



**HAL**  
open science

## On the long time behavior of potential MFG

Marco Masoero

► **To cite this version:**

Marco Masoero. On the long time behavior of potential MFG. Probability [math.PR]. Université Paris sciences et lettres, 2019. English. NNT : 2019PSLED057 . tel-03222267

**HAL Id: tel-03222267**

**<https://theses.hal.science/tel-03222267>**

Submitted on 10 May 2021

**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.



**THÈSE DE DOCTORAT**

**DE L'UNIVERSITÉ PSL**

Préparée à Université Paris-Dauphine

**On the long time behavior of potential MFG**

Soutenue par  
MASOERO Marco  
Le 21/11/2019

École doctorale n°543

**ÉCOLE DOCTORALE DE  
DAUPHINE**

Spécialité  
**Sciences**

**Composition du jury :**

CAMILLI Fabio Professeur, Sapienza Università di Roma	<i>Rapporteur</i>
GOMES Diogo Professeur, King Abdullah University of Science and Technology	<i>Rapporteur</i>
CANNARSA Piermarco Professeur, Università degli Studi di Roma "Tor Vergata"	<i>Examineur</i>
CARLIER Guillaume Professeur, Université Paris-Dauphine	<i>Examineur</i>
TONON Daniela Maître de conférences, Université Paris-Dauphine	<i>Examineur</i>
CARDALIAGUET Pierre Professeur, Université Paris-Dauphine	<i>Directeur de thèse</i>
ACHDOU Yves Professeur, Université Paris-Diderot	<i>Président du jury</i>



*alla mia mamma e al mio papà.*



# Acknowledgements

I would like to start this thesis by mentioning the people that made it possible through these complicated years. First of all I would like to express my gratitude towards my supervisor Pierre Cardaliaguet for his continuous and precious support even when I moved back to Italy. I would like to thank him also for the very interesting research subject and for all the advise and help that guided me through this topic that was completely new to me.

I thank the CEREMADE which is made up of kind and welcoming people that I am grateful to have met. In particular Marie Belle, Isabelle Bellier and Cesar Faivre that were always available and the group of PhD students, the old and the new ones. A special thanks goes to William for all the times we grumbled together.

I thank Diogo Gomes and Fabio Camilli for kindly taking the time to read my thesis and Guillaume Carlier, Daniela Tonon, Piermarco Cannarsa and Yves Achdou for participating in the defense committee.

I thank all my friends who gave me fundamental support in these years.

Finally I thank all my family, especially my brothers and sisters because they taught me that *nel male e del male si può ridere*. (D. Masoero)



# Contents

<b>Introduction</b>	<b>v</b>
0.1 The MFG system and the potential setting . . . . .	v
0.1.1 The idea behind Mean field Games . . . . .	v
0.1.2 Potential MFG . . . . .	viii
0.2 The long time behavior and the weak KAM theory . . . . .	x
0.2.1 The long time behavior . . . . .	x
0.2.2 Weak KAM theory . . . . .	xii
0.3 Contributions . . . . .	xvi
0.3.1 On the long time convergence of potential MFG . . . . .	xvii
0.3.2 Weak KAM theory for potential MFG . . . . .	xix
0.3.3 Convergence of the solutions of the MFG discounted Hamilton-Jacobi equation . . . . .	xxi
<b>1 On the long time convergence of potential MFG</b>	<b>1</b>
1.1 Ergodic limit value . . . . .	6
1.1.1 Minimizers and dynamic programming principle for $\mathcal{V}_\delta$ . . . . .	6
1.1.2 Existence of a corrector . . . . .	9
1.1.3 Convergence of $\mathcal{U}(t, \cdot)/t$ and $\delta\mathcal{V}_\delta(\cdot)$ . . . . .	16
1.1.4 Another representation for $\lambda$ . . . . .	17
1.2 Projected Mather set and Calibrated curves . . . . .	19
1.2.1 Calibrated Curves . . . . .	19
1.2.2 The projected Mather set . . . . .	21
1.2.3 Compactness of the projected Mather set . . . . .	23
1.2.4 Minimal invariant set and Ergodicity . . . . .	24
1.3 The role of Monotonicity . . . . .	25
1.3.1 The convex case: $\lambda = \bar{\lambda}$ . . . . .	25
1.3.2 A non convex example where $-\lambda < -\bar{\lambda}$ . . . . .	26
1.4 Appendix . . . . .	29
<b>2 Weak KAM theory for potential MFG</b>	<b>39</b>
2.1 Assumptions and preliminary results . . . . .	43
2.1.1 Notation and assumptions . . . . .	43
2.1.2 Corrector functions and the limit value $\lambda$ . . . . .	45
2.2 A dual problem . . . . .	46
2.3 The $N$ -particle problem. . . . .	49
2.4 On the support of the Mather measures . . . . .	54
2.5 The long time behavior of potential MFG . . . . .	56
2.5.1 Convergence of $\mathcal{U}^T(0, \cdot) + \lambda T$ . . . . .	57



2.5.2	Convergence of optimal trajectories . . . . .	64
2.6	Appendix . . . . .	65
2.6.1	Proof of Proposition 2.3.4 . . . . .	65
2.6.2	A viscosity solution property . . . . .	69
2.6.3	Smooth test functions . . . . .	70
<b>3</b>	<b>Convergence of the MFG discounted HJ equation</b>	<b>73</b>
3.1	Introduction . . . . .	74
3.2	Assumptions and preliminary results . . . . .	76
3.2.1	Notation and assumptions . . . . .	76
3.2.2	Definitions and preliminary results . . . . .	77
3.3	Convergence of $\bar{\mathcal{V}}_\delta$ . . . . .	79
3.3.1	Measures induced by optimal trajectories . . . . .	79
3.3.2	Lower bound . . . . .	81
3.3.3	Smooth Mather measures . . . . .	85
3.3.4	Conclusion . . . . .	87

# Introduction

The aim of this chapter is give some context to the thesis and to present our contributions. We will start with a brief account on the theory of Mean Field Games (MFG). We will introduce the so-called MFG system focusing in particular on the second order case. Afterwards, we will present the potential setting, a special class of MFG models which have a variational formulation. We will then touch upon the weak KAM theory focusing on the main results and ideas that we used in this thesis. Eventually, in the last part of the introduction, we will summarize our contributions which address the problem of the long time behavior of potential MFG.

## 0.1 The MFG system and the potential setting

### 0.1.1 The idea behind Mean field Games

The theory of MFG was introduced in 2006 by Lasry and Lions [61, 62] and, independently, by Huang, Caines and Malhamé [57]. The purpose of this theory is to analyze models of differential games where a large number of small players interact through their repartition density. This means that an agent does not observe every other player individually but she also considers their collective behavior. From the mathematical point of view, the mean field approach drastically reduces the complexity of the systems that have to be analyzed. If we had to keep track of every individual interaction then, for large number of players, we would need to impose a sheer amount of conditions which would lead to intractable mathematical models. On the other hand, to be sure that the mean field approach is meaningful, the large number of players is not enough. Generally speaking, the most important hypotheses that have to be matched are the following: the players have to be sufficiently "small", in the sense that no one individually has an impact on the whole system; the players are identical so that they can be described by a representative agent; lastly, they must be indistinguishable, meaning that if one switches two players then the payoffs of the others do not change.

Keeping in mind this general idea, we now spend a few words to explain (at least at the heuristic level) how the Nash equilibria of some differential games involving an infinite number of players are described by the solutions of the MFG system

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

We now consider a representative agent who has to solve an optimal control problem that depends on the distribution of the other players. We suppose that the dynamic of this agent,

starting at time  $t \in [0, T]$  from  $x \in \mathbb{R}^d$ , evolves according to the stochastic differential equation

$$\begin{cases} dX_s = \alpha(s)ds + \sqrt{2}dB_s \\ X_t = x \end{cases} \quad (1)$$

where  $B_s$  is a standard  $d$ -dimensional Brownian motion and  $\{\alpha(s)\}_{s \in [t, T]}$  is a drift adapted to the filtration induced by  $B_s$ . The agent, at any time  $s \in [t, T]$ , has control over the drift  $\alpha(s)$ . The agent faces a control problem that depends on the evolution of the other players. Therefore, the first step is to guess how they are going to behave in the future. Let us suppose that her guess is represented by the time dependent probability measure  $m(s)$ , then she has to solve the following problem

$$\inf_{\alpha} \mathbb{E} \left[ \int_t^T L(X_s, \alpha(s)) + F(X_s, m(s))ds + G(X_T, m(T)) \right]. \quad (2)$$

While  $L$  is a convex Lagrangian as in the standard optimal control theory, the functions  $F$  and  $G$  embed the dependence of the agent's payoff on the guess  $m(s)$ . Standard arguments in optimal control theory ensure that, if we define the value function

$$u(t, x) = \inf_{\alpha} \mathbb{E} \left[ \int_t^T L(X_s, \alpha(s)) + F(X_s, m(s))ds + G(X_T, m(T)) \right],$$

then the optimal control  $\bar{\alpha}$  is defined by  $\bar{\alpha}(s, x) = -D_p H(x, Du(s, x))$  and the value function  $u$  solves the following Hamilton-Jacobi equation

$$\begin{cases} -\partial_t u(s) - \Delta u + H(x, Du(s)) = F(x, m(s)) \\ u(T, x) = G(x, m(T)), \end{cases}$$

where  $H(x, p) = \inf_{a \in \mathbb{R}^d} -a \cdot p - L(x, a)$ .

Let us now suppose that also the other players make the guess  $m(s)$  and that their initial distribution is described by the probability measure  $m_0$ . Then, as they are identical, they all face the same minimization problem whose solution is the optimal control  $\bar{\alpha}$ .

If they all play the control  $\bar{\alpha}$ , their respective Brownian motions are independent and  $\bar{m}(s)$  is the effective distribution of players at time  $s \in [0, T]$ , then  $\bar{m}(s) = \mathcal{L}(\bar{X}_s)$ , where  $\mathcal{L}(\bar{X}_s)$  is the law of the stochastic process  $\bar{X}_s$  which is the solution of

$$\begin{cases} dX_s = \bar{\alpha}(s)ds + \sqrt{2}dB_s \\ X_0 = Z_0, \quad \mathcal{L}(Z_0) = m_0. \end{cases}$$

It is standard that  $\bar{m}(s)$  verifies the following Fokker-Plank equation

$$\begin{cases} -\partial_t \bar{m} + \Delta \bar{m} + \operatorname{div}(\bar{\alpha} \bar{m}) = 0 & \text{in } [t, T] \times \mathbb{R}^d \\ \bar{m}(0) = m_0. \end{cases}$$

This model is in equilibrium when the guess  $m$  coincides with the real evolution  $\bar{m}$ , so that the strategy  $\bar{\alpha}$  is the optimal response to what players actually observe. Therefore, the Nash equilibrium is described by the MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{R}^d \times [t, T] \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{R}^d \times [t, T] \\ m(t) = m_0, u(T, x) = G(x, m(T)) & \text{in } \mathbb{R}^d. \end{cases} \quad (3)$$

Similarly one can define the *Ergodic MFG* system. Let us suppose that our infinitesimal player moves again according to (1) but instead of facing the optimal control problem (2), she has to minimize

$$\liminf_{T \rightarrow +\infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T L(X_s, \alpha(s)) + F(X_s, m) ds \right]$$

among all possible controls  $\alpha$ . In this case the probability measure  $m$ , that appears in the coupling function  $F$ , represents the guess of the agent on the limit distribution of players when the horizon  $T$  goes to  $+\infty$ . Then, if she starts from position  $x$ , her value function will be

$$u(x) = \inf_{\alpha} \liminf_{T \rightarrow +\infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T L(X_s, \alpha(s)) + F(X_s, m) ds \right].$$

Using again standard argument from the theory of optimal control, we can prove that the optimal strategy is  $\alpha^*(x) = -D_p H(x, Du(x))$  and that there exists a  $\bar{\lambda} \in \mathbb{R}$  such that  $(u, \bar{\lambda})$  solves

$$-\bar{\lambda} - \Delta u + H(x, Du) = F(x, m). \quad (4)$$

Again, if  $\bar{m}$  is the actual limit configuration of players when they all share the same guess  $m$  and play the strategy  $\alpha^*$ , we say that we have a Nash equilibrium when the guess  $m$  coincides with  $\bar{m}$ . One can prove that  $\bar{m}$  solves

$$\begin{cases} -\Delta \bar{m} - \operatorname{div}(D_p H(x, Du) \bar{m}) = 0 & \text{in } \mathbb{R}^d \\ \int d\bar{m} = 1. \end{cases}$$

Therefore, the Nash equilibrium is described by ergodic MFG system

$$\begin{cases} -\bar{\lambda} - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{R}^d \\ -\Delta m - \operatorname{div}(D_p H(x, Du) m) = 0 & \text{in } \mathbb{R}^d. \end{cases} \quad (5)$$

We are interested in the ergodic MFG because one natural guess is that the solutions of the MFG system tend to stabilize and to converge to the solutions of the ergodic one. Even though under specific structural assumptions (namely monotonicity) this convergence holds, we will prove that it is not in general the case.

In the past decade the literature on MFG thrived and this theory moved towards different directions. The existence of solutions of the MFG system under various assumptions is studied in [20, 21, 24, 16, 55]. For MFG as limit of N-player differential games one can look at [19]. Carmona and Delarue [30, 31] focused on the probabilistic approach. For the numerical analysis of MFG models we refer to [3, 2, 14, 12, 29]. The theory has been fruitful also in terms of applications: a couple of examples are [56, 25, 1] in economics and [37] in engineering. Moreover, new classes of MFG have appeared in the last few years as, for instance, the obstacle problem [8] and the MFG of control [52].

Keep in mind that the models that can be analyzed through the MFG system were the first to be introduced but they do not represent the most general case. This means that the MFG system can be derived only under specific assumptions. A most important one is that the noises that the players face are all independent. Therefore, nothing like common noises can be taken into account through MFG systems like (3). To look at MFG models with common noise one has either to introduce stochastic MFG as in [58, 32] or to work at the level of the *Master Equation*. For this latter approach we refer to [19] and [30].

### 0.1.2 Potential MFG

In this thesis we will focus on a special class of MFG that goes under the name of *potential MFG* (or variational MFG). These models, which were initially introduced by Lasry and Lions in [62], are those MFG whose associated MFG system can be derived as an optimality condition of a particular minimization problem. If the coupling function  $F$  and the terminal cost  $G$  admit potentials, which we respectively denote with  $\mathcal{F}$  and  $\mathcal{G}$ , then, for any  $t < T < +\infty$ , one can define the following control problem

$$\mathcal{U}^T(t, m_0) = \inf_{(m, w)} J(m, w) \quad (6)$$

where

$$J(m, w) = \int_t^T \int_{\mathbb{R}^d} H^* \left( x, -\frac{dw(s)}{dm(s)}(x) \right) dm(s) + \mathcal{F}(m(s)) ds + \mathcal{G}(m(T)), \quad (7)$$

$(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m + \operatorname{div} w = 0$  with  $m(t) = m_0$  and  $H^*$  is the Fenchel conjugate of  $H$  with respect to the second variable.

If  $(\bar{m}, \bar{w})$  minimizes  $\mathcal{U}^T(t, m_0)$  then, formally, one has that  $\bar{w} = -\bar{m} D_p H(x, D\bar{u})$  where  $(\bar{u}, \bar{m})$  solves

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

The variational nature of potential MFG endows these models with a powerful structure that allows to go deeper in the analysis of the solutions of the MFG system. See for instance [21, 70, 20] for existence results or [63, 72, 27] for regularity results that rely on this variational structure.

Potential MFG are generally divided in two macro groups depending on the type of couplings  $F$  and  $G$ , which can be either local or non-local. The coupling is said to be local when the functions  $F$  and  $G$  depend pointwisely on the density of the distribution of players  $m$ . Namely,  $F(m) = f(m(x))$  for a function  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ . In this case the potential  $\mathcal{F}$  is defined as

$$\mathcal{F}(m) = \int_{\mathbb{R}^d} \int_0^{m(x)} f(s) ds dx$$

and analogously for  $G$ . There are several papers that analyze this kind of potential MFG, see for instance [21, 27, 54, 70, 7]. In this setting one generally looks for weak solutions in Sobolev spaces.

In this thesis we focus on potential MFG with regularizing non-local couplings. In this case the functions  $F$  and  $G$  are defined directly on the space of Borel probability measures  $\mathcal{P}(\mathbb{R}^d)$  and we say that we are in the potential setting if there exist two functions  $\mathcal{F}$  and  $\mathcal{G}$  such that

$$\frac{\delta \mathcal{F}}{\delta m}(m, x) = F(m, x) \quad \text{and} \quad \frac{\delta \mathcal{G}}{\delta m}(m, x) = G(m, x).$$

The notion of derivative is the following.

**Definition 0.1.1.** *We say that  $\Phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$  is  $C^1$  if there exists a continuous function  $\frac{\delta \Phi}{\delta m} : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that*

$$\Phi(m_1) - \Phi(m_2) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta \Phi}{\delta m}((1-t)m_2 + tm_1, x)(m_1 - m_2)(dx) dt, \quad \forall m_1, m_2 \in \mathcal{P}(\mathbb{R}^d).$$

This definition is taken from [19] where one can find further properties.

Provided that  $\mathcal{F}$  and  $\mathcal{G}$  are sufficiently smooth, the solutions of the associated MFG system are smooth as well thanks to standard estimates for solutions of uniform parabolic equations.

The relation between (3) and the minimizers of (7) is explained by Proposition 0.1.2 for which we first state some hypothesis

- The functions  $\mathcal{F}$  and  $\mathcal{G}$  on  $\mathcal{P}(\mathbb{R}^d)$  belong to  $C^1(\mathcal{P}(\mathbb{R}^d))$ . Moreover, their derivatives  $F, G$  on  $\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d$  are smooth in the second variable.
- The Hamiltonian  $H$  is smooth and uniformly convex, in the sense that there exists  $C > 0$  such that

$$C^{-1}I_d \leq D_{pp}H \leq CI_d.$$

**Proposition 0.1.2.** *Let assume that the above hypotheses on the coupling functions and on the Hamiltonian hold true. If  $m_0 \in \mathcal{P}(\mathbb{R}^d)$  and  $(\bar{m}, \bar{w})$  is a minimizer of*

$$\inf_{(m,w)} J(m, w) = \inf_{(m,w)} \int_t^T \int_{\mathbb{R}^d} H^* \left( x, -\frac{dw(s)}{dm(s)}(x) \right) dm(s) + \mathcal{F}(m(s))ds + \mathcal{G}(m(T)),$$

where the infimum is taken over the set of couples  $(m, w)$  which solves in  $[t, T] \times \mathbb{R}^d$  the Fokker Plank equation  $-\partial_t m + \Delta m + \operatorname{div} w = 0$  with  $m(t) = m_0$ , then,  $\bar{m} \in C^{1,2}(\mathbb{R}^d \times (t, T])$  and there exists  $\bar{u} \in C^{1,2}(\mathbb{R}^d \times [t, T])$  such that  $(\bar{u}, \bar{m})$  is a classical solution of

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

*Proof.* When the domain is  $d$ -dimensional flat torus  $\mathbb{T}^d$  instead of the whole space  $\mathbb{R}^d$ , the proof, which relies on Fenchel-Rockafellar duality theorem, can be found in [11, Proposition 3.1] (this proof includes also the one for the existence of a minimizer  $(\bar{m}, \bar{w})$ ).  $\square$

Similarly to the time dependent MFG system, the ergodic MFG system also admits a variational formulation. In this case the minimization reads

$$\inf_{(m,w)} \int_{\mathbb{R}^d} H^* \left( x, -\frac{dw}{dm}(x) \right) dm + \mathcal{F}(m), \quad (8)$$

where  $(m, w)$  solves  $\Delta m - \operatorname{div} w = 0$  in the sense of distributions.

One natural issue is the uniqueness of solutions of system (3) and system (5). It has been clear since the very first papers that the natural assumption to be imposed to have the uniqueness is the so called Lasry-Lions monotonicity assumption (introduced in [61, 62]).

**Definition 0.1.3.** *We say that a function  $F : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}$  verifies the Lasry-Lions monotonicity assumption if, for any  $m_1, m_2 \in \mathcal{P}(\mathbb{R}^d)$*

$$\int_{\mathbb{R}^d} (F(m_1, x) - F(m_2, x)) d(m_1(x) - m_2(x)) \geq 0.$$

In the potential setting, if  $F$  and  $G$  verify the monotonicity assumption then the minimization problem (7) is convex and admits a unique minimizer which in turn implies that the solution of the MFG system is unique as well. Note that the uniqueness result under monotonicity assumption is true even outside the potential setting. The proof of the following proposition does not require  $F$  and  $G$  to have a potential. The main argument of the proof, which we skip, can be found already in the original works by Lasry and Lions [61, 62].

**Proposition 0.1.4.** *Let  $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  be a uniformly convex and smooth Hamiltonian. If  $F$  and  $G$  are continuous and verify the monotonicity assumption then there exists at most one solution of (3).*

## 0.2 The long time behavior and the weak KAM theory

### 0.2.1 The long time behavior

A standard question, that arises from the dynamical nature of MFG, is how players behave when their time horizon converges to infinity. This problem was initially considered in [64] and then discussed in several different contexts in [23, 18, 22, 26, 50]. The forward-backward structure of the MFG system makes it a tricky question. Indeed, while from the point of the model the meaning of the long time behavior of a mass of players is clear, when one looks at the MFG system this is no longer the case. The system is made of a Fokker-Plank equation that goes forward in time and a Hamilton-Jacobi that goes backward, hence, one has to carefully define the notion of convergence. Moreover, a crucial role is played by the monotonicity assumption (Definition 0.1.3). When this assumption is in place the system enjoys some sort of convexity which gives at the same time uniqueness of solutions and strong uniform estimates from energy equalities (we refer to [23]).

We start summarizing two of the most important results regarding the long time behavior of second order MFG when the monotonicity assumption holds.

A first possibility is to look at the average convergence. In [23], the authors defined

$$U^T(t, x) = u^T(tT, x) \quad \text{and} \quad M^T(t) = m^T(tT),$$

where  $(u^T, m^T)$  is a solution of

$$\begin{cases} -\partial_t u^T - \Delta u^T + \frac{1}{2}|Du^T|^2 = F(m^T, x) & \text{in } [0, T] \times \mathbb{T}^d \\ -\partial_t m^T + \Delta m^T + \operatorname{div}(m^T Du^T) = 0 & \text{in } [0, T] \times \mathbb{T}^d \\ m^T(0) = m_0, u^T(T) = u_f & \text{in } \mathbb{T}^d. \end{cases}$$

They proved that, under the monotonicity assumption, if  $(\bar{\lambda}, \bar{u}, \bar{m})$  is the unique solution of the ergodic MFG system (5) (with  $H(x, \cdot) = |\cdot|^2/2$ ), then

$$\|m^T(t) - \bar{m}\|_{C^{2,\alpha}(\mathbb{T}^d)} + \|Du^T(t) - D\bar{u}\|_{C^{2,\alpha}(\mathbb{T}^d)} \leq C \left( e^{-k(t-T)} + e^{-kt} \right) \quad (9)$$

for certain  $C, k > 0$  and  $\alpha \in (0, 1)$ . Moreover,  $U^T$  converges locally uniformly to the function  $(t, x) \mapsto -\bar{\lambda}t$  and  $M^T$  converges in  $L^1((0, 1) \times \mathbb{T}^d)$  to  $\bar{m}$ .

The rate of convergence (9), which is sometimes called *turnpike* property, shows that the convergence can be expected only far from the initial and terminal condition. Note also that at this point the convergence of the value function holds either on average or for its gradient.

Afterwards, in [28] the result was pushed further. In this paper the authors managed to understand how the function  $u^T(0, \cdot) - \bar{\lambda}T$  behaves when  $T \rightarrow +\infty$ . The solution was found at the level of master equation. If one define the function  $U : (-\infty, 0] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  by  $U(-T, x, m_0) = u^T(0, x)$  where  $(u^T, m^T)$  is the solution of the MFG system (3) with initial condition  $m(0) = m_0$  then,  $U$  solves the following equation set in  $(-\infty, 0] \times \mathcal{P}(\mathbb{T}^d)$

$$\begin{cases} -\partial_t U - \Delta_x U + H(x, D_x U(t, x, m)) - \int_{\mathbb{T}^d} \operatorname{div}_y D_m U(t, x, m, y) m(dy) \\ \quad + \int_{\mathbb{T}^d} D_m U(t, x, m, y) \cdot D_p H(y, D_x U(t, y, m)) m(dy) = F(x, m) \\ U(0, x, m) = G(x, m). \end{cases} \quad (10)$$

Even though we will not discuss the master equation, which goes far beyond the scope of this introduction, it is worthwhile to mention that the above equation has a "standard" backward structure. Therefore, we have a much more natural notion of long time convergence. What is proved in [28] is that  $U(-T, \cdot, \cdot) + \bar{\lambda}T$  converges to a function  $\chi : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  which solves the *ergodic master equation*

$$\begin{cases} \bar{\lambda} - \Delta_x \chi + H(x, D_x \chi(x, m)) - \int_{\mathbb{T}^d} \operatorname{div}_y D_m \chi(x, m, y) m(dy) \\ \quad + \int_{\mathbb{T}^d} D_m \chi(x, m, y) \cdot D_p H(y, D_x \chi(y, m)) m(dy) = F(x, m). \end{cases}$$

Consequently, one has that  $u^T(0, \cdot) - T\bar{\lambda}$  converges to  $\chi(\cdot, m_0)$ . A most important aspect is that, while  $m^T$  converges always to the invariant measure  $\bar{m}$ , one cannot expect that  $u^T(0, \cdot) - T\bar{\lambda}$  converges to  $\bar{u}$  (where  $(\bar{u}, \bar{m})$  is the solution of the ergodic MFG system (4)). Indeed, the function  $\chi$  keeps track of the initial condition and in general, if  $m_0 \neq \bar{m}$ ,  $\chi(\cdot, m_0)$  is different from  $\bar{u}$ .

Outside the monotonicity assumption, the MFG system behaves quite differently and the existing results are much weaker than the ones that we presented for the monotone case. A first difficulty is the multiplicity of solutions of both the time-dependent MFG system and the ergodic one. It is clear from [28] that, even under the monotonicity assumption, working at the level of solutions of the MFG system might not be enough to understand their long time behavior. Indeed, only after lifting the analysis from the MFG system to the master equation the authors managed to characterize the limit of  $u^T(0, \cdot) - T\bar{\lambda}$ . On the other hand, the definition of  $U$ , solution of (10), is subordinated to the uniqueness of the solution of the MFG system which cannot be expected if monotonicity is not imposed.

Moreover, while in the monotone case the value  $\bar{\lambda}$  such that  $(\bar{\lambda}, \bar{u}, \bar{m})$  is a solution of the ergodic MFG system and the value  $\lambda$ , for which  $u^T(0, \cdot) - T\lambda$  has a limit, coincide, outside monotonicity we will prove that it is not always the case.

A further difference is the following one. In general, if we are in the potential setting we know that if  $(m, w)$  is a minimizer of  $\mathcal{U}^T(t, m_0)$  (defined in (6)) then  $w(s, x) = -m(s, x)D_p H(x, Du(s, x))$  where  $(u, m)$  is a solution of the MFG system (3). In the monotone case the previous implication goes also in the other direction, which means that if you have a solution  $(u, m)$  of (3) then  $(m, -mD_p H(x, Du))$  is a minimizer for  $\mathcal{U}^T(t, m_0)$  (see for instance [21] for this result in the local setting).

Aside from the ones of this thesis, the results on the long time behavior of non monotone MFG that I am aware of can be found in [49, 34, 33]. Before we comment these papers, we recall that a precursor result has already appeared in [56]. In this paper the authors constructed a periodic solution in a MFG models which mimics the Mexican wave in a stadium ("la ola"). Coming back to [49, 34, 33], we first point out that these papers, in contrast with [23] and [28], rather look for the existence of periodic solutions than study the long time behavior of the finite horizon model.

In [49], Gomes and Sedjero analyze a forward-forward non monotone MFG system with congestion. Even though it is a quite different framework with respect to the standard MFG model, they provided the first example of traveling wave in MFG. More related to the setting of this thesis are the results in [33, 34]. In [34], by the means of bifurcation method, Cirant proved



the existence of infinite branches that emanate from the constant trivial one, considering the MFG system

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = f(m) & \text{in } [0, T] \times \Omega \\ -\partial_t m + \Delta m + \operatorname{div}(mDu) = 0 & \text{in } [0, T] \times \Omega \end{cases} \quad (11)$$

with a non monotone local coupling  $f$ . In the finite horizon regime, this particular solutions exhibit an oscillatory behavior when they are close to the trivial constant one. Moreover, through numerical simulations the author suggested that this oscillatory behavior survives when the horizon goes to infinity.

Afterward, in [33], Cirant and Nurbekyan, using again a bifurcation method, managed to prove the existence of truly periodic solutions for a non monotone configuration of the system (11).

Note that, as we mentioned, the previous results are concerned with the existence of periodic solutions of the MFG system and they do not directly address the problem of the long time convergence. Moreover, even if the setting related to the periodic solution in [33] is of potential type, the periodic trajectory built through iterated bifurcations is no longer associated to the minimum of the correspondent minimization problem.

## 0.2.2 Weak KAM theory

We introduce the main ideas of the weak KAM theory that we borrowed for this thesis. As the weak KAM theory is not itself the main subject of this dissertation, we will only touch upon those results that have been useful for our purpose. For a full overview one can look at Fathi's book [42] or his seminal papers [39, 40, 41] where we took most of the results we discuss in this section. This brief introduction is also inspired by Evans' notes on the topic [38].

Hereafter,  $M$ ,  $TM$  and  $T^*M$  will be respectively a smooth compact manifold without boundaries, its tangent bundle and its cotangent bundle. The function  $H : T^*M \rightarrow \mathbb{R}$  is an Hamiltonian such that  $H$  is smooth, superlinear and strictly convex on the fibers, which means that

- $\lim_{|p| \rightarrow +\infty} \frac{H(x,p)}{|p|} = +\infty$ ,
- the matrix  $\partial_{pp}^2 H(x, \cdot)$  is positive definite.

