frac{t}{T}}_{L^q(\mathbb{T}^d)}\|m^T(\cdot)\|^{\frac{t}{T}}_{L^q(\mathbb{T}^d)}, \quad \forall\, t \in [0,T]. \tag{20}$$

Proof. First of all, notice that if f is smooth and positive, then $\ln f$ is convex if and only if

$$(\ln f)'' = \left(\frac{f'}{f}\right)' = \frac{f''f - (f')^2}{f^2} \geq 0; \tag{21}$$

that is,

$$f''f \geq (f')^2. \tag{22}$$

First, we consider the case $1 \leq q < \infty$. We begin by computing

$$P(z) = U'(z)z - U(z) = (qz^{q-1})z - z^q = (q-1)z^q.$$

Then, plug $U(z) = z^q$ into (19) to get

$$\begin{aligned}\frac{d^2}{dt^2}\int m(x,t)^q \geq \int \left(q - 1 + \frac{1}{d}\right)(q-1)m^q\operatorname{div}(D_pH)^2 + \\ + q(q-1)m^{q-1}g'(m)DmD^2_{pp}HDm \geq (q-1)^2\int m^q\operatorname{div}(D_pH)^2.\end{aligned}$$

Thus,
$$\left(\frac{d}{dt}\int m^q\right)^2 = \left((q-1)\int m^q\,\mathrm{div}(D_pH)\right)^2$$
$$\leq (q-1)^2\left(\int m^q\right)\left(\int m^q\,\mathrm{div}(D_pH)^2\right) \leq \left(\int m^q\right)\left(\frac{d^2}{dt^2}\int m^q\right).$$

The preceding inequality combined with (22) shows that $\ln\left(\int m^q\right)$ is convex. Therefore,

$$\ln\left(\int m(x,t)^q\right) \leq \left(1-\frac{t}{T}\right)\ln\left(\int m^0(x)^q\right) + \frac{t}{T}\ln\left(\int m^T(x)^q\right)$$
$$= \ln\left(\left(\int m^0(x)^q\right)^{1-\frac{t}{T}}\left(\int m^T(x)^q\right)^{\frac{t}{T}}\right).$$

Therefore,
$$\int m(x,t)^q \leq \left(\int m^0(x)^q\right)^{1-\frac{t}{T}}\left(\int m^T(x)^q\right)^{\frac{t}{T}}.$$

Exponentiating the previous inequality to $\frac{1}{q}$ to obtain the result.

Finally, we address the case $q=\infty$. Because $\mathcal{L}^d(\mathbb{T}^d) = 1 < \infty$, we can pass to the limit in (20) as $q \to \infty$ to derive the estimate for the supremum. □

Remark 3.3. For $g(m) = 0$ and $H(p) = \frac{|p|^2}{2} + \bar{H}$, $\bar{H} \in \mathbb{R}$, solutions of (5) are displacement interpolants between the initial density, m^0, and the terminal density, m^T. Therefore, Proposition 3.2 gives both the log convexity of L^q norms and L^∞ bounds for the optimal transport problem, provided the initial and terminal densities are bounded.

For certain choices of g and H, the preceding estimate can be improved even further if $1 < q < \infty$, For example, here, we examine the case $g'(m) \geq Cm^\alpha, C > 0, \alpha \in \mathbb{R}$ and H uniformly convex. When $\alpha < 0$, we assume $m(x,t) > 0$ everywhere.

Lemma 3.4. *Let $m(x,t)$ be as in Theorem 1.1 and suppose that*

$$\int_{\mathbb{T}^d} m(x,t)\,dx = 1$$

for all $t \in [0,T]$. Assume also that $\|m^0\|_{L^q}, \|m^T\|_{L^q} > 1, g'(m) \geq Cm^\alpha, C > 0, \alpha \in \mathbb{R}$ and H is uniformly convex. Then, for all $1 < q < \infty$,

$$\|m(\cdot,t)\|_{L^q(\mathbb{T}^d)} < \|m^0(\cdot)\|_{L^q(\mathbb{T}^d)}^{1-\frac{t}{T}}\|m^T(\cdot)\|_{L^q(\mathbb{T}^d)}^{\frac{t}{T}}, \quad \forall\, t \in (0,T).$$

