Brake orbits and heteroclinic connections for first order Mean Field Games

Annalisa Cesaroni and Marco Cirant

Abstract

We consider first order variational MFG in the whole space, with aggregative interactions and density constraints, having stationary equilibria consisting of two disjoint compact sets of distributions with finite quadratic moments. Under general assumptions on the interaction potential, we provide a method for the construction of periodic in time solutions to the MFG, which oscillate between the two sets of static equilibria, for arbitrarily large periods. Moreover, as the period increases to infinity, we show that these periodic solutions converge, in a suitable sense, to heteroclinic connections. As a model example, we consider a MFG system where the interactions are of (aggregative) Riesz-type.

AMS-Subject Classification. 91A13 37J50 37J45 49Q20 35Q91

Keywords. Aggregating interaction potential, Wasserstein spaces, Infinite-dimensional Hamiltonian systems, Optimal transport, Variational methods.

Contents

1 Introduction ............................................. 1
1.1 A model problem .................................. 6
2 The Wasserstein spaces ................................. 8
3 Brake periodic solutions ............................... 10
4 Heteroclinic connections ............................... 16
5 A model problem ...................................... 20
  5.1 Standing assumptions on W and K .................. 20
  5.2 Assumptions (REF), (BDD) ....................... 22
  5.3 Continuity properties of W: assumptions (lsc) and (CON) . 23
  5.4 Minimizers for the stationary problem: the assumption (Z) .. 25
References .............................................. 27

1 Introduction

Mean field games (MFG) theory describes interactions among a large numbers of indistinguishable rational individuals, in which a generic agent optimizes some functional depending both on its dynamical state and on the average collective behavior, represented by the density of the overall population. In an equilibrium regime, the optimal dynamics of the average agent is consistent with the collective evolution. Such equilibria can be described by a system of coupled PDEs, a backward Hamilton-Jacobi equation characterizing the value function of the average
agent, and a forward continuity equation modelling the evolution of the population density, that is (in the model case of first order MFG with quadratic Hamiltonian)

\[
\begin{align*}
-\partial_t u + \frac{1}{2} |\nabla u|^2 &= f(x, m), & \text{in } (0, T) \times \mathbb{R}^d \\
\partial_t m - \text{div}(m \nabla u) &= 0, & \text{in } (0, T) \times \mathbb{R}^d \\
m &\geq 0, & \int_{\mathbb{R}^d} m(t, x) dx = 1.
\end{align*}
\]

(1.1)

Usually the system is coupled with initial/final time conditions. This theory has been introduced in the mathematical community by Lasry and Lions in [20, 21] and since then, there has been a large development of the subject in the literature.

Here, we will focus on the widely studied class of potential (or variational [4]) MFG: these are MFG systems that can be derived as optimality conditions of suitable optimal control problems. Precisely, we assume that \(f(x, m)\) is the derivative of a potential \(\mathcal{W}\) defined on the space of Borel probability measures \(\mathcal{P}(\mathcal{R}^d)\), that is \(f(x, m) = \frac{\partial}{\partial m} \mathcal{W}(m) \in C(\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d))\), where

\[
\lim_{h \to 0^+} \frac{\mathcal{W}(m + h(m' - m)) - \mathcal{W}(m)}{h} = \int_{\mathbb{R}^d} f(x, m) d(m' - m)(x)
\]

for all \(m, m' \in \mathcal{P}(\mathbb{R}^d)\). In this case, the PDE system (1.1) formally appears as the first order condition for critical points of the following energy functional:

\[
J_T(m, v) := \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |v(t, x)|^2 m(t, dx) dt + \int_0^T \mathcal{W}(m(t)) dt,
\]

(1.2)

to be computed among all possible evolutions of the mass distributions, that is among all couples \((m, v)\) such that \(m_t - \text{div}(mv) = 0\) in the distributional sense, where \(m(t) \in \mathcal{P}(\mathbb{R}^d)\) for all \(t\), and the velocity field \(v \in L^2(dt \otimes m(t, dx))\). It is well-known that when \(\mathcal{W} \equiv 0\), and \(m(0), m(T)\) are given, this is the so-called fluid mechanics formulation of the Monge-Kantorovich mass transfer problem introduced by Benamou and Brenier [3], which leads to the dynamic characterization of the \(L^2\)-Kantorovich-Rubinstein-Wasserstein distance \(d_2\) between measures in \(\mathcal{P}_2(\mathbb{R}^d)\) (those with finite quadratic moments in \(\mathbb{R}^d\) see Definition [22] and [2] [29] for a general discussion). The similarities between the Benamou-Brenier formulation of optimal transport and MFG have been already explored in the study of first-order MFG systems, and we refer to [17, 18, 22, 27].

Throughout the work, we will construct (constrained) critical points of \(J_T\), rather than producing solutions to PDE systems of the form (1.1). These critical points \((\bar{m}, \bar{v})\) give rise to mean field Nash equilibria, in the following sense: for any admissible competitor \((m, v)\),

\[
\int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |v|^2 m + \int_0^T f(\bar{m}) m \geq \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |v|^2 \bar{m} + \int_0^T f(\bar{m}) \bar{m},
\]

that in turn can be shown to provide solutions \((\bar{m}, \bar{v})\) to (1.1) (see Remark [3, 4] for further considerations).

Another crucial observation is that variational MFG systems of the form (1.1) can be interpreted as Hamiltonian systems on the infinite dimensional metric space \(\mathcal{P}_2(\mathbb{R}^d)\), endowed with the distance \(d_2\). In addition, the energy \(J_T(m, v)\) defined (1.2) can be rewritten via the Benamou-Brenier formula [3] as an energy on the space of trajectories \(C([0, T], \mathcal{P}_2(\mathbb{R}^d))\), as follows:

\[
J_T(m) = \int_0^T \frac{1}{2} |m'_{t}|^2 + \mathcal{W}(m(t)) dt,
\]

(1.3)

where \(|m'_t|\) is the metric derivative of the curve with respect to the Wasserstein distance \(d_2\), see [2]. In such a form, \(J_T\) is reminiscent of standard action functionals appearing in Hamiltonian mechanics.
We will also make use, as in the work by Benamou and Brenier, of the standard change of variables which replaces velocity by momentum, i.e. \((m, w) = (m, \nu m)\). The energy \((1.2)\) then becomes, in a generalized sense,

\[
I_T(m, w) = \int_0^T \int_{\mathbb{R}^d} \frac{1}{2} \frac{dw}{dt \otimes m(t, dx)} \left( m(t, dx) dt + \int_0^T \mathcal{W}(m) dt \right),
\]

to be computed on the set

\[
\mathcal{K} := \left\{ (m, w) \mid m \in C(\mathbb{R}, \mathcal{P}_2(\mathbb{R}^d)), \right. \\
w \text{ is a Borel } d\text{-vector measure on } \mathbb{R} \times \mathbb{R}^d, \text{ absolutely continuous w.r.t. } dt \otimes m(t, dx), \\
-\partial_t m + \text{div}(w) = 0 \text{ in the sense of distributions,} \\
\left. \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{1}{2} \frac{dw}{dt \otimes m(t, dx)} \left( m(t, dx) \right) < \infty \text{ for all } -\infty < t_1 < t_2 < \infty \right\}.
\]

The two energies \((1.2)\) and \((1.4)\) are equivalent, see [4]. Note that under these new variables the differential constraints become linear, that is \(m_1 - \text{div } w = 0\), and moreover the function \((m, w) \mapsto \frac{|w|^2}{m}\) (extended to 0 where \(m = 0\)) is a convex function. In the following, we are going to consider only solutions to the MFG which correspond to minimizers of \((1.4)\), in some suitable subset of \(\mathcal{K}\).

An interesting issue in MFG is the description of the long time behavior of equilibria, that is: given some information of the system at initial and final time, say at \(t = 0\) and \(t = T\), such as the population distribution \(m\) and/or the final cost \(u\), is it possible to describe \(m\) (and \(u\)) at intermediate times? A natural goal would be to characterize attractors that are approached by \(m\) as \(T \to \infty\). A large part of the literature in this direction is devoted to congestion type games, that are games in which players prefer sparsely populated areas of the state space. This is typically translated in the assumption that \(\mathcal{W}\) is convex, or equivalently that the interaction cost \(f(x, m)\) is monotone increasing with respect to the mass distribution:

\[
\int_{\mathbb{R}^d} (f(x, m) - f(x, m')) d(m - m')(x) \geq 0 \quad \forall m, m' \in \mathcal{P}_2(\mathbb{R}^d).
\]

We point out that this condition does not imply that the functional \(\mathcal{W}\) is geodesically convex in \(\mathcal{P}_2\) (see [2] [29]): geodesic convexity of \(\mathcal{W}\) and monotonicity of \(f\) are actually unrelated conditions. Under this monotonicity assumption, one expects in general uniqueness of the equilibria of the game, and some further regularity properties. The long time behavior of the system is quite well understood (at least when the state space is the flat torus): in a long time horizon, solutions of the MFG approach the (unique) stationary equilibrium, which is attractive for the evolutive system. Moreover, the stationary equilibrium is provided by the unique minimizer of \(\mathcal{W}\). We refer to the recent paper [8] and references therein for more details.

On the other hand, without the monotonicity assumption, the long time behavior is much less understood and very few is known about long time patterns. The second author obtained recently some results for second order MFG in the flat torus with anti-monotone interactions, that is assuming that \(-f(x, m)\) is monotone increasing. In particular in [13] (see also [14]), it is provided the construction, using bifurcation arguments, of an infinite number of branches of non-trivial solutions which exhibit an oscillatory (in time) behavior, and emanating from a trivial stationary solution (also for the case of two populations of players, which is non-variational in general). Finally, we recall also that in [24], by using weak KAM methods in an infinite dimensional setting, it is provided an example of a second order MFG with non monotone interaction cost, settled in the periodic torus, for which solutions in the long time horizon do not converge to the stationary state (see also [6] for further results). Long time pattern formation has also been explored in MFG models arising in socioeconomics [19] [30] [31].

In this paper, we analyze long time patterns for another class of aggregating first order (deterministic) MFG, defined on the whole space and without periodicity conditions. We consider
the anti-monotone case, that is when players are attracted towards crowded areas. The presence of such aggregating interaction forces naturally leads the population distribution to develop singularities. Against this force, we put a density constraint, i.e., we impose that the distribution of players \(m(t)\) must have a density which does not exceed some given value \(\rho\), that is, for all \(t\),

\[
m(t) \in \mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d) := \{ m \in \mathcal{P}_2(\mathbb{R}^d) : \exists \ 0 \leq \tilde{m} \leq \rho \ \text{a.e. on } \mathbb{R}^d \text{ s.t. } m = \tilde{m}dx \}
\]

(with a slight abuse of notation, we will often identify \(m\) with its density \(\tilde{m}\)) and so we restrict the set \(\mathcal{K}\) defined in (1.5) to

\[
\mathcal{K}^\rho := \{(m,w) \in \mathcal{K}, \ m(t) \in \mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d), \ \forall t \in \mathbb{R}\}. \quad (1.7)
\]

This constraint models an environment with finite capacity. Alternatively, it could be regarded as an infinite cost paid by players that try to cluster over saturated regions (hard congestion).

We mention that first order \(\text{MFG}\) with density constraints have been studied, in the monotone case, in [7], where connections with variational models for the incompressible Euler’s equations à la Brenier are also discussed (see also [22]). Another effect against concentration could be dissipation, that may appear as a viscosity term in the continuity equation for \(m\). This setting has been considered recently in [9], where stationary solutions to second order aggregating \(\text{MFG}\) are constructed; concentration phenomena and selection type results when the dissipation term vanishes are also shown.

Throughout the paper, we assume the following general conditions on the interaction potential \(W: \mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d) \to [0, \infty)\). First of all we assume that \(\min_{\mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d)} W\) exists, and without loss of generality that \(\min_{\mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d)} W = 0\). We suppose in addition that minima of \(W\) consists of two disjoint compact subsets of \(\mathcal{P}_2(\mathbb{R}^d)\), that is

\[
\exists \mathcal{M}^+, \mathcal{M}^- \subset \mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d) \ s.t. \ d_2(\mathcal{M}^+, \mathcal{M}^-) := 2\delta_0 > 0, \ \text{and} \ W(m) = 0 \Leftrightarrow m \in \mathcal{M}^\pm. \quad (Z)
\]

(where \(\mathcal{M}^\pm = \mathcal{M}^+ \cup \mathcal{M}^-\)). We assume some standard lower semi-continuity (in a topology which is slightly weaker than the one of \(\mathcal{P}_2\))

for any \(p < 2\), \(\{m_n\} \subset \mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d)\), if \(\lim d_\rho(m_n, m) = 0\) then \(\liminf_n W(m_n) \geq W(m),\) (isc)

which will be needed to construct minimizers of (1.4). Note that lower semi-continuity of the kinetic part term in \(J_T\) is standard by convexity (see Proposition 2.7). Some coercivity of \(W\) in \(\mathcal{P}_2(\mathbb{R}^d)\) will be also needed: there exists \(C_W > 0\) such that for all \(m \in \mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d)\)

\[-C_W + C_W^{-1} \int_{\mathbb{R}^d} |x|^2 m(x) dx \leq W(m) \leq C_W \left(1 + \int_{\mathbb{R}^d} |x|^2 m(x) dx\right). \quad (\text{BDD})\]

Note that (BDD) implies that \(W\) has compact sublevel sets in \(\mathcal{P}_p(\mathbb{R}^d)\) for every \(p < 2\), see Lemma 2.4 and Remark 2.5 but not necessarily for \(p = 2\).

We finally assume the following continuity property in \(\mathcal{P}_2(\mathbb{R}^d)\) close to the zero level-set: for any \(\{m_n\} \subset \mathcal{P}_{2,\rho}^\rho(\mathbb{R}^d)\),

\[
\text{if } \lim_n W(m_n) = 0, \ \text{then} \ \lim d_2(m_n, \mathcal{M}^\pm) = 0. \quad (\text{CON})
\]

Note that if \(W\) is assumed to be lower semicontinuous and with compact sublevel sets in \(\mathcal{P}_2(\mathbb{R}^d)\), then (CON) follows directly from (Z).

