



HAL
open science

Convergence of some Mean Field Games systems to aggregation and flocking models

Martino Bardi, Pierre Cardaliaguet

► **To cite this version:**

Martino Bardi, Pierre Cardaliaguet. Convergence of some Mean Field Games systems to aggregation and flocking models. *Nonlinear Analysis*, 2021, 204, pp.112199. hal-02536846

HAL Id: hal-02536846

<https://hal.science/hal-02536846>

Submitted on 8 Apr 2020

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

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

Convergence of some Mean Field Games systems to aggregation and flocking models

Martino Bardi*and Pierre Cardaliaguet†

April 8, 2020

Abstract

For two classes of Mean Field Game systems we study the convergence of solutions as the interest rate in the cost functional becomes very large, modeling agents caring only about a very short time-horizon, and the cost of the control becomes very cheap. The limit in both cases is a single first order integro-partial differential equation for the evolution of the mass density. The first model is a 2nd order MFG system with vanishing viscosity, and the limit is an aggregation equation. The result has an interpretation for models of collective animal behaviour and of crowd dynamics. The second class of problems are 1st order MFGs of acceleration and the limit is the kinetic equation associated to the Cucker-Smale model. The first problem is analyzed by PDE methods, whereas the second is studied by variational methods in the space of probability measures on trajectories.

Contents

1	Convergence for classical MFG systems	5
1.1	The convergence results	5
1.2	Proof of Theorem 1.1	6
1.3	Examples	9
1.3.1	The aggregation equation	9
1.3.2	Models of crowd dynamics	10
1.3.3	On uniqueness of solutions	10
2	Convergence for some MFGs of acceleration towards the Cucker-Smale model	11
2.1	The convergence result	11
2.2	Proof of the convergence result	13

Introduction

The aim of this work is to discuss, in some particular settings, how models involving crowds of rational agents continuous in space-time can degenerate to agent based models as the agents become less and less rational. The models of rational agents used in this paper are the Mean Field Games (MFG), introduced by Lasry and Lions [24] (see also [23]). They

*Department of Mathematics “T. Levi-Civita”, University of Padova, Via Trieste 63, 35121 Padova, Italy - bardi@math.unipd.it

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

describe optimal control problems with infinitely many infinitesimal agents who interact through their distribution.

Our results are inspired on one hand by the last part of [8], in which the authors show how to derive a McKean-Vlasov equation from a mean field game system and, on the other hand, by [21] (see also [4]) which discusses how multi-agent control problems in which the players have limiting anticipation converge to aggregation models. Let us briefly recall the content of both papers. In [8], the authors study MFG systems of the form

$$\begin{cases} -\partial_t u_\lambda - \nu \Delta u_\lambda + H(x, Du_\lambda, m_\lambda(t)) + \lambda u = 0 & \text{in } \mathbb{R}^d \times (0, T), \\ \partial_t m_\lambda - \nu \Delta m_\lambda - \operatorname{div}(m_\lambda D_p H(x, Du_\lambda, m_\lambda(t))) = 0 & \text{in } \mathbb{R}^d \times (0, T), \\ u_\lambda(T, x) = u_T(x), m_\lambda(0) = m_0 & \text{in } \mathbb{R}^d. \end{cases} \quad (1)$$

Here $\nu > 0$ is fixed, $\lambda > 0$ is a large parameter which describes the impatience of the players and $H = H(x, p, m)$ is the Hamiltonian of the problem which includes interaction terms between the players. Under suitable assumptions on the data, [8] states that, as λ tends to ∞ and up to subsequences, $u_\lambda, Du_\lambda \rightarrow 0$ and m_λ converges to a solution of the McKean-Vlasov equation

$$\begin{cases} \partial_t m - \nu \Delta m - \operatorname{div}(m D_p H(x, 0, m(t))) = 0 & \text{in } \mathbb{R}^d \times (0, T), \\ m(0) = m_0 & \text{in } \mathbb{R}^d. \end{cases}$$

Possible variants and extensions (to MFG models with relative running costs and to higher order approximation) are also discussed in [8].

Although [21] shares some common features with [8], it is quite different. It proposes a deterministic model in which, as in [8], the agents have little rationality, in the sense that they anticipate only on a short horizon (here through time discretization). On the other hand, and this is in contrast with [8], the agents are supposed to pay little for their move. The paper [21] explains, at least at a heuristic level, that the optimal feedback control of each agent should converge to the gradient descent of the running cost, which the authors call ‘‘Best Reply Strategy’’. They also discuss the limit of the distribution of agents as their number goes to infinity and the related 1st order McKean-Vlasov equation.

In the present paper we consider a continuous time variant of the model in [21] which contains its two main features: the fact that the players minimize a cost on a very short horizon, that we model as in [8] by a large discount factor, and the fact that they pay little for their moves. To fit also better with aggregation or kinetic models, we work with problems with a vanishing viscosity ($\nu = \nu_\lambda \rightarrow 0^+$ as $\lambda \rightarrow +\infty$) and in infinite horizon. In particular, our result makes rigorous the approach of [21].

We prove two convergence results. In the first one, our model (in its simplest version) takes the form

$$\begin{cases} -\partial_t u_\lambda - \nu_\lambda \Delta u_\lambda + \lambda u_\lambda + \frac{\lambda}{2} |Du_\lambda|^2 = F(x, m_\lambda(t)) & \text{in } \mathbb{R}^d \times (0, +\infty) \\ \partial_t m_\lambda - \nu_\lambda \Delta m_\lambda - \operatorname{div}(m_\lambda \lambda Du_\lambda) = 0 & \text{in } \mathbb{R}^d \times (0, +\infty) \\ m_\lambda(0) = m_0, & \text{in } \mathbb{R}^d, \quad u_\lambda \text{ bounded.} \end{cases} \quad (2)$$

Under some natural assumptions on F (typically, continuous on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ and uniformly Lipschitz continuous and semi-concave in the space variable), we show that, as λ tends to infinity (meaning that players become more and more myopic and that their control is increasingly cheap) and along subsequences, m_λ converges to a solution m of the aggregation model

$$\begin{cases} \partial_t m - \operatorname{div}(m D_x F(x, m)) = 0 & \text{in } \mathbb{R}^d \times (0, +\infty) \\ m(0) = m_0, & \text{in } \mathbb{R}^d. \end{cases} \quad (3)$$

Moreover, the optimal feedback $-\lambda Du_\lambda$ for the generic agent in the MFG converges a.e. to the vector field $-D_x F(\cdot, m)$, giving the gradient descent of the running cost corresponding to the limit distribution of agents m . To compare with [8], let us note that, in the case where

$H(x, p, m) = \frac{1}{2}|p|^2 - F(x, m)$, the limit of (1) is a simple diffusion equation, while in our setting this limit is non trivial.

The limit equation in (3) covers most examples of the so-called Aggregation Equation

$$\partial_t m + \operatorname{div} \left(m \int_{\mathbb{R}^d} K(x-y)m(y) dy \right) = 0,$$

because the kernel of the convolution is usually the gradient of a potential, $K = -Dk$. This equation describes the collective behavior of various animal populations, its derivation and the choice of the kernel are based on phenomenological considerations, see, e.g., [30, 6] and the references therein. In Subsection 1.3 we show that the examples of Aggregation Equation most studied in the mathematical biology literature fit the assumptions of our convergence theorem, as well as some known models of crowd dynamics. Therefore our result gives a further justification of such models within the framework of dynamic games with a large number of players.

Our second result concerns (first order) MFG of acceleration [1, 12], formally written in the form

$$\begin{cases} -\partial_t u_\lambda + \lambda u_\lambda - v \cdot D_x u_\lambda + \frac{\lambda}{2} |D_v u_\lambda|^2 = F(x, v, m_\lambda(t)) & \text{in } (0, +\infty) \times \mathbb{R}^{2d} \\ \partial_t m_\lambda + v \cdot D_x m_\lambda - \operatorname{div}_v (m_\lambda \lambda D_v u_\lambda) = 0 & \text{in } (0, +\infty) \times \mathbb{R}^{2d} \\ m_\lambda(0) = m_0, & \text{in } \mathbb{R}^{2d}. \end{cases} \quad (4)$$

for which we prove the convergence to kinetic equations of the form

$$\begin{cases} \partial_t m + v \cdot D_x m - \operatorname{div}_v (m D_v F(x, v, m)) = 0 & \text{in } (0, +\infty) \times \mathbb{R}^{2d}, \\ m(0) = m_0, & \text{in } \mathbb{R}^{2d}, \end{cases} \quad (5)$$

as $\lambda \rightarrow +\infty$. To fix the ideas we work in the case where the coupling term F corresponds to the Cucker-Smale model:

$$F(x, v, m(t)) = k * m(x, v, t) = \int_{\mathbb{R}^{2d}} k(x-y, v-v_*) m(y, v_*, t) dy dv_*,$$

where

$$k(x, v) = \frac{|v|^2}{(\alpha + |x|^2)^\beta}, \quad \alpha > 0, \beta \geq 0$$

and v represents the velocity of the generic agent. Note that, in contrast with the first result, the coupling function F is no longer globally Lipschitz continuous: as we explain below, this is a source of major issues and it obliges us to change completely the analysis.

Let us briefly explain the mechanism of proofs and the differences with the existing literature. In [8], the rough idea is that u_λ converges to 0 and therefore Du_λ converges to 0 as well. In addition, the fact that the diffusion is nondegenerate ($\nu > 0$) provides $C^{2+\alpha, 1+\alpha/2}$ bounds on u_λ and m_λ , thanks to which one can pass to the limit.

For our first result, (Theorem 1.1, on the convergence of (2) to (3)), we have to use a different argument. The key idea is that λu_λ behaves like $F(x, m_\lambda)$, because $\lambda^{-1} F(x, m_\lambda)$ is almost a solution to (2). Therefore λDu_λ is close to $DF(x, m_\lambda)$, which explains the limit equation (3). Compared to [8], an additional difficulty comes from the lack of (uniform in λ) smoothness of the solutions, since we have no diffusion term in the limit equation. In particular, the product $m_\lambda \lambda Du_\lambda$ has to be handled with care, since m_λ could degenerate as a measure while Du_λ could become singular. We overcome this issue by proving a uniform semi-concavity of λu_λ , which provides at the same time the L^1_{loc} convergence of λDu_λ and, thanks to an argument going back to [24] (see also [16]) a (locally in time) uniform L^∞ bound on the density of m_λ , and hence a weak-* convergence of m_λ .

For the second result (Theorem 2.2, on the convergence of (4) to (5)), the fact that the coupling function F grows in a quadratic way with respect to the (moment of) the measure prevents us from using fixed point techniques (as in [1, 12]) to show the existence of

a solution to the MFG system (4) and to obtain estimates on the solution (this would also be the case in the presence of a viscous term). This obliges us to give up the PDE approach of the previous set-up and to use variational techniques, first suggested for MFG problems in [24] and largely developed since then: see, for instance, among many other contributions, [5, 14, 15, 17, 26]. For that very same reason, we have to work with a finite horizon problem and with initial measure having a compact support. In contrast with the first result, we do not prove the convergence of *all* the solutions of the MFG system, but only for the ones which minimize the energy written formally as

$$\int_0^T e^{-\lambda t} \int_{\mathbb{R}^{2d}} \left(\frac{1}{2\lambda} |\alpha(x, v, t)|^2 + \int_{\mathbb{R}^{2d}} k(x - x_*, v - v_*) m(dx_*, dv_*, t) \right) m(dx, dv, t) dt \quad (6)$$

where $\partial_t m + v \cdot D_x m + \operatorname{div}_v(m\alpha) = 0$. We formulate this problem in the space of probability measures on curves $(\gamma, \dot{\gamma})$, and the main technique of proof consists in obtaining estimates on the solution based on the dynamic programming principle in such space. This is reminiscent of ideas developed in [29] that we discuss below. Such an approach naturally involves weak solution of the MFG system and does not require the initial measure m_0 to be absolutely continuous. In this case the natural notion of solution for the limit equation (27) is the measure-valued solution developed in [10] for (5).

We could also have developed this second approach for the first type of results (i.e., the convergence of (2) to (3)), assuming that the coupling function F derives from an energy (the so-called potential mean field games):

$$F(x, m) = \frac{\delta F}{\delta m}(m, x)$$

(see [2] for the notion of derivative). Then it is known [24] that minimizers $(m_\lambda, \alpha_\lambda)$ of the problem

$$\inf \left\{ \int_0^{+\infty} e^{-\lambda t} \left(\int_{\mathbb{R}^d} \frac{1}{2} |\alpha|^2 dx + \lambda \mathcal{F}(m(t)) \right) dt, \quad \partial_t m + \operatorname{div}(m\alpha) = 0, \quad m(0) = m_0 \right\}, \quad (7)$$

are solutions to the MFG system (2) (with $\nu_\lambda = 0$ and if \mathcal{F} is smooth enough) in the sense that there exists u_λ such that (u_λ, m_λ) solves (2) and $\alpha_\lambda = -\lambda D u_\lambda$. The convergence of minimizers, as $\lambda \rightarrow +\infty$, is studied in the nice paper [29], where this convergence is called “Weighted Energy-Dissipation”: the authors prove that, under suitable assumptions on the function \mathcal{F} (which allow for singular coupling functions), minimizers converge to a solution of the gradient flow associated to \mathcal{F} , i.e., at least at a formal level, to a solution of (3). Let us note that, in contrast with our setting, the solution of the limit equation can be singular and that [29] works in general metric spaces. It would be interesting to understand the precise interpretation of the results of [29] in terms of limits of MFGs, but this exceeds the scope of the present paper. Note however that our second result (i.e., the convergence of (4) to (5)) does not fit in the framework of [10]. Indeed, the key idea of [10] is that m_λ is a gradient flow for the value function associated with Problem (7); as this value function converges to \mathcal{F} , (m_λ) has to converge to the gradient flow for \mathcal{F} , which is precisely m ; this gradient flow structure is completely lost in our framework of MFG of acceleration (4): we have therefore to design a different approach.

