DETERMINISTIC MEAN FIELD GAMES WITH JUMPS AND MIXED VARIATIONAL STRUCTURE

ANNETTE DUMAS, FILIPPO SANTAMBROGIO

ABSTRACT. We consider in this paper some Mean field Game problems where the agents have an individual cost composed of an action penalizing their motion and a running cost depending on the density of all players. This running cost involves two parts: one which is variational but possibly non-smooth (typically, a local cost of the form $f'(\rho(x))$ where $\rho(x)$ is the density of the distribution of players at point x) and one which is smooth but non-variational (typically, a non-local cost involving convolutions). In order to show that this is a very general framework, we consider both the case where the action is of kinetic energy type, i.e. it is an integral of a power of the velocity, and the case where it counts the number of jumps of the curve, more adapted to some real estate and urban planning models. We provide an existence result of a "formal" equilibrium by Kakutani's fixed point theorem. "Formal" means that we do not prove that such an equilibrium is a measure on curves concentrated on optimal curves but only that it solves a variational problem whose optimality conditions formally correspond to this. We then rigorously prove that optimizers of this variational problem are indeed concentrated on optimal curves for an individual problem. Both the existence and these optimality conditions for the minimizers are studied in the kinetic and in the jumps case. We then prove that the opimization among measures on curves (the Lagrangian framework) can be reduced to a minimization among curves of measures (the Eulerian framework) by proving a representation of the action functional. This is classical in the kinetic case and involves the Wasserstein distances W_p, while it lets the total variation norm appear in the jump case. We then prove that, under some assumptions, the solution of the variational problem expressed in Eulerian language depends in a Lipschitz continuous way on the data, which can prove that the fixed point argument for the equilibrium can be reformulated as the fixed point of a Lipschitz uni-valued map. Under smallness assumptions on some data, this becomes a contraction and the equilibrium can be found by Banach fixed point. This allows for efficient numerical computations, based on the solution of a non-smooth convex optimization problem, which we present at the end of the paper.

1. INTRODUCTION

The goal of this paper is to study some classes of Mean Field Games where each negligible player chooses a trajectory $\gamma : [0,T] \to \Omega$ and the cost that she pays is of the form

$$A(\gamma) + \int_0^T W_{[\rho]}(t,\gamma(t))dt + \psi(\gamma(T)), \tag{1}$$

where A is a functional defined on the set of possible trajectories and penalizing their variations, and $W_{[\rho]}$ is a running cost acting on the position of the player (i.e., $\gamma(t)$) but depending itself on the overall distribution of the players (i.e., ρ).

We will consider the case where $W_{[\rho]}$ can be decomposed into two parts, one with a variational structure and one with a more continuous dependence on the data, i.e.

$$W_{[\rho]}(t,x) = \frac{\delta I}{\delta m}[\rho](x) + F(t,\rho,x).$$

Here $I : \mathscr{P}(\Omega) \to \overline{\mathbb{R}}$ is a functional defined on the space of measures and $\frac{\delta I}{\delta m}[\rho]$ is its first variation at a measure ρ, characterized by

$$\lim_{\varepsilon \to 0} \frac{I(\rho + \varepsilon\chi) - I(\rho)}{\varepsilon} = \int \frac{\delta I}{\delta m}[\rho](x) d\chi(x)$$

for suitable perturbations χ. This part of the function W is the one which has a variational structure.

The function $F : [0,T] \times \mathscr{P}(\Omega) \times \Omega \to \mathbb{R}$ is supposed to be continuous in all its variables, when endowing the space of measures with the topology of weak-* convergence (of course the continuity in t is not strictly needed, and Caratheodory functions would be suitable as well). It is important to note that a typical example of function which can play the role of $\frac{\delta I}{\delta m}$ but not of F is, for instance, the one associating with every absolutely continuous measure ρ (identified with its density) and a point x a quantity of the form $f(\rho(x))$: indeed, this is not continuous in x in general (since we did not require the density to be continuous) and not in ρ neither (since the weak-* convergence is not enough to let the density at each point converge). On the other hand, several examples exist of terms which are continuous but do not admit a variational structure, such as $f((\eta * \rho)(x))$, or, to make things more complicated, $(\rho, x) \mapsto u(x)$, where

Universite Claude Bernard Lyon 1, ICJ UMR5208, CNRS, École Centrale de Lyon, INSA Lyon, Université Jean Monnet, 69622 Villeurbanne cedex, France, dumas,santambrogio@math.univ-lyon1.fr.

u is defined as the distance function to a given target computed according to a metric depending on ρ, i.e. the viscosity solution to, for instance, a PDE of the form $|\nabla u| = g(\eta * \rho)$ with Dirichlet conditions $u = 0$ on the target set.

The goal in the theory of MFG is to find an equilibrium configuration, i.e. a measure Q on the space $\mathscr{C}$ of curves, representing the distribution of players among possible strategies (i.e., paths), which is concentrated on curves which optimize the cost given in (1) when taking $\rho_t := (e_t)_\# Q$ (e_t denotes the evaluation map, so that ρ_t is the distribution of players in the state space Ω at time t). When $I = 0$ the existence of an equilibrium can be obtained using the Kakutani's multivalued fixed point theorem, in the following way: with every Q we associate the curve of measures ρ_t, then the set of optimal curves for (1), and then the set $\mathscr{S}(Q)$ of measures $\tilde{Q}$ concentrated on optimal curves and such that $(e_0)_\#\tilde{Q}$ is equal to the (prescribed) initial distribution ρ_0. The multivalued map $Q \mapsto \mathscr{S}(Q)$ is convex-valued and, due to the continuity of F, its graph is closed. It is not hard to prove that all the other assumptions of Kakutani's theorem are satisfied and obtain the existence of a $\bar{Q}$ such that $\bar{Q} \in \mathscr{S}(\bar{Q})$, i.e. of an equilibrium.

This cannot be performed in the presence of a non-continuous running cost $\delta I/\delta m$. Yet, when $I \neq 0$ but $F = 0$, it is possible to use a variational strategy. Indeed, formally, one can easily see that any minimizer of the cost

$$Q \mapsto \mathscr{U}(Q) := \int_{\mathscr{C}} A(\gamma)dQ(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma) + \int_0^T I((e_t)_\# Q)dt$$

is indeed an equilibrium (this means that this game is a *potential game*, see [25]; see also the recent preprint [14]). A very rigorous statement in order to explain this will be presented in Section 3, since some regularity issues arise. Indeed, in many cases the quantity $W_{[\rho]}$ would not be well-defined at every point, if we think at the examples where it depends on the value of the density $\rho_t(x)$, which is itself only defined for a.e. x. Choosing a precise representative and proving a rigorous statement about the optimality of almost every trajectory will be the object of Section 3, depending on different choices of the penalization A.

One of the main goals of this paper is to discuss fixed-point strategies for the existence of equilibria when the two parts of W, i.e. $\delta I/\delta m$ and F, coexist. In this case we look for a measure $\bar{Q}$ which optimizes

$$Q \mapsto \mathscr{U}_{\bar{Q}}(Q) := \int_{\mathscr{C}} A(\gamma)dQ(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma) + \int_0^T I[t,\bar{Q}]((e_t)_\# Q)dt$$

where

$$I[t,\bar{Q}](\rho) := I(\rho) + \int_\Omega F(\bar{\rho}_t, x)d\rho(x),$$

the measure $\bar{\rho}_t$ being defined via $\bar{\rho}_t = (e_t)_\#\bar{Q}$. This equivalently means that $\bar{Q}$ optimizes

$$Q \mapsto J_{\bar{Q}}(Q) := \int_{\mathscr{C}} A(\gamma)d(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma) + \int_0^T I((e_t)_\# Q)dt + \int_{\mathscr{C}}\int_0^T F(t,\bar{\rho}_t,\gamma(t))dtdQ(\gamma).$$

We can also study the existence of such a $\bar{Q}$ via a use of Kakutani's fixed point theorem, and the convexity of I is required in order to deal with a convex-valued map (the map associating with every $\bar{Q}$ the set of optimizers Q).

Section 2 is indeed devoted to the proof of the existence of a measure $\bar{Q}$ with optimizes $J_{\bar{Q}}$, i.e. a fixed point. Some technical assumptions are required, and we will prove that they are satisfied in the two main cases which are the object of this paper:

- the case where the penalization A stands for the number of jumps of a piecewise constant curve (*"the jump case"*); this case, which was the motivation of [13], has been inspired by real estate models where agents are inhabitants who can move inside a urban area, but do not do it continuously; in this case a natural space of trajectories would be the space of BV curves from $[0,T]$ to Ω;
- the case where the penalization A is the kinetic energy, or a similar integral quantity (*"the kinetic case*); in this case the space of trajectories with finite cost is the space of $W^{1,p}$ curves from $[0,T]$ to Ω and the cost is given by $A(\gamma) = \int \frac{1}{p}|\gamma'(t)|^p dt$.

Yet, when looking for an equilibrium we want a measure $\bar{Q}$ concentrated on curves which are optimal in the individual problem (1), and the equivalence between minimizing a global energy and being concentrated on optimal curves is easy from an informal point of view but not always easy to be made rigorous. Section 3 is devoted to this equivalence. This question was already studied, in a MFG setting, in [10] with a technique borrowed from [1]. The very same technique cannot be used in our setting in the jump case, because of the lack of separability of the BV space. This requires to develop a slightly different technique, which can then be applied to the kinetic case as well. In this case, we obtain a proof of the equivalence between minimizers and equilibria under some conditions on the different parameters involved (the exponents, and the dimension) which already played a role in the first papers on MFG.

Sections 2 and 3 are thus concerned with a Lagrangian description of the equilibria. In the language of mathematical physics, this means describing the motion by following each particle (here, particles are rational agents). Since they are assumed to be indistinguishable, instead of providing the trajectory of each of them, this is done by providing the information of how much mass follows each possible trajectory, which means a measure on the set $\mathscr{C}$ of paths. In Section 4 we translate this into an Eulerian framework, i.e. describing the motion with quantities which are associated to points in space-time. The main quantities are then the density $\rho(t,x)$ and the velocity field $v(t,x)$. Yet, the density plays the most important role here, both since in some cases the velocity can be deduced from it through the theory of curves in the Wasserstein space, and since in the jump case the velocity is not a well-defined quantity. We then look for a curve of measures $(\rho_t)_t$ instead of a measure on curves Q. In Section 4 we prove the equivalence, in particular in the non-trivial jump case, between the minimization of an energy on Q and that of an energy on ρ.

After this equivalence has been established, we look again for fixed points, but using the Eulerian description: this means that we look for a curve of measures $\bar\rho$ which minimizes an action including a cost of the form $\rho \mapsto \int_0^T \int_\Omega F(t,\bar\rho,x)d\rho_t(x)$. We study the Lipschitz constant, in the space $L^2([0,T]\times\Omega)$, of the map associating the optimal solution ρ with the data $\bar\rho$. We prove that such a map is in some cases a contraction, thus allowing to use Banach's fixed point theorem.

Such a contractive fixed point approach can also be used for numerical purposes, which is what we do in Section 6, in particular in the jump case. We adapt the techniques of [13], based on non-smooth optimization algorithms, and obtain numerical simulations of equilibrium evolutions in various examples.

2. Mixed MFG

In this section, we prove the existence of an equilibrium measure $\bar Q$, characterized as a fixed point of a multi-valued function. First, let us introduce the notations and the problem solved by each agent. The metric space $(\mathscr{C},\mathrm{d})$ will designate the set of curves followed by an agent in the space $\Omega \subset \mathbb{R}^d$ with $d \geq 1$ (of course, it would also be possible to consider other domains such as the torus). The set Ω is supposed to be compact. On the space $\mathscr{C}$ we will define an action $A : \mathscr{C} \to [0,+\infty]$ which represents the cost for moving. The choice of $\mathscr{C}$ is not very important since it is more important to know the set of curves with finite action. In particular, we require $\{\gamma \in \mathscr{C} \,:\, A(\gamma) < +\infty\} \subset BV([0,T])$. We recall that we have

$$\mathrm{BV}([0,T],\Omega) = \{\gamma\colon [0,T] \to \bar\Omega \,:\, TV(\gamma,[0,T]) := \sup_{0=t_0<\ldots<t_n=T} \sum_{i=1}^n |\gamma(t_i)-\gamma(t_{i-1})| < \infty \}.$$

For BV curves we choose a particular representative: for simplicity we can take the left-continuous representative (we will explain in a while why left-continuity instead of right-continuity). This representative is well-defined and unique. With this choice, curves with finite action will always be considered to be defined pointwisely.

To be explicit, the main setting that we consider is the one where we take for $\mathscr{C}$ the set of measurable curves defined on $[-1,T]$, valued into Ω, and constant on $[-1,0]$. These curves are identified by a.e. equivalence, and we endow it with the L^1 distance. This space is a complete metric Polish space. We spend few words on this choice of modeling:

- Even if curves with finite action are all BV we do not want to choose $\mathscr{C} = \mathrm{BV}$ since this would suggest to endow it with the BV norm, which would make our compactness results fail.
- We do not want neither to choose $\mathscr{C} = \mathrm{BV}$ and endow it with the L^1 norm since this would not make it a complete metric space and some measure-theoretic tools that we use require to be in a Polish space.
- The choice of the left-continuous representative together with the choice of extending the curves to $[-1,T]$ instead of only $[0,T]$ allows to give a precise meaning of the initial point $\gamma(0)$, but also allows for immediate jumps at $t = 0^+$. Another possibility would have been to require curves to be right-continuous at $t = 0$ and, instead of imposing initial data of the form $\gamma(0) = x_0$ penalizing the difference $|\gamma(0)-x_0|$ as done in [13], but this would have made the presentation much heavier. On the other hand, we never impose Dirichlet boundary conditions at $t = T$, so there is no need to do the same for the final point $\gamma(T)$ and we can just think that the final penalization ψ includes the possibility to jump at $t = T$.
- Topologically, the only effect of extending curves in a constant way to $[-1,0]$ is that the convergence in $\mathscr{C}$ means convergence for a.e. $t \in (0,T]$ and at $t = 0$.
- When only continuous curves have finite action (we will see later that we consider two cases: the jump case where the BV setting is the adapted one, and the kinetic case where curves with finite action are continuous) it would also be possible to choose $\mathscr{C}$ to be the set of continuous curves valued into Ω, endowed with the uniform convergence instead of the L^1 convergence, and many difficulties would disappear.

For all $t \in [0,T]$, the evaluation map e_t is defined as follows:

$$\begin{aligned} e_t :\quad \mathscr{C} &\to \Omega \\ \gamma &\mapsto \gamma(t). \end{aligned}$$

Let $m_0 \in \mathscr{P}(\Omega)$ be a probability measure over the space Ω. The set $\mathscr{P}_{m_0}(\mathscr{C})$ will be the set of probability measures Q over $\mathscr{C}$ such that $\int_{\mathscr{C}} A dQ < +\infty$ and such that the push-forward measure by the evaluation map satisfies $e_0\#Q = m_0$.

We recall that for all $t \in [0,T]$, the push-forward measure $e_t\#Q$ is characterized by

$$\forall \varphi \in C_b(\Omega), \quad \int_\Omega \varphi(x)d(e_t\#Q)(x) = \int_{\mathscr{C}} \varphi(\gamma(t))dQ(\gamma),$$

where $C_b(\Omega)$ is the set of continuous and bounded functions on Ω. The measures $e_t\#Q$ are well-defined whenever $\int A dQ < +\infty$ since in this case Q is concentrated on left-continuous curves.

Let $\bar{Q} \in \mathscr{P}_{m_0}(\mathscr{C})$. Each agent whose starting point is x_0 seeks to solve the following optimization problem:

$$\min_{\{\gamma\in\mathscr{C};\ \gamma(0)=x_0\}} J_{(\bar{Q})}(\gamma) := A(\gamma) + \int_0^T \frac{dI}{dm}(e_t\#\bar{Q})(\gamma(t))dt + \int_0^T F(t,\gamma(t),e_t\#\bar{Q})dt + \psi(\gamma(T)). \tag{2}$$

In this section, we make the following hypotheses:

- (H_1) A is lower semi-continuous and coercive (its sublevel sets are compact in $\mathscr{C}$), and $A(\gamma) = 0$ for any constant curve γ (this simplifying assumption is the reason not to include $\psi(\gamma(T))$ in the action A); the function ψ is also lower semi-continuous and we suppose $\psi \geq 0$;
- (H_2) $I\colon \mathscr{P}(\Omega) \to [0,+\infty]$ is convex, l.s.c. for the weak convergence, and admits a first variation dI/dm; we also assume $I(m_0) < +\infty$;
- (H_3) $F\colon [0,T]\times\Omega\times\mathscr{P}(\Omega) \to \mathbb{R}$ is continuous with respect to its three variables, and we suppose $F \geq 0$;
- (H_4) the space $\mathscr{C}$ is such that if $Q_n \overset{*}{\rightharpoonup} Q$ in $\mathscr{P}_{m_0}(\mathscr{C})$ with $\int A dQ_n \leq C$, then up to a subsequence, we have $e_t\#Q_n \underset{n}{\overset{*}{\rightharpoonup}} e_t\#Q$, for a.e t.

The goal is to find a measure $\bar{Q} \in \mathscr{P}_{m_0}(\mathscr{C})$ whose support is included in the set of optimal trajectories of $J_{(\bar{Q})}$. This is equivalent to satisfying the following inequality:

$$\forall Q \in \mathscr{P}_{m_0}(\mathscr{C}),\ \int_{\mathscr{C}} J_{(\bar{Q})}(\gamma)d\bar{Q}(\gamma) \leq \int_{\mathscr{C}} J_{(\bar{Q})}(\gamma)dQ(\gamma). \tag{3}$$

One can see that (3) is the optimality condition of an optimization problem. In particular, if $\bar{Q}$ minimizes

$$Q \mapsto \mathscr{U}_{\bar{Q}}(Q) = \int_{\mathscr{C}} A(\gamma)dQ(\gamma) + \int_0^T \int_\Omega I(e_t\#Q)(x)dxdt + \int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\bar{Q})dt\ dQ(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma),$$

then it formally satisfies (3). A rigorous statement will be proved in Section 3, but here it motivates the study of the problem $\min_Q \mathscr{U}_{\bar{Q}}(Q)$. The goal then becomes to find $\bar{Q}$ such that $\bar{Q}$ is a minimizer of $\mathscr{U}_{\bar{Q}}$.