It is clear that minima \(\mathcal{M}^\pm\) of \(W\) are stationary solutions/equilibria, namely minimizers of the energy \(J_T\). The main goal of this work is show that the \(\text{MFG}\) problem has other equilibria that exhibit peculiar patterns. First, we construct \textit{periodic in time} critical points of \(J_T\), that oscillate between stationary solutions (brake orbits). Then, we construct \textit{heteroclinic connections}, that are,
with a slight abuse of notation, solutions to the MFG problem which are defined for all times, and approach \( \mathcal{M}^- \) as \( t \to -\infty \) and \( \mathcal{M}^+ \) at \( +\infty \) (see Definition 4.1). We will exploit the fact that the potential \( W \) in the energy (1.4) is assumed to be, roughly speaking, a double-well potential in \( \mathcal{P}_{2p}(\mathbb{R}^d) \). Written in the form (1.3), the energy can be interpreted as an action functional on the space of continuous curves with values in the metric space \( \mathcal{P}_2(\mathbb{R}^d) \), and is reminiscent of classical variational problems for finite-dimensional Hamiltonian systems. There is a huge literature (see the survey [28] and references therein) on the construction of periodic or heteroclinic trajectories in Hamiltonian systems by means of variational techniques. Among periodic solutions, the so-called brake orbits are widely studied; these are \( T \)-periodic curves \( m^T \) such that

\[
m^T \left( \frac{T}{4} - t \right) = m^T \left( \frac{T}{4} + t \right), \tag{1.8}
\]

so a brake orbit basically always travels along the same trajectory back and forth in \( T/2 \)-time (note that the speed \( m' (\pm T/4) \) vanishes). Brake orbits are periodic critical points of the action functional (1.3) (with Morse index 1 in the context of periodic perturbations) and not global minimizers. To mode out this instability, some symmetry can be added to the system. Here, we assume that there exists a reflection \( \gamma : \mathbb{R}^d \to \mathbb{R}^d \), such that

\[
W(\gamma m) = W(m). \tag{REF}
\]

Before stating our results, we recall that other extensions to the infinite dimensional setting of these kind of constructions has been considered quite recently in the literature. The existence of heteroclinic connections in the general framework of metric spaces has been provided in [25], under the assumption that the potential \( W \) has a finite numbers of zeros. The result is obtained by a different procedure, namely by re-parametrizing the action functional (1.3) to a length functional in the metric space: then an heteroclinic connection is a geodesic with respect the new length functional. Another class of infinite dimensional problems is related to functionals \( W \) defined on Hilbert spaces (such as \( H^1(\Omega) \), with appropriate boundary conditions) and \( W(u) = \| \nabla u \|^2_{L^2(\Omega)} + \int_{\Omega} W(x, u) dx \), where \( W(x, \cdot) \) is a double well potential. In [1] (see also references therein) the authors prove the existence of brake orbits and also convergence to heteroclinic connections as the period goes to infinity by minimizing the action functional among curves with prescribed energy. Analogous results have been proved in [16], with a different approach: instead of minimizing the action functional with fixed mechanical energy, the author minimize it on a set of \( T \)-periodic maps with fixed \( T > 0 \). In this paper, we follow the same approach as in [16], and as far as we know, similar constructions for MFG systems have never been studied.

The first main result is about construction of brake orbits, and it is proved in Section 3. We introduce the sets of curves on which we minimize our functional

\[
K^{p,S}_{T} := \{ (m, w) \in K^p, T\text{-periodic, } m \left( \frac{T}{4} + t \right) = m \left( \frac{T}{4} - t \right), m(-t) = \gamma m(t), \forall t \in \mathbb{R} \} \tag{1.9}
\]

where \( K^p \) is defined in (1.7). Observe that, due to the symmetry assumption (REF), we have that the two symmetry conditions \( m(T/4 + t) = m(T/4 - t) \), \( m(-t) = \gamma m(t) \) appearing in (1.9) are natural, in the sense that minimizers in \( K^{p,S}_{T} \) are also critical points in the larger set

\[
K^{p}_{T} := \{ (m, w) \in K^p, \text{ T-periodic} \} \tag{1.10}
\]

see Remark 3.4.

**Theorem 1.1.** Assume **Z**, **BDD**, **CON** and **REF**. Let \( q \in (0, q_0) \), where \( q_0 \) is defined in **Z**. Then there exists \( \bar{T} = \bar{T}(q) > 4 \) such that, for any \( T \geq \bar{T} \), there exists a \( T \)-periodic minimizer \( (m^T, w^T) \in K^{p,S}_{T} \) of the functional \( J_T \) in (1.4) which satisfies

\[
\begin{align*}
&d_2(m^T(t), \mathcal{M}^+) < q \quad \forall t \in \left( s, \frac{T}{2} - s \right) \\
&d_2(m^T(t), \mathcal{M}^-) < q \quad \forall t \in \left( -\frac{T}{2} + s, -s \right),
\end{align*}
\]
for some $0 < s < C = C(q)$ (note that $C$ does not depend on $T$). Moreover $\Bar{T}(q) \to +\infty$ as $q \to 0$.

Note that the transition phase between (neighborhoods of) the two steady states is of order $C(q)$, and remains bounded as $T \to \infty$. This is a key point in obtaining the second main result, which is about construction of heteroclinic solutions and convergence of brake orbits to heteroclinics; this is proved in Section 4.

To this aim, we introduce the energy on the whole space:

$$J(m, w) = \int_{-\infty}^{+\infty} \int_{\mathbb{R}^d} \frac{1}{2} |\frac{d}{dt} m(t, dx)|^2 m(t, dx) dt + \int_{-\infty}^{+\infty} \mathcal{W}_0(m(t)) dt.$$ 

and the sets of curves

$$\mathcal{K}^{p, S} := \{(m, w) \in \mathcal{K}^p : m(-t) = \gamma_0 m(t) \forall t, J(m, w) < +\infty\}.$$ 

Note that, see Lemma 4.2, if $(m, w) \in \mathcal{K}^{p, S}$, then

$$\lim_{t \to \pm \infty} d_2 (m(t), \mathcal{M}^\pm) = 0.$$ 

Moreover, arguing as in Remark 3.4 minimizers in $\mathcal{K}^{p, S}$ are also minimizers in $\mathcal{K}^p$. We have the following result.

**Theorem 1.2.** Assume [Z], [lsc], [BDD], [CON] and [REF].

a) There exist minimal heteroclinic connections, that is $(m, w) \in \mathcal{K}^{p, S}$ such that $J(m, w) = \min_{\mathcal{K}^{p, S}} J$.

b) For any $T > 0$, let $(m^T, w^T) \in \mathcal{K}_T^{p, S}$ be a minimizer of $J_T$ constructed in Theorem 1.1. Then

$$\lim_{T \to +\infty} d_2^2 \left( m^T \left( \frac{T}{4} \right), \mathcal{M}^+ \right) = 0 = \lim_{T \to +\infty} d_2^2 \left( m^T \left( -\frac{T}{4} \right), \mathcal{M}^- \right),$$

and up to passing to subsequences $T_n \to +\infty$, there holds

$$m^{T_n}(t) \to m(t) \in C(\mathbb{R}, \mathcal{P}_2(\mathbb{R}^d)) \text{ locally uniformly in } C(\mathbb{R}, \mathcal{P}_p(\mathbb{R}^d)) \text{ for all } p < 2$$

and $w^{T_n} \rightharpoonup w$ weakly in $L^2([-L, L] \times \mathbb{R}^d)$ for all $L > 0$, where $(m, w)$ is a minimal heteroclinic connection. Finally, $J(m, w) = \frac{1}{2} \lim_{T \to +\infty} J_T(m^T, w^T)$.

We make a few final remarks in light of the two results. As we previously mentioned, the unique minimizer of $\mathcal{W}$ is an attractor of MFG equilibria under the monotonicity assumption [1.6]. If one drops [1.6], the picture may change substantially. Heteroclinics produced here connect two different minimizers of $\mathcal{W}$; hence, the state of the system can be arbitrarily close to a minimum (with respect to $d_2$) of $\mathcal{W}$, and converge to a different steady state as $t \to \infty$. A further study of stability of minimizers of $\mathcal{W}$ can be matter of future work.

Note again that minimizers $(m, w)$ obtained in Theorems 1.1 and 1.2 provide solutions to MFG systems of the form (1.1) in a suitable weak sense. The connection between the variational formulation and the PDE system for first order problems has been extensively studied in [7], and adaptions to our framework may require minor technical work (see also Remark 3.4).

### 1.1 A model problem

Finally we present a model problem, where our results apply. We consider a vatiational MFG where the potential term $\mathcal{W}$ is given by

$$\mathcal{W}(m) = \int_{\mathbb{R}^d} W(x) m(dx) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(|x-y|)m(dx)m(dy).$$

(1.12)
Note that in this case $f(x,m) = \delta_m W = W(x) - 2 \int_{\mathbb{R}^d} K(|x-y|)m(dy)$. The first part of the energy is a potential energy, where $W : \mathbb{R}^d \to [0, +\infty)$ is a “double-well” confining function, vanishing on two disjoint balls, invariant by the reflection $\gamma$, and quadratically increasing at infinity, see assumption (5.2). The function $W$ models a spatial preference for the area where aggregation of the crowd takes place. The second part of the energy is an interaction energy, modeled through the interaction kernel $-K$. $K$ is assumed to be positive definite, radially symmetric, locally integrable and increasing at zero (in an appropriate sense), see (5.5), (5.6). In particular a model class of such interaction kernels $K$ is given by the Riesz kernels

$$K(|x-y|) = \frac{1}{|x-y|^{d-\alpha}}, \quad \text{with } \alpha \in (0, d).$$

Note that energies like (1.12) have been recently studied extensively, as they are directly connected to a class of self-assembly/aggregation models which have received much attention, see e.g. [10] and references therein.

It is possible to show, see Section 5, that under the previous assumptions, $W$ defined in (1.12) satisfies (BDD), (isc), (CON), (REF). Regarding the general assumption (Z), we provide the following characterization of minimizers of (1.12) in Section 5.4.

**Theorem 1.3.** Under the assumptions (5.2), (5.5), (5.6), there exist minimizers of (1.12) in $\mathcal{M}^\pm \subset \mathcal{P}_{2,\rho}(\mathbb{R}^d)$, and all the minimizers are given by characteristic functions (multiplied by $\rho$) of compact sets in $\mathbb{R}^d$. If, in addition, the flat zones of $W$ are sufficiently large in terms of $\rho$, i.e. (5.4) holds, then all the minimizers are characteristic functions of balls (multiplied by $\rho$), and consists of two compact disjoint sets $\mathcal{M}^\pm \subset \mathcal{P}_{2,\rho}(\mathbb{R}^d)$.

So, in the case described in Theorem 1.3 Theorems 1.1 and Theorem 1.2 apply and we may construct brake orbits and heteroclinic solutions.

Some interesting issues, in our opinion, are left open for this model problem. In particular, we know by Theorem 1.3 that stationary minimal solutions to the MFG, so in particular couples $(m,0) \in \mathcal{K}$ which minimize the energy (1.4), are given by $m = \rho \chi_E$, where $E$ is a compact set. A natural question is whether or not other (evolutive) equilibria enjoy these two features (that is, have compact support and are evolving characteristic functions). In other words, given a periodic brake orbit $(m^T, w^T)$ as in Theorem 1.1 or a minimal heteroclinic connection $(m,w)$ as in Theorem 1.2 is it true that $m^T(t), m(t)$ are characteristic functions (multiplied by $\rho$) of a family of evolving compact sets $E_t$ for all times ? At the moment a full answer to this question seems far to be understood.

Another natural related problem is the discrete (in space) version of the game: MFG can be interpreted indeed as limiting models for large populations of interacting agents, where any given individual is affected by the averaged state of the other individuals. In the work in preparation [10], we will consider the analogous variational problem involving the energy (1.2) for a finite number of interacting particles, where the density constraint appears as a bound from below on the minimal distance between particles (being in turn inversely proportional to the number of particles $N$). First of all, we formalize the connection between the discrete $N$-particles problem and the continuous MFG model by proving a $\Gamma$-convergence type result, as $N \to +\infty$, of the energies, in the same spirit of [13]. Moreover, we show that for the $N$-particle system, at least in the 1-dimensional case, periodic minimizers are compactly supported, and particles minimize reciprocal distances. This will give a partial answer to our question (again, at least in dimension one), namely we will provide the existence of limiting brake orbits for the continuous problem that are time-dependent characteristic functions.

**Acknowledgements**

The authors are partially supported by the Fondazione CaRiPaRo Project “Nonlinear Partial Differential Equations: Asymptotic Problems and Mean-Field Games”, and are members of...
Notation
We will denote by $B(x,r) \subset \mathbb{R}^d$ the ball centered at $x$ and with radius $r$, $B_r = B(0,r)$ and $\omega_d = |B_1|$. For any measurable set $E \subset \mathbb{R}^d$, we define $\chi_E$ to be the characteristic function of $E$. $P(\mathbb{R}^d)$, $P_p(\mathbb{R}^d)$ and $P_{2,p}(\mathbb{R}^d)$ are (sub)sets of Borel probability measures defined below (see Section 2). For any set $\mathcal{M} \subset P_2(\mathbb{R}^d)$, $d_2(\mu, \mathcal{M}) = \inf_{m \in \mathcal{M}} d_2(\mu, m)$.

2 The Wasserstein spaces
We introduce some notions for calculus in Wasserstein spaces that will be useful in the following. For a general reference on these results we refer to [2], [29]. First, let $P(\mathbb{R}^d)$ be the space of Borel probability measures on $\mathbb{R}^d$, endowed with the topology of narrow convergence, that is:

**Definition 2.1.** Let $\mu_k, \mu \in P(\mathbb{R}^d)$. We say that $\mu_k \to \mu$ narrowly if

$$\lim_k \int_{\mathbb{R}^d} g(x) \mu_k(dx) = \int_{\mathbb{R}^d} g(x) \mu(dx) \quad \forall g \in C_b(\mathbb{R}^d),$$

where $C_b(\mathbb{R}^d)$ is the space of continuous and bounded functions on $\mathbb{R}^d$.

Note that this notion of convergence is equivalent to the one of convergence in the sense of distributions (see [2] Remark 5.1.6)).

**Definition 2.2.** Let $p \geq 1$. The Wasserstein space of Borel probability measures with bounded $p$-moments is defined by

$$P_p(\mathbb{R}^d) = \left\{ \mu \in P(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} |x|^p d\mu(x) < +\infty \right\}.$$

The Wasserstein space can be endowed with the $p$-Wasserstein distance

$$d_p(\mu, \nu)^p = \inf \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x - y|^p d\pi(x,y) \mid \pi \in \pi(\mu, \nu) \right\} \quad (2.1)$$

where $\pi(\mu, \nu)$ is the set of Borel probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ such that $\pi(A \times \mathbb{R}^d) = \mu(A)$ and $\pi(\mathbb{R}^d \times A) = \nu(A)$ for any Borel set $A \subset \mathbb{R}^d$. When $p = 1$ and $\mu, \nu$ have compact support, 1-Wasserstein distance has also the following dual representation

$$d_1(\mu, \nu) = \sup \left\{ \int_{\mathbb{R}^d} \phi(x)d(\mu - \nu)(x) \mid \phi : \mathbb{R}^d \to \mathbb{R}, \text{Lip}(\phi) \leq 1 \right\} \quad (2.2)$$

Note that $P_p(\mathbb{R}^d) \subset P_q(\mathbb{R}^d)$ for $p < q$, and by Jensen inequality, $d_p(\mu, \nu) \leq d_q(\mu, \nu)$ for $p < q$. We then recall the following results about narrow convergence and convergence in Wasserstein spaces.