Notation

For any $p \geq 1$ we denote by $\mathcal{P}_p(\mathbb{R}^d)$ (or, in short \mathcal{P}_p) the set of Borel probability measures with finite p -order moment M_p :

$$M_p(m) := \int_{\mathbb{R}^d} |x|^p m(dx).$$

The sets $\mathcal{P}_p(\mathbb{R}^d)$ are endowed with the corresponding Wasserstein distance. Given a positive constant κ , we denote by $\mathcal{M}_{p,\kappa}(\mathbb{R}^d)$ the set of measures $m \in \mathcal{P}_p(\mathbb{R}^d)$ absolutely continuous

with respect to the Lebesgue measure and with a density bounded by κ . We set $\mathcal{M}_p(\mathbb{R}^d) := \bigcup_{\kappa>0} \mathcal{M}_{p,\kappa}(\mathbb{R}^d)$. In Section 2 we will also use, for $m \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$,

$$M_{2,v}(m) := \int_{\mathbb{R}^{2d}} |v|^2 m(dx, dv).$$

Acknowledgment

The first-named author is member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM); he was partially supported by the research project "Nonlinear Partial Differential Equations: Asymptotic Problems and Mean-Field Games" of the Fondazione CaRiPaRo. The second author was partially supported by the ANR (Agence Nationale de la Recherche) project ANR-12-BS01-0008-01, by the CNRS through the PRC grant 1611 and by the Air Force Office for Scientific Research grant FA9550-18-1-0494. This work started during the visit of the second author at Padova University: the University is warmly thanked for this hospitality.

1 Convergence for classical MFG systems

In this section we consider MFG systems of the form

$$\begin{cases} -\partial_t u_\lambda - \nu_\lambda \Delta u_\lambda + \lambda u_\lambda + \lambda^{-1} H(\lambda D u_\lambda, x) = F(x, m_\lambda(t)) & \text{in } \mathbb{R}^d \times (0, +\infty) \\ \partial_t m_\lambda - \nu_\lambda \Delta m_\lambda - \operatorname{div}(m_\lambda D_p H(\lambda D u_\lambda, x)) = 0 & \text{in } \mathbb{R}^d \times (0, +\infty) \\ m_\lambda(0) = m_0, & \text{in } \mathbb{R}^d, \end{cases} \quad (8)$$

where $\lambda > 0$, $\nu_\lambda > 0$ and $\nu_\lambda \rightarrow 0$ as $\lambda \rightarrow +\infty$. Our aim is to prove the convergence (up to a subsequence) of m_λ as $\lambda \rightarrow +\infty$ to a solution m of

$$\begin{cases} \partial_t m - \operatorname{div}(m D_p H(D_x F(x, m(t)), x)) = 0 & \text{in } \mathbb{R}^d \times (0, +\infty), \\ m(0) = m_0, & \text{in } \mathbb{R}^d. \end{cases} \quad (9)$$

and to show also that

$$\lambda u_\lambda(x, t) \rightarrow F(x, m(t)) \quad \text{loc. uniformly}, \quad \lambda D u_\lambda(x, t) \rightarrow D_x F(x, m(t)) \quad \text{a.e.}$$

1.1 The convergence results

We work under the following conditions: we assume that the initial measure m_0 satisfies

$$m_0 \in \mathcal{P}_2(\mathbb{R}^d) \text{ is absolutely continuous with a bounded density.} \quad (10)$$

The kind of costs we are interested in are non-local and regularizing. A possible assumptions on F is that $F : \mathbb{R}^d \times \mathcal{M}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ is continuous in a suitable topology, has a linear growth and is Lipschitz continuous and semi-concave in x . More precisely, we suppose the existence of a constant $C_o \geq 1$ such that:

$$\begin{aligned} & \text{For any } \kappa > 0, \text{ the restrictions of } F \text{ and } D_x F \text{ to } \mathbb{R}^d \times \mathcal{M}_{1,\kappa}(\mathbb{R}^d) \\ & \text{are continuous in both variables for the topology of } \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d), \end{aligned} \quad (11)$$

$$|F(x, m)| \leq C_o(1 + |x|), \quad |F(x, m) - F(y, m)| \leq C_o|x - y|, \quad (12)$$

$$F(x + h, m) + F(x - h, m) - 2F(x, m) \leq C_o|h|^2, \quad \forall m \in \mathcal{M}_1(\mathbb{R}^d), \quad (13)$$

(recall that $\mathcal{M}_{p,\kappa}(\mathbb{R}^d)$ and $\mathcal{M}_p(\mathbb{R}^d)$ are defined in the introduction). We assume that $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is convex with respect to the first variable and satisfies,

$$-C_o \leq H(p, x) \leq C_o(1 + |p|^2), \quad D_{pp}^2 H(p, x) \geq C_o^{-1} I_d, \quad (14)$$

$$|H(p, x) - H(p, y)| + |D_p H(p, x) - D_p H(p, y)| \leq C_o |x - y|(1 + |p|), \quad (15)$$

$$|H(p, x) - H(q, x)| \leq C_o |p - q|(1 + |p| + |q|), \quad (16)$$

$$H(p, x + h) + H(p, x - h) - 2H(p, x) \geq -C_o |h|^2 (1 + |p|). \quad (17)$$

Note that, if H is smooth, then conditions (15), (16) and (17) can be equivalently rewritten as

$$|D_x H(p, x)| + |D_{px} H(p, x)| \leq C_o (1 + |p|),$$

$$|D_p H(p, x)| \leq C_o (1 + |p|), \quad D_{xx}^2 H(p, x) \geq -C_o (1 + |p|).$$

Theorem 1.1. *Assume (10), (11), (12), (13), (14), (15), (16) and (17). Let (u_λ, m_λ) be a solution to (8). Then (m_λ) is relatively compact in $C^0([0, T], \mathcal{P}_1(\mathbb{R}^d))$ and is bounded in $L^\infty(\mathbb{R}^d \times [0, T])$ for any $T > 0$. Moreover, the limit m , as $\lambda_n \rightarrow +\infty$, of any converging subsequence (m_{λ_n}) in $C^0([0, T], \mathcal{P}_1(\mathbb{R}^d))$ is a solution of (9) in the sense of distributions and*

$$\lambda_n u_{\lambda_n}(x, t) \rightarrow F(x, m(t)) \quad \text{locally uniformly and} \quad \lambda_n D u_{\lambda_n}(x, t) \rightarrow DF(x, m(t)) \quad \text{a.e.}$$

The existence of a solution to (8) under the assumptions above can be established by standard arguments, using the estimates in Section 1.2 below, Remark 1.4. A typical example of a Hamiltonian satisfying our assumptions in

$$H(x, p) = -v(x) \cdot p + \frac{1}{2}|p|^2,$$

where the vector field $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is bounded and with bounded first and second order derivatives.

Remark 1.1. In the case of deterministic MFGs, $\nu_\lambda = 0$ for all λ , the solution (u_λ, m_λ) is not smooth and the proof of convergence by PDE methods is harder. We can prove a result analogous to Theorem 1.1 under the additional assumption that $\|F(\cdot, m)\|_{C^2} \leq C$ for all $m \in \mathcal{M}_1(\mathbb{R}^d)$, and the support of m_0 is compact, using the methods of [13]. In this case we can also prove that $m_\lambda(t)$ has a support uniformly bounded for $t \in [0, T]$: a result of this kind is proved in Section 2 for the MFGs of acceleration. Then we expect uniqueness in the limit equation: see the next remark.

Remark 1.2. In addition to the assumptions of Theorem 1.1 suppose that the vector field G appearing in the limit equation (9), $G(x, m) := -D_p H(D_x F(x, m), x)$, is such that, for all $m \in \mathcal{M}_1(\mathbb{R}^d)$, $x \mapsto G(x, m)$ is C^1 and

$$|G(x, m) - G(y, m)| \leq C_1 |x - y|, \quad \|G(\cdot, m) - G(\cdot, \bar{m})\|_\infty \leq C_1 \mathbf{d}_1(m, \bar{m}),$$

where \mathbf{d}_1 is the 1-Wasserstein distance. Then it is proved in [27] that there is a unique solution m of (9) with compact support in x . Therefore, under these additional conditions, the whole family m_λ converges to m as $\lambda \rightarrow +\infty$, as in the problem of Section 2. For instance, the support of m is compact in x in deterministic MFGs, if m_0 has compact support, see the preceding remark.

Remark 1.3. The case of $\nu_\lambda \rightarrow \nu_\infty > 0$ can be treated as in the proof of Theorem 1.1 and leads in the limit to the viscous Fokker-Planck equation

$$\partial_t m - \nu_\infty \Delta m - \operatorname{div}(m D_p H(D_x F(x, m(t)), x)) = 0 \quad \text{in } \mathbb{R}^d \times (0, +\infty).$$

1.2 Proof of Theorem 1.1

In this part, assumptions (10), (11), (12), (13), (14), (15), (16) and (17) are in force. We start with some estimates for a solution to (8).

Proposition 1.2. *Let (u_λ, m_λ) be a solution of (8). Then $|u_\lambda(x, t)| \leq \lambda^{-1} \tilde{C}(1 + |x|)$ for some constant \tilde{C} independent of $\lambda \geq 1 + 4d + 16C_0^2$ such that $\nu_\lambda \leq 1$.*

Proof. We note that $w^\pm(x, t) := \pm\lambda^{-1}\tilde{C}(1 + |x|^2)^{1/2}$ is a supersolution (for +) and a subsolution (for -) of (8) for a suitable \tilde{C} . Let us determine C such that $w = -\lambda^{-1}C(1 + |x|^2)^{1/2}$ is a subsolution, the other case being easier. By the growth assumptions (12) and (14), and for $\nu_\lambda \leq 1$,

$$\begin{aligned} & -\partial_t w - \nu_\lambda \Delta w + \lambda w + \lambda^{-1}H(\lambda Dw, x) - F(x, m_\lambda(t)) \leq \\ & \frac{Cd\nu_\lambda}{\lambda(1 + |x|^2)^{1/2}} - C(1 + |x|^2)^{1/2} + \frac{1}{\lambda}H\left(\frac{-Cx}{(1 + |x|^2)^{1/2}}, x\right) - F(x, m_\lambda(t)) \leq \\ & \frac{Cd}{\lambda} - \frac{C}{2}(1 + |x|) + \frac{C_0}{\lambda}\left(1 + \frac{C^2|x|^2}{(1 + |x|^2)}\right) + C_0(1 + |x|). \end{aligned}$$

If we choose $C = 4C_0$ the right hand side can be bounded above by

$$\frac{4C_0d}{\lambda} - C_0 + C_0\frac{1 + 16C_0^2}{\lambda} \leq 0$$

if $\lambda \geq 1 + 4d + 16C_0^2$. ■

Proposition 1.3. *Let (u_λ, m_λ) be a solution of (8). Then $\|Du_\lambda\|_\infty \leq 4\lambda^{-1}C_o$ for $\lambda \geq 2C_o$.*

Proof. We use an a priori estimate, proving that, if u_λ is Lipschitz continuous and if H and $(x, t) \rightarrow F(x, m_\lambda(t))$ are smooth, then u_λ satisfies the required estimate. One can then complete the proof easily, approximating the HJ equation by HJ equations with smooth and globally Lipschitz continuous Hamiltonians and right-hand sides and passing to the limit. We omit this last part which is standard and proceed with the argument.