Proof. We select $f(t)=\int m(x,t)^q$, to which corresponds $P(z)=(q-1)z^q$, and use (19) and (21) to get the inequality

$$\frac{d^2}{dt^2}\ln\left(\int m^q\right)\geq\frac{\left(\int m^q\right)\left(\int P'(m)g'(m)DmD^2_{pp}HDm\right)}{\left(\int m^q\right)^2}$$

$$\geq C\,\frac{\int m^{q-1+\alpha}|Dm|^2}{\int m^q}=\begin{cases}C\,\frac{\int|D(m^{\frac{q+1+\alpha}{2}})|^2}{\int m^q}, & \alpha\neq -q-1\\ C\,\frac{\int|D\ln(m)|^2}{\int m^q}, & \alpha=-q-1\end{cases}$$

Because m integrates to 1, Jensen's inequality implies $\|m(\cdot,t)\|_{L^q(\mathbb{T}^d)}=1$ if and only if $m(x,t)=1$ for all $x\in\mathbb{T}^d$. Therefore, $\|m(\cdot,\bar{t})\|_{L^p}>1$ if and only if $\|m(\cdot,t)\|_{L^q(\mathbb{T}^d)}$ is strictly convex in a neighborhood of $\bar{t}$. Because $\|m^0\|_{L^p}>1$, then $t\mapsto\|m(\cdot,t)\|_{L^p}$ is strictly convex in a neighborhood of 0. Analogously, $t\mapsto\|m(\cdot,t)\|_{L^p}$ is strictly convex in a neighborhood of $t=T$. Therefore, the inequality in (20) is strict for all $t\in(0,T)$. □

3.3. Convexity in dimension 1

Finally, we address the one-dimensional case, $d=1$. A direct computation shows that the convexity of U implies the convexity of $t\mapsto\int_0^1 U(m(x,t))dx$. Accordingly, convexity holds for functions of the form $U(z)=(z+\varepsilon)^{-q}$, $q\geq 0,\varepsilon>0$; that is,

$$\int_0^1\frac{1}{(m(x,t)+\varepsilon)^q}\,dx\leq\left(1-\frac{t}{T}\right)\int_0^1\frac{1}{(m^0(x)+\varepsilon)^q}dx+\frac{t}{T}\int_0^1\frac{1}{(m^T(x)+\varepsilon)^q}\,dx.$$

Now, raising both sides to the power $\frac{1}{q}$ and bounding the r.h.s, we get

$$\begin{aligned}\|(m(\cdot,t)+\varepsilon)^{-1}\|_{L^q}&\leq\left(\left(1-\frac{t}{T}\right)\int_0^1\frac{1}{(m^0(x)+\varepsilon)^q}\,dx+\frac{t}{T}\int_0^1\frac{1}{(m^T(x)+\varepsilon)^q}\,dx\right)^{\frac{1}{p}}\\ &\leq\max\left\{\int_0^1\frac{1}{(m^0(x)+\varepsilon)^q}\,dx,\int_0^1\frac{1}{(m^T(x)+\varepsilon)^q}\,dx\right\}^{\frac{1}{q}}\\ &=\max\{\|(m^0(\cdot)+\varepsilon)^{-1}\|_{L^q},\|(m^T(\cdot)+\varepsilon)^{-1}\|_{L^q}\}.\end{aligned}$$

By letting $\varepsilon\to 0$ and then $q\to\infty$, we get

$$\|m(\cdot,t)^{-1}\|_{L^\infty}\leq\max\{\|m^0(\cdot)^{-1}\|_{L^\infty},\|m^T(\cdot)^{-1}\|_{L^\infty}\}.$$

Finally, we invert the above inequality to get quasi-concavity for the infimum

$$\inf m(\cdot,t)\geq\min\{\inf m^0(\cdot),\inf m^T(\cdot)\}.$$

4. Extension to first-order MFG with congestion

In MFG with congestion, the Hamiltonian-Jacobi equation depends on the inverse of the density, $m(x,t)$. Here, we study MFGs with Hamiltonians $H(p) = \frac{|p|^\beta}{\beta}$ and with a congestion exponent $\alpha > 0$:

$$\begin{cases} -u_t + m^{\alpha(1-\beta)}\frac{|Du|^\beta}{\beta} = g(m) \\ m_t - \operatorname{div}\left(m^{1+\alpha(1-\beta)}Du|Du|^{\beta-2}\right) = 0 \quad (x,t) \in \mathbb{T}^d \times (0,T). \end{cases} \tag{23}$$