To find such a measure $\bar{Q}$, we will define an application H that associates with any Q_0 the set of minimizers of $\min_Q \mathscr{U}_{Q_0}(Q)$, and we will prove that it admits a fixed point $\bar{Q}$ thanks to Kakutani's fixed point theorem.

We recall the following useful fact.

Lemma 2.1. *If $(\mathscr{C},d)$ is a metric space, $f\colon \mathscr{C} \to \mathbb{R}$ is a l.s.c. function bounded from below, and $(Q_N)_N \subset \mathscr{P}(\mathscr{C})$ is a sequence of probability measures narrowly converging towards Q, then we have*

$$\int_{\mathscr{C}} f dQ \leq \liminf_N \int_{\mathscr{C}} f dQ_N.$$

For the proof, we refer to [28, Lemma 1.6].

Theorem 2.2. *Under the assumptions (H1–4), the multifunction*

$$\begin{aligned} H\colon \mathscr{P}_{m_0}(\mathscr{C}) &\rightrightarrows \mathscr{P}_{m_0}(\mathscr{C}) \\ \tilde{Q} &\longmapsto \underset{Q,e_0\#Q=m_0}{\operatorname{argmin}}\ \mathscr{U}_{\tilde{Q}}(Q) \\ &= \underset{Q,e_0\#Q=m_0}{\operatorname{argmin}} \left\{ \int_{\mathscr{C}} A(\gamma)dQ(\gamma) + \int_0^T I(e_t\#Q)dt + \int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q})dt\ dQ(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma) \right\} \end{aligned}$$

admits a fixed point.

Proof. To prove the theorem, we will apply Kakutani's fixed point theorem.

First of all, we choose a set $\Gamma \subset \mathscr{P}_{m_0}(\mathscr{C})$ which is compact for the narrow convergence and invariant for this multifunction.

For each C, let us define the set $\Gamma(C) := \{Q \in \mathscr{P}_{m_0}(\mathscr{C}) : \int A dQ \leq C\}$. The coercivity of A (Assumption (H1)) implies that this set is compact for the narrow convergence of probability measures on $\mathscr{C}$. We have to choose a suitable value of C such that for any $Q, \tilde{Q}$ with $Q \in H(\tilde{Q})$ we have $Q \in \Gamma(C)$. To do this, let us choose a reference measure Q_{still}, in order to have a uniform bound on $\int_{\mathscr{C}} A dQ$. We take as Q_{still} a measure concentrated on constant curves, i.e. the image of m_0 through the map associating with every $x \in \Omega$ the constant curve equal to x. For this measure, the quantity

$\mathscr{U}_{\tilde{Q}}(Q_{still})$ is bounded by a number C_0 independent of $\tilde{Q}$ (we use here $A(\gamma) = 0$ for constant curves, and $I(m_0) < +\infty$; the only dependence of $\mathscr{U}_{\tilde{Q}}(Q_{still})$ in terms of $\tilde{Q}$ is in the part with F, which is bounded). Since for every $Q \in H(\tilde{Q})$ we have

$$\int_{\mathscr{C}} A(\gamma)dQ(\gamma) \leq \mathscr{U}_{\tilde{Q}}(Q) \leq \mathscr{U}_{\tilde{Q}}(Q_{still}) \leq C_0,$$

we obtain $\int A(\gamma)dQ(\gamma) \leq C_0$. We then choose $C = C_0$ and from now on will set all our analysis in the set $\Gamma := \Gamma(C_0)$.

We then have a multifunction $H : \Gamma \to \Gamma$ and we need to verify the three following conditions:

(i) *For each measure $\tilde{Q} \in \Gamma$, the set $H(\tilde{Q})$ is convex.*

This is a consequence of the fact that the functional which is minimized over Q is convex in Q, and it is actually the only reason to require the convexity of I.

(ii) *For each $\tilde{Q} \in \Gamma$, the set $H(\tilde{Q})$ is non-empty.*

We have to prove the existence of a minimizer in the problem defining $H(\tilde{Q})$ and, as we said, we can restrict the minimization to the set Γ. Let $(Q_n)_n$ be a minimizing sequence of $\inf_{Q\in\Gamma} \mathscr{U}_{\tilde{Q}}(Q)$. Since Γ is compact for the narrow convergence of probability measures on $\mathscr{C}$, there exists a subsequence $(Q_{n_k})_k$ converging narrowly towards $Q \in \Gamma$. We re-extract a subsequence in order to apply, later on (H_4).

Let us estimate $\liminf_k \mathscr{U}_{\tilde{Q}}(Q_{n_k})$ as $k \to \infty$.

Since A and ψ are l.s.c. and bounded from below (the boundedness from below is a consequence of lower semi-continuity and coercivity for A), we have by Lemma 2.1,

$$\int_{\mathscr{C}} A(\gamma)dQ(\gamma) \leq \liminf_k \int_{\mathscr{C}} A(\gamma)dQ_{n_k}(\gamma) \tag{4}$$

$$\text{and } \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma) \leq \liminf_k \int_{\mathscr{C}} \psi(\gamma(T))dQ_{n_k}(\gamma). \tag{5}$$

Using (H_4) and after extracting a subsequence, we have $e_t\#Q_{n_k} \underset{*}{\rightharpoonup} e_t\#Q$ for a.e. t. By the lower semi-continuity of I and Fatou's Lemma, we obtain

$$\begin{aligned} \int_0^T I(e_t\#Q)\,dt &\leq \int_0^T \liminf_k I(e_t\#Q_{n_k})\,dt \\ &\leq \liminf_k \int_0^T I(e_t\#Q_{n_k})\,dt, \end{aligned} \tag{6}$$

and since F is continuous in its variables, we have

$$\int_{\mathscr{C}} \int_0^T F(t,\gamma(t),e_t\#\tilde{Q})dt\; dQ_{n_k}(\gamma) \underset{k\to\infty}{\longrightarrow} \int_{\mathscr{C}} \int_0^T F(t,\gamma(t),e_t\#\tilde{Q})dt\; dQ(\gamma). \tag{7}$$

Gathering the results (4), (6), (7) and (5) gives

$$\mathscr{U}_{\tilde{Q}}(Q) \leq \liminf_k \mathscr{U}_{\tilde{Q}}(Q_{n_k}) \leq \inf_{Q\in\Gamma} \mathscr{U}_{\tilde{Q}}(Q),$$

which proves that Q is a minimizer of $\mathscr{U}_{\tilde{Q}}$ and consequently, $H(\tilde{Q})$ is non-empty.

(iii) *The graph of H is closed.*

Let $(\tilde{Q}_n)_n$ be a sequence narrowly converging towards $\tilde{Q}_\infty$ and $Q_n \in H(\tilde{Q}_n)$ a sequence converging towards Q_∞. We will prove $Q_\infty \in H(\tilde{Q}_\infty)$.

Since $Q_n \in H(\tilde{Q}_n)$, we have for all $Q \in \Gamma$

$$\begin{aligned} &\int_{\mathscr{C}} A(\gamma)dQ_n(\gamma) + \int_0^T I(e_t\#Q_n)dt + \int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_n)dt\; dQ_n(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ_n(\gamma) \\ &\leq \int_{\mathscr{C}} A(\gamma)dQ(\gamma) + \int_0^T I(e_t\#Q)dt + \int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_n)dt\; dQ(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma). \end{aligned} \tag{8}$$

For each term in (8), we consider its limit when $n \to +\infty$.

Let us begin with $\int A(\gamma)dQ_n(\gamma)$ and $\int \psi(\gamma(T))dQ_n(\gamma)$. Again, by Lemma 2.1 we have

$$\int_{\mathscr{C}} A(\gamma)dQ_\infty(\gamma) \leq \liminf_n \int_{\mathscr{C}} A(\gamma)dQ_n(\gamma)$$

$$\text{and } \int_{\mathscr{C}} \psi(\gamma(T))dQ_\infty(\gamma) \leq \liminf_n \int_{\mathscr{C}} \psi(\gamma(T))dQ_n(\gamma).$$

For the second term, by Hypothesis (H_4) (up to extracting a subsequence), the lower semi-continuity of I and Fatou's Lemma, we have the inequality

$$\int_0^T I(e_t\#Q_\infty)dt \le \int_0^T \liminf_n I(e_t\#Q_n)dt \le \liminf_n \int_0^T I(e_t\#Q_n)dt.$$

When it comes to $\int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_n)dt\; dQ_n(\gamma)$, the continuity of F gives

$$\int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_n)dt\; dQ_n(\gamma) \underset{n\to+\infty}{\longrightarrow} \int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_\infty)dt\; dQ_\infty(\gamma).$$

In the right-hand side of (8), only one term depends on n. The dominated convergence theorem yields

$$\int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_n)dt\; dQ(\gamma) \underset{n}{\longrightarrow} \int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_\infty)dt\; dQ(\gamma).$$

Finally, passing to the limit $\liminf_n$ in (8) gives

$$\begin{aligned}&\int_{\mathscr{C}} A(\gamma)dQ_\infty(\gamma)+\int_0^T I(e_t\#Q_\infty)dt+\int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_\infty)dt\; dQ_\infty(\gamma)+\int_{\mathscr{C}}\psi(\gamma(T))dQ_\infty(\gamma)\\ &\le \int_{\mathscr{C}} A(\gamma)dQ(\gamma)+\int_0^T I(e_t\#Q)dt+\int_{\mathscr{C}}\int_0^T F(t,\gamma(t),e_t\#\tilde{Q}_\infty)dt\; dQ(\gamma)+\int_{\mathscr{C}}\psi(\gamma(T))dQ(\gamma).\end{aligned}$$

Since this inequality is true for all $Q\in\Gamma$, Q_∞ is a minimizer of $\mathscr{U}_{\tilde{Q}_\infty}$, namely $Q_\infty \in H(\tilde{Q}_\infty)$. This proves that the graph of H is closed.

Since the conditions (i), (ii), (iii) are verified, by Kakutani's theorem, we can conclude that H admits a fixed point in Γ. □

2.1. **The jump case.** The theorem 2.2 can be applied to the case where $A=S$, where the functional S describes the number of *jumps* of a curve γ. As we said, we choose $\mathscr{C}$ to be the space of measurable curves (defined on $[-1,T]$ but constant on $[-1,0]$) equipped with the L^1-norm and $S\colon \mathscr{C}\to[0,\infty]$ is defined via

$$S(\gamma)=\inf\{\#\{t\ ;\ \tilde{\gamma}\text{ is discontinuous in }t\ \}\ ;\ \tilde{\gamma}=\gamma\text{ a.e. and }\tilde{\gamma}\text{ is piecewise constant}\}.$$

Of course, $S(\gamma)=+\infty$ whenever γ does not admit a piecewise constant representative.

We note that we have

$$TV(\gamma;[0,T])\le \operatorname{diam}(\Omega)S(\gamma),$$

where $\operatorname{diam}(\Omega)$ is the diameter of the compact domain Ω.

The following proposition shows that the function S is lower semi-continuous and coercive:

Proposition 2.3. $S\colon\mathscr{C}\to\mathbb{N}\cup\{+\infty\}$ *is l.s.c. and coercive.*

Proof. We need to prove that for all $\alpha\in\mathbb{R}$, the set $E_\alpha:=\{\gamma\in\mathscr{C}\ ;\ S(\gamma)\le\alpha\}$ is compact. This proves at the same time lower semi-continuity (for which the closedness of E_α would have been enough) and coercivity.

We assume $\alpha\ge0$ since otherwise $E_\alpha=\emptyset$. Let $(\gamma_n)_n$ be a sequence in E_α. Take $\tilde{\gamma}_n$ a representative of γ_n which realizes exactly $S(\gamma_n)$ jumps. Up to extracting a subsequence there exists a number N such that $N-1\le\alpha$ and each $\tilde{\gamma}_n$ is of the form $\tilde{\gamma}_n=\sum_{i=0}^{N-1} z_i^n \mathbb{1}_{[a_i^n,a_{i+1}^n)}$. It is then possible to extract converging subsequences from $(z_i^n)_n$ and $(a_i^n)_n$ and prove that $\tilde{\gamma}_n$ converges to a curve γ of the form $\gamma=\sum_{i=0}^{N-1} z_i\mathbb{1}_{[a_i,a_{i+1})}$. This convergence will be very strong: from the convergence of the values z_i^n taken by the curves and of the intervals on which these values are taken we can deduce that we have a.e. and L^1 convergence (but not uniform convergence). Since γ is piecewise constant and defined using N intervals we have $S(\gamma)\le N-1\le\alpha$ (the inequality in $S(\gamma)\le N$ could be strict, since we do not know whether the points z_i are distinct). This proves the compactness of E_α since we extracted a converging subsequence, and the limit still belongs to the same set. □

There remains to verify that Assumption (H_4) holds true:

Proposition 2.4. *Let $(Q_n)_n$ be a sequence narrowly converging towards Q in $\mathscr{P}(\mathscr{C})$ with $\int_{\mathscr{C}} S(\gamma)dQ_n(\gamma)\le C<+\infty$. Then there exists a subsequence Q_{n_k} such that*

$$e_t\#Q_{n_k}\underset{k}{\overset{*}{\rightharpoonup}} e_t\#Q,\ \textit{a.e. } t.$$

Proof. For each n, let us define a measure μ_n on $[0,T]$ via

$$\mu_n(I) := \int_{\mathscr{C}} TV(\gamma;I)dQ_n(\gamma)$$

for every open interval I. The total mass of μ_n is equal to $\int_{\mathscr{C}} TV(\gamma;[0,T])dQ_n(\gamma) \leq \mathrm{diam}(\Omega)\int_{\mathscr{C}} S(\gamma)dQ_n(\gamma) \leq C$. The sequence μ_n is then bounded in the space of positive measures on $[0,T]$ and we can extract a converging subsequence $\mu_{n_k} \xrightarrow[k]{*} \mu$. We will then choose the corresponding subsequence Q_{n_k}.

We now claim that the convergence $e_t\#Q_{n_k} \xrightarrow[k]{*} e_t\#Q$ occurs for each t which is not an atom of μ, i.e. for all but a countable quantity of t.

We fix a test function $\phi \in \mathrm{Lip}(\Omega)$ and we need to consider $\int_{\mathscr{C}} \phi(\gamma(t))dQ_n(\gamma)$. The function $\mathscr{C} \ni \gamma \mapsto \Phi(\gamma) := \phi(\gamma(t))$ is not continuous, so this term does not pass easily to the limit. For this reason we fix $\varepsilon > 0$ and consider instead the function $\mathscr{C} \ni \gamma \mapsto \Phi_\varepsilon(\gamma) := \fint_{t-\varepsilon}^{t+\varepsilon} \phi(\gamma(s))ds$, which is continuous for the L^1 convergence of γ. In particular, we have $\lim_n \int_{\mathscr{C}} \Phi_\varepsilon(\gamma)dQ_n(\gamma) = \int_{\mathscr{C}} \Phi_\varepsilon(\gamma)dQ(\gamma)$.

Our goal is to prove $\lim_k \int_{\mathscr{C}} \Phi(\gamma)dQ_{n_k}(\gamma) = \int_{\mathscr{C}} \Phi(\gamma)dQ(\gamma)$. We use $|\Phi(\gamma)-\Phi_\varepsilon(\gamma)| \leq \mathrm{Lip}(\phi)TV(\gamma;[t-\varepsilon,t+\varepsilon])$ which allows to obtain

$$\left|\int_{\mathscr{C}} \Phi dQ_{n_k} - \int_{\mathscr{C}} \Phi_\varepsilon dQ_{n_k}\right| \leq \mathrm{Lip}\phi\, \mu_{n_k}([t-\varepsilon,t+\varepsilon]), \quad \left|\int_{\mathscr{C}} \Phi dQ - \int_{\mathscr{C}} \Phi_\varepsilon dQ\right| \leq \mathrm{Lip}\phi\, \mu([t-\varepsilon,t+\varepsilon])$$

This implies

$$\left|\int_{\mathscr{C}} \Phi dQ_{n_k} - \int_{\mathscr{C}} \Phi dQ\right| \leq \mathrm{Lip}\phi\, \left(\mu([t-\varepsilon,t+\varepsilon]) + \mu_{n_k}([t-\varepsilon,t+\varepsilon])\right) + \left|\int_{\mathscr{C}} \Phi_\varepsilon dQ_{n_k} - \int_{\mathscr{C}} \Phi_\varepsilon dQ\right|.$$

Using $\limsup_k \mu_{n_k}(E) \leq \mu(E)$, which holds for every closed set E, we obtain

$$\limsup_k \left|\int_{\mathscr{C}} \Phi dQ_{n_k} - \int_{\mathscr{C}} \Phi dQ\right| \leq 2\mathrm{Lip}\phi\, \mu([t-\varepsilon,t+\varepsilon]).$$

The arbitrariness of ε shows the desired limit as soon as $\mu(\{t\}) = 0$.

□

All the assumptions for Theorem 2.2 are verified, so we can deduce that there exists a measure $\bar{Q} \in \Gamma$ which is a fixed point, i.e. it minimizes $\mathscr{U}_{\bar{Q}}$ in the case $A = S$.