We can associate to the Hamilton  $H$  the Lagrangian  $L : TM \rightarrow \mathbb{R}$  defined by

$$L(x, v) = \max_{p \in T^*M} p(v) - H(x, p), \quad \forall (x, v) \in TM,$$

which is smooth and strictly convex as well. We also introduce the Lax-Oleinik semigroup  $(T^{-t})_{t>0}$ . If  $u : M \rightarrow \mathbb{R}$  is a function and  $t > 0$ , then

$$T^{-t}u(x) = \inf_{\gamma} u(\gamma(0)) + \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds, \quad (12)$$

where the infimum is taken over all possible absolutely continuous curves  $\gamma : [0, t] \rightarrow M$  such that  $\gamma(t) = x$ . An important object in this framework are the extremal curves. Given two points  $x, y \in M$ , we say that an absolutely curve  $\bar{\gamma} : [0, t] \rightarrow M$  is an extremal curve if  $\bar{\gamma}$  minimizes

$$\inf_{\gamma} \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds$$

among all possible absolutely continuous curves  $\gamma$  such that  $\gamma(0) = x$  and  $\gamma(t) = y$ . It is easy to prove that if  $\bar{\gamma}$  is a minimizer of (12) then  $\bar{\gamma}$  is an extremal curve for the endpoints  $x$  and  $\bar{\gamma}(0)$ .

An other fundamental notion in the weak KAM theory is the one of domination. We say that  $u$  is dominated by  $(L, \alpha)$  with  $\alpha \in \mathbb{R}$ , if for any  $t > 0$  and any absolutely continuous curve  $\gamma$  on  $M$  we have

$$u(\gamma(t)) - u(\gamma(0)) \leq \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds + \alpha t.$$

We define the *Mañé's critical value*  $\alpha[0]$  as the infimum over all  $\alpha \in \mathbb{R}$  such that there exists a function  $u : M \rightarrow \mathbb{R}$  which is dominated by  $(L, \alpha)$ .

A first major result of the weak KAM theory is the following theorem.

**Theorem 0.2.1.** *There exists a function  $u^- : M \rightarrow \mathbb{R}$  such that, for any  $t > 0$ ,*

$$u^- = T^{-t}u^- + \alpha[0]t. \quad (13)$$

Moreover, if for a given  $\alpha \in \mathbb{R}$  there exists  $u : M \rightarrow \mathbb{R}$  such that, for any  $t > 0$ ,

$$u = T^{-t}u + \alpha t,$$

then  $\alpha = \alpha[0]$ .

If  $u$  verifies (13) then we call it *weak KAM solution* of the critical equation

$$H(x, d_x u) = \alpha[0]. \quad (14)$$

The connection between the critical equation (14) and the fixed points of the Lax-Oleinik semi-group relies in the following property.

**Proposition 0.2.2.** *If  $u : M \rightarrow \mathbb{R}$  is dominated by  $(L, \alpha)$  then  $u$  is a viscosity subsolution of*

$$H(x, d_x u) = \alpha.$$

Moreover, if  $u$  is weak KAM solution in the sense of Theorem 0.2.1 then  $u$  is a viscosity solution of

$$H(x, d_x u) = \alpha[0].$$

The characterization of the Mañé's critical value  $\alpha[0]$  of Theorem 0.2.1 is not the only possible and for the purpose of this thesis it is useful to introduce two other points of view.

For any  $(x, v) \in TM$  we define the Euler-Lagrangian flow  $\phi_t(x, v)$  as follows. Let  $\mathbf{x}$  be the solution of

$$\begin{cases} -\frac{d}{dt}(D_v L(\mathbf{x}, \dot{\mathbf{x}})) + D_x L(\mathbf{x}, \dot{\mathbf{x}}) = 0 \\ \mathbf{x}(0) = x, \dot{\mathbf{x}}(0) = v. \end{cases} \quad (15)$$

We set  $\phi_t(x, v) = (\mathbf{x}(t), \dot{\mathbf{x}}(t))$ . We say that a measure  $\mu$  on the tangent bundle  $TM$  is invariant under the Euler-Lagrangian flow if, for any continuous and bounded function  $\Psi$  on  $TM$ , we have

$$\int_{TM} \Psi(x, v) d\mu(x, v) = \int_{TM} \Psi(\phi_t(x, v)) d\mu(x, v), \quad \forall t > 0.$$

Mather, in his seminal paper [69], proved the following result.

**Theorem 0.2.3.** *The Mañé critical value  $\alpha[0]$  verifies the following relation*

$$-\alpha[0] = \inf_{\mu} \int_{TM} L(x, v) d\mu(x, v), \quad (16)$$

where the infimum is taken over the set of invariant probability measures  $\mu$ .

If  $\mu$  minimizes (16), we say that  $\mu$  is a Mather measure. We define the *Mather set*  $\mathcal{M}$  as follows

$$\mathcal{M} = \overline{\bigcup_{\mu} \text{supp } \mu},$$

the union over all Mather measure  $\mu$ . We also define the projected Mather set  $\tilde{\mathcal{M}}$  as the projection of  $\mathcal{M}$  on  $M$ .

One goal of the weak KAM theory is to understand the relations between the Mather set  $\mathcal{M}$  and the weak KAM solutions (which are the fixed points of the Lax-Oleinik semigroup). Two important examples are the following propositions (which can be both found in [42]).

**Proposition 0.2.4.** *If  $u$  is a weak KAM solution, then  $u$  is derivable at any point  $x \in \tilde{\mathcal{M}}$ .*

**Proposition 0.2.5.** *Let  $V$  be an open neighborhood of  $\mathcal{M}$  in  $TM$ . There exists a time  $t(V)$  such that, if  $t \geq t(V)$  and  $\gamma : [0, t] \rightarrow M$  is an extremal curve, then there exists  $s \in [0, t]$  such that  $(\gamma(s), \dot{\gamma}(s)) \in V$ .*

The preliminary result to prove Proposition 0.2.5 is that if  $\gamma$  is an extremal curve then  $\gamma$  is a solution of the Euler-Lagrange equation (15). Using this basic fact one can build a Mather measure  $\mu$  starting from the optimal curves of  $T^{-t}u(x)$  as follows. Let  $\gamma^t : [0, t] \rightarrow M$  be an optimal trajectory for  $T^{-t}u(x)$  and  $\mu^t$  be the probability measure over  $TM$  defined by

$$\int_{TM} f(x, v) d\mu^t(x, v) = \frac{1}{t} \int_0^t f(\gamma^t(s), \dot{\gamma}^t(s)) ds \quad (17)$$

for any continuous bounded function  $f$  on  $TM$ . Then, if  $\mu$  is a weak limit of  $\mu^t$  for  $t \rightarrow +\infty$ , the probability measure  $\mu$  is a Mather measure.

The above results are some of the key points to prove the following theorem which is the reason why we decided to look at the weak KAM theory to try to understand the long time behavior of potential MFG.

**Theorem 0.2.6.** *For any  $u \in C(M)$  the limit  $t \rightarrow +\infty$  of  $T_t^- u + t\alpha[0]$  exists. Moreover, this limit is a weak KAM solution.*

We continue this brief account on the weak KAM theory saying a few words on a close topic which is the Lions–Papanicolaou–Varadhan theory (see [66]).

Fathi’s proof of Theorem 0.2.1 relies on a abstract result regarding the existence of a common fixed point for certain classes of non-expansive maps on Banach spaces (see [39]). This approach does not fit well in the context of MFG and here the Lions–Papanicolaou–Varadhan approach helps to overcome this issue.

In this part we suppose that the manifold  $M$  is the  $d$ -dimensional flat torus  $\mathbb{T}^d$ . Combining Theorem 0.2.1 with Proposition 0.2.2 one can define  $\alpha[0]$  as the unique value for which there exists a viscosity solution of

$$H(x, Du) = \alpha[0] \quad \forall x \in \mathbb{T}^d. \quad (18)$$

This theory is an alternative and purely PDE-based technique to prove the existence of a solution  $u$  of the critical equation (18) and the critical value  $\alpha[0]$ . Even though it is not essential for this thesis, it is worthwhile to mention that the Lions–Papanicolaou–Varadhan theory goes in the direction of Hamilton–Jacobi homogenization and one generally considers (18) adding a new variable  $P \in \mathbb{R}^d$ . We define the function  $P \mapsto \alpha[P]$  as follows. For each  $P \in \mathbb{R}^d$ ,  $\alpha[P] \in \mathbb{R}$  is the only constant for which the following equation admits a viscosity solution

$$H(x, P + Du) = \alpha[P] \quad \forall x \in \mathbb{T}^d. \quad (19)$$

The function  $\alpha[\cdot]$  is called *effective Hamiltonian*.

The idea of Lions, Papanicolaou and Varadhan was to consider the discounted Hamilton–Jacobi equation

$$\delta u + H(x, P + Du) = 0 \quad \forall x \in \mathbb{T}^d. \quad (20)$$

It is standard from viscosity solutions theory that the above equation enjoys the comparison principle and therefore, for any  $\delta > 0$ , the solution  $u_\delta$  of (20) is unique. As long as the Hamiltonian verifies the superlinearity and the convex hypothesis that we discussed at the beginning of this section, the solution  $u_\delta$  admits a variational representation:

$$u_\delta(x) = \inf_{\gamma} \int_0^{+\infty} e^{-\delta t} L_P(\gamma(s), \dot{\gamma}(s)) ds, \quad (21)$$

where the infimum is taken over all absolutely curves  $\gamma : [0, +\infty) \rightarrow \mathbb{T}^d$  such that  $\gamma(0) = x$  and  $L_P(x, \cdot)$  is the Lagrangian associated to the Hamiltonian  $H(x, P + \cdot)$ . Using the superlinearity and the convexity of  $H$  it is not hard to prove through the representation (21) and the equation (20) that there exists a constant  $C$ , independent of  $\delta$ , such that

$$|Du_\delta| + |\delta u_\delta| \leq C.$$

The direct consequence is that, up to subsequence,  $u_\delta - u_\delta(\bar{x}) \rightarrow u$  and  $\delta u_\delta \rightarrow \alpha[P]$  where  $u$  and  $\alpha[P]$  solves in viscosity sense (19). A standard application of the comparison principle verified by  $u_\delta$  ensures the uniqueness (and therefore the full convergence of  $\delta u_\delta$  to)  $\alpha[P]$ .

It was only in 2016 that Davini, Fathi, Iturriaga, Zavidovique [35] proved the full convergence of  $u_\delta$  to a solution  $u_0$  of the critical equation. We conclude this introduction on the weak KAM theory commenting this last result.

Let  $M$  be again any smooth compact manifold without boundaries. We define  $\mathcal{F}^-$  as the set of viscosity subsolutions  $u : M \rightarrow \mathbb{R}$  of the critical equation

$$H(x, Du) = \alpha[0] \quad \forall x \in M, \quad (22)$$

such that, for any Mather measure  $\mu$  (see (16)),

$$\int_{TM} u(x) d\mu(x, v) \leq 0. \quad (23)$$

The main result of [35] is the following.

**Theorem 0.2.7.** *Let  $u_\delta$  be the solution of the discounted Hamilton–Jacobi equation*

$$\delta u(x) + H(x, Du(x)) = \alpha[0] \quad \forall x \in M, \quad (24)$$

*then  $u_\delta$  uniformly converges to a viscosity solution  $u_0$  of the critical equation (22) which is characterized by*

$$u_0(x) = \sup_{u \in \mathcal{F}^-} u(x).$$

Note that in the right hand side of the discounted equation (24) appears the Mañé critical value  $\alpha[0]$  while previously we had zero. The reason of this translation is that  $\alpha[0]$  is the only constant such that the solutions  $u_\delta$  of (24) are uniformly bounded with respect to  $\delta$ .

Let  $u_0$  be a uniform limit, up to subsequence, of  $u_\delta$ . To prove that  $u_0(x) \leq \sup_{u \in \mathcal{F}^-} u(x)$  the idea is to show that  $u_0 \in \mathcal{F}^-$ . This comes directly from the discounted equation (24). Indeed, using (24) one can prove that, for any Mather measure  $\mu$  and any  $\delta > 0$ ,

$$\delta \int_{TM} u_\delta(x) d\mu(x, v) \leq 0.$$

Letting  $\delta \rightarrow 0$  proves that  $u_0 \in \mathcal{F}^-$  and so that  $u_0(x) \leq \sup_{u \in \mathcal{F}^-} u(x)$ .

If we consider a minimizer  $\gamma_\delta^x : [0, +\infty) \rightarrow M$  of the variational representation (21) of  $u_\delta(x)$ , we can associate a measure  $\mu_\delta^x$  on  $TM$  as follows

$$\int_{TM} f(x, v) d\mu_\delta^x(x, v) = \delta \int_0^{+\infty} e^{-\delta s} f(\gamma_\delta^x(s), \dot{\gamma}_\delta^x(s)) ds$$

for any continuous bounded function  $f$  on  $TM$ . The key ingredient for the opposite inequality is that any weak limit  $\mu^x$  of  $\mu_\delta^x$  is a Mather measure. Then, comparing the discounted equation (24) with the critical equation (22), one can prove that, for any  $w \in \mathcal{F}^-$  and any  $\delta > 0$ ,

$$u_\delta(x) \geq w(x) - \int_{TM} w(y) d\mu_\delta^x(y, v) \quad \forall x \in M.$$

Passing to the limit  $\delta \rightarrow 0$  (possibly up to subsequence), we get

$$u_0(x) \geq w(x) - \int_{TM} w(y) d\mu^x(y, v) \quad \forall x \in M.$$

As  $\mu^x$  is a Mather measure and  $w \in \mathcal{F}^-$ , if we plug (23) into the above inequality we find that

$$u_0(x) \geq w(x) \quad \forall x \in M,$$

which proves that  $u_0(x) \geq \sup_{w \in \mathcal{F}^-} w(x)$  and so the result.

We conclude mentioning that in [35] the authors give a second characterization of  $u_0$  through the Peierls barrier but we will not discuss it.

### 0.3 Contributions

In this thesis we address the problem of the long time behavior of non local potential MFG by the means of weak KAM theory. The link between potential MFG and weak KAM theory is possible thanks to the fact that MFG can be seen as infinite dimensional Hamiltonian systems. In particular, if  $(u, m) \in C^2(\mathbb{T}^d) \times \mathcal{P}(\mathbb{T}^d)$  and we define the quantity

$$c(u, m) = \int_{\mathbb{T}^d} \left( H(x, Du(x)) - \Delta u(x) \right) m(dx) - \mathcal{F}(m(t)),$$

then, if  $(u, m)$  is a solution of the MFG system,  $c(u(t), m(t))$  is constant in time. The aim this thesis is to provide a set of tools and results, which do not depend on the monotonicity assumption (Definition 0.1.3), that might be useful to tackle the problem of the long time convergence in a more systematic way. Our results are collected in the next three chapters.

In Chapter 1 we look at the time dependent minimization problem  $\mathcal{U}^T$  (defined in (6)) and we analyze the relationship between the limit behavior of  $\mathcal{U}^T(0, \cdot)/T$  when  $T \rightarrow +\infty$  and the value

of the stationary minimization problem (8). We first show that, as the time horizon goes to  $+\infty$ , the value of the time dependent optimal control problem converges to a limit  $-\lambda$  independent from the initial condition. Then, if we denote with  $-\bar{\lambda}$  the value of the stationary one, in general we have that  $\lambda \geq \bar{\lambda}$ . Moreover, we provide a class of explicit examples where the strict inequality  $\lambda > \bar{\lambda}$  holds true. This will imply that in those cases the trajectories minimizing the energy of the time-dependent MFG system do not converge to static equilibria.

In Chapter 2 we develop the counterpart for potential mean field games of the convergence result of the Lax-Oleinik semigroup proved in [41]. As a consequence, we have the convergence of the solutions of the time dependent MFG system associated to optimal trajectories of  $\mathcal{U}^T$ . The main result is that  $\mathcal{U}^T(0, \cdot) - \lambda T$  uniformly converges to a corrector function (which is a fixed point of the Lax-Oleinik semigroup). In addition, we show a mean field limit for the ergodic constant associated with the corresponding Hamilton-Jacobi equation.

In Chapter 3 we consider the solution  $\mathcal{V}_\delta$  of the discounted Hamilton-Jacobi equation in the Wasserstein space arising from potential MFG and we prove its full convergence to a corrector function  $\chi_0$ . We follow the structure of the proof of the analogous result in the finite dimensional setting provided by Davini, Fathi, Iturriaga, Zavidovique in 2017. We characterize the limit  $\chi_0$  through a particular set of smooth Mather measures. A major point, that distinguishes the techniques deployed in the standard setting from the ones that we use here, is the lack of mollification in the Wasserstein space.

In the last part of this introduction we discuss in more details the results that we sketched above. In particular we will focus on differences and the difficulties that occur when one transposes the standard finite dimensional weak KAM theory into the infinite dimensional framework of potential MFG.

### 0.3.1 On the long time convergence of potential MFG

One of the main goal of Chapter 1 is to recover, in the context of potential MFG, the results of Theorem 0.2.1 on the characterization of the Mañé's critical value and the existence of fixed points of the Lax-Oleinik semigroup.

Let us first see how in practice the MFG minimization problem  $\mathcal{U}^t$  is structurally similar to the standard Lax-Oleinik semigroup  $T^{-t}$  that we discussed in the previous section. For any function  $\Phi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ , we can define the semigroup  $\{\tau_h\}_{h>0}$  as follows

$$\tau_h \Phi(m_0) = \inf_{(m,w)} \left\{ \int_0^h \left( \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(s)}{dm(s)}(x) \right) dm(s) + \mathcal{F}(m(s)) ds \right) ds + \Phi(m(h)) \right\},$$

where  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m + \operatorname{div} w = 0$  with  $m(0) = m_0$ . Therefore, if the MFG minimization problem  $\mathcal{U}^t(0, m_0)$  is defined by

$$\mathcal{U}^t(0, m_0) = \inf_{(m,w)} \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(s)}{dm(s)}(x) \right) dm(s) + \mathcal{F}(m(s)) ds + \mathcal{G}(m(t)),$$

where  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m + \operatorname{div} w = 0$  with  $m(0) = m_0$ , then  $\mathcal{U}^t(0, m_0) = \tau_t \mathcal{G}(m_0)$ . According to these definitions, we can analogously introduce a critical value  $\lambda \in \mathbb{R}$  and the notion of fixed point of the semigroup  $\tau$ . Namely, we say that  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is a fixed point of the semigroup  $\tau$  if, for any  $t > 0$  and any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,

$$\chi(m_0) = \inf_{(m,w)} \left( \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{w(s)}{m(s)} \right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t. \quad (25)$$

where  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m + \operatorname{div} w = 0$  with  $m(0) = m_0$ . We call *corrector function* any function  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  which verifies the above dynamic programming principle.

Large part of Chapter 1 is devoted to the proof of the existence and uniqueness of such a critical value  $\lambda \in \mathbb{R}$ . Afterward, we compare this critical value  $\lambda$  with the ergodic one  $\bar{\lambda}$  (which is part of the solution of the ergodic MFG system (5)) to derive some information on the long time behavior  $\mathcal{U}^t$ .

As we mentioned in Section 0.2.2, the structure of the standard proof of Theorem 0.2.1 relies on arguments that can hardly be transposed in the MFG setting. Therefore, as in the Lions–Papanicolaou–Varadhan theory we introduce the discounted infinite horizon problem  $\mathcal{V}_\delta$ . For any  $\delta > 0$  and any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ , we define

$$\mathcal{V}_\delta(m_0) = \inf_{(m, w)} \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt,$$

where  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m - \operatorname{div} w = 0$  with  $m(0) = m_0$ . If  $(\bar{m}, \bar{w})$  is a minimizer of  $\mathcal{V}_\delta(m_0)$  then  $\bar{w}(t, x) = -D_p H(x, D\bar{u}(t, x))$  where  $(\bar{m}, \bar{w})$  solves

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, +\infty) \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, +\infty) \\ m(t) = m_0, u \in L^\infty([0, +\infty) \times \mathbb{T}^d). \end{cases}$$

When  $\delta > 0$ , the Hamilton–Jacobi equation in the above system enjoys the comparison principle that we use to prove that  $\bar{u}$  is bounded in  $C^{1,2}([0, +\infty) \times \mathbb{T}^d)$  uniformly with respect to  $\delta$ . Those estimates, which imply through the drift  $D_p H(x, D\bar{u})$  uniform regularity on  $\bar{m}$  as well, are the key point to prove that the family of function  $\{\mathcal{V}_\delta(\cdot)\}_\delta$  is uniformly Lipschitz continuous. At this point it is straightforward that, up to subsequence,  $\delta \mathcal{V}_\delta \rightarrow -\lambda$  and that, for any fixed measure  $\bar{m}_0$ ,  $\mathcal{V}_\delta(\cdot) - \mathcal{V}_\delta(m_0) \rightarrow \chi(\cdot)$ , where  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is a Lipschitz function which verifies the dynamic programming principle (25).

Once that we have these convergence, the next step is the following theorem.

**Theorem 0.3.1.** *The limit value  $-\lambda$  is uniquely defined and  $\delta \mathcal{V}_\delta(\cdot) \rightarrow -\lambda$  depends neither on a subsequence nor on the initial condition. Moreover,  $\frac{1}{T} \mathcal{U}^T(0, \cdot)$  uniformly converges to  $-\lambda$  when  $T$  goes to  $+\infty$ .*

In the second part of the chapter we focus on the relation between the critical value  $\lambda$  and the ergodic value  $\bar{\lambda}$ . We recall that  $\bar{\lambda}$  is defined by

$$-\bar{\lambda} = \inf_{(m, w)} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm + \mathcal{F}(m). \quad (26)$$

It is always true that  $\lambda \geq \bar{\lambda}$  and, moreover, one can easily prove that, under monotonicity assumption,  $\lambda = \bar{\lambda}$ . Indeed, if monotonicity is in place, the minimization problem that defines  $\mathcal{U}^T$  is convex with respect to  $(m, w)$ . Therefore, one just need to apply the Jensen’s inequality to find that, at the limit, stationary configurations are more efficient than the other ones. Note that, in this case, we recover part of the result of [23]. More interestingly, we build a class of explicit examples where  $\lambda > \bar{\lambda}$ .

To understand why the relationship between  $\lambda$  and  $\bar{\lambda}$  is important to analyze the long time behavior of potential MFG, we need to introduce the notion of *projected Mather*  $\mathcal{M}$ . The projected Mather set  $\mathcal{M}$  is the set of probability measures contained in a calibrated curve. We

say that  $(\bar{m}, \bar{w})$  is a calibrated curve associated to a corrector function  $\chi$  if it is defined for any time  $t \in \mathbb{R}$  and if, for any  $t_1, t_2 \in \mathbb{R}$ ,  $(\bar{m}, \bar{w})$  is optimal for the dynamic programming principle, i.e.

$$\chi(\bar{m}(t_1)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{\bar{w}(s)}{\bar{m}(s)} \right) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + \chi(\bar{m}(t_2)) + \lambda(t_2 - t_1).$$

When  $\lambda > \bar{\lambda}$ ,  $\mathcal{M}$  cannot contain any stationary curve. Moreover, as this set is compact, calibrated curves can not even approach a static configuration. Indeed, due to compactness, the stationary minimizers of (26) lay at positive distance to the projected Mather set. While in this chapter we only prove that  $\mathcal{M}$  contains the limit points of the optimal trajectories of  $\mathcal{V}_\delta$ , in the next one we get the same result for the optimal trajectories of  $\mathcal{U}^T$ .

### 0.3.2 Weak KAM theory for potential MFG

This chapter is devoted to the proof of the convergence of  $\mathcal{U}^T$ . The main results are collected in the following two theorems.

**Theorem 0.3.2.** *When  $T \rightarrow +\infty$ ,  $\mathcal{U}^T(0, \cdot) + \lambda T$  uniformly converges to a corrector function  $\chi$ .*

**Theorem 0.3.3.** *If  $(m^T, w^T)$  is an optimal trajectory for  $\mathcal{U}^T(-T, m_0)$ , then  $(m^T, w^T)$  converges, up to subsequence and locally uniformly, to a calibrated curve  $(m, w)$ . Consequently,  $m(t) \in \mathcal{M}$  for any  $t \in \mathbb{R}$ .*

Even if Theorem 0.3.2 is the transposition, in the MFG setting, of Fathi's famous convergence result for Hamilton-Jacobi equations [41] (Theorem 0.2.6), its proof presents several additional difficulties. In particular, in the standard setting many proofs rely on approximation arguments. Namely, when one has a viscosity solution of the critical Hamilton-Jacobi equation then its mollification not only approximates the solution but it is itself an approximated solution of the equation. Such a tool does not exist in the context of Wasserstein spaces (even though it is still possible to approximate functions with smooth ones).

Anyway, the structure of the proof of Theorem 0.3.1 is still inspired by the one Theorem 0.2.6. One first important step is a further characterization of the critical value  $\lambda$ . Let us set

$$I := \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} (H(y, D_m \Phi(m, y)) - \mathcal{F}(m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy).$$

Then, by duality techniques, one can check the following equality

$$-I = \min_{(\mu, p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \left( H^* \left( y, \frac{dp_1}{dm \otimes \mu} \right) + \mathcal{F}(m) \right) m(dy) \mu(dm),$$

where the minimum is taken over  $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$  and  $p_1$ , Borel vector measure on  $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$ , such that  $p_1$  is absolutely continuous with respect to the measure  $dm \otimes \mu := m(dy) \mu(dm)$  and such that  $(\mu, p_1)$  is closed, in the sense that, for any  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ ,

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi(m, y) m(dy) \mu(dm) = 0.$$

We call Mather measure any couple  $(\mu, p_1)$  which minimizes the dual problem. This value  $I$  does actually coincide with the critical value  $\lambda$ , but to prove this equality is not as straightforward as in the finite dimensional setting. While it is not hard to show that  $I \geq \lambda$ , the opposite inequality is more subtle. One has to construct a smooth subsolution of the ergodic problem (27) without



using mollification arguments. To overcome this issue we pass through a finite particle system. Let  $(v^N, \lambda^N) \in C^2((\mathbb{T}^d)^N) \times \mathbb{R}$  be a solution of

$$-\sum_{i=1}^N \Delta_{x_i} v^N(\mathbf{x}) + \frac{1}{N} \sum_{i=1}^N H(x_i, N D_{x_i} v^N(\mathbf{x})) = \mathcal{F}(m_{\mathbf{x}}^N) + \lambda^N$$

and we define the smooth function  $W^N$  on  $\mathcal{P}(\mathbb{T}^d)$  by

$$W^N(m) := \int_{(\mathbb{T}^d)^N} v^N(x_1, \dots, x_N) \prod_{i=1}^N m(dx_i).$$

Then,  $W^N$  verifies

$$-\int_{\mathbb{T}^d} \operatorname{div}_y D_m W^N(m, y) m(dy) + \int_{\mathbb{T}^d} H(y, D_m W^N(m, y), m) m(dy) \leq \mathcal{F}(m) + \lambda^N + o(N),$$

which proves that  $I \leq \liminf \lambda^N$ . Using Bernstein type estimates ([67] for the original argument), we get

$$N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 \leq \bar{C}_1 \quad \forall \mathbf{x} \in (\mathbb{T}^d)^N.$$

As a consequence there exists a subsequence  $\{N_k\}_k$  such that  $v^{N_k}$  uniformly converges to a Lipschitz function  $V$  on  $\mathcal{P}(\mathbb{T}^d)$ . Adapting the results of [59], which is concerned with the connection between optimal control of McKean-Vlasov dynamics and the limit behavior of large number of interacting controlled state processes, we can prove that, for any accumulation point  $\lambda^*$  of  $\{\lambda^{N_k}\}_k$ ,

$$V(m_0) = \inf_{(m, \alpha)} \left( \int_0^T \int_{\mathbb{T}^d} H^* \left( x, \frac{dw}{dm}(s) \right) dm(s) + \mathcal{F}(m(s)) ds + V(m(T)) \right) + \lambda^* T,$$

where  $(m, w)$  verifies the usual constraint. As the above dynamic programming principle is verified for any  $T > 0$ ,  $\lambda^* = \lambda$  and, consequently,  $\lambda = \lim_N \lambda^N$ . We recall that we have that  $\lambda \leq I$  and  $I \leq \liminf \lambda^N$ . Then  $I = \lambda$  because  $\lambda = \lim_N \lambda^N$ .

The second difficulties that we faced is related to the connection between the points of the support of Mather measures and the critical equation

$$\int_{\mathbb{T}^d} (H(y, D_m \chi(m, y)) - \operatorname{div}_y D_m \chi(m, y)) m(dy) = \mathcal{F}(m) + \lambda \quad \text{in } \mathcal{P}(\mathbb{T}^d). \quad (27)$$

In the standard setting these points verifies the critical equation. In our framework, due to the divergence term in (27), we do not know if this is the case.

The idea is to focus only a specific class of Mather measures defining the notion of *smooth* Mather measures. These are Mather measures whose support points have regular densities (so that on these points the critical equation (27) is well defined).

As in the standard setting, to any sequence of minimizers  $(m^T, w^T)$  of  $\mathcal{U}^T$  we can associate a Mather measure (the construction is analogous to the in in (17)). Thanks to the regularity of minimal trajectories we have that this limit measure is indeed a smooth Mather measure. Using this regularity, one can prove that the support points of such a measure (weakly) verify the critical equation (27).

Once that this link between the critical equation and the limit behavior of  $\mathcal{U}^T$  is proven, the rest of the proof follows closely Fathi's argument in [41].

### 0.3.3 Convergence of the solutions of the MFG discounted Hamilton-Jacobi equation

In this final chapter we answer a question more connected to the weak KAM approach to potential MFG in itself than to MFG as a theory to model differential games. In Chapter 1 we proved the existence of a corrector  $\chi$  and we characterized the critical value  $\lambda$  through an homogenization argument. We defined

$$\mathcal{V}_\delta(m_0) = \inf_{(m,w)} \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt,$$

where  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m - \operatorname{div} w = 0$  with  $m(0) = m_0$  and we proved that  $\delta \mathcal{V}_\delta(\cdot)$  uniformly converges to  $-\lambda$  and that, for any fixed measure  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,  $\mathcal{V}_\delta(\cdot) - \mathcal{V}_\delta(m_0)$  uniformly converges, up to subsequence, to a corrector function  $\chi$ . As we have seen in Section 0.2.2, a natural question is whether the convergence of  $\mathcal{V}_\delta(\cdot) - \mathcal{V}_\delta(m_0)$  holds only up to subsequence or not.