with $g\colon \mathbb{R}^+ \to \mathbb{R}$ smooth and non-decreasing. A similar but distinct system was examined in [12] in the context of nolinear mobility continuity equations (also see [7]). The analog of (23) for $\beta = 2$ is the system

$$\begin{cases} -u_t + (1-\alpha)m^{-\alpha}\frac{|Du|^2}{2} = g(m) \\ m_t - \operatorname{div}\left(m^{1-\alpha}Du\right) = 0, \end{cases}$$

whereas the system we study here is

$$\begin{cases} -u_t + m^{-\alpha}\frac{|Du|^2}{2} = g(m) \\ m_t - \operatorname{div}\left(m^{1-\alpha}Du\right) = 0. \end{cases}$$

The authors of [12] established the convexity of similar functionals

Theorem 4.1. *Let $m, u \in C^\infty(\mathbb{T}^d \times [0,T])$, $m > 0$, be periodic solutions of the first-order MFG with congestion,* (23)*. If*

$$\beta \geq 2, \quad q + 2\alpha(1-\beta) \geq 0 \quad \textit{and} \quad 1 - \frac{1-\frac{1}{d}}{q+2\alpha(1-\beta)} - \frac{\alpha(\beta-1)}{2} \geq 0 \tag{24}$$

or

$$1 < \beta < 2, \quad q + 2\alpha(1-\beta) \geq 0 \quad \textit{and} \quad 1 - \frac{1-\frac{1}{d}}{q+2\alpha(1-\beta)} - \frac{\alpha}{2} \geq 0, \tag{25}$$

then

$$t \mapsto \int_{\mathbb{T}^d} m(x,t)^q dx \quad \textit{is convex.} \tag{26}$$

Proof. First, we compute

$$\begin{aligned}
&\frac{d}{dt}\int m^q = q\int m^{q-1}m_t = q\int m^{q-1}\operatorname{div}(m^{1+\alpha(1-\beta)}Du|Du|^{\beta-2}) \\
&= q\int m^{q-1}\Big((1+\alpha(1-\beta))m^{\alpha(1-\beta)}DmDu|Du|^{\beta-2} + m^{1+\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})\Big) \\
&= q\int (1+\alpha(1-\beta))m^{q-1+\alpha(1-\beta)}DmDu|Du|^{\beta-2} + m^{q+\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2}) \\
&= q\int \frac{(1+\alpha(1-\beta))}{q+\alpha(1-\beta)}D(m^{q+\alpha(1-\beta)})Du|Du|^{\beta-2} + m^{q+\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})
\end{aligned}$$

$$= q\int\left(1-\frac{1+\alpha(1-\beta)}{q+\alpha(1-\beta)}\right)m^{q+\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})$$

$$= \int\frac{q(q-1)}{q+\alpha(1-\beta)}m^{q+\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2}).$$

Thus, we have $\displaystyle\frac{1}{q(q-1)}\frac{d^2}{dt^2}\int m^q =$

$$= \int\overbrace{m^{q+\alpha(1-\beta)-1}m_t\operatorname{div}(Du|Du|^{\beta-2})}^{A}+\overbrace{\frac{1}{q+\alpha(1-\beta)}m^{q+\alpha(1-\beta)}\operatorname{div}((Du|Du|^{\beta-2})_t)}^{B}.$$

Now, we expand and integrate by parts

$$\begin{aligned}
A &= \int m^{q+\alpha(1-\beta)-1}m_t\operatorname{div}(Du|Du|^{\beta-2})\\
&= \int m^{q+\alpha(1-\beta)-1}\operatorname{div}(m^{1+\alpha(1-\beta)}Du|Du|^{\beta-2})\operatorname{div}(Du|Du|^{\beta-2})\\
&= \int m^{q+\alpha(1-\beta)-1}\operatorname{div}(Du|Du|^{\beta-2})\cdot\\
&\qquad\cdot\Big((1+\alpha(1-\beta))m^{\alpha(1-\beta)}DmDu|Du|^{\beta-2}+m^{1+\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})\Big)\\
&= \int(1+\alpha(1-\beta))m^{q+2\alpha(1-\beta)-1}DmDu|Du|^{\beta-2}\operatorname{div}(Du|Du|^{\beta-2})+\\
&\qquad\qquad+m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&= \int\frac{(1+\alpha(1-\beta))}{q+2\alpha(1-\beta)}D(m^{q+2\alpha(1-\beta)})Du|Du|^{\beta-2}\operatorname{div}(Du|Du|^{\beta-2})+\\
&\qquad\qquad+m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&= \int-\frac{(1+\alpha(1-\beta))}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\operatorname{div}\Big(Du|Du|^{\beta-2}\operatorname{div}(Du|Du|^{\beta-2})\Big)+\\
&\qquad\qquad+m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&= \int\left(1-\frac{(1+\alpha(1-\beta))}{q+2\alpha(1-\beta)}\right)m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&\qquad\qquad-\frac{(1+\alpha(1-\beta))}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}D(\operatorname{div}(Du|Du|^{\beta-2}))Du|Du|^{\beta-2}.
\end{aligned}$$