2.2. **The kinetic case.** Another application of Theorem 2.2 can be obtained by using $A = K$ with the function K defined as follows

$$K(\gamma) = \begin{cases} \int_0^T \frac{|\dot{\gamma}(t)|^p}{p} dt & \text{if } \gamma \in W^{1,p}([0,T]), \\ +\infty & \text{if not,} \end{cases} \tag{9}$$

for a given exponent p with $1 < p < \infty$. In this case we can either choose $\mathscr{C}$ to be again the same space as in the jump case, or even choose $\mathscr{C} = C^0([0,T];\Omega)$ endowed with the sup distance, i.e. with the topology of uniform convergence. This second choice simplifies some arguments and make this section independent of the previous one.

Proposition 2.5. *The application $K\colon \mathscr{C} \to \mathbb{R}$ is l.s.c. and coercive.*

Proof. It suffices to show that for all $\alpha \in \mathbb{R}$, the lower level-sets $E_\alpha = \{\gamma \in W^{1,p}([0,T]); \int_0^T \frac{|\dot{\gamma}(t)|^p}{p} dt \leq \alpha\}$ are compact for the uniform convergence (which also implies compactness for the L^1 convergence, in case the reader prefers to use $\mathscr{C}$ as in the previous section on the jump case).

If $\alpha < 0$, then $E_\alpha = \emptyset$.

Now suppose $\alpha \geq 0$. Let $(\gamma_n)_n \subset E_\alpha$ be a sequence, which we can suppose bounded in $W^{1,p}([0,T])$ (we also use the compactness of Ω in oder to bound the L^p norm of γ_n and not only of its derivative). By the Sobolev inequality we obtain for all $n \in \mathbb{N}$,

$$[\gamma_n]_{C^{0,1-\frac{1}{p}}([0,T])} = \sup_{s\neq t} \frac{|\gamma_n(t)-\gamma_n(s)|}{|t-s|^{1-1/p}} \leq \|\dot{\gamma}_n\|_{L^p([0,T])}.$$

Hence, the sequence $(\gamma_n)_n$ is equicontinuous and we can apply Arzelà-Ascoli's theorem. Thus, there exists a subsequence $(\gamma_{n_k})_k$ that converges uniformly towards a limit curve γ. It is also possible to extract a further subsequence so that $\dot{\gamma}_{n_k}$ weakly-* converges in L^p to a function, which can only be (identifying the limit with the distributional limit) $\dot{\gamma}$. The lower semi-continuity of the L^p-norm shows that we have

$$\int_0^T \frac{|\dot{\gamma}(t)|^p}{p} dt \leq \liminf_n \int_0^T \frac{|\dot{\gamma}_n(t)|^p}{p} dt \leq \alpha,$$

which proves that the limit γ belongs to E_α and hence the compactness of E_α. □

The last point to check is if assumption (H_4) is verified.

Proposition 2.6. *Let $(Q_n)_n$ be a sequence narrowly converging towards Q in $\mathscr{P}_{m_0}(\mathscr{C})$. Then*

$$e_t\#Q_n \xrightharpoonup[n]{*} e_t\#Q, \forall t.$$

Proof. We remind that in this case the space $\mathscr{C}$ has been endowed with the topology of uniform convergence. Hence, for each t, differently from what we faced in the proof of Proposition 2.4, the map $e_t : \mathscr{C} \to \Omega$ is continuous. Continuous maps transform narrowly converging sequences on $\mathscr{C}$ into narrowly converging sequences on Ω, hence the result is straightforward. □

Since all the hypotheses for Theorem 2.2 are verified, we can conclude also in this case the existence of a fixed point $\bar{Q} \in \Gamma$ that minimizes $\mathscr{U}_{\bar{Q}}$.

3. Equivalence between equilibria and optimizers

In the previous section, we showed that there exists a fixed point, i.e. a measure $\bar{Q}$ which minimizes $\mathscr{U}_{\bar{Q}}$. However, this does not prove yet that $\bar{Q}$ is a Nash equilibrium in the sense that it is supported on the set of optimal trajectories for the individual cost $J_{(\bar{Q})}$. The difficulty in proving this fact comes from the variational part dI/dm. What we need to prove is that, for a fixed function $V : [0,T]\times\Omega \to \mathbb{R}$, whenever V is exogenous or is obtained as $(t,x) \mapsto F(t,x,e_t\#Q)$, the measure Q which minimizes

$$\mathscr{U}(Q) = \int_{\mathscr{C}} A(\gamma)dQ(\gamma) + \int_0^T I(e_t\#Q)dt + \int_{\mathscr{C}}\int_0^T V(t,\gamma(t))dt\, dQ(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma) \tag{10}$$

is concentrated on curves which minimize

$$J_W(\gamma) := A(\gamma) + \int_0^T W(t,\gamma(t))dt + \psi(\gamma_T) \quad \text{for } W(t,x) = \frac{dI}{dm}[e_t\#Q](x) + V(t,x).$$

Note that we already used in (2) the notation $J_{(Q)}$. Compared to the above notation J_W by $J_{(Q)}$ we mean J_W for a specific function W depending on Q. We prefer here to have two independent notations since we will need to choose a specific function W to use.

The result will be proven in the case where we take

$$I(\rho) = \begin{cases} \int_\Omega f(\rho(x))dx & \text{if } \rho \ll \mathscr{L}^d, \\ +\infty & \text{if not.} \end{cases} \tag{11}$$

This functional is lower semi-continuous for the weak-* convergence whenever f is convex and superlinear.

In this case, the function dI/dm equals $f'(\rho)$ and is thus a priori only defined a.e. In general, this requires to choose a precise representative of it, but a first result that we will show will concern the case where *by chance* $f'(\rho)$ is a continuous function.

The regularity of the density ρ can be studied as an independent question and is better analyzed if the variational problem solved by Q is re-written as a problem involving ρ, hence in Eulerian terms. The equivalence between these two approaches will be addressed in the next section in the more delicate case $A = S$ (the jump case).

In this case we can see that minimizing $\mathscr{U}$ is equivalent to solving

$$\min_\rho \int_0^T \int_\Omega (|\partial_t\rho| + V\rho + f(\rho))dxdt + \Psi_0(\rho(0)) + \int_\Omega \psi d\rho(T) \tag{12}$$

for a suitable $\Psi_0 : \mathscr{P}(\Omega) \to \mathbb{R}$. This allows to use the results in [13], which indeed provide regularity of the optimal ρ (under suitable assumptions, the solution $\bar{\rho}$ of (12) is Lipschitz in time valued in $L^2(\Omega)$ and its regularity in space depends on the regularity of V).

Let us begin with the next Proposition which provides the optimality condition for Problem (10).

Proposition 3.1. *Let $\bar{Q}$ be a solution of Problem* (10) *in the case where I is given by* (11). *Let $Q \in \Gamma$ be a measure such that, setting $\rho(t,\cdot) = e_t\#Q$, we have*

$$\int_0^T \int_\Omega f(\rho(t,x))dxdt < \infty. \tag{13}$$

Then we have

$$\int_{\mathscr{C}} A(\gamma)d\bar{Q}(\gamma)+\int_0^T\int_\Omega \big(V(t,x)+f'(\bar\rho(t,x))\big)\,\bar\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))d\bar{Q}(\gamma)$$
$$\leq \int_{\mathscr{C}} A(\gamma)dQ(\gamma)+\int_0^T\int_\Omega \big(V(t,x)+f'(\bar\rho(t,x))\big)\,\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ(\gamma). \tag{14}$$

Proof. Given $\varepsilon\in[0,1]$, let us define $Q_\varepsilon=\bar{Q}+\varepsilon(Q-\bar{Q})$ and

$$u(\varepsilon)=\int_{\mathscr{C}} A(\gamma)dQ_\varepsilon(\gamma)+\int_0^T\int_\Omega V(t,x)d(e_t\#Q_\varepsilon)(x)dt+\int_0^T\int_\Omega f(e_t\#Q_\varepsilon)(x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ_\varepsilon(\gamma).$$

By the optimality of $\bar{Q}$, we have $u(\varepsilon)\geq u(0)$. We want to compute $u'(0)$. All terms in u are affine in ε except for $\varepsilon\mapsto\int_0^T\int_\Omega f(e_t\#Q_\varepsilon)(x)dxdt=\int_0^T\int_\Omega f((1-\varepsilon)\bar\rho+\varepsilon\rho))dxdt$ (notice $\rho_t:=e_t\#Q$ and $\bar\rho_t:=e_t\#\bar{Q}$). Yet, the convexity of f provides

$$\frac{f((1-\varepsilon)\bar\rho+\varepsilon\rho)-f(\bar\rho)}{\varepsilon}\leq f(\rho)-f(\bar\rho).$$

The limit as $\varepsilon\to0^+$ of the l.h.s. is $f'(\bar\rho)(\rho-\bar\rho)$. The above inequality, together with the integrability condition (13), show in particular that we have $f'(\bar\rho)(\rho-\bar\rho)<+\infty$ a.e. and that the positive part of this function is L^1. Then, we can obtain by Fatou's lemma

$$\int f'(\bar\rho)(\rho-\bar\rho)\geq\lim_{\varepsilon\to0}\int\frac{f((1-\varepsilon)\bar\rho+\varepsilon\rho))-f(\bar\rho)}{\varepsilon}.$$

Therefore we have:

$$0\leq u'(0)\leq\int_{\mathscr{C}} A(\gamma)d(Q-\bar{Q})(\gamma)+\int_0^T\int_\Omega V(t,x)d(e_t\#(Q-\bar{Q}))(x)dt$$
$$+\int_0^T\int_\Omega f'(e_t\#\bar{Q})(x)d(e_t\#(Q-\bar{Q}))(x)\,dt+\int_{\mathscr{C}}\psi(\gamma(T))d(Q-\bar{Q})(\gamma).$$

This inequality is exactly (14). □

We first re-write the inequality (14) in a more suitable way using the notation J_W. We observe that we have, for any $Q\in\Gamma$ with $\rho_t:=e_t\#Q$,

$$\int_{\mathscr{C}} A(\gamma)dQ(\gamma)+\int_0^T\int_\Omega W(t,x)\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ(\gamma)=\int_{\mathscr{C}} J_W(\gamma)dQ(\gamma).$$

This means that the Proposition 3.1 provides the inequality

$$\int_{\mathscr{C}} J_W(\gamma)d\bar{Q}(\gamma)\leq\int_{\mathscr{C}} J_W(\gamma)dQ(\gamma)$$

for $W=V+f'(\bar\rho)$ and any Q satisfying the condition $\int_0^T\int_\Omega f(\rho(t,x))dxdt<+\infty$ (where $\rho_t:=e_t\#Q$). Note that because both $\bar{Q}$ and Q are such that the measures $\bar\rho_t$ and ρ_t are absolutely continuous for every t, modifying W into another function which is equal to W a.e. does not change the validity of this inequality. On the other hand, what we would like would be to prove that $\bar{Q}$ is actually concentrated on curves which minimize J_W starting from their starting point. This condition, instead, could depend on the choice of the representative of W since curves are negligible.

We prove now an integral characterization of this condition.

Lemma 3.2. *Given a measurable function $W:[0,T]\times\Omega\to\mathbb{R}$, a measure $\bar{Q}\in\mathscr{P}_{m_0}(\mathscr{C})$ is concentrated on the curves γ such that the following condition holds*

$$\forall\omega\in\mathscr{C}\text{ s.t. }\gamma(0)=\omega(0),\ J_W(\gamma)\leq J_W(\omega)$$

if and only if it satisfies

$$\int_{\mathscr{C}} J_W d\bar{Q}\leq\int_{\mathscr{C}} J_W dQ\quad\text{for every }Q\in\mathscr{P}_{m_0}(\mathscr{C}). \tag{15}$$

Proof. Let us define a value function φ through $\varphi(x_0):=\inf_{\omega:\,\omega(0)=x_0}J_W(\omega)$.

First, we suppose that $\bar{Q}$ satisfies (15). Then, for every $\varepsilon>0$ we can choose for every x_0 a curve ω_{x_0} such that $\omega_{x_0}(0)=x_0$ and $J_W(\omega_{x_0})\leq\varphi(x_0)+\varepsilon$. Choosing the map $x_0\mapsto\omega_{x_0}$ in a measurable way (this is possible by measurable selection arguments, for which we refer to the seminal book [11]) and taking the image of m_0 through such a map we obtain a measure $\tilde{Q}\in\mathscr{P}_{m_0}(\mathscr{C})$ such that

$$\int_{\mathscr{C}} J_W d\tilde{Q}\leq\int\varphi(\gamma(0))d\tilde{Q}(\gamma)+\varepsilon=\int_\Omega\varphi dm_0+\varepsilon.$$

We then have

$$\int_{\mathscr{C}} J_W d\bar{Q} \leq \int_{\mathscr{C}} J_W d\tilde{Q} \leq \int_{\Omega} \varphi + \varepsilon.$$

The number ε being arbitrarily small, we also obtain $\int_{\mathscr{C}} J_W d\bar{Q} \leq \int_{\Omega} \varphi$. Yet, we have $J_W(\gamma) \geq \varphi(\gamma(0))$, which shows that this inequality is an equality. Then, we can see that $\bar{Q}$ is concentrated on curves satisfying $J_{\bar{Q}}(\gamma) = \varphi(\gamma(0))$, i.e. on optimal curves, which proves the claim.

For the converse implication, suppose that $\bar{Q}$ is concentrated on optimal curves. Then we have, for $\bar{Q}-$a.e. γ, $J_W(\gamma) = \varphi(\gamma(0))$ and $\int_{\mathscr{C}} J_W d\bar{Q} = \int_{\mathscr{C}} \varphi(\gamma(0)) d\bar{Q} = \int_{\Omega} \varphi dm_0$. On the other hand, for any other measure Q we have $\int_{\mathscr{C}} J_W dQ \geq \int_{\mathscr{C}} \varphi(\gamma(0)) dQ = \int_{\Omega} \varphi dm_0$, which proves (15). □

The goal becomes now to choose a function W which is a precise representative of the function $V + f'(\bar{\rho})$ and for which we have the validity of (15). The main difficulty is to prove that this inequality is valid for every Q and not only for those satisfying the integrability condition $\int_0^T \int_{\Omega} f(\rho(t,x)) dx dt < +\infty$.

We will do this in the two cases that we are considering, i.e. the jump case and the kinetic case. We start from the jump case where, under suitable conditions, it is possible to prove, thanks to regularity results contained in [13], that $\bar{\rho}$ is continuous. In this case there is no ambiguity, and we can take $W = V + f'(\bar{\rho})$ just choosing such a continuous representative.

Before analyzing the two cases and proving that optimizers are indeed equilibria in the sense that we explained, we want to explain why this section is called "equivalence between equilibria and optimizers" while we only prove that optimizers are equilibria. Indeed, the converse implication is straightforward because of the convexity of the functional $\mathscr{U}$. Proposition 3.1 provides necessary optimality conditions for the minimization of $\mathscr{U}$, consisting in inequality (14) which has to be satisfied for those Q for which (13) holds. Actually, because of convexity, this necessary optimality condition is also sufficient, and every equilibrium satisfies it (and even more, because it satisfies (14) even without (13)).

3.1. **The jump case.**

Let us consider the case $A(\gamma) = S(\gamma)$ defined in Section 2.1. It is important to have in mind the results of [13], where we proved that the solution $\bar{\rho}$ of the problem (12) is Lipschitz in time and that according to the regularity of V and the conditions at time 0 and T, $\bar{\rho}$ can be either continuous or bounded. Because of these results, we will provide statements under the additional assumption that $\bar{\rho}$ is either continuous or bounded. In order to illustrate the strategy, we start from the case where it is continuous.

Theorem 3.3. *Suppose that Ω is either the d-dimensional torus or a compact Lipschitz domain in $\mathbb{R}^d$. Let $\bar{Q}$ be such that* (14) *holds for all Q satisfying* (13). *Define $\bar{\rho}_t := e_t \# \bar{Q}$, suppose that the measures $\bar{\rho}_t$ are all absolutely continuous, and call $\bar{\rho}(t,x)$ their densities. Suppose that the function $\bar{\rho} : [0,T] \times \Omega \to \mathbb{R}$ is continuous (and hence bounded). Then we have*

$$\begin{aligned} (16) \qquad \forall Q \in \Gamma, \int_{\mathscr{C}} S(\gamma) d\bar{Q}(\gamma) + \int_0^T \int_{\Omega} \left(V(t,x) + f'(\bar{\rho}(t,x)) \right) \bar{\rho}(t,x) dx dt + \int_{\mathscr{C}} \psi(\gamma(T)) d\bar{Q}(\gamma) \\ \leq \int_{\mathscr{C}} S(\gamma) dQ(\gamma) + \int_0^T \int_{\Omega} \left(V(t,x) + f'(\bar{\rho}(t,x)) \right) \rho(t,x) dx dt + \int_{\mathscr{C}} \psi(\gamma(T)) dQ(\gamma). \end{aligned}$$

We observe that, in this statement, the continuity of $\bar{\rho}$ in both variables is not really needed, and continuity in x for a.e. t together with a unifom L^∞ bound would be enough.

Proof. Let us start from the case where Ω is the torus. Let Q be in Γ. Choose $r > 0$. For all $y \in B_r = B(0,r) \subset \mathbb{R}^d$, we define the function

$$T_y \colon \mathscr{C} \to \mathscr{C},\ \gamma \mapsto \left(t \mapsto \begin{cases} \gamma(t), & \text{if } t \leq t_0(\gamma), \\ \gamma(t) + y, & \text{if } t > t_0(\gamma), \end{cases} \right)$$

where $t_0(\gamma)$ is the time at which the curve γ has its first jump or discontinuity (this notion is well-defined for γ such that $A(\gamma) < +\infty$); otherwise one can define $t_0(\gamma) = T$, or not define $t_0(\gamma)$ and$T_y(\gamma)$ at all). This function is a translation of a vector y of the curves in Ω starting from the time the curve jumps (note that this very construction requires to use the torus, otherwise $\gamma(t) + y$ could be outside the domain; the general case of other domains will be considered later). This construction preserves left-continuity and leaves the curves unchanged at time 0, i.e

$$\forall \gamma \in \mathscr{C},\ T_y(\gamma)(0) = \gamma(0).$$

In addition, by the construction of T_y, the function T_y does not add new jumps to the curves, so the following inequality holds:

$$(17) \qquad S(T_y(\gamma)) \leq S(\gamma).$$

For fixed $r>0$, we define the measure Q_r such that

$$Q_r = \fint_{B_r} T_y \# Q dy,$$

so the inequality (17) implies

$$\int_{\mathscr{C}} S(\gamma) dQ_r(\gamma) \le \int_{\mathscr{C}} S(\gamma) dQ(\gamma) \le C. \tag{18}$$

Next, we have for all $\varphi \in C_b(\Omega)$,

$$\int_\Omega \varphi(x) d(e_0\#Q_r)(x) = \int_{\mathscr{C}} \fint_{B_r} \varphi(\gamma(0)) dy\, dQ(\gamma) = \int_\Omega \varphi(x) dm_0(x),$$

so $e_0\#Q_r = m_0$, which means that the initial condition is verified.