Given a direction $\xi \in \mathbb{R}^d$ with $|\xi| \leq 1$, let $w := Du_\lambda \cdot \xi$. Then w satisfies

$$-\partial_t w - \nu_\lambda \Delta w + \lambda w + D_p H(\lambda Du_\lambda, x) \cdot Dw + \lambda^{-1}D_\xi H(\lambda Du_\lambda, x) = D_\xi F(x, m(t)).$$

In view of assumptions (12) and (15) we have therefore

$$-\partial_t w - \nu_\lambda \Delta w + \lambda w + D_p H(\lambda Du_\lambda, x) \cdot Dw - C_o\lambda^{-1}(1 + \lambda\|Du_\lambda\|_\infty) \leq C_o.$$

So by the maximum principle we have

$$Du_\lambda \cdot \xi = w \leq \lambda^{-1}C_o(1 + \lambda^{-1} + \|Du_\lambda\|_\infty).$$

Taking the supremum over $|\xi| \leq 1$, gives the result for λ larger than $2C_o$. ■

Proposition 1.4. *Let (u_λ, m_λ) be a solution of (8). Then $D^2u_\lambda \leq \lambda^{-1}\tilde{C}$, where \tilde{C} does not depend on $\lambda \geq 2C_o$.*

Proof. Here again we focus on a priori estimates for smooth data. Given a direction $\xi \in \mathbb{R}^d$ with $|\xi| \leq 1$, let $w := Du_\lambda \cdot \xi$ and $z := D^2u_\lambda \xi \cdot \xi$. Then

$$\begin{aligned} & -\partial_t z - \nu_\lambda \Delta z + \lambda z + D_p H(\lambda Du_\lambda, x) \cdot Dz + 2D_{p,\xi} H(\lambda Du_\lambda, x) \cdot Dw \\ & + \lambda D_{pp} H(\lambda Du_\lambda, x) Dw \cdot Dw + \lambda^{-1}D_{\xi\xi} H(\lambda Du_\lambda, x) = D_{\xi\xi}^2 F(x, m_\lambda(t)). \end{aligned}$$

Since F is semiconcave in x (13), the right hand side $D_{\xi\xi}^2 F$ is bounded above by C_0 . The uniform bound on λDu_λ proved in Proposition 1.3 and the assumption (15) imply

$$2D_{p,\xi} H(\lambda Du_\lambda, x) \cdot Dw \geq -2C_0(1 + \lambda|Du_\lambda|)|Du_\lambda| \geq -C_1.$$

The same bound and the assumption (17) imply

$$\lambda^{-1}D_{\xi\xi} H(\lambda Du_\lambda, x) \geq -C_0(1 + \lambda|Du_\lambda|) \geq -C_2.$$

Since H is convex in p , $D_{pp}^2 H \geq 0$ and we infer that z satisfies

$$-\partial_t z - \nu_\lambda \Delta z + \lambda z + D_p H(\lambda D u_\lambda, x) \cdot Dz \leq \tilde{C},$$

where the constant \tilde{C} does not depend on λ and $|\xi| \leq 1$. We conclude again by the maximum principle. \blacksquare

Proposition 1.5. *Let (u_λ, m_λ) be a solution of (8). For any $T > 0$, the family (m_λ) satisfies*

$$\sup_{\lambda \geq 2C_0} \sup_{t \in [0, T]} \int_{\mathbb{R}^d} |x|^2 m_\lambda(x, t) dx < +\infty,$$

is relatively compact in $C^0([0, T], \mathcal{P}_1)$ and bounded in $L^\infty(\mathbb{R}^d \times [0, T])$.

Proof. We do the proof again for smooth data. For the bound on the second order moment of $m_\lambda(t)$ on $[0, T]$ we recall that $m_\lambda(t)$ is the law $\mathcal{L}(X_t)$ of the solution X_t of the SDE

$$dX_t = -D_p H(\lambda D u_\lambda(X_t), X_t) dt + \sqrt{2\nu_\lambda} dW_t, \quad \mathcal{L}(X_0) = m_0,$$

where W_t is a standard Brownian motion. Since the vector field $D_p H(\lambda D u_\lambda, x)$ is uniformly bounded by Proposition 1.3, we have $\mathbb{E}[|X_t|^2] \leq C(\mathbb{E}[|X_0|^2] + 1)e^{Ct}$. Then

$$M_2(m_\lambda(t)) = \int_{\mathbb{R}^d} |x|^2 m_\lambda(x, t) dx = \mathbb{E}|X_t|^2 \leq C(M_2(m_0) + 1)e^{Ct}, \quad \forall t \leq T.$$

For the L^∞ bound on m_λ , we rewrite the equation of m_λ as

$$\begin{aligned} \partial_t m_\lambda - \nu_\lambda \Delta m_\lambda - m_\lambda \operatorname{Tr}(D_{pp} H(\lambda D u_\lambda, x) D^2 u_\lambda + D_{px} H(\lambda D u_\lambda, x)) \\ - D m_\lambda \cdot D_p H(D u_\lambda, x) = 0 \quad \text{in } \mathbb{R}^d \times (0, +\infty) \end{aligned}$$

where, by convexity of H , (14) and Proposition 1.4 on the one hand, and by (15) and Proposition 1.3 on the other hand, we have

$$\operatorname{Tr}(D_{pp} H(\lambda D u_\lambda, x) D^2 u_\lambda) \leq C \text{ and } \operatorname{Tr}(D_{px} H(\lambda D u_\lambda, x)) \leq C,$$

where C does not depend on λ . Therefore, by the maximum principle again, the L^∞ norm of m_λ has at most an exponential growth in time, uniform with respect to λ . \blacksquare

Proof of Theorem 1.1. By Proposition 1.5, (m_λ) is relatively compact in $C^0([0, T], \mathcal{P}_1(\mathbb{R}^d))$ and is bounded in $L^\infty(\mathbb{R}^d \times [0, T])$ for any $T > 0$. Let (m_{λ_n}) be a converging subsequence in $C^0([0, T], \mathcal{P}_1)$ for any $T > 0$. Then (m_{λ_n}) converges to m in L^∞ -weak-* on $\mathbb{R}^d \times [0, T]$ for any $T > 0$. In particular, by our continuity assumption on F in (11), the maps $(x, t) \rightarrow F(x, m_{\lambda_n}(t))$ and $(x, t) \rightarrow D_x F(x, m_{\lambda_n}(t))$ converge locally uniformly to the maps $(x, t) \rightarrow F(x, m(t))$ and $(x, t) \rightarrow D_x F(x, m(t))$ respectively.

As u_λ solves (8), $w_\lambda := \lambda u_\lambda$ solves

$$-\lambda^{-1} \partial_t w_\lambda - \lambda^{-1} \nu_\lambda \Delta w_\lambda + w_\lambda + \lambda^{-1} H(D w_\lambda, x) = F(x, m_\lambda(t)) \quad \text{in } \mathbb{R}^d \times (0, +\infty).$$

Hence the half-relaxed limits w^* and w_* of (w_λ) (which are locally uniformly bounded in view of Proposition 1.2) are respectively sub- and super-solutions of the trivial equation

$$w = F(x, m(t)) \quad \text{in } \mathbb{R}^d \times (0, +\infty).$$

This proves the locally uniform convergence of $(\lambda_n u_{\lambda_n})$ to $F(x, m)$.

Next we use Theorem 3.3.3 in [11]. By Proposition 1.2 (λu_λ) is uniformly locally bounded, and by Proposition 1.4 it is uniformly semi-concave in space (locally in time). Then any sequence $(\lambda_n u_{\lambda_n})$ has a subsequence such that $(\lambda_n D u_{\lambda_n})$ converges to $D_x F(x, m)$ a.e. and therefore also in $L^1_{loc}(\mathbb{R}^d \times [0, +\infty))$. One easily derives from this that m solves (9) in the sense of distribution. \blacksquare

Remark 1.4. The existence of a solution (u_λ, m_λ) of the system (8) can be proved by approximating with solutions of the following system with finite time-horizon

$$\begin{cases} -\partial_t u^T - \nu_\lambda \Delta u^T + \lambda u^T + \lambda^{-1} H(\lambda D u^T, x) = F(x, m^T(t)) & \text{in } \mathbb{R}^d \times (0, T) \\ \partial_t m^T - \nu_\lambda \Delta m^T - \operatorname{div}(m^T D_p H(\lambda D u^T, x)) = 0 & \text{in } \mathbb{R}^d \times (0, T) \\ u^T(T) = 0, \quad m^T(0) = m_0, & \text{in } \mathbb{R}^d. \end{cases} \quad (18)$$

The existence of a solution (u^T, m^T) for fixed $\lambda > 0$ follows from standard argument (see for instance Lions' course of Nov. 12, 2010 [25]). The estimates of Propositions 1.2, 1.3, 1.4, and 1.5 hold for (u^T, m^T) with the same proof (using comparison principles for Cauchy problems with constant terminal data). Then there is enough compactness to pass to the limit as $T \rightarrow +\infty$, as in the proof of Theorem 1.1, and see that the limit satisfies (8).

1.3 Examples

In this section we present several examples of coupling functions F of the form

$$F(x, m) = k * m(x, t) = \int_{\mathbb{R}^d} k(x - y) m(dy), \quad (19)$$

where the convolution kernel k can take different forms and is at least globally Lipschitz continuous.

$$H(p) = |p|^2/2 - v(x) \cdot p,$$

with the vector field v bounded together with its first and second derivatives. Then

$$D_p H(D_x F(x, m)) = \int_{\mathbb{R}^d} Dk(x - y) m(dy) - v(x),$$

and the limit equation (9) becomes

$$\begin{cases} \partial_t m + \operatorname{div}(m(v - Q[m])) = 0 & \text{in } \mathbb{R}^d \times (0, +\infty), \quad Q[m](x, t) = \int_{\mathbb{R}^d} Dk(x - y) m(y, t) dy, \\ m(0) = m_0, & \text{in } \mathbb{R}^d. \end{cases} \quad (20)$$

Note that the condition (11) is satisfied. In addition, we suppose that k is bounded, which implies condition (12), and the semi-concavity of k , which ensure condition (13). Under these assumptions Theorem 1.1 holds. Next we review some special cases that arise in applications.

1.3.1 The aggregation equation

The special case of (20) with $v \equiv 0$ is often called the aggregation equation. For suitable choices of the kernel k it models the collective behaviour of groups of animals, see, e.g., [9, 30] and the references therein. Most kernels used in the aggregation models are of the form $k(x) = \phi(|x|)$ with ϕ smooth but $\phi'(0)$ not necessarily 0, so k can be not differentiable in the origin. However, most of them satisfy the assumptions above.

Example 1.1. The kernel

$$k(x) = \alpha e^{-a|x|}, \quad a > 0, \quad (21)$$

considered in [9, 30] (see also the references therein), is bounded, globally Lipschitz continuous, and semiconcave if $\alpha > 0$. Note that the case $\alpha > 0$ describes repulsion among individuals at all distances, because $k(x) = \phi(|x|)$ and $\phi'(r) < 0$ implies repulsion. The case $\alpha < 0$, describing attraction, does not fit into our theory because $k \sim |x|$ near 0, so it is not semiconcave, which is consistent with the fact that solutions of the aggregation equation (20) are known to blow up in finite time for suitable initial data (at least in dimension $d = 1$, see [9]).

Example 1.2. The kernel

$$k(x) = -|x|e^{-a|x|}, \quad a > 0, \quad (22)$$

considered in [9] is also bounded, globally Lipschitz continuous and semiconcave because $k \sim -|x|$ near 0. Note that this kernel describes repulsion at small distance and attraction at distance $|x| > 1/a$. Our theory is consistent with the global existence of solutions of the aggregation equation (20) in this case, at least for $d = 1$, proved in [9].

Example 1.3. To model repulsion at short distance and attraction at medium range, decaying at infinite, a commonly used kernel is the so-called Morse potential

$$k(x) = e^{-|x|} - Ge^{-|x|/L}, \quad 0 < G < 1, \quad L > 1, \quad (23)$$

see [6] and the references therein. It is again bounded and globally Lipschitz continuous. It is also semiconcave because $k \sim 1 - G + |x|(G/L - 1)$ near 0, and $G/L - 1 < 0$.

1.3.2 Models of crowd dynamics

There is a large and fast growing literature on models of the interactions among pedestrians, see the survey in the book [20]. They split into first order models, where the velocity of the pedestrian is a prescribed function of the density of individuals and position, and second order models, where the acceleration is prescribed. In the next Section 2 we study second order models, focusing on the celebrated Cucker-Smale model of flocking, see Remark 2.4 for more references on crowd dynamics.

A first order model fitting in the assumptions of the present section is the one proposed in [19], where the velocity of each agent at position x and time t is of the form $v(x) - Q[m(t)](x)$, v being the desired velocity of the pedestrian, and the other term Q accounting for the interaction with the other agents. If we assume that Q does not depend on the angular focus of the walker in position x , then the model in [19] can be written as

$$Q[m(t)](x) = \int_{\mathbb{R}^d} Dk(x-y)m(y,t)dy, \quad k(x) = \phi(|x|)$$

with $\phi \in Lip([0, +\infty))$, decreasing in $(0, r)$, increasing in (r, R) , and constant in $[R, +\infty)$, so with a behavior similar to the Morse kernel (23) and to (22). If we take $\phi \in C^2((0, +\infty))$ with ϕ'' bounded, then F given by (19) satisfies the assumptions of Theorem 1.1.