Before simplifying B, we compute $(Du|Du|^{\beta-2})_t$:

$$\begin{aligned}
(Du|Du|^{\beta-2})_t &= |Du|^{\beta-2}Du_t+Du(\beta-2)|Du|^{\beta-4}DuDu_t\\
&= (\mathcal{I}|Du|^{\beta-2}+(\beta-2)|Du|^{\beta-4}Du\otimes Du)Du_t = RDu_t,
\end{aligned}$$

where $\mathcal{I}$ is the identity matrix and $R = (\mathcal{I}|Du|^{\beta-2} + (\beta-2)|Du|^{\beta-4}Du \otimes Du) = D_{pp}H$. Using the preceding identity, we expand B as follows

$$\begin{aligned}
B &= \frac{1}{q+\alpha(1-\beta)} \int m^{q+\alpha(1-\beta)} \operatorname{div}((Du|Du|^{\beta-2})_t) \\
&= -\int m^{q+\alpha(1-\beta)-1} Dm (Du|Du|^{\beta-2})_t \\
&= -\int m^{q+\alpha(1-\beta)-1} DmRDu_t \\
&= -\int m^{q+\alpha(1-\beta)-1} DmR \left(D\left(m^{\alpha(1-\beta)} \frac{|Du|^\beta}{\beta} \right) - g'(m)Dm \right) \\
&= \int \overbrace{-m^{q+\alpha(1-\beta)-1} DmRD\left(m^{\alpha(1-\beta)} \frac{|Du|^\beta}{\beta} \right)}^{C} + \overbrace{m^{q+\alpha(1-\beta)-1} g'(m)DmRDm}^{D}.
\end{aligned}$$

Because $\beta \geq 1$, we get $D \geq 0$ as follows

$$\begin{aligned}
D &= \int m^{q+\alpha(1-\beta)-1} g'(m) Dm(\mathcal{I}|Du|^{\beta-2} + (\beta-2)|Du|^{\beta-4}Du \otimes Du)Dm \\
&= \int m^{q+\alpha(1-\beta)-1} g'(m)(|Dm|^2|Du|^{\beta-2} + (\beta-2)|Du|^{\beta-4}|DuDm|^2) \\
&\geq \int m^{q+\alpha(1-\beta)-1} g'(m)(|DmDu|^2|Du|^{\beta-4} + (\beta-2)|Du|^{\beta-4}|DuDm|^2) \\
&= \int (\beta-1) m^{q+\alpha(1-\beta)-1} g'(m)|DmDu|^2|Du|^{\beta-4},
\end{aligned}$$

where we used $|DmDu| \leq |Dm||Du|$. Concerning C, we expand further the expression

$$\begin{aligned}
C &= -\int m^{q+\alpha(1-\beta)-1} DmRD\left(m^{\alpha(1-\beta)} \frac{|Du|^\beta}{\beta} \right) \\
&= -\int m^{q+\alpha(1-\beta)-1} DmR \cdot \\
&\qquad \cdot \left(\alpha(1-\beta) m^{\alpha(1-\beta)-1} \frac{|Du|^\beta}{\beta} Dm + m^{\alpha(1-\beta)} |Du|^{\beta-2} D^2uDu \right) \\
&= \int \overbrace{\alpha(\beta-1) m^{q+2\alpha(1-\beta)-2} DmRDm \frac{|Du|^\beta}{\beta}}^{E} - \\
&\qquad - \overbrace{m^{q+2\alpha(1-\beta)-1} DmRD^2uDu|Du|^{\beta-2}}^{F}.
\end{aligned}$$