**Lemma 2.3.** Let $\mu_k, \mu \in P(\mathbb{R}^d)$ such that $\mu_k \to \mu$ narrowly.

(i) Let $g : \mathbb{R}^d \to [0, +\infty]$ be lower semicontinuous. Then

$$\liminf_k \int_{\mathbb{R}^d} g(x)d\mu_k(x) \geq \int_{\mathbb{R}^d} g(x)d\mu(x).$$
(ii) Let \( g : \mathbb{R}^d \to [0, +\infty) \), continuous and \( \mu_k \)-integrable, be such that

\[
\limsup_k \int_{\mathbb{R}^d} g(x) d\mu_k(x) \leq \int_{\mathbb{R}^d} g(x) d\mu(x) < \infty.
\]

Then, \( g \) is uniformly integrable with respect to \( \mu_k \), that is

\[
\lim_{R \to +\infty} \sup_k \int_{\{x \mid g(x) \geq R\}} g(x) d\mu_k(x) = 0.
\]

**Proof.** We refer to [2, Lemma 5.1.7].

**Lemma 2.4.** \( \mathcal{P}_p(\mathbb{R}^d) \) endowed with the \( p \)-Wasserstein distance is a separable complete metric space. A set \( \mathcal{M} \subset \mathcal{P}_p(\mathbb{R}^d) \) is relatively compact if and only if it has uniformly integrable \( p \)-moments, that is

\[
\lim_{R \to +\infty} \sup_{\mu \in \mathcal{M}} \int_{\mathbb{R}^d \setminus B(0, R)} |x|^p d\mu(x) = 0.
\]

Let now \( \mu_k, \mu \in \mathcal{P}_p(\mathbb{R}^d) \) for some \( p \geq 1 \). Then the statements below are equivalent:

(i) \( d_p(\mu_k, \mu) \to 0 \)

(ii) \( \mu_k \to \mu \) narrowly and \( \mu_k \) have uniformly integrable \( p \)-moments.

Finally, for any \( v \in \mathcal{P}_p(\mathbb{R}^d) \), the map \( \mu \to d_p(\mu, v) \) is lower semicontinuous with respect to narrow convergence.

**Proof.** We refer to [2, Prop. 7.1.5]. Note that if \( \mathcal{M} \) has uniformly integrable \( p \)-moments then it is tight, i.e. for all \( \varepsilon > 0 \) there exists \( K_\varepsilon \subseteq \mathbb{R}^d \) compact for which \( \sup_{\mu \in \mathcal{M}} \int_{\mathbb{R}^d \setminus K_\varepsilon} d\mu(x) \leq \varepsilon \).

The lower semicontinuity of the Wasserstein distance is proved in [2, Proposition 7.1.3].

**Remark 2.5.** Note that, if for some \( q > p \),

\[
\sup_{\mu \in \mathcal{M}} \int_{\mathbb{R}^d} |x|^q d\mu(x) < +\infty
\]

then \( \mathcal{M} \) has uniformly integrable \( p \)-moments.

Finally, we introduce a subspace of regular measures as follows.

**Definition 2.6.** We define \( \mathcal{P}^r_{2, p}(\mathbb{R}^d) \) to be the set of measures belonging to \( \mathcal{P}_2(\mathbb{R}^d) \) and having density in \( L^\infty(\mathbb{R}^d) \), with \( L^\infty(\mathbb{R}^d) \) norm bounded by \( \rho > 0 \):

\[
\mathcal{P}^r_{2, p}(\mathbb{R}^d) = \{ m \in \mathcal{P}_2(\mathbb{R}^d) : \exists 0 \leq \tilde{m} \leq \rho \text{ a.e. on } \mathbb{R}^d \text{ s.t. } m = \tilde{m} dx\}.
\]

Note that elements of \( \mathcal{P}^r_{2, p}(\mathbb{R}^d) \) “see” also the topology induced by the \( L^\infty \)-norm. We recall in particular the notion of weak* convergence in \( L^\infty \), that is: for \( \mu_k, \mu \in L^\infty(\mathbb{R}^d) \), \( \mu_k \) is said to converge to \( \mu \) weak* in \( L^\infty \) if

\[
\lim_k \int_{\mathbb{R}^d} g(x) \mu_k(dx) = \int_{\mathbb{R}^d} g(x) \mu(dx) \quad \forall g \in L^1(\mathbb{R}^d).
\]

We now make a few considerations on the kinetic part of the energy in (1.4), that is on the functional

\[
(m, w) \mapsto \int_{\mathbb{R}^d} \frac{1}{2} \left| \frac{d}{dt} \otimes m(t, dx) \right|^2 m(t, dx) dt,
\]

which can be defined in general for couples \( (m, w) \), \( m \in C(\mathbb{R}, \mathcal{P}_1(\mathbb{R}^d)) \) and \( w \) a Borel \( d \)-vector measure on \( \mathbb{R} \times \mathbb{R}^d \), absolutely continuous w.r.t. \( dt \otimes m(t, dx) \). These properties are indeed part
of the definition of admissible couples \((m, w) \in \mathcal{K}\). Throughout the paper, \(m(t)\) will further satisfy the \(L^\infty\) constraint \(m(t) \in \mathcal{P}_{2,\rho}(\mathbb{R}^d)\). We immediately note that if \((m, w) \in \mathcal{K}^p\), then \(w\) has a density which is in \(L^2_{\text{loc}}(\mathbb{R}, L^2(\mathbb{R}^d))\), that is

\[
\int_{t_1}^{t_2} \int_{\mathbb{R}^d} |w|^2 dx dt = \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{w(t,x)}{m(t,x)} m^2(t,x) dx dt \leq \rho \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \left| \frac{w(t,x)}{m(t,x)} \right|^2 m(t,x) dx dt.
\]

Moreover, by Hölder inequality and recalling that \(m(t) \in \mathcal{P}_2(\mathbb{R}^d)\), we have

\[
\left( \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |x||w(t,x)| dx dt \right)^2 \leq \left( \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |x|^2 m(t,x) dx dt \right) \left( \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \left| \frac{w(t,x)}{m(t,x)} \right|^2 m(t,x) dx dt \right) < \infty.
\]

We now state a lower semi-continuity result (which could be stated for weaker convergence in the variables \(m, w\), but it will be used below in the present form).

**Proposition 2.7.** Suppose that \(m_n \to m\) in \(C(\mathbb{R}, \mathcal{P}_p(\mathbb{R}^d))\) for some \(p \geq 1\), \(w_n \to w\) (weakly) in \(L^2(\Omega)\), and \(m_n(t), m(t)\) are absolutely continuous with respect to the Lebesgue measure for all \(t, n\). Then,

\[
\int_{t_1}^{t_2} \int_{\mathbb{R}^d} \left| \frac{w(t,x)}{m(t,x)} \right|^2 m(t,x) dx dt \leq \liminf_{n \to \infty} \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \left| \frac{w_n(t,x)}{m_n(t,x)} \right|^2 m_n(t,x) dx dt.
\]

**Proof.** See [29, Proposition 5.18].

Finally, we recall the following uniform continuity property of elements belonging to \(\mathcal{K}\), that will be useful in the sequel.

**Proposition 2.8.** Let \((m, w) \in \mathcal{K}\), as defined in (1.5). Then

\[
d^2(\tau, \sigma, m(t)) \leq |t-s| \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{1}{2} \left| \frac{dw}{dt} \otimes m(\tau, dx) \right|^2 m(\tau, dx) d\tau, \quad \forall t, s \in (t_1, t_2).
\]

**Proof.** This can be proved using Hölder inequality and [2 Thm 8.3.1].

## 3 Brake periodic solutions

In this section we prove Theorem 1.1. Denote by \(\Omega = (-\infty, \infty) \times \mathbb{R}^d\) and by \(C_T((\mathbb{R}, \mathcal{P}_2(\mathbb{R}^d))\) for any \(T > 0\), the subset of \(T\)-periodic elements of \(C((-\infty, \infty), \mathcal{P}_2(\mathbb{R}^d))\).

We provide a preliminary result on existence of geodesics in \(\mathcal{K}^p\) based on displacement convexity introduced by McCann [25].

**Lemma 3.1.** Let \(t_1 < t_2\), and \(m_1, m_2 \in \mathcal{P}_{2,\rho}(\mathbb{R}^d)\). Then, there exists a couple \((m, w) \in \mathcal{K}^p\), as defined in (1.7), such that

\[
m(t_1) = m_1, \quad m(t_2) = m_2, \quad \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{w(t,x)}{m(t,x)} m(t,x) dx dt = \frac{d_2^2(m(t_1), m(t_2))}{t_2 - t_1},
\]

for all \(t_1 \leq t \leq t_2 \leq t_2\).

**Proof.** Let \(\hat{m}\) be the unique constant speed geodesic \(\hat{m} \in AC([0,1], \mathcal{P}_2(\mathbb{R}^d))\) (see [2, Section 7.2]) connecting \(m_1\) and \(m_2\) (i.e. \(m(0) = m_1, m(1) = m_2\)), which satisfies for all \(0 \leq s \leq t \leq 1\)

\[
d_2(m(s), m(t)) = (t-s)d_2(m_1, m_2).
\]

The functional \(m \mapsto ||m||_{L^\infty(\mathbb{R}^d)}\) is geodesically convex in \(\mathcal{P}_2(\mathbb{R}^d)\) (see [29, Proposition 7.29]), namely \(m\) is in \(L^\infty(\mathbb{R}^d)\) for every \(s\), and it satisfies

\[
||\hat{m}(t)||_{L^\infty(\mathbb{R}^d)} \leq \max\{||m_1||_{L^\infty(\mathbb{R}^d)}, ||m_2||_{L^\infty(\mathbb{R}^d)}\} = \rho.
\]
Being \( m(t) \) a constant speed geodesic connecting \( m_1 \) and \( m_2 \),
\[
d_2^2(\hat{m}(s_1), \hat{m}(s_2)) = (s_2 - s_1) \int_{s_1}^{s_2} |\hat{m}'|^2(s) \, ds \quad \text{for all } 0 \leq s_1 \leq s_2 \leq 1. \tag{3.1}\]

By \cite{2} Thm 8.3.1] we get for a.e. \( s \in (0, 1) \) the existence of a vector field \( \hat{\delta}(s) \in L^2(\hat{m}(s); \mathbb{R}^d) \) such that \(-\partial_t \hat{m} + \text{div} (\hat{\delta} \hat{m}) = 0\) is satisfied in the distributional sense, and for a.e. \( s \)
\[
|\hat{m}'|(s) = \left( \int_{\mathbb{R}^d} |\hat{\delta}(s, x)|^2 \hat{m}(s, x) \, dx \right)^{1/2}.
\]

Hence, substituting \( |\hat{m}'|(s) \) into (3.1) and setting \( \hat{\omega} = \hat{\delta} \hat{m} \) (on the set \( \{ m > 0 \} \), and identically zero elsewhere), we obtain
\[
d_2^2(\hat{m}(s_1), \hat{m}(s_2)) = (s_2 - s_1) \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \frac{|\hat{\omega}(s, x)|^2 \hat{m}(s, x)}{\hat{m}(s, x)} \, dx \, ds.
\]

We then have that the couple \((\hat{m}, \hat{\omega})\) belongs to \( \mathcal{K} \). To obtain the required couple \((m, w)\) it is enough to perform a linear change of variables, i.e.
\[
m(t, x) := \hat{m} \left( \frac{t - t_1}{t_2 - t_1} x \right), \quad w(t, x) := \frac{1}{t_2 - t_1} \hat{\omega} \left( \frac{t - t_1}{t_2 - t_1} x \right).
\]

Finally we extend \( m(t, x) \) to all \( t \in \mathbb{R} \) by setting \( m(t, x) = m_1(x) \), \( w(t, x) = 0 \) for \( t < t_1 \), and \( m(t, x) = m_2(x) \), \( w(t, x) = 0 \) for \( t > t_2 \).

Now we need a technical lemma about positivity properties of the functional \( \mathcal{W} \) outside \( \mathcal{M}^\pm \).

**Lemma 3.2.** For any \( q > 0 \), we have
\[
\inf \left\{ \mathcal{W}(m) \mid m \in \mathcal{P}_{2,p}(\mathbb{R}^d) , \, d_2(m, \mathcal{M}^\pm) \geq q \right\} =: \delta(q, \mathcal{W}) = \delta > 0. \tag{3.2}
\]

**Proof.** Assume by contradiction that there exists \( q > 0 \) for which \( \delta = 0 \). We consider \( m_n \in \mathcal{P}_{2,p}(\mathbb{R}^d) \) such that \( q \leq d_2(m_n, \mathcal{M}^\pm) \) and \( 0 \leq \mathcal{W}_0(m_n) \leq 1/n \). By the lower bound in the assumption (BDD) and Lemma 2.4] we have that the sublevel set \( \mathcal{W}(m) \leq 1 \) is compact in \( \mathcal{P}_{p}(\mathbb{R}^d) \), for any \( p < 2 \), so by (lsc) and (Z), we conclude that
\[
\lim_n \mathcal{W}(m_n) = 0.
\]

Hence, by the continuity property (CON) we get that \( \lim_n d_2(m_n, \mathcal{M}^\pm) = 0 \), which is in contradiction with the fact that \( d_2(m_n, \mathcal{M}^\pm) \geq q \).

We will show in the next crucial lemma 3.3] that if \( (m, w) \in \mathcal{K}^p \) has bounded energy and \( m(t) \) is sufficiently close to \( \mathcal{M}^+ \) (resp. \( \mathcal{M}^- \)) at some times \( t_1, t_2 \), then actually it is close to \( \mathcal{M}^+ \) (resp. \( \mathcal{M}^- \)) in the whole time interval; otherwise, it is indeed possible to modify it and decrease the energy. The lemma is based on a cut argument, which has been already used in the analysis of periodic orbits and heteroclinic connections for Hamiltonian systems, see e.g. [16] Lemma 2.1].