1.3.3 On uniqueness of solutions

If we assume in addition that $k \in C^2(\mathbb{R}^d)$ with D^2k bounded, then the limit equation (20) has a unique solution with compact support in space, as observed in Remark 1.2. This occurs, for instance, in Section 1.3.2 if $\phi \in C^2([0, +\infty))$ and $\phi'(0) = \phi''(0) = 0$ (recall that $k(x) = \phi(|x|)$). Uniqueness is also known for the aggregation equation with kernels like those of Section 1.3.1: see [9, 18] and the references therein. However, we expect uniqueness of solutions to the Mean-Field Game system (2) with F given by (19) only for the exponential kernel (21), and not in all other models where there is attraction among individuals in some range of densities. In fact, the uniqueness of solutions in Mean Field Games is strongly connected with a property of monotonicity of F discovered by Lasry and Lions [24]. For coupling functions of the form (19) such monotonicity is equivalent to the property that k is a positive semidefinite kernel, namely,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} k(x-y)v(y)v(x) dydx \geq 0 \quad \forall v.$$

This property is deeply studied and has several characterizations. If $k(x-y) = \psi(|x-y|^2)$ with $\psi \in C^\infty((0, +\infty))$ and continuous in 0, then it is known that k is a positive semidefinite

kernel if and only if ψ is completely monotone, namely, $\psi' \leq 0$ and all other derivatives have alternating signs [22]. In all examples describing attraction it occurs that ϕ' , and therefore ψ' , is instead positive in some range. Then the MFG is not expected to have a unique solution and our result also says that the distance among the possibly multiple solutions of the MFG system tends to 0 as λ becomes large.

2 Convergence for some MFGs of acceleration towards the Cucker-Smale model

For $\lambda > 0$ and $0 < T < +\infty$, we now consider the MFG systems of acceleration, which is written in a formal way as:

$$\begin{cases} -\partial_t u_\lambda + \lambda u_\lambda - v \cdot D_x u_\lambda + \frac{\lambda}{2} |D_v u_\lambda|^2 = F(x, v, m_\lambda(t)) & \text{in } \mathbb{R}^{2d} \times (0, T) \\ \partial_t m_\lambda + v \cdot D_x m_\lambda - \operatorname{div}_v(m_\lambda \lambda D_v u_\lambda) = 0 & \text{in } \mathbb{R}^{2d} \times (0, +\infty) \\ m_\lambda(0) = m_0, \quad u_\lambda(x, v, T) = 0 & \text{in } \mathbb{R}^{2d}. \end{cases} \quad (24)$$

Here the space variables are denoted by (x, v) , with $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$. System (24) models a Nash equilibrium of a game in which the (small) players, given the flow $(m_\lambda(t))$ of probability measures on \mathbb{R}^{2d} , try to minimize over γ the quantity

$$\int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 + F(\gamma(t), \dot{\gamma}(t), m_\lambda(t)) \right) dt,$$

while the flow $(m_\lambda(t))$ is the evolution of the positions and the velocities of the players when they play in an optimal way.

We assume that the coupling function F is a cost associated to the Cucker-Smale model:

$$F(x, v, m(t)) = k * m(x, v, t) = \int_{\mathbb{R}^{2d}} k(x - y, v - v_*) m(y, v_*, t) dy dv_*, \quad k(x, v) = \frac{|v|^2}{g(x)}, \quad (25)$$

where $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is bounded below by a positive constant, is even, smooth and such that $|Dg|/g$ is globally bounded. For instance,

$$g(x) = (\alpha + |x|^2)^\beta, \quad \alpha > 0, \beta \geq 0. \quad (26)$$

In this case $D_v k(x, v) = \frac{2v}{g(x)}$ and so

$$D_v F(x, v, m(t)) = (D_v k) * m(x, v, t) = \int_{\mathbb{R}^{2d}} 2 \frac{(v - v_*)}{g(x - y)} m(y, v_*, t) dy dv_*.$$

The aim of this section is to show that $m_\lambda \rightarrow m$ as $\lambda \rightarrow +\infty$, where m solves the continuous version of the Cucker-Smale model:

$$\begin{cases} \partial_t m + v \cdot D_x m - \operatorname{div}_v(m D_v F(x, v, m)) = 0 & \text{in } \mathbb{R}^{2d} \times (0, +\infty), \\ m(0) = m_0, & \text{in } \mathbb{R}^{2d}. \end{cases} \quad (27)$$

2.1 The convergence result

Throughout this section, we assume that m_0 and F satisfy the following conditions:

$$m_0 \in \mathcal{P}(\mathbb{R}^{2d}) \text{ has a compact support,} \quad (28)$$

and

$$F \text{ is given by (25) where } g : \mathbb{R}^d \rightarrow \mathbb{R} \text{ is bounded below by a positive constant,} \quad (29)$$

is even, smooth, and $|Dg|/g$ is globally bounded.

Let us start by describing what we mean by a weak (variational) solution of the MFG problem. Let $\Gamma = C^1([0, T], \mathbb{R}^d)$ endowed with usual C^1 norm and $\mathcal{P}(\Gamma)$ be the set of Borel probability measures on Γ . We consider, for $\eta \in \mathcal{P}(\Gamma)$,

$$\mathcal{J}_\lambda(\eta) = \int_\Gamma \int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 dt \eta(d\gamma) + \int_0^T e^{-\lambda t} \mathcal{F}(m^\eta(t)) dt,$$

where $m^\eta(t) = \tilde{e}_t \# \eta$ (with $\tilde{e}_t : \Gamma \rightarrow \mathbb{R}^{2d}$, $\tilde{e}_t(\gamma) = (\gamma(t), \dot{\gamma}(t))$) and

$$\mathcal{F}(m) = \frac{1}{2} \int_{\mathbb{R}^{4d}} k(x - x_*, v - v_*) m(dx, dv) m(dx_*, dv_*) \quad \forall m \in \mathcal{P}(\mathbb{R}^{2d}).$$

Lemma 2.1. *For any $\lambda > 0$, there exists at least a minimizer $\bar{\eta}_\lambda$ of \mathcal{J}_λ under the constraint $\tilde{e}_0 \# \bar{\eta}_\lambda = m_0$. It is a weak solution of the MFG problem of acceleration, in the sense that, for $\bar{\eta}_\lambda$ -a.e. $\bar{\gamma} \in \Gamma$,*

$$\begin{aligned} & \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\bar{\gamma}}(t)|^2 + F(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), m^{\bar{\eta}_\lambda}(t)) \right) dt \\ &= \inf_{\gamma \in H^2, (\gamma(0), \dot{\gamma}(0)) = (\bar{\gamma}(0), \dot{\bar{\gamma}}(0))} \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 + F(\gamma(t), \dot{\gamma}(t), m^{\bar{\eta}_\lambda}(t)) \right) dt. \end{aligned} \quad (30)$$

The link between the equilibrium condition (30) and the MFG system (24) is the following: if we set

$$u_\lambda(x, v, s) = \inf_{(\gamma(s), \dot{\gamma}(s)) = (x, v)} \int_s^T e^{-\lambda(t-s)} \left(\frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 + F(\gamma(t), \dot{\gamma}(t), m^{\bar{\eta}_\lambda}(t)) \right) dt,$$

then the pair $(u_\lambda, m^{\bar{\eta}_\lambda})$ is (at least formally) a weak solution of (24), in the sense that u_λ is a viscosity solution to the first equation in (24) while $m^{\bar{\eta}_\lambda}$ is a solution in the sense of distribution of the second equation in (24). Existence of a solution to the equilibrium condition (30) for more general MFG systems is obtained in [12], however under a much more restrictive growth condition on F . In addition, [1, 12] show that there exists a weak solution to the MFG system of acceleration (24).

We postpone the (quite classical) proof of Lemma 2.1 to the next section and proceed with the notion of solution for the kinetic equation (27). Following [10], we say a map $m \in C^0([0, T], \mathcal{P}_2(\mathbb{R}^d))$ is a measure-valued solution to (27) if $m(t) = P^{x,v}(t) \# m_0$ where $P^{x,v}(t) = (P_1^{x,v}(t), P_2^{x,v}(t)) \in \mathbb{R}^d \times \mathbb{R}^d$ solves the ODE

$$\begin{cases} \frac{d}{dt} P_1^{x,v}(t) = P_2^{x,v}(t), \\ \frac{d}{dt} P_2^{x,v}(t) = -D_v F(P_1^{x,v}(t), P_2^{x,v}(t), m(t)), \\ P^{x,v}(0) = (x, v). \end{cases} \quad (31)$$

In [10], the authors propose several conditions under which such a measure-valued solution exists and is unique. This include the case of the Cucker-Smale model studied here, under the assumption that m_0 has a compact support.

Our main result is the following:

Theorem 2.2. *Let $\bar{\eta}_\lambda$ be a minimizer of \mathcal{J}_λ under the constraint $\tilde{e}_0 \# \bar{\eta}_\lambda = m_0$. Then $(m^{\bar{\eta}_\lambda})$ converges as $\lambda \rightarrow +\infty$ to the unique measure-valued solution to (27) in $C_{loc}^0([0, T], \mathcal{P}_2(\mathbb{R}^{2d}))$.*

Remark 2.1. Note that we do not prove the convergence of all the equilibria $(\bar{\eta}_\lambda)$ of (30), but only for the minimizers of \mathcal{J}_λ . The reason is that we were not able to obtain enough estimates for the other equilibria.

2.2 Proof of the convergence result

Before starting the proof, let us note that, by our assumptions, there is a constant $C_0 > 0$ such that

$$g \geq C_0^{-1}, \quad 0 \leq F \leq C_0(1 + |v|^2 + M_{2,v}(m)), \quad \text{where } M_{2,v}(m) := \int_{\mathbb{R}^{2d}} |v|^2 m(dx, dv). \quad (32)$$

$$|D_x F(x, v, m)| \leq C_0 F(x, v, m), \quad |D_v F(x, v, m)| \leq C_0 F^{1/2}(x, v, m). \quad (33)$$

Indeed,

$$|D_x F(x, v, m)| \leq \int_{\mathbb{R}^{2d}} |Dg(x - x_*)| \frac{|v - v_*|^2}{(g(x - x_*))^2} m(dx_*, dv_*) \leq \|Dg/g\|_\infty F(x, v, m),$$

while, as $g \geq c$ (for some $c > 0$),

$$\begin{aligned} |D_v F(x, v, m)| &\leq \int_{\mathbb{R}^{2d}} \frac{2|v - v_*|}{g(x - x_*)} m(dx_*, dv_*) \\ &\leq \left(\int_{\mathbb{R}^{2d}} \frac{|v - v_*|^2}{g(x - x_*)} m(dx_*, dv_*) \right)^{1/2} \left(\int_{\mathbb{R}^{2d}} \frac{4}{g(x - x_*)} m(dx_*, dv_*) \right)^{1/2} \leq 2c^{-1/2} F^{1/2}(x, v, m). \end{aligned}$$

Throughout the proof (and unless specified otherwise), C denotes a constant which may vary from line to line and depends only on T , d , m_0 and the constant C_0 in (32) and (33).

Let us now explain the existence of a minimizer for \mathcal{J}_λ .

Proof of Lemma 2.1. Let $\varepsilon > 0$ and η_ε be ε -optimal in Problem (30). We define $\eta \in \mathcal{P}(\Gamma)$ by

$$\int_\Gamma \phi(\gamma) \eta(d\gamma) = \int_{\mathbb{R}^{2d}} \phi(t \rightarrow x + tv) m_0(dx, dv) \quad \forall \phi \in C_b^0(\Gamma).$$

Let $\pi_2 : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$ defined by $\pi_2(x, v) = v$. Then $\pi_2 \# m^\eta(t) = \pi_2 \# m_0$ for any $t \in [0, T]$ because, for any $\phi \in C_b^0(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \phi(v) \pi_2 \# m^\eta(dv, t) = \int_\Gamma \phi(\dot{\gamma}(t)) \eta(d\gamma) = \int_{\mathbb{R}^{2d}} \phi\left(\frac{d}{dt}(t \rightarrow x + tv)\right) m_0(dx, dv) = \int_{\mathbb{R}^d} \phi(v) \pi_2 \# m_0(dv).$$

Hence, by ε -optimality of η_ε ,

$$\mathcal{J}_\lambda(\eta_\varepsilon) \leq \varepsilon + \mathcal{J}_\lambda(\eta) = \varepsilon + \int_\Gamma \int_0^T e^{-\lambda t} \mathcal{F}(m^\eta(t)) dt,$$

where, for any $t \geq 0$, and as $\pi_2 \# m^\eta(t) = \pi_2 \# m_0$,

$$\mathcal{F}(m^\eta(t)) \leq C_0 \int_{\mathbb{R}^{2d}} |v - v_*|^2 m^\eta(x, v, t) m^\eta(x_*, v_*, t) \leq 2C_0 \int_{\mathbb{R}^{2d}} |v|^2 m^\eta(x, v, t) = 2C_0 M_{2,v}(m_0).$$

This shows that

$$\mathcal{J}_\lambda(\eta_\varepsilon) \leq \varepsilon + 2\lambda^{-1} C_0 M_{2,v}(m_0).$$

As \mathcal{F} is nonnegative, this implies that

$$\int_\Gamma \int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\dot{\gamma}(t)|^2 dt \eta_\varepsilon(d\gamma) \leq \mathcal{J}_\lambda(\eta_\varepsilon) \leq \varepsilon + 2\lambda^{-1} C_0 M_{2,v}(m_0).$$

As m_0 has a compact support (say contained in B_{R_0}) and the set

$$\{\gamma \in \Gamma, |(\gamma(0), \dot{\gamma}(0))| \leq R_0, \int_0^T e^{-\lambda t} |\dot{\gamma}(t)|^2 dt \leq C\}$$

is compact in Γ for any C , we conclude that the family (η_ε) is tight. By lower semi-continuity of \mathcal{J}_λ we can then conclude that there exists a minimizer $\bar{\eta}_\lambda$ of \mathcal{J}_λ under the (closed) constraint $\bar{e}_0 \# \bar{\eta}_\lambda = m_0$. Note for later use that, in view of the above estimates,

$$\int_\Gamma \int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 dt \bar{\eta}_\lambda(d\gamma) \leq 2\lambda^{-1} C_0 M_{2,v}(m_0),$$

so that, as m_0 has a compact support,

$$\sup_{t \in [0, T]} M_{2,v}(m^{\bar{\eta}_\lambda}(t)) \leq C_\lambda, \quad (34)$$

for some constant C_λ depending on m_0 , C_0 and λ .