For each t, let us define the subsets

$$\mathscr{A}(t) = \{\gamma \in \mathscr{C};\ t \le t_0(\gamma)\} \ \text{ and } \ \mathscr{B}(t) = \{\gamma \in \mathscr{C};\ t > t_0(\gamma)\}. \tag{19}$$

Let t be in $[0,T]$. For all positive functions $\varphi \in C_b(\Omega)$, we have

$$\begin{aligned}
\int_\Omega \varphi(x) d(e_t\#Q_r)(x) &= \int_{\mathscr{C}} \fint_{B_r} \varphi(\gamma(t)) \mathbb{1}_{\mathscr{A}(t)}(\gamma) dy\, dQ(\gamma) + \int_{\mathscr{C}} \fint_{B_r} \varphi(\gamma(t)+y) \mathbb{1}_{\mathscr{B}(t)}(\gamma) dy\, dQ(\gamma) \\
&= \int_{\mathscr{A}(t)} \varphi(\gamma(0)) dQ(\gamma) + \int_{\mathscr{C}} \fint_{B_r} \varphi(\gamma(t)+y) \mathbb{1}_{\mathscr{B}(t)}(\gamma) dy\, dQ(\gamma) \\
&\le \int_\Omega \varphi(x) dm_0(x) + \int_{\mathscr{C}} \fint_{B_r} \varphi(\gamma(t)+y) dy\, dQ(\gamma) \\
&= \int_\Omega \varphi(x) dm_0(x) + \int_\Omega \int_{\mathbb{R}^d} \frac{\varphi(x+y)}{|B_r|} \mathbb{1}_{B_r}(-y) dy\, d(e_t\#Q)(x) \\
&\underset{u=x+y}{=} \int_\Omega \varphi(x) dm_0(x) + \int_\Omega \int_{\mathbb{R}^d} \frac{\varphi(u)}{|B_r|} \mathbb{1}_{B_r}(x-u) du\, \rho(t,x) dx \\
&= \int_\Omega \varphi(x) dm_0(x) + \int_{\mathbb{R}^d} \varphi(u) \int_\Omega \frac{\mathbb{1}_{B_r}(u-x)}{|B_r|} \rho(t,x) dx du \\
&= \int_\Omega \varphi(x) dm_0(x) + \int_{\mathbb{R}^d} \varphi(u) \left(\frac{\mathbb{1}_{B_r}}{|B_r|} * \rho_t \right)(u) du.
\end{aligned} \tag{20}$$

Therefore, for almost every $x \in \Omega$, we have

$$\rho_r(t,x) := (e_t\#Q_r)(x) \le m_0(x) + \frac{1}{|B_r|}.$$

The right-hand side is independent of t and of the form $m_0 + C$. By assumption, m_0 is bounded, and so is ρ_r. Hence, we obtain

$$\int_0^T \int_\Omega f(\rho_r(t,x)) dx dt < \infty. \tag{21}$$

Therefore, we can apply Proposition 3.1 so as to write

$$\begin{aligned}
\forall r>0, \int_{\mathscr{C}} S(\gamma) d\bar{Q}(\gamma) + \int_0^T \int_\Omega \big(V(t,x) + f'(\bar\rho(t,x))\big) \bar\rho(t,x) dx dt + \int_{\mathscr{C}} \psi(\gamma(T)) d\bar{Q}(\gamma) \\
\le \int_{\mathscr{C}} S(\gamma) dQ_r(\gamma) + \int_0^T \int_\Omega \big(V(t,x) + f'(\bar\rho(t,x))\big) \rho_r(t,x) dx dt + \int_{\mathscr{C}} \psi(\gamma(T)) dQ_r(\gamma).
\end{aligned} \tag{22}$$

We now verify that we have $\rho_r \overset{*}{\rightharpoonup} \rho$ by using again the sets $\mathscr{A}(t)$ and $\mathscr{B}(t)$, so that we will be able to pass to the limit in (22). We fix a time $t \in [0,T]$: for all test function ϕ which are Lipschitz continuous on Ω, we have

$$\begin{aligned}
\int_\Omega \phi d(e_t\#Q_r) &= \fint_{B_r} dy \int_{\mathscr{C}} \phi(\gamma(t)) d(T_y\#Q)(\gamma) \\
&= \int_{\mathscr{A}(t)} \phi(\gamma(t)) dQ(\gamma) + \int_{\mathscr{B}(t)} \fint_{B_r} \phi(\gamma(t)+y) dy dQ(\gamma).
\end{aligned}$$

Since ϕ is Lipschitz continuous and Q is a probability measure, we have

$$\left|\int_{\mathscr{B}(t)} \fint_{B_r} \phi(\gamma(t)+y)dydQ(\gamma) - \int_{\mathscr{B}(t)} \phi(\gamma(t))dQ(\gamma)\right| \leq \mathrm{Lip}(\phi)r \to 0.$$

This shows that we have $\int_\Omega \phi d(e_t\#Q_r) \to \int_\Omega \phi d(e_t\#Q)$, i.e. $\rho_r(t) \overset{*}{\rightharpoonup} \rho(t)$ for every t. Then, using the continuity of V, ψ and $\bar{\rho}$, we finally obtain (using dominated convergence in time)

$$\begin{aligned} &\lim_{r\to 0} \int_0^T \int_\Omega \left(V(t,x)+f'(\bar{\rho}(t,x))\right)\rho_r(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))dQ_r(\gamma) \\ &= \int_0^T \int_\Omega \left(V(t,x)+f'(\bar{\rho}(t,x))\right)\rho(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma). \end{aligned} \tag{23}$$

Combining the inequality (18) and the limit (23), we have the desired inequality:

$$\begin{aligned} &\int_{\mathscr{C}} S(\gamma)d\bar{Q}(\gamma) + \int_0^T \int_\Omega \left(V(t,x)+f'(\bar{\rho}(t,x))\right)\bar{\rho}(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))d\bar{Q}(\gamma) \\ &\leq \limsup_{r\to 0} \int_{\mathscr{C}} S(\gamma)dQ_r(\gamma) + \int_0^T \int_\Omega \left(V(t,x)+f'(\bar{\rho}(t,x))\right)\rho_r(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))dQ_r(\gamma) \\ &\leq \int_{\mathscr{C}} S(\gamma)dQ(\gamma) + \int_0^T \int_\Omega \left(V(t,x)+f'(\bar{\rho}(t,x))\right)\rho(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma), \end{aligned}$$

which ends the proof in the case of the torus.

Let us consider now a Lipschitz domain Ω. For $r_0 > 0$ small enough, there exists a map $\Pi : \Omega_{r_0} \to \Omega$ (where $\Omega_r := \{x \in \mathbb{R}^d : d(x,\Omega) \leq r\}$) with the following properties:

- $\Pi(x) = x$ for all $x \in \Omega$;
- there exist a constant C_1 such that $|\Pi(x)-x| \leq C_1 r$ for all $x \in \Omega_r$:
- there exist a constant C_2 such that for any measure on Ω_{r_0} with density bounded by a constant M, its image through Π is a measure on Ω with density bounded by C_2M.

Note that we do not need Π to be continuous and that it can be defined using charts where Ω is bounded by a graph: if locally we have $\Omega = \{x = (x', x_d) \in \mathbb{R}^d : x_d > h(x')\}$ we can define $\Pi(x', x_d) = (x', 2h(x') - x_d)$ for points $x \notin \Omega$. The whole domain can be covered by a finite number of these charts, and for points in the intersection of different charts we can choose arbitrarily which definition of Π to choose. The constant C_2 would depend on the maximal number of intersecting charts, in this case.

We replace now the map T_y used in the case of the torus with a map $\tilde{T}_y$ defined via $\tilde{T}_y(\gamma)(t) = \Pi(T_y(\gamma)(t))$. We thus obtain a family of measures $\tilde{Q}_r$. Note that it is still true that we have $S(\tilde{T}_y(\gamma)) \leq S(\gamma)$. We then go on with the same procedure as before, defining $\tilde{\rho}_r(t) := e_t\#\tilde{Q}_r = \Pi_\#(\rho_r(t))$. The densities of the measures ρ_r were bounded by $||m_0||_{L^\infty} + \frac{1}{|B_r|}$, thus the densities of $\tilde{\rho}_r$ will be bounded by another constant, but the argument does not change. Moreover, we will again have $\tilde{\rho}_r(t) \overset{*}{\rightharpoonup} \rho(t)$. This is due to the fact that $\rho_r(t)$ is supported in Ω_r for every t, and hence for every Lipschitz test function ϕ we have $\left|\int \phi d\tilde{\rho}_r - \int \phi d\rho_r\right| \leq C\mathrm{Lip}\phi\, r \to 0$. This allows to conclude the argument in the very same way as in the case of the torus. □

We now consider a variant, where we only suppose $\bar{\rho}$ to be bounded instead of continuous. First of all, we consider the map Π defined on the set Ω_{r_0} at the end of the proof of the previous theorem and we define, for $r < r_0$ the following quantities

$$\forall x \in \Omega,\ \left[f'(\bar{\rho})\right]_r(t,x) := \fint_{B_r} f'(\bar{\rho}(t,\Pi(x+y)))dy, \quad \widehat{f'(\bar{\rho})}(t,x) := \limsup_{r\to 0} \left[f'(\bar{\rho})\right]_r(t,x). \tag{24}$$

Of course the map Π is useless in the case of the torus and could be omitted for any internal point x in the definition of $\widehat{f'(\bar{\rho})}(t,x)$ and in the definition of $[f'(\bar{\rho})]_r(t,x)$ as soon as $r < d(x,\partial\Omega)$. Also note that we have $\widehat{f'(\bar{\rho})} = f'(\bar{\rho})$ if $\bar{\rho}$ is a continuous function and that in any case the equality $\widehat{f'(\bar{\rho})} = f'(\bar{\rho})$ holds a.e. (more precisely, at any interior point x which is a Lebesgue point of $f'(\bar{\rho})$; this only excludes a negligible set of points, and the boundary, which is also negligible).

Theorem 3.4. *Suppose that Ω is either the d-dimensional torus or a compact Lipschitz domain in $\mathbb{R}^d$. Let $\bar{Q}$ be such that* (14) *holds for all Q satisfying* (13). *Suppose that the densities of all the measures $e_t\#\bar{Q}$, denoted by $\bar{\rho}(t,\cdot)$, are* *bounded by a common constant. Then for all $Q \in \Gamma$, we have*

$$\int_{\mathscr{C}} S(\gamma)d\bar{Q}(\gamma) + \int_0^T \int_\Omega \left(V(t,x) + \widehat{f'(\bar{\rho})}(t,x)\right)\bar{\rho}(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))d\bar{Q}(\gamma) \tag{25}$$

$$\leq \int_{\mathscr{C}} S(\gamma)dQ(\gamma) + \int_0^T \int_\Omega \left(V(t,x) + \widehat{f'(\bar{\rho})}(t,x)\right)\rho(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma). \tag{26}$$

Proof. The proof is similar to Theorem 3.3. The main difference is that the function $\widehat{f'(\bar{\rho})}$ is no longer continuous in space, but only bounded. We directly treat the case where we use the map Π, as it is now quite standard. We will omit, by the way, the symbol $\tilde{}$.

We start from the same construction: for all $r > 0$, we define as in Theorem 3.3 the measures

$$Q_r = \fint_{B_r} T_y \# Q dy.$$

As a consequence, $e_0 \# Q_r = m_0$ and $\int_{\mathscr{C}} S(\gamma)dQ_r(\gamma) \leq \int_{\mathscr{C}} S(\gamma)dQ(\gamma) \leq C$. We have seen that the densities ρ_r are bounded by a constant (by $||m_0|| + \frac{1}{|B_r|}$ in the case of the torus – note that in this case as well our assumption implies that m_0 is bounded – or by a multiple of this number in case we need to use Π). This yields

$$\int_0^T \int_\Omega f(\rho_r(t,x))dxdt < \infty.$$

By applying Proposition 3.1, we obtain

$$\begin{aligned}\forall r > 0, \int_{\mathscr{C}} S(\gamma)d\bar{Q}(\gamma) + \int_0^T \int_\Omega \left(V(t,x) + f'(\bar{\rho}(t,x))\right)\bar{\rho}(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))d\bar{Q}(\gamma) \\ \leq \int_{\mathscr{C}} S(\gamma)dQ_r(\gamma) + \int_0^T \int_\Omega \left(V(t,x) + f'(\bar{\rho}(t,x))\right)\rho_r(t,x)dxdt + \int_{\mathscr{C}} \psi(\gamma(T))dQ_r(\gamma),\end{aligned}$$

and we need to pass to the limit in the right-hand side. We use again the sets $\mathscr{A}(t)$ and $\mathscr{B}(t)$ defined in (19). The only term which has to be treated differently can be dealt with via the definition of $\widehat{f'(\bar{\rho})}$. Indeed, we have

$$\begin{aligned}\int_0^T \int_\Omega f'(\bar{\rho}(t,x))\rho_r(t,x)dxdt &= \int_0^T \int_{\mathscr{A}(t)} f'(\bar{\rho}(t,\gamma(t)))dQ(\gamma)dt + \int_0^T \int_{\mathscr{B}(t)} \fint_{B_r} f'(\bar{\rho}(t,\Pi(\gamma(t)+y)))dydQ(\gamma)dt \\ &= \int_0^T \int_{\mathscr{A}(t)} f'(\bar{\rho}(t,\gamma(t)))dQ(\gamma)dt + \int_0^T \int_{\mathscr{B}(t)} \left[f'(\bar{\rho})\right]_r(t,\gamma(t))dQ(\gamma)dt.\end{aligned}$$

The first term on the right-hand side does not depend on r. Moreover we note that we have

$$\int_0^T \int_{\mathscr{A}(t)} f'(\bar{\rho}(t,\gamma(t)))dQ(\gamma)dt = \int_0^T \int_{\mathscr{A}(t)} \widehat{f'(\bar{\rho})}(t,\gamma(t))dQ(\gamma)dt.$$

This is due to the a.e. equality $\widehat{f'(\bar{\rho})} = f'(\bar{\rho})$ and to the fact that $e_t \# (Q \mathbb{1}_{\mathscr{A}(t)})$ is absolutely continuous (it is indeed a measure bounded from above by m_0). For the second term we can use a reverse Fatou's Lemma together with $\limsup_{r\to 0} [f'(\bar{\rho})]_r = \widehat{f'(\bar{\rho})}$. For this, we need to upper bound all these functions, but the assumption guarantees that they are all bounded by $f'(M)$, where $M := \sup \bar{\rho}$. We then obtain

$$\limsup_{r\to 0} \int_0^T \int_{\mathscr{B}(t)} \left[f'(\bar{\rho})\right]_r(t,\gamma(t))dQ(\gamma)dt \leq \int_0^T \int_{\mathscr{B}(t)} \widehat{f'(\bar{\rho})}(t,\gamma(t))dQ(\gamma)dt$$

and hence

$$\limsup_{r\to 0} \int_0^T \int_\Omega f'(\bar{\rho}(t,x))\rho_r(t,x)dxdt \leq \int_0^T \int_\Omega \widehat{f'(\bar{\rho})}(t,x)\rho(t,x)dxdt.$$

The other terms (including the integral of the potential V, which is now separated from $f'(\bar\rho)$) are treated exactly as in Theorem 3.3. By combining the different terms we obtain

$$\begin{aligned}
&\int_{\mathscr{C}} S(\gamma)d\bar{Q}(\gamma)+\int_0^T\int_\Omega \left(V(t,x)+f'(\bar\rho(t,x))\right)\bar\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))d\bar{Q}(\gamma)\\
&\le \limsup_{r\to0}\int_{\mathscr{C}} S(\gamma)dQ_r(\gamma)+\int_0^T\int_\Omega \left(V(t,x)+f'(\bar\rho(t,x))\right)\rho_r(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ_r(\gamma)\\
&\le \int_{\mathscr{C}} S(\gamma)dQ(\gamma)+\int_0^T\int_\Omega \left(V(t,x)+\widehat{f'(\bar\rho)}(t,x)\right)\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ(\gamma).
\end{aligned}$$

We then conclude by replacing $f'(\bar\rho)$ with $\widehat{f'(\bar\rho)}$ in the left-hand side, since these two functions agree a.e. and for a.e. t the measure $\bar\rho = e_t\#\bar{Q}$ is absolutely continuous. □

We can then conclude about the equilibrium properties of the measure $\bar{Q}$ that we obtained as a fixed point in the previous section.