**Lemma 3.3.** Let \( 0 < t_1 < t_2 < T \). Let \( (m, w) \in \mathcal{K}^p \) and let assume that \( (m, w) \) satisfies
\[
\int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{|w(t, x)|^2}{m(t, x)} \, m(t, x) \, dx \, dt \leq C'.
\]

Then, for all \( q \in (0, q_0] \), where \( q_0 \) is as in (Z), there exists \( q' = q'(q, C', \mathcal{W}) \) such that, if there exists \( \hat{m}_+ \in \mathcal{M}^+ \) such that
\[
d_2(m(t_i), \mathcal{M}^+) = d_2(m(t_i), \hat{m}_+) \leq q', \quad \text{for } i = 1, 2,
\]
\[
d_2(m(t_i), \mathcal{M}^+) > q \quad \text{for some } \, t^* \in (t_1, t_2),
\]

\[
\int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{|w(t, x)|^2}{m(t, x)} \, m(t, x) \, dx \, dt \leq C'.
\]
then there exists \((\mu, v) \in K^p\) such that \((m(t,x), w(t,x)) = (\mu(t,x), v(t,x))\) for \(t \in \mathbb{R} \setminus (t_1, t_2)\) and which satisfies
\[
d_2(\mu(t), M^+) < q \quad \text{for all} \quad t \in (t_1, t_2),
\]
and
\[
\int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{|v(t,x)|^2}{\mu(t,x)} \, dx \, dt + \int_{t_1}^{t_2} W(\mu(t)) \, dt < \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{|w(t,x)|^2}{m(t,x)} \, dx \, dt + \int_{t_1}^{t_2} W(m(t)) \, dt.
\]

**Proof.** For any \(0 < q' < q/2\) set
\[
\tau_1 = \max\{t > t_1 : d_2(m(s), M^+) \leq q, \quad \text{for all} \quad s \leq t\},
\]
\[
\tau_1' = \max\{t < t_1 : d_2(m(t), M^+) \leq q'\}.
\]
It holds \(t_1 \leq \tau_1' < \tau_1 < t^* < t_2\), and \(q' \leq d_2(m(t), M^+) \leq q\) for all \(t \in [\tau_1', \tau_1]\). Note that, by (2.3) and the triangle inequality,
\[
(C'(\tau_1 - t_1))^{1/2} \geq d_2(m(\tau_1), m(t_1)) \geq d_2(m(\tau_1), \bar{m}^+) - d_2(m(t_1), \bar{m}^+)
\]
\[
\quad \geq d_2(\mu(t_1), M^+) - d_2(m(t_1), M^+) \geq q - q' > \frac{q}{2},
\]
hence
\[
t_2 > \tau_1 > t_1 + \frac{q^2}{4C'}.
\]

We construct \((\mu, v)\) as follows. Choose \(0 < q' < \frac{1}{2} \min\left\{\frac{q^2}{4C'}, q\right\}\). By means of Lemma 3.1 there are two couples \((m_1, w_1)\) and \((m_2, w_2)\) belonging to \(K^p\) which connect \(m(t_1)\) to \(\bar{m}^+\) at time \(t_1 + q'\) and \(\bar{m}^+\) at time \(t_2 - q'\) to \(m(t_2)\), respectively. Set then
\[
\mu(t) := \begin{cases} m_1(t) & t \in [t_1, t_1 + q'], \\ \bar{m}^+ & t \in [t_1 + q', t_2 - q'], \\ m_2(t) & t \in [t_2 - q', t_2], \\ m(t) & \text{otherwise} \end{cases}, \quad v(t) := \begin{cases} w_1(t) & t \in [t_1, t_1 + q'], \\ 0 & t \in [t_1 + q', t_2 - q'], \\ w_2(t) & t \in [t_2 - q', t_2], \\ w(t) & \text{otherwise} \end{cases}
\]
The constraint \((\mu, v) \in K^p\) is easily verified. Note that for \(t \in [t_1, t_1 + q']\), by Lemma 3.1 and by the fact that \(d_2^2(m(t_1), \bar{m}^+) \leq (q')^2\), we get
\[
\frac{d_2^2(\mu(t), M^+)}{2} \leq \frac{d_2^2(\mu(t), \bar{m}^+)}{2} = \frac{d_2^2(m_1(t), m_1(t_1 + q'))}{2}
\]
\[
\quad = (t_1 + q' - t) \int_{t}^{t_1 + q'} \frac{w_1(t,x)^2}{m_1(t,x)} \int_{\mathbb{R}^d} \frac{m_1(t,x)}{m_1(t,x)} \, dx \, d\tau \leq \frac{d_2^2(m(t_1), \bar{m}^+)}{2} \leq (q')^2.
\]
The same inequalities hold on $t \in [t_2 - q', t_2]$, hence
\[ d_2(\mu(t), M^+) \leq q' \quad \text{for all } t \in (t_1, t_2). \] (3.3)

Since $M^+$ is compact,
\[ \int_{\mathbb{R}^d} |x|^2 \mu(t, x) \, dx = d_2(\mu(t), \delta_0) \leq d_2(\mu(t), M^+) + d_2(M^+, \delta_0) \leq \epsilon + q, \]
for some $\epsilon > 0$.

Therefore, by Lemma 3.1 and the growth assumption on $W$ given by (BDD), we get
\[ \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{|v|^2}{\mu} \, dx \, dt + \int_{t_1}^{t_2} W(\mu(t)) \, dt \]
\[ = \int_{t_1}^{t_1 + q'} \int_{\mathbb{R}^d} \frac{|v|^2}{m_1} \, dx \, dt + \int_{t_1}^{t_2} W(m_1(t)) \, dt + \int_{t_2 - q'}^{t_2} \int_{\mathbb{R}^d} \frac{|v|^2}{m_2} \, dx \, dt + \int_{t_2 - q'}^{t_2} W(m_2(t)) \, dt \]
\[ = \frac{d_2^2(m(t_1), m^+)}{q'} + \int_{t_1}^{t_1 + q'} W(m_1(t)) \, dt + \frac{d^2(m(t_2), m^+)}{q'} + \int_{t_2 - q'}^{t_2} W(m_2(t)) \, dt \]
\[ \leq \frac{d_2^2(m(t_1), m^+)}{q'} + \int_{t_1}^{t_1 + q'} C_W(1 + \epsilon + q) \, dt \]
\[ + \frac{d_2^2(m(t_2), m^+)}{q'} + \int_{t_2 - q'}^{t_2} C_W(1 + \epsilon + q) \, dt \leq 2q'(1 + \epsilon + q). \]

We now introduce a further intermediate time $\tilde{t}_1 := \max\{t < \tau_t : d_2(m(t), M^+) \leq q/2\}$. It holds $\tau_t < \tilde{t}_1 < \tau_t$, and $q/2 \leq d_2(m(t), M^+) \leq q$ for all $t \in [\tilde{t}_1, \tau_t]$. By the triangular inequality and the compactness of $M^+$, recalling the definition of $q_0$, we get $d_2(m(t), M^-) \geq 2q_0 - d_2(m(t), M^+) \geq 2q_0 - q \geq q$ for all $t \in [\tilde{t}_1, \tau_t]$. Therefore, by Lemma 3.2, we get that there exists $\delta = \delta(q/2, W)$ such that $W(m(t)) \geq \delta > 0$ for all $t \in [\tilde{t}_1, \tau_t]$. By (2.3)
\[ C'(\tau_t - \tilde{t}_1) \geq d_2^2(m(\tilde{t}_1), m(\tilde{t}_1)). \]

Recall that $(m, w) \in \mathcal{K}$, so [2 Theorem 8.3.1] guarantees that $\int_{\mathbb{R}^d} \frac{|w(t)|^2}{m(t)} \, dx \geq (|m'(t)|)^2$ for a.e. $t$. Hence, by Young’s inequality and the triangle inequality,
\[ \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \frac{|w|^2}{m} \, dx \, dt + \int_{t_1}^{t_2} W(m(t)) \, dt \geq \int_{\tilde{t}_1}^{\tau_t} \int_{\mathbb{R}^d} \frac{|w|^2}{m} \, dx \, dt + \int_{\tilde{t}_1}^{\tau_t} W(m(t)) \]
\[ \geq \sqrt{2} \int_{\tilde{t}_1}^{\tau_t} \left( \int_{\mathbb{R}^d} \frac{|w|^2}{m} \, dx \right)^{1/2} \sqrt{W(m(t))} \, dt \geq \sqrt{2 \delta} \int_{\tilde{t}_1}^{\tau_t} |m'(t)| \, dt = \sqrt{2 \delta} d_2(m(\tilde{t}_1), m(\tau_t)) \]
\[ \geq \sqrt{2 \delta} (d_2(m(\tilde{t}_1), M^+) - d_2(m(\tilde{t}_1), M^+)) \geq \frac{q}{2} \sqrt{2 \delta}. \]

Combining this inequality with (3.4) we complete the proof of the lemma, decreasing eventually $q'$ so that $2q'(1 + C_W + C_W q) < \frac{q}{2} \sqrt{2 \delta}$. \hfill \Box

We are now ready to construct $T$-periodic minimizers of $f_T$. We restrict the class $\mathcal{K}$ to flows of probability measures that are $T$-periodic and enjoy additional symmetries, so we introduce the set $K_T^{S}$ as defined in (1.9). We observe that the second symmetry constraint $m(-t) = \gamma m(t)$ rules out orbits which remain for all time in $M^+$ or in $M^-$. The first symmetry constraint $m(T/4 + t) = m(T/4 - t)$ is due to the fact that we are looking for brake periodic orbits, which oscillate twice in a period between $M^+$ and $M^-$. Note that we are using the notation $\gamma m(t) = \gamma m(t)$; since $m(t)$ has a density, this means that $\gamma m(t, x) = m(-t, \gamma(x))$ a.e.

We provide now the proof of the first main result, that is Theorem 1.1
Proof of Theorem 1.1

Step 1: Energy bounds. Choose any \( m_0 \in \mathcal{P}_p^0(\mathbb{R}^d) \) with compact support such that \( m_0 = \gamma m_0 \). Observe that \( d_2 \) is preserved by the transformation \( \gamma \) and define

\[
d := d_2(m_0, \mathcal{M}^+) = d_2(m_0, \mathcal{M}^-).
\]

Let \( \tilde{m}_+ \in \mathcal{M}^+ \), such that \( d_2(m_0, \mathcal{M}^+) = d_2(m_0, \tilde{m}_+) \). So \( d_2(m_0, \mathcal{M}^-) = d_2(m_0, \gamma \tilde{m}_+) \). By Lemma 3.1 there exists a couple \((m, w) \in \mathcal{K}^p \) that connects \( m_0 \) at time \( t = 0 \) to \( \tilde{m}_+ \) at time \( t = 1 \).

Let \( T > 4 \), and for \( t \in [0, T/2] \),

\[
\tilde{m}(t) := \begin{cases} 
  m(t) & t \in [0, 1], \\
  \tilde{m}_+ & t \in [1, T/2 - 1], \\
  m(T/2 - t) & t \in [T/2 - 1, T/2].
\end{cases}
\]

Observe that \( d_2(\tilde{m}(t), \mathcal{M}^+) \leq d_2(\tilde{m}(t), \tilde{m}_+) \leq d \) for all \( t \in [0, T/2] \). On the interval \([-T/2, 0]\), \((\tilde{m}, \tilde{w})\) can be extended symmetrically:

\[
(\tilde{m}(t), \tilde{w}(t)) := (\gamma \tilde{m}(-t), -\gamma \tilde{w}(-t)),
\]

Finally, \((\tilde{m}, \tilde{w})\) can be extended periodically over the whole time interval, so \((\tilde{m}, \tilde{w}) \in \mathcal{K}^{p,S}_T \).

Moreover we compute, recalling Lemma 3.1 and the growth condition \( \text{BDD} \) on \( W \),

\[
0 \leq \int_T (\tilde{m}, \tilde{w}) = 4 \int_0^1 \int_{\mathbb{R}^d} |\tilde{w}(t, x)|^2 \tilde{m}(t, x) \, dx \, dt + 4 \int_0^1 W(\tilde{m}(t)) \, dt \\
\leq 4d^2 + 4C_W(1 + d_2(\tilde{m}(t), \delta_0)) \leq 4d^2 + 4C_W(1 + d + d_2(\mathcal{M}^+, \delta_0)) = C'.
\]

Note that \( C' > 0 \) does not depend on \( T \). We may then suppose that along any minimizing sequence \((m_n, w_n)\),

\[
\int_0^T \int_{\mathbb{R}^d} |w_n(t, x)|^2 m_n(t, x) \, dx \, dt \leq \int_T (m_n, w_n) \leq C'.
\]

Step 2: Minimizing sequences can be chosen to be close to \( \mathcal{M}^\pm \). Pick any minimizing sequence \((m_n, w_n) \in \mathcal{K}^{p,S}_T \) of \( I_T \). Fix now \( n \in \mathbb{N} \). Let \( q \in (0, q_0) \), and \( 0 < q' < q \) be as in Lemma 3.2 (with \( C' \) as in (3.6)).

Note that the triangle inequality, the invariance of \( d_2 \) under \( \gamma \), \( m_n(0) = \gamma m_n(0) \), \( \mathcal{M}^+ = \gamma \mathcal{M}^- \) imply that \( d_2(\mathcal{M}^+, m_n(0)) = d_2(\mathcal{M}^-, m_n(0)) \) and then

\[
2q' < 2q_0 = d_2(\mathcal{M}^+, \mathcal{M}^-) \leq 2d_2(\mathcal{M}^+, m_n(0)) = 2d_2(\mathcal{M}^-, m_n(0)).
\]

Let \( \delta(q') = \inf_{m \in \mathcal{P}_p^0(\mathbb{R}^d), d_2(m, \mathcal{M}^+) \geq q'} W_0(m) > 0 \), as in Lemma 3.2. Let \( 0 \leq s \leq T \). Note that if \( d_2(m, \mathcal{M}^+) \geq q' \) for all \( t \in [0, s] \), then this implies

\[
s \delta(q') \leq \int_0^s W(m_n(t)) \, dt \leq I_T (m_n, w_n) \leq C'.
\]

Hence, for \( T > s := C'(\delta(q'))^{-1} \), by continuity of \( m_n(\cdot) \), since \( d_2(m_n(0), \mathcal{M}^+) > q' \), there exists \( s \in (0, \bar{s}) \) such that

\[
\delta_2(m_n(t), \mathcal{M}^+) > q' \text{ for all } t \in [0, s) \text{ and } d_2(m_n(s), \mathcal{M}^+) = q'.
\]

Let \( \bar{m} \in \mathcal{M}^+ \cup \mathcal{M}^- \) such that \( d_2(m_n(s), \bar{m}) = q' \). We may assume without loss of generality that \( \bar{m} \in \mathcal{M}^+ \) (the proof is completely analogous if \( \bar{m} \in \mathcal{M}^- \)).

So \( d_2(m_n(s), \mathcal{M}^+) = d_2(m_n(s), \bar{m}) = q' \). Note that by symmetry of \( m_n(t) \) we also have \( d_2(m_n(T/2 - s), \mathcal{M}^+) = d_2(m_n(T/2 - s), \bar{m}) = q' \). Hence, if \( d_2(m_n(t), \mathcal{M}^+) > q \) for some
$t \in (s, T/2 - s)$, by Lemma 3.3 it is possible to modify $(m_n, w_n)$ in $(s, T/2 - s)$ to construct a competitor $(\mu_n, v_n)$ with $f_T(\mu_n, v_n) < f_T(m_n, w_n)$. Therefore, we can further restrict the minimization process to competitors $(m, w) \in \mathcal{K}_T$ that satisfy for some $s$

$$
\begin{align*}
&\left\{d_2(m(t), \mathcal{M}^+) < q \quad \forall t \in \left(s, \frac{T}{2} - s\right) \\
&d_2(m(t), \mathcal{M}^-) < q \quad \forall t \in \left(s - \frac{T}{2}, -s\right) \right. .
\end{align*}
$$

(3.7)

Note that $0 < s \leq \bar{s} := C'(\delta(q'))^{-1}$ and that $T > C'(\delta(q'))^{-1} \to +\infty$ as $q \to 0$.