Next we show equality (30). Let γ_0 belong to the support of $\bar{\eta}_\lambda$ and set $(x_0, v_0) = (\gamma_0(0), \dot{\gamma}_0(0))$. Fix $\gamma_1 \in H^2([0, T], \mathbb{R}^d)$ with $(\gamma_1(0), \dot{\gamma}_1(0)) = (x_0, v_0)$. For $\varepsilon, \delta > 0$, let $E_\varepsilon = \{\gamma \in \Gamma, \|\gamma - \gamma_0\|_{C^1} \leq \varepsilon\}$, $\tilde{m}_\varepsilon = \tilde{e}_t \# (\bar{\eta}_\lambda|_{E_\varepsilon})$ and define $\eta_{\varepsilon, \delta}$ as the Borel measure on Γ by

$$\int_\Gamma \phi(\gamma) \eta_{\varepsilon, \delta}(d\gamma) = \int_{E_\varepsilon} \phi(\gamma) \bar{\eta}_\lambda(d\gamma) + (1-\delta) \int_{E_\varepsilon} \phi(\gamma) \bar{\eta}_\lambda(d\gamma) + \delta \int_{\mathbb{R}^{2d}} \phi(\gamma_1 + (x - x_0 + t(v - v_0))) \tilde{m}_\varepsilon(dx, dv, 0)$$

for any $\phi \in C_b^0(\Gamma)$. Let $\hat{m}_\varepsilon(t)$ be the Borel measure on \mathbb{R}^{2d} defined by

$$\int_{\mathbb{R}^{2d}} \phi(x, v) \hat{m}_\varepsilon(dx, dv, t) = \int_{\mathbb{R}^{2d}} \phi(\gamma_1(t) + x - x_0, \dot{\gamma}_1(t) + v - v_0) \tilde{m}_\varepsilon(dx, dv, 0), \quad \forall \phi \in C_b^0(\mathbb{R}^{2d}).$$

We note that

$$m^{\eta_{\varepsilon, \delta}}(t) = m^{\bar{\eta}_\lambda}(t) + \delta(\hat{m}_\varepsilon(t) - \tilde{m}_\varepsilon(t)), \quad m^{\eta_{\varepsilon, \delta}}(0) = m_0. \quad (35)$$

Hence, testing the optimality of $\bar{\eta}_\lambda$ for \mathcal{J}_λ against $\eta_{\varepsilon, \delta}$ and using the definition of $\eta_{\varepsilon, \delta}$, we obtain

$$\begin{aligned} & \delta \int_{E_\varepsilon} \int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 dt \bar{\eta}_\lambda(d\gamma) + \int_0^T e^{-\lambda t} \mathcal{F}(m^{\bar{\eta}_\lambda}(t)) dt \\ & \leq \delta \left(\int_{\mathbb{R}^{2d}} \tilde{m}_\varepsilon(dx, dv, 0) \right) \left(\int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}_1(t)|^2 dt \right) + \int_0^T e^{-\lambda t} \mathcal{F}(m^{\eta_{\varepsilon, \delta}}(t)) dt. \end{aligned}$$

By definition of \mathcal{F} and the fact that k is even, (35) implies that

$$\begin{aligned} & \delta \int_{E_\varepsilon} \int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 dt \bar{\eta}_\lambda(d\gamma) \leq \delta \left(\int_{\mathbb{R}^{2d}} \tilde{m}_\varepsilon(dx, dv, 0) \right) \left(\int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}_1(t)|^2 dt \right) \\ & + \delta \int_0^T e^{-\lambda t} \int_{\mathbb{R}^{4d}} k(x - x_*, v - v_*) m^{\bar{\eta}_\lambda}(dx_*, dv_*, t) (\hat{m}_\varepsilon - \tilde{m}_\varepsilon)(dx, dv, t) dt \\ & + \frac{\delta^2}{2} \int_0^T e^{-\lambda t} \int_{\mathbb{R}^{4d}} k(x - x_*, v - v_*) (\hat{m}_\varepsilon - \tilde{m}_\varepsilon)(dx_*, dv_*, t) (\hat{m}_\varepsilon - \tilde{m}_\varepsilon)(dx, dv, t) dt. \end{aligned}$$

We divide by $\delta > 0$ and let $\delta \rightarrow 0$ to obtain, using the definition of F :

$$\begin{aligned} & \int_{E_\varepsilon} \int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 dt \bar{\eta}_\lambda(d\gamma) \leq \left(\int_{\mathbb{R}^{2d}} \tilde{m}_\varepsilon(dx, dv, 0) \right) \left(\int_0^T e^{-\lambda t} \frac{1}{2\lambda} |\ddot{\gamma}_1(t)|^2 dt \right) \\ & + \int_0^T e^{-\lambda t} \int_{\mathbb{R}^{2d}} F(x, v, m^{\bar{\eta}_\lambda}(t)) (\hat{m}_\varepsilon(dx, dv, t) - \tilde{m}_\varepsilon(dx, dv, t)) dt. \end{aligned}$$

Rearranging, we find by the definition of \tilde{m}_ε and \hat{m}_ε :

$$\int_{E_\varepsilon} \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 + F(\gamma(t), \dot{\gamma}(t), m^{\bar{\eta}_\lambda}(t)) \right) dt \bar{\eta}_\lambda(d\gamma) \quad (36)$$

$$\leq \int_{\mathbb{R}^{2d}} \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}_1(t)|^2 + F(\gamma_1(t) + x - x_0, \dot{\gamma}_1(t) + v - v_0, m^{\bar{\eta}_\lambda}(t)) \right) dt \tilde{m}_\varepsilon(dx, dv, 0) \quad (37)$$

Fix $\kappa > 0$ small. By lower-semicontinuity on Γ of the functional

$$\gamma \rightarrow \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 + F(\gamma(t), \dot{\gamma}(t), m^{\bar{\eta}^\lambda}(t)) \right) dt,$$

we have, for any $\varepsilon > 0$ small enough, that, for any $\gamma \in E_\varepsilon$,

$$\begin{aligned} & \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}_0(t)|^2 + F(\gamma_0(t), \dot{\gamma}_0(t), m^{\bar{\eta}^\lambda}(t)) \right) dt \\ & \leq \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 + F(\gamma(t), \dot{\gamma}(t), m^{\bar{\eta}^\lambda}(t)) \right) dt + \kappa. \end{aligned}$$

On the other hand, by the regularity of F in (33) and the bound on $M_{2,v}(m^{\bar{\eta}^\lambda}(t))$ in (34), we have, for $|(x, v)| \leq \varepsilon$ and $\varepsilon \in (0, 1)$,

$$\begin{aligned} & \int_0^T e^{-\lambda t} F(\gamma_1(t) + x - x_0, \dot{\gamma}_1(t) + v - v_0, m^{\bar{\eta}^\lambda}(t)) dt \\ & \leq \int_0^T e^{-\lambda t} F(\gamma_1(t), \dot{\gamma}_1(t), m^{\bar{\eta}^\lambda}(t)) dt + C(\gamma_1, \lambda)\varepsilon. \end{aligned}$$

Plugging the inequalities above into (36) gives

$$\begin{aligned} & \bar{\eta}_\lambda(E_\varepsilon) \left(\int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}_0(t)|^2 + F(\gamma_0(t), \dot{\gamma}_0(t), m^{\bar{\eta}^\lambda}(t)) \right) dt + \kappa \right) \\ & \leq \left(\int_{\mathbb{R}^{2d}} \tilde{m}_\varepsilon(dx, dv, 0) \right) \left(\int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}_1(t)|^2 + F(\gamma_1(t), \dot{\gamma}_1(t), m^{\bar{\eta}^\lambda}(t)) \right) dt + C(\gamma_1, \lambda)\varepsilon \right). \end{aligned}$$

As $\bar{\eta}_\lambda(E_\varepsilon) = \left(\int_{\mathbb{R}^{2d}} \tilde{m}_\varepsilon(dx, dv, 0) \right)$, we can divide the inequality above by this quantity (which is positive since γ_0 is in the support of $\bar{\eta}_\lambda$) and then let $\varepsilon \rightarrow 0$, $\kappa \rightarrow 0$ to obtain

$$\begin{aligned} & \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}_0(t)|^2 + F(\gamma_0(t), \dot{\gamma}_0(t), m^{\bar{\eta}^\lambda}(t)) \right) dt \\ & \leq \int_0^T e^{-\lambda t} \left(\frac{1}{2\lambda} |\ddot{\gamma}_1(t)|^2 + F(\gamma_1(t), \dot{\gamma}_1(t), m^{\bar{\eta}^\lambda}(t)) \right) dt, \end{aligned}$$

which gives (30). ■

From now on we fix $\bar{\eta}_\lambda$ a minimizer of \mathcal{J}_λ under the constraint $\tilde{e}_0 \# \bar{\eta}_\lambda = m_0$ and set

$$u_\lambda(x, v, s) = \inf_{\gamma \in H^2, (\gamma(s), \dot{\gamma}(s)) = (x, v)} \int_s^T e^{-\lambda(t-s)} \left(\frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 + F(\gamma(t), \dot{\gamma}(t), m^{\bar{\eta}^\lambda}(t)) \right) dt.$$

We now note that this value function is bounded:

Lemma 2.3. *We have*

$$\mathcal{J}_\lambda(\bar{\eta}_\lambda) \leq 2C_0\lambda^{-1}M_{2,v}(m_0), \quad (38)$$

and, for any $0 \leq s \leq t \leq T$,

$$M_{2,v}(m^{\bar{\eta}^\lambda}(t)) = \int_{\mathbb{R}^{2d}} |v|^2 m^{\bar{\eta}^\lambda}(dx, dv, t) \leq 2(1 + 4C_0\lambda^{-1}e^{\lambda(t-s)})M_{2,v}(m^{\bar{\eta}^\lambda}(s)) \quad (39)$$

and

$$0 \leq u_\lambda(x, v, s) \leq C\lambda^{-1}(1 + |v|^2 + M_{2,v}(m^{\bar{\eta}^\lambda}(s))).$$

Remark 2.2. We use here the fact that we work in a finite horizon problem to obtain the last inequality from (39): see the end of the proof.

Proof. The key point of the proof consists in refining the estimate (34) obtained in the proof of Lemma 2.1. For this we need to introduce a few notations. Given $s \in [0, T]$, let $\Gamma_s = C^1([s, T], \mathbb{R}^d)$ and, for $\eta \in \mathcal{P}(\Gamma_s)$,

$$\mathcal{J}_{\lambda,s}(\eta) = \int_{\Gamma_s} \int_s^T e^{-\lambda(t-s)} \frac{1}{2\lambda} |\ddot{\gamma}(t)|^2 dt \eta(d\gamma) + \int_s^T e^{-\lambda(t-s)} \mathcal{F}(m^\eta(t)) dt,$$

By dynamic programming principle (see Lemma 2.4 below), the restriction $\bar{\eta}_{\lambda,s}$ of $\bar{\eta}_\lambda$ defined by

$$\int_{\Gamma_s} \phi(\gamma) \bar{\eta}_{\lambda,s}(d\gamma) = \int_{\Gamma} \phi(\gamma|_{[s,T]}) \bar{\eta}_\lambda(d\gamma) \quad \forall \phi \in C_b^0(\Gamma_s),$$

is a minimizer of $\eta \rightarrow \mathcal{J}_{\lambda,s}(\eta)$ under the constraint $\tilde{e}_s \# \eta = \tilde{e}_s \# \bar{\eta}_\lambda$.

Defining $\eta \in \mathcal{P}(\Gamma_s)$ by

$$\int_{\Gamma_s} \phi(\gamma) \eta(d\gamma) = \int_{\mathbb{R}^{2d}} \phi(t \rightarrow x + tv) m^{\bar{\eta}_\lambda}(dx, dv, s) \quad \forall \phi \in C_b^0(\Gamma_s),$$

we obtain

$$\mathcal{J}_{\lambda,s}(\bar{\eta}_{\lambda,s}) \leq \mathcal{J}_\lambda(\eta) = \int_{\Gamma} \int_s^T e^{-\lambda(t-s)} \mathcal{F}(m^\eta(t)) dt,$$

where, as in the proof of Lemma 2.1, for any $t \geq s$,

$$\mathcal{F}(m^\eta(t)) \leq C_0 \int_{\mathbb{R}^{2d}} |v - v_*|^2 m^\eta(x, v, t) m^\eta(x_*, v_*, t) \leq 2C_0 M_{2,v}(m^{\bar{\eta}_\lambda}(s)).$$

This shows that

$$\mathcal{J}_{\lambda,s}(\bar{\eta}_{\lambda,s}) \leq 2\lambda^{-1} C_0 M_{2,v}(m^{\bar{\eta}_\lambda}(s)) \quad (40)$$

and inequality (38) holds if we choose $s = 0$.