The first term above simplifies to

$$
\begin{aligned}
E = \int \alpha(\beta-1)m^{q+2\alpha(1-\beta)-2}Dm(\mathcal{I}|Du|^{\beta-2} +& \\
+ (\beta-2)|Du|^{\beta-4}Du\otimes Du)Dm\frac{|Du|^{\beta}}{\beta}& \\
= \int \frac{\alpha(\beta-1)}{\beta}m^{q+2\alpha(1-\beta)-2}|Dm|^2|Du|^{2\beta-2} +& \\
+ \frac{\alpha(\beta-1)(\beta-2)}{\beta}m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4}& \\
\geq \int \frac{\alpha(\beta-1)^2}{\beta}m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4}.&
\end{aligned}
$$

Finally, F becomes

$$
\begin{aligned}
F &= \int -m^{q+2\alpha(1-\beta)-1}Dm\Big(\mathcal{I}|Du|^{\beta-2} + (\beta-2)|Du|^{\beta-4}Du\otimes Du\Big)D^2uDu|Du|^{\beta-2} \\
&= \int -\frac{1}{q+2\alpha(1-\beta)}D(m^{q+2\alpha(1-\beta)})\cdot \\
&\qquad \cdot\Big(D^2uDu|Du|^{2\beta-4} + (\beta-2)Du\otimes DuD^2uDu|Du|^{2\beta-6}\Big) \\
&= \int \frac{1}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\operatorname{div}(D^2uDu|Du|^{2\beta-4} + \\
&\qquad + (\beta-2)Du\otimes DuD^2uDu|Du|^{2\beta-6}) \\
&= \int \frac{1}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\cdot \\
&\qquad \cdot\operatorname{div}\Big(\Big(D^2u|Du|^{\beta-2} + (\beta-2)|Du|^{\beta-4}Du\otimes(D^2uDu)\Big)Du|Du|^{\beta-2}\Big) \\
&= \int \frac{1}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\operatorname{div}\Big(D(Du|Du|^{\beta-2})Du|Du|^{\beta-2}\Big) \\
&= \int \frac{1}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\operatorname{tr}(D(Du|Du|^{\beta-2})^2) + \\
&\qquad + \frac{1}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}D(\operatorname{div}(Du|Du|^{\beta-2})Du|Du|^{\beta-2}).
\end{aligned}
$$

We add all terms and simplify, to conclude

$$
\begin{aligned}
\frac{d^2}{dt^2}\int m^q \geq \int \left(1 - \frac{(1+\alpha(1-\beta))}{q+2\alpha(1-\beta)}\right)m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2& \\
- \frac{(1+\alpha(1-\beta))}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}D(\operatorname{div}(Du|Du|^{\beta-2}))Du|Du|^{\beta-2}&
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\operatorname{tr}(D(Du|Du|^{\beta-2})^2)\\
&+\frac{1}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}D(\operatorname{div}(Du|Du|^{\beta-2})Du|Du|^{\beta-2})\\
&+\frac{\alpha(\beta-1)^2}{\beta}m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4}\\
&+(\beta-1)m^{q+\alpha(1-\beta)-1}g'(m)|DmDu|^2|Du|^{\beta-4}\\
\geq\int&\left(1-\frac{(1+\alpha(1-\beta)-\frac{1}{d})}{q+2\alpha(1-\beta)}\right)m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&+\overbrace{\frac{\alpha(\beta-1)}{p+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}D(\operatorname{div}(Du|Du|^{\beta-2}))Du|Du|^{\beta-2}}^{G}\\
&+\frac{\alpha(\beta-1)^2}{\beta}m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4}\\
&+(\beta-1)m^{q+\alpha(1-\beta)-1}g'(m)|DmDu|^2|Du|^{\beta-4},
\end{aligned}
\tag{27}
$$

where we used Lemma 3.1 to estimate

$$
\operatorname{tr}(D(Du|Du|^{\beta-2})^2)\geq\frac{1}{d}\operatorname{div}(Du|Du|^{\beta-2})^2.
$$