Step 3: Existence of a minimizer. By the growth condition (BDD), we get that there exists $t_n \in [0, T]$ such that $m_n(t_n)$ are uniformly bounded in $\mathcal{P}_2(\mathbb{R}^d)$ with respect to $n$. Moreover by (2.3)

$$
d_2^2(m_n(t), m_n(s)) \leq C |t - s|
$$

(3.8)
for all $t, s \in [0, T]$. This implies that $(m_n)$ is uniformly continuous as a sequence of $\mathcal{P}_2(\mathbb{R}^d)$-valued periodic functions, and

$$
\sup_n \sup_{t \in [0, T]} \int_{\mathbb{R}^d} |x|^2 m_n(x, t) dx < \infty.
$$

(3.9)

Therefore, by Ascoli-Arzelà theorem and Lemma 2.4 $(m_n)$ has a subsequence (still denoted by $(m_n)$) which converges in $C(\mathbb{R}, \mathcal{P}_p(\mathbb{R}^d))$ for all $p < 2$ to some $m^T \in C_T(\mathbb{R}, \mathcal{P}_p(\mathbb{R}^d))$. Due to the lower semicontinuity (lsc), and the growth assumption (BDD) of $\mathcal{W}$, we get that $m \in C_T(\mathbb{R}, \mathcal{P}_2(\mathbb{R}^d))$. Note that by convergence in $C(\mathbb{R}, \mathcal{P}_p(\mathbb{R}^d))$ symmetry properties pass to the limit. Moreover also (3.7) passes to the limit, due to lower semicontinuity of $d_2$ with respect to narrow convergence, see Lemma 2.4.

Finally, $(m_n)$ is bounded in $L^\infty(Q)$, so we can extract a further subsequence that converges $L^\infty(Q)$-weak-* to $m^T$, and $0 \leq m^T(t, x) \leq \rho$ a.e.

Regarding $(w_n)$, we have

$$
\int_0^T \int_{\mathbb{R}^d} |w_n(t, x)|^2 dx dt \leq \rho \int_0^T \int_{\mathbb{R}^d} \frac{|w_n(t, x)|^2}{m_n(t, x)} dx dt,
$$

(10.30)
hence $w_n$ converges weakly (up to a subsequence) in $L^2(Q)$ to some $w^T$.

It is easy to check that $-\partial_t m^T + \text{div}(w^T) = 0$ in the distributional sense. So we are just left to check that $(m^T, w^T)$ minimizes $\mathcal{W}$. We use the lower semicontinuity of the kinetic part of the energy recalled in Proposition 2.7 and for the potential part, we use the lower semicontinuity (lsc) of $\mathcal{W}$ and Fatou lemma.

We end this section with a remark on the optimality conditions.

Remark 3.4. Assume now that $\mathcal{W}$ has a derivative $f(x, m) = \frac{\partial}{\partial m} \mathcal{W}(m) \in C(\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d))$. Given any minimizer $(\bar{m}, \bar{w})$ of $\mathcal{W}$ in $\mathcal{K}_T^{p,s}$ as in Theorem 1.1, it is possible to show by convexity of $(m, w) \mapsto \frac{|w|^2}{m}$ and arguing as in [15] that for all $(m, w) \in \mathcal{K}_T^{p,s}$

$$
\int_0^T \int_{\mathbb{R}^d} \left| \frac{w(t, x)}{\bar{m}(t, x)} \right|^2 \bar{m}(t, x) + f(x, \bar{m}(t)) \bar{m}(t, x) dx dt \leq \int_0^T \int_{\mathbb{R}^d} \left| \frac{w(t, x)}{m(t, x)} \right|^2 m(t, x) + f(x, m(t)) m(t, x) dx dt.
$$

(11.31)

Such a minimality of $(\bar{m}, \bar{w})$ can be regarded as mean field Nash equilibrium property. We show below that the minimization property (3.11) can be extended to the more general class of non-symmetric competitors $(m, w) \in \mathcal{K}_T^p$. Therefore, following [7], the fact that $(\bar{m}, \bar{w})$ satisfies (3.11) for all $(m, w) \in \mathcal{K}_T^p$ could be used as a starting point to derive optimality conditions, that
are of the form \( \{1,1\} \). We mention that additional “pressure” terms and an ergodic constant may appear in the Hamilton-Jacobi equation, due to density constraints and T-periodicity. In any case, no further multipliers related to \( m(T/4 + t) = m(T/4 - t) \), \( m(-t) = \gamma m(t) \) appear in view of the symmetry assumption (Ref. 1).

We show just that the symmetry condition \( m(-t) = \gamma m(t) \) on competitors is natural and can be dropped (then arguing analogously, it can be shown the symmetry constraint around \( T/4 \) is also natural). Indeed, for \((m, w) \in K^0_T \) satisfying \( m(T/4 + t) = m(T/4 - t) \) only, let

\[
\tilde{m}(t) = \frac{1}{2} m(t) + \frac{1}{2} \gamma m(-t), \quad \tilde{w}(t) = \frac{1}{2} w(t) + \frac{1}{2} \gamma w(-t).
\]

Note that \( \mathcal{W}(m) = \mathcal{W}(\gamma m) \) yields \( f(x, m) = f(\gamma(x), \gamma m) \) (recall that \( f = \delta_m \mathcal{W} \)), and therefore, since \( m(t, x) = \tilde{m}(-t, \gamma(x)) \) via a change of variables and convexity,

\[
\begin{align*}
\int_0^T \int_{\mathbb{R}^d} \frac{|w(t, x)|^2}{m(t, x)} + f(x, \tilde{m}(t)) m(t, x) dx dt &= \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} \frac{|w(t, x)|^2}{m(t, x)} + f(x, \tilde{m}(t)) m(t, x) dx dt + \\
& \quad + \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} \frac{|w(-t, \gamma(x))|^2}{m(-t, \gamma(x))} + f(\gamma(x), \tilde{m}(-t)) m(-t, \gamma(x)) dx dt \\
& \quad \geq \int_0^T \int_{\mathbb{R}^d} \frac{|\tilde{w}(t, x)|^2}{\tilde{m}(t, x)} + f(x, \tilde{m}(t)) \tilde{m}(t, x) dx dt.
\end{align*}
\]

Then, since we have that \( (\tilde{m}, \tilde{w}) \in K^0 T \),

\[
\int_0^T \int_{\mathbb{R}^d} \frac{|w(t, x)|^2}{m(t, x)} + f(x, \tilde{m}(t)) m(t, x) dx dt \geq \int_0^T \int_{\mathbb{R}^d} \frac{|\tilde{w}(t, x)|^2}{\tilde{m}(t, x)} + f(x, \tilde{m}(t)) \tilde{m}(t, x) dx dt.
\]

4 Heteroclinic connections

In this section we provide the proof of the second main result, that is Theorem 1.2.

We introduce our definition of heteroclinic connection. First of all we recall the definition of the energy on the whole space:

\[
f(m, w) = \int_{-\infty}^{+\infty} \int_{\mathbb{R}^d} \frac{1}{2} \left( \frac{dw}{dt} \otimes m(t, dx) \right)^2 m(t, dx) dt + \int_{-\infty}^{+\infty} \mathcal{W}(m(t)) dt.
\] (4.1)

Recall that \( f = \delta_m \mathcal{W} \), and couples \((m, w) \in K^p \) are absolutely continuous with respect to the Lebesgue measure \( dt \otimes dx \).

**Definition 4.1.** Let \((\tilde{m}, \tilde{w}) \in K^p \). We say that \((\tilde{m}, \tilde{w}) \) is a heteroclinic connection for the MFG if \( \lim_{t \to -\infty} d_2(m, \mathcal{M}^-) = 0 = \lim_{t \to +\infty} d_2(m, \mathcal{M}^+) \), and \((\tilde{m}, \tilde{w}) \) satisfies

\[
\int_{-\infty}^{+\infty} \int_{\mathbb{R}^d} \frac{|\tilde{w}(t, x)|^2}{\tilde{m}(t, x)} \tilde{m}(t, x) + f(x, \tilde{m}(t)) \tilde{m}(t, x) dx dt \leq \int_{-\infty}^{+\infty} \int_{\mathbb{R}^d} \frac{|w(t, x)|^2}{m(t, x)} m(t, x) + f(x, m(t)) m(t, x) dx dt \quad \forall (m, w) \in K^p.
\]

We start observing that if \((m, w) \) has bounded energy, then \( m \) should approach at \( \pm \infty \) the stationary sets \( \mathcal{M}^\pm \).

**Lemma 4.2.** Let \((m, w) \in K^p \), and let \( C > 0 \) such that \( f(m, w) = C < +\infty \). Then

\[
\lim_{t \to +\infty} d_2(m(t), \mathcal{M}^+) = 0 \quad \text{and} \quad \lim_{t \to -\infty} d_2(m(t), \mathcal{M}^-) = 0
\]
or viceversa.
The function \( t \in (s_n, t_n) \to m(t) \) is a continuous function with value in \( \mathcal{P}_2(\mathbb{R}^d) \). Therefore there exists \( t_n \in (s_n, t_n) \) such that
\[
d_2(m(t_n), \mathcal{M}^+) = q_0.
\]
Again by triangular inequality we get that
\[
d_2(m(t_n), \mathcal{M}^-) \geq d_2(\mathcal{M}^+, \mathcal{M}^-) - d_2(m(t_n), \mathcal{M}^+) \geq 2q_0 - q_0 = q_0.
\]
And this, again, would contradict the boundedness of the energy.

Therefore, we get that either for all \( t_n \to +\infty \), we have that \( \lim n d_2(m(t_n), \mathcal{M}^+) = 0 \), or for all \( t_n \to +\infty \), we have that \( \lim n d_2(m(t_n), \mathcal{M}^-) = 0 \). This implies in particular the conclusion. \( \square \)

We provide now the existence of a solution to the problem
\[
J(\tilde{m}, \tilde{w}) = \min_{m, w \in K^{\tilde{s}}} J(m, w).
\]
(4.3)

Then, arguing as in Remark 3.4 one can show that any minimizer is an heteroclinic connection, in the sense of Definition 4.1.

Proof of Theorem 1.2

Proof of item a). We use similar arguments to those in the proof of Theorem 1.1.

Step 1: energy bounds. First of all we show that \( K^{\tilde{s}} \neq \emptyset \). Choose \( m_0 \in \mathcal{P}_{2,\rho}(\mathbb{R}^d) \) with compact support such that \( m_0 = \gamma m_0 \) and let
\[
d := d_2(m_0, \mathcal{M}^+) = d_2(m_0, \mathcal{M}^-).
\]
Let $\tilde{m}_+ \in M^+$ such that $d_2(m_0, \tilde{m}_+) = d$. By Lemma 3.1, there exists a couple $(m, w) \in K^p$ that connects $m_0$ at time $t = 0$ to $\tilde{m}_+$ at time $t = 1$.

$$\tilde{m}(t) := \begin{cases} m(t) & t \in [0, 1], \\
\tilde{m}_+ & t \in [1, +\infty), \end{cases} \quad \tilde{w}(t, x) := \begin{cases} w(t, x) & t \in [0, 1], \\
0 & t \in [1, +\infty). \end{cases}$$

Observe that $d_2(\tilde{m}(t), M^+) \leq d$ for all $t \in [0, 1]$. We extend $(\tilde{m}, \tilde{w})$ on $(-\infty, 0)$ symmetrically:

$$(\tilde{m}(t), \tilde{w}(t)) := (\gamma \tilde{m}(-t), -\gamma \tilde{w}(-t)),$$

Note that

$$J(\tilde{m}, \tilde{w}) = 2 \int_0^1 \int_{\mathbb{R}^d} \frac{\tilde{w}(t, x)^2}{\tilde{m}(t, x)} \tilde{m}(t, x) \, dx \, dt + 2 \int_0^1 \mathcal{W}(\tilde{m}(t)) \, dt \leq \frac{C'}{2},$$

where $C'$ is defined in (3.5).

**Step 2: limit of minimizing sequences.** We consider now a minimizing sequence $(m_n, w_n) \in K^{p,s}$ such that $J(m_n, w_n) \leq \frac{C'}{2}$. By the growth condition (BDD) on $\mathcal{W}$, since $\int_0^1 \mathcal{W}(m_n) \, dt < C'/2$, there exists $t_n \in [0, 1]$ such that $m_n(t_n)$ is uniformly bounded in $\mathcal{W}_2(\mathbb{R}^d)$. By (2.3), we get that $(m_n) \subset C(\mathbb{R}, \mathcal{W}_2(\mathbb{R}^d))$ is equicontinuous. So, by the triangle inequality, we get that $m_n(t)$ is uniformly bounded in $\mathcal{W}_2(\mathbb{R}^d)$ for all $t$. By Ascoli-Arzelà theorem and Lemma 2.3 up to extracting a subsequence and to a diagonalization procedure, we get that $m_n$ converges uniformly in $C([-L, L, \mathcal{W}_p(\mathbb{R}^d))$ for all $p < 2$ and all $L > 0$, to some $\bar{m} \in C(\mathbb{R}, \mathcal{W}_p(\mathbb{R}^d))$.

Again by lower semicontinuity (Lsc), and the growth condition (BDD), there holds that $\bar{m} \in C(\mathbb{R}, \mathcal{W}_2(\mathbb{R}^d))$. Moreover $\bar{m}(-t) = \gamma \bar{m}(t)$ since symmetry properties pass to the limit, and we can extract a further subsequence that converges also in $L^\infty([-L, L] \times \mathbb{R}^d)$-weak-* to $\bar{m}$, so $0 \leq \bar{m}(x, t) \leq \rho$ a.e.. Finally, reasoning as in (3.10), we get that $w_n$ converges weakly (up to the extraction of a subsequence and a diagonalization procedure) in $L^2([-L, L] \times \mathbb{R}^d)$ to some $\bar{w}$ for every $L > 0$. In particular we get that $-\partial_t \bar{m} + \text{div}(\bar{w}) = 0$ in distributional sense in $(-\infty, +\infty) \times \mathbb{R}^d$.

**Step 3: finite energy.** We fix $L > 0$. By the lower semicontinuity properties and Fatou lemma, we get that for every $L > 0$,

$$0 \leq \int_{-L}^L \int_{\mathbb{R}^d} \frac{\tilde{w}(t, x)^2}{\tilde{m}(t, x)} \tilde{m}(t, x) \, dx \, dt + \int_{-L}^L \mathcal{W}(\tilde{m}(t)) \, dt \leq \liminf_{n} J(m_n, w_n) : = \eta \leq \frac{C'}{2}$$

and so again by Fatou lemma

$$0 \leq J(\bar{m}, \bar{w}) \leq \eta = \inf_{(\mu, v) \in K^{p,s}} J(\mu, v) \leq \frac{C'}{2}.$$

This implies that $(m, w) \in K^{\infty}$ and moreover that $(m, w)$ is a minimizer.

**Proof of item b.**

**Step 4:** limit of $m^T(\pm T/4)$ as $T \to +\infty$. First of all, by Theorem 1.1, observe that for all $\varepsilon > 0$ small and $T \geq \max(4s(\varepsilon), T(\varepsilon))$ (using the same notation as in Theorem 1.1), there exists a minimizer $(m^\varepsilon, w^\varepsilon)$ such that

$$0 \leq d^2_2 \left( m^T(t), M^+ \right) \leq \varepsilon \quad \forall t \in \left( s(\varepsilon), \frac{T}{2} - s(\varepsilon) \right).$$

The conclusion follows observing that $\frac{T}{4} \in \left( s(\varepsilon), \frac{T}{2} - s(\varepsilon) \right)$ and that $\tilde{T}(\varepsilon) \to +\infty$ as $\varepsilon \to 0$.