Next we note that $M_{2,v}(m^{\bar{\eta}_\lambda}(t))$ is finite: we have, for $\bar{\eta}_\lambda$ -a.e. $\bar{\gamma}$, and any $0 \leq s \leq t \leq T$,

$$|\dot{\bar{\gamma}}(t) - \dot{\bar{\gamma}}(s)| \leq \left(\int_s^t e^{-\lambda(\tau-s)} |\ddot{\bar{\gamma}}(\tau)|^2 d\tau \right)^{1/2} \left(\int_s^t e^{\lambda(\tau-s)} d\tau \right)^{1/2},$$

so that (by the elementary inequality $a^2 - 2b^2 \leq 2|a - b|^2$),

$$|\dot{\bar{\gamma}}(t)|^2 \leq 2|\dot{\bar{\gamma}}(s)|^2 + 2\lambda^{-1} e^{\lambda(t-s)} \left(\int_s^t e^{-\lambda(\tau-s)} |\ddot{\bar{\gamma}}(\tau)|^2 d\tau \right).$$

Integrating with respect to $\bar{\eta}_{\lambda,s}$ gives, using (40) in the last inequality,

$$\begin{aligned} \int_{\mathbb{R}^{2d}} |v|^2 m^{\bar{\eta}_\lambda}(dx, dv, t) &= \int_{\Gamma} |\dot{\bar{\gamma}}(t)|^2 \bar{\eta}_\lambda(d\bar{\gamma}) = \int_{\Gamma_s} |\dot{\bar{\gamma}}(t)|^2 \bar{\eta}_{\lambda,s}(d\bar{\gamma}) \\ &\leq 2 \int_{\Gamma_s} |\dot{\bar{\gamma}}(s)|^2 \bar{\eta}_{\lambda,s}(d\bar{\gamma}) + 2\lambda^{-1} e^{\lambda(t-s)} \int_{\Gamma_s} \int_s^T e^{-\lambda(\tau-s)} |\ddot{\bar{\gamma}}(\tau)|^2 d\tau \bar{\eta}_{\lambda,s}(d\bar{\gamma}) \\ &\leq 2 \int_{\mathbb{R}^{2d}} |v|^2 m^{\bar{\eta}_\lambda}(dx, dv, s) + 4e^{\lambda(t-s)} \mathcal{J}_{\lambda,s}(\bar{\eta}_{\lambda,s}) \\ &\leq 2(1 + 4C_0 \lambda^{-1} e^{\lambda(t-s)}) M_{2,v}(m^{\bar{\eta}_\lambda}(s)). \end{aligned}$$

This proves (39). Finally, using $\gamma(t) = x + (t-s)v$ as a test function for $u_\lambda(x, v, s)$, we have:

$$\begin{aligned} u_\lambda(x, v, s) &\leq \int_s^T e^{-\lambda(t-s)} F(x + (t-s)v, v, m^{\bar{\eta}_\lambda}(t)) dt \\ &\leq \int_s^T e^{-\lambda(t-s)} C_0 (1 + |v|^2 + M_{2,v}(m^{\bar{\eta}_\lambda}(t))) dt, \end{aligned}$$

which gives the result thanks to (39). Note that if we were working with an infinite horizon problem, the right-hand side of the inequality above could be unbounded. \blacksquare

Lemma 2.4. *Under the notation of the proof of Lemma 2.3 and for any $s \in [0, T)$, $\bar{\eta}_{\lambda, s}$ is a minimizer of $\eta \rightarrow \mathcal{J}_{\lambda, s}(\eta)$ under the constraint $\tilde{e}_s \# \eta = \tilde{e}_s \# \bar{\eta}_{\lambda}$.*

Proof. Let us set, for $m \in \mathcal{P}(\mathbb{R}^{2d})$ and $s \in [0, T)$,

$$\mathcal{V}_{\lambda}(m, s) = \inf\{\mathcal{J}_{\lambda, s}(\eta), \eta \in \mathcal{P}(\Gamma_s), \tilde{e}_s \# \eta = m\}.$$

We claim that

$$\begin{aligned} \mathcal{V}_{\lambda}(m_0, 0) &= \inf_{\eta \in \mathcal{P}(\Gamma), \tilde{e}_0 \# \eta = m_0} \int_0^s e^{-\lambda\tau} \left(\int_{\Gamma} \frac{1}{2\lambda} |\ddot{\gamma}(\tau)|^2 \eta(d\gamma) + \mathcal{F}(m^{\eta}(\tau)) \right) d\tau + e^{-\lambda s} \mathcal{V}_{\lambda}(m^{\eta}(s), s) \\ &= \int_0^s e^{-\lambda\tau} \left(\int_{\Gamma} \frac{1}{2\lambda} |\ddot{\gamma}(\tau)|^2 \bar{\eta}_{\lambda}(d\gamma) + \mathcal{F}(m^{\bar{\eta}_{\lambda}}(\tau)) \right) d\tau + e^{-\lambda s} \mathcal{V}_{\lambda}(m^{\bar{\eta}_{\lambda}}(s), s), \end{aligned} \quad (41)$$

which proves the lemma. The proof of (41) is a straightforward application of the usual techniques of dynamic programming, the only point being to be able to concatenate at time s two measures $\eta_1 \in \mathcal{P}(\Gamma)$ and $\eta_2 \in \mathcal{P}(\Gamma_s)$ such that $m := \tilde{e}_s \# \eta_1 = \tilde{e}_s \# \eta_2$. For this, let us denote by $\gamma_1 \wedge \gamma_2$ (for $\gamma_1 \in \Gamma$ and $\gamma_2 \in \Gamma_s$ such that $(\gamma_1(s), \dot{\gamma}_1(s)) = (\gamma_2(s), \dot{\gamma}_2(s))$) the map in Γ such that

$$\gamma_1 \wedge \gamma_2(t) = \begin{cases} \gamma_1(t) & \text{if } t \in [0, s], \\ \gamma_2(t) & \text{if } t \in [s, T]. \end{cases}$$

In order to define the concatenation $\eta_1 \wedge \eta_2$, we disintegrate η_1 (respectively η_2) with respect to the measure m . We have

$$\eta_1(d\gamma) = \int_{\mathbb{R}^{2d}} \eta_{1,x,v}(d\gamma) m(dx, dv) \quad (\text{resp. } \eta_2(d\gamma) = \int_{\mathbb{R}^{2d}} \eta_{2,x,v}(d\gamma) m(dx, dv)),$$

where for m -a.e. (x, v) and for $(\eta_{1,x,v} + \eta_{2,x,v})$ -a.e. γ , one has $(\gamma(s), \dot{\gamma}(s)) = (x, v)$. We then define $\eta_1 \wedge \eta_2 \in \mathcal{P}(\Gamma)$ by

$$\int_{\Gamma} \phi(\gamma) (\eta_1 \wedge \eta_2)(d\gamma) = \int_{\mathbb{R}^{2d}} \int_{\Gamma \times \Gamma_s} \phi(\gamma_1 \wedge \gamma_2) \eta_{1,x,v}(d\gamma_1) \eta_{2,x,v}(d\gamma_2) m(dx, dv) \quad \forall \phi \in C_b^0(\Gamma).$$

By construction we have $m^{\eta_1 \wedge \eta_2}(t) = m^{\eta_1}(t)$ if $t \in [0, s]$, $m^{\eta_1 \wedge \eta_2}(t) = m^{\eta_2}(t)$ if $t \in [s, T]$ and

$$\begin{aligned} &\int_0^T e^{-\lambda\tau} \left(\int_{\Gamma} \frac{1}{2\lambda} |\ddot{\gamma}(\tau)|^2 (\eta_1 \wedge \eta_2)(d\gamma) + \mathcal{F}(m^{\eta_1 \wedge \eta_2}(\tau)) \right) d\tau \\ &= \int_0^s e^{-\lambda\tau} \left(\int_{\Gamma} \frac{1}{2\lambda} |\ddot{\gamma}(\tau)|^2 \eta_1(d\gamma) + \mathcal{F}(m^{\eta_1}(\tau)) \right) d\tau \\ &\quad + e^{-\lambda s} \int_s^T e^{-\lambda(\tau-s)} \left(\int_{\Gamma_s} \frac{1}{2\lambda} |\ddot{\gamma}(\tau)|^2 \eta_2(d\gamma) + \mathcal{F}(m^{\eta_2}(\tau)) \right) d\tau. \end{aligned}$$

The rest of the proof of (41) follows then the usual arguments of dynamic programming. \blacksquare

As u_{λ} is the value function of an optimal control problem with smooth (in space) coefficients, it is locally Lipschitz continuous. We now evaluate its derivative with respect to v :

Lemma 2.5. *For any $\varepsilon > 0$, $\lambda \geq \varepsilon^{-1}$, we have*

$$|D_v u_{\lambda}(x, v, s)| \leq C_1 (\lambda^{-1/2} u_{\lambda}^{1/2}(x, v, s) + \varepsilon u_{\lambda}(x, v, s)) \quad \text{for a.e. } (x, v, s) \in \mathbb{R}^{2d} \times [0, T - \varepsilon],$$

where $C_1 = C_0 + 4$.

Proof. Let $\varepsilon > 0$, (x, v, s) be a point of differentiability of u_λ with $s \in [0, T - \varepsilon]$. Let $z^\varepsilon : [0, +\infty) \rightarrow \mathbb{R}$ be defined by $z^\varepsilon(t) = t - \frac{2t^2}{\varepsilon} + \frac{t^3}{\varepsilon^2}$ on $[0, \varepsilon]$ and $z^\varepsilon(t) = 0$ on $[\varepsilon, +\infty)$. Then $z^\varepsilon(0) = z^\varepsilon(\varepsilon) = \dot{z}^\varepsilon(\varepsilon) = 0$, $\dot{z}^\varepsilon(0) = 1$ and $z^\varepsilon \in \dot{H}^2([0, +\infty))$. Therefore, if $\bar{\gamma}$ is optimal for $u_\lambda(x, v, s)$, we have, for any $h \in \mathbb{R}^d$ and using $t \rightarrow \bar{\gamma}(t) + z^\varepsilon(t-s)h$ as a competitor in $v_\lambda(x, v+h, s)$:

$$\begin{aligned} & u_\lambda(x, v+h, s) \\ & \leq \int_s^T e^{-\lambda(t-s)} \left(\frac{1}{2\lambda} |\ddot{\bar{\gamma}}(t) + \ddot{z}^\varepsilon(t-s)h|^2 + F(\bar{\gamma}(t) + z^\varepsilon(t-s)h, \dot{\bar{\gamma}}(t) + \dot{z}^\varepsilon(t-s)h, m^{\bar{\eta}_\lambda}(t)) \right) dt \\ & \leq u_\lambda(x, v) + \int_s^{s+\varepsilon} e^{-\lambda(t-s)} \left(\frac{1}{\lambda} \ddot{\bar{\gamma}}(t) \cdot (\ddot{z}^\varepsilon(t-s)h) + \frac{1}{2\lambda} |\ddot{z}^\varepsilon(t-s)|^2 |h|^2 \right. \\ & \quad \left. + \int_0^1 (D_x F \cdot (z^\varepsilon(t-s)h) + D_v F \cdot (\dot{z}^\varepsilon(t-s)h)) d\tau \right) dt \end{aligned}$$

where for simplicity we have omitted the argument $(\bar{\gamma}(t) + \tau z^\varepsilon(t-s)h, \dot{\bar{\gamma}}(t) + \tau \dot{z}^\varepsilon(t-s)h, m^{\bar{\eta}_\lambda}(t))$ after $D_x F$ and $D_v F$. Dividing by $|h|$ and letting $h \rightarrow 0$ shows that

$$|D_v u_\lambda(x, v, s)| \leq \int_s^{s+\varepsilon} e^{-\lambda(t-s)} \left(\frac{1}{\lambda} |\ddot{\bar{\gamma}}(t)| |\ddot{z}^\varepsilon(t-s)| + |D_x F| |z^\varepsilon(t-s)| + |D_v F| |\dot{z}^\varepsilon(t-s)| \right) dt,$$

where, from now on, F , $D_x F$ and $D_v F$ have for argument $(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), m^{\bar{\eta}_\lambda}(t))$. Recalling (33) and the expression of z^ε we get

$$\begin{aligned} |D_v u_\lambda(x, v, s)| & \leq \lambda^{-1} \left(\int_s^{s+\varepsilon} e^{-\lambda(t-s)} |\ddot{\bar{\gamma}}(t)|^2 dt \right)^{1/2} \left(\int_s^{s+\varepsilon} e^{-\lambda(t-s)} |\ddot{z}^\varepsilon(t-s)|^2 dt \right)^{1/2} \\ & \quad + C_0 \varepsilon \int_s^{s+\varepsilon} e^{-\lambda(t-s)} F dt + C_0 \int_s^{s+\varepsilon} e^{-\lambda(t-s)} F^{1/2} dt \\ & \leq \frac{1}{\lambda} \left(\frac{16}{\varepsilon^2} \frac{1 - e^{-\lambda\varepsilon}}{\lambda} \right)^{1/2} u_\lambda^{1/2}(x, v, s) + C_0 \varepsilon u_\lambda(x, v, s) + C_0 \lambda^{-1/2} u_\lambda^{1/2}(x, v, s). \end{aligned}$$

So, if $\lambda \geq \varepsilon^{-1}$, we obtain

$$|D_v u_\lambda(x, v, s)| \leq (C_0 + 4) \lambda^{-1/2} u_\lambda^{1/2}(x, v, s) + C_0 \varepsilon u_\lambda(x, v, s). \quad \blacksquare$$

Lemma 2.6. *Let $\bar{\gamma}$ be optimal for $u_\lambda(x, v, 0)$. Then we have, for any $t \in [0, T - \varepsilon]$,*

$$|\ddot{\bar{\gamma}}(t)| \leq 2C_1 \left(\lambda^{1/2} u_\lambda^{1/2}(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), t) + \varepsilon \lambda u_\lambda(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), t) \right),$$

where C_1 is the constant in Lemma 2.5.