In this chapter we prove, in the context of potential MFG, the analogous of Theorem 0.2.7. Even though we go through the same steps and the structure of the chapter is very similar to the one of [35], the proofs are quite different. In [35] Davini, Fathi, Iturriaga and Zavidovique work mostly at the level of viscosity solutions using often mollification arguments. In our case this is not possible and we have to work with optimal trajectories. Moreover, this lack of regularity brought to a slightly different characterization of the limit corrector  $\chi_0$ . Indeed, in [35] the set of subsolution  $\mathcal{F}^-$  is tested against any Mather measure (we refer to (23)), while in our case we consider only *smooth* Mather measures (note that this is not the same notion of *smooth* Mather measure of Chapter 2).

Before stating the main result we need to introduce few objects. We define  $\mathcal{S}^-$  the set of subsolutions  $\chi$  of

$$-\int_{\mathbb{T}^d} \operatorname{div}_y D_m \chi(m, y) m(dy) + \int_{\mathbb{T}^d} H(y, D_m \chi(m, y)) m(dy) - \mathcal{F}(m) = \lambda, \quad m \in \mathcal{P}(\mathbb{T}^d)$$

such that  $\int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu(dm)$  for any  $\nu \in \mathcal{M}_\nu$ . We say that a  $\nu \in \mathcal{M}_\nu$  if it is induced by an optimal trajectories for  $\mathcal{V}_\delta$ : if  $(m_\delta, \alpha_\delta)$  is an optimal trajectory for  $\mathcal{V}_\delta(m_0)$  we define  $\nu_\delta^{m_0} \in \mathcal{P}(\mathbb{T}^d \times C^1(\mathbb{T}^d, \mathbb{R}^d))$  as follows

$$\int_{\mathcal{P}(\mathbb{T}^d) \times C^1(\mathbb{T}^d, \mathbb{R}^d)} f(m, \alpha) \nu_\delta^{m_0}(dm, d\alpha) = \delta \int_0^\infty e^{-\delta s} f(m^\delta(s), \alpha^\delta(s)) ds.$$

If  $\nu^{m_0}$  is the limit of  $\nu_\delta^{m_0}$  and  $m_0$  has smooth density, then  $\nu^{m_0} \in \mathcal{M}_\nu$ . Such a limit  $\nu^{m_0}$  is a Mather measure. The main result of the chapter is the following theorem.

**Theorem 0.3.4.** *The function  $\mathcal{V}_\delta + \frac{\lambda}{\delta}$  uniformly converges to a corrector  $\chi_0$ , which is defined by*

$$\chi_0(m) = \sup_{\chi \in \mathcal{S}^-} \chi(m).$$



## Chapter 1

# On the long time convergence of potential MFG

We look at the long time behavior of potential Mean Field Games (briefly MFG) using some standard tools from weak KAM theory. Potential MFGs are those models where the MFG systems associated can be derived as optimality conditions of suitable optimal control problems on the Fokker-Plank equation. In particular we analyze the relationship between the limit behavior of the time dependent one, whose optimality condition corresponds with the finite horizon MFG system, and the stationary one, whose optimality condition is the ergodic MFG system. We first show that, as the time horizon goes to  $+\infty$ , the value of the time dependent optimal control problem converges to a limit  $-\lambda$ . Then, if we denote with  $-\bar{\lambda}$  the value of the stationary one, in general we have that  $\lambda \geq \bar{\lambda}$ . Moreover, we provide a class of explicit examples where the strict inequality  $\lambda > \bar{\lambda}$  holds true. This will imply that the trajectories of the time-dependent MFG system do not converge to static equilibria.

This chapter was published in *Nonlinear Differential Equations and Applications* (NoDEA).

## Introduction

Mean Field Games were first introduced by Lasry and Lions [61, 62] and, simultaneously, by Huang, Caines and Malamh e [57]. This theory is a branch of the broader theory of Dynamic Games and it is devoted to those models where infinitely many players interact strategically with each other.

In many cases the Nash equilibria of those games can be analyzed through the solutions of the, so called, MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{R}^d \times [0, T] \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{R}^d \times [0, T] \\ m(0) = m_0, u(T, x) = G(x, m(T)) & \text{in } \mathbb{R}^d. \end{cases}$$

with unknown the couple  $(u, m)$ . We can think at  $m(t)$  as the distribution of players at time  $t$  and  $u(t, x)$  as the value function of any infinitesimal player starting from  $x$  at time  $t$ .

The aim of this paper is to shed some light on the long time behavior of potential MFG when monotonicity is not in place. The long time behavior and the existence of solutions which are periodic in time have been subject of several papers starting from [65] and the Mexican wave model in [56] to more recent results in [18, 22, 23, 50], but in these papers either monotonicity was assumed or the MFG was not of potential type. In the last few years, the first results in the direction of periodic solutions of non monotone MFG have appeared. The first example of periodic solutions is due to Gomes and Sedjro in [49], even though in the quite different framework of one-dimensional first order systems with congestion. More recently, in the setting of second order MFG, Cirant [34] suggested the existence of non monotone configurations under which oscillatory behaviors were to be expected. Afterwards in [33], with Nurbekyan, they proved, through bifurcation methods, the existence of a path of branches which corresponds to a periodic trajectory. The main difference with our work is the choice of the class of solutions. In our case we look at paths which are energy minimizers whereas, in their, it might not be the case. It is worthwhile to mention that the long time problem for MFG is deeply connected to the problem of long time convergence of solutions of Hamilton-Jacobi equations. In this latter context a similar approach, based on the non linear adjoint method, has been studied in both first order and second order case in [13].

Potential MFG are those games whose MFG system can be derived as optimality condition of the following minimization problem

$$\mathcal{U}(T, m_0) = \inf_{(m, w)} \int_0^T \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt + \mathcal{G}(m(T)),$$

where  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m + \operatorname{div} w = 0$  with  $m(0) = m_0$  and the functions  $F$  and  $G$  in the MFG system are respectively the derivatives with respect to the measure of  $\mathcal{F}$  and  $\mathcal{G}$ . These games have been largely studied (see Lasry and Lions [62] for existence results and, among others [11, 21, 70] for further properties) and the connection between the long time behavior of potential MFG and their variational structure was already highlighted in [44]. Anyway, so far, not much is known outside the assumption of monotonicity, where Cardaliaguet, Lasry, Lions and Porretta [22] proved the convergence to the ergodic system

$$\begin{cases} -\bar{\lambda} - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \\ \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d. \end{cases}$$

We show that in general this is not the case, even in the very regular setting of non local coupling. We look at the problem from the point of view of weak KAM theory. Note that the link between the two theories is not new and it was already proposed by Cardaliaguet [18] in the first order monotone case, even though in a different manner and with quite different purposes. Moreover, the idea of reformulating some results of the weak KAM theory to tackle infinite dimensional problems was already used in several works. See for instance [45, 46, 48, 47] where the authors consider the cell problem for infinite dimensional systems of particles on the torus. Although not in the MFG setting, those paper address issues that look very much like the one considered in this work, especially if instead of parabolic constraints we were to consider first order ones.

The paper is divided in three sections. In the first one we prove the convergence of  $T^{-1}\mathcal{U}(T, \cdot)$  when  $T$  goes to infinity. The method we use is directly inspired by Lions, Papanicolaou, and Varadhan [66]: instead of looking directly at  $\lim_T T^{-1}\mathcal{U}$  we define the infinite horizon, discounted problem

$$\mathcal{V}_\delta(m_0) = \inf_{(m,w)} \int_0^\infty e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w(t)}{m(t)} \right) dm(t) + \mathcal{F}(m(t)) dt$$

where  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m - \operatorname{div} w = 0$  with  $m(0) = m_0$  and we prove that  $\lim_\delta \delta \mathcal{V}_\delta(\cdot) = -\lambda$  when  $\delta \rightarrow 0^+$  and that this limit is uniform with respect to the initial distribution. A key assumption is the boundedness of the second derivative of  $F(x, m)$  with respect to the state variable. This gives uniform semiconcavity estimates of the solutions of the MFG system associated to the discounted minimization problem. The existence of the limit  $\lim_\delta \delta \mathcal{V}_\delta(\cdot)$  implies the existence of the limit  $\lim_T T^{-1}\mathcal{U}$  and the two must coincide.

As byproduct, we have the existence of a corrector function  $\chi$  on the space of measures which enjoys the following dynamic programming principle

$$\chi(m_0) = \inf_{(m,w)} \left( \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{w(s)}{m(s)} \right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t,$$

where, again,  $(m, w)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m - \operatorname{div} w = 0$  with  $m(0) = m_0$ .

The second section is devoted to the study of the set of corrector functions. A corrector is any continuous function on the space of measures which verifies the dynamic programming principle above. Both the terminology and the techniques are borrowed from weak KAM theory, in particular we rely on Fathi's book [42], along with his seminal papers [39, 40, 41]. In principle, as in the standard weak KAM theory, the corrector functions verify an HJB equation in the space of probability measure. In this work nothing is said about this property which is the subject of a paper that is still in progress.

Particular interest is given to the projected Mather set which is the set of probability measures contained in a calibrated curve. We say that  $(\bar{m}, \bar{w})$  is a calibrated curve associated to a corrector function  $\chi$  if it is defined for any time  $t \in \mathbb{R}$  and if, for any  $t_1, t_2 \in \mathbb{R}$ ,  $(\bar{m}, \bar{w})$  is optimal for the dynamic programming principle, i.e.

$$\chi(\bar{m}(t_1)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{\bar{w}(s)}{\bar{m}(s)} \right) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + \chi(\bar{m}(t_2)) + \lambda(t_2 - t_1).$$

These curves play a fundamental role to understand the long time behavior of these MFG. They are indeed the attractors of the dynamics which minimize the discounted, infinite horizon MFG.

In the third section we focus on the relation between the limit value  $\lambda$  and the ergodic value  $\bar{\lambda}$ , associated to the stationary MFG, defined by

$$-\bar{\lambda} = \inf_{(m,w)} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm + \mathcal{F}(m).$$

We propose two examples which highlight how important is the structure of the coupling function  $F(x, \cdot)$  in the dynamic of potential MFG. In the first example we impose monotonicity and we recover part of the results in [22]. In this case the limit value and the ergodic one coincide.

On the other hand, in the second example, the minimization problems are no longer convex and we can prove that  $\lambda > \bar{\lambda}$ . This means that the solutions of the MFG system can not converge to a stationary equilibrium. The fact that  $\lambda > \bar{\lambda}$  implies that the energy of the finite horizon game goes below the energy of the stationary one. Looking at the projected Mather set we can say even more. As this set is compact and it can not contain any stationary curve, calibrated curves can not even approach any static configuration.

### Acknowledgement

I would like to thank Pierre Cardaliaguet (Paris Dauphine) for the fruitful discussions all along this work and Marco Cirant (Università di Padova) for a crucial hint which resulted in Lemma 1.4.2. Moreover, I wish to express gratitude to the anonymous referees for their very careful reading of the manuscript and their valuable advices.

I was partially supported by the ANR (Agence Nationale de la Recherche) project ANR-16-CE40-0015-01.

### Assumptions and definitions

We work on the  $d$ -dimensional flat torus  $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$  to avoid to deal with boundary conditions and to set the problem on a compact domain. Moreover, we set  $\mathcal{G} \equiv 0$ . However, a more general  $\mathcal{G}$  would not add further difficulties, provided that it verifies suitable regularity conditions like, for instance, the ones verified by  $\mathcal{F}$ .

**Notation:** We denote by  $\mathcal{P}(\mathbb{T}^d)$  the set of Borel probability measures on  $\mathbb{T}^d$ . This is a compact, complete and separable set when endowed with the 1-Wasserstein distance  $\mathbf{d}(\cdot, \cdot)$ . We define  $\mathcal{M}(\mathbb{T}^d; \mathbb{R}^d)$  the set of Borel vector valued measures  $w$  with finite mass  $|w|$ . Let  $m$  be a Borel measure over  $[0, T] \times \mathbb{T}^d$ , with first marginal the Lebesgue measure  $dt$  over  $[0, T]$ , then with  $\{m(t)\}_{t \in [0, T]}$  we denote the disintegration of  $m$  with respect to  $dt$ . We will always consider measures  $m$  such that  $m(t)$  is a probability measure on  $\mathbb{T}^d$  for any  $t \in [0, T]$ .

If  $m$  is such a measure, then  $L_m^2([0, T] \times \mathbb{T}^d)$  is the set of  $m$ -measurable functions  $f$  such that the integral of  $|f|^2 dm(t)$  over  $[0, T] \times \mathbb{T}^d$  is finite. Analogously for  $L_m^2(\mathbb{T}^d)$  and  $L_m^2(\mathbb{T}^d; \mathbb{R}^d)$ , where in the latter case we consider vector valued functions.

We use throughout the paper the notion of derivative for functions defined on  $\mathcal{P}(\mathbb{T}^d)$  introduced in [19]. We say that  $\Phi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is  $C^1$  if there exists a continuous function  $\frac{\delta \Phi}{\delta m} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}$  such that

$$\Phi(m_1) - \Phi(m_2) = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta \Phi}{\delta m}((1-t)m_2 + tm_1, x)(m_1 - m_2)(dx)dt, \quad \forall m_1, m_2 \in \mathcal{P}(\mathbb{T}^d).$$

As this derivative is defined up to an additive constant, we use the standard normalization

$$\int_{\mathbb{T}^d} \frac{\delta \Phi}{\delta m}(m, x)m(dx) = 0.$$

**Assumptions:** We impose the following assumptions on the Hamiltonian  $H$  and the coupling function  $F$  so that we can derive uniform estimates on the solutions of the MFG system.

1.  $H : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is of class  $C^2$ ,  $p \mapsto D_{pp}H(x, p)$  is Lipschitz continuous, uniformly with respect to  $x$ . Moreover there exists  $\bar{C} > 0$  that verifies

$$\bar{C}^{-1}I_d \leq D_{pp}H(x, p) \leq \bar{C}I_d, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d \quad (1.1)$$

and  $\theta \in (0, 1)$ ,  $C > 0$  such that the following conditions hold true

$$|D_{xx}H(x, p)| \leq C(1 + |p|)^{1+\theta}, \quad |D_{x,p}H(x, p)| \leq C(1 + |p|)^\theta, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d. \quad (1.2)$$

Note that (1.1) implies that  $H^*$ , the Fenchel conjugate of  $H$  with respect to the second variable, verifies (1.1) as well for a possible different positive constant. Moreover, it also guarantees that  $H^*$  is bounded from below. Note also that (1.2) is a strong restriction on the class of Hamiltonians that fit the assumptions. In particular we rule out quite natural Hamiltonians as

$$H(x, p) = g(x)|p|^2,$$

regardless of the regularity of  $g$ .

2.  $\mathcal{F} : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is of class  $C^1$ . Its derivative  $F : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is twice differentiable in  $x$  and  $D_{xx}^2 F$  is bounded. Examples of non monotone coupling functions which verify such conditions can be found in Lemma 1.4.3 and Lemma 1.4.4.

We recall that, if  $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$ , the 1-Wasserstein distance is defined by

$$d(\mu, \nu) = \sup \left\{ \int_{\mathbb{T}^d} \phi(x) d(\mu - \nu)(x) \mid \text{continuous } \phi : \mathbb{T}^d \rightarrow \mathbb{R}, \text{Lip}(\phi) \leq 1 \right\}.$$

**Minimization Problems:** Under the above assumptions, we can introduce two minimization problems. Each one of those will be proposed in two different but equivalent forms. The first one is

$$\mathcal{U}(T, m_0) = \inf_{(m, \alpha)} \int_0^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm(t) + \mathcal{F}(m(t)) dt, \quad m_0 \in \mathcal{P}(\mathbb{T}^d)$$

where  $m \in C^0([0, T], \mathcal{P}(\mathbb{T}^d))$ ,  $\alpha \in L_m^2([0, T] \times \mathbb{T}^d, \mathbb{R}^d)$  and the following equation is verified in the sense of distributions

$$\begin{cases} -\partial_t m + \Delta m + \text{div}(m\alpha) = 0 & \text{in } [0, T] \times \mathbb{T}^d \\ m(0) = m_0 & \text{in } \mathbb{T}^d. \end{cases} \quad (1.3)$$

Equivalently (see [11] for more details),

$$\mathcal{U}(T, m_0) = \inf_{(m, w) \in \mathcal{E}_2^T(m_0)} \int_0^T \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt, \quad m_0 \in \mathcal{P}(\mathbb{T}^d)$$

where  $\mathcal{E}_2^T(m_0)$  is the set of Borel measures  $(m, w)$  over  $[0, T] \times \mathbb{T}^d$  such that the first marginal of both  $m$  and  $w$  is the Lebesgue measure  $dt$  over  $[0, T]$  and, if  $m(t)$  and  $w(t)$  are the disintegrations of  $m$  and  $w$  with respect to  $dt$ , then  $m(t) \in \mathcal{P}(\mathbb{T}^d)$  and  $w \in \mathcal{M}(\mathbb{T}^d; \mathbb{R}^d)$ . Moreover, we require that  $w$  is absolutely continuous with respect to  $m$ , its density  $dw/dm$  belongs to  $L_m^2([0, T] \times \mathbb{T}^d)$  and  $-\partial_t m + \Delta m - \text{div} w = 0$  is verified in the sense of distributions with initial condition  $m(0) = m_0$ .



The second minimization problem reads

$$\mathcal{V}_\delta(m_0) = \inf_{(m,\alpha)} \int_0^\infty e^{-\delta t} \int_{\mathbb{T}^d} H^*(x, \alpha) dm(t) + \mathcal{F}(m(t)) dt, \quad m_0 \in \mathcal{P}(\mathbb{T}^d)$$

where  $\delta > 0$ ,  $m \in C^0([0, +\infty), \mathcal{P}(\mathbb{T}^d))$ ,  $\alpha \in L^2_{m,\delta}([0, +\infty) \times \mathbb{T}^d, \mathbb{R}^d)$ , that is  $L^2_m$  with weight  $e^{-\delta t}$ , and  $(m, \alpha)$  verifies (1.3) in  $[0, +\infty) \times \mathbb{T}^d$ . Equivalently,  $\mathcal{V}_\delta$  can be defined as

$$\mathcal{V}_\delta(m_0) = \inf_{(m,w) \in \mathcal{E}_2^\delta(m_0)} \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt, \quad m_0 \in \mathcal{P}(\mathbb{T}^d) \quad (1.4)$$

where  $\mathcal{E}_2^\delta(m_0)$  is defined as  $\mathcal{E}_2(m_0)$  with the only difference that we ask  $dw/dm$  to be  $L^2$  in  $[0, +\infty) \times \mathbb{T}^d$  with respect to  $e^{-\delta t} m(t)$ . For convenience we introduce the functional on  $\mathcal{E}_2^\delta(m_0)$

$$J_\delta(m_0, m, w) = \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt,$$

so that  $\mathcal{V}_\delta(m_0) = \inf_{(m,w)} J_\delta(m_0, m, w)$ .

We also define the ergodic value  $\bar{\lambda}$  as follows

$$-\bar{\lambda} = \inf_{m,\alpha} \int_{\mathbb{T}^d} H^*(x, \alpha) dm + \mathcal{F}(m) \quad (1.5)$$

where  $(m, \alpha) \in \mathcal{P}(\mathbb{T}^d) \times L^2_m(\mathbb{T}^d; \mathbb{R}^d)$  verifies in the sense of distributions  $\Delta m + \operatorname{div}(m\alpha) = 0$  in  $\mathbb{T}^d$ .

Or, equivalently,

$$-\bar{\lambda} = \inf_{(m,w) \in \mathcal{E}} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw}{dm}(x) \right) dm(x) + \mathcal{F}(m) \quad (1.6)$$

where  $\mathcal{E}$  is the set of  $(m, w) \in \mathcal{P}(\mathbb{T}^d) \times \mathcal{M}(\mathbb{T}^d; \mathbb{R}^d)$  such that  $w$  is absolutely continuous with respect to  $m$ , its density  $dw/dm$  belongs to  $L^2_m(\mathbb{T}^d)$  and  $\Delta m - \operatorname{div} w = 0$  is verified in the sense of distributions.

Throughout the paper we will use the constant  $C > 0$  which may change from line to line.

## 1.1 Ergodic limit value

### 1.1.1 Minimizers and dynamic programming principle for $\mathcal{V}_\delta$

We start proving that the minimization problem (1.4) admits a minimizer and we also give a characterization of such a minimizer in terms of solutions of the associated MFG system.

**Proposition 1.1.1.** *For any  $\delta > 0$  and any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,  $\mathcal{V}_\delta(m_0)$  admits a minimizer  $(m, w)$ . Moreover there exists  $u \in C^{1,2}([0, +\infty) \times \mathbb{T}^d)$  and  $m \in C^0([0, +\infty) \times \mathcal{P}(\mathbb{T}^d))$  solutions of*

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = \frac{\delta \mathcal{F}}{\delta m}(x, m) := F(x, m) & \text{in } \mathbb{T}^d \times [0, +\infty) \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, +\infty) \\ m(0) = m_0, u \in L^\infty([0, +\infty) \times \mathbb{T}^d) \end{cases} \quad (1.7)$$

such that  $w = -m D_p H(x, Du)$ .

*Proof.* First of all we show that  $\mathcal{V}_\delta(m_0)$  is finite and that it is bounded by a constant  $K_\delta$  independent of  $m_0$ . We can always use as competitor for  $\mathcal{V}_\delta(m_0)$  the couple  $(\mu, 0)$  where  $\mu$  is the solution of

$$\begin{cases} -\partial_t \mu + \Delta \mu = 0 & \text{in } \mathbb{T}^d \times [0, +\infty) \\ \mu(0) = m_0 & \text{in } \mathbb{T}^d. \end{cases}$$

Given that  $\mathcal{F}$  is bounded, if we use  $(\mu, 0)$  as a competitor, we get

$$\mathcal{V}_\delta(m_0) \leq \int_0^{+\infty} e^{-\delta t} \left( \sup_{x \in \mathbb{T}^d} H^*(x, 0) + \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \mathcal{F}(m) \right) dt := K_\delta.$$

We fix a minimizing sequence  $(m_n, w_n)$  and we denote with  $\{w_n(t)\}_t$  the disintegration of  $w_n$  with respect to the Lebesgue measure  $dt$  over  $\mathbb{R}^+$ . If we use (1.1), we have that there exists a constant  $M_\delta > 0$  that does not depend on  $m_0$  and  $n$ , such that

$$\int_0^{+\infty} e^{-\delta t} dt \int_{\mathbb{T}^d} d|w_n(t)| \leq M_\delta$$

and

$$\int_0^{+\infty} e^{-\delta t} dt \int_{\mathbb{T}^d} \left| \frac{dw_n(t)}{dm_n(t)} \right|^2 dm_n(t) \leq M_\delta.$$

Hence, for any fixed  $k > 0$

$$\int_0^k dt \int_{\mathbb{T}^d} d|w_n(t)| \leq M_\delta e^{\delta k} \quad (1.8)$$

and

$$\left\| \frac{dw_n(t)}{dm_n(t)} \right\|_{L^2_{m_n}([0, k] \times \mathbb{T}^d)}^2 \leq M_\delta e^{\delta k}. \quad (1.9)$$

Following Lemma 4.1 in [21], we get that for any  $t, s \in [0, k]$

$$\mathbf{d}(m_n(t), m_n(s)) \leq C_\delta^k |t - s|^{1/2}, \quad (1.10)$$

where  $C_\delta^k$  depends only on  $\delta$  and  $k$ . Inequality (1.10) tells us that  $\{m_n\}_n$  is uniformly bounded in  $C^{1/2}([0, k], \mathcal{P}(\mathbb{T}^d))$ . We have then that  $m_n$  converges uniformly on any compact set to a limit  $\bar{m} \in C^0([0, +\infty), \mathcal{P}(\mathbb{T}^d))$ . Thanks to the bounds (1.8) we also know that  $w_n$  weakly converges in  $\mathcal{M}(I \times \mathbb{T}^d; \mathbb{R}^d)$  to a certain  $\bar{w}$  on any bounded interval  $I \subset [0, +\infty)$ . We argue also that  $\bar{w}$  is absolutely continuous with respect to  $\bar{m}$ , which comes from the uniform boundedness on compact sets of  $\|dw_n/dm_n\|_{L^2_{m_n}}$  ensured by (1.9).

As the couple  $(\bar{m}, \bar{w})$  belongs to  $\mathcal{E}_2^\delta(m_0)$  we have that  $J_\delta(m_0, \bar{m}, \bar{w}) < +\infty$ . This means that

$$J_\delta(m_0, \bar{m}, \bar{w}) = \lim_{k \rightarrow +\infty} \int_0^k e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{d\bar{w}(t)}{d\bar{m}(t)}(x) \right) d\bar{m}(t) + \mathcal{F}(\bar{m}(t)) dt$$

Note also that the functional is bounded from below, so there exists a constant  $C_\delta$  such that, for any  $(m, w) \in \mathcal{E}_2^\delta(m_0)$ ,

$$\begin{aligned} & \int_k^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt \\ & \geq e^{-\delta k} \int_0^{+\infty} e^{-\delta t} \inf_{(x,p) \in \mathbb{T}^d \times \mathbb{R}^d} H^*(x, p) + \inf_{m \in \mathcal{P}(\mathbb{T}^d)} \mathcal{F}(m) dt = e^{-\delta k} C_\delta. \end{aligned}$$

Therefore,

$$J_\delta(m_0, m_n, w_n) \geq \int_0^k e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw_n(t)}{dm_n(t)}(x) \right) dm_n(t) + \mathcal{F}(m_n(t)) dt + e^{-\delta k} C_\delta.$$

Thanks to the convergence of  $(m_n, w_n)$  on compact sets, we can pass to the limit in  $n$  and we get

$$\mathcal{V}_\delta(m_0) \geq \int_0^k e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{d\bar{w}(t)}{d\bar{m}(t)}(x) \right) d\bar{m}(t) + \mathcal{F}(\bar{m}(t)) dt + C_\delta e^{-\delta k}.$$

Taking the limit on  $k$  we finally get that  $\mathcal{V}_\delta(m_0) \geq J(m_0, \bar{m}, \bar{w})$ .

The proof of the second statement is a direct application of Fenchel-Rockafellar duality theorem that we present for completeness, even though it follows closely the proof of Lemma 4.2 in [21].

Using the convexity of  $H^*$  and the regularity of  $\mathcal{F}$ , we can show that if  $(\bar{m}, \bar{w})$  is a minimizer for  $\mathcal{V}_\delta(m_0)$  then it must be a minimizer for  $\bar{J}_\delta : \mathcal{E}_2^\delta(m_0) \rightarrow \mathbb{R}$ , which reads

$$\bar{J}_\delta(m, w) = \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + F(x, \bar{m}(t)) dm(t) dt.$$

A detailed proof of this statement can be found in [11].

As  $\bar{J}_\delta$  convex, we can see it as the dual problem (in the sense of Fenchel-Rockafellar) of the following one

$$\inf_{u \in C_b^2([0, +\infty) \times \mathbb{T}^d)} \left\{ - \int_{\mathbb{T}^d} u(0, x) dm_0(x) \text{ where } -\partial_t u - \Delta u + \delta u + H(x, Du) \leq F(x, \bar{m}) \right\}, \quad (1.11)$$

which means that

$$- \inf_{-\partial_t u - \Delta u + \delta u + H(x, Du) \leq F(x, \bar{m})} \int_{\mathbb{T}^d} u(0, x) dm_0(x) = - \min_{(m, w) \in \mathcal{E}_2^\delta(m_0)} \bar{J}_\delta(m, w). \quad (1.12)$$

The proof of the above result is exhibit in Appendix in Lemma 1.4.1.

Note that the equation

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = F(x, \bar{m}) & \text{in } [0, +\infty) \times \mathbb{T}^d \\ u \in L^\infty([0, +\infty) \times \mathbb{T}^d) \end{cases}$$

enjoys the comparison principle and so it admits a unique solution  $\bar{u} \in C_b^2([0, +\infty) \times \mathbb{T}^d)$  which has to be the minimizer of (1.11). Indeed, let  $u$  be such that

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) \leq F(x, \bar{m}) & \text{in } [0, +\infty) \times \mathbb{T}^d \\ u \in L^\infty([0, +\infty) \times \mathbb{T}^d), \end{cases}$$

then, by the comparison principle,  $u(0, x) \leq \bar{u}(0, x)$  for any  $x \in \mathbb{T}^d$ .

To prove that  $\bar{w} = -\bar{m} D_p H(x, D\bar{u})$  we start summing the two minimization problems in duality so that we get

$$\int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{d\bar{w}(t)}{d\bar{m}(t)}(x) \right) d\bar{m}(t) + F(x, \bar{m}(t)) d\bar{m}(t) dt - \int_{\mathbb{T}^d} \bar{u}(0, x) dm_0(x) = 0.$$

As  $\bar{u}$  solves  $-\partial_t \bar{u} - \Delta \bar{u} + \delta \bar{u} + H(x, D\bar{u}) = F(x, \bar{m})$ , the above relation becomes

$$\int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{d\bar{w}(t)}{d\bar{m}(t)}(x) \right) - \partial_t \bar{u}(t, x) - \Delta \bar{u}(t, x) + \delta \bar{u}(t, x) + H(x, D\bar{u}(t, x)) d\bar{m}(t) dt$$

$$- \int_{\mathbb{T}^d} \bar{u}(0, x) dm_0(x) = 0,$$

which can be rewrite as

$$\int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{d\bar{w}(t)}{d\bar{m}(t)}(x) \right) + H(x, D\bar{u}(t, x)) d\bar{m}(t) dt$$

$$\int_0^{+\infty} \int_{\mathbb{T}^d} \partial_t \left( e^{-\delta t} \bar{u}(t, x) \right) - \Delta \left( e^{-\delta t} \bar{u}(t, x) \right) d\bar{m}(t) dt - \int_{\mathbb{T}^d} \bar{u}(0, x) dm_0(x) = 0.$$

If we use the Fokker-Plank equation verified by  $(\bar{m}, \bar{w})$ , we end up with

$$\int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{d\bar{w}(t)}{d\bar{m}(t)}(x) \right) + H(x, D\bar{u}(t, x)) d\bar{m}(t) dt + \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} D\bar{u}(t, x) d\bar{w}(t, x) = 0.$$

As  $H$  is uniformly convex, the above relation holds true only if  $\bar{w} = -\bar{m} D_p H(x, D\bar{u})$   $\bar{m}$ -a.e.. This means that  $\bar{w}$  is bounded which in turn implies that  $\bar{m} > 0$  so that  $\bar{w} = -\bar{m} D_p H(x, D\bar{u})$  is verified everywhere.  $\square$

In the following lemma we present the dynamic programming principle verified by  $\mathcal{V}_\delta$ . We placed the proof in the appendix because, although in a different framework, it relies on standard arguments in optimal control theory (see for instance [15] for the finite dimension setting).