**Step 5:** equicontinuity of $m^T$ and passage to the limit. Let fix $q \in (0, q_0)$ and let $(m^T, w^T) \in K^{p,s}$ be a minimizer of $f_T$ constructed in Theorem 1.1 with $T > \tilde{T}(q)(\geq 4)$. First of all observe that by (3.6) there exists $C'$ independent of $T$ such that $0 \leq f_T(m^T, w^T) \leq C'$, and so in particular by
we get that \((m^T(\cdot))_T \subset C(\mathbb{R}, P_2(\mathbb{R}^d))\) is equicontinuous. By the growth condition (BDD), since \(\int_0^1 \mathcal{W}(m^t) dt \leq C\), there exists \(t(T) \in [0, 1]\) such that \(m^T(t(T))\) is bounded in \(P_2(\mathbb{R}^d)\), uniformly with respect to \(T\).

By (2.3) and triangular inequality we conclude that for all \(t \in [0, T]\), \(m^T(t)\) is bounded in \(P_2(\mathbb{R}^d)\), uniformly with respect to \(T\). By Ascoli-Arzelà theorem and Lemma 2.3 we get that up to extracting a subsequence \(T_n \rightarrow +\infty\) and using a diagonalization procedure, we get that \(m^{T_n}\) converges uniformly in \(C([-L, L], P_p(\mathbb{R}^d))\) for all \(p < 2\) and all \(L > 0\), to some \(m \in C(\mathbb{R}, P_2(\mathbb{R}^d))\), which a posteriori, due to (isc) and (BDD), satisfies \(m \in C(\mathbb{R}, P_2(\mathbb{R}^d))\). Moreover \(m(-t) = \gamma m(t)\) since symmetry properties pass to the limit, and we can extract a further subsequence that converges also in \(L^\infty([-L, L] \times \mathbb{R}^d)\)-weak-* to \(m\), and \(0 \leq m(x, t) \leq \rho\) a.e.. Finally, reasoning as in (3.10), we get that \(w\) converges weakly (up to the extraction of a subsequence and a diagonalization procedure) in \(L^2([-L, L] \times \mathbb{R}^d)\) to some \(w\) for every \(L > 0\).

In particular we get that \(-\partial_t m + \text{div}(w) = 0\) in distributional sense in \((-\infty, +\infty) \times \mathbb{R}^d\).

It is immediate to check that the same argument applies to every limit point of \((m^T, w^T)\).

**Step 6: finite energy of the limit points \((m, w)\).** Let \((m, w)\) the limit of \((m^{T_n}, w^{T_n})\) as obtained in the previous step. Fix now \(L > 0\) and let \(n_0\) such that \(T_n \geq 4L\) for all \(n \geq n_0\). By the lower semicontinuity properties and Fatou lemma, we get that for every \(L > 0\), we get that

\[
0 \leq \int_{-L}^L \int_{\mathbb{R}^d} \frac{w(t, x)}{m(t, x)}^2 m(t, x) \, dx \, dt + \int_{-L}^L \mathcal{W}(m(t)) \, dt
\]

\[
\leq \liminf_n \int_{-L}^L \int_{\mathbb{R}^d} \frac{w^{T_n}(t, x)}{m^{T_n}(t, x)}^2 m^{T_n}(t, x) \, dx \, dt + \int_{-L}^L \liminf_n \mathcal{W}(m^{T_n}(t)) \, dt
\]

\[
\leq \liminf_n \int_{-L}^L \int_{\mathbb{R}^d} \frac{w^{T_n}(t, x)}{m^{T_n}(t, x)}^2 m^{T_n}(t, x) \, dx \, dt + \liminf_n \int_{-L}^L \mathcal{W}(m^{T_n}(t)) \, dt
\]

\[
= \frac{1}{2} \liminf_n \int_{-L}^{T_n/4} \int_{-L/4}^{L/4} \frac{w^{T_n}(t, x)}{m^{T_n}(t, x)}^2 m^{T_n}(t, x) \, dx \, dt + \int_{-L/4}^{T_n/4} \mathcal{W}(m^{T_n}(t)) \, dt
\]

and so by Fatou lemma

\[
0 \leq J(m, w) \leq \frac{1}{2} C'.
\]

This, along with the properties of limit points proved in Step 5, implies that \((m, w) \in \mathcal{K}_\infty\).

**Step 7: \((m, w)\) is a solution of (4.3).** For every converging subsequence, since \(J_{T_n}(m^{T_n}, w^{T_n})\) is equibounded, up to passing to a further subsequence we may assume that \(\lim_n J_{T_n}(m^{T_n}, w^{T_n}) = e > 0\). Arguing as above it is immediate to check that

\[
e = \lim_n J_{T_n}(m^{T_n}, w^{T_n}) \geq 2J(m, w).
\]

We claim that

\[
e \leq 2J(\bar{m}, \bar{w})
\]

where \((\bar{m}, \bar{w})\) is a minimizer constructed in item a). If the claim is true, then we have that

\[
2J(m, w) \leq e \leq 2J(\bar{m}, \bar{w})
\]

which implies immediately that \((m, w)\) is a minimizer and moreover that \(e = 2J(\bar{m}, \bar{w})\).

Assume by contradiction that for some \(\delta > 0\), there holds

\[
e > 2J(\bar{m}, \bar{w}) + \delta.
\]

Let us fix \(T_n\) and consider \((\bar{m}, \bar{w})\) restricted to \([-\frac{T_n}{4}, \frac{T_n}{4}]\). Extend them to \([-\frac{T_n}{4}, 3\frac{T_n}{4}]\) by putting

\[
\bar{m}_n \left( t + \frac{T_n}{4} \right) := \bar{m} \left( \frac{T_n}{4} - t \right), \quad \bar{w}_n \left( t + \frac{T_n}{4} \right) := -\bar{w} \left( \frac{T_n}{4} - t \right)
\]

for \(t \in [0, \frac{T_n}{4}]\) and then extend them
periodically in \( \mathbb{R} \). It is easy to check that \( \gamma \bar{m}_n(t) = \bar{m}_n(-t) \). So \((\bar{m}_n, \bar{w}_n) \in \mathcal{K}_{T_n} \) and therefore

\[
J_{T_n}(m_{T_n}, w_{T_n}) \leq J_{T_n}(\bar{m}_n, \bar{w}_n) = 2 \int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} \int_{\mathbb{R}^d} \frac{1}{\bar{m}(t, x)}^2 \bar{m}(t, x) \, dx + \mathcal{W}(\bar{m}(t)) \, dt \leq 2 J(\bar{m}, \bar{w}) < e - \delta.
\]

Taking \( n \) sufficiently large, this gives a contradiction with the fact that \( e = \lim_n J_{T_n}(m_{T_n}, w_{T_n}) \).

\[ \square \]

5 A model problem

In this section we describe a model to which previous results apply. We define on \( \mathcal{P}_{2,\rho}^\star(\mathbb{R}^d) \) (for any fixed \( \rho > 0 \)) the following potential energy

\[
\mathcal{W}(m) = \int_{\mathbb{R}^d} W(x) m(dx) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(|x - y|) m(dx) \, m(dy). \tag{5.1}
\]

First of all we describe our main assumptions on \( K \) and \( W \) and then we check all the conditions that are needed in Theorems 1.1, 1.2. Note that, as we will see below, \( \mathcal{W}(m) \) has minimizers on \( \mathcal{P}_{2,\rho}^\star(\mathbb{R}^d) \), but \( \min_{\mathcal{P}_{2,\rho}^\star(\mathbb{R}^d)} \mathcal{W} < 0 \). Therefore, to apply Theorems 1.1, 1.2 one just needs to add to \( \mathcal{W} \) the renormalization constant \( \min_{\mathcal{P}_{2,\rho}^\star(\mathbb{R}^d)} \mathcal{W} \), that is to consider

\[
\mathcal{W}_0(m) = \mathcal{W}(m) - \min_{\mathcal{P}_{2,\rho}^\star(\mathbb{R}^d)} \mathcal{W}.
\]

5.1 Standing assumptions on \( W \) and \( K \)

We start describing the assumptions on the local energy \( \int_{\mathbb{R}^d} W(x) \, dm(x) \). Let \( W : \mathbb{R}^d \to [0, +\infty) \) be a confining double-well potential such that

\[
\begin{align*}
W &\in C(\mathbb{R}^d) \text{ is non-negative,} \\
\exists C > 0, \text{ such that } C^{-1} |x|^2 - C \leq W(x) \leq C|x|^2 + C \\
\exists a^+, a^- \in \mathbb{R}^d, r_0 > 0 \text{ such that } B(a^+, r_0) \cap B(a^-, r_0) = \emptyset \\
\text{and } W(x) = 0 \iff x \in B(a^+, r_0) \cup B(a^-, r_0).
\end{align*} \tag{5.2}
\]

Note that we require \( W \) to have two disjoint flat regions \( B(a^\pm, r_0) \). Moreover, we assume that \( W \) is invariant under a reflection \( \gamma : \mathbb{R}^d \to \mathbb{R}^d \), that is

\[
W(x) = W(\gamma(x)) \quad x \in \mathbb{R}^d. \tag{5.3}
\]

In particular this implies that \( \gamma(a^+) = a^- \). Finally, we assume that the plateaus of \( W \) are sufficiently large with respect to the density constraint \( \rho \), in the following sense:

\[
\rho \geq \rho_0 := \frac{1}{\omega_d r_0^d}, \tag{5.4}
\]

where \( r_0 \) is defined in (5.2). See Figure 2 for an example of \( W \) satisfying our assumptions.

We describe now the assumptions on the interaction energy \( -\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(|x - y|) \, m(dx) \, m(dy) \) and some basic properties. We consider a radially symmetric interaction kernel \( K(|x|) \), where \( K : [0, +\infty) \to [0, +\infty) \) is a function such that

\[
\begin{align*}
r \mapsto r^{d-1} K(r) &\in L^1_{loc}([0, +\infty), [0, +\infty)), \text{ } K \text{ is nonincreasing,} \\
\lim_{r \to 0} K(r) - K(t + r) &> 0 \text{ for every } t \text{ and } \lim_{r \to +\infty} K(r) = 0.
\end{align*} \tag{5.5}
\]
Moreover we assume that $K$ is positive definite, which means that
\[ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(y)K(|x-y|)dxdy \geq 0 \quad \text{for all } f \in L^1(\mathbb{R}^d) \tag{5.6} \]
and $\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(y)K(|x-y|)dxdy = 0$ if and only if $f = 0$.

We define the energy interaction functional for $f \in L^1(\mathbb{R}^d)$
\[ \mathcal{I}(f) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(y)K(|x-y|)dxdy - \int_{\mathbb{R}^d} f(x)V_f(x)dx \tag{5.7} \]
where $V_f$ is the interaction potential
\[ V_f(x) = f \ast K(x) = \int_{\mathbb{R}^d} f(y)K(|x-y|)dy. \tag{5.8} \]

**Remark 5.1.** Note that if $K$ is positive definite, then the map $f \mapsto V_f = K \ast f$ is monotone increasing in the sense that
\[ \int_{\mathbb{R}^d} (V_{f_1} - V_{f_2})(f_1 - f_2)dx \geq 0 \quad \forall f_1, f_2 \in L^1(\mathbb{R}^d). \]

We recall a well-known result on the interaction potential.

**Lemma 5.2.** Assume (5.5). Let $f \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Then $V_f \in C(\mathbb{R}^d)$ and $\lim_{|x|\to+\infty} V_f(x) = 0$.

**Proof.** We sketch the proof. First of all we prove that $V_f$ is continuous (the fact that it is well defined is straightforward). Let $\eta : [0, +\infty) \to [0, 1]$ be a smooth function such that $\eta = 0$ in $[0, 1]$ and $\eta = 1$ in $[2, +\infty)$. Define $K_\epsilon(|x|) = K(|x|)\eta(|x|/\epsilon)$ for $\epsilon > 0$. Observe that $K_\epsilon \in L^\infty$ and define $V_\epsilon^f = K_\epsilon \ast f$. Then $V_\epsilon^f \in C(\mathbb{R}^d)$ since $|V_\epsilon^f(x + h) - V_\epsilon^f(x)| \leq \|K_\epsilon\|_\infty \|f(\cdot + h) - f(\cdot)\|_1 \to 0$ as $|h| \to 0$. Moreover observe that $V_\epsilon^f$ converges uniformly to $V_f$ as $\epsilon \to 0$ since
\[ |V_\epsilon^f(x) - V_f(x)| \leq \int_{B(0,2\epsilon)} K(|y|)|f(x-y)|dy \leq \|f\|_\infty \int_{B(0,2\epsilon)} K(|y|)dy. \]

Now we prove that $V$ is vanishing at $\infty$. Observe that for $R > 1$,
\[ V_f(x) = \int_{B(0,1/R)} K(|y|)f(x-y)dy + \int_{B(0,R)\setminus B(0,1/R)} K(|y|)f(x-y)dy + \int_{\mathbb{R}^d\setminus B(0,R)} K(|y|)f(x-y)dy. \]

By our assumptions on $K$, we get that $| \int_{\mathbb{R}^d\setminus B(0,R)} K(|y|)f(x-y)dy | \leq K(R)\|f\|_1 \to 0$ and $| \int_{B(0,1/R)} K(|y|)f(x-y)dy | \leq \|f\|_\infty \int_{B(0,1/R)} K(|y|)dy \to 0$ as $R \to +\infty$. Therefore for every $\eta > 0$, there exists $R_\eta$, not depending on $x$, such that
\[ \left| V_f(x) - \int_{B(0,R)\setminus B(0,1/R)} K(|y|)f(x-y)dy \right| \leq \eta \quad \forall R \geq R_\eta. \]

Finally, we fix $R > R_\eta$, and we observe that for all $|x| > 2R$, we get
\[
\left| \int_{B(0,R)\setminus B(0,1/R)} K(|y|)f(x-y)dy \right| = \left| \int_{B(x,R)\setminus B(x,1/R)} K(|y-x|)f(y)dy \right| \\
\leq K(1/R) \int_{B(x,R)} |f(y)|dy \leq K(1/R) \int_{\mathbb{R}^d\setminus B(0,|x|/2)} |f(y)|dy \to 0 \quad \text{as } |x| \to +\infty.
\]

So there exists $R' > 2R$ such that $\left| \int_{B(0,R')\setminus B(0,1/R')} K(|y|)f(x-y)dy \right| \leq \eta$ if $|x| > R'$.

In conclusion for all $\eta > 0$ there exists $R' > 0$ such that $|V_f(x)| \leq 2\eta$ for all $|x| > R'$. \(\square\)
Note that by the Riesz rearrangement inequality (see [23]) for every \( f \in L^1(\mathbb{R}^d) \) such that \( f \geq 0 \),
\[
\mathcal{I}(f) \leq \mathcal{I}(f^*) ,
\] (5.9)
where \( f^* \) is the spherical rearrangement of \( f \), that is
\[
f^*(x) = \int_0^{+\infty} \chi_{\{y \mid f(y) > t\}}(x) dt \quad \text{where} \quad \{y \mid f(y) > t\}^* = B(0,r), \text{with} \ \omega_d r^d = |\{y \mid f(y) > t\}|.
\]

We recall a well known result, see [23].