Corollary 3.5. *Suppose that the measures $e_t\#\bar{Q}=\bar\rho(t,\cdot)$ are all bounded by a same constant. Then, the measure $\bar{Q}$ is concentrated on the curves $\bar\gamma$ which are minimizers, for fixed starting point, of the action function J_W with $W = V+\widehat{f'(\bar\rho)}$, i.e. on curves $\bar\gamma$ such that*

$$\begin{aligned}
\forall\omega\in\mathscr{C}\text{ with }\omega(0)=\bar\gamma(0)\quad & S(\bar\gamma)+\int_0^T\left[V(t,\bar\gamma(t))+\widehat{f'(\bar\rho)}(t,\bar\gamma(t))\right]dt+\psi(\bar\gamma(T))\\
&\le S(\omega)+\int_0^T\left[V(t,\omega(t))+\widehat{f'(\bar\rho)}(t,\omega(t))\right]dt+\psi(\omega(T)).
\end{aligned}$$

The proof is a combination of the previous results. the case where $\bar\rho$ is continuous is included in this same statement, as in this case we simply have $\widehat{f'(\bar\rho)}=f'(\bar\rho)$.

3.2. **The kinetic case.** In this part, we consider the case $A(\gamma)=K(\gamma)$ (see (9)). The goal is to perform a similar computation as in the case $A=S$, i.e. showing by a convolution argument that the inequality (16) actually holds for all $Q\in\Gamma$.

The regularity theory for the kinetic case is a classical issue in MFG theory, and one of the strongest results is the one contained in [23], which proves L^∞ bounds on $\bar\rho$. Continuity for the density of $\bar\rho$ is a much harder task, so we will only rely on its boundedness. We refer to [23] for the proofs of this boundedness, and to [29] for an overview of the applications of such a result to the equivalence between optimizers and equilibria. Indeed, a general strategy first introduced by [1] for fluid mechanics applications allowed to prove the optimality of $\bar{Q}$-almost every curve within a restricted class of curves on which the maximal function of $f'(\bar\rho)$ is integrable in time. This has later been used in [10] in the framework of MFG, following essentially the same ideas. As explained in [29], the upper bounds on $\bar\rho$ allow to state a much simpler result, showing optimality in the class of all $W^{1,p}$ curves. The proof of [1] and [10] is based on a multiple approximation procedure strongly relying on the separability of the space $W^{1,p}$ and on the continuity of A for the strong $W^{1,p}$ convergence. Such a procedure was not possible for the jump case, which is naturally set in the space BV which is non-separable. This motivated the different approach that we presented in the previous subsection. A variant of such approach allows to obtain the following result in the kinetic case. Such a result recovers already known results, and our approach requires a strong assumption on the growth of the function f in terms of the exponent p in the kinetic energy and on the space dimension d. Yet, it is interesting to observe that such a numeric assumption is not new in the setting of deterministic (first-order) MFG (see, for instance, [9]). Moreover, such a proof is considerably simpler than the one in [1, 10].

Compared to the theorems in the jump case, we need to require extra regularity on the domain Ω. We need indeed an extra property for the map Π: we need to require Π to be $(1+Cr)$-Lipschitz continuous on Ω_r. The reason for this is that we need now to preserve Sobolev bounds on the composition of a curve with Π. It is possible to guarantee this by requiring Ω to be C^2 and defining $\Pi(x)=2P_\Omega(x)-x$, where P_Ω is the projection onto Ω, well-defined for smooth domains in a neighborhood of Ω. This map is a sort of reflection across the boundary. The Jacobian matrix of this map Π is symmetric and its eigenvalues are very close to -1 in the direction of $x-P_\Omega(x)$ and to 1 in the orthogonal directions. The precise bounds depend on r and on the maximal curvature of $\partial\Omega$. We leave the precise computations to the reader, who can also look at [12, Chapter 3].

Theorem 3.6. *Suppose that Ω is either the torus or a C^2 compact domain in $\mathbb{R}^d$. Let $\bar{Q}$ be such that* (14) *holds for all Q satisfying* (13). *Suppose that the densities of all the measures $e_t\#\bar{Q}$, denoted by $\bar\rho(t,\cdot)$, are bounded by a common* *constant. Let $f\colon \mathbb{R}\to\mathbb{R}$ be a C^1 convex function such that $f(u)\le C(u^q+1)$. Let p, q and d verify $q<1+\frac{p'}{d}$ where p' is the conjugate exponent of p. Let $\widehat{f'(\bar\rho)}$ be defined as in* (24)*. Then for all $Q\in\Gamma$, we have*

$$\begin{aligned}\int_{\mathscr{C}} K(\gamma)d\bar Q(\gamma)+\int_0^T\int_\Omega\Big(V(t,x)+\widehat{f'(\bar\rho)}(t,x)\Big)\bar\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))d\bar Q(\gamma)\\ \le\int_{\mathscr{C}} K(\gamma)dQ(\gamma)+\int_0^T\int_\Omega\Big(V(t,x)+\widehat{f'(\bar\rho)}(t,x)\Big)\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ(\gamma).\end{aligned}\tag{27}$$

Proof. Let $Q\in\Gamma$ and $\rho(t,\cdot):=e_t\#Q$. Let $r>0$ and $\eta(t)=t^\alpha$. For all $y\in B(0,1)\subset\mathbb{R}^d$, we define the function

$$\begin{aligned}T_y\colon W^{1,p}([0,T]) &\to W^{1,p}([0,T]),\\ \gamma &\mapsto (t\mapsto \Pi(\gamma(t)+r\eta(t)y)).\end{aligned}$$

Then, we have $\|T_y(\gamma)'\|_{L^p([0,T])}\le(1+Cr)(\|\dot\gamma\|_{L^p([0,T])}+\|r\dot\eta y\|_{L^p([0,T])})$ (where the multiplying factor $1+Cr$ comes from the composition with Π), so a necessary condition for $T_y(\gamma)'$ to be in $L^p([0,T])$ is that $\dot\eta$ is in $L^p([0,T])$ and this is true if $p(\alpha-1)>-1$, i.e $\alpha>1-1/p$.

Next, we define similarly to Theorem 3.3 the sequence of measures $(Q_r)_r$:

$$Q_r=\fint_{B(0,1)}T_y\#Qdy,$$

and we denote by $\rho_r(t,\cdot)=e_t\#Q_r$.

We now perform similar computations as those of (20). We have

$$e_t\#Q_r=\Pi\#\left((e_t\#Q)*\frac{\mathbb{1}_{B_{r\eta(t)}}}{|B_{r\eta(t)}|}\right).$$

Computing the volume of the ball in $\mathbb{R}^d$ gives $|B_{r\eta(t)}|=c(r\eta(t))^d$ for some $c>0$. Then we have

$$\|e_t\#Q_r\|_{L^\infty(\Omega)}\le C_2\|e_t\#Q\|_{L^1(\Omega)}\left\|\frac{\mathbb{1}_{B_{r\eta(t)}}}{|B_{r\eta(t)}|}\right\|_{L^\infty(\Omega)}\le C(r\eta(t))^{-d}.$$

Since $\eta(t)\to0$ as $t\to0$, the L^∞ norm will not be uniformly bounded, but we are actually interested in the L^q norm. We recall that for every probability density u we have

$$||u||_{L^q}=\left(\int u^qdx\right)^{1/q}\le\left(||u||_{L^\infty}^{q-1}\int udx\right)^{1/q}\le||u||_{L^\infty}^{\frac{q-1}{q}}.$$

Thus,

$$\|e_t\#Q_r\|_{L^q(\Omega)}\le C(r\eta(t))^{\frac{d(1-q)}{q}}.$$

We want to impose the integrability condition $\int_0^T\int_\Omega f(\rho_r)dxdt<+\infty$, which, thanks to the assumption on f, becomes $\int_0^T||\rho_r(t)||_{L^q}^qdt<+\infty$. For this we need to require another condition on α, i.e. $\alpha<\frac{1}{d(q-1)}$.

Such an α exists, because of the assumption on p and q. Indeed, we assumed $q<1+\frac{p'}{d}$, which is equivalent to $\frac{1}{p'}:=1-\frac1p<\frac{1}{d(q-1)}$.

The condition (13) is now satisfied and we obtain by Proposition 3.1 the inequality

$$\begin{aligned}\int_{\mathscr{C}} K(\gamma)d\bar Q(\gamma)+\int_0^T\int_\Omega\big(V(t,x)+f'(\bar\rho(t,x))\big)\bar\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))d\bar Q(\gamma)\\ \le\int_{\mathscr{C}} K(\gamma)dQ_r(\gamma)+\int_0^T\int_\Omega\big(V(t,x)+[f'(\bar\rho)]_r(t,x)\big)\rho_r(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ_r(\gamma).\end{aligned}\tag{28}$$

We want to take the limit as $r\to0$ in the right-hand side. The last term can be treated just by noting that we have $Q_r\overset{*}{\rightharpoonup}Q$. The term in the middle can be treated by taking its limsup exactly as in the proof of Theorem 3.4. We are only left to consider the first term, which, in the proof of Theorem 3.4, was dealt with by noting that we had $S(T_y(\gamma))\le S(\gamma)$.

Here this inequality is no longer true, but almost. Indeed, we noted that we have the inequality $\|T_y(\gamma)'\|_{L^p([0,T])}\le(1+Cr)\|\dot\gamma\|_{L^p([0,T])}+\|r\dot\eta y\|_{L^p([0,T])}$. This means $K(T_y(\gamma))\le(1+Cr)^{1/p}((K(\gamma)^{1/p}+Cr)^p$. We then write

$$\int_{\mathscr{C}}K(\gamma)dQ_r(\gamma)=\fint_{B(0,1)}\int_{\mathscr{C}}K(T_y(\gamma))dQ(\gamma)\le(1+Cr)^{1/p}\fint_{B(0,1)}\int_{\mathscr{C}}(K(\gamma)^{1/p}+Cr)^pdQ(\gamma),$$

and we can take the limit by dominated convergence once we note that we have $(K(\gamma)^{1/p}+Cr)^p \le C(K(\gamma)+1)$ and K is supposed to be $Q-$integrable. We then obtain

$$\int_{\mathscr{C}} K(\gamma)d\bar{Q}(\gamma)+\int_0^T\int_\Omega \left(V(t,x)+f'(\bar{\rho}(t,x))\right)\bar{\rho}(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))d\bar{Q}(\gamma)$$
$$\le \int_{\mathscr{C}} K(\gamma)dQ(\gamma)+\int_0^T\int_\Omega \left(V(t,x)+\widehat{f'(\bar{\rho})}(t,x)\right)\rho(t,x)dxdt+\int_{\mathscr{C}}\psi(\gamma(T))dQ_r(\gamma)$$

and we conclude by replacing once more $f'(\bar{\rho})$ with $\widehat{f'(\bar{\rho})}$ in the left-hand side, since these functions agree a.e. □

Remark 3.7. *We note that in the above proof the condition on p,q and d is required to find a function $\eta \in W^{1,p}([0,T])$ with $\eta(0)=0$ and such that $\eta^{(1-q)d}$ is integrable. The difficulty is due to the fact that this integrability condition requires η not to tend to 0 too fast and the Sobolev regularity of η imposes to tend to 0 fast enough... If we could remove the condition $\eta(0)=0$ we could just use $\eta(t)=1$ which satisfies both conditions at the same time. The reason for imposing $\eta(0)=0$ lies in the need to preserve the initial condition of the curves, but actually we only need $e_0\#Q_r=m_0$. Hence, an example where the condition on p,q and d could be dropped is the case where Ω is the torus and m_0 is the uniform measure on it. In this case with $\eta(t)=1$ the measure $e_0\#Q_r$ would be the convolution of m_0 with the uniform measure on the ball, and thus it would also be equal to m_0. It would be interesting to investigate whether smoothness properties on m_0, which do not allow to say that the convolution equals m_0 but that it is close enough to it, would be enough to "correct" the error in the initial condition and obtain the same result under less stringent assumptions on p,q and d.*

Again, we can then conclude about the equilibrium properties of the measure $\bar{Q}$ that we obtained as a fixed point in the previous section.

Corollary 3.8. *Suppose that the measures $e_t\#\bar{Q}=\bar{\rho}(t,\cdot)$ are all bounded by a same constant. Then, the measure $\bar{Q}$ is concentrated on the curves $\bar{\gamma}$ which are minimizers, for fixed starting point, of the action function J_W with $W = V+\widehat{f'(\bar{\rho})}$, i.e. such that such*

$$\forall\omega\in\mathscr{C} \text{ with } \omega(0)=\bar{\gamma}(0) \quad K(\bar{\gamma})+\int_0^T\left[V(t,\bar{\gamma}(t))+\widehat{f'(\bar{\rho})}(t,\bar{\gamma}(t))\right]dt+\psi_T(\bar{\gamma}(T))$$
$$\le K(\omega)+\int_0^T\left[V(t,\omega(t))+\widehat{f'(\bar{\rho})}(t,\omega(t))\right]dt+\psi_T(\omega(T)).$$

4. EQUIVALENCE BETWEEN EULERIAN AND LAGRANGIAN FORMULATIONS

The goal of this section is to prove that, instead of considering a minimization problem among measures on curves (the *Lagrangian* viewpoint), it is possible to consider a minimization problem among curves of measures (the *Eulerian* viewpoint). In the kinetic case this is well-known and stems out of the following result from optimal transport theory.

Theorem 4.1. *Let $\mathscr{C}$ be the space of continuous curves valued into a domain Ω and K be the kinetic energy defined in* (9). *Take $Q\in\mathscr{P}(\mathscr{C})$ and define $\rho_t := e_t\#Q$. Then we have*

$$\int_0^T \frac{1}{p}|\rho'|_{W_p}(t)^p dt \le \int_{\mathscr{C}} K(\gamma)dQ(\gamma),$$

where $|\rho'|_{W_p}$ denotes the metric derivative of the curve ρ in the Wasserstein space W_p. In particular, we also have the existence of a time-dependent velocity field v such that $v_t\in L^p(\rho_t)$ for a.e. t and

$$\partial_t\rho+\nabla\cdot(\rho v)=0, \quad \int_0^T\int_\Omega \frac{1}{p}|v_t|^p d\rho_t dt \le \int_{\mathscr{C}} K(\gamma)dQ(\gamma).$$

Conversely, given a curve of measures $t\mapsto\rho_t\in\mathscr{P}(\Omega)$ which is absolutely continuous for the W_p distance, there exists a measure $Q\in\mathscr{P}(\mathscr{C})$ such that $\rho_t=e_t\#Q$ and

$$\int_0^T \frac{1}{p}|\rho'|_{W_p}(t)^p dt \ge \int_{\mathscr{C}} K(\gamma)dQ(\gamma).$$

The above result is classical in the theory of optimal transport, is essentially taken from [22], and can also be deduced combining the results presented in [28, Chapter 5].

One can informally understand why the Eulerian equivalent of $\int KdQ$ involves the power p of the metric derivative in W_p. Indeed, we can see $K(\gamma)$ as a limit (or a sup) of $\frac{1}{p}\sum_k(t_{k+1}-t_k)\left|\frac{|\gamma(t_{k+1})-\gamma(t_k)|}{t_{k+1}-t_k}\right|^p$ when the partitions $(t_k)_k$ become finer and finer. The minimal value of optimal transport problems with cost $(x,y)\mapsto|x-y|^p$ is the power p of the Wasserstein distance W_p. Hence, we obtain as an Eulerian equivalent of $\int KdQ$ a sum of the form $\sum_k \frac{1}{p}(t_{k+1}-t_k)\frac{W_p^p(\rho_{t_{k+1}},\rho_{t_k})}{|t_{k+1}-t_k|^p}$, which is itself a discretization of $\frac{1}{p}\int_0^T|\rho'|_{W_p}(t)^p dt$.

If we want to do a similar procedure for the jump case, i.e. with S instead of K, we first note that $S(\gamma)$ can be discretized as follows

$$S(\gamma)=\sup_{t_0<t_1<\cdots<t_N\subset[0,T]}\sum_k \mathbb{1}_{\gamma(t_{k+1})\neq\gamma(t_k)}, \tag{29}$$

where, of course, one has to choose the left-continuous representative of γ (otherwise we need to take an infimum among representatives). Then, we notice two points. The first point is the absence of the dependence in $(t_{k+1}-t_k)$, corresponding to the case $p=1$: as a consequence, we will not have a kinetic energy but a length, for a certain distance. The other is the fact that $(x,y)\mapsto\mathbb{1}_{x\neq y}$ is a distance (the so-called discrete distance) and that the associated Wasserstein distance is the total variation distance between measures, as explained in the following lemma, whose proof can be found, for instance, in [30].

Lemma 4.2. *For all μ and ν in $\mathscr{P}(\Omega)$,*

$$\inf_{\pi\in\Pi(\mu,\nu)}\int_{\Omega\times\Omega}\mathbb{1}_{x\neq y}d\pi(x,y)=\sup_{f:\sup f-\inf f\leq 1}\int_\Omega fd(\mu-\nu)=\sup_{f:|f|\leq 1/2}\int_\Omega fd(\mu-\nu)=\frac{1}{2}\|\mu-\nu\|_{\mathscr{M}}$$

where $\Pi(\mu,\nu)$ is the set of probability measures over $\Omega\times\Omega$ whose marginals are μ and ν and $||\cdot||_{\mathscr{M}}$ is the norm on the space of measures, i.e. the total variation of (signed) measures.

For these reasons, we will consider the length of a curve ρ in the space $\mathscr{P}(\Omega)$ computed according to the distance $||\cdot||_{\mathscr{M}}$. We will denote it by $\mathbf{L}$.

$$\mathbf{L}(\rho):=\sup_{0=t_0<\cdots<t_i<\cdots<t_N=T}\sum_{i=0}^{N-1}\|\rho(t_{i+1})-\rho(t_i)\|_{\mathscr{M}}.$$

Very informally we can write

$$\mathbf{L}(\rho)=\int_0^T\int_\Omega|\dot\rho(t,x)|dxdt$$

since, when all the measures ρ_t are absolutely continuous, the total variation distance coincides with the L^1 distance, and computing a length corresponds to computing the integral of the norm of the derivative.

The next proposition gives one inequality which is of interest for the equivalence of a Lagrangian and of an Eulerian problem.