We only need to bound G from below. So, we integrate it by parts and use Cauchy-Schwarz inequality to conclude that

$$
\begin{aligned}
G&=\int\frac{\alpha(\beta-1)}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}Du|Du|^{\beta-2}D(\operatorname{div}(Du|Du|^{\beta-2}))\\
&=\int-\frac{\alpha(\beta-1)}{q+2\alpha(1-\beta)}\operatorname{div}(m^{q+2\alpha(1-\beta)}Du|Du|^{\beta-2})\operatorname{div}(Du|Du|^{\beta-2})\\
&=\int-\frac{\alpha(\beta-1)}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2 \qquad (28)\\
&\quad-\alpha(\beta-1)m^{q+2\alpha(1-\beta)}(m^{-1}DmDu|Du|^{\beta-2})\operatorname{div}(Du|Du|^{\beta-2})\\
&\geq\int-\frac{\alpha(\beta-1)}{q+2\alpha(1-\beta)}m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&\quad-\frac{\alpha(\beta-1)}{2}m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4}\\
&\quad-\frac{\alpha(\beta-1)}{2}m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&=\int\left(\frac{\alpha(1-\beta)}{q+2\alpha(1-\beta)}-\frac{\alpha(\beta-1)}{2}\right)m^{q+2\alpha(1-\beta)}\operatorname{div}(Du|Du|^{\beta-2})^2\\
&\quad-\frac{\alpha(\beta-1)}{2}m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4}.
\end{aligned}
$$

We use this last term into (27) to get

$$\begin{aligned}
&\int \left(1-\frac{(1+\alpha(1-\beta)-\frac{1}{d})}{q+2\alpha(1-\beta)}\right) m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2 + \\
&\quad + \left(\frac{\alpha(1-\beta)}{q+2\alpha(1-\beta)} - \frac{\alpha(\beta-1)}{2}\right) m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2 - \\
&\quad - \frac{\alpha(\beta-1)}{2} m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4} + \\
&\quad + \frac{\alpha(\beta-1)^2}{\beta} m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4} + \\
&\quad + (\beta-1)m^{q+\alpha(1-\beta)-1}g'(m)|DmDu|^2|Du|^{\beta-4} \\
&= \int \left(1-\frac{1-\frac{1}{d}}{q+2\alpha(1-\beta)} - \frac{\alpha(\beta-1)}{2}\right) m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2 + \\
&\quad + \frac{\alpha(\beta-1)(\beta-2)}{2\beta} m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4} + \\
&\quad + (\beta-1)m^{q+\alpha(1-\beta)-1}g'(m)|DmDu|^2|Du|^{\beta-4}.
\end{aligned}$$

From the above inequality, we see that if (26) holds, then $\frac{d^2}{dt^2}\int m^q \geq 0$. This concludes the proof for the case $\beta \geq 2$.

For $1 < \beta < 2$, we revisit the expression for G in (28). Then, we modify the Cauchy-Schwarz inequality to include the term $(\beta-1)$; that is,

$$\begin{aligned}
&\int -\frac{\alpha(\beta-1)}{q+2\alpha(1-\beta)} m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2 - \\
&\quad - \alpha m^{q+2\alpha(1-\beta)}((\beta-1)m^{-1}DmDu|Du|^{\beta-2}) \operatorname{div}(Du|Du|^{\beta-2}) \\
&\geq \int -\frac{\alpha(\beta-1)}{q+2\alpha(1-\beta)} m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2 - \\
&\quad - \frac{\alpha(\beta-1)^2}{2} m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4} - \frac{\alpha}{2} m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2. \\
&= \int \left(\frac{\alpha(1-\beta)}{q+2\alpha(1-\beta)} - \frac{\alpha}{2}\right) m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2 \\
&\quad - \frac{\alpha(\beta-1)^2}{2} m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\frac{1}{q(q-1)}\int m^q \geq &\int \left(1-\frac{1-\frac{1}{d}}{q+2\alpha(1-\beta)} - \frac{\alpha}{2}\right) m^{q+2\alpha(1-\beta)} \operatorname{div}(Du|Du|^{\beta-2})^2 \\
&+ \frac{\alpha(\beta-1)^2(2-\beta)}{2\beta} m^{q+2\alpha(1-\beta)-2}|DmDu|^2|Du|^{2\beta-4} \\
&+ (\beta-1)m^{q+\alpha(1-\beta)-1}g'(m)|DmDu|^2|Du|^{\beta-4},
\end{aligned}$$

which is non-negative if (25) holds. □

The idea used in the case without congestion to prove log-convexity of L^q norms fails in this case. However, we can still prove quasi-convexity of the L^∞ norm. In the next corollary, we identify couples $(\alpha, \beta) \in \mathbb{R}^+ \times (1, \infty)$ such that the set of values $q \geq 1$ satisfying either (24) or (25) is unbounded. Subsequently, we obtain a uniform bound on the density of solutions of (23).