**Lemma 5.3.** Assume (5.5). Let \( r_\rho = (1/\rho \omega_d)^{1/d} \). There holds
\[
\sup_{m \in \mathcal{P}_{2,\rho}^*(\mathbb{R}^d)} \mathcal{I}(m) = \mathcal{I}(\rho \chi_{B_\rho}) > 0.
\]

**Proof.** It follows from the Riesz rearrangement inequality and the fact that minimizers of \( \mathcal{I}(f) \) in \( \mathcal{P}_{2,\rho}^*(\mathbb{R}^d) \) are characteristic functions, as it can be proven looking at the second variation of the functional (see for a similar argument the following Proposition 5.9). \( \square \)

**Remark 5.4.** Note that, due to the fact that \( W(x) \geq 0 \) and to Lemma 5.3 we get that for all \( m \in \mathcal{P}_{2,\rho}^*(\mathbb{R}^d) \) there holds
\[
-W(\rho \chi_{B_\rho}) \leq W(m) \leq \int_{\mathbb{R}^d} W(x)m(x) .
\]

### 5.2 Assumptions (REF, BDD)

We check that \( W \) defined in (5.1), under the standing assumptions in Section 5.1, satisfies the growth condition (BDD), and the reflection invariance (REF).

**Proposition 5.5.** Under the assumptions (5.2), (5.3), (5.5), (5.6), the functional \( W \) in (5.1) satisfies (REF, BDD).

Since \( W \) and \( W_0 \) differ by a constant, the same conclusion holds for \( W_0 \). Moreover, note that
\[
\frac{\partial}{\partial m} W(m) = W(x) - \int_{\mathbb{R}^d} K(|x - y|) m(dy) .
\]
Hence, as a direct consequence of the positivity of \( K \) assumed in (5.6), \( W \) (and \( W_0 \)) is “aggregating”, namely it satisfies \( \int_{\mathbb{R}^d} (f(x,m) - f(x,m')) d|m - m'| \leq 0 \) for all \( m, m' \).

**Proof.** We observe that by (5.3) and the symmetry properties of \( K \),
\[
W(\gamma(m)) = \int_{\mathbb{R}^d} W(x) \gamma m(dx) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(|x - y|) \gamma m(dx) \gamma m(dy) = \int_{\mathbb{R}^d} W(\gamma x) m(dx) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(|\gamma x - \gamma y|) m(dx) m(dy) = W(m)
\]
which is (REF). Finally, by (5.2) and Lemma 5.3 we get that for all \( m \in \mathcal{P}_{2,\rho}^*(\mathbb{R}^d) \), there holds
\[
C^{-1} \int_{\mathbb{R}^d} |x|^2 m(dx) - C - W(\rho \chi_{B_\rho}) \leq W(m) \leq C \int_{\mathbb{R}^d} |x|^2 m(dx) + C,
\]
which is (BDD). \( \square \)
5.3 Continuity properties of $\mathcal{W}$: assumptions (lsc) and (CON)

We provide continuity and semicontinuity properties of $\mathcal{W}$ (and $\mathcal{W}_0$). Let us first check that $\mathcal{W}$ is lower semicontinuous with respect narrow convergence, which implies (lsc).

**Proposition 5.6.** The functional $\mathcal{W}$ satisfies (lsc). In particular let $m_k, m$ be Borel probability measures on $\mathbb{R}^d$ such that $m_k \rightarrow m$ narrowly. Then the following holds.

(i) $$\liminf_{k} \int_{\mathbb{R}^d} \mathcal{W}(x)m_k(dx) \geq \int_{\mathbb{R}^d} \mathcal{W}(x)m(dx)$$

where $\mathcal{W}$ satisfies (5.2).

(ii) If moreover $m_k, m \in \mathcal{P}_{\rho}(\mathbb{R}^d)$, then

$$\lim_{k} \mathcal{I}(m_k) = \mathcal{I}(m)$$

where $\mathcal{I}$ has been introduced in (5.7) and $K$ satisfies (5.5).

**Proof.** The lower semicontinuity (lsc) follows from (i) and (ii), recalling the characterization of convergence in $\mathcal{P}_{\rho}(\mathbb{R}^d)$ given in Lemma 2.4. (i) follows by Lemma 2.3. We sketch briefly for completeness the proof of (ii). Similar arguments have been used in [12, Lemma 3.3].

We recall a well known inequality (see [12]), identifying $m_k, m$ with their densities:

$$\left| (\mathcal{I}(m_k))^{1/2} - (\mathcal{I}(m))^{1/2} \right| \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (m_k(x) - m(x))(m_k(y) - m(y))K(|x - y|)dxdy^{1/2}. $$

We fix $R > 0$ and we write, recalling the conditions on $K$,

$$\left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (m_k(x) - m(x))(m_k(y) - m(y))K(|x - y|)dxdy \right|$$

$$\leq \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (m_k(x) - m(x))(m_k(y) - m(y))K(|x - y|)\chi_{|x-y| \leq R}dxdy \right|$$

$$+ \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (m_k(x) - m(x))(m_k(y) - m(y))K(|x - y|)\chi_{|x-y| > R}dxdy \right|$$

$$\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (m_k(x) - m(x))(m_k(y) - m(y))K(|x - y|)\chi_{|x-y| \leq R}dxdy + K(R). \quad (5.10)$$

Observe now that $K(|x|)\chi_{|x| \leq R} \in L^1(\mathbb{R}^d)$. We define

$$F_k(x) = \int_{\mathbb{R}^d} m_k(y)K(|x - y|)\chi_{|x-y| \leq R}dy, \quad \text{and} \quad F(x) = \int_{\mathbb{R}^d} m(y)K(|x - y|)\chi_{|x-y| \leq R}dy$$

and we observe that $F_k, F \in L^1(\mathbb{R}^d)$.

We rewrite

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (m_k(x) - m(x))(m_k(y) - m(y))K(|x - y|)\chi_{|x-y| \leq R}dxdy$$

$$= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (m_k(x) - m(x))F(x)dx + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} m_k(x)(F_k(x) - F(x))dx. \quad (5.11)$$

Observe that since $m_k \rightarrow m$ narrowly and $\mu_k, m \leq \rho$, then $m_k \rightarrow m$ weak* in $L^\infty$, due to density of continuous functions in $L^1$.

Therefore

$$\lim_{k} \int_{\mathbb{R}^d} (m_k(x) - m(x))F(x)dx = 0. \quad (5.12)$$
Moreover \( \lim_k F_k(x) = F(x) \) for a.e. \( x \) and
\[
\|F_k\|_1 = \|K(|x|)\chi_{|x| \leq R}\|_1 \to \|F\|_1 = \|K(|x|)\chi_{|x| \leq R}\|_1.
\]
Therefore by Fatou lemma, \( F_k \to F \) in \( L^1(\mathbb{R}^d) \), which implies, recalling that \( m_k \leq \rho \) that
\[
\lim_k \int_{\mathbb{R}^d} (F_k(x) - F(x))m_k(x)dx = 0. \tag{5.13}
\]
So, using (5.12), (5.13) in (5.10) and recalling that \( K(R) \to 0 \) as \( R \to +\infty \), we get the conclusion. \( \square \)

We observe the following fact about uniformly integrability of narrowly convergent sequences of measures.

**Lemma 5.7.** Let \( W \) is as in (5.2) and \( \mu_k, \mu \in \mathcal{P}_2(\mathbb{R}^d) \) such that \( \mu_k \to \mu \) narrowly. Then, \( \lim_k d_2(\mu_k, \mu) = 0 \) if and only if \( \lim_k \int_{\mathbb{R}^d} W(x)\mu_k(dx) = \int_{\mathbb{R}^d} W(x)\mu(dx) \).

**Proof.** We observe that, by Lemma 2.4, \( \lim_k d_2(\mu_k, \mu) = 0 \) is equivalent to the fact that \( \mu_k \) has uniformly integrable 2-moments, that is
\[
\lim_{R \to +\infty} \sup_k \int_{\mathbb{R}^d \setminus B(0,R)} |x|^2 d\mu_k(x) = 0. \tag{5.14}
\]

Let \( R > 0 \), sufficiently large, such that \( RC - C^2 > 1 \) and \( RC^{-1} > 3 \), where \( C \) is the constant appearing in (5.2). We denote \( A_R := \{ W(x) \geq R \} \). Then by (5.2) we get \( \mathbb{R}^d \setminus B(0, \sqrt{RC^{-1} + 1}) \subseteq A_R \subseteq \mathbb{R}^d \setminus B(0, \sqrt{RC - C^2}) \). Then, recalling (5.2), we get
\[
\frac{C}{2} \sup_k \int_{\mathbb{R}^d \setminus B(0, \sqrt{RC^{-1} + 1})} |x|^2 d\mu_k(x) \leq \sup_k \int_{\mathbb{R}^d \setminus B(0, \sqrt{RC^{-1} + 1})} (C|x|^2 - C)d\mu_k(x)
\]
\[
\leq \sup_k \int_{\mathbb{R}^d \setminus B(0, \sqrt{RC^{-1} + 1})} W(x)d\mu_k(x) \leq \sup_k \int_{A_R} W(x)d\mu_k(x) \leq \sup_k \int_{\mathbb{R}^d \setminus B(0, \sqrt{RC - C^2})} W(x)d\mu_k(x)
\]
\[
\leq \sup_k \int_{\mathbb{R}^d \setminus B(0, \sqrt{RC - C^2})} (C^{-1}|x|^2 + C)d\mu_k(x) \leq (C^{-1} + C) \sup_k \int_{\mathbb{R}^d \setminus B(0, \sqrt{RC - C^2})} |x|^2 d\mu_k(x).
\]
Sending \( R \to +\infty \), we get that \( \mu_k \) has uniformly integrable 2-moments, that is (5.14) holds, if and only if \( W \) is uniformly integrable with respect to \( \mu_k \), that is
\[
\lim_{R \to +\infty} \sup_k \int_{\{W(x) \geq R\}} W(x)d\mu_k(x) = 0. \tag{5.15}
\]

Now observe that if \( \mu_k \to \mu \) narrowly, then \( \lim_k \int_{\{W(x) \leq R\}} W(x)d\mu_k(x) = \int_{\{W(x) \leq R\}} d\mu(x) \) for all \( R > 0 \). Therefore if \( \mu_k \to \mu \) narrowly, then \( \lim_k \int_{\mathbb{R}^d} W(x)d\mu_k(x) = \int_{\mathbb{R}^d} W(x)d\mu(x) \) if and only if (5.15) holds. This gives the conclusion. \( \square \)

We finally prove the continuity property of \( \mathcal{W} \) that will entail the property (CON) for \( \mathcal{W}_0 \).

**Proposition 5.8.** Assume that \( m_k \in \mathcal{P}^r_{2,\rho}(\mathbb{R}^d) \) is such that \( \lim_k \mathcal{W}(m_k) = -\mathcal{I}(\rho\chi_{B_R}) \), with the notation of Lemma 5.3. Then up to passing to a subsequence there exists \( m \in \mathcal{P}^r_{2,\rho}(\mathbb{R}^d) \) such that \( \mathcal{W}(m) = -\mathcal{I}(\rho\chi_{B_R}) \) and \( \lim_n d_2(m_n, m) = 0 \).

**Proof.** By the growth condition (BDD), we get that \( m_n \) are uniformly bounded in \( \mathcal{P}_2(\mathbb{R}^d) \). So, by Lemma 2.3 up to passing a subsequence we get that there exists \( m \in \mathcal{P}_p(\mathbb{R}^d) \) such that \( m_n \to m \) in \( \mathcal{P}_p(\mathbb{R}^d) \) for all \( p < 2 \), and a posteriori by the growth condition (5.2), and the lower semicontinuity in Proposition 5.6 we have that \( \mathcal{W}(m) \leq -\mathcal{I}(\rho\chi_{B_R}) \), and so, recalling Remark 5.4 \( \mathcal{W}(m) = -\mathcal{I}(\rho\chi_{B_R}) \). Since, by Proposition 5.6 \( \lim_n \mathcal{I}(m_n) = \mathcal{I}(m) \) this implies in particular that \( \lim_n \int_{\mathbb{R}^d} W(x)m_k(dx) = \int_{\mathbb{R}^d} W(x)m(dx) \), so by Lemma 5.7 \( \lim_n d_2(m_n, m) = 0 \). \( \square \)
5.4 Minimizers for the stationary problem: the assumption (Z)

We start proving existence and qualitative properties of minimizers of $\mathcal{W}(m)$ in the set $\mathcal{P}^\prime_{\mathcal{Z},\mu}(\mathbb{R}^d)$. Then, in Proposition 5.11, we show that (Z) is satisfied by $\mathcal{W}$ (up to a renormalization constant $\mathcal{I}(\rho\chi_{B_\rho})$) under the standing assumptions.

The first result is about qualitative properties of minimizers, by looking at the first and second variation of the functional.

Proposition 5.9. Assume (5.5) and (5.2). Let $m \in \mathcal{P}^\prime_{\mathcal{Z},\mu}(\mathbb{R}^d)$ be a minimizer of the functional $\mathcal{W}$. Then there exists a bounded, measurable set $E \subseteq \mathbb{R}^d$ such that $m = \rho\chi_E$.

Proof. The proof is based on analogous arguments as in [11, Prop. 5.4, Thm 5.7].

We start showing that $m$ has bounded support. We compute the first variation of the functional $\mathcal{W}$ as in [11, Lemma 5.3]. We sketch it briefly. First of all we introduce the following sets $S = \{x \mid m(x) = \rho\}$, $N = \{x \mid m(x) = 0\}$, where $S,N$ are subsets of the set of density points of $m$. Pick any $\psi, \phi \in L^1(\mathbb{R}^d, [0,\rho])$ such that $\int_{\mathbb{R}^d} \phi(x)dx = \int_{\mathbb{R}^d} \psi(x)dx$, $\psi = 0$ a.e. in $S$ and $\phi = 0$ a.e. in $N$. Then $m + \lambda(\psi - \phi) \in \mathcal{P}^\prime_{\mathcal{Z},\mu}(\mathbb{R}^d)$ for $\lambda > 0$ sufficiently small. Using the fact that $\mathcal{W}(m) \leq \mathcal{W}(m + \lambda(\psi - \phi))$, we get, sending $\lambda \to 0^+$,

$$\int_{\mathbb{R}^d} (-2V_m(x) + W(x))(\psi(x) - \phi(x))dx \geq 0 \quad (5.16)$$

where $V_m$ is the potential of $m$ defined in (5.8). If we choose $\psi, \phi$ such that $\psi = \phi = 0$ in $N \cup S$, then we can exchange the role of $\phi, \psi$ in (5.16) and obtain

$$\int_{\mathbb{R}^d \setminus (S \cup N)} (-2V_m(x) + W(x))(\psi(x) - \phi(x))dx = 0$$

for all $\phi - \psi \in L^1(\mathbb{R}^d)$, such that $\phi - \psi = 0$ in $S \cup N$, with $\int_{\mathbb{R}^d}(\phi - \psi)dx = 0$.