Proposition 4.3. *For all measure $Q\in\mathscr{P}(\mathscr{C})$, setting for all $t\in[0,T]$, $e_t\#Q=\rho_t$, we have*

$$\mathbf{L}(\rho)\leq 2\int_{\mathscr{C}}S(\gamma)dQ(\gamma).$$

Proof. Let $t_0<\cdots<t_i<\cdots<t_N$ be a subdivision of $[0,T]$. For all $i\in\{0,\cdots,N-1\}$, if φ_i is a function such that $\|\varphi_i\|_\infty\leq 1$, we have

$$\begin{aligned}\sum_{i=0}^{N-1}\int_\Omega\varphi_i d(\rho(t_{i+1})-\rho(t_i))&=\sum_{i=0}^{N-1}\int_{\mathscr{C}}(\varphi_i(\gamma(t_{i+1}))-\varphi_i(\gamma(t_i)))dQ(\gamma)\\&\leq 2\int_{\mathscr{C}}\sum_{i=0}^{N-1}\mathbb{1}_{\gamma(t_i)\neq\gamma(t_{i+1})}dQ(\gamma)\leq 2\int_{\mathscr{C}}S(\gamma)dQ(\gamma).\end{aligned}$$

The term in the right-hand side of the inequality is independent of φ_i, so we can take the supremum over the functions bounded by 1. We obtain

$$\sum_{i=0}^{N-1}\|\rho(t_{i+1})-\rho(t_i)\|_{\mathscr{M}}\leq 2\int_{\mathscr{C}}S(\gamma)dQ(\gamma).$$

Next, by taking the supremum over the set of subdivisions of $[0,T]$, we obtain the desired result, by definition of length. □

In the following proposition, we construct a sequence of measures $(Q_N)_N$ which are transport plans for the optimal transport problem with $c(x,y)=\mathbb{1}_{x\neq y}$:

Lemma 4.4. *Given $N+1$ measures $\rho_0,\dots,\rho_N \in \mathscr{P}(\Omega)$ there exists a measure $\pi^N \in \mathscr{P}(\Omega^{N+1})$ which has marginals $\rho_0,\dots,\rho_N \in \mathscr{P}(\Omega)$ and such that*

$$\int_{\Omega^{N+1}} \sum_{k=0}^{N-1} \mathbb{1}_{x_k \neq x_{k+1}} d\pi^N(x_0,\dots,x_N) = \frac{1}{2}\sum_{k=0}^{N-1} \|\rho_k - \rho_{k+1}\|_{\mathscr{M}}. \tag{30}$$

Proof. By Lemma 4.2, we know that for each k we have

$$\inf_{\pi \in \Pi(\rho_k,\rho_{k+1})} \int_{\Omega\times\Omega} \mathbb{1}_{x\neq y} d\pi(x,y) = \frac{1}{2}\|\rho_k - \rho_{k+1}\|_{\mathscr{M}}.$$

Since the cost $(x,y)\mapsto \mathbb{1}_{x\neq y}$ is lower semi-continuous, we can choose for each k a measure $\pi_k \in \mathscr{P}(\Omega\times\Omega)$ attaining such an infimum. The second marginal of π_k equals the first marginal of π_{k+1}. Hence it is possible, applying several times the well-known gluing lemma about the composition of transport plans (see [28, Lemma 5.5]), to build a measure $\pi^N \in \mathscr{P}(\Omega^{N+1})$ such that, for each $k \leq N-1$, the projection of this measure onto the k-th and $(k+1)$-th coordinate (i.e. its image through the map $(x_0,\dots,x_N)\mapsto(x_k,x_{k+1})$) equals π_k. Such a measure satisfies

$$\int_{\Omega^{N+1}} \mathbb{1}_{x_k \neq x_{k+1}} d\pi^N(x_0,\dots,x_N) = \frac{1}{2}\|\rho_k - \rho_{k+1}\|_{\mathscr{M}}$$

and, summing over k, we obtain the claim. □

Lemma 4.5. *Let $[-1,T] \ni t \mapsto \rho(t)$ be a left-continuous curve valued into $\mathscr{P}(\Omega)$, constant on $[-1,0]$, and such that $\rho(0) = m_0$. Let E be a finite subset of $(-1,T)\setminus\{0\}$. Then, there exists a measure $Q \in \mathscr{P}_{m_0}(\mathscr{C})$ such that*

$$\int_{\mathscr{C}} S dQ \leq \frac{1}{2}\mathbf{L}(\rho) \quad \text{ and } e_t\#Q = \rho_t \text{ for all } t \in E.$$

Proof. We order the points in E and call them $t_1,t_2,\dots,t_{N-1}$. We add, if needed, $t_0 = -1$ and $t_N = T$. There is an index j such that $t_j \leq 0$ and $t_{j+1} > 0$. We take $\rho_i = \rho(t_i)$ and apply the previous lemma. Then, we consider a map $Z : \Omega^{N+1} \to \mathscr{C}$, where the points of Ω^{N+1} are denoted by $(x_0,x_1,\dots,x_N)$. We define Z in this way: $Z(x)(t) = x_i$ for every $t \in (t_{i-1},t_i]\cap(0,T]$ (which requires $i \geq j+1$); $Z(x)(t) = x_j$ for every $t \in (t_j,0]$; $Z(x)(t) = x_i$ for every $t \in (t_{i-1},t_i]$ for $i \leq j$ and $Z(x)(-1) = x_0$. Then, we take $Q = Z\#\pi$ and we see that it satisfies the condition of the claim, since $Z(x)(t_i) = x_i$ and $S(Z(x)) = \sum_{k=0}^{N-1}\mathbb{1}_{x_k\neq x_{k+1}}$. □

Proposition 4.6. *Let $[-1,T] \ni t \mapsto \rho(t)$ be a left-continuous curve valued into $\mathscr{P}(\Omega)$ and constant on $[-1,0]$, with $\rho(0) = m_0$. Then there exists a measure $Q \in \mathscr{P}_{m_0}(\mathscr{C})$ such that*

$$\int_{\mathscr{C}} S dQ \leq \frac{1}{2}\mathbf{L}(\rho) \quad \text{ and } e_t\#Q = \rho_t \text{ for all } t.$$

Proof. First, note that there is nothing to prove if $\mathbf{L}(\rho) = +\infty$. We then assume $\mathbf{L}(\rho) < +\infty$ and consider the measure $\bar\mu$ on $[-1,T]$ defined via $\bar\mu(I) = TV(\rho;I)$ for every open interval I, where the total variation is to be intended when endowing $\mathscr{P}(\Omega)$ with the distance induced by the $||\cdot||_{\mathscr{M}}$ norm. We have $\bar\mu([-1,T]) = \mathbf{L}(\rho) < +\infty$. Let us consider a dense countable set $E_\infty = \{t_k, k\in\mathbb{N}\}$ of points in $(-1,T)\setminus\{0\}$ which are not atoms of $\bar\mu$. For every N, we take $t_1,t_2,\dots,t_N$ and we reorder them, calling t_k^N the ordered sequence that we obtain, characterized by $t_i^N < t_{i+1}^N$ and $\{t_i^N\}_{i=1,\dots,N} = \{t_i\}_{i=1,\dots,N}$. We then apply Lemma 4.5 to $E = E_N = \{t_i^N\}_{i=1,\dots,N}$ and we obtain the existence of a curve $Q_N \in \mathscr{P}_{m_0}(\mathscr{C})$ such that

$$\int_{\mathscr{C}} S dQ \leq \frac{1}{2}\mathbf{L}(\rho) \quad \text{and } e_t\#Q = \rho_t \text{ for all } t = 0, t_1^N,\dots,t_N^N, T.$$

We can say, with our previous notation, $Q_N \in \Gamma(C)$, and such a set is compact for the narrow convergence, so that we can extract a narrowly converging subsequence $Q_{N_k} \overset{*}{\rightharpoonup} Q$. By construction, for each $t \in E$ and each N large enough (depending on t) we have $e_t\#Q_N = \rho(t)$. We want now to apply Lemma 2.4 in order to obtain $e_t\#Q_N \overset{*}{\rightharpoonup} e_t\#Q$, which would imply $e_t\#Q = \rho(t)$. This is not straightforward since Lemma 2.4 only guarantees this convergence up to subsequences (but this is not a problem, we could extract once more) and for a.e. t and we do not know if these t are concerned. Here comes the assumption $\bar\mu(\{t\}) = 0$. Recall the measures μ_N defined in the proof of Lemma 2.4: they were defined via $\mu_N(I) = \int TV(\gamma;I)dQ_N(\gamma)$ but we can say

$$\mu_N(I) = \sum_{k:t_k^N,t_{k+1}^N\in I} \int \mathbb{1}_{\gamma(t_k^N)\neq\gamma(t_{k+1}^N)} dQ_N(\gamma) = \sum_{k:t_k^N,t_{k+1}^N\in I} ||\rho(t_k^N) - \rho(t_{k+1}^N)||_{\mathscr{M}} \leq \bar\mu(I).$$

This implies that any narrow limit μ of a subsequence of the μ_N should satisfy $\mu \leq \bar\mu$. In particular, any t which is not an atom for $\bar\mu$ is not an atom for μ neither, and the convergence $e_t\#Q_N \overset{*}{\rightharpoonup} e_t\#Q$ holds along such a subsequence.

We then obtained the existence of a measure Q which, by semi-continuity of S, satisfies

$$\int_{\mathscr{C}} S(\gamma) \leq \frac{1}{2}\mathbf{L}(\rho) \quad \text{ and } e_t\#Q = \rho_t \text{ for all } t \in E,$$

where E is a dense set. We now want to extend this to any $t \in (-1,T]$ using left-continuity. We remind that we made the choice to use left-continuous curves and that we compute the evaluations e_t using such representatives. In particular, we have, when $s \to t^-$, the narrow convergence $e_s\#Q \overset{*}{\rightharpoonup} e_t\#Q$. This can be seen by testing against a test function $\phi \in C_b(\Omega)$:

$$\int \phi(\gamma(s))dQ(\gamma) \to \int \phi(\gamma(t))dQ(\gamma),$$

this convergence being justified by the dominated convergence $\phi(\gamma(s)) \to \phi(\gamma(t))$, which itself comes from the left-continuity of γ. We also have, again for $s \to t^-$, the convergence $\rho(s) \overset{*}{\rightharpoonup} \rho(t)$, which is due to the assumption that ρ is left-continuous.

As a consequence, approximating an arbitrary $t \in (-1,T]$ with $s < t$, $s \in E$, we obtain $e_t\#Q = \rho(t)$ and this concludes the proof. □

We then obtain the following equality.

Corollary 4.7. *For every curve $t \mapsto \rho(t)$ valued in $\mathscr{P}(\Omega)$ and left-continuous we have*

$$\frac{1}{2}\mathbf{L}(\rho) = \inf\left\{ \int_{\mathscr{C}} S(\gamma)dQ(\gamma) \ : \ Q \in \mathscr{P}(\mathscr{C}), e_t\#Q = \rho(t) \textit{ for all } t \right\}.$$

Proof. The left-hand side is bounded by the right-hand side using Proposition 4.3 and the converse inequality is deduced from Proposition 4.6. □

This allows to state the equivalence, both in the jump case and in the kinetic case, of a Lagrangian and an Eulerian problem.

Theorem 4.8. *Given $V : [0,T] \times \Omega \to \mathbb{R}$, an action A equal either to S or to K, and a function $I : \mathscr{P}(\Omega) \to [0+\infty]$, the Lagrangian problem*

$$\min\left\{ \mathscr{U}(Q) := \int_{\mathscr{C}} A(\gamma)dQ(\gamma) + \int_0^T I(e_t\#Q)dt + \int_{\mathscr{C}} \int_0^T V(t,\gamma(t))dt \ dQ(\gamma) + \int_{\mathscr{C}} \psi(\gamma(T))dQ(\gamma) \right\}$$

is equivalent to the following Eulerian problem

$$\min \mathbf{U}(\rho) := \mathbf{A}(\rho) + \int_0^T I(\rho(t))dt + \int_0^T \int_\Omega V(t,x)d\rho(t)dt + \int_\Omega \psi d\rho(T),$$

where the minimization is performed among left-continuous curves $t \mapsto \rho(t)$ with $\rho(0) = m_0$, and the action functional $\mathbf{A}$ is given by $\mathbf{A} = \frac{1}{2}\mathbf{L}$ if $A = S$ or $\mathbf{A}(\rho) = \int_0^T \frac{1}{p}|\rho'|_{W_p}(t)^p dt$ if $A = K$.

The equivalence means that the minimal values are the same and that from an optimal Q we can find an optimal ρ by taking $\rho(t) = e_t\#Q$ and from an optimal ρ we can find an optimal Q by taking the one provided by Proposition 4.6.

Proof. Take an optimal $\bar{Q}$ and construct $\bar{\rho}$ as described in the statement. Since all terms but the first one in the functional only depend on the measures $e_t\#Q$, then clearly we have, using Proposition 4.3,

$$\min \mathbf{U} \leq \mathbf{U}(\bar{\rho}) \leq \mathscr{U}(\bar{Q}) = \min \mathscr{U}. \tag{31}$$

Then, take an optimal $\bar{\rho}$ and construct $\bar{Q}$ via Proposition 4.6. The same proposition also shows the inequality needed to obtain

$$\min \mathscr{U} \leq \mathscr{U}(\bar{Q}) \leq \mathbf{U}(\bar{\rho}) = \min \mathbf{U}. \tag{32}$$

Putting together (31) and (32), we obtain at the same time the equality of the minimal values and the optimality of $\bar{\rho}$ in (31) and $\bar{Q}$ in (32). □

5. Banach fixed point

In Section 2 we presented the fixed point problem where we look for a measure $\bar{Q}$ which is a minimizer of the functional $\mathscr{U}_{\bar{Q}}$ and in Section 3 we explained how to prove that such a minimizer is indeed concentrated on curves optimizing $J_{(\bar{Q})}$, i.e. it is a Nash equilibrium. Then, in section 4 we explained how to translate the optimization problem from the Lagrangian language to the Eulerian one.

We want now to consider the fixed-point problem in Eulerian language. First, we define an operator $\mathscr{F}$ on the set of curves valued into $\mathscr{P}(\Omega)$ via

$$\mathscr{F}(\bar{\rho}) := \mathrm{argmin}_\rho \ \mathbf{A}(\rho) + \int_0^T I(\rho(t))dt + \int_0^T \int_\Omega V(t,\bar{\rho}(t),\cdot)d\rho(t)dt + \int_\Omega \psi d\rho(T),$$

where the minimization is performed among left-continuous curves $t \mapsto \rho(t)$ with $\rho(0) = m_0$. Note that the minimizer is unique as soon as I is strictly convex. Then, thanks to the results of the previous sections, we look for a curve $\bar{\rho}$ such that $\bar{\rho} = \mathscr{F}(\bar{\rho})$.

We will consider the case where I is given by (11) and, more generally, we study the dependence of the optimizer in terms of the potential appearing in the linear part of the minimization problem. Let us define a map $\mathscr{O}$ associating with every potential $V : [0,T]\times\Omega \to \mathbb{R}$ the solution ρ of

$$\min \ \mathbf{A}(\rho) + \int_0^T \int_\Omega f(\rho(t,x))dxdt + \int_0^T \int_\Omega V(t,x)\rho(t,x)dxdt + \int_\Omega \psi d\rho(T).$$

We will prove that this map is Lipschitz continuous for the L^2 norm in time-space and, for this, we first need to recall the notion of proximal operator.

Definition 5.1. *Given a Hilbert space H and a function $h\colon H \to (-\infty,\infty]$, the proximal operator of h is defined by*

$$\mathrm{prox}_h(v) = \mathrm{argmin}_{u\in H}\left\{h(u) + \frac{1}{2}\|u-v\|^2\right\}.$$

The prox operator has the property to be $1-$Lipschitz continuous.

Proposition 5.2 (nonexpansivity property)**.** *Let h be a proper closed and convex function. Then for any $v_1, v_2 \in H$ we have*

$$\|\,\mathrm{prox}_h(v_1) - \mathrm{prox}_h(v_2)\| \le \|v_1 - v_2\|.$$

Proof. See for example [2, Theorem 6.42] or the seminal paper [26]. □

We will prove that the operator $\mathscr{O}$ defined by

$$\mathscr{O}(V) := \mathrm{argmin}_\rho \mathbf{A}(\rho) + \int_0^T \int_\Omega f(\rho(t,x))dxdt + \int_0^T \int_\Omega V(t,x)\rho(t,x)dxdt + \int_\Omega \psi d\rho(T)$$

is Lipschitz continuous for the L^2 norm by showing that it can be seen as a proximal operator of a convex function. First, we analyze the convexity of $\mathbf{A}$. We recall the definition of $\mathbf{A}$ in the two cases

$$\text{(jump case) } \mathbf{A}(\rho) = \mathbf{L}(\rho) \qquad \text{(length for the } \mathscr{M} \text{ norm);}$$

$$\text{(kinetic case) } \mathbf{A}(\rho) = \int_0^T \frac{1}{p}|\rho'|_{W_p}(t)^p dt = \inf_{v:\,\partial_t\rho+\nabla\cdot(\rho v)=0} \int_0^T \int_\Omega \frac{1}{p}|v|^p d\rho(t)dt \qquad \text{(action in the } W_p \text{ space).}$$

Lemma 5.3. *Given $\rho_0, \rho_1 : [0,T] \to \mathscr{P}(\Omega)$ two curves of measures, we have*

$$\mathbf{A}(\rho_\lambda) \le (1-\lambda)\mathbf{A}(\rho_0) + \lambda\mathbf{A}(\rho_1),$$

where $\rho_\lambda(t) := (1-\lambda)\rho_0(t) + \lambda\rho_1(t)$.