Remark 2.3. In fact we expect that $\ddot{\bar{\gamma}}(t) = -\lambda D_v u_\lambda(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), t)$ for any $t \in (0, T]$, which would imply the lemma (without the “2” in the right-hand side) thanks to Lemma 2.5. This equality is known to hold in several frameworks [3, 11], but we are not aware of a reference for our precise setting. The estimate in Lemma 2.6, much simpler to prove, suffices however for our purpose.

Proof. As $\bar{\gamma}$ is a minimizer of a calculus of variation problem with smooth coefficients and with quadratic growth, it is known that $\bar{\gamma}$ satisfies the Euler-Lagrange equation

$$\frac{d^2}{dt^2} (\lambda^{-1} e^{-\lambda t} \ddot{\bar{\gamma}}_\lambda(t)) = \frac{d}{dt} (e^{-\lambda t} D_v F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t))) - e^{-\lambda t} D_x F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)).$$

Therefore $\bar{\gamma}$ is actually of class H^4 and, in particular, C^3 .

For $h > 0$ small, let $\gamma_h(s) = \bar{\gamma}(t) + (s-t)\dot{\bar{\gamma}}(t)$. By dynamic programming principle and the optimality of $\bar{\gamma}$ we have:

$$\begin{aligned} u_\lambda(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), t) &= \int_t^{t+h} e^{-\lambda(s-t)} \left(\frac{1}{2\lambda} |\ddot{\bar{\gamma}}(s)|^2 + F(\bar{\gamma}(s), \dot{\bar{\gamma}}(s), m^{\bar{\eta}_\lambda}(s)) \right) ds + e^{-\lambda h} u_\lambda(\bar{\gamma}(t+h), \dot{\bar{\gamma}}(t+h), t+h) \\ &\leq \int_t^{t+h} e^{-\lambda(s-t)} F(\gamma_h(s), \dot{\gamma}_h(s), m^{\bar{\eta}_\lambda}(s)) ds + e^{-\lambda h} u_\lambda(\gamma_h(t+h), \dot{\gamma}_h(t+h), t+h). \end{aligned} \quad (42)$$

Note that, by C^3 regularity of $\bar{\gamma}$, $|\bar{\gamma}(t+h) - \gamma_h(t+h)| \leq C_\gamma h^2$ (where, here and below, C_γ depends here on γ and on λ). So, as u_λ is locally Lipschitz continuous and $\dot{\gamma}_h(t+h) = \dot{\bar{\gamma}}(t)$, we get

$$\begin{aligned} u_\lambda(\gamma_h(t+h), \dot{\gamma}_h(t+h), t+h) - u_\lambda(\bar{\gamma}(t+h), \dot{\bar{\gamma}}(t+h), t+h) \\ \leq u_\lambda(\bar{\gamma}(t+h), \dot{\bar{\gamma}}(t), t+h) - u_\lambda(\bar{\gamma}(t+h), \dot{\bar{\gamma}}(t+h), t+h) + C_\gamma h^2. \end{aligned}$$

Still by C^3 regularity we also have $|\dot{\bar{\gamma}}(t+h) - \dot{\bar{\gamma}}(t) - \ddot{\bar{\gamma}}(t)h| \leq C_\gamma h^2$. Now the bound on $D_v u_\lambda$ of Lemma 2.5 yields (setting $(x, v) = (\bar{\gamma}(t), \dot{\bar{\gamma}}(t))$)

$$\begin{aligned} u_\lambda(\gamma_h(t+h), \dot{\gamma}_h(t+h), t+h) - u_\lambda(\bar{\gamma}(t+h), \dot{\bar{\gamma}}(t+h), t+h) \\ \leq C_1(\lambda^{-1/2} u_\lambda^{1/2}(x, v, t) + \varepsilon u_\lambda(x, v, t)) |\ddot{\bar{\gamma}}(t)| h + C_\gamma h^2. \end{aligned}$$

Plugging this inequality into (42) gives, after dividing by h and letting $h \rightarrow 0$,

$$\begin{aligned} \frac{1}{2\lambda} |\ddot{\bar{\gamma}}(t)|^2 + F(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), m^{\bar{\eta}_\lambda}(t)) &\leq F(\gamma_h(t), \dot{\gamma}_h(t), m^{\bar{\eta}_\lambda}(t)) \\ &\quad + C_1(\lambda^{-1/2} u_\lambda^{1/2}(x, v, t) + \varepsilon u_\lambda(x, v, t)) |\ddot{\bar{\gamma}}(t)|. \end{aligned}$$

Recalling that $(\gamma_h(t), \dot{\gamma}_h(t)) = (\bar{\gamma}(t), \dot{\bar{\gamma}}(t))$ gives the result. \blacksquare

Lemma 2.7. *There exists $\varepsilon_0 > 0$ and a constant $C > 0$ such that, for any $\varepsilon \in (0, \varepsilon_0]$, any $\lambda \geq \varepsilon^{-1} \vee 1$ and any $t \in [0, T - \varepsilon]$, the support of $m^{\bar{\eta}_\lambda}(t)$ is contained in B_C and*

$$\|\ddot{\bar{\gamma}}\|_{L^\infty([0, T - \varepsilon])} \leq C \quad \text{for } \bar{\eta}_\lambda - \text{a.e. } \gamma.$$

In particular, $(\bar{\eta}_\lambda)$ is tight and the family $(m^{\bar{\eta}_\lambda}(t))$ is relatively compact in $C^0([0, T], \mathcal{P}_2(\mathbb{R}^{2d}))$.

Proof. We have, by Lemmata 2.3 and 2.6, for any $\varepsilon > 0$ and $\lambda \geq \varepsilon^{-1}$, and for $\bar{\eta}_\lambda$ -a.e. γ and a.e. $t \in [0, T - \varepsilon]$,

$$\begin{aligned} |\ddot{\bar{\gamma}}(t)| &\leq 2C_1(\lambda^{1/2} u_\lambda^{1/2}(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), t) + \lambda \varepsilon u_\lambda(\bar{\gamma}(t), \dot{\bar{\gamma}}(t), t)) \\ &\leq C(1 + |\dot{\bar{\gamma}}(t)| + M_{2,v}^{1/2}(m^{\bar{\eta}_\lambda}(t))) + C\varepsilon(1 + |\dot{\bar{\gamma}}(t)|^2 + M_{2,v}(m^{\bar{\eta}_\lambda}(t))). \end{aligned} \quad (43)$$

Let us set

$$R_\lambda(t) = \inf\{r > 0, \text{Spt}(m^{\bar{\eta}_\lambda}(t)) \subset \mathbb{R}^d \times B_r\}.$$

We note that R_λ is upper semi-continuous. We now show that R_λ is finite on a maximal time interval $[0, \tau_\lambda)$, with $\tau_\lambda > 0$, with either $\tau_\lambda = T - \varepsilon$ or $\lim_{t \rightarrow \tau_\lambda^-} R_\lambda(t) = +\infty$. For the proof of this fact, λ is fixed and all constants depend on λ unless specified otherwise. By (39) and (43), we have, for $0 \leq s \leq t \leq T - \varepsilon$ and $\bar{\eta}_\lambda$ -a.e. $\bar{\gamma}$,

$$\begin{aligned} |\ddot{\bar{\gamma}}(t)| &\leq C(1 + |\dot{\bar{\gamma}}(t)| + \lambda^{-1/2} e^{\lambda(t-s)/2} M_{2,v}^{1/2}(m^{\bar{\eta}_\lambda}(s))) \\ &\quad + C\varepsilon(1 + |\dot{\bar{\gamma}}(t)|^2 + \lambda^{-1} e^{\lambda(t-s)} M_{2,v}(m^{\bar{\eta}_\lambda}(s))). \end{aligned}$$

Then, as $M_{2,v}(m^{\bar{\eta}^\lambda}(s)) \leq CR_\lambda^2(s)$ for some constant C depending on dimension only,

$$\begin{aligned} |\ddot{\gamma}(t)| &\leq C(1 + |\dot{\gamma}(t)| + \lambda^{-1/2}e^{\lambda(t-s)/2}R_\lambda(s)) \\ &\quad + C\varepsilon(1 + |\dot{\gamma}(t)|^2 + \lambda^{-1}e^{\lambda(t-s)}R_\lambda^2(s)). \end{aligned} \quad (44)$$

So, if $R_\lambda(s)$ is finite for some s and $\lambda \geq 1$, $\varepsilon \leq 1$, one can find K depending only on $R_\lambda(s)$ and the constant C in (44) such that

$$|\dot{\gamma}(t)| \leq |\dot{\gamma}(s)| + K \int_0^t (1 + |\dot{\gamma}(\tau)| + \varepsilon|\dot{\gamma}(\tau)|^2)d\tau.$$

Then we can compare $|\dot{\gamma}(t)|$ with the solution of the ODE

$$\dot{\phi} = K(1 + \phi + \varepsilon\phi^2), \quad \phi(s) = |\dot{\gamma}(s)|,$$

which is given by

$$\phi(t) = \Phi_\varepsilon^{-1}\left(\Phi_\varepsilon(|\dot{\gamma}(s)|) + K(t-s)\right),$$

where

$$\Phi_\varepsilon(r) = \int_0^r \frac{1}{1 + \tau + \varepsilon\tau^2}d\tau.$$

So one can find $\varepsilon_0, \sigma > 0$ depending only on K such that, for all $\varepsilon \in (0, \varepsilon_0]$,

$$|\dot{\gamma}(t)| \leq \phi(t) \leq R_\lambda(s) + 1, \quad \forall t \in [s, s + \sigma],$$

for any $\bar{\gamma} \in H^2$ satisfying (44) and $|\dot{\gamma}(s)| \leq R_\lambda(s)$. As, by definition of R_λ , $m^{\bar{\eta}^\lambda}(s)$ has a support contained in $\mathbb{R}^d \times B_{R_\lambda(s)}$, this shows that $m^{\bar{\eta}^\lambda}(t)$ has a support contained in $\mathbb{R}^d \times B_{R_\lambda(s)+1}$ for any $t \in [s, s + \sigma]$. In particular, as m_0 has a compact support, $R_\lambda(0)$ is finite and thus $R_\lambda(t)$ is finite at least on a small time interval $[0, \sigma]$ for some $\sigma > 0$. We denote by $[0, \tau_\lambda)$ the maximal time interval on which R_λ is finite. Let us assume that $\tau_\lambda < T - \varepsilon$. Let $t_n \rightarrow \tau_\lambda^-$. If $(R_\lambda(t_n))$ remains bounded by a constant M , then by the above argument R_λ is bounded by $M + 1$ on $[\tau_\lambda, \tau_\lambda + \sigma]$ for some $\sigma > 0$ (depending on M), which contradicts the definition of τ_λ . Hence $\lim_{t \rightarrow \tau_\lambda^-} R_\lambda(t) = +\infty$. So we have proved that R_λ is finite on a maximal time interval $[0, \tau_\lambda)$, with $\tau_\lambda > 0$, with either $\tau_\lambda = T - \varepsilon$ or $\lim_{t \rightarrow \tau_\lambda^-} R_\lambda(t) = +\infty$.