**Lemma 1.1.2.** *The function  $\mathcal{V}_\delta$  verifies the dynamic programming principle, which reads*

$$\mathcal{V}_\delta(m_0) = \inf_{(m, \alpha) \in \mathcal{E}_\delta^2(m_0)} \left( \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta t} \mathcal{V}_\delta(m(t)) \right).$$

### 1.1.2 Existence of a corrector

The main result of this section is Theorem 1.1.5 where we show that the function  $\mathcal{V}_\delta(\cdot)$  is uniformly Lipschitz with respect to  $\delta$ . As a consequence, we have Proposition 1.1.6 which claims on one side that the limit  $\lim_{\delta \rightarrow 0} \delta \mathcal{V}_\delta(m_0)$  is well defined and it is uniform in  $m_0$  and, on the other, that, up to subsequence, also  $\mathcal{V}_\delta(\cdot) - \mathcal{V}_\delta(m_0)$  converges to a continuous function  $\chi$ . In Lemma 1.1.7 we prove that  $\chi$  enjoys the dynamic programming principle and, therefore, we have the existence of a corrector function.

The idea behind the proof of Theorem 1.1.5 is the following: we want to prove that there exists a constant  $\bar{K} > 0$  independent of  $\delta$ , such that

$$|\mathcal{V}_\delta(m_2^0) - \mathcal{V}_\delta(m_1^0)| \leq \bar{K} \mathbf{d}(m_1^0, m_2^0).$$

We fix an horizon  $T > 0$ , to be chosen later, and we take  $(m_1(\cdot), \alpha_1(\cdot))$  a minimizer for  $\mathcal{V}_\delta(m_1^0)$ . We consider any couple  $(m_2, \alpha_2)$  such that (1.3) is verified in  $[0, T]$ ,  $m_2(0) = m_2^0$ ,  $m_2(T) = m_1(T)$  and  $m_2 \equiv m_1$ ,  $\alpha_2 \equiv \alpha_1$  in  $[T, \infty)$  (for the existence of such a  $(m_2, \alpha_2)$  see the construction in (1.17) and below). The couple  $(m_2, \alpha_2)$  is admissible and

$$\mathcal{V}_\delta(m_2^0) \leq \int_0^T e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(s) + \mathcal{F}(m_2(s)) ds + e^{-\delta T} \mathcal{V}_\delta(m_1(T)).$$

Therefore,

$$\mathcal{V}_\delta(m_2^0) - \mathcal{V}_\delta(m_1^0) \leq \int_0^T e^{-\delta t} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(t) - H^*(x, \alpha_1) dm_1(t) + \mathcal{F}(m_2(t)) - \mathcal{F}(m_1(t)) dt. \quad (1.13)$$

In order to prove the continuity of  $\mathcal{V}_\delta$  with respect to the initial data we need to introduce some standard estimates on the solutions of the MFG system (1.7).

**Lemma 1.1.3.** *There exists  $C > 0$  independent of  $m_0, \delta$  such that, if  $(u, m)$  is a classical solution of (1.7), then*

- $\|Du\|_{L^\infty([0, +\infty) \times \mathbb{T}^d)} \leq C$
- $|D^2u(s, \cdot)| \leq C$  for any  $s \in [0, +\infty)$
- $d(m(s), m(l)) \leq C|l - s|^{1/2}$  for any  $l, s \in [0, +\infty)$

Consequently, we also have that  $|\partial_t u(s, \cdot)| \leq C$  for any  $s \in [0, +\infty)$ .

*Proof.* The proof follows closely the one proposed in [28] and it relies on semiconcavity estimates for the value function  $u$ . We recall that if  $\phi \in C^\infty(\mathbb{T}^d)$  then

$$\|D\phi\|_{L^\infty(\mathbb{T}^d)} \leq d^{\frac{1}{2}} \sup_{x \in \mathbb{T}^d, |\xi| \leq 1} D^2\phi(x)\xi \cdot \xi. \quad (1.14)$$

We first prove the result for  $u^T : [0, T] \times \mathbb{T}^d \rightarrow \mathbb{R}$ , solution of

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, T] \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, T] \\ m(t) = m_0, u(T, x) = 0 & \text{in } \mathbb{T}^d \end{cases}$$

We consider  $\xi \in \mathbb{R}^d$ ,  $|\xi| \leq 1$ , that maximizes  $\sup_{t,x} D^2u^T(t, x)\xi \cdot \xi = M$  and we look at the equation solved by  $w(t, x) = D^2u^T(t, x)\xi \cdot \xi$  deriving twice in space the HJB equation in (1.7):

$$\begin{aligned} & -\partial_t w - \Delta w + \delta w + D_{\xi\xi}H(x, Du) + 2D_{\xi p}H(x, Du) \cdot D^2u\xi \\ & + D_{pp}H(x, Du)D^2u\xi \cdot D^2u\xi + D_pH(x, Du) \cdot Dw = D_{\xi\xi}^2F(x, m). \end{aligned}$$

The maximum of  $w$  can be achieved either at  $t = T$ , but using the terminal condition of  $u^T$  we get  $M = 0$ , or at a point  $(s, x)$  in the interior. In this case, if we use hypothesis (1.1) and (1.2) on  $H$  and the boundedness of  $D_{xx}^2F$ , then, at the maximum  $(s, x)$ , we get the following inequality

$$\delta M - C(1 + |Du|)^{1+\theta} - 2C(1 + |Du|)^\theta |D^2u\xi| + \bar{C}^{-1}|D^2u\xi|^2 \leq C.$$

By Cauchy-Schwarz inequality we have that  $M = w(s, x) \leq |D^2u(s, x)\xi|$ . If we also plug (1.14) we get, for a possible different constant  $C$

$$-C(1 + M)^{1+\theta} - 2C(1 + M)^{2\theta} + \bar{C}^{-1}M^2 \leq C.$$

Given that  $\theta < 1$  the above inequality ensures that  $M$  is bounded by a constant that does not depend on  $m_0, \delta$  and  $T$ . The bound on  $\|Du^T\|_{L^\infty([0, T] \times \mathbb{T}^d)}$  follows from (1.14). Now that we proved that  $Du^T$  is bounded so that Theorem V 5.4 in [60] gives us the boundedness of  $D^2u^T$ . Note that the estimates on  $Du^T$  and  $D^2u^T$  imply directly from the HJB equation that  $\partial_t u^T$  is bounded as well. As all the estimates are independent of  $T$ , if we look at the sequence of  $u^T$  we have that, on any compact set,  $u^T$  is uniformly bounded and continuous. This means that  $u^T$  converges to  $u$  solution of the HJB equation on  $[0, +\infty)$  and the same estimates hold true for  $u$ .

Furthermore, it implies that also  $D_pH(x, Du)$ , the drift of the Fokker Planck equation, is uniformly bounded. Standard results on SDEs (for instance Lemma 3.4 in [17]) ensure the Holder continuity, with respect to the Wasserstein distance, of  $s \mapsto m(s)$  uniformly with respect to  $m_0$  and  $\delta$ .  $\square$

We now fix  $m_1^0, m_2^0 \in \mathcal{P}(\mathbb{T}^d)$ . According to (1.13) we have

$$\mathcal{V}_\delta(m_2^0) - \mathcal{V}_\delta(m_1^0) \leq$$

$$\int_0^T e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(s) - H^*(x, DH(x, Du_1)) dm_1(s) + \mathcal{F}(m_2(s)) - \mathcal{F}(m_1(s)) ds, \quad (1.15)$$

where  $(u_1, m_1)$  is a solution of (1.7) related to a minimizer  $(m_1, w_1)$  of  $\mathcal{V}_\delta(m_1^0)$  that we found in Proposition 1.1.1. The couple  $(m_2, \alpha_2)$  is such that (1.3) is verified in  $[0, T]$  with  $m_2(0) = m_2^0$  and  $m_2(T) = m_1(T)$ . The key point to prove the continuity of  $\mathcal{V}_\delta$  is to construct a suitable  $(m_2, \alpha_2)$ . We first consider  $\tilde{m}_2$  solution of

$$\begin{cases} -\partial_t \tilde{m}_2 + \Delta \tilde{m}_2 + \operatorname{div}(\tilde{m}_2 D_p H(x, Du_1)) = 0 & \text{in } [0, T] \times \mathbb{T}^d \\ \tilde{m}_2(0) = m_2^0 \end{cases} \quad (1.16)$$

and then we set

$$m_2(s, x) = \begin{cases} \tilde{m}_2(s, x), & \text{if } s \in (0, h] \\ \frac{\tau+h-s}{\tau} \tilde{m}_2(s, x) + \frac{s-h}{\tau} m_1(s, x), & \text{if } s \in [h, h+\tau] \\ m_1(s, x) & \text{if } s \in [\tau+h, T], \end{cases} \quad (1.17)$$

where  $h$  and  $\tau$  will be chosen later. Note that, thanks to the boundedness of their drifts, Corollary 6.3.2 in [10] ensures that  $\tilde{m}_2$  and  $m_1$  have a density for any  $s > 0$ . What we still need is to define  $\alpha_2$  in  $[h, h+\tau]$ . We compute the equation verified by  $m_2$  in  $[h, h+\tau]$  and, using (1.7) and (1.16), we get

$$\partial_t m_2 - \Delta m_2 = \frac{\tau+h-s}{\tau} \operatorname{div}(\tilde{m}_2 D_p H(x, Du_1)) + \frac{s-h}{\tau} \operatorname{div}(m_1 D_p H(x, Du_1)) + \frac{m_1 - \tilde{m}_2}{\tau}$$

that is, by linearity,

$$\partial_t m_2 - \Delta m_2 = \operatorname{div}(m_2 D_p H(x, Du_1)) + \frac{m_1 - \tilde{m}_2}{\tau}.$$

Let  $\zeta : [0, h+\tau] \times \mathbb{T}^d \rightarrow \mathbb{R}$  be the solution to

$$\begin{cases} \Delta \zeta = m_1 - \tilde{m}_2, & \text{in } [0, h+\tau] \times \mathbb{T}^d \\ \int_{\mathbb{T}^d} \zeta(s, x) = 0. \end{cases} \quad (1.18)$$

We can now define the drift  $\alpha_2$  as follows:  $\alpha_2 = D_p H(x, Du_1) + \frac{D\zeta}{m_2 \tau}$  in  $[h, h+\tau]$  and  $\alpha_2 = D_p H(x, Du_1)$  elsewhere. As  $(m_2, \alpha_2)$  verifies the Fokker-Plank equation by construction with  $m_2(0) = m_2^0$ , it is admissible. For the continuity of  $\mathcal{V}_\delta$  we still need estimates on the drift  $\alpha_2$ . We prove those estimates in the next lemma using the regularity of the solutions of the adjoint of the Fokker-Plank equation.

**Lemma 1.1.4.** *For any time  $s < h+\tau$ , there exists a constant  $K_s > 0$ , bounded for  $s > \varepsilon > 0$ , such that*

$$\|D\zeta(s, \cdot)\|_{L^2(\mathbb{T}^d)} \leq K_s \mathbf{d}(m_1^0, m_2^0).$$

The constant  $K_s$  is independent of  $m_1^0, m_2^0$ .

*Proof.* We first note that, if we multiply (1.18) by  $\zeta$  and we use Cauchy-Schwarz inequality on the right hand side, we get

$$\|D\zeta(s)\|_{L^2(\mathbb{T}^d)}^2 \leq \|m_1(s) - \tilde{m}_2(s)\|_{L^2(\mathbb{T}^d)} \|\zeta(s)\|_{L^2(\mathbb{T}^d)}.$$

Now Pointcaré-Wirtinger inequality gives us

$$\|D\zeta(s)\|_{L^2(\mathbb{T}^d)} \leq C \|m_1(s) - \tilde{m}_2(s)\|_{L^2(\mathbb{T}^d)}. \quad (1.19)$$

If we define  $\mu(s) = m_1(s) - \tilde{m}_2(s)$  then  $\mu$  verifies the following equation

$$\begin{cases} -\partial_t \mu + \Delta \mu + \operatorname{div}(\mu D_p H(x, Du_1)) = 0 & \text{in } [0, s] \times \mathbb{T}^d \\ \mu(0) = m_1^0 - m_2^0 & \text{in } \mathbb{T}^d. \end{cases} \quad (1.20)$$

We now fix a  $\bar{\phi} \in L^2(\mathbb{T}^d)$  and we consider the adjoint backward equation

$$\begin{cases} -\partial_t \phi - \Delta \phi + D_p H(x, Du_1) D\phi = 0 & \text{in } [0, s] \times \mathbb{T}^d \\ \phi(s, x) = \bar{\phi}(x) & \text{in } \mathbb{T}^d. \end{cases} \quad (1.21)$$

Given that  $D_p H(x, Du)$  is bounded, if  $\phi$  is the solution of (1.21), then there exists a constant  $K_s$  (Theorem 11.1 in [60]), such that

$$\|\phi(0)\|_{C^{1+\alpha}(\mathbb{T}^d)} \leq K_s \|\bar{\phi}\|_{L^2(\mathbb{T}^d)}. \quad (1.22)$$

As the equation (1.21) is the adjoint of (1.20),

$$\int_{\mathbb{T}^d} \phi(s) \mu(s) dx = \int_{\mathbb{T}^d} \phi(0) \mu(0) dx.$$

We now plug in the initial and terminal conditions and we estimate the right-hand side as follows

$$\int_{\mathbb{T}^d} \bar{\phi}(x) (m_1(s)(dx) - \tilde{m}_2(s)(dx)) = \int_{\mathbb{T}^d} \phi(0, x) (m_1^0(dx) - m_2^0(dx)) \leq \|D\phi(0)\|_{L^\infty} \mathbf{d}(m_1^0, m_2^0).$$

If we use the interior estimate (1.22) on the right-hand side and we take the supremum over  $\|\bar{\phi}\|_{L^2} \leq 1$ , we finally end up with

$$\|m_1(s) - \tilde{m}_2(s)\|_{L^2} \leq K_s \mathbf{d}(m_1^0, m_2^0).$$

If we plug the last inequality into (1.19), we get the result.  $\square$

**Theorem 1.1.5.** *The family of functions  $\{\mathcal{V}_\delta(\cdot)\}_\delta$  is uniformly Lipschitz continuous.*

*Proof.* Let  $\alpha_1(t, x) = D_p H(x, Du_1)$  for any  $t \in [0, h + \tau]$ . We consider the same  $(m_2, \alpha_2)$  that we defined earlier:  $m_2$  is defined in (1.17),  $\alpha_2 = \alpha_1 + \frac{D\zeta}{m_2 \tau}$  in  $[h, h + \tau]$  and  $\alpha_2 = \alpha_1$  elsewhere, where  $\zeta$  solves (1.18). According to (1.15) we have

$$\begin{aligned} \mathcal{V}_\delta(m_2^0) - \mathcal{V}_\delta(m_1^0) &\leq \int_0^h e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_1) d(m_2 - m_1) + \\ &+ \int_h^{h+\tau} e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2 - H^*(x, \alpha_1) dm_1 + \int_0^{h+\tau} \mathcal{F}(m_2(s)) - \mathcal{F}(m_1(s)) ds. \end{aligned} \quad (1.23)$$

Thanks to the convexity of  $H$ , if we run a second order expansion of  $H^*(x, \cdot)$  around  $\alpha_1$  and we use the growth condition (1.1) on  $D_{pp}H$ , we can estimate the term  $H^*(x, \alpha_2(s))$  for any time  $s \in [h, h + \tau]$  as follows

$$\int_{\mathbb{T}^d} H^*(x, \alpha_2(s)) dm_2(s) \leq \int_{\mathbb{T}^d} H^*(x, \alpha_1(s)) dm_2(s) + \frac{1}{\tau} |D_p H^*(x, \alpha_1(s))| |D\zeta(s)| + \bar{C} \frac{|D\zeta(s)|^2}{\tau^2 m_2(s)} dx. \quad (1.24)$$

We recall that, as the drift  $D_p H(x, Du_1)$  is continuous and bounded, according to Theorem 2.2.1 [9], the measure  $m_1$  has a density  $m_1(s, x)$  for any  $s > 0$ , then, using Theorem 2.5.1 in [9], for any  $l \in (0, s)$ , we have

$$m_1(s, x) > m_1(l, x_0) e^{-Q(1 + \frac{1}{s-l} + \frac{1}{l})},$$

where  $Q$  does not depend on  $m_1^0$ ,  $l$  and  $s$ . As  $\mathbb{T}^d$  is bounded with measure  $|\mathbb{T}^d| = 1$ , for any  $l > 0$ , there exists a  $x_0$  such that  $m_1(l, x_0) > 1/2$ . Given that the same holds true for  $\tilde{m}_2$  then, for any  $s \in [h, h + \tau]$ , the definition of  $m_2$  in (1.17) implies that

$$m_2(s, x) > \frac{1}{2} e^{-Q(1 + \frac{1}{s-l} + \frac{1}{l})} \quad \forall l \in (0, h).$$

For  $l = h/2$  we obtain that the infimum, with respect to  $s$ , in the right-hand side is achieved when  $s = h$ . Thus,

$$m_2(s, x) > \frac{1}{2} e^{-Q(1 + \frac{4}{h})}. \quad (1.25)$$

We can now plug (1.24) and (1.25) into (1.23), which becomes

$$\begin{aligned} \mathcal{V}_\delta(m_2^0) - \mathcal{V}_\delta(m_1^0) &\leq \int_0^{h+\tau} e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_1) d(m_2 - m_1) + \mathcal{F}(m_2(s)) - \mathcal{F}(m_1(s)) ds \\ &\quad + \int_h^{h+\tau} e^{-\delta s} \int_{\mathbb{T}^d} \frac{C_2}{\tau} |D\zeta| + 2 \frac{C}{\tau^2} |D\zeta|^2 e^{Q(1 + \frac{4}{h})} dx ds. \end{aligned}$$

Using the bounds on  $Du_1$  found in Lemma 1.1.3, Lemma 1.1.4 and the regularizing property of  $\mathcal{F}$ , we get

$$\begin{aligned} \mathcal{V}_\delta(m_2^0) - \mathcal{V}_\delta(m_1^0) &\leq \\ C \int_0^{h+\tau} \mathbf{d}(m_2(s), m_1(s)) ds &+ \int_h^{h+\tau} C \mathbf{d}(m_1^0, m_2^0) K_s + 2 \frac{C}{\tau^2} \mathbf{d}^2(m_1^0, m_2^0) e^{Q(1 + \frac{4}{h})} K_s^2 ds \leq \\ C \int_0^{h+\tau} \mathbf{d}(m_2(s), m_1(s)) ds &+ C \mathbf{d}(m_1^0, m_2^0) e^{Q(1 + \frac{4}{h})} \int_h^{h+\tau} \left( 1 + \frac{K_s}{\tau^2} \mathbf{d}(m_1^0, m_2^0) \right) K_s ds \leq \\ C \int_0^{h+\tau} \mathbf{d}(m_2(s), m_1(s)) ds &+ \frac{C}{\tau^2} \mathbf{d}(m_1^0, m_2^0) e^{Q(1 + \frac{4}{h})} \int_h^{h+\tau} K_s^2 ds. \end{aligned} \quad (1.26)$$

In the last inequality we neglected the terms which go to infinity slower than  $K_s^2$  and which vanish faster than  $\mathbf{d}(m_1^0, m_2^0)$ . Note that the constant  $K_s$  might explode when  $s$  goes to 0 but, otherwise, it is bounded. Therefore, as  $h > 0$ , there is no problem of integrability for the term  $\int_h^{h+\tau} K_s^2$ .

We now focus on the first term in the above inequality. In order to estimate  $\mathbf{d}(m_1(s), m_2(s))$ , we have to look at the SDEs verified by the stochastic processes whose laws are  $m_1$  and  $m_2$ . We first recall that an equivalent formulation of the 1-Wasserstein distance between two probability measures  $\mu$  and  $\nu$  is

$$\mathbf{d}(\mu, \nu) = \inf_{\gamma} \left\{ \int_{\mathbb{T}^d \times \mathbb{T}^d} |x - y| d\gamma(x, y) \text{ s.t. } \pi_1 \gamma = \mu, \pi_2 \gamma = \nu \right\}. \quad (1.27)$$



We consider a standard probability space  $(\Omega, \mathcal{G}, \mathbb{P})$  and two random variables  $Z^1, Z^2$  such that  $\mathcal{L}(Z^i) = m_i^0$  and  $\mathbb{E} [|Z^2 - Z^1|] = \mathbf{d}(m_1^0, m_2^0)$ . Therefore,  $m_1$  and  $m_2$  are the laws of the processes defined by the following SDEs

$$\begin{cases} dX_s^i = \alpha_i(t, X_s)ds + \sqrt{2}dB_s \\ X_0^i = Z^i. \end{cases}$$

Using the definition of distance in (1.27), we have

$$\mathbf{d}(m_1(s), m_2(s)) \leq \mathbb{E} [|X_s^2 - X_s^1|] \leq \mathbb{E} \left[ |Z^2 - Z^1| + \int_0^s |\alpha_2(l, X_l^2) - \alpha_1(l, X_l^1)| dl \right].$$

We first split  $\int_0^{h+\tau} \mathbf{d}(m_2(s), m_1(s))ds$  in the sum of the integrals on the intervals  $[0, h]$  and  $[h, h + \tau]$ . For any  $s \in [0, h]$ ,  $\alpha_1(l, x) = D_p H(x, Du_1(t, x)) = \alpha_2(l, x)$ , then

$$\mathbf{d}(m_1(s), m_2(s)) \leq \mathbf{d}(m_1^0, m_2^0) + \mathbb{E} \left[ \int_0^s |D_p H(X_l^2, Du_1(l, X_l^2)) - D_p H(X_l^1, Du_1(l, X_l^1))| dl \right].$$

Hypothesis (1.1), (1.2) and Lemma 1.1.3 ensure that both  $p \rightarrow D_p H(x, p)$  and  $x \rightarrow D_p H(x, Du_1(l, x))$  are Lipschitz continuous, hence

$$\mathbf{d}(m_1(s), m_2(s)) \leq \mathbf{d}(m_1^0, m_2^0) + C \int_0^s \mathbf{d}(m_1(l), m_2(l))dl.$$

If we apply Gronwall's inequality, then for any  $s \in [0, h]$

$$\mathbf{d}(m_1(s), m_2(s)) \leq \mathbf{d}(m_1^0, m_2^0)e^{Cs}. \quad (1.28)$$

We now look at  $\int_h^{h+\tau} \mathbf{d}(m_1(s), m_2(s))ds$ . According to the definition of  $\alpha_2$ , for  $s \in [h, h + \tau]$ , we have

$$\begin{aligned} \mathbf{d}(m_1(s), m_2(s)) &\leq \mathbf{d}(m_1(h), m_2(h)) + \\ &+ \mathbb{E} \left[ \int_h^s \left| D_p H(X_l^2, Du_1(l, X_l^2)) + \frac{D\zeta(l, X_l^2)}{\tau m_2(l, X_l^2)} - D_p H(X_l^1, Du_1(l, X_l^1)) \right| dl \right]. \end{aligned} \quad (1.29)$$

Using (1.28) on  $\mathbf{d}(m_1(h), m_2(h))$  and splitting the last term, we get that (1.29) is smaller than

$$\begin{aligned} \mathbf{d}(m_1^0, m_2^0)e^{Ch} + \mathbb{E} \left[ \int_h^s |D_p H(X_l^2, Du_1(l, X_l^2)) - D_p H(X_l^1, Du_1(l, X_l^1))| dl \right] \\ + \mathbb{E} \left[ \int_h^s \left| \frac{D\zeta(l, X_l^2)}{\tau m_2(l, X_l^2)} \right| dl \right]. \end{aligned}$$

If we use again that  $x \rightarrow D_p H(x, Du_1(l, x))$  is Lipschitz continuous, we get

$$\mathbf{d}(m_1(s), m_2(s)) \leq \mathbf{d}(m_1^0, m_2^0)e^{Ch} + \frac{C}{\tau} \mathbb{E} \left[ \int_h^s \left| \frac{D\zeta(l, X_l^2)}{m_2(l, X_l^2)} \right| dl \right] + \int_h^s \mathbf{d}(m_1(l), m_2(l))dl. \quad (1.30)$$

Thanks to estimates (1.25) on  $m_2$  we can use Tonelli's theorem and switch the integral with the expectation. Using Lemma 1.1.4, we eventually have

$$\mathbb{E} \left[ \int_h^s \left| \frac{D\zeta(l, X_l^2)}{m_2(l, X_l^2)} \right| dl \right] = \int_h^s \int_{\mathbb{T}^d} |D\zeta(l, x)| dx dl \leq \mathbf{d}(m_1^0, m_2^0) \int_h^s K_l dl.$$

If we plug the last inequality into (1.30), we can use again Gronwall's inequality so that for  $s \in [h, h + \tau]$

$$\mathbf{d}(m_1(s), m_2(s)) \leq \left( e^{Cs} + \frac{e^{s-h}}{\tau} \int_h^s K_l dl \right) \mathbf{d}(m_1^0, m_2^0). \quad (1.31)$$

We can now suppose  $h = \tau = 1$  and plugging (1.28) and (1.31) into (1.26), we finally get that, for a given constant  $C$  depending on all the other ones

$$\mathcal{V}_\delta(m_2^0) - \mathcal{V}_\delta(m_1^0) \leq C \mathbf{d}(m_1^0, m_2^0) \left( \int_0^2 e^{Cs} ds + \int_1^2 e^{s-1} \int_1^s K_l dl ds + e^{5Q} \int_1^2 K_s^2 ds \right).$$

We recall that the constant  $K_s$  of Lemma 1.1.4 is bounded when  $h$  is not close to 0 (Theorem 11.1 in [60]). The infimum in the expression above is finite and none of the constants therein depends on  $\delta$ . Therefore,  $\{\mathcal{V}_\delta\}_\delta$  is uniformly  $\bar{K}$ -Lipschitz with

$$\bar{K} = C \left( \int_0^2 e^{Cs} ds + \int_1^2 e^{s-1} \int_1^s K_l dl ds + e^{5Q} \int_1^2 K_s^2 ds \right).$$

□

**Proposition 1.1.6.** *For any fixed  $\eta \in \mathcal{P}(\mathbb{T}^d)$  there exists a subsequence  $\delta_n \rightarrow 0$ , such that  $\mathcal{V}_{\delta_n}(\cdot) - \mathcal{V}_{\delta_n}(\eta)$  uniformly converges to a function  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  when  $n \rightarrow +\infty$ . Moreover,  $\delta_n \mathcal{V}_{\delta_n}$  uniformly converges to a constant  $-\lambda$*

*Proof.* The continuity proved in Theorem 1.1.5 ensures the boundedness of  $\mathcal{V}_\delta(\cdot) - \mathcal{V}_\delta(\eta)$ . Indeed we have  $|\mathcal{V}_\delta(\cdot) - \mathcal{V}_\delta(\eta)| \leq \bar{K} \mathbf{d}(\cdot, \eta)$ . As  $\mathcal{P}(\mathbb{T}^d)$  is compact, the right hand side is bounded by a constant  $K$ . Arzelà-Ascoli theorem ensures that there exists a subsequence  $\delta_n \rightarrow 0$  such that  $\mathcal{V}_{\delta_n}(\cdot) - \mathcal{V}_{\delta_n}(\eta)$  converges to a continuous function  $\chi$ .

We now want to prove that  $\delta \mathcal{V}_\delta$  is a bounded function. We fix a measure  $\mu \equiv 1$ , then we define the control  $(m, \alpha)$  as follows:  $m(t) = \mu$  and  $\alpha(t) = 0$  for all  $t \in [0, +\infty)$ . The control is admissible, therefore we have

$$\delta \mathcal{V}_\delta(\mu) \leq \delta(H^*(x, 0) + \mathcal{F}(\mu)) \int_0^\infty e^{-\delta s} ds = H^*(x, 0) + \mathcal{F}(\mu).$$

Given that  $H^*$  and  $\mathcal{F}$  are bounded from below, then

$$\delta \mathcal{V}_\delta(\mu) \geq \delta \left( \inf_{(x,a) \in \mathbb{T}^d \times \mathbb{R}^d} H^*(x, a) + \inf_{m \in \mathcal{P}(\mathbb{T}^d)} \mathcal{F}(m) \right) \int_0^\infty e^{-\delta s} ds = \inf_{(x,a)} H^*(x, a) + \inf_m \mathcal{F}(m).$$

Therefore,  $\delta \mathcal{V}_\delta(\mu)$  is uniformly bounded in  $\delta$ . If we fix any other measure  $m_0$  we can use again the uniform continuity of  $\mathcal{V}_\delta$  to get that  $|\delta \mathcal{V}_\delta(m_0) - \delta \mathcal{V}_\delta(\mu)| \leq \delta K$  that in turn tells us that  $\delta \mathcal{V}_\delta(\cdot)$  is a sequence of uniformly continuous functions. Using again Arzelà-Ascoli theorem we get that  $\delta_n \mathcal{V}_{\delta_n}$  uniformly converges to a function  $\Psi$  (we can suppose  $\delta_n$  to be the same subsequence that we identified earlier). Moreover, we have  $|\delta_n \mathcal{V}_{\delta_n}(\cdot) - \delta_n \mathcal{V}_{\delta_n}(\mu)| \leq \delta_n K$ . Taking the limit we get  $|\Psi(\cdot) - \Psi(\mu)| \leq 0$  so that  $\delta_n \mathcal{V}_{\delta_n}$  converges to the constant function  $\Psi(\mu) := -\lambda$ .