Proof. We have in both the jump and the kinetic case $\mathbf{A}(\rho) = \min\{\int_{\mathscr{C}} A(\gamma)dQ(\gamma) : e_t\#Q = \rho(t)\}$. Take Q_0, Q_1 optimal in this definition for ρ_0, ρ_1, respectively. Define $Q_\lambda := (1-\lambda)Q_0 + \lambda Q_1$. We have then $e_t\#Q_\lambda = \rho_\lambda(t)$ and hence

$$\mathbf{A}(\rho_\lambda) \le \int_{\mathscr{C}} A(\gamma)dQ_\lambda(\gamma) = (1-\lambda)\mathbf{A}(\rho_0) + \lambda\mathbf{A}(\rho_1). \quad \square$$

We can now use these notions to establish the Lipschitz dependence of the optimal ρ in terms of the data V. Before stating the main estimates, we need to be clear about what is meant by *Lipschitz constant*. In the following, we will use maps defined on a subset of $L^2([0,T]\times\Omega)$ and valued into $L^2([0,T]\times\Omega)$, and denote by $\mathrm{Lip}_{L^2([0,T]\times\Omega)}$ its Lipschitz constant w.r.t. to the L^2 norm in space-time; we will also later use maps defined on a subset of $L^2(\Omega)$ and valued into $L^2(\Omega)$, and denote by $\mathrm{Lip}_{L^2(\Omega)}$ its Lipschitz constant w.r.t. to the L^2 norm in space only. It will also happen later that we consider curves $t \mapsto \rho(t)$ and we will consider their Lipschitz constant as a map from time to $L^2(\Omega)$, which will be denoted by Lip_{t,L^2}.

Proposition 5.4. *Suppose that $\Psi_0 \colon L^1(\Omega) \to \mathbb{R}$ and $\Psi_T \colon L^1(\Omega) \to \mathbb{R}$ are convex and lower semi-continuous on $L^1(\Omega)$. Suppose that $f \colon \mathbb{R} \to \mathbb{R}$ is c_0-convex, i.e suppose that $s \mapsto f(s) - \frac{c_0}{2}s^2$ is convex. Then, the map $\mathscr{O} : L^2([0,T]\times\Omega) \to L^2([0,T]\times\Omega)$ defined via*

$$\mathscr{O}(V) := \mathrm{argmin}_\rho\, \mathbf{A}(\rho) + \int_0^T \int_\Omega (V\rho + f(\rho))dxdt + \Psi_0(\rho(0)) + \Psi_T(\rho(T))$$

satisfies

$$\mathrm{Lip}_{L^2([0,T]\times\Omega)}(\mathscr{O}) \le \frac{1}{c_0}.$$

Note that the functional Ψ_0 can encode the constraint $\rho(0) = m_0$.

Proof. Let us rewrite the functional $\mathscr{F}$ as a proximal operator. Since f is c_0-convex, we note by g the funtion such that $f(s) = g(s) + \frac{c_0}{2}s^2$. We then denote by $\mathbf{G}$ the function defined via

$$\mathbf{G}(\rho) := \mathbf{A}(\rho) + \int_0^T \int_\Omega g(\rho)dxdt + \Psi_0(\rho(0)) + \Psi_T(\rho(T))$$

and we note that $\mathbf{G}$ is convex. We have the following equalities:

$$\begin{aligned}
\mathscr{O}(V) &= \mathrm{argmin}_\rho\, \mathbf{G}(\rho) + \int_0^T \int_\Omega (V\rho + \frac{c_0}{2}\rho^2)dxdt \\
&= \mathrm{argmin}_\rho\, \mathbf{G}(\rho) + \int_0^T \int_\Omega \frac{c_0}{2}\left|\rho + \frac{V}{c_0}\right|^2 dxdt \\
&= \mathrm{argmin}_\rho\, \mathbf{G}(\rho) + \frac{c_0}{2}\left\|\rho + \frac{V}{c_0}\right\|^2_{L^2([0,T]\times\Omega)} \\
&= \mathrm{prox}_{G/c_0}\left(\frac{-V}{c_0}\right).
\end{aligned}$$

Thus, by the 1-Lipschitz property of the proximal operator displayed in Proposition 5.2 we obtained the desired estimate. □

Thanks to these results, we can bound the Lipschitz constant of $\mathscr{F}$.

Proposition 5.5. *Suppose that for each $t \in [0,T]$ the map $\rho \mapsto F(t,\rho,\cdot)$ is L-Lipschitz continuous on $L^2(\Omega)$. Then we have*

$$\mathrm{Lip}_{L^2([0,T]\times\Omega)}(\mathscr{F}) \le \frac{L}{c_0}.$$

Proof. Let us take ρ_0, ρ_1 and define, for $i = 0,1$, the functions $V_i(t,x) := F(t,\rho_i(t),x)$. We have $\mathscr{F}(\rho_i) = \mathscr{O}(V_i)$, hence

$$\|\mathscr{F}(\rho_0) - \mathscr{F}(\rho_1)\|_{L^2([0,T]\times\Omega)} = \|\mathscr{O}(V_0) - \mathscr{O}(V_1)\|_{L^2([0,T]\times\Omega)} \le \frac{1}{c_0}\|V_0 - V_1\|_{L^2([0,T]\times\Omega)}.$$

We then compute

$$\|V_0 - V_1\|^2_{L^2([0,T]\times\Omega)} = \int_0^T \|V_0(t,\cdot) - V_1(t,\cdot)\|^2_{L^2(\Omega)} dt \le \int_0^T L^2\|\rho_0(t,\cdot) - \rho_1(t,\cdot)\|^2_{L^2(\Omega)} dt = L^2\|\rho_0 - \rho_1\|^2_{L^2([0,T]\times\Omega)},$$

which allows to conclude. □

As a consequence, we obtain the following

Theorem 5.6. *If the assumptions of Proposition 5.4 hold and if $L < c_0$, then $\mathscr{F}$ admits a unique fixed point which is the limit of any sequence $(\mathscr{F}^n(\rho_0))_n$ with $\rho_0 \in L^2([0,T]\times\Omega)$.*

Proof. By Proposition 5.5 and $\frac{L}{c_0} < 1$, we see that $\mathscr{F}$ is a contraction. By Banach's fixed point theorem, there exists a unique $\bar\rho$ such that $\mathscr{F}(\bar\rho) = \bar\rho$ and such a fixed point is the limit of the sequence $(\mathscr{F}^n(\rho_0))_n$ for all $\rho_0 \in L^2([0,T]\times\Omega)$ whatever is ρ_0. □

Some examples fit into the framework of the above theorem.

- If $F(t,\rho,x) = V_0(t,x) + g(\eta * \rho(x))$ then we can use $L = \mathrm{Lip}(g)||\eta||_{L^1}$;

- If $F(t,\rho,x) = V_0(t,x) + u_\rho$, where u_ρ is the viscosity solution of $|\nabla u| = g(\eta * \rho)$ with $u = 0$ on $\partial\Omega$ (for a given non-decreasing function g) then we can use $L = \mathrm{Lip}(\log g)||\eta||_{L^2} g(||\eta||_{L^\infty})$;
- the last example can be made more explicit when $\Omega = [a,b]$ in 1D, and we can take $F(t,\rho,x) = V_0(t,x) + u_\rho$, where u_ρ is given by $u_\rho(x) = \min\{\int_a^x g(\rho(y))dy; \int_x^b g(\rho(y))dy\}$ or we can replace, in the integrals, ρ with $\eta * \rho$ (this is not strictly necessary since in dimension one $\rho \in L^2$ already guarantees $u_\rho \in C^0$, but without the convolution this example does not fit the assumptions of Section 2). Then we can use $L = \mathrm{Lip}(g)\frac{(b-a)}{\pi}$, where the constant π comes from the sharp Poincaré constant on $[a,b]$.

Many other examples can be cooked up with convolutions or similar tools. We also want to underline an example which actually does not fit our general framework but could be treated similarly. Indeed, it is possible to also consider the case where $V(t,x)$ actually depends on the whole history $(\rho(t))_t$ and not only on $\rho(t)$. An interesting example is the following one:

$$F_{[\rho]}(t,x) = \frac{1}{t}\int_0^t \rho(s,x)ds.$$

In this example, the running cost paid by an agent at time t depends on the average of the density she saw at the same point in the past. Of course it is also possible to add an exogenous cost V_0 thus getting

$$F_{[\rho]}(t,x) = V_0(t,x) + \frac{1}{t}\int_0^t \rho(s,x)ds$$

and making the problem non-autonomous. It is not possible to treat this MFG using Proposition 5.5 but one can directly look at the Lipschitz dependence of F in terms of ρ in the time-space norm $L^2([0,T]\times\Omega)$. Thanks to the Hardy inequality (see [15, 16]) we have

$$||F_0 - F_1||_{L^2([0,T]\times\Omega)} \le 2||\rho_0 - \rho_1||_{L^2([0,T]\times\Omega)},$$

so that the map $\mathscr{F}$ is a contraction as soon as $c_0 > 2$.

We finish this section with an observation which is specific to the jump case $A = S$. Indeed, in this setting we know from [13] that the solution to any variational problem of the form

$$\min\ \mathbf{A}(\rho) + \int_0^T\int_\Omega f(\rho(t,x))dxdt + \int_0^T\int_\Omega V(t,x)\rho(t,x)dxdt + \int_\Omega \psi d\rho(T)$$

is Lipschitz in time valued in $L^2(\Omega)$ whenever $t \mapsto V(t,\cdot)$ is Lipschitz in time valued in the same space and f is uniformly convex. More precisely we have

$$\mathrm{Lip}_{t,L^2}(\rho) = \sup_{t\in[0,T]} \|\dot\rho(t,\cdot)\|_{L^2(\Omega)} \le \frac{\mathrm{Lip}_{t,L^2}(V)}{c_0}. \tag{33}$$

We then obtain the following result

Proposition 5.7. *Suppose $F(t,\rho,x) = V_0(t,x) + F_1(\rho,x)$ and suppose that the map $\rho \mapsto F_1(\rho,\cdot)$ is L-Lipschitz continuous on $L^2(\Omega)$ and that V_0 is Lipschitz continuous valued into $L^2(\Omega)$. Suppose that f is c_0-convex and $L < c_0$. Then the unique fixed point satisfies*

$$\mathrm{Lip}_{t,L^2}(\bar\rho) \le \frac{\mathrm{Lip}_{t,L^2}V_0}{c_0 - L}.$$

In particular, if F does not depend explicitly on time (i.e. $V_0 = 0$), then $\bar\rho$ is constant in time.

Proof. Using (33) we obtain

$$\mathrm{Lip}_{t,L^2}(\bar\rho) \le \frac{1}{c_0}\mathrm{Lip}_{t,L^2}(V), \text{ where } V(t,x) = V_0(t,x) + F_1(\bar\rho(t),x).$$

We then have

$$\mathrm{Lip}_{t,L^2}V \le \mathrm{Lip}_{t,L^2}V_0 + \mathrm{Lip}_{t,L^2}\left(t \mapsto F_1(\bar\rho(t),\cdot)\right) \le \mathrm{Lip}_{t,L^2}V_0 + L\cdot\mathrm{Lip}_{t,L^2}(\bar\rho).$$

Using $L < c_0$ allows to obtain the claim. □

6. Numerical Simulations

In order to illustrate the equilibria of the MFG we described so far, we perform numerical simulations. The main tool will be the algorithm already used in [13] to solve the non-smooth optimization problem which is solved by the equilibria. The algorithm which was chosen was the (Fast) Dual Proximal Gradient Method from [2]. The main difference here is that we consider mixed MFG, which are not completely variational. We will focus on the problem where $f(\rho) = c_0\frac{\rho^2}{2}$ and we add a coefficient $\lambda > 0$ in front of $\iint |\dot\rho|$ so that we can control the importance of jumping in the model. Roughly, the more λ is large, the more expensive it becomes to jump. The main problem will be

$$\min \int_0^T \int_\Omega \left(\lambda|\dot\rho(t,x)| + F(t,\bar\rho(t),x)\rho(t,x) + c_0\frac{\rho(t,x)^2}{2} \right) dxdt + \psi_0(\rho(0)) + \psi(\rho(T)) \tag{34}$$

and we look for a choice of a curve $\bar\rho$ such that the minimizer is exactly $\bar\rho$. As explained in Section 5, we can recover such a curve as a fixed point of an operator in $L^2([0,T]\times\Omega)$, and this operator is contractant under some assumptions. We will hence stick to the case where F is of the following form

$$F(t,\rho,x) = V_0(t,x) + V[\rho](x),$$

with $\mathrm{Lip}_{L^2(\Omega)}V < c_0$ in order to guarantee the existence of a Banach fixed point. This means that we consider the operator $\mathscr{F}$ described in Section 5 and we compute the sequence $(\mathscr{F}^k(\rho^0))_k$. At each iteration k for a given ρ^k, we use the FDPG method to compute the optimizer, i.e. the next iteration $\rho^{k+1} = \mathscr{F}(\rho^k)$, with the initialization at ρ^k.

Of course, the optimization in the functional space $L^2([0,T]\times\Omega)$ requires a discretization, both in time and space. For simplicity we only consider $\Omega = [0,S]$ to be a one-dimensional space. The sets $\{t_0,\dots,t_K\}$ and $\{x_0,\dots,x_N\}$ will be the regular subdivisions of respectively $[0,T]$ and $[0,S]$. The constants $h = t_1 - t_0$ and $l = x_1 - x_0$ represents the steps of the subdivions. The integral in (34) will be approximated by the left-rectangle method and the solutions will be of the form $(\rho(t_i,x_j))_{\substack{0\le i\le K-1\\ 0\le j\le N-1}}$ in the vector space $\mathbb{R}^{K\cdot N}$.

In this section, we present three examples of functions $V(\rho) = V_i(\rho) + V_0$ for $i = 1,2,3$. It is important to consider a non-autonomous term V_0 in order to have a more dynamic solution (see the conclusion of Proposition 5.7). The function V_0 is defined as

$$V_0(t,x) := \left(\frac{S}{2} + 3\cos\left(\frac{4\pi}{T}t\right) - x\right)^2, \tag{35}$$

so that at time t, it is more interesting to be at position $\frac{S}{2} + 3\cos\left(\frac{4\pi}{T}t\right)$.

Parameter	Value
T	10
S	10
K	1000
N	1000
a	200
h	0.01
L	$6/(c_0h)$
λ	0.2
c_0	3

Table 1. Parameters for the numerical simulation of the solution in examples (i) and (iii).

(i) $V_1[\rho](t,x) = \left(⨍_{x-\delta}^{x+\delta}\rho(t,y)dy\right)^2$.

In this example V_1 is a function of a convolution of ρ (here, the convolution with an indicator function). From a modeling point of view, this means that the running cost at a point depends on the mass of ρ in a fixed neighborhood of such a point. Note that, in many cases, using $V_1[\rho] = \eta * \rho$ would be the first variation of a functional (if η is symmetric it is the first variation of $\rho \mapsto \iint \rho(x)\eta(x-y)\rho(y)dydx$) but this fails when it is composed with a nonlinear function such as the square here, which makes this term non-variational (it would become variational again if convolved a second time, in the form $\eta * (f'(\eta*\rho))$, which is the first variation of $\rho \mapsto \int f(\eta*\rho)$). To be more precise, the expression of V_1 has to be modified when $x \pm \delta$ is out of bounds of $[0,S]$ for a given $\delta > 0$, so a more precise definition

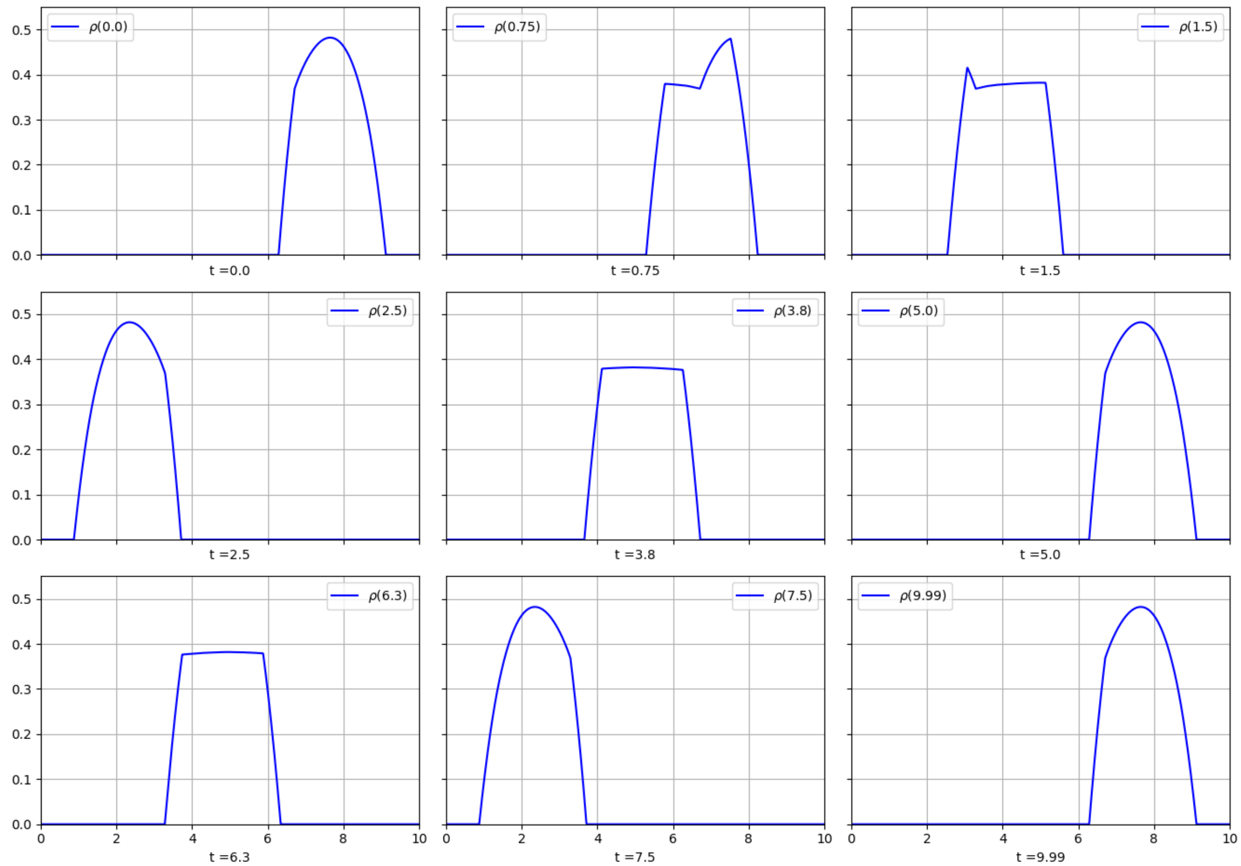

FIGURE 1. Simulation of the fixed point of Example (i) at times $0, 0.75, 1.5, 2.5, 3.8, 5, 6.3, 7.5, 9.99$ for $V(\rho) \equiv V_1(\rho) + V_0$ with V_0 defined in (35) and parameters from Table 1 and $\delta = \frac{1}{2}$.

is the following:

$$V_1(\rho)(t,x) = \begin{cases} \left(⨍_0^{x+\delta} \rho(t,y)dy \right)^2, & \text{if } x \in [0,\delta], \\ \left(⨍_{x-\delta}^{x+\delta} \rho(t,y)dy \right)^2, & \text{if } x \in]\delta, S-\delta], \\ \left(⨍_{x-\delta}^{S} \rho(t,y)dy \right)^2, & \text{if } x \in]S-\delta, S]. \end{cases}$$

The aspect of the solution in shown in Figure 1.