This implies by the fundamental lemma of calculus of variations, that there exists a constant $c$ such that

$$-2V_m(x) + W(x) = c \quad x \in \mathbb{R}^d \setminus (N \cup S). \quad (5.17)$$

Using this fact in (5.16) we get, taking $\psi = 0$ in $S \cup N$, and observing that $\int_{\mathbb{R}^d \setminus (S \cup N)} \psi dx = \int_{\mathbb{R}^d} \phi dx$,

$$0 \leq \int_S (2V_m(x) - W(x))\phi(x)dx + c \int_{\mathbb{R}^d \setminus (S \cup N)} (\psi - \phi)dx = \int_S (2V_m(x) - W(x) + c)\phi(x)dx$$

an analogously, taking $\phi = 0$ in $S \cup N$,

$$0 \leq \int_N (-2V_m(x) + W(x))\psi(x)dx + c \int_{\mathbb{R}^d \setminus (S \cup N)} (\psi - \phi)dx = \int_N (-2V_m(x) + W(x) - c)\psi(x)dx.$$

This implies that

$$\begin{cases} -2V_m(x) + W(x) \geq c & x \in N \\ -2V_m(x) + W(x) \leq c & x \in S, \end{cases} \quad (5.18)$$

Recalling that $W$ is coercive (see assumption (5.2)) and $V_m$ vanishes at infinity, see Lemma 5.2, we conclude from (5.17), (5.18) that necessarily $\mathbb{R}^d \setminus N$, that is the support of $m$, has to be bounded.

Now we show that $m(x) \in \{0,\rho\}$ for a.e. $x$. We compute the second variation of the functional as in [11, Lemma 5.5]. We take $\xi \in L^1(\mathbb{R}^d, [-1,1])$ such that $\xi = 0$ in $N \cup S$ and $\int_{\mathbb{R}^d} \xi dx = 0$. Then for $\lambda$ small we get that $m + \lambda \xi \in \mathcal{P}^\prime_{\mathcal{Z},\mu}(\mathbb{R}^d)$ and using the minimality of $m$ and (5.16), we get

$$\int_{\mathbb{R}^d \setminus (N \cup S)} \int_{\mathbb{R}^d \setminus (N \cup S)} \xi(x)\xi(y)K(|x-y|)dxdy \leq 0. \quad (5.19)$$
Assume now by contradiction that there are two Lebesgue points \( x, y \) of \( m \) such that \( 0 < m(x), m(y) < \rho \). Let \( d = |x - y| \). It is possible to find, for \( 0 < \varepsilon < d/4 \) sufficiently small, \( A(x, \varepsilon) \subseteq B(x, \varepsilon) \cap \mathbb{R}^d \setminus (N \cup S) \), \( A(y, \varepsilon) \subseteq B(y, \varepsilon) \cap \mathbb{R}^d \setminus (N \cup S) \), so that \( d(A(x, \varepsilon), A(y, \varepsilon)) \geq d/2 \), and moreover \( |A(x, \varepsilon)| = |A(y, \varepsilon)| \). Observe that if either \( t, z \in A(x, \varepsilon) \subseteq B(x, \varepsilon) \) or \( t, z \in A(y, \varepsilon) \subseteq B(y, \varepsilon) \), then \( |t - z| \leq 2\varepsilon \), and if \( t \in A(x, \varepsilon), z \in A(y, \varepsilon) \), then \( |t - z| \geq d(A(x, \varepsilon), A(y, \varepsilon)) \geq d/2 \).

We define \( \xi = \chi_{A(x, \varepsilon)} - \chi_{A(y, \varepsilon)} \), and in (5.19) we find, using the fact that \( K \) is decreasing, that

\[
0 \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(x)\xi(y)K(|x - y|)dxdy
= \int_{A(x, \varepsilon)} \int_{A(x, \varepsilon)} K(|t - z|)dtdz + \int_{A(y, \varepsilon)} \int_{A(y, \varepsilon)} K(|t - z|)dtdz - \int_{A(x, \varepsilon)} \int_{A(y, \varepsilon)} K(|t - z|)dtdz
\geq K(2\varepsilon)|A(x, \varepsilon)|^2 + K(2\varepsilon)|A(y, \varepsilon)|^2 - 2K(d/2)|A(x, \varepsilon)||A(y, \varepsilon)|
= |A(x, \varepsilon)|^2(K(2\varepsilon) - K(d/2)) > 0,
\]

which gives a contradiction. \( \square \)

We provide now the existence and characterization of minimizers to \( \mathcal{W} \).

**Theorem 5.10.** Under the assumptions (5.5), (5.2), the problem

\[
\min_{m \in \mathcal{P}_{2p}(\mathbb{R}^d)} \mathcal{W}(m)
\]

admits at least one solution. Each solution is given by \( \rho\chi_E \) for some measurable set \( E \) such that \( |E| = \rho^{-1} \). Moreover if \( \rho\chi_E \) is a minimizer then also \( \rho\gamma_E \) is a minimizer.

**Proof.** The result is an application of the direct method in calculus of variations. By Remark 5.4, \( \inf_{\mathcal{P}_{2p}(\mathbb{R}^d)} \mathcal{W} \geq -\mathcal{I}(\rho\chi_{B_\rho}) \). Let \( m_n \) be a minimizing sequence. By (BDD) and Lemma 2.4, up to a subsequence, there exists \( m \) such that \( m_n \rightarrow m \) narrowly and also weak* in \( L^\infty \). Again by the growth condition (BDD) and the lower semicontinuity property (5c), \( m \in \mathcal{P}_2(\mathbb{R}^d) \). Moreover \( \|m\|_\infty \leq \|m_n\|_\infty \leq \rho \), and so \( m \in \mathcal{P}_{2p}(\mathbb{R}^d) \) and, again by Proposition 5.6, \( \lim_n \mathcal{W}(m_n) \geq \mathcal{W}(m) \), which implies that \( m \) is a minimizer. Finally by Proposition 5.9 \( m = \rho\chi_E \) for some bounded measurable set \( E \). The fact that \( \rho\chi_{\gamma E} \) is still a minimizers comes from (5.3).

**Proposition 5.11.** Assume (5.2), (5.3) and (5.4). Then \( \inf_{\mathcal{P}_{2p}(\mathbb{R}^d)} \mathcal{W} = -\mathcal{I}(\rho\chi_{B_\rho}) \) for \( r_\rho = (\omega d\rho)^{-1/d} \), and all the minimizers of (5.1) are given by \( \mathcal{M}^+ \cup \mathcal{M}^- \), where \( \mathcal{M}^- = \gamma\mathcal{M}^+ \) and

\[
\mathcal{M}^+ = \{\rho\chi_E \text{ where } E = B(x', (\omega d\rho)^{-1/d}) \subseteq B(a^+, r) \text{ for some } x' \in \mathbb{R}^d\}.
\]

If \( \rho = \rho_0 \), then \( \mathcal{M}^+ = \{\rho\chi_{B(a^+, r)}\} \).

\( \mathcal{M}^+ \) and \( \mathcal{M}^- \) are compact subsets of \( \mathcal{P}_2(\mathbb{R}^d) \) and \( d_2(\mathcal{M}^+, \mathcal{M}^-) > 0 \).

**Proof.** Let \( \rho\chi_E \) be a minimizer. We get, recalling Remark 5.4

\[
-\mathcal{I}(\rho\chi_{B(0,r_\rho)}) \leq \rho \int_E \mathcal{W}(x)dx - \mathcal{I}(\rho\chi_E) \leq \rho \int_{B(a^+,r_\rho)} \mathcal{W}(x)dx - \mathcal{I}(\rho\chi_{B(0,r_\rho)}).
\]

Note that under assumption (5.3), \( \omega d\rho \leq \omega dr \), then \( \mathcal{W}(\rho\chi_{B(a^+, r)}) = -\mathcal{I}(\rho\chi_{B(0,r_\rho)}) \), and so \( \rho\chi_{B(a^+, r_\rho)} \) are minimizers. Moreover, due to Lemma 5.3 \( \int_E \mathcal{W}(x)dx = 0 \) for every \( E \) such that \( \rho\chi_E \) is a minimizer. If \( r_\rho < r \), there are infinitely many minimizers, which are given by all possible balls \( B(x', r_\rho) \subseteq B(a^+, r) \), whereas if \( r_\rho = r \), the only minimizers are \( B(a^+, r) \).

The compactness of \( \mathcal{M}^\pm \) is straightforward. To evaluate \( d_2(\mathcal{M}^+, \mathcal{M}^-) \), we make use of the standard duality formula for \( d_1 \) (note that elements of \( \mathcal{M}^\pm \) have bounded support), see
e.g. [2] equation (7.1.2)]. Namely, for any \( \bar{m}^+ \in \mathcal{M}^+ \) and \( \bar{m}^- \in \mathcal{M}^- \), \( \bar{m}^+ = \rho \chi_{\mathcal{E}^+} \), where \( \mathcal{E}^+ = B(x', r_\rho) \subseteq B(a^+, r) \) and \( \mathcal{E}^- = B(y', r_\rho) \subseteq B(a^-, r) \), and

\[
d_2(\bar{m}^+, \bar{m}^-) \geq d_1(\bar{m}^+, \bar{m}^-) = \sup \left\{ \int_{\mathbb{R}^d} q d(\bar{m}^+ - \bar{m}^-) \mid q : \mathbb{R}^d \to \mathbb{R}; \text{1-Lipschitz} \right\}. \tag{5.21}
\]

Assume that the reflection \( \gamma \) is given by \( \gamma(x_1, x_2, \cdots, x_d) = (-x_1, x_2, \cdots, x_d) \) (the general case can be treated analogously). Since \( B(a^+, r_0) \cap B(a^-, r_0) = \emptyset \) and \( (a^+_1) = -(a^-_1) \), assuming without loss of generality that \( (a^+_1) > 0 \), we have \( (a^+_1) - r_0 > 0 \) (otherwise \( B(a^+, r_0) \) and \( B(a^-, r_0) \) would have non-empty intersection). Picking \( \varphi(x) = x_1 \) in (5.21) gives

\[
d_2(\bar{m}^+, \bar{m}^-) \geq \int_{B(x', r_\rho)} x_1 \, dx - \int_{B(y', r_\rho)} x_1 \, dx = |B(x', r_\rho)| x_1 - |B(y', r_\rho)| y_1 = \rho^{-1}(x'_1 - y'_1)
\geq 2\rho^{-1}(a^+_1 - r_0 + r_\rho),
\]

since \( B(x', r_\rho) \subseteq B(a^+, r) \) and \( B(y', r_\rho) \subseteq B(a^-, r) \), yielding the conclusion.

Therefore,

\[
\mathcal{W}_0(m) = W(m) + \mathcal{I}(\rho \chi_{\mathcal{B}_r}).
\]

The proof of Theorem 5.13 then follows by Theorem 5.10 and Proposition 5.11. Moreover,

**Corollary 5.12.** Under the assumptions (5.2), (5.3), (5.4), (5.5), (5.6), \( \mathcal{W}_0 \) satisfies (Z), (Isq), (BDD), (CON) and (REF).

**References**

[1] F. Alessio and P. Montecchiari. Brake orbit solutions for semilinear elliptic systems with asymmetric double well potential. *J. Fixed Point Theory Appl.*, 19(1):691–717, 2017.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

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

[4] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. *Active particles. Vol. 1. Advances in theory, models, and applications*, pages 141–171, 2017.

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

[6] P. Cardaliaguet and M. Mastöro. Weak kamin theory for potential mfg. *Journal of Differential Equations*, 2019.
[7] P. Cardaliaguet, A. R. Mézés, and F. Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM J. Control Optim.*, 54(5):2672–2709, 2016.

[8] P. Cardaliaguet and A. Porretta. Long time behavior of the master equation in mean field game theory. *Anal. PDE*, 12(6):1397–1453, 2019.

[9] A. Cesaroni and M. Cirant. Concentration of ground states in stationary mean-field games systems. *Anal. PDE*, 12(3):737–787, 2019.

[10] A. Cesaroni and M. Cirant. Periodic orbits and $\Gamma$-convergence for interacting particle systems with aggregation. *in preparation*, 2019.

[11] A. Cesaroni and M. Novaga. The isoperimetric problem for nonlocal perimeters. *Discrete Contin. Dyn. Syst. Ser. S*, 11(3):425–440, 2018.

[12] R. Choksi, R. C. Fetecau, and I. Topaloglu. On minimizers of interaction functionals with competing attractive and repulsive potentials. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 32(6):1283–1305, 2015.

[13] M. Cirant. On the existence of oscillating solutions in non-monotone mean-field games. *J. Differential Equations*, 266(12):8067–8093, 2019.

[14] M. Cirant and L. Nurbekyan. The variational structure and time-periodic solutions for mean-field games systems. *Minimax Theory Appl.*, 3(2):227–260, 2018.

[15] M. Fornasier, S. Lisini, C. Orrieri, and G. Savarè. Mean-field optimal control as $\Gamma$-limit of finite agent controls. *European Journal of Applied Mathematics*, pages 1–34, 2019.

[16] G. Fusco, G. F. Gronchi, and M. Novaga. Existence of periodic orbits near heteroclinic connections. *Minimax Theory Appl.*, 4(1):113–149, 2019.

[17] D. A. Gomes and T. Seneci. Displacement convexity for first-order mean-field games. *Minimax Theory Appl.*, 3(2):261–284, 2018.

[18] P. J. Graber, A. R. Mézáros, F. J. Silva, and D. Tonon. The planning problem in mean field games as regularized mass transport. *Calc. Var. Partial Differential Equations*, 58(3):Art. 115, 28, 2019.

[19] M.-O. Hongler. Exactly solvable gaussian and non-gaussian mean-field games and collective swarms dynamics. *arXiv preprint*, https://arxiv.org/abs/1811.07891, 2018.

[20] 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.

[21] 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.

[22] H. Lavenant and F. Santambrogio. Optimal density evolution with congestion: $L^\infty$ bounds via flow interchange techniques and applications to variational mean field games. *Comm. Partial Differential Equations*, 43(12):1761–1802, 2018.

[23] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997.

[24] M. Masoero. On the long time convergence of potential MFG. *NoDEA Nonlinear Differential Equations Appl.*, 26(2):Art. 15, 45, 2019.

[25] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[26] A. Monteil and F. Santambrogio. Metric methods for heteroclinic connections in infinite dimensional spaces. *Indiana Univ. Math. J.*, to appear, 2019. *arXiv preprint*, https://arxiv.org/abs/1709.02117.

[27] C. Orrieri, A. Porretta, and G. Savarè. A variational approach to the mean field planning problem. *J. Funct. Anal.*, 277(6):1868–1957, 2019.

[28] P. H. Rabinowitz. Periodic solutions of Hamiltonian systems: a survey. *SIAM J. Math. Anal.*, 13(3):343–352, 1982.

[29] 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.

[30] D. Ullmo, I. Swiecicki, and T. Gobron. Quadratic mean field games. *arXiv preprint*, https://arxiv.org/abs/1708.07730, 2017.
[31] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Bifurcation analysis of a heterogeneous mean-field oscillator game model. In Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference, CDC-ECC 2011, pages 3895–3900, 2011.