□

**Lemma 1.1.7.** *Dynamic programming principle for  $\chi$ : for any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and  $t > 0$ ,*

$$\chi(m_0) = \inf_{(m, \alpha)} \left( \int_0^t H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t \quad (1.32)$$

where  $m \in C^0([0, t], \mathcal{P}(\mathbb{T}^d))$ ,  $\alpha \in L_m^2([0, t] \times \mathbb{T}^d, \mathbb{R}^d)$  and the pair  $(m, \alpha)$  solves in the sense of distributions  $-\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0$  with initial condition  $m_0$ .

*Proof.* In Proposition 1.1.6 we proved the convergence of  $\mathcal{V}_\delta(\cdot) - \mathcal{V}_\delta(\eta)$  to  $\chi(\cdot)$  along the subsequence  $\{\delta_n\}_n$ , for a fixed measure  $\eta$ . Hereafter,  $\{\delta_n\}_n$  and  $\eta$  will be the ones identified in that proposition.

We know from Proposition 1.1.1 that, for any  $\delta > 0$ , there exists a solution  $(u_\delta, m_\delta)$  to (1.7) such that

$$\mathcal{V}_\delta(m_0) = \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_\delta) dm_\delta(s) + \mathcal{F}(m_\delta(s)) ds + e^{-\delta t} V_\delta(m_\delta(t)),$$

where  $\alpha_\delta = D_p H(x, Du_\delta)$ . If we take the expansion of  $e^{-\delta t}$  and we subtract on both sides  $\mathcal{V}_\delta(\eta)$  we get

$$\mathcal{V}_\delta(m_0) - \mathcal{V}_\delta(\eta) = \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_\delta) dm_\delta(s) + \mathcal{F}(m_\delta(s)) ds + (1 - t\delta + o(t\delta)) V_\delta(m_\delta(t)) - V_\delta(\eta). \quad (1.33)$$

We recall that the estimates in Lemma 1.1.3 are uniform in  $\delta$ . Then,  $D_p H(x, Du_{\delta_n})$  converges uniformly to a function  $\alpha$  and, as  $m_{\delta_n}$  is uniformly bounded in  $C^{1/2}([0, t], \mathcal{P}(\mathbb{T}^d))$ , it uniformly converges to a time dependent measure  $m$ . We can now take the limit  $n \rightarrow +\infty$  and, using that  $\mathcal{V}_{\delta_n}(\cdot) - \mathcal{V}_{\delta_n}(\eta) \rightarrow \chi(\cdot)$  and  $\delta_n \mathcal{V}_{\delta_n}(\cdot) \rightarrow -\lambda$ , we get

$$\chi(m_0) = \int_0^t \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + \lambda t + \chi(m_t)$$

In order to show that  $(\alpha, m)$  is optimal, we fix a competitor  $(\beta, \mu)$ . According to the dynamic programming principle of  $\mathcal{V}_\delta$ , if we plug  $(\beta, \mu)$  into (1.33), we get

$$\mathcal{V}_\delta(m_0) - \mathcal{V}_\delta(\eta) \leq \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \beta) d\mu(s) + \mathcal{F}(\mu(s)) ds + (1 - t\delta + o(t\delta)) V_\delta(\mu(t)) - V_\delta(\eta).$$

Taking again the limit on the subsequence  $\{\delta_n\}_n$  we eventually have that

$$\chi(m_0) \leq \int_0^t \int_{\mathbb{T}^d} H^*(x, \beta) d\mu(s) + \mathcal{F}(\mu(s)) ds + \lambda t + \chi(\mu_t),$$

which proves the result. □

### 1.1.3 Convergence of $\mathcal{U}(t, \cdot)/t$ and $\delta \mathcal{V}_\delta(\cdot)$

In this section we establish a Tauberian-type result where we prove that the limit of  $\delta \mathcal{V}_\delta(\cdot)$  coincides with the one of  $\mathcal{U}(t, \cdot)/t$  when  $t \rightarrow +\infty$ .

**Theorem 1.1.8.** *The limit value  $-\lambda$  is uniquely defined and  $\delta \mathcal{V}_\delta(\cdot) \rightarrow -\lambda$  does not depend on a subsequence. Moreover,  $\frac{1}{T} \mathcal{U}(T, \cdot)$  uniformly converges to  $-\lambda$  when  $T$  goes to  $+\infty$ .*

*Proof.* Let  $\{\delta_n\}_n$  such that  $\delta_n \mathcal{V}_{\delta_n} \rightarrow -\lambda$  and  $\mathcal{V}_{\delta_n}(\cdot) - \mathcal{V}_{\delta_n}(\eta) \rightarrow \chi(\cdot)$ . As  $\chi$  is a continuous function on the compact set  $\mathcal{P}(\mathbb{T}^d)$ , there exists a constant  $C > 0$  such that  $0 \leq \chi(m) + C$  for any  $m \in \mathcal{P}(\mathbb{T}^d)$ . If  $(m(t), w(t)) \in \mathcal{E}_2^T(m_0)$ , then

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt \\ & \leq \int_0^T \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt + \chi(m(T)) + \lambda T - \lambda T + C. \end{aligned}$$

Taking the infimum over  $\mathcal{E}_2^T$ , the definition of  $\mathcal{U}(T, m_0)$  and the dynamic programming principle of  $\chi$  lead to

$$\mathcal{U}(T, m_0) \leq \chi(m_0) - \lambda T + C.$$

As the constant  $C$  does not depend on  $m_0$  and  $T$ , if we divide on both sides by  $T$  and we take the limit  $T \rightarrow +\infty$ , we get

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \mathcal{U}(T, m_0) \leq -\lambda.$$

The other inequality is analogous. We just need to take a  $C_2 > 0$  such that  $0 \geq \chi(m) - C_2$  and repeat the same computation.

Note that the limit  $\mathcal{U}(T, \cdot)/T \rightarrow -\lambda$  is uniform and does not depend on the subsequence  $\delta_n$  or the function  $\chi$ . Therefore, the limit  $\delta \mathcal{V}_\delta$  is uniquely defined.  $\square$

We conclude the section showing that our limit value  $\lambda$  is never lower than the ergodic one  $\bar{\lambda}$  defined in (1.5).

**Proposition 1.1.9.** *Under the above assumptions,  $\lambda \geq \bar{\lambda}$ .*

*Proof.* We know that the convergence of  $\mathcal{U}(\cdot, T)/T$  is uniform, therefore, if  $(m, \alpha)$  is an admissible couple for the static problem, we can use it as competitor for  $\mathcal{U}(m, T)$ . So,

$$\frac{1}{T} \mathcal{U}(T, m) \leq \frac{1}{T} \int_0^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm + \mathcal{F}(m) dt = \int_{\mathbb{T}^d} H^*(x, \alpha) dm + \mathcal{F}(m).$$

If we take the infimum over all admissible static  $(m, \alpha)$  we get

$$\frac{1}{T} \mathcal{U}(T, m) \leq -\bar{\lambda}.$$

Letting  $T$  go to  $+\infty$ , we get the result.  $\square$

One of the main goals of this paper is to understand under which assumptions the inequality above is an equality and when, instead, it is a strict inequality. This latter case is the one we are most interested in and it is addressed in Section 1.3.

#### 1.1.4 Another representation for $\lambda$

We can now introduce a third representation for  $\lambda$ , inspired again by classic results on weak KAM theory (see for instance [39]), which consists in minimizing over paths with fixed endpoints.

Let  $\Pi_T(m_0, m_1)$  be the set of  $(m, \alpha) \in C^0([0, T], \mathcal{P}(\mathbb{T}^d)) \times L_m^2([0, T] \times \mathbb{T}^d, \mathbb{R}^d)$  such that  $(m, \alpha)$  solves the usual Fokker-Plank equation  $-\partial_t m + \Delta m - \operatorname{div}(m\alpha) = 0$  with the extra constraint  $m(0) = m_0$  and  $m(T) = m_1$ . Note that, due to the smoothing property of the parabolic constraint, not for every  $m_1$  we can find such a path, so  $\Pi_T(m_0, m_1)$  might be the empty set. In particular, the  $L_m^2$  integrability of the drift  $\alpha$  prevents the target measure  $m_1$  from being too singular. For instance if  $m_1$  is a Dirac delta, then  $\Pi_T(m_0, m_1)$  is the empty set.

**Proposition 1.1.10.** *Let  $m_0, m_1 \in \mathcal{P}(\mathbb{T}^d)$ . If  $m_1$  has a density in  $H^1(\mathbb{T}^d)$  and there exists an  $\varepsilon > 0$  such that  $m_1 > \varepsilon$  almost everywhere then*

$$-\lambda = \lim_{T \rightarrow \infty} \frac{1}{T} \inf_{\Pi_T(m_0, m_1)} \int_0^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds.$$

*Proof.* Let  $m_0$  and  $m_1$  be as above and  $(\bar{m}, \bar{\alpha})$  be optimal for  $\mathcal{U}(T, m_0)$ . We extend  $(\bar{m}, \bar{\alpha})$  in  $[0, T+1]$  as follows: for any  $t \in [T, T+1]$  we define  $\bar{\alpha}(t, x) = \bar{\alpha}(T, x)$  and  $\bar{m}(t, x)$  as the solution of  $-\partial_t \bar{m} + \Delta \bar{m} + \operatorname{div}(\bar{\alpha} \bar{m}) = 0$  with  $\bar{m}(T, x) = \bar{m}(T, x)$ . Note that  $\bar{\alpha}$  is continuous and bounded in  $[0, T+1]$ , therefore, the estimates (1.25) still apply.

We now define a path from  $m_0$  to  $m_1$  as follows:

$$m_2(s, x) = \begin{cases} \bar{m}(s, x) & s \in [0, T] \\ (T+1-s)\bar{m}(s, x) + (s-T)m_1(x) & s \in [T, T+1] \end{cases}.$$

Let also  $\zeta(s, x)$  be solution of  $-\Delta \zeta(s, x) = m_1(x) - \bar{m}(s, x)$  with  $\int_{\mathbb{T}^d} \zeta = 0$ . We can define the control

$$\alpha_2(s, x) = \begin{cases} \bar{\alpha}(s, x) & s \in [0, T] \\ \bar{\alpha}(T, x) - \frac{(s-T)\bar{\alpha}(T, x)m_1(x) + D\zeta(s, x) + (s-T)Dm_1(x)}{(T+1-s)\bar{m}(s, x) + (s-T)m_1(x)} & s \in [T, T+1]. \end{cases}$$

The couple  $(m_2, \alpha_2)$  belongs to  $\Pi_T(m_0, m_1)$ . From the definition of  $\mathcal{U}$  we deduce that

$$\begin{aligned} \frac{1}{T+1} \inf_{m \in \mathcal{P}(\mathbb{T}^d)} \mathcal{U}(T+1, m) &\leq \frac{1}{T+1} \inf_{\Pi_{T+1}(m_0, m_1)} \int_0^{T+1} \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds \\ &\leq \frac{1}{T+1} \int_0^{T+1} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(s) + \mathcal{F}(m_2(s)) ds \\ &= \frac{T}{T+1} \left( \frac{1}{T} \mathcal{U}(T, m_0) \right) + \frac{1}{T+1} \int_T^{T+1} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(s) + \mathcal{F}(m_2(s)) ds. \end{aligned} \quad (1.34)$$

If we prove that  $\frac{1}{T+1} \int_T^{T+1} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(s) + \mathcal{F}(m_2(s)) ds$  converges to zero we have the result. Indeed, if we let  $T$  go to  $+\infty$ , according to Theorem 1.1.8, we have

$$-\lambda \leq \lim_{T \rightarrow \infty} \frac{1}{T+1} \inf_{\Pi_{T+1}(m_0, m_1)} \int_0^{T+1} \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds \leq -\lambda.$$

We now focus on the last part in (1.34). Given that  $\mathcal{F}(m_2)$  is uniformly bounded, we look at the first term.

$$\begin{aligned} &\int_T^{T+1} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(s) \\ &\leq C \int_T^{T+1} \int_{\mathbb{T}^d} \frac{|\bar{\alpha}(T, x)m_2(s) + (s-T)\bar{\alpha}(T, x)m_1(s) + D\zeta(s, x) + (s-T)Dm_1(s)|^2}{m_2^2(s)} + 1 dm_2(s) ds \\ &\leq C \int_T^{T+1} \int_{\mathbb{T}^d} \frac{(|\bar{\alpha}(T, x)m_2| + |\bar{\alpha}(T, x)m_1| + |D\zeta(s, x)| + |Dm_1|)^2}{m_2} dx ds + C \end{aligned} \quad (1.35)$$

If we use the hypothesis on  $m_1$  and the estimates (1.25) on  $\bar{m}$  with  $h = T+1$  and  $l = T-1$ , we get that  $m_2 \geq \tau$  for a certain  $\tau > 0$  independent of  $T$ . Lemma 1.1.3 ensures that  $\bar{\alpha}$  is uniformly bounded by a constant  $K$  independent of  $T$ . Therefore, (1.35) is lower than

$$\frac{1}{\tau} \left( K \|m_2\|_{L^2(\mathbb{T}^d \times [T, T+1])} + K \|m_1\|_{L^2(\mathbb{T}^d)} + \|D\zeta\|_{L^2(\mathbb{T}^d \times [T, T+1])} + \|Dm_1\|_{L^2(\mathbb{T}^d)} \right)^2 + C.$$

Thanks again to the boundedness of  $\bar{\alpha}$ , standard result on parabolic equations tell us that  $\bar{m}(s)$  (which is defined at the beginning of the proof) is uniformly bounded from above in  $[T, T + 1]$ . Hence,  $\|D\zeta\|_{L^2} \leq C\|m_1\|_{L^2}$  and  $\|m_2\|_{L^2(\mathbb{T}^d \times [T, T+1])} \leq C\|m_1\|_{L^2(\mathbb{T}^d)} + C_2$ . Thus

$$\int_T^{T+1} \int_{\mathbb{T}^d} H^*(x, \alpha_2) dm_2(s) \leq M\|m_1\|_{H^1(\mathbb{T}^d)}^2 + M_2$$

where neither  $M$  nor  $M_2$  depends on  $T$ . Dividing by  $T + 1$  and taking the limit completes the proof.  $\square$

## 1.2 Projected Mather set and Calibrated curves

### 1.2.1 Calibrated Curves

We borrow again some tools and some notations from the weak KAM theory (see Chapter 4 of [42]) and in particular we will focus on the notion of calibrated curve. Before introducing this notion, we look back to the dynamic programming principle verified by corrector functions, which reads

$$\chi(m_0) = \inf_{(m,w) \in \mathcal{E}_2^t(m_0)} \left( \int_0^t H^* \left( x, -\frac{w}{m}(s) \right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t.$$

As the function  $\chi$  is continuous, standard arguments show that, for any fixed  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and  $t > 0$ , there exists a solution  $(m_1, w_1) \in \mathcal{E}_2^t(m_0)$  to the minimization problem described above. It is easy to construct a new trajectory  $(\bar{m}, \bar{w})$ , defined on  $[0, +\infty)$  such that  $(\bar{m}, \bar{w}) = (m_1, w_1)$  on  $[0, T]$  and for any  $\tau > 0$ , it verifies

$$\chi(m_0) = \lambda \tau + \int_0^\tau \int_{\mathbb{T}^d} H^* \left( x, -\frac{\bar{w}(s)}{\bar{m}(s)} \right) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + \chi(\bar{m}(\tau)).$$

Indeed, you just need to attach to  $(m_1, w_1)$  an optimal trajectory for  $\chi(m_1(t))$  on the interval  $[t, 2t]$ . You end up with a new trajectory  $(m_2, w_2)$ , defined on  $[0, 2t]$ , which is optimal for  $\chi(m_1)$ . We now repeat the construction attaching to  $(m_2, w_2)$  a new branch which is optimal for  $\chi(m_1(2t))$  on  $[2t, 3t]$  so that now,  $(m_3, w_3)$  is defined on  $[0, 3t]$  and is still optimal for  $\chi(m_0)$ . Iterating, we can extend  $(m_1, w_1)$  to any interval  $[0, nt]$  and eventually, at the limit, we get the trajectory  $(\bar{m}, \bar{w})$ , defined on  $[0, +\infty)$ , that we were looking for.

We now prove that any of these trajectories is associated to a MFG system.

**Proposition 1.2.1.** *Let  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,  $\chi$  be a corrector function and  $(\bar{m}, \bar{w})$  be a minimizing trajectory on  $[0, +\infty)$  defined as above. Then,  $\bar{m} \in C^{1,2}([0, +\infty) \times \mathbb{T}^d)$  and there exists a function  $\bar{u} \in C^{1,2}([0, +\infty) \times \mathbb{T}^d)$  such that  $\bar{w} = -\bar{m}D_p H(x, D\bar{u})$  where  $(\bar{m}, \bar{u})$  solves*

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } [0, +\infty) \times \mathbb{T}^d \\ -\partial_t m + \Delta m + \operatorname{div}(mD_p H(x, Du)) = 0 & \text{in } [0, +\infty) \times \mathbb{T}^d \\ m(0) = m_0. \end{cases}$$

**Remark 1.2.2.** *Due to the lack of regularity of  $\chi$  we can not derive the MFG system as optimal condition for the minimization problem (1.36). Indeed, if  $\chi$  were  $C^1$  we would derive typical MFG system with terminal condition  $u(t) = \delta\chi(m(t))/\delta m$  but this latter term is not well defined.*

*Proof.* The proof relies on the same arguments as in Proposition 1.1.1. Let  $(\bar{m}, \bar{w})$  be as in the hypothesis. Then it verifies the Fokker-Plank equation and it is a minimizer of the problem

$$\inf_{(m,w) \in \mathcal{E}_2^t(m_0)} \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)). \quad (1.36)$$

As  $(\bar{m}, \bar{w})$  is optimal for the minimization problem above, then it must be also optimal for the following MFG planning problem

$$\inf_{(m,w) \in \Pi(m_0, \bar{m}(t))} \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \mathcal{F}(m(s)) ds,$$

where  $\Pi(m_0, \bar{m}(t))$  is the set of  $(m, w) \in \mathcal{E}_2^t(m_0)$  that solves the usual Fokker-Plank equation on  $[0, t]$  with the constraints  $m(0) = m_0$  and  $m(t) = \bar{m}(t)$ . We want to prove that  $\bar{w} = -\bar{m} D_p H(x, D\bar{u})$  where  $(\bar{m}, \bar{u})$  solves in classical sense

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } [0, t] \times \mathbb{T}^d \\ \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } [0, t] \times \mathbb{T}^d \\ m(0) = m_0, m(t) = \bar{m}(t). \end{cases}$$

We argue again as in Proposition 1.1.1. According to Proposition 3.1 in [11],  $(\bar{m}, \bar{w})$  minimizes also the following convex problem

$$\inf_{(m,w) \in \Pi(m_0, \bar{m}(t))} \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) dm(s) ds. \quad (1.37)$$

This problem admits a dual formulation, in the sense of Fenchel Rockafellar Theorem, which reads

$$\inf_{\psi \in \bar{\mathcal{K}}} \left\{ \int_{\mathbb{T}^d} \psi(x, t) d\bar{m}(t) - \int_{\mathbb{T}^d} \psi(x, 0) dm_0 \right\}$$

where  $\bar{\mathcal{K}}$  is the set of  $\psi \in C^{1,2}([0, t] \times \mathbb{T}^d)$  such that  $-\partial_t \psi - \Delta \psi + H(x, D\psi) \leq F(x, \bar{m})$ . A full justification of the result above can be found again in [21].

In the definition of the dual problem we can replace  $\bar{\mathcal{K}}$  with  $\mathcal{K}$ , where  $\mathcal{K}$  is the set of  $u \in C^{1,2}([0, t] \times \mathbb{T}^d)$  such that  $-\partial_t u - \Delta u + H(x, Du) = F(x, \bar{m})$ . Indeed, if  $\psi$  verifies  $-\partial_t \psi - \Delta \psi + H(x, D\psi) \leq F(x, \bar{m})$ , we can always consider  $u \in C^{1,2}([0, t] \times \mathbb{T}^d)$  solution of

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, \bar{m}) & \text{in } [0, t] \times \mathbb{T}^d \\ u(x, t) = \psi(x, t) & \text{in } \mathbb{T}^d \end{cases}$$

Thanks to the comparison principle we have that  $u(0, x) \geq \psi(0, x)$ , thus

$$\inf_{\psi \in \bar{\mathcal{K}}} \left\{ \int_{\mathbb{T}^d} \psi(x, 0) d\bar{m}(t) - \int_{\mathbb{T}^d} \psi(x, 0) dm_0 \right\} \geq \inf_{u \in \mathcal{K}} \left\{ \int_{\mathbb{T}^d} u(x, t) d\bar{m}(t) - \int_{\mathbb{T}^d} u(x, 0) dm_0 \right\} \quad (1.38)$$

The opposite inequality holds by inclusion. Lemma 3.2 in [21] and Proposition 3.1 in [11], which rely on the Fenchel-Rockafellar Theorem, ensure that, if  $(\bar{m}, \bar{w})$  is a minimizer of (1.37) and  $\bar{u} \in \mathcal{K}$  is a minimizer of the dual problem, then

$$\int_{\mathbb{T}^d} u(x, t) d\bar{m}(t) - \int_{\mathbb{T}^d} u(x, 0) dm_0 + \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{\bar{w}}{\bar{m}} \right) d\bar{m}(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) d\bar{m}(s) ds = 0.$$

This implies that  $\bar{w} = -\bar{m}D_pH(x, \bar{u})$ . As a consequence, we have that  $\bar{m}$  is driven by a smooth drift and so, by Schauder theory,  $\bar{m} \in C^{1,2}((0, t] \times \mathbb{T}^d)$ . In particular, given that  $t$  is arbitrary, then  $\bar{m} \in C^{1,2}((0, +\infty) \times \mathbb{T}^d)$ .

We assumed that the minimization problem on the right-hand side of (1.38) admitted a solution. It is indeed the case and the proof of this result is developed in Lemma 1.4.2 in appendix.  $\square$

**Remark 1.2.3.** *Note that the convex duality used in Proposition 1.2.1 between (1.37) and (1.38), i.e*

$$\min_{(m,w) \in \Pi(m_0, \bar{m}(t))} \int_0^t \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) dm(s) ds =$$

$$\inf_{u \in \mathcal{K}} \left\{ \int_{\mathbb{T}^d} u(x, t) d\bar{m}(t) - \int_{\mathbb{T}^d} u(x, 0) dm_0 \right\},$$

holds true independently from the existence of minimizers for the latter one and, therefore, independently from Lemma 1.4.2.

We can now introduce the notion of calibrated curve. Let  $\mathcal{E}_2^\infty$  be the set of  $(m(t), w(t)) \in \mathcal{P}(\mathbb{T}^d) \times \mathcal{M}(\mathbb{T}^d; \mathbb{R}^d)$  such that  $m \in C^0(\mathbb{R}, \mathcal{P}(\mathbb{T}^d))$ ,  $w$  is absolutely continuous with respect to  $m$ , its density  $dw/dm$  belongs to  $L^2_{m,loc}(\mathbb{R} \times \mathbb{T}^d)$  and  $-\partial_t m + \Delta m - \operatorname{div} w = 0$  is verified in the sense of distributions.

**Definition 1.2.4.** *We say that  $(\bar{m}, \bar{w}) \in \mathcal{E}_2^\infty$  is a calibrated curve if there exists a continuous function  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  which verifies the dynamic programming principle (1.32) and  $(\bar{m}, \bar{w})$  is optimal for  $\chi$ : for any  $t_1 < t_2 \in \mathbb{R}$*

$$\chi(\bar{m}(t_1)) = \lambda(t_2 - t_1) + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{\bar{w}(s)}{\bar{m}(s)} \right) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + \chi(\bar{m}(t_2)).$$

A direct consequence of Proposition 1.2.1 is the following result which tells that calibrated curves are smooth and associated to MFG systems defined for any time  $t \in \mathbb{R}$ .

**Proposition 1.2.5.** *If  $(m, w) \in \mathcal{E}_2^\infty$  is a calibrated curve, then  $m \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$  and there exists a function  $u \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$  such that  $w = -mD_pH(x, Du)$  where  $(m, u)$  solves*

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{R} \times \mathbb{T}^d \\ -\partial_t m + \Delta m + \operatorname{div}(mD_pH(x, Du)) = 0 & \text{in } \mathbb{R} \times \mathbb{T}^d \end{cases}$$

## 1.2.2 The projected Mather set

**Definition 1.2.6.** *We say that  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  belongs to the projected Mather set  $\mathcal{M} \subset \mathcal{P}(\mathbb{T}^d)$  if there exists a calibrated curve  $(m(t), w(t))$  such that  $m(0) = m_0$ .*

Note that, if from  $m_0$  starts a calibrated curve  $m(t)$ , then, by translation,  $m(t) \in \mathcal{M}$  for any  $t \in \mathbb{R}$ .

**Proposition 1.2.7.** *There exists a calibrated curve and, consequently, the projected Mather set  $\mathcal{M}$  is not empty.*



*Proof.* We fix a smooth density  $m_0$  and we look at the  $\delta$ -discounted problem (1.4) which reads

$$\mathcal{V}_\delta(m_0) = \inf_{(m,w)} \int_0^\infty e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(t) + \mathcal{F}(m(t)) dt.$$

We recall that  $\mathcal{V}_\delta$  satisfies the dynamic programming principle

$$\mathcal{V}_\delta(m_0) = \inf_{m,w} \int_0^T e^{-\delta s} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta T} \mathcal{V}_\delta(m(T)),$$

where the infimum is taken over  $(m, w) \in \mathcal{E}_2^\delta(m_0)$ . We already know that the solution of the minimization problem corresponds to a couple  $(\bar{m}_\delta^T, -\bar{m}_\delta^T D_p H(x, D\bar{u}_\delta^T))$  where  $(\bar{m}_\delta^T, \bar{u}_\delta^T)$  solves

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, +\infty) \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, +\infty) \\ m(0) = m_0 & \text{in } \mathbb{T}^d. \end{cases} \quad (1.39)$$

Note that, as the initial condition is smooth, the solution  $(\bar{m}_\delta^T, \bar{u}_\delta^T)$  is smooth as well.

We define the new couple  $(m_\delta^T, w_\delta^T)$  as  $m_\delta^T(t, x) = \bar{m}_\delta^T(t+T, x)$  and  $w_\delta^T(t, x) = \bar{w}_\delta^T(t+T, x)$  so that our problem is set on  $[-T, +\infty)$ . We now want to prove that, when we take the limit  $T \rightarrow +\infty$ , our sequence converges to a couple  $(m_\delta, w_\delta)$  defined on  $\mathbb{R} \times \mathbb{T}^d$  such that the Fokker-Plank equation is still verified. We proved in Lemma 1.1.3 that the drift  $D_p H(x, Du_\delta^T)$  is uniformly bounded in  $T$ , therefore,  $m_\delta^T$  is the solution of a Fokker-Plank equation with bounded and smooth drift. This means that  $m_\delta^T$  is uniformly bounded in  $C^{1,2}([-T+1, +\infty) \times \mathbb{T}^d)$ .

This implies that, at least on compact subsets of  $\mathbb{R} \times \mathbb{T}^d$ , when we take the limit  $T \rightarrow +\infty$ , we have, up to a subsequence, uniform convergence of  $m_\delta^T$  to a limit  $m_\delta$ .

The same convergence holds true also for  $w_\delta^T$ . Indeed, in Lemma 1.1.3 we proved also the uniform boundedness of  $D^2 u_\delta^T$  and  $\partial_t u_\delta^T$  that implies the uniform continuity and the uniform boundedness of  $w_\delta^T$ . The convergence  $(m_\delta^T, w_\delta^T)$  to  $(m_\delta, w_\delta)$  ensures that the couple  $(m_\delta, w_\delta)$  verifies the Fokker-Plank equation on  $\mathbb{R} \times \mathbb{T}^d$ .

We fix two different times  $t_1 < t_2$ . For sufficiently large  $T$ , the interval  $[t_1, t_2]$  is included in  $[-T, +\infty)$ . If we apply the dynamic programming principle for  $\mathcal{V}_\delta$ , we get

$$\mathcal{V}_\delta(m_\delta^T(t_1)) = \int_{t_1}^{t_2} e^{-\delta(s-t_1)} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w_\delta^T}{m_\delta^T} \right) dm_\delta^T(s) + \mathcal{F}(m_\delta^T(s)) ds + e^{-\delta(t_2-t_1)} \mathcal{V}_\delta(m_\delta^T(t_2)).$$

We can now take the limit of  $T \rightarrow +\infty$  in the above expression and we find that  $(m_\delta, w_\delta)$  verifies

$$\mathcal{V}_\delta(m_\delta(t_1)) = \int_{t_1}^{t_2} e^{-\delta(s-t_1)} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w_\delta}{m_\delta} \right) dm_\delta(s) + \mathcal{F}(m_\delta(s)) ds + e^{-\delta(t_2-t_1)} \mathcal{V}_\delta(m_\delta(t_2)).$$

for any  $t_1 < t_2 \in \mathbb{R}$ .