(ii) $V_2(\rho)(t,x) = \min\left(\int_0^x \arctan(\rho(t,y))dy, \int_x^S \arctan(\rho(t,y))dy \right)$.

This second example models a problem where players want to be ready to escape the domain, exiting either from one side or from the other, but the cost to escape depends on congestion. More precisely, a part of the cost is a congested distance to the boundary in the spirit of the Hughes'model for crowd motion (see [18]). We compute the congestion in terms of a non-linear and bounded function of the density (here, the arctangent), so that the resulting function is Lipschitz in x whatever is ρ. In the present case, $V_2 \colon L^2([0,T] \times \Omega) \to L^2([0,T] \times \Omega)$ is a 1-Lipschitz operator (because arctan is 1-Lipschitz), so the parameter $c_0 = 3$ is more than suitable for $\mathscr{F}$ to be a contraction. This example was already mentioned in Section 5 and can be seen as a 1D case of the model where $V(\rho) = V_0 + u_\rho$ with u_ρ the viscosity solution to $|\nabla u| = g(\eta \star \rho)$ with Dirichlet boundary conditions $u = 0$ on $\partial\Omega$ (here we omit the convolution, or, equivalently, we choose $\eta = \delta_0$).

Compared to the previous example, we keep the same values of the parameters in Table 1, but we change the value of λ, taking $\lambda = 1$ (instead of 0.2) in order to observe any effect of an increased cost of jumping.

The profile of the solution is shown in Figure 2.

The increased cost of jumping has influenced the solution to be less flexible than the one in Figure 1. Starting from $t = 0$, as the density moves from the right to the left, the population is more concentrated near the positions where the minimum of V_0 should approach during a large amount of time, namely $x = 7$ and $x = 3$, whereas the density is less important in places where the minimum of V_0 only passes through, e.g. $x = 5$. This behavior may be caused by the cost

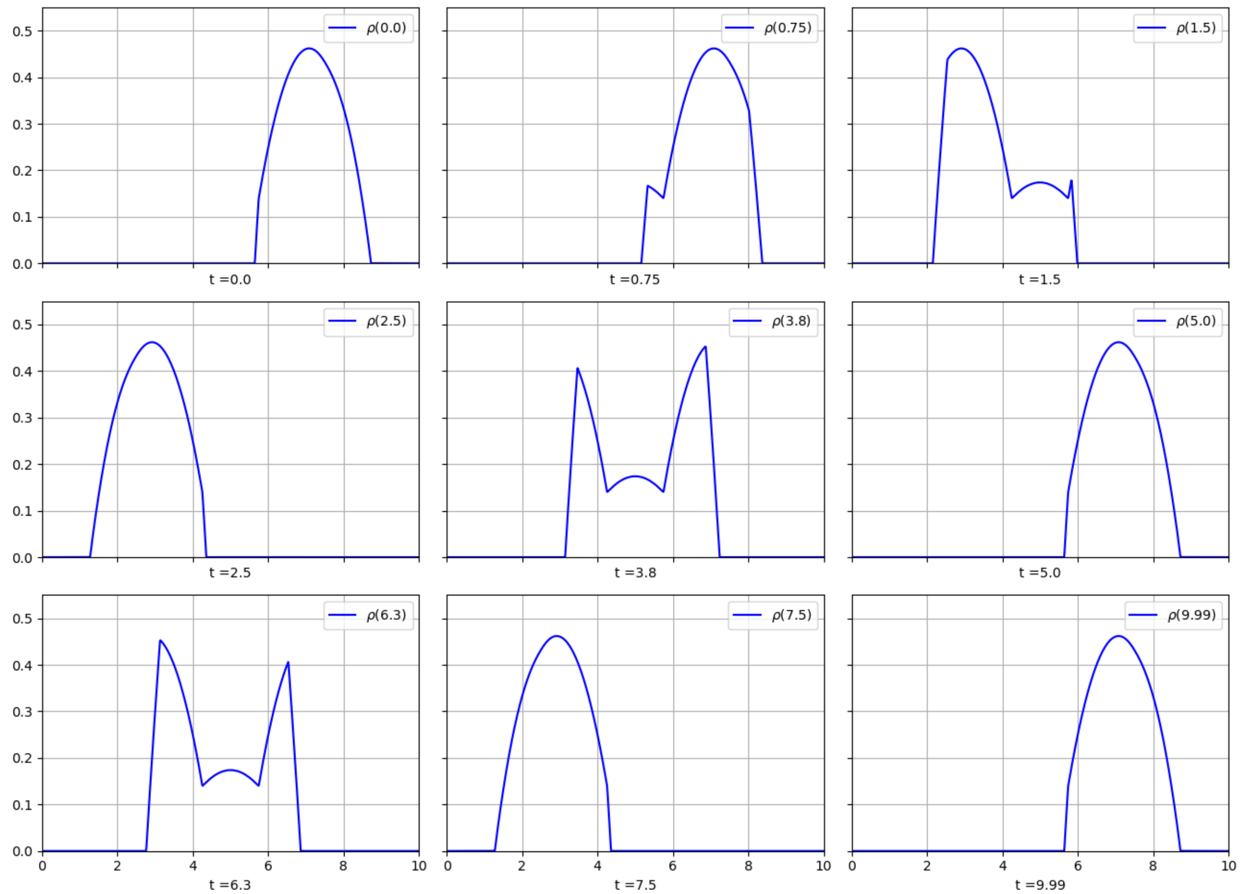

FIGURE 2. Simulation of the fixed point of Example (ii) at times $0, 0.75, 1.5, 2.5, 3.8, 5, 6.3, 7.5, 9.99$ for $V(\rho) \equiv V_2(\rho) + V_0$ with V_0 defined in (35) and parameters from Table 1 except for parameter $\lambda = 1$.

of jumping which is higher than the one in Figure 1. Indeed, the agents anticipate the position at which it will cost less to live, i.e. near the minimum of V_0, in order to minimize the number of jumps.

(iii) $V_3(\rho)(t,x) = \fint_0^t \rho(s,x)ds$.

This example corresponds to the last possibility evoked in Section 5 and is also non-variational. As we already pointed out, the function V_3 is 2-Lipschitz in $L^2([0,T] \times \Omega)$ by Hardy's inequality. Therefore, the parameter c_0 is chosen to be strictly larger than 2. One shall notice that in this case, V_3 at time t does not only depend on ρ_t, but on the whole history. From modeling point of view we consider that agents pay attention to the past experience in estimating the running cost at (t,x).

The profile of the fixed point ρ of $\mathscr{F}$ is displayed in Figure 3 at different times t. One can notice first that at each time the density is concentrated near the minimum of $V_0(t)$, i.e. $\frac{S}{2} + 3\cos\left(\frac{4\pi}{T}t\right)$, which moves from the right to the left and vice versa periodically.

Second, the effect of the term $\fint_0^t \rho(s,x)ds$ can be observed for example at time $t = 0.75$ where the density is reduced at position for example $x = 7$ and a portion of the density directly go to $x = 6$. In fact, the term $\fint_0^t \rho(s,x)ds$ keeps in memory a mean in time of the population who visited position x and a player may not visit a place which was too crowded in the past. In addition, the cost λ for jumping may affect the decision of the players which causes the population to anticipate and go to less crowded places.

REFERENCES

[1] L. AMBROSIO, A. FIGALLI, Geodesics in the space of measure-preserving maps and plans, *Arch. Rational Mech. Anal.*, 194 (2009), 421-462.
[2] A. BECK, *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, 2017.
[3] A. BECK AND M. TEBOULLE, A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems, *SIAM Journal on Imaging Sciences* Vol. 2, Iss. 1, 183-202 (2009)

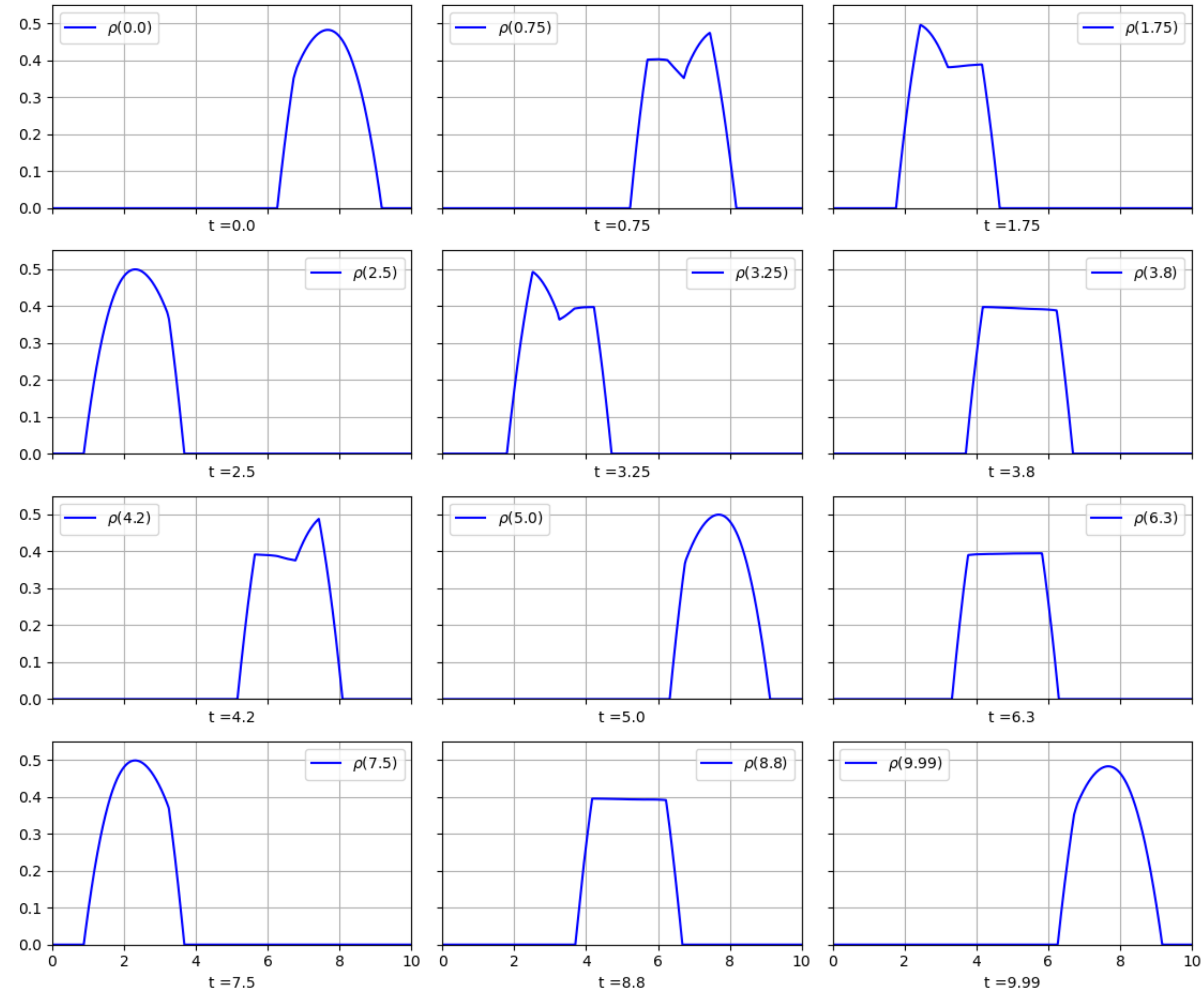

FIGURE 3. Simulation of the fixed point of Example (iii) at times $t = 0, 0.75, 1.75, 2.5, 3.25, 3.8, 4.2, 5, 6.3, 7.5, 8.8, 9.99$ for $V(\rho) \equiv V_3(\rho) + V_0$ with V_0 defined in (35) and parameters from Table 1.

[4] J.-D. BENAMOU AND Y. BRENIER, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84, 375–393, 2000.

[5] J. D. BENAMOU, G. CARLIER, F. SANTAMBROGIO, Variational Mean Field Games, in *Active Particles, Volume 1: Theory, Models, Applications*, edited by N. Bellomo, P. Degond, E. Tadmor, 141–171, 2017.

[6] A. BENSOUSSAN, J.-L. LIONS, *Contrôle impulsionnel et inéquations variationnelles*, Gauthier-Villars, Paris, 1982.

[7] G. BUTTAZZO, C. JIMENEZ, E. OUDET, An Optimization Problem for Mass Transportation with Congested Dynamics *SIAM J. Control Optim.* 48 (2010), 1961–1976.

[8] P. CARDALIAGUET, Notes on Mean Field Games (from P.-L. Lions' lectures at Collège de France), available at `https://www.ceremade.dauphine.fr/ cardalia/`

[9] P. CARDALIAGUET, P. J. GRABER. Mean field games systems of first order. *ESAIM: Control, Optimisation and Calculus of Variations* 21.3 (2015): 690-722.

[10] P. CARDALIAGUET, A. R. MÉSZÁROS, F. SANTAMBROGIO First order Mean Field Games with density constraints: pressure equals price., *SIAM J. Contr. Opt.*, 54(5), 2672-2709, 2016.

[11] C. CASTAING AND M. VALADIER *Convex analysis and measurable multifunctions*. Lecture Notes in Mathematics, **580**. Springer-Verlag, Berlin-New York (1977).

[12] A. DUMAS, *Deterministic Mean Field Games with Jumps*, PhD thesis, Univ. Lyon 1, in preparation.

[13] A. DUMAS, F. SANTAMBROGIO, *Optimal trajectories in L^1 and under L^1 penalizations*. To appear in *Comptes Rendus Math.*

[14] P. J. GRABER Remarks on potential mean field games, preprint, 2024, arxiv 2405.15921

[15] G. H. HARDY, Note on a theorem of Hilbert, *Mathematische Zeitschrift*, vol. 6, nos 3-4, 1920, p. 314-317

[16] G. H. HARDY, J. E. LITTLEWOOD J.E. AND G. PÓLYA, *Inequalities*, 2nd ed, Cambridge University Press, 1952

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

[18] R. L. HUGHES, A continuum theory for the flow of pedestrian, *Transport. Res. Part B* 36 (2002) 507–535.
[19] J.-M. LASRY, P.-L. LIONS, Jeux à champ moyen. I. Le cas stationnaire, *C. R. Math. Acad. Sci. Paris*, 343 (2006), No. 9, 619–625.
[20] J.-M. LASRY, P.-L. LIONS, Jeux à champ moyen. II. Horizon fini et contrôle optimal, *C. R. Math. Acad. Sci. Paris*, 343 (2006), No. 10, 679–684.
[21] J.-M. LASRY AND P.-L. LIONS, Mean-Field Games, *Japan. J. Math*. 2, 229–260, 2007.
[22] S. Lisini, Characterization of absolutely continuous curves in Wasserstein spaces, Calc. Var. Partial Differential Equations **28** (2007), 85–120.
[23] H. LAVENANT, F. SANTAMBROGIO. Optimal density evolution with congestion: L^∞ bounds via flow interchange techniques and applications to variational Mean Field Games. *Communications in Partial Differential Equations*, Vol 43 no 12 (2018), p. 1761–1802.
[24] P.-L. LIONS, Series of lectures on Mean Filed Games, *Collège de France*, Paris, 2006-2012, video-recorderd and available at the web page `http://www.college-de-france.fr/site/audio-video/`
[25] D. Monderer and L.S. Shapley: Potential games. *Games and Economic Behavior* **14**, 124–143 (1996).
[26] J.J. MOREAU, *Proximité et dualité dans un espace hilbertien.* Bulletin de la Société Mathématique de France, Tome 93 (1965), pp. 273-299. doi : 10.24033/bsmf.1625
[27] A. PROSINSKI, F. SANTAMBROGIO Global-in-time regularity via duality for congestion-penalized Mean Field Games. *Stochastics* Vol 89, 2017 - Issue 6-7: Proceedings of the Hammamet Conference, 19-23 October 2015, 923–942.
[28] F. SANTAMBROGIO *Optimal Transport for Applied Mathematicians* in *Progress in Nonlinear Differential Equations and Their Applications* 87, Birkhäuser Basel (2015).
[29] F. SANTAMBROGIO, Lecture notes on Variational Mean Field Games, *Mean Field Games – Cetraro, Italy, 2019*, Cardaliaguet and Porretta (Eds), Springer, C.I.M.E. Foundation Subseries.
[30] C. VILLANI, *Topics in Optimal Transportation*. Graduate Studies in Mathematics, AMS, 2003.