By definition of $m^{\bar{\eta}^\lambda}(t)$, for any $\delta > 0$ and $t \in [0, \tau_\lambda)$ there exists $\bar{\gamma} \in \Gamma$ in the support of $\bar{\eta}_\lambda$ such that $|\dot{\gamma}(t)| \geq R_\lambda(t) - \delta$. Thus

$$R_\lambda(t) - \delta \leq |\dot{\gamma}(t)| \leq |\dot{\gamma}(0)| + \int_0^t |\ddot{\gamma}(s)|ds.$$

As $(\bar{\gamma}(t), \dot{\gamma}(t))$ belongs to the support of $m^{\bar{\eta}^\lambda}(t)$ for any t , we get by (43) and the definition of R_λ :

$$\begin{aligned} R_\lambda(t) - \delta &\leq |\dot{\gamma}(0)| + C \int_0^t (1 + |\dot{\gamma}(s)| + M_{2,v}^{1/2}(m^{\bar{\eta}^\lambda}(s)))ds \\ &\quad + C\varepsilon \int_0^t (1 + |\dot{\gamma}(s)|^2 + M_{2,v}(m^{\bar{\eta}^\lambda}(s)))ds \\ &\leq R_0 + C \int_0^t (1 + R_\lambda(s) + \varepsilon R_\lambda^2(s))ds. \end{aligned}$$

As δ is arbitrary, this proves that

$$R_\lambda(t) \leq R_0 + C \int_0^t (1 + R_\lambda(s) + \varepsilon R_\lambda^2(s))ds \quad \forall t \in [0, \tau_\lambda).$$

Arguing as above we get

$$R_\lambda(t) \leq \Phi_\varepsilon^{-1}(\Phi_\varepsilon(R_0) + Ct).$$

For all $\varepsilon > 0$ small enough (but independent of λ) and $\lambda \geq \varepsilon^{-1} \vee 1$, we have therefore that R_λ is bounded by a constant C independent of λ on $[0, \tau_\lambda]$. Thus $\tau_\lambda = T - \varepsilon$ and R_λ is bounded by C on $[0, T - \varepsilon]$.

This estimate gives immediately the bound on $|\dot{\gamma}|$ and therefore, by (43), the bound on $|\ddot{\gamma}|$ for $\bar{\eta}_\lambda$ -a.e. $\bar{\gamma}$. As m_0 has a compact support, this also implies that the $m_\lambda(t)$ have a support contained in a ball B_C , where C is independent of λ and t . In addition the sequence $\bar{\eta}_\lambda$ is tight.

Finally, we have, for any $0 \leq s \leq t \leq T - \varepsilon$,

$$\begin{aligned} \mathbf{d}_1(m^{\bar{\eta}_\lambda}(s), m^{\bar{\eta}_\lambda}(t)) &= \int_\Gamma (|\bar{\gamma}(t) - \bar{\gamma}(s)|^2 + |\dot{\bar{\gamma}}(t) - \dot{\bar{\gamma}}(s)|^2)^{1/2} \bar{\eta}_\lambda(d\bar{\gamma}) \\ &\leq C(t-s)^{1/2} \int_\Gamma \left(\int_s^t |\ddot{\bar{\gamma}}(\tau)|^2 d\tau \right)^{1/2} \bar{\eta}_\lambda(d\bar{\gamma}) \leq C(t-s)^{1/2}. \end{aligned}$$

As the $(m^{\bar{\eta}_\lambda}(t))$ have a support which is uniformly bounded, this shows that it is a relatively compact sequence in $C^0([0, T], \mathcal{P}_2(\mathbb{R}^d))$. \blacksquare

We are now ready to prove the main result:

Proof of Theorem 2.2. In view of Lemma 2.7, $(\bar{\eta}_\lambda)$ is tight and we can consider a subsequence $(\bar{\eta}_{\lambda_n})$ which converges weakly to some η in $\mathcal{P}(\Gamma)$. Then $(m^{\bar{\eta}_{\lambda_n}}(t))$ converges in $C^0([0, T], \mathcal{P}_2(\mathbb{R}^{2d}))$ to $m = \tilde{e}_t \# \eta$. Our aim is to prove that m is a measure valued solution to the kinetic equation (27).

For this we identify the $\limsup(\text{Spt}(\bar{\eta}_{\lambda_n}))$. Let us recall that, by Lemma 2.1, for $\bar{\eta}_\lambda$ -a.e. $\bar{\gamma}_\lambda$, $\bar{\gamma}_\lambda$ minimizes problem (30). Hence by the Euler equation we have that $\bar{\gamma}_\lambda$ is of class H^4 and for a.e. $t \in [0, T]$,

$$\frac{d^2}{dt^2}(\lambda^{-1} e^{-\lambda t} \ddot{\bar{\gamma}}_\lambda(t)) = \frac{d}{dt} (e^{-\lambda t} D_v F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t))) - e^{-\lambda t} D_x F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)).$$

We rewrite this equality as

$$\begin{aligned} \ddot{\bar{\gamma}}_\lambda(t) + D_v F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)) &= \lambda^{-1} \left(-\lambda^{-1} \ddot{\bar{\gamma}}_\lambda^{(iv)}(t) + 2\ddot{\bar{\gamma}}_\lambda(t) + \frac{d}{dt} D_v F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)) \right. \\ &\quad \left. - D_x F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)) \right). \end{aligned}$$

We integrate this equation by parts against a test function $z \in C_c^\infty((0, T), \mathbb{R}^d)$ to get

$$\begin{aligned} &\int_0^T \left(-\dot{\bar{\gamma}}_\lambda(t) \cdot \dot{z}(t) + D_v F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)) \cdot z(t) \right) dt \\ &= \lambda^{-1} \int_0^T \left(\lambda^{-1} \dot{\bar{\gamma}}_\lambda(t) \cdot \ddot{z}(t) + 2\dot{\bar{\gamma}}_\lambda(t) \cdot \dot{z}(t) - D_v F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)) \cdot \dot{z}(t) \right. \\ &\quad \left. - D_x F(\bar{\gamma}_\lambda(t), \dot{\bar{\gamma}}_\lambda(t), m^{\bar{\eta}_\lambda}(t)) \cdot z(t) \right) dt. \end{aligned}$$

By Lemma 2.7 $(\bar{\gamma}_\lambda)$ is relatively compact in Γ , and for any sequence $\lambda_n \rightarrow +\infty$ we can extract a subsequence such that $\bar{\gamma}_{\lambda_n} \rightarrow \gamma \in \Gamma$ and $m^{\bar{\eta}_{\lambda_n}} \rightarrow m \in C^0([0, T], \mathcal{P}_2(\mathbb{R}^{2d}))$. Therefore

$$\int_0^T \left(-\dot{\gamma}(t) \cdot \dot{z}(t) + D_v F(\gamma(t), \dot{\gamma}(t), m(t)) \cdot z(t) \right) dt = 0, \quad \forall z \in C_c^\infty((0, T), \mathbb{R}^d),$$

which means that it is a solution to

$$\dot{\gamma}(t) = -D_v F(\gamma(t), \dot{\gamma}(t), m(t)).$$

In other words, $(\gamma(t), \dot{\gamma}(t)) = P^{x,v}(t)$, where P is defined by (31) and $(x, v) = (\gamma(0), \dot{\gamma}(0))$. By Lemma 2.7 we can also extract a further subsequence such that $\bar{\eta}_{\lambda_n} \rightharpoonup \eta \in \mathcal{P}(\Gamma)$. As the support of η consists of solutions to (31) and $\tilde{e}_0 \# \eta = m_0$, we have

$$\eta = \int_{\mathbb{R}^{2d}} \delta_{P^{x,v}} m_0(dx, dv),$$

so that

$$m(t) = \tilde{e}_t \# \eta = P^{x,v}(t) \# m_0.$$

Hence m is the measure-valued solution to (27). Following [10] this solution is unique. We have proved therefore that any converging subsequence of the relatively compact family $(m^{\bar{\eta}_\lambda})$ has for limit the unique solution m to (27): the entire sequence converges. \blacksquare

Remark 2.4. The Cucker-Smale model is usually associated to the collective animal behaviour, such as flocking of birds or swarming of insects. However, similar models where the acceleration of the agents is prescribed have been proposed for describing the dynamics of crowds of pedestrians, and some of them fit in our results. We refer to the book [20], in particular the section on mesoscopic or kinetic models, and to the recent survey paper [28], where they are called social forces models.

References

- [1] Achdou, Y., Mannucci, P., Marchi, C., Tchou, N. (2019). Deterministic mean field games with control on the acceleration. arXiv:1908.03330, to appear in NoDEA Nonlinear Differential Equations Appl.
- [2] Ambrosio, L., Gigli, N., & Savaré, G. (2008). GRADIENT FLOWS: IN METRIC SPACES AND IN THE SPACE OF PROBABILITY MEASURES. Second edition. Birkhäuser Verlag, Basel, 2008.
- [3] Bardi, M., Capuzzo-Dolcetta, I. (2008). OPTIMAL CONTROL AND VISCOSITY SOLUTIONS OF HAMILTON-JACOBI-BELLMAN EQUATIONS. Birkhäuser Boston, Inc., Boston, MA, 1997.
- [4] Barker, M. (2019). From mean field games to the best reply strategy in a stochastic framework. J. Dyn. Games 6 (2019), no. 4, 291-314.
- [5] Benamou, J. D., Carlier, G., Santambrogio, F. (2017). Variational mean field games. In Active Particles, Volume 1 (pp. 141-171). Birkhäuser, Cham.
- [6] Bernoff, A. J.; Topaz, C. M.: Nonlocal aggregation models: a primer of swarm equilibria. SIAM Rev. 55 (2013), 709–747
- [7] Bertozzi, A. L.; Garnett, J. B.; Laurent, T.: Characterization of radially symmetric finite time blowup in multidimensional aggregation equations. SIAM J. Math. Anal. 44 (2012), 651–681
- [8] C. Bertucci, P.-L. Lions, J.-M. Lasry: Some remarks on Mean Field Games, to appear in Comm. Partial Differential Equations.
- [9] Bodnar, M.; Velazquez, J. J. L. An integro-differential equation arising as a limit of individual cell-based models. J. Differential Equations 222 (2006), 341–380.
- [10] Canizo, J. A., Carrillo, J. A., Rosado, J. (2011). A well-posedness theory in measures for some kinetic models of collective motion. Mathematical Models and Methods in Applied Sciences, 21(03), 515-539.

- [11] Cannarsa, P., Sinestrari, C. (2004). *Semiconcave functions, Hamilton-Jacobi equations, and optimal control* (Vol. 58). Springer Science & Business Media.
- [12] Cannarsa, P., Mendico, C. (2019). Mild and weak solutions of Mean Field Games problem for linear control systems. arXiv preprint arXiv:1907.02654.
- [13] P. Cardaliaguet, Notes on mean field games. Technical report.
- [14] Cardaliaguet, P. (2015). Weak solutions for first order mean field games with local coupling. In *Analysis and geometry in control theory and its applications* (pp. 111-158). Springer, Cham.
- [15] Cardaliaguet, P., Graber, P. J. (2015). Mean field games systems of first order. *ESAIM: Control, Optimisation and Calculus of Variations*, 21(3), 690-722.
- [16] P. Cardaliaguet, S. Hadikhanloo (2017). Learning in mean field games: the fictitious play. *ESAIM: Control, Optimisation and Calculus of Variations*, 23(2), 569-591.
- [17] Cardaliaguet, P., Mészáros, A. R., Santambrogio, F. (2016). First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization*, 54(5), 2672-2709.
- [18] Carrillo, J. A., Rosado: Uniqueness of bounded solutions to aggregation equations by optimal transport methods. *European Congress of Mathematics*, 3-16, Eur. Math. Soc., Zürich, 2010.
- [19] E. Cristiani, B. Piccoli, A. Tosin: Multiscale modeling of granular flows with application to crowd dynamics. *Multiscale Model. Simul.* 9(1) (2011), 155-182.
- [20] E. Cristiani, B. Piccoli, A. Tosin: *Multiscale modeling of pedestrian dynamics*. Springer, Cham, 2014.
- [21] P. Degond, M. Herty, J.G. Liu: Mean field games and model predictive control. *Commun. Math. Sci.* 15 (2017), 1403–1422.
- [22] G.E. Fasshauer, *Meshfree Approximations Methods with Matlab*, World Scientific, Singapore, 2007.
- [23] Huang, M., Malhamé, R. P., Caines, P. E. (2006). Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3), 221-252.
- [24] J.-M. Lasry and P.-L. Lions, Mean field games, *Japanese Journal of Mathematics*, 2 (2007), pp. 229–260.
- [25] Lions, P. L. (2010). *Cours au Collège de France*. Available at www.college-de-france.fr.
- [26] Orrieri, C., Porretta, A., Savaré, G. (2019). A variational approach to the mean field planning problem. *Journal of Functional Analysis*, 277(6), 1868-1957.
- [27] B. Piccoli, F. Rossi, Transport equation with nonlocal velocity in Wasserstein spaces: convergence of numerical schemes. *Acta Appl. Math.* 124 (2013), 73-105.
- [28] B. Piccoli, F. Rossi, *Measure-theoretic models for crowd dynamics in "Crowd Dynamics Volume 1 - Theory, Models, and Safety Problems"*, N. Bellomo and L. Gibelli Eds, Birkhauser, 2018.
- [29] R. Rossi, G. Savaré, A. Segatti, U. Stefanelli (2019). Weighted Energy-Dissipation principle for gradient flows in metric spaces. *Journal de Mathématiques Pures et Appliquées*, 127, 1-66.
- [30] Topaz, C.M.; Bertozzi, A. L.; Lewis, M. A. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.* 68 (2006), 1601–1623.