Corollary 4.2. *Let* (u, m) *be classical solutions of* (23). *If*

$$\left(\beta \geq 2 \quad and \quad \alpha < \frac{2}{\beta - 1}\right) \quad or \quad \left(1 < \beta < 2 \quad and \quad \alpha < 2\right)$$

then, for every $d \geq 1$,

$$\|m(\cdot, t)\|_{L^\infty(\mathbb{T}^d)} \leq \max\{\|m(\cdot)^0\|_{L^\infty(\mathbb{T}^d)}, \|m(\cdot)^T\|_{L^\infty(\mathbb{T}^d)}\}. \tag{29}$$

Proof. First, we examine the case $\beta \geq 2$. Because $\alpha < \frac{2}{\beta-1}$, there is $\varepsilon > 0$ such that $\alpha = \frac{2}{\beta-1}(1 - \varepsilon)$. We use this value to simplify the following expression,

$$1 - \frac{1 - \frac{1}{d}}{q + 2\alpha(1 - \beta)} - \frac{\alpha(\beta - 1)}{2} = \varepsilon - \frac{1 - \frac{1}{d}}{q + 2\alpha(1 - \beta)}.$$

As $q \to \infty$, the r.h.s of the above identity converges to ε for all $d \geq 1$. Thus, there exists $Q > 0$ such that, for all $q > Q$,

$$1 - \frac{1 - \frac{1}{d}}{q + 2\alpha(1 - \beta)} - \frac{\alpha(\beta - 1)}{2} > 0.$$

Moreover, upon taking Q large enough, we can assume that $q + 2\alpha(1 - \beta) > 0$ for all $q > Q$. By Theorem (4.1), $t \mapsto \int m(x, t)^q dx$ is convex for all $q > Q$. Following similar computations to the ones in Section 3.3, we get

$$\|m(\cdot, t)\|_{L^\infty(\mathbb{T}^d)} \leq \max\{\|m(\cdot)^0\|_{L^\infty(\mathbb{T}^d)}, \|m(\cdot)^T\|_{L^\infty(\mathbb{T}^d)}\}.$$

Analogously, in the case $1 < \beta < 2$, we use (25), set $\alpha = 2(1 - \varepsilon)$, and follow the same reasoning as for $\beta \geq 2$ to obtain (29). □

Remark 4.3. In dimension $d = 1$, because

$$\frac{1 - \frac{1}{d}}{q + 2\alpha(1 - \beta)} = 0$$

for all $q > 2\alpha(\beta - 1)$, we do not need to use the ε argument. Therefore, (29) holds even in the case $\beta \geq 2$ and $\alpha = \frac{2}{\beta-1}$ or $1 \leq \beta < 2$ and $\alpha = 2$.

References

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

[2] Y. Achdou, A. Porretta: *Mean field games with congestion*, Ann. Inst. H. Poincaré Anal. Non Linéaire 35(2) (2018) 443–480.

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

[4] A. Bensoussan, J. Frehse, P. Yam: *Mean Field Games and Mean Field Type Control Theory*, Springer Briefs in Mathematics, Springer, New York (2013).

[5] Y. Brenier: *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. 44(4) (1991) 375–417.

[6] P. Cardaliaguet: *Notes on Mean-Field Games*, Lecture given at Tor Vergata, Rome, April-May 2010, https://www.ceremade.dauphine.fr/∼cardaliaguet/MFG20130420.pdf (2011).

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

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

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

[10] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, A. Porretta: *Long time average of mean field games*, Networks and Heterogeneous Media, 7(2) (2012) 279–301.

[11] P. Cardaliaguet, A. Mészáros, F. Santambrogio: *First order mean field games with density constraints: pressure equals price*, SIAM J. Control Optim. 54(5) (2016) 2672–2709.

[12] J. A. Carrillo, S. Lisini, G. Savaré, D. Slepcev: *Nonlinear mobility continuity equations and generalized displacement convexity*, J. Funct. Anal. 258(4) (2010) 1273–1309.