As the function  $u_\delta^T$  is uniformly bounded in  $T$  we have also uniform convergence of  $u_\delta^T$  to a function  $u_\delta$ . We can then pass to the limit in the MFG system (1.39) and the couple  $(u_\delta, m_\delta)$  solves

$$\begin{cases} -\partial_t u_\delta - \Delta u_\delta + \delta u_\delta + H(x, Du_\delta) = F(x, m_\delta) & \text{in } \mathbb{T}^d \times \mathbb{R} \\ -\partial_t m_\delta + \Delta m_\delta + \operatorname{div}(m_\delta D_p H(x, Du_\delta)) = 0 & \text{in } \mathbb{T}^d \times \mathbb{R}. \end{cases}$$

As in [5], in order to let  $\delta \rightarrow 0$ , we need to define  $\bar{u}_\delta(t, x) = u_\delta(t, x) - u_\delta(0, 0)$  and  $\theta_\delta = u_\delta(0, 0)$ . The couple  $(\bar{u}_\delta, m_\delta)$  solves

$$\begin{cases} -\partial_t \bar{u}_\delta - \Delta \bar{u}_\delta + \delta \bar{u}_\delta + \delta \theta_\delta + H(x, D\bar{u}_\delta) = F(x, m_\delta) & \text{in } \mathbb{T}^d \times \mathbb{R} \\ -\partial_t m_\delta + \Delta m_\delta + \operatorname{div}(m_\delta D_p H(x, D\bar{u}_\delta)) = 0 & \text{in } \mathbb{T}^d \times \mathbb{R} \\ \bar{u}_\delta(0, x) = u_\delta(0, x) - u_\delta(0, 0) & \text{in } \mathbb{T}^d. \end{cases}$$

We restrict ourselves to the subsequence  $\{\delta_n\}_n$  identified in the proof of Lemma 1.1.7. Using again the uniform estimates on  $D\bar{u}_\delta$ , we have that  $\bar{u}_\delta(0, x)$  is uniformly bounded which implies the boundedness of  $\delta\theta_\delta$ . Moreover, thanks to the bounds on  $D^2 u_\delta$  and  $\partial_t u_\delta$ ,  $\bar{u}_\delta$  is also uniformly continuous and the same holds true for  $m_\delta$ . We can then pass to the limit on any compact set and  $\bar{u}_{\delta_n} \rightarrow u$ ,  $m_{\delta_n} \rightarrow m$  and  $\delta_n \theta_{\delta_n} \rightarrow \theta$  where  $(u, m, \theta)$  solves

$$\begin{cases} -\partial_t u - \Delta u + \theta + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times \mathbb{R} \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times \mathbb{R} \end{cases}$$

As we can always replace  $u(t, x)$  with  $u(t, x) - \theta t$  we can suppose  $\theta = 0$ . The convergences above give us also the uniform convergence on compact sets (up to subsequence) of the couple  $(m_{\delta_n}, w_{\delta_n}) = (m_{\delta_n}, -m_{\delta_n} D_p H(x, D\bar{u}_{\delta_n}))$  to  $(m, w) = (m, -m D_p H(x, Du))$  which solves the usual Fokker-Plank equation.

Let now  $\eta \in \mathcal{P}(\mathbb{T}^d)$  be the measure identified in the proof of Lemma 1.1.7. Then

$$\begin{aligned} & \mathcal{V}_{\delta_n}(m_{\delta_n}(t_1)) - \mathcal{V}_{\delta_n}(\eta) = \\ & \int_{t_1}^{t_2} e^{-\delta_n(s-t_1)} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w_{\delta_n}}{m_{\delta_n}} \right) dm_{\delta_n}(s) + \mathcal{F}(m_{\delta_n}(s)) ds + e^{-\delta_n(t_2-t_1)} \mathcal{V}_{\delta_n}(m_{\delta_n}(t_2)) - \mathcal{V}_{\delta_n}(\eta). \end{aligned}$$

Given the continuity of  $\mathcal{V}_\delta$ , the uniform convergence of  $m_\delta$  and  $w_\delta$ , we can pass to the limit in  $n$  and we finally get that for any interval  $[t_1, t_2]$  the couple  $(m, w)$  verifies the Fokker-Plank equation on  $\mathbb{R}$  and

$$\chi(m(t_1)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t_2)) + \lambda(t_2 - t_1).$$

In particular we found a calibrated curve and, for any  $t \in \mathbb{R}$ ,  $m(t)$  belongs to the projected Mather set  $\mathcal{M}$ . □

### 1.2.3 Compactness of the projected Mather set

In Proposition 1.2.1 we proved that, if  $\chi$  is a corrector function and  $(m, w)$  is a trajectory starting from  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  which is optimal for the dynamic programming principle of  $\chi$ , then  $(m, w)$  is associated to a MFG system which enjoys the estimates we proved in Lemma 1.1.3. Therefore, a completely analogous proof to the one given in Theorem 1.1.5 gives the following result.

**Proposition 1.2.8.** *The set of corrector functions is uniformly Lipschitz continuous.*

We can now prove the compactness of the projected Mather set  $\mathcal{M}$ .

**Proposition 1.2.9.** *The projected Mather set  $\mathcal{M}$  is compact*

*Proof.* Let  $m_n \in \mathcal{M}$  such that  $m_n^0 \rightarrow m_0$ . Let  $(m_n(t), w_n(t))$  be the calibrated curve starting from  $m_n^0$ . For any  $t_1, t_2$  we know that  $(m_n(t), w_n(t))$  verifies

$$\chi_n(m_n(t_1)) \geq \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w_n}{m_n} \right) dm_n(s) + \mathcal{F}(m_n(s)) + \lambda(t_2 - t_1) + \chi_n(m_n(t_2)). \quad (1.40)$$

We know from Proposition 1.2.8 that the set  $\{\chi_n\}_n$  is uniformly Lipschitz. If we replace  $\chi_n$  with  $\chi_n(\cdot) - \chi_n(\eta)$ , then  $\{\chi_n\}_n$  is also bounded and thus compact. Therefore, we can pick a subsequence such that  $\chi_n$  converges to a function  $\chi$ .

Given that  $\chi_n$  are uniformly bounded, then

$$\int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w_n}{m_n} \right) dm_n(s) + \mathcal{F}(m_n(s)) ds \leq C.$$

The constant  $C$  does not depend on  $n$  and, therefore,  $\int_{t_1}^{t_2} |w_n|$  is uniformly bounded as well. As we argued in Proposition 1.1.1, this implies that  $\{m_n\}_n$  is uniformly bounded in  $C^{1/2}([t_1, t_2], \mathcal{P}(\mathbb{T}^d))$  for any  $t_1, t_2$ . We have then that  $m_n$  converges uniformly on any compact set to a limit  $m \in C^0(\mathbb{R}, \mathcal{P}(\mathbb{T}^d))$  and the same holds true for  $w_n$  in  $\mathcal{M}(\mathbb{R} \times \mathbb{T}^d; \mathbb{R}^d)$ , therefore  $(m, w)$  solves in the sense of distributions the usual FP equation on  $\mathbb{R}$ . By weak lower semicontinuity of the integral part in (1.40) and the uniform convergence of  $\chi_n$  we get that

$$\chi(m(t_1)) \geq \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t_2)) + \lambda(t_2 - t_1)$$

with  $m(0) = m_0$  because  $m_n^0 \rightarrow m_0$ . The opposite inequality is true by the dynamic programming principle and so this proves that  $m_0 \in \mathcal{M}$  and, eventually, that  $\mathcal{M}$  is closed.  $\square$

#### 1.2.4 Minimal invariant set and Ergodicity

We say that a closed subset  $\mathcal{C}$  of  $\mathcal{M}$  is invariant if, for any  $m_0 \in \mathcal{C}$ , there exists a calibrated curve  $(m, w)$  such that  $m(0) = m_0$  and  $m(t) \in \mathcal{C}$  for any  $t \in \mathbb{R}$ . Note that, despite this notion of invariance formally plays the role of the invariance under the Euler-Lagrangian flow in the standard setting of weak KAM theory and Aubri-Mather theory in finite dimension, the forward-backward structure of the MFG system does not allow to define solutions with initial conditions for both the distribution of players  $m(t)$  and the value function  $u(t)$  and this prevents from defining any sensible notion of flow.

We say that an invariant set  $\mathcal{C}$  is minimal if  $\mathcal{C}$  does not contains any proper closed invariant subset.

**Lemma 1.2.10.** *There exists a minimal set  $\mathcal{N}$ .*

We do not present the proof which is a standard application of Zorn's Lemma (see for instance [36]).

**Proposition 1.2.11.** *Let  $\mathcal{N}$  be a minimal invariant set. If  $m_0 \in \mathcal{N}$  and  $\{m(t), t \in \mathbb{R}\}$  is a calibrated curve such that  $m(0) = m_0$ , then  $\{m(t), t \in \mathbb{R}\}$  is dense in  $\mathcal{N}$ .*

*Proof.* The proof is analogous to the one we used to prove that  $\mathcal{M}$  is closed. We define  $\mathcal{C}$  the closure of the trajectory  $\{m(t) \mid t \in \mathbb{R}\}$ .  $\mathcal{C}$  is a closed subset of  $\mathcal{N}$ , in order to prove that it coincides with  $\mathcal{N}$  we just need to prove that it is invariant or, in other words, that, if  $\bar{m} \in \mathcal{C}$ , then also a calibrated curve passing through  $\bar{m}$  belongs to  $\mathcal{C}$ .

Let  $\bar{m}$  be the limit of  $m_n = m(t_n) \in \{m(t), t \in \mathbb{R}\}$  and  $\{m_n(t)\}$  their corresponding calibrated curves. If  $w_n(t)$  are the control associated to the calibrated curve  $m_n(t)$  then, as in Proposition 1.2.9, we get that  $\|w_n\|_{L^1}$  is uniformly bounded on any compact set. As we already pointed out, it implies the uniform convergence of  $m_n(t)$  on compact sets. If  $\bar{m}(t)$  is the trajectory to which  $m_n(t)$  converges, then it must be a calibrated curve starting from  $\bar{m}$  because we imposed that  $m_n \rightarrow \bar{m}$ . This means that for any  $s \in \mathbb{R}$ ,  $\bar{m}(s)$  is the limit of  $m_n(s)$ . As  $\mathcal{C}$  is closed and  $m_n(s) \in \mathcal{C}$ , then  $\bar{m}(s) \in \mathcal{C}$ .

We proved that  $\mathcal{C}$  is a not empty, invariant, closed subset of  $\mathcal{N}$ . By the minimality of  $\mathcal{N}$  the two sets must coincide.  $\square$

### 1.3 The role of Monotonicity

So far, the hypothesis on  $\mathcal{F}$  were mostly about its regularity and no structural assumptions were imposed. On the other hand, when we are interested in understanding whether the limit value  $\lambda$  coincides with  $\bar{\lambda}$ , the ergodic one, the structure of  $\mathcal{F}$  does actually play a fundamental role. In the next section we impose convexity and, as it was already proved in [22], we get that  $\lambda = \bar{\lambda}$ . More interestingly, in Section 1.3.2 we provide a class of explicit examples where  $\lambda > \bar{\lambda}$  and, therefore, there is not convergence of the time dependent MFG system to the ergodic one.

#### 1.3.1 The convex case: $\lambda = \bar{\lambda}$

In this section we show that, under the so-called monotonicity assumption, these two values are the same. We say that the coupling function  $F$  verifies the monotonicity assumption if

$$\int_{\mathbb{T}^d} (F(x, m_1) - F(x, m_2)) d(m_1 - m_2) \geq 0. \quad (1.41)$$

We introduce the functional  $J^T(m_0, \cdot, \cdot) : \mathcal{E}^T(m_0) \rightarrow \mathbb{R}$  defined by

$$J^T(m_0, m, w) = \int_0^T \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t)) dt,$$

so that

$$\mathcal{U}(T, m_0) = \inf_{(m, w) \in \mathcal{E}_2^T(m_0)} J^T(m_0, m, w).$$

Under the monotonicity assumption (1.41), the functional  $J^T$  is convex, therefore we can easily prove that  $\lambda = \bar{\lambda}$ . We recall that

$$-\bar{\lambda} = \inf_{(m, w) \in \mathcal{E}} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw}{dm}(x) \right) dm(x) + \mathcal{F}(m).$$

**Proposition 1.3.1.** *Under the above assumptions  $\lambda = \bar{\lambda}$ .*

*Proof.* In order to prove the proposition we use the representation of  $\lambda$  that we discussed in Proposition 1.1.10:

$$-\lambda = \lim_{T \rightarrow \infty} \frac{1}{T} \inf_{\Pi_T(m_0, m_1)} \int_0^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds.$$

Given that  $\lambda$  does not depend on the initial value  $m_0$ , we take  $m_0 = m_1$ . We can also suppose that  $m_0$  is smooth and bounded from below by a positive constant, so that we can

apply Proposition 1.1.10. We now consider any admissible  $(m^T, w^T)$  for  $\Pi(m_0, m_0)$  and we define  $\bar{m}^T = \frac{1}{T} \int_0^T m^T dt$  and  $\bar{w}^T = \frac{1}{T} \int_0^T w^T dt$ . Given that the  $m^T(0) = m^T(T) = m_0$ , the couple  $(\bar{m}^T, \bar{w}^T)$  verifies, in the sense of distributions,  $-\Delta \bar{m} + \operatorname{div}(\bar{w}) = 0$  and it is an admissible competitor for the stationary problem.

Now we just need to apply Jensen's inequality to get

$$J^T(m_0, m^T, w^T) = \frac{1}{T} \int_0^T \int_{\mathbb{T}^d} H^*(x, w^T/m^T) dm^T + \mathcal{F}(m^T) dt \geq \int_{\mathbb{T}^d} H^*(x, \bar{w}^T/\bar{m}^T) d\bar{m}^T + \mathcal{F}(\bar{m}^T).$$

If we take the infimum over  $(m^T, w^T) \in \Pi_T(m_0, m_0)$  and we take the limit in  $T$ , we end up with

$$-\lambda \geq \lim_{T \rightarrow \infty} \inf_{\Pi_T(m_0, m_0)} \int_{\mathbb{T}^d} H^*(x, \bar{w}^T/\bar{m}^T) d\bar{m}^T + \mathcal{F}(\bar{m}^T) \geq -\bar{\lambda}.$$

We already proved the opposite inequality in Proposition 1.1.9, which we recall that it holds true also outside the monotonicity assumption.  $\square$

As  $\lambda = \bar{\lambda}$ , the dynamic programming principle for  $\chi$  reads

$$\chi(m(t_1)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*\left(x, -\frac{w}{m}\right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t_2)) + \bar{\lambda}(t_2 - t_1).$$

If  $(\bar{m}, \bar{w})$  is a minimizer for the static MFG problem then, if we define  $(m(t), w(t)) = (\bar{m}, \bar{w})$  for any  $t \in \mathbb{R}$ , the relation above holds true. This means that the constant trajectory  $(\bar{m}, \bar{w})$  is a calibrated curve so that  $\bar{m} \in \mathcal{M}$ . Moreover, as the calibrated curve is stationary, the singleton  $\mathcal{N} = \{\bar{m}\}$  is a minimal invariant set because it is closed, invariant and it cannot contain any proper subset.

### 1.3.2 A non convex example where $-\lambda < -\bar{\lambda}$

We now present an example where the non convexity of  $\mathcal{F}$  leads to an ergodic configuration where the limit value  $-\lambda$  is strictly lower than the ergodic one  $-\bar{\lambda}$ . A straightforward consequence will be that there cannot be any stationary calibrated curve.

We fix  $e_d \in \mathbb{R}^d \setminus \{0\}$  a unit vector parallel to one of the axes and we identify  $\mathbb{T}^d$  with  $\mathbb{T}^{d-1} \times \mathbb{T}$  where  $\mathbb{T}$  is the torus identified by the direction  $e_d$ . We fix also  $H$  such that  $H^*$  verifies the following conditions:  $H^*(x, p) > 0$  for any  $x \in \mathbb{T}^d$ ,  $p \neq -e_d$  and  $H^*(x, -e_d) = 0$  for any  $x \in \mathbb{T}^d$ . The assumption (1.1) on the hamiltonian  $H$  implies that there exist two constants  $C_1 > 0$ ,  $C_2 > 0$  such that

$$C_1 I_d \leq D_{pp} H^*(x, p) \leq C_2 I_d \quad \forall x \in \mathbb{T}^d, \forall p \in \mathbb{R}^d. \quad (1.42)$$

Let us define the set  $\mathcal{A} \subset \mathcal{P}(\mathbb{T}^d)$  as the set of  $\mu \in \mathcal{P}(\mathbb{T}^d)$  for which there exists  $\mu' \in \mathcal{P}(\mathbb{T}^{d-1})$  such that  $\mu = \mu' \otimes dx_d$ , where  $dx_d$  is the Lebesgue measure on  $\mathbb{T}$ . Note that  $\mu \in \mathcal{A}$  if and only if  $\operatorname{div}(e_d \mu) = 0$ .

We fix  $m_0 : \mathbb{T}^d \rightarrow \mathbb{R}$  a smooth, strictly positive density such that  $m_0 \notin \mathcal{A}$ . A measure  $m$  belongs to the set  $\mathcal{B}$  if there exists  $z \in \mathbb{T}^d$  such that  $m(\cdot) = m_0(\cdot + z)$ . As  $\mathcal{A}$  and  $\mathcal{B}$  are closed and disjoint, they are separated by a positive distance  $\varepsilon > 0$ .

We choose  $\mathcal{F} : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  such that  $\mathcal{F} \geq 0$ ,  $\mathcal{F} \equiv 2$  in  $\mathcal{A}$  and  $\mathcal{F} \equiv 0$  in  $\mathcal{B}$ . The existence of such a function is ensured by Lemma 1.4.3 and Lemma 1.4.4 in Appendix. They also guarantee that we can choose  $\mathcal{F}$  such that it verifies the regularity assumptions that were in place in the previous sections.

We recall that the functional  $J^T(\mu, \cdot, \cdot)$  is defined on  $\mathcal{E}_2^T(\mu)$  by

$$J^T(\mu, m, w) = \int_0^T \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + \mathcal{F}(m(t))dt.$$

In this section we add in the definition of  $\mathcal{E}_2^T(\mu)$  a viscosity constant  $\sigma > 0$  so that  $(m, w)$  verifies  $-\partial_t m + \sigma \Delta m - \operatorname{div}(w) = 0$ .

We also define the ergodic functional  $J : \mathcal{E} \rightarrow \mathbb{R}$  as follows

$$J(m, w) = \int_{\mathbb{T}} H^* \left( x, -\frac{dw}{dm}(x) \right) dm + \mathcal{F}(m), \quad (1.43)$$

where in this framework,  $(m, w)$  verifies  $\sigma \Delta m - \operatorname{div} w = 0$ . According to the definition of  $\bar{\lambda}$  in (1.6), we have

$$-\bar{\lambda} = \inf_{(m, w) \in \mathcal{E}} J(m, w) \quad (1.44)$$

**Proposition 1.3.2.** *There exists a  $\sigma_0 > 0$  such that for any  $\sigma \in (0, \sigma_0]$  we have  $-\lambda < -\bar{\lambda}$ .*

*Proof.* We define

$$m(t, x) = m_0(x - e_d t),$$

and

$$w(t, x) = e_d m(t, x) + \sigma Dm_0(x - e_d t).$$

The couple  $(m, w)$  belongs to  $\mathcal{E}_2^T(m_0)$ , so  $-\lambda \leq J^T(m, w)$ . By definition of  $\mathcal{F}$ , we know that  $\mathcal{F}(m(t)) = 0$  for any time  $t$ . Moreover, since  $D_{pp}L \leq C_2 I_d$  with  $0 = H^*(x, -e_d) \leq L(x, \alpha)$ , we have  $H^*(x, \alpha) \leq \frac{1}{2}C_2|\alpha + e_d|^2$ . Thus

$$-\lambda \leq \frac{1}{T} \int_0^T \int_{\mathbb{T}^d} \frac{C_2}{2} \left| \frac{w(t, x)}{m(t, x)} - e_d \right|^2 m(t, x) dx dt = \frac{C_2}{2} \sigma^2 \int_{\mathbb{T}^d} \frac{|Dm_0(x)|^2}{m_0(x)} dx = \sigma^2 I,$$

where  $I = \frac{C_2}{2} \int_{\mathbb{T}^d} \frac{|Dm_0(x)|^2}{m_0(x)} dx$ .

We now focus on the static case. We recall that the differential constraint on  $J$  is  $-\sigma \Delta m + \operatorname{div} w = 0$ . By standard arguments we have that there exists a minimizer  $(\bar{m}, \bar{w})$  of (1.43). As in Proposition 1.1.1,  $(\bar{m}, \bar{w})$  must also minimize

$$\bar{J}(m, w) = \int_{\mathbb{T}} H^* \left( x, -\frac{dw}{dm}(x) \right) dm + F(x, \bar{m}) dm. \quad (1.45)$$

As in the proof of Proposition 1.1.1 we can define the dual problem of (1.45) in the sense of Fenchel-Rockafellar, which reads

$$\inf_{(u, \Lambda) \in C^2(\mathbb{T}^d) \times \mathbb{R}} \{ \Lambda \text{ s.t. } -\Lambda - \Delta u + H(x, Du) \leq F(x, \bar{m}) \}.$$

Thanks to the regularity of  $F(\cdot, \bar{m})$  we have a smooth solution  $(\bar{u}, \bar{\lambda})$  which solves  $-\bar{\lambda} + \Delta \bar{u} + H(x, D\bar{u}) = F(x, \bar{m})$ . By duality, if we argue as in Proposition 1.1.1 (see again [11]), we get that  $\bar{w} = -\bar{m} D_p H(x, D\bar{u})$ , so that, by Schauder theory,  $\bar{m}$  is smooth and bounded from below.

We can now estimate  $-\bar{\lambda}$ . Thanks to the regularity of  $(\bar{m}, \bar{w})$ , the parabolic constraint ensures that  $\bar{w} = \sigma D\bar{m} + \zeta$  where  $\zeta$  is a smooth, divergence free vector field. If  $(\bar{m}, \bar{w}) = (\bar{m}, \sigma D\bar{m} + \zeta)$  is a minimizer of (1.43) and we use the growth assumption (1.42), we have

$$\int_{\mathbb{T}^d} H^* \left( x, \frac{\sigma D\bar{m} + \zeta}{\bar{m}} \right) \bar{m}(dx) + \mathcal{F}(\bar{m}) \geq \frac{C_1}{2} \int_{\mathbb{T}^d} \left| \frac{\sigma D\bar{m} + \zeta}{\bar{m}} + e_d \right|^2 \bar{m}(dx) + \mathcal{F}(\bar{m}). \quad (1.46)$$

Note that, as  $\operatorname{div}\zeta = 0$  and  $\bar{m}$  is smooth and bounded from below,  $\int_{\mathbb{T}^d} \zeta \cdot D\bar{m}/\bar{m} = 0$ . Indeed,

$$\int_{\mathbb{T}^d} \frac{D\bar{m}}{\bar{m}} \zeta dx = \int_{\mathbb{T}^d} D(\ln(\bar{m})) \zeta dx = - \int_{\mathbb{T}^d} \ln(\bar{m}) \operatorname{div}(\zeta) dx = 0.$$

Therefore, if we expand the square in (1.46), we get

$$\begin{aligned} & \int_{\mathbb{T}} H^* \left( x, \frac{\sigma D\bar{m} + \zeta}{\bar{m}} \right) \bar{m}(dx) \\ & \geq \frac{C_1}{2} \int_{T^d} \sigma^2 \frac{|D\bar{m}|^2}{\bar{m}} + \left| \frac{\zeta}{\bar{m}} + e_d \right|^2 \bar{m} dx \geq \frac{C_1}{2} \int_{T^d} \left| \frac{\zeta}{\bar{m}} + e_d \right|^2 \bar{m} dx. \end{aligned} \quad (1.47)$$

Plugging (1.47) into (1.44) we eventually find that

$$-\bar{\lambda} \geq \frac{C_1}{2} \int_{T^d} \left| \frac{\zeta}{\bar{m}} + e_d \right|^2 \bar{m} dx + \mathcal{F}(\bar{m}).$$

The right-hand side of the above inequality is bounded from below by a positive constant independent of  $\sigma$ . Indeed, we know that, for any  $\sigma$ ,  $\bar{m} > 0$ , so

$$-\bar{\lambda} \geq \frac{C_1}{2} \inf_{(m, \xi)} \int_{T^d} \left| \frac{\xi}{m} + e_d \right|^2 m dx + \mathcal{F}(m)$$

where the infimum is taken over all the probability densities  $m > 0$  and the free divergence vectors  $\xi$ . Here,  $m$  and  $\xi$  do not verify the elliptic constraint, therefore we lose the dependence on  $\sigma$ .

Let  $(m_n, \xi_n)$  be a minimizing sequence and  $m \in \mathcal{P}(\mathbb{T}^d)$  be the limit of  $m_n$  (the existence of  $m$  is guaranteed by the compactness of  $\mathcal{P}(\mathbb{T}^d)$ ). If the infimum were zero then  $\operatorname{div}(m_n e_d) \rightarrow \operatorname{div}(m e_d) = 0$ . Indeed, as both the addends should converge to zero, for any test function  $\varphi$ , we have

$$\begin{aligned} & \left| \int_{\mathbb{T}^d} \operatorname{div}(m_n e_d) \varphi dx \right| = \left| \int_{\mathbb{T}^d} m_n e_d \cdot D\varphi dx \right| = \left| \int_{\mathbb{T}^d} (m_n e_d + \xi_n) \cdot D\varphi dx \right| \\ & \leq \left( \int_{\mathbb{T}^d} \frac{|m_n e_d + \xi|^2}{m_n} dx \right)^{1/2} \left( \int_{\mathbb{T}^d} |D\varphi|^2 m_n dx \right)^{1/2} \rightarrow 0 \end{aligned}$$

On the other hand, if  $\operatorname{div}(m e_d) = 0$ , then, by construction of  $\mathcal{F}$ , we have  $\mathcal{F}(m_n) \rightarrow \mathcal{F}(m) = 2$ . Therefore, there exists a constant  $K > 0$  independent of  $\sigma$  such that  $-\bar{\lambda} > K$ .

We can conclude the proof choosing  $\sigma$  small enough such that

$$0 \leq -\lambda \leq \sigma^2 I < K \leq -\bar{\lambda}.$$

□

**Proposition 1.3.3.** *Under the hypothesis of Subsection 1.3.2 the projected Mather set  $\mathcal{M}$  does not contain any stationary calibrated curve.*

*Proof.* We recall that if  $m(t)$  is a calibrated curve then

$$\chi(m(t_1)) - \chi(m(t_2)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \mathcal{F}(m(s)) ds + \lambda(t_2 - t_1).$$

If  $m(t)$  is constantly equal to  $\bar{m}$  then we have that  $\int_{\mathbb{T}} L(x, -\bar{w}/\bar{m}) dm + \mathcal{F}(\bar{m}) = -\bar{\lambda}$ . As  $-\lambda < -\bar{\lambda}$ , it implies

$$\chi(\bar{m}) - \chi(\bar{m}) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{\bar{w}}{\bar{m}} \right) d\bar{m}(s) + \mathcal{F}(\bar{m}) ds + \lambda(t_2 - t_1) = (t_2 - t_1)(-\bar{\lambda} + \lambda) > 0,$$

so the contradiction. □

## 1.4 Appendix

We present here the proof of the duality (1.12), that we used in Proposition 1.1.1, and the proof of Lemma 1.1.2.

**Lemma 1.4.1.** *Let  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and  $(\bar{m}, \bar{w}) \in \mathcal{E}_\delta^2(m_0)$  be a minimizer of  $\mathcal{V}_\delta(m_0)$ . Then, the following equality holds true*

$$\begin{aligned} & \inf_{u \in C_b^2([0, +\infty) \times \mathbb{T}^d)} \left\{ - \int_{\mathbb{T}^d} u(0, x) dm_0(x) \text{ where } -\partial_t u - \Delta u + \delta u + H(x, Du) \leq F(x, \bar{m}) \right\} \\ &= \min_{(m, w) \in \mathcal{E}_\delta^2(m_0)} \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + F(x, \bar{m}(t)) dm(t) dt. \end{aligned} \quad (1.48)$$

*Proof.* The proof is a careful adjustment of the one in [21, Lemma 4.2]. It relies on the Fenchel-Rockafellar theorem. In our setting we have an additional difficulty which comes from the fact that we are working with trajectories and measures defined on the unbounded set  $[0, +\infty) \times \mathbb{T}^d$ . Therefore, one has to be careful when it comes to handle the dual space of continuous functions.

To overcome this issue we approximate the infinite horizon minimization problem with the following one

$$\inf_{(m, w) \in \mathcal{E}_\delta^t(m_0)} \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(s)}{dm(s)}(x) \right) dm(s) + F(x, \bar{m}(s)) dm(s) ds. \quad (1.49)$$

Now that we have a problem defined on the compact set  $[0, t] \times \mathbb{T}^d$  we can apply the Fenchel-Rockafellar theorem to prove that (1.49) is the dual problem of

$$\inf_{u \in \mathcal{K}^t} \left\{ - \int_{\mathbb{T}^d} u(0, x) dm_0(x) \right\}, \quad (1.50)$$

where  $\mathcal{K}^t$  is the set of  $u \in C^{1,2}([0, t] \times \mathbb{T}^d)$  such that  $-\partial_t u - \Delta u + \delta u + H(x, Du) \leq F(x, \bar{m})$  with terminal condition  $u(t, x) = 0$  for any  $x \in \mathbb{T}^d$ .

To get this result, we first need to introduce several objects. We set  $E_0 = C^{1,2}([0, t] \times \mathbb{T}^d)$  and  $E_1 = C([0, t] \times \mathbb{T}^d) \times C([0, t] \times \mathbb{T}^d; \mathbb{R}^d)$ . We define the linear functional  $\Lambda : E_0 \rightarrow E_1$  as follows

$$\Lambda(\phi) = (e^{-\delta s}(\partial_t \phi + \Delta \phi - \delta \phi), e^{-\delta t} D\phi).$$



We define also  $\mathcal{I} : E_0 \rightarrow \mathbb{R}$  and  $\mathcal{L} : E_1 \rightarrow \mathbb{R}$  as

$$\mathcal{I}(\phi) = - \int_{\mathbb{T}^d} \phi(0, x) dm_0(x) + \Omega_t(\phi), \quad \mathcal{L}(\alpha, \beta) = \int_0^t \int_{\mathbb{T}^d} \chi_s(x, \alpha(s, x), \beta(s, x)) ds,$$

where

$$\Omega_t(\phi) = \begin{cases} 0 & \text{if } \phi(t, x) = 0 \quad \forall x \in \mathbb{T}^d \\ +\infty & \text{otherwise} \end{cases}$$

and  $\chi_s : \mathbb{T}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  is

$$\chi_s(x, a, b) = \begin{cases} 0 & \text{if } -e^{\delta s} a + H(x, e^{\delta s} b) \leq F(x, \bar{m}(s)) \\ +\infty & \text{otherwise.} \end{cases} \quad (1.51)$$

If, for any  $\phi \in E_0$ , we set  $\mathcal{A}_t(\phi) = \mathcal{I}(\phi) + \mathcal{L}(\Lambda(\phi))$ , then

$$\inf_{\phi \in E_0} \mathcal{A}_t(\phi) = \inf_{u \in \mathcal{K}^t} \left\{ - \int_{\mathbb{T}^d} u(0, x) dm_0(x) \right\}. \quad (1.52)$$

Both  $\mathcal{I}$  and  $\mathcal{L}$  are convex, lower semicontinuous and proper. The linear functional  $\Lambda$  is bounded and if  $\varepsilon > 0$  and  $u_\varepsilon$  is the solution of

$$\begin{cases} -\partial_t u_\varepsilon - \Delta u_\varepsilon + \delta u_\varepsilon + H(x, Du_\varepsilon) = F(x, \bar{m}) - \varepsilon & \text{in } [0, t] \times \mathbb{T}^d \\ u_\varepsilon(t, x) = 0 & \text{in } \mathbb{T}^d, \end{cases}$$

then, in a neighborhood of  $\Lambda(u_\varepsilon)$ ,  $\mathcal{L}$  is constantly equal to zero and, hence, it is continuous. As the qualification hypothesis are verified, the Fenchel-Rockafellar theorem states that

$$\inf_{\phi \in E_0} \mathcal{A}_t(\phi) = - \min_{(m, w) \in E'_1} \mathcal{I}^*(\Lambda^*(m, w)) + \mathcal{L}^*(-(m, w)) := - \min_{(m, w) \in E'_1} \mathcal{B}_t(m, w),$$

where  $E'_1$  is the set of vector valued Radon measures  $(m, w)$  over  $[0, t] \times \mathbb{T}^d$ .

If we use the definition of Fenchel conjugate and adjoint operator, we get

$$\mathcal{I}^*(\Lambda^*(m, w)) = \sup_{\phi \in E_1} \{ \langle \Lambda^*(m, w), \phi \rangle - \mathcal{I}(\phi) \} = \sup_{\phi \in E_1} \{ \langle (m, w), \Lambda \phi \rangle - \mathcal{I}(\phi) \}.$$

We can suppose that  $\phi(t, x) = 0$  for any  $x \in \mathbb{T}^d$ , otherwise we had  $\mathcal{I}^*(\Lambda^*(m, w)) = -\infty$ . Therefore, we get that

$$\begin{aligned} \mathcal{I}^*(\Lambda^*(m, w)) &= \sup_{\phi \in E_1} \left\{ \int_0^t \int_{\mathbb{T}^d} e^{-\delta s} (\partial_t \phi(s) - \delta \phi(s) + \Delta \phi(s)) dm(s) ds + \int_0^t \int_{\mathbb{T}^d} e^{-\delta s} D \phi dw(s) ds \right. \\ &\quad \left. + \int_{\mathbb{T}^d} \phi(0) dm_0 \right\} \\ &= \sup_{\phi \in E_1} \left\{ \int_0^t \int_{\mathbb{T}^d} \partial_t (e^{-\delta s} \phi(s)) + \Delta (e^{-\delta s} \phi(s)) dm(s) + \int_0^t \int_{\mathbb{T}^d} D (e^{-\delta s} \phi(s)) dw(s) ds \right. \\ &\quad \left. + \int_{\mathbb{T}^d} \phi(0) dm_0 \right\}. \end{aligned}$$

The above expression is equal to zero when  $m(0) = m_0$  and  $(m, w)$  solves  $\partial_t m - \Delta m + \text{div} w = 0$  in the sense of distributions and  $+\infty$  otherwise.

We now turn to  $\mathcal{L}^*(-(m, w))$ . We compute the Fenchel conjugate of  $\chi_s(x, \cdot, \cdot)$ , defined in (1.51), that is

$$\chi_s^*(p, q) = \sup_{(a, b) \in \mathbb{R} \times \mathbb{R}^d} \{ap + bq - \chi_s(x, a, b)\}.$$

If  $p > 0$ , then  $\chi_s^*(x, p, q) = +\infty$ , so we suppose  $p \leq 0$ . If we use the definition of  $\chi_s$ , we get that the above expression is equal to

$$\begin{aligned} & \sup_{b \in \mathbb{R}^d} \left\{ e^{-\delta s} H(x, e^{\delta s} b) p - e^{-\delta s} F(x, \bar{m}(s)) p + bq \right\} \\ &= -pe^{-\delta s} \sup_{b \in \mathbb{R}^d} \left\{ \left( -\frac{q}{p} e^{\delta s} b - H(x, e^{\delta s} b) \right) \right\} - e^{-\delta s} p F(x, \bar{m}(s)) \\ &= -e^{-\delta s} p H^* \left( x, -\frac{q}{p} \right) - e^{-\delta s} p F(x, \bar{m}(s)). \end{aligned}$$

So, we finally have that

$$\mathcal{B}_t((m, w)) = \begin{cases} \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw}{dm} \right) + F(x, \bar{m}) dm & \text{if } (m, w) \in \mathcal{E}_2^t(m_0) \\ +\infty & \text{otherwise} \end{cases}$$

and, eventually,

$$\min_{(m, w) \in \mathcal{E}_1^t} \mathcal{B}_t((m, w)) = \min_{(m, w) \in \mathcal{E}_2^t(m_0)} \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(s)}{dm(s)}(x) \right) dm(s) + F(x, \bar{m}(s)) dm(s) ds.$$

To conclude the proof we need to argue that (1.52) and (1.49) converge respectively to the first and the second term of (1.48).

We recall that  $(\bar{m}, \bar{w})$  is the minimizer of  $\mathcal{V}_\delta(m_0)$  around which we linearize the non convex term of  $\mathcal{V}_\delta(m_0)$ . Thus,  $(\bar{m}, \bar{w})$  is also a minimizer, among all  $(m, w) \in \mathcal{E}_\delta^2(m_0)$ , of the linearized problem

$$\bar{J}_\delta(m, w) = \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} H^* \left( x, -\frac{dw(t)}{dm(t)}(x) \right) dm(t) + F(x, \bar{m}(t)) dm(t) dt.$$

We fix  $(m^t, w^t)$  a minimizer of  $\mathcal{B}_t$ . Note that  $\mathcal{B}_t$  is the restriction on  $[0, t]$  of the integral that defines  $\bar{J}_\delta$ . Therefore, the estimates (1.8), (1.9) and (1.10), that we proved for minimizing sequences of  $\mathcal{V}_\delta(m_0)$ , hold true for  $(m^t, w^t)$  and they are uniform in  $t$ . This ensures that, up to subsequence,  $(m^t, w^t)$  converges to  $(\hat{m}, \hat{w})$  and

$$\lim_{t \rightarrow +\infty} \mathcal{B}_t(m^t, w^t) = \bar{J}_\delta(\hat{m}, \hat{w}).$$

Let us suppose that  $(\hat{m}, \hat{w})$  is not optimal, i.e.  $\bar{J}_\delta(\hat{m}, \hat{w}) = \bar{J}_\delta(\bar{m}, \bar{w}) + 3\varepsilon$ , for a given  $\varepsilon > 0$ . Then, there exists a sufficiently large  $t$  such that  $\mathcal{B}_t(m^t, w^t) \geq \bar{J}_\delta(\bar{m}, \bar{w}) + 2\varepsilon$  and

$$|\bar{J}_\delta(\bar{m}, \bar{w}) - \mathcal{B}_t(\bar{m}, \bar{w})| = \int_t^{+\infty} e^{-\delta s} \int_{\mathbb{T}^d} H^* \left( x, -\frac{d\bar{w}(s)}{d\bar{m}(s)}(x) \right) d\bar{m}(s) + F(x, \bar{m}(s)) d\bar{m}(s) ds \leq \varepsilon.$$

This implies that

$$\mathcal{B}_t(\bar{m}, \bar{w}) + \varepsilon \leq \mathcal{B}_t(m^t, w^t),$$

which is in contradiction with the assumption that  $(m^t, w^t)$  is a minimizer of  $\mathcal{B}_t$ .

The last piece missing is to prove that

$$\begin{aligned} & \lim_{t \rightarrow +\infty} \inf_{u \in \mathcal{K}^t} \left\{ - \int_{\mathbb{T}^d} u(0, x) dm_0(x) \right\} \\ = & \inf_{u \in C_b^2([0, +\infty) \times \mathbb{T}^d)} \left\{ - \int_{\mathbb{T}^d} u(0, x) dm_0(x) \text{ where } -\partial_t u - \Delta u + \delta u + H(x, Du) \leq F(x, \bar{m}) \right\}. \end{aligned} \quad (1.53)$$

Note that the equation

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = F(x, \bar{m}) & \text{in } [0, t] \times \mathbb{T}^d \\ u(t, x) = 0 & \text{in } \mathbb{T}^d \end{cases} \quad (1.54)$$

enjoys the comparison principle and so it admits a unique solution  $\bar{u}_t \in C^{1,2}([0, t] \times \mathbb{T}^d)$  which is also the unique minimizer of (1.50). Indeed, let  $u$  be such that

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) \leq F(x, \bar{m}) & \text{in } [0, t] \times \mathbb{T}^d \\ u(t, x) = 0 & \text{in } \mathbb{T}^d, \end{cases}$$

then, by the comparison principle,  $u(0, x) \leq \bar{u}_t(0, x)$  for any  $x \in \mathbb{T}^d$ .

In the proof of Lemma 1.1.3 we showed that, if  $\bar{u}_t$  is a solution of (1.54), then  $\partial_t \bar{u}_t$ ,  $D\bar{u}_t$  and  $D^2 \bar{u}_t$  are uniformly bounded with respect to  $t$ . If we use these bounds we get that, as long as  $\delta > 0$ ,  $\bar{u}_t$  is uniformly bounded as well. This implies that, up to subsequence,  $\bar{u}_t$  uniformly converges on compact sets to  $\bar{u} \in C_b^{1,2}([0, +\infty) \times \mathbb{T}^d)$ , where  $\bar{u}$  is the solution of

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = F(x, \bar{m}) & \text{in } [0, +\infty) \times \mathbb{T}^d \\ u \in L^\infty([0, +\infty) \times \mathbb{T}^d). \end{cases}$$

The same argument we propose for  $\bar{u}_t$  gives us that  $\bar{u}$  is the unique minimizer of (1.53). Moreover, as the convergence of  $\bar{u}_t$  to  $\bar{u}$  is uniform on compact sets, then

$$\lim_{t \rightarrow +\infty} \int_{\mathbb{T}^d} \bar{u}_t(0, x) dm_0(x) = \int_{\mathbb{T}^d} \bar{u}(0, x) dm_0(x),$$

which concludes the proof. □

We now provide the argument that proves the dynamic programming principle of  $\mathcal{V}_\delta$ .

*Proof of Lemma 1.1.2.* We want to prove that

$$\mathcal{V}_\delta(m_0) = \inf_{(m, \alpha) \in \mathcal{E}_2^\delta(m_0)} \left( \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta t} \mathcal{V}_\delta(m(t)) \right). \quad (1.55)$$

We start showing that the left-hand side of (1.55) is greater than the right-hand side. For any  $(m, \alpha) \in \mathcal{E}_2^\delta(m_0)$  we define  $(m^t, \alpha^t) \in \mathcal{E}_2^\delta(m(t))$  as  $(m^t(s), \alpha^t(s)) = (m(s+t), \alpha(s+t))$ . Then, using the definition of  $\mathcal{V}_\delta$  we get

$$\mathcal{V}_\delta(m_0) = \inf_{(m, \alpha) \in \mathcal{E}_2^\delta(m_0)} \left( \int_0^\infty e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds \right)$$

$$\begin{aligned}
&= \inf_{(m,\alpha)} \left( \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds \right. \\
&\quad \left. + \int_t^\infty e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds \right) \\
&= \inf_{(m,\alpha)} \left( \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds \right. \\
&\quad \left. + e^{-\delta t} \int_0^\infty e^{-\delta l} \int_{\mathbb{T}^d} H^*(x, \alpha^t(l)) dm^t(l) + \mathcal{F}(m^t(l)) dl \right) \\
&\geq \inf_{(m,\alpha) \in \mathcal{E}_2^\delta(m_0)} \left[ \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds \right. \\
&\quad \left. + e^{-\delta t} \inf_{(m,\alpha) \in \mathcal{E}_2^\delta(m(t))} \left( \int_0^\infty e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds \right) \right] \\
&= \inf_{(m,\alpha) \in \mathcal{E}_2^\delta(m_0)} \left( \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta t} \mathcal{V}_\delta(m(t)) \right).
\end{aligned}$$

We prove the inverse inequality by contradiction, i.e. we suppose that there exist  $\varepsilon > 0$ ,  $(m, \alpha) \in \mathcal{E}_2^\delta(m_0)$  and  $(\bar{m}, \bar{\alpha}) \in \mathcal{E}_2^\delta(m(t))$  such that

$$\mathcal{V}_\delta(m_0) - \varepsilon \geq$$

$$\int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta t} \int_0^\infty e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(s)) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds.$$

We can now define a new competitor  $(m_1, w_1)$  for  $\mathcal{V}_\delta(m_0)$  as follows

$$(m_1(s), w_1(s)) = \begin{cases} (m(s), w(s)) & \text{if } s < t \\ (\bar{m}(s-t), \bar{w}(s-t)) & \text{if } s \geq t. \end{cases}$$

As  $(m_1, w_1) \in \mathcal{E}_2^\delta(m_0)$  we get immediately the contradiction, indeed

$$\begin{aligned}
\mathcal{V}_\delta(m_0) &\leq \int_0^\infty e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm_1(s) + \mathcal{F}(m_1(s)) ds \\
&= \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha_1) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta t} \int_0^\infty e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(s)) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds. \\
&\leq \mathcal{V}_\delta(m_0) - \varepsilon.
\end{aligned}$$

□

In the next lemma we prove the result used in Proposition 1.2.5, which is part of an on-going work with Marco Cirant. The following lemma addresses the problem of existence of minimizers in the context of dual representation for Hamilton-Jacobi equations. A similar approach was used in [51] in the case of  $F = 0$  but with a more general viscosity and a wider range of Hamiltonians. Moreover, in the first order case, it is a natural problem that arises in the context of optimal transport. See for instance [6].

**Lemma 1.4.2.** *Let  $(\bar{m}, \bar{w})$  be as in Proposition 1.2.5. For any  $0 \leq t_1 < t_2$ , the minimization problem*

$$\bar{A}_{t_1}^{t_2} = \inf_{u \in \mathcal{K}} A_{t_1}^{t_2}(u) = \inf_{u \in \mathcal{K}} \left\{ \int_{\mathbb{T}^d} u(x, t_2) d\bar{m}(t_2) - \int_{\mathbb{T}^d} u(x, t_1) d\bar{m}(t_1) \right\}$$

where  $\mathcal{K}$  is the set of  $u \in C^{1,2}([t_1, t_2] \times \mathbb{T}^d)$  such that  $-\partial_t u - \Delta u + H(x, Du) = F(x, \bar{m})$  and  $\int_{\mathbb{T}^d} u(t_1, x) dx = 0$ , admits a solution.

*Proof.* The difficulties of this minimization problem comes from the fact that, a priori, we have no regularity on the measure  $\bar{m}$ , which does not allow us to directly get the compactness of the minimizing sequence that we need. On the other hand the dynamic programming principle of  $\chi$  and some local in time semiconcavity estimates help to overcome this obstacle.

Let us recall that the dynamic programming principle for  $\chi$  reads

$$\chi(m_0) = \inf_{(m, w)} \left( \int_0^t H^* \left( x, -\frac{w}{m} \right) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t$$

and that the following convex duality holds true (see Remark 1.2.3)

$$\begin{aligned} \inf_{(m, w) \in \Pi(\bar{m}(t_1), \bar{m}(t_2))} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) dm(s) ds \\ = - \inf_{u \in \mathcal{K}} \left\{ \int_{\mathbb{T}^d} u(x, t_2) d\bar{m}(t_2) - \int_{\mathbb{T}^d} u(x, t_1) d\bar{m}(t_1) \right\}. \end{aligned}$$

First of all we prove that for any  $t_1 < t_2 < t_3$  we have that  $\bar{A}_{t_1}^{t_3} = \bar{A}_{t_1}^{t_2} + \bar{A}_{t_2}^{t_3}$ . Indeed, using the duality between the two minimization problems, we have

$$\begin{aligned} \bar{A}_{t_1}^{t_2} + \bar{A}_{t_2}^{t_3} &= \inf_{(m, w) \in \Pi(\bar{m}(t_1), \bar{m}(t_2))} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) dm(s) ds \\ &+ \inf_{(m, w) \in \Pi(\bar{m}(t_2), \bar{m}(t_3))} \int_{t_2}^{t_3} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) dm(s) ds. \end{aligned}$$

If we use the dynamic programming principle of  $\chi$  and the fact that  $(\bar{m}, \bar{w})$  is optimal, the expression above is equal to

$$\begin{aligned} \int_{t_1}^{t_3} \int_{\mathbb{T}^d} H^* \left( x, -\frac{\bar{w}}{\bar{m}} \right) d\bar{m}(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) d\bar{m}(s) ds = \\ \inf_{(m, w) \in \Pi(\bar{m}(t_1), \bar{m}(t_3))} \int_{t_1}^{t_3} \int_{\mathbb{T}^d} H^* \left( x, -\frac{w}{m} \right) dm(s) + \int_{\mathbb{T}^d} F(x, \bar{m}(s)) dm(s) ds = \bar{A}_{t_1}^{t_3} \end{aligned}$$

We claim now and we prove later that, if  $\{u_n\}_n$  is a minimizing sequence for  $\bar{A}_{t_1}^{t_3}$ , then  $u_n$  uniformly converges to a function  $u \in C^{1,2}([t_1, t_3] \times \mathbb{T}^d)$  on any  $[t_1, t]$  with  $t < t_3$  and  $u$  is admissible for  $\bar{A}_{t_1}^t$ . This implies that the function  $u$  is a minimizer for  $\bar{A}_{t_1}^t$  and in particular for  $t = t_2$ . If we suppose that  $A_{t_1}^{t_2}(u) = \bar{A}_{t_1}^{t_2} + \varepsilon$ , then we have

$$A_{t_1}^{t_3}(u_n) = A_{t_1}^{t_2}(u_n) + A_{t_2}^{t_3}(u_n) \geq A_{t_1}^{t_2}(u_n) + \bar{A}_{t_2}^{t_3}.$$

If we take the limit  $n \rightarrow \infty$  on both side, the uniform convergence of  $u_n$  on  $[t_1, t_2]$  and the fact that  $u_n$  is a minimizing sequence for  $\bar{A}_{t_1}^{t_3}$  give us

$$\bar{A}_{t_1}^{t_3} \geq A_{t_1}^{t_2}(u_n) + \bar{A}_{t_2}^{t_3} = A_{t_1}^{t_2}(u) = \bar{A}_{t_1}^{t_2} + \varepsilon + \bar{A}_{t_2}^{t_3} = \bar{A}_{t_1}^{t_3} + \varepsilon$$

which is impossible and so  $u$  has to be a minimizer for  $\bar{A}_{t_1}^{t_2}$ .

We now prove our claim and we show that the set of functions  $u \in C^{1,2}([t_1, t_3] \times \mathbb{T}^d)$  which solves

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, \bar{m}) & \text{in } [t_1, t_3] \times \mathbb{T}^d \\ \int_{\mathbb{T}^d} u(t_1, x) dx = 0 \end{cases} \quad (1.56)$$

is uniformly bounded in  $C^1([t_1, \tau] \times \mathbb{T}^d)$  for any  $\tau < t_3$ . This gives the local convergence that we used earlier. Without loss of generality we can suppose  $t_1 = 0$  and  $t_3 = T$ . As in Lemma 1.1.3 we argue by semiconcavity.

We consider  $\xi \in \mathbb{R}^d$ ,  $|\xi| \leq 1$  and we look at the equation solved by  $w(t, x) = D^2 u^T(t, x) \xi \cdot \xi$ . We now define  $\bar{w}(t, x) = w(t, x) \eta(t)$ , where  $\eta$  is the cutoff function  $\eta(t) = (t - T)^2$ . We choose  $\xi$  such that it maximizes  $\sup_{t,x} \bar{w}(t, x)$ .

If we derive twice in space the HJB equation in (1.56), then  $\bar{w}$  solves

$$\begin{aligned} & -\partial_t \bar{w} - w \eta' - \Delta \bar{w} + D_{\xi\xi} H(x, Du) \eta + 2D_{\xi p} H(x, Du) \cdot D^2 u \xi \eta \\ & + D_{pp} H(x, Du) D^2 u \xi \cdot D^2 u \xi \eta + D_p H(x, Du) \cdot D \bar{w} = D_{\xi\xi}^2 F(x, m) \eta. \end{aligned}$$

The cutoff function ensures the existence of a positive interior maximum of  $\bar{w}$ . At the maximum, using also the boundedness of  $D_{pp}^2 H$ , the above equation implies

$$-w \eta' - K + 2D_{\xi p} H(x, Du) \cdot D^2 u \xi \eta + \bar{C}^{-1} |D^2 u \xi|^2 \eta \leq D_{\xi\xi}^2 F(x, m) \eta.$$

Rearranging the terms and using the boundedness of  $D_{\xi\xi}^2 F$  we get

$$|D^2 u \xi|^2 \eta \leq w \eta' + C + 2C |D^2 u \xi| \eta.$$

As  $\eta' = 2\eta^{1/2}$  we can apply the Young's inequality so that  $\eta' w \leq \eta/2 |D^2 u \xi|^2 + 4$  and

$$\frac{1}{2} |D^2 u \xi|^2 \eta \leq C + 2C |D^2 u \xi| \eta,$$

which in turn gives

$$|D^2 u \xi|^2 \eta \leq C.$$

If  $w^+$  and  $\bar{w}^+$  are the positive parts of  $w$  and  $\bar{w}$ , then we have our semiconcavity estimates because on  $[0, \tau] \times \mathbb{T}^d$

$$(w^+ \eta)^2 \leq (\bar{w}^+)^2 \leq |D^2 u \xi|^2 \eta^2 \leq M.$$

Note that  $M = M(\tau)$  and it diverges when  $\tau \rightarrow T$ . On the other hand the estimates above, along with (1.14), gives uniform bounds on  $\|Du\|_\infty$  on  $[0, \tau]$  with  $\tau < T$ .

Integrating in space the HJB equation we get

$$|\partial_t \int_{\mathbb{T}^d} u dx| \leq \int_{\mathbb{T}^d} |H(x, Du)| + |F(x, \bar{m})| dx \leq C(\tau).$$

As  $\int_{\mathbb{T}^d} u(0, x) dx = 0$ , the above inequality ensures that  $|\int_{\mathbb{T}^d} u(t, x) dx| \leq C(\tau)$  for any  $t \leq \tau$ . This gives us  $\text{osc}(u(t, \cdot)) \leq |\int_{\mathbb{T}^d} u(t, x) dx| + C \sup_x |Du(t, x)| \leq C(\tau)$ .

As in Lemma 1.1.3, the boundedness of space derivatives implies also that  $|\partial_t u(t)| \leq C(\tau)$  for any  $t \leq \tau$  and so the claim.  $\square$

Here we provide the proof of the existence of the smooth functions that we used in Subsection 1.3.2.

**Lemma 1.4.3.** *For any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and any  $\epsilon > 0$ , there exists  $\Phi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  of class  $C^1$  such that  $\Phi(m_0) = 1$  and  $\Phi = 0$  on  $B_\epsilon^c(m_0)$ . Moreover, we can choose  $\Phi$  such that*

$$\|D_m \Phi\|_\infty \leq 10/\epsilon.$$

and with  $D_x D_m \Phi$  bounded.

*Proof.* Let  $E$  be the compact set of 1-Lipschitz continuous maps on  $\mathbb{T}^d$  vanishing at 0. Let  $(\phi_n)$  be a dense family in  $E$  consisting of smooth maps. For  $N$  large, we consider

$$\Psi_N(m) = \sup_{n=1, \dots, N} \int_{\mathbb{T}^d} \phi_n(m - m_0).$$

Then  $(\Psi_N)$  is a family which is uniformly Lipschitz continuous in  $\mathcal{P}(\mathbb{T}^d)$  and converges to  $\mathbf{d}_1(\cdot, m_0)$ . So, for  $\eta > 0$  small there exists  $N$  large enough such that

$$\|\Psi_N - \mathbf{d}_1(\cdot, m_0)\|_\infty \leq \eta.$$

Next we approximate the sup in the definition of  $\Psi_N$ . We consider

$$\Psi_N^\delta(m) = \delta \log \left( \sum_{n=1}^N \exp \left\{ \delta^{-1} \int_{\mathbb{T}^d} \phi_n(m - m_0) \right\} \right).$$

Recall that

$$\Psi_N(m) \leq \Psi_N^\delta(m) \leq \delta \ln(N) + \Psi_N(m).$$

Note that  $\Psi_N^\delta$  is  $C^1$ , with

$$D_m \Psi_N^\delta(m, x) = \left( \sum_{n=1}^N \exp \left\{ \delta^{-1} \int_{\mathbb{T}^d} \phi_n(m - m_0) \right\} \right)^{-1} \sum_{n=1}^N \exp \left\{ \delta^{-1} \int_{\mathbb{T}^d} \phi_n(m - m_0) \right\} D \phi_n(x). \quad (1.57)$$

Note that  $D_m \Psi_N^\delta(m, x)$  is a convex combination of  $D \phi_n(x)$ , so that

$$|D_m \Psi_N^\delta(m, x)| \leq \sup_n |D \phi_n(x)| \leq 1.$$

For  $\delta > 0$  small (depending on  $N$ ), we have

$$\|\Psi_N^\delta - \mathbf{d}_1(\cdot, m_0)\|_\infty \leq 2\eta.$$

In particular, for  $\epsilon > 0$ , choose  $\eta = \epsilon/5$ : then

$$\inf_{m \in B_\epsilon^c(m_0)} \Psi_N^\delta(m) \geq \epsilon - 2\eta = 3\epsilon/5 \quad \text{and} \quad \Psi_N^\delta(m_0) \leq 2\eta = 2\epsilon/5.$$

Moreover, if we derive (1.57) in space we get

$$D_x D_m \Psi_N^\delta(m, x) = \left( \sum_{n=1}^N \exp \left\{ \delta^{-1} \int_{\mathbb{T}^d} \phi_n(m - m_0) \right\} \right)^{-1} \sum_{n=1}^N \exp \left\{ \delta^{-1} \int_{\mathbb{T}^d} \phi_n(m - m_0) \right\} D^2 \phi_n(x).$$

Note that  $D_x D_m \Psi_N^\delta(m, x)$  is a convex combination of  $D^2 \phi_n(x)$ . Therefore, there exists a constant  $C_N$  such that

$$|D_x D_m \Psi_N^\delta(m, x)| \leq \sup_n |D^2 \phi_n(x)| \leq C_N.$$

To complete the result, define a map  $\zeta_\epsilon = \mathbb{R} \rightarrow [0, 1]$  smooth and nonincreasing, with  $\zeta_\epsilon(s) = 0$  if  $s \geq 3\epsilon/5$  and  $\zeta_\epsilon(s) = 1$  for  $s \leq 2\epsilon/5$ . We can choose  $\|\zeta'_\epsilon\|_\infty \leq 10/\epsilon$ . The map  $\Phi = \zeta_\epsilon \circ \Psi_N^\delta$  satisfies the claim.  $\square$

**Lemma 1.4.4.** *Let  $A$  and  $B$  be closed subsets of  $\mathcal{P}(\mathbb{T}^d)$  with an empty intersection. Then there exists a  $C^1$  map  $\Phi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  such that  $\Phi = 1$  on  $A$ ,  $\Phi = 0$  on  $B$  and  $D_x D_m \Phi$  bounded.*

*Proof.* Let  $\epsilon > 0$  be the minimal distance between  $A$  and  $B$ :

$$\epsilon := \inf_{m \in A, m' \in B} \mathbf{d}_1(m, m') > 0.$$

Let  $(m_n)$  be dense in  $A$  and  $\Phi_n : \mathcal{P}(\mathbb{T}^d) \rightarrow [0, 1]$  be associated with  $m_n$  as in Lemma 1.4.3:  $\Phi_n(m_n) = 1$ ,  $\Phi_n = 0$  in  $B_\epsilon^c(m_n)$  and  $\|D_x D_m \Phi_n\|$  bounded. For  $\delta > 0$  small and  $N$  large, let us set

$$\Psi_N^\delta(m) = \delta \log \left( \sum_{n=1}^N \exp \{ \delta^{-1} \Phi_n(m) \} \right).$$

Note that  $\Psi_N^\delta$  is  $C^1$  with

$$D_x D_m \Psi_N^\delta(m, y) = \left( \sum_{n=1}^N \exp \{ \delta^{-1} \Phi_n(m) \} \right)^{-1} \sum_{n=1}^N \exp \{ \delta^{-1} \Phi_n(m) \} D_x D_m \Phi_n(m, y).$$

In particular,

$$|D_x D_m \Psi_N^\delta(m, x)| \leq \sup_n |D^2 \phi_n(x)| \leq C_N.$$

For  $m \in B$  we have  $\Phi_n(m) = 0$ , so that  $\Psi_N^\delta(m) = \delta \ln(N)$ . As  $(m_n)$  is dense, we can choose, for  $\eta > 0$ ,  $N$  large enough so that

$$\max_{m \in A} \min_{n=1, \dots, N} \mathbf{d}_1(m, m_n) \leq \eta.$$

Then, for  $m \in A$ , there exists  $n \in \{1, \dots, N\}$  with  $\mathbf{d}_1(m, m_n) \leq \eta$ , so that (by Lipschitz continuity of  $\Phi_n$ )

$$\Phi_n(m) \geq \Phi_n(m_n) - 10\epsilon^{-1} \mathbf{d}_1(m, m_n) \geq 1 - 10\epsilon^{-1} \eta.$$

Thus

$$\Psi_N^\delta(m) \geq \delta \log \left( \exp \{ \delta^{-1} \Phi_n(m) \} \right) \geq 1 - 10\epsilon^{-1} \eta.$$

We now choose  $\eta > 0$  such that  $1 - 10\epsilon^{-1} \eta = 2/3$  (which in turns fixes  $N$ ), and then  $\delta > 0$  small such that  $\delta \ln(N) \leq 1/3$ . Then we have

$$\inf_{m \in A} \Psi_N^\delta(m) \geq 2/3 \quad \text{and} \quad \sup_{m \in B} \Psi_N^\delta(m) \leq 1/3.$$

Then conclusion follows easily.  $\square$

On the long time convergence of potential MFG





## Chapter 2

# Weak KAM theory for potential MFG

We develop the counterpart of weak KAM theory for potential mean field games. This allows to describe the long time behavior of time-dependent potential mean field game systems. Our main result is the existence of a limit, as time tends to infinity, of the value function of an optimal control problem stated in the space of measures. In addition, we show a mean field limit for the ergodic constant associated with the corresponding Hamilton-Jacobi equation.