[13] J. A. Carrillo, D. Slepčev: *Example of a displacement convex functional of first order*, Calc. Var. Partial Diff. Equations 36(4) (2009) 547–564.

[14] J. Dolbeault, B. Nazaret, G. Savaré: *A new class of transport distances between measures* Calc. Var. Partial Diff. Equations 34(2) (2009) 193–231.

[15] D. Evangelista, R. Ferreira, D. Gomes, L. Nurbekyan, V. Voskanyan: *First-order, stationary mean-field games with congestion*, Nonlinear Analysis (2018), to appear.

[16] D. Evangelista, D. Gomes: *On the existence of solutions for stationary mean-field games with congestion*, J. Dyn. Diff. Equations (2016) 1–24.

[17] D. Gomes, L. Nurbekyan, M. Prazeres: *Explicit solutions of one-dimensional, first-order, stationary mean-field games with congestion*, in: *2016 IEEE 55th Conference on Decision and Control*, 4534–4539 (2016).

[18] D. Gomes, E. Pimentel: *Time dependent mean-field games with logarithmic nonlinearities*, SIAM J. Math. Anal. 47(5) (2015) 3798–3812.

[19] D. Gomes, E. Pimentel: *Local regularity for mean-field games in the whole space*, Minimax Theory Appl. 1(1) (2016) 065–082.

[20] D. Gomes, E. Pimentel, H. Sánchez-Morgado: *Time-dependent mean-field games in the subquadratic case*, Comm. Partial Diff. Equations 40(1) (2015) 40–76.

[21] D. Gomes, E. Pimentel, H. Sánchez-Morgado: *Time-dependent mean-field games in the superquadratic case*, ESAIM Control Optim. Calc. Var. 22(2) (2016) 562–580.

[22] D. Gomes, E. Pimentel, V. Voskanyan: *Regularity Theory for Mean-Field Game Systems*, Springer Briefs in Mathematics, Springer, New York (2016).

[23] D. Gomes, J. Saúde: *Mean field games models – a brief survey*, Dyn. Games Appl. 4(2) (2014) 110–154.

[24] D. Gomes, M. Sedjro: *One-dimensional, forward-forward mean-field games with congestion*, Discrete Contin. Dyn. Syst. Ser. S 11(5) (2018) 901–914.

[25] J. Graber: *Weak solutions for mean field games with congestion*, preprint, arXiv: 1503.04733 (2015).

[26] J. Graber, A. Mészáros: *Sobolev regularity for first order mean field games*, preprint, arXiv: 1708.06190 (2017).

[27] M. Huang, P. E. Caines, R. P. Malhamé: *Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria*, IEEE Trans. Automat. Control 52(9) (2007) 1560–1571.

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

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

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

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

[32] H. Lavenant, F. Santambrogio: *Optimal density evolution with congestion: L^∞ bounds via flow interchange techniques and applications to variational mean field games*, arXiv: 1705.05658 (2017), to appear in Comm. Partial Diff. Equations (2018).

[33] P. L. Lions: *Collège de France Course on Mean-Field Games*, https://www.college-de-france.fr/site/en-pierre-louis-lions (2007-2011).

[34] R. McCann: *A convexity principle for interacting gases*, Advances in Math. 128(1) (1997) 153–179.

[35] R. J. McCann: *Displacement concavity of Boltzmann's entropy characterizes positive energy in general relativity*, preprint, arXiv: 1808.01536 (2018).

[36] L. Nurbekyan: *One-dimensional, non-local, first-order stationary mean-field games with congestion: a Fourier approach*, Discrete Contin. Dyn. Syst. Ser. S 11(5) (2018) 963–990.

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

[38] A. Porretta: *Weak solutions to Fokker-Planck equations and mean field games*, Arch. Ration. Mech. Anal. 216(1) (2015) 1–62.

[39] F. Santambrogio: *Optimal Transport for Applied Mathematicians*, Progress in Nonlinear Differential Equations and their Applications 87, Birkhäuser/Springer, Cham (2015).

[40] B. Schachter: *A new class of first order displacement convex functionals*, SIAM J. Math. Analysis 50(2) (2018) 1779–1789.

[41] C. Villani: *Topics in Optimal Transportation*, Graduate Studies in Mathematics 58, American Mathematical Society, Providence (2003).

[42] C. Villani: *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften 338, Springer, Berlin et al. (2009).