This chapter is a joint work with Pierre Cardaliaguet and it was accepted for publication in *Journal of Differential Equations* (JDE).

## Introduction

The theory of mean field games (MFG), introduced simultaneously and independently by Lasry and Lions [61, 62] and Huang, Caines and Malhamé [57], is devoted to the analysis of models where a large number of players interact strategically with each other. Under suitable assumptions, the Nash equilibria of those games can be analyzed through the solutions of the, so-called, MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, T] \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, T] \\ m(0) = m_0, u(T, x) = G(x, m(T)) & \text{in } \mathbb{T}^d. \end{cases} \quad (2.1)$$

The first unknown  $u(t, x)$  is the value function of an infinitesimal player starting from  $x$  at time  $t$  while the second one,  $m(t)$ , describes the distribution of players at time  $t$ . The maps  $F, G : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  (where  $\mathcal{P}(\mathbb{T}^d)$  is the set of Borel probability measures on the torus  $\mathbb{T}^d$ ) describe the interactions between players.

In this paper we investigate the limit behavior, as the horizon  $T$  tends to infinity, of this system. This is a very natural question, especially when one looks at those models as dynamical systems.

One natural guess is that the system simplifies in large times and converges to a time independent model, called the ergodic MFG system:

$$\begin{cases} -\bar{\lambda} - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d, \\ \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d. \end{cases} \quad (2.2)$$

There is a relatively wide evidence of this phenomenon, starting from [65] and the Mexican wave model in [56] to more recent contributions in [18, 22, 23, 50].

All these papers, however, rely on a structure property, the so-called monotonicity assumption, which is seldom met in practice. More recently, the problem of understanding what happens in the non-monotone setting has been addressed in several papers. Gomes and Sedjro [49] found the first example of periodic solutions in the context of one-dimensional first order system with congestion. Cirant in [34] and Cirant and Nurbekyan in [33] forecast and then proved the existence of periodic solutions for a specific class of second order MFG systems (with quadratic Hamiltonian). These periodic trajectories were built through a bifurcation method in a neighborhood of a simple solution. Note that these examples show that the ergodic MFG system is not always the limit of the time-dependent ones. In [68] the second author gave additional evidence of this phenomena using ideas from *weak KAM theory* [39, 40, 41]. The main interest of the approach is that it allows to study the question for a large class of MFG systems, *potential MFG systems*.

We say that a MFG system like (2.1) is of potential type if it can be derived as optimality condition of the following optimal control problem on the Fokker-Plank equation

$$\mathcal{U}^T(t, m_0) = \inf_{(m, \alpha)} \int_t^T \int_{\mathbb{T}^d} H^*(x, \alpha(s, x)) dm(s) + \mathcal{F}(m(s)) dt + \mathcal{G}(m(T)), \quad (2.3)$$

where  $(m, \alpha)$  verifies the Fokker-Plank equation  $-\partial_t m + \Delta m + \operatorname{div}(\alpha m) = 0$  with  $m(t) = m_0$  and  $\mathcal{F}$  and  $\mathcal{G}$  are respectively the potentials of the functions  $F$  and  $G$  that appear in (2.1). Since the very beginning, this class of models has drawn a lot of attention. [62] first explained the mechanism behind the minimizing problem (2.3) and the MFG system (2.1) and, since then, the literature on potential MFG thrived. See for instance [11, 21, 44, 70] for the use of

theses techniques to build solutions and analyse their long-time behavior under a monotonicity assumption.

In the present paper, we investigate the behavior, as  $T \rightarrow +\infty$ , of the solutions to the mean field games system (2.1) which are minimizers of (2.3). It is a continuation of [68], which started the analysis of the convergence of the time-dependent, non-monotone, potential MFG systems through weak KAM techniques. We believe that these techniques lead to a more fundamental understanding of long time behavior for potential MFG. When the powerful tools of the weak KAM theory can be deployed, one can look at this problem in a more systematic way. Unlike the PDEs techniques that were so far used, this approach does not depend on the monotonicity of the system. A key point is that the weak KAM theory, exploiting the Hamiltonian structure of potential MFG, gives us a clear understanding of the limit object that the trajectories minimize when the time goes to infinity. We draw fully from both Fathi's seminal papers [39, 40, 41] and his book [42]. Several objects defined along the paper and the very structure of many proofs will sound familiar for who is acquainted with weak KAM theory. Nonetheless, it is not always straightforward to transpose these techniques into the framework of MFG and it often requires more effort than in the standard case. It is worthwhile to mention that infinite dimensional weak KAM theorems are not new, especially in the context of Wasserstein spaces: see for instance [45, 46, 47, 48]. These papers do not address the MFG problem but they surely share the same inspiration.

Let us now present our main results and discuss the strategy of proofs. As we have anticipated, the starting point of this paper are some results proved in [68]. The first one is the existence of the ergodic constant  $\lambda$ , such that

$$\frac{\mathcal{U}^T(0, \cdot)}{T} \longrightarrow -\lambda,$$

where  $\mathcal{U}^T$  is defined in (2.3). The second one is the existence of corrector functions. We say that a continuous function  $\chi$  on  $\mathcal{P}(\mathbb{T}^d)$  is a corrector function if it verifies the following dynamic programming principle

$$\chi(m_0) = \inf_{(m, \alpha)} \left( \int_0^t H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t \quad (2.4)$$

where  $(m, \alpha)$  solves in the sense of distributions  $-\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0$  with initial condition  $m(0) = m_0$ . At the heuristic level, this amounts to say that  $\chi$  solves the ergodic problem

$$\int_{\mathbb{T}^d} (H(y, D_m \chi(m, y)) - \operatorname{div}_y D_m \chi(m, y)) m(dy) = \mathcal{F}(m) + \lambda \quad \text{in } \mathcal{P}(\mathbb{T}^d). \quad (2.5)$$

(the notion of derivative  $D_m \chi$  is described in Section 2.1 below).

The main results of this paper are Theorem 2.5.7 and Theorem 2.5.9. The first one states that  $\mathcal{U}^T(0, \cdot) + \lambda T$  uniformly converges to a corrector function while the second one ensures that this convergence does not hold only at the level of minimization problems but also when it comes to optimal trajectories. In particular, Theorem 2.5.9 says that optimal trajectories for  $\mathcal{U}^T(0, \cdot) + \lambda T$  converge to calibrated curves (i.e., roughly speaking, to global minimizers of (2.4)). Let us recall that, in [68], the second author provides examples in which the calibrated curves stay away from the solutions of the MFG ergodic system (2.2). In that framework, our result implies that no solution to the MFG system (2.1) obtained as minimizers of (2.3) converges to a solution of the MFG ergodic system.

The convergence of  $\mathcal{U}^T(0, \cdot) + \lambda T$  to a corrector is of course the transposition, in our setting, of Fathi's famous convergence result for Hamilton-Jacobi equations [41]. The basic strategy of

proof is roughly the same. Here, the additional difficulty lies in the fact that, in our infinitely dimensional framework, the Hamiltonian in (2.5) is neither first order nor “uniformly elliptic” (cf. the term in divergence in (2.5)).

We overcome this difficulty by introducing two main ideas that we now describe. As in [41], we start with further characterizations of the limit value  $\lambda$ . Let us set

$$I := \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} (H(y, D_m \Phi(m, y)) - \mathcal{F}(m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy).$$

Then, by duality techniques, one can check the following equality (Proposition 2.2.2)

$$-I = \min_{(\mu, p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \left( H^* \left( y, \frac{dp_1}{dm \otimes \mu} \right) + \mathcal{F}(m) \right) m(dy) \mu(dm), \quad (2.6)$$

where  $(\mu, p_1)$  are closed measures, in the sense that, for any  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ ,

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi(m, y) m(dy) \mu(dm) = 0.$$

We call Mather measure any couple  $(\mu, p_1)$  which minimizes the dual problem (on this terminology, see Remark 2.5.1).

One key step is to show that  $I = \lambda$ . While it is easy to prove that  $I \geq \lambda$  (Proposition 2.2.1), the opposite inequality is trickier. One has to construct a smooth subsolution of the ergodic problem (2.5) in a context where there is no “classical” convolution. The idea is to look at a finite particle system on  $(\mathbb{T}^d)^N$ . A similar idea was used in [46] for first order problems on the  $L^2(0, 1)$ -torus. The main difference with [46] is that, for first order problems, the particle system is embedded into the continuous one, which is not the case for problems with diffusion. The argument of proof is therefore completely different. We set  $(v^N, \lambda^N) \in C^2((\mathbb{T}^d)^N) \times \mathbb{R}$  solution of

$$- \sum_{i=1}^N \Delta_{x_i} v^N(\mathbf{x}) + \frac{1}{N} \sum_{i=1}^N H(x_i, N D_{x_i} v^N(\mathbf{x})) = \mathcal{F}(m_{\mathbf{x}}^N) + \lambda^N.$$

Note that, in contrast with [46], the constant  $\lambda^N$ , here, depends on  $N$ . Our first main idea is to introduce the smooth function on  $\mathcal{P}(\mathbb{T}^d)$

$$W^N(m) := \int_{(\mathbb{T}^d)^N} v^N(x_1, \dots, x_N) \prod_{i=1}^N m(dx_i)$$

and to show that it satisfies in  $\mathcal{P}(\mathbb{T}^d)$

$$- \int_{\mathbb{T}^d} \operatorname{div}_y D_m W^N(m, y) m(dy) + \int_{\mathbb{T}^d} H(y, D_m W^N(m, y), m) m(dy) \leq \mathcal{F}(m) + \lambda^N + o(N),$$

which implies that  $I \leq \liminf_N \lambda^N$ . To conclude that  $I = \lambda$  we proved that  $\lambda^N \rightarrow \lambda$ . The proof of this result is organized in two steps. The first one (Lemma 2.3.2) is inspired by [67] and consists in deriving, through Berstein’s method, estimates on  $v^N$  of the form

$$N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 \leq \bar{C}_1 \quad \forall \mathbf{x} \in (\mathbb{T}^d)^N$$

and to derive from this the fact that  $v^N$  uniformly converges to a Lipschitz function  $V$  on  $\mathcal{P}(\mathbb{T}^d)$ . In the second one (Proposition 2.3.4) we adapt the argument of [59], which is concerned with

the connection between optimal control of McKean-Vlasov dynamics and the limit behavior of large number of interacting controlled state processes, to show that if  $\lambda^*$  is an accumulation point of  $\lambda^N$  then

$$V(m_0) = \inf_{(m, \alpha)} \left( \int_0^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + V(m(T)) \right) + \lambda^* T,$$

where  $(m, \alpha)$  solves in the sense of distributions  $-\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0$  with initial condition  $m_0$ . Consequently  $\lambda^* = \lambda$  and so  $\lambda^N \rightarrow \lambda$ .

The next difficulty is that the Hamiltonian appearing in (2.5) is singular (because of the divergence term). This prevents us to say, as in the classical setting, that Mather measures are supported by graphs on which the Hamilton-Jacobi is somehow satisfied. To overcome this issue, we introduce our second main idea, the notion of “smooth” Mather measures (measures supported by “smooth” probability measures). We prove that limits of minimizers for  $\mathcal{U}^T$  provide indeed “smooth” Mather measures and that, if  $(\mu, p_1)$  is a “smooth” Mather measure then, if we set

$$q_1(y, m) := D_a H^* \left( y, \frac{dp_1}{dm \otimes \mu}(y, m) \right),$$

we have, for  $\mu$ -a.e.  $m \in \mathcal{P}(\mathbb{T}^d)$ ,

$$\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} H(y, q_1(y, m)) m(dy) = \mathcal{F}(m) + \lambda. \quad (2.7)$$

(see Proposition 2.4.2). Note that (2.7) is a kind of reformulation of the ergodic equation (2.5), in which  $q_1 = D_m \chi$  and where the divergence term is integrated by parts. The rest of the proof is more standard and does not bring new difficulties compared to [41].

Let us briefly describe the organization of the paper. In Section 2.1, we fix the main notation and assumption and collect the results of [68] that we sketched above. Section 2.2 and 2.3 focus on further characterizations of the limit value  $\lambda$ . In particular, in Section 2.2, we prove that (2.6) and  $I \geq \lambda$  hold, while Section 2.3 is devoted to the analysis of the particle system and the proof that  $I = \lambda$ . Section 2.4 gives a closer look to Mather measures and explains (2.7). Section 2.5 contains Theorem 2.5.7 and Theorem 2.5.9 and their proofs.

**Acknowledgement:** The first author was partially supported by the ANR (Agence Nationale de la Recherche) project ANR-16-CE40-0015-01 and by the AFOSR grant FA9550-18-1-0494. We would like to thank Marco Cirant who carefully read this manuscript and pointed to a blunder in the previous version of Section 2.3.

## 2.1 Assumptions and preliminary results

The aim of this preliminary section is twofold. Firstly, we introduce the notation and the assumptions that we will use throughout the paper. Then, we collect some results from [68] which are the starting point of this work.

### 2.1.1 Notation and assumptions

We work on the  $d$ -dimensional flat torus  $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$  to avoid boundary conditions and to set the problem on a compact domain. We denote by  $\mathcal{P}(\mathbb{T}^d)$  the set of Borel probability measures

on  $\mathbb{T}^d$ . This is a compact, complete and separable set when endowed with the 1-Wasserstein distance  $\mathbf{d}(\cdot, \cdot)$ . Let  $m$  be a Borel measure over  $[t, T] \times \mathbb{T}^d$ , with first marginal the Lebesgue measure  $ds$  over  $[t, T]$ , then with  $\{m(s)\}_{s \in [t, T]}$  we denote the disintegration of  $m$  with respect to  $dt$ . We will always consider measures  $m$  such that  $m(s)$  is a probability measure on  $\mathbb{T}^d$  for any  $s \in [t, T]$ .

If  $m$  is such a measure, then  $L_m^2([t, T] \times \mathbb{T}^d)$  is the set of  $m$ -measurable functions  $f$  such that the integral of  $|f|^2 dm(s)$  over  $[t, T] \times \mathbb{T}^d$  is finite.

We use throughout the paper the notion of derivative for functions defined on  $\mathcal{P}(\mathbb{T}^d)$  introduced in [19]. We say that  $\Phi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is  $C^1$  if there exists a continuous function  $\frac{\delta\Phi}{\delta m} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}$  such that

$$\Phi(m_1) - \Phi(m_2) = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta\Phi}{\delta m}((1-t)m_2 + tm_1, x)(m_1 - m_2)(dx)dt, \quad \forall m_1, m_2 \in \mathcal{P}(\mathbb{T}^d).$$

As this derivative is defined up to an additive constant, we use the standard normalization

$$\int_{\mathbb{T}^d} \frac{\delta\Phi}{\delta m}(m, x)m(dx) = 0. \quad (2.8)$$

We recall that, if  $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$ , the 1-Wasserstein distance is defined by

$$\mathbf{d}(\mu, \nu) = \sup \left\{ \int_{\mathbb{T}^d} \phi(x) d(\mu - \nu)(x) \mid \text{continuous } \phi : \mathbb{T}^d \rightarrow \mathbb{R}, \text{Lip}(\phi) \leq 1 \right\}.$$

**Assumptions:** Throughout the paper the following conditions will be in place.

1.  $H : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is of class  $C^2$ ,  $p \mapsto D_{pp}H(x, p)$  is Lipschitz continuous, uniformly with respect to  $x$ . Moreover, there exists  $\bar{C} > 0$  that verifies

$$\bar{C}^{-1}I_d \leq D_{pp}H(x, p) \leq \bar{C}I_d, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d$$

and  $\theta \in (0, 1)$ ,  $C > 0$  such that the following conditions hold true:

$$|D_x H(x, p)| \leq C + C|p| \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d \quad (2.9)$$

and

$$|D_{xx}H(x, p)| \leq C(1 + |p|)^{1+\theta}, \quad |D_{x,p}H(x, p)| \leq C(1 + |p|)^\theta, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d.$$

2.  $\mathcal{F} : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is of class  $C^2$ . Its derivative  $F : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is twice differentiable in  $x$  and  $D_{xx}^2 F$  is bounded. Examples of non monotone coupling functions which verify such conditions can be found in [68].

Note that some of the above assumptions will not be used explicitly in this paper but have been used in [68] to prove results that we will assume to hold true. (The only differences are assumption (2.9) and that here we need  $\mathcal{F}$  to be of class  $C^2$ , in order to build the  $C^2$  test functions of Lemma 2.6.2, while in [68]  $\mathcal{F}$  was required only to be  $C^1$ ).

Very often in the text, we do not need to work explicitly on  $\mathcal{F}$ , so, in order to have a lighter notation, we incorporate  $\mathcal{F}$  in the Hamiltonian defining, for any  $(x, p, m) \in \mathbb{T}^d \times \mathbb{R}^d \times \mathcal{P}(\mathbb{T}^d)$ ,

$$\mathcal{H}(x, p, m) := H(x, p) - \mathcal{F}(m). \quad (2.10)$$

We denote with  $\mathcal{H}^*$  the Fenchel conjugate of  $\mathcal{H}$  with respect to the second variable. Then, for any  $(x, a, m) \in \mathbb{T}^d \times \mathbb{R}^d \times \mathcal{P}(\mathbb{T}^d)$ ,

$$\mathcal{H}^*(x, a, m) = H^*(x, a) + \mathcal{F}(m).$$

We can now introduce the standard minimization problem in potential MFG:

$$\mathcal{U}^T(t, m_0) = \inf_{(m, \alpha)} \int_t^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds, \quad m_0 \in \mathcal{P}(\mathbb{T}^d)$$

where  $m \in C^0([t, T], \mathcal{P}(\mathbb{T}^d))$ ,  $\alpha \in L_m^2([t, T] \times \mathbb{T}^d, \mathbb{R}^d)$  and the following equation is verified in the sense of distributions

$$\begin{cases} -\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0 & \text{in } [t, T] \times \mathbb{T}^d \\ m(t) = m_0 & \text{in } \mathbb{T}^d. \end{cases} \quad (2.11)$$

### 2.1.2 Corrector functions and the limit value $\lambda$

Here we collect the results already proved in [68] that we will use. A most important one is the following.

**Theorem 2.1.1.** *The function  $\frac{1}{T}\mathcal{U}^T(t, \cdot)$  uniformly converges to a limit value  $-\lambda$  when  $T$  goes to  $+\infty$ .*

The second result that we will use is the existence of corrector functions.

**Definition 2.1.2.** *We say that  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is a corrector function if, for any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and any  $t > 0$ ,*

$$\chi(m_0) = \inf_{(m, \alpha)} \left( \int_0^t \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t, \quad (2.12)$$

where  $m \in C^0([0, t], \mathcal{P}(\mathbb{T}^d))$ ,  $\alpha \in L_m^2([0, t] \times \mathbb{T}^d, \mathbb{R}^d)$  and the pair  $(m, \alpha)$  solves in the sense of distributions  $-\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0$  with initial condition  $m_0$ .

**Proposition 2.1.3.** *The set of corrector functions is not empty and uniformly Lipschitz continuous. In addition, if  $\chi$  is a continuous map on  $\mathcal{P}(\mathbb{T}^d)$  for which equality (2.12) holds for some constant  $\lambda'$  and for any  $t > 0$ , then  $\lambda' = \lambda$ .*

A last notion that will come at hand is the one of calibrated curve:

**Definition 2.1.4.** *We say that  $(\bar{m}, \bar{\alpha})$ , which satisfies (2.11) for any  $t \in \mathbb{R}$ , is a calibrated curve if there exists a corrector function  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  such that  $(\bar{m}, \bar{\alpha})$  is optimal for  $\chi$ : for any  $t_1 < t_2 \in \mathbb{R}$*

$$\chi(\bar{m}(t_1)) = \lambda(t_2 - t_1) + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(s)) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + \chi(\bar{m}(t_2)).$$

The set of calibrated curves verifies the following property.

**Proposition 2.1.5.** *The set of calibrated curves is not empty. Moreover, if  $(m, \alpha)$  is a calibrated curve, then  $m \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$  and there exists a function  $u \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$  such that  $\alpha = D_p H(x, Du)$  where  $(u, m)$  solves*

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



## 2.2 A dual problem

In this section we introduce the two usual characterizations of the constant  $\lambda$ :  $\lambda$  is expected to be the smallest constant for which there exists a smooth sub-corrector and  $-\lambda$  is the smallest value of the Lagrangian when integrated against suitable “closed” measures. The goal of this section is to show that both problems are in duality and have the same value  $I$ . We postpone the analysis of the equality  $I = \lambda$  to the next section.

We start with the HJ equation which we write in variational form.

$$I := \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m \Phi(m, y), m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy), \quad (2.13)$$

where by  $C^{1,1}(\mathcal{P}(\mathbb{T}^d))$  we mean the set of maps  $\Phi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  such that  $D_m \Phi$  and  $D_y D_m \Phi$  are continuous. We recall that  $\mathcal{H}$  is defined in (2.10). Let us start with a comparison between  $I$  and  $\lambda$ .

**Proposition 2.2.1.** *We have  $I \geq \lambda$ .*

*Proof.* Let  $\epsilon > 0$  and  $\Phi$  be such that

$$\int_{\mathbb{T}^d} (\mathcal{H}(y, D_m \Phi(m, y), m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy) \leq I + \epsilon \quad \forall m \in \mathcal{P}(\mathbb{T}^d).$$

Let  $(\bar{m}, \bar{\alpha})$  be a calibrated curve and  $\chi$  be a corrector function associated with  $(\bar{m}, \bar{\alpha})$ . By definition of calibrated curve,  $(\bar{m}, \bar{\alpha})$  verifies

$$-\partial_t \bar{m}(t) + \Delta \bar{m}(t) + \operatorname{div}(\bar{\alpha}(t) \bar{m}(t)) = 0, \quad \forall t \in \mathbb{R}$$

and

$$\chi(\bar{m}(0)) = \int_0^T \int_{\mathbb{T}^d} \mathcal{H}^*(y, \bar{\alpha}(t, y), \bar{m}(t)) \bar{m}(t, dy) + \lambda T + \chi(\bar{m}(T)). \quad (2.14)$$

As  $\Phi$  is smooth, we get

$$\begin{aligned} \Phi(\bar{m}(T)) - \Phi(\bar{m}(0)) &= \int_0^T \frac{d}{dt} \Phi(\bar{m}(t)) dt = \int_0^T \int_{\mathbb{T}^d} \frac{\delta \Phi}{\delta m}(\bar{m}(t), y) \partial_t \bar{m}(t) dy dt \\ &= \int_0^T \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi(\bar{m}(t), y) \bar{m}(t, dy) dt - \int_0^T \int_{\mathbb{T}^d} D_m \Phi(\bar{m}(t), y) \cdot \bar{\alpha}(t, y) \bar{m}(t, dy) dt \\ &\geq -(I + \epsilon)T + \int_0^T \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m \Phi(\bar{m}(t), y), \bar{m}(t)) - D_m \Phi(\bar{m}(t), y) \cdot \bar{\alpha}(t, y)) \bar{m}(t, dy) dt \\ &\geq -(I + \epsilon)T - \int_0^T \int_{\mathbb{T}^d} \mathcal{H}^*(y, \bar{\alpha}(t, y), \bar{m}(t)) \bar{m}(t, dy). \end{aligned}$$

Using (2.14), we end up with

$$\begin{aligned} \Phi(\bar{m}(T)) - \Phi(\bar{m}(0)) &\geq -(I + \epsilon)T - \int_0^T \int_{\mathbb{T}^d} \mathcal{H}^*(y, \bar{\alpha}(t, y), \bar{m}(t)) \bar{m}(t, dy) \\ &= -(I + \epsilon)T + \chi(\bar{m}(T)) - \chi(\bar{m}(0)) + \lambda T. \end{aligned}$$

Using the fact that  $\chi$  and  $\Phi$  are bounded, we divide both sides by  $T$  and we conclude that  $\lambda \leq I$  by letting  $T \rightarrow +\infty$  and  $\epsilon \rightarrow 0$ .  $\square$

Next we reformulate  $I$  in terms of "closed measures".

**Proposition 2.2.2.** *We have*

$$-I = \min_{(\mu, p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}, m\right) m(dy) \mu(dm), \quad (2.15)$$

where the minimum is taken over  $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$  and  $p_1$ , Borel vector measure on  $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$ , such that  $p_1$  is absolutely continuous with respect to the measure  $dm \otimes \mu := m(dy)\mu(dm)$  and such that  $(\mu, p_1)$  is closed, in the sense that, for any  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ ,

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi(m, y) m(dy) \mu(dm) = 0. \quad (2.16)$$

In analogy with weak KAM theory, we call a measure  $(\mu, p_1)$  satisfying (2.16) a closed measure and we introduce following definition.

**Definition 2.2.3.** *We say that a closed measure  $(\mu, p_1)$  is a Mather measure if it is minimizer of (2.15).*

*Proof.* As usual we can rewrite  $I$  as

$$I = \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \sup_{\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))} \int_{\mathcal{P}(\mathbb{T}^d)} \left( \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m \Phi(m, y), m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy) \right) \mu(dm).$$

We claim that

$$I = \max_{\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))} \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \int_{\mathcal{P}(\mathbb{T}^d)} \left( \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m \Phi(m, y), m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy) \right) \mu(dm). \quad (2.17)$$

Indeed,  $\mathcal{P}(\mathcal{P}(\mathbb{T}^d))$  is a compact subspace of  $\mathcal{M}(\mathcal{P}(\mathbb{T}^d))$  and, for any fixed  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$  the function on  $\mathcal{M}(\mathcal{P}(\mathbb{T}^d))$  defined by

$$\mu \mapsto \int_{\mathcal{P}(\mathbb{T}^d)} \left( \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m \Phi(m, y), m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy) \right) \mu(dm)$$

is continuous and concave (as it is linear). On the other hand, when we fix  $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ , the function on  $C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ , defined by

$$\Phi \mapsto \int_{\mathcal{P}(\mathbb{T}^d)} \left( \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m \Phi(m, y), m) - \operatorname{div}_y D_m \Phi(m, y)) m(dy) \right) \mu(dm),$$

is continuous with respect to the uniform convergence in  $C^{1,1}(\mathcal{P}(\mathbb{T}^d))$  and convex due to the convexity of  $\mathcal{H}$ . Therefore, the hypothesis of Sion's min-max Theorem are fulfilled and (2.17) holds true.

We now define the continuous linear map  $\Lambda : C^{1,1}(\mathcal{P}(\mathbb{T}^d)) \rightarrow (C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d \times C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)$  by

$$\Lambda(\Phi) = (D_m \Phi, \operatorname{div}_y D_m \Phi).$$

From now on we fix a maximizer  $\mu$  for (2.17) and we define  $E := C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ ,  $F := (C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d \times C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)$  and

$$f(\Phi) = 0, \quad g(a, b) = \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} (\mathcal{H}(y, a(m, y), m) - b(m, y)) m(dy) \mu(dm), \quad \forall (a, b) \in F.$$

We note that

$$I = \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \{f(\Phi) + g(\Lambda\Phi)\}.$$

To use the Fenchel-Rockafellar theorem we need to check the transversality conditions. These hypothesis are easily verified, indeed, both  $f$  and  $g$  are continuous and, therefore, proper functions. The function  $f$  is convex because it is linear and so is  $g$ , due to the convexity of Hamiltonian  $\mathcal{H}$ . Moreover, it comes directly from its definition that  $\Lambda$  is a bounded linear functional on  $E$ . Then, the Fenchel-Rockafellar Theorem states that

$$I = - \min_{(p_1, p_2) \in F'} \{f^*(-\Lambda^*(p_1, p_2)) + g^*(p_1, p_2)\}.$$

Note that

$$F' = (\mathcal{M}(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d \times \mathcal{M}(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d),$$

that  $f^*(q) = 0$  if  $q = 0$ ,  $f^*(q) = +\infty$  otherwise. So, for any  $(p_1, p_2) \in F'$ ,

$$\begin{aligned} f^*(-\Lambda(p_1, p_2)) &= \sup_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} -\langle \Lambda^*(p_1, p_2), \Phi \rangle - f(\Phi) = \sup_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} -\langle (p_1, p_2), \Lambda(\Phi) \rangle \\ &= \sup_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} - \int_{(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)} D_m \Phi(m, y) \cdot p_1(dm, dy) - \int_{(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)} \operatorname{div}_y D_m \Phi(m, y) p_2(dm, dy) \end{aligned}$$

which is 0 if, for any  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ ,

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi(m, y) p_2(dm, dy) = 0,$$

and  $+\infty$  otherwise. On the other hand,

$$\begin{aligned} g^*(p_1, p_2) &= \sup_{(a, b) \in F} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} b(m, y) p_2(dm, dy) \\ &\quad - \int_{\mathcal{P}(\mathbb{T}^d)} \left( \int_{\mathbb{T}^d} (\mathcal{H}(y, a(m, y), m) - b(m, y)) m(dy) \right) \mu(dm). \end{aligned}$$

So, if  $g^*(p_1, p_2)$  is finite, one must have that  $p_2(dm, dy) = -m(dy)\mu(dm)$  and that  $p_1$  is absolutely continuous with respect to the measure  $dm \otimes \mu := m(dy)\mu(dm)$ . Indeed, if  $p_1$  were not absolutely continuous with respect to  $dm \otimes \mu$ , we could find a sequence of continuous functions  $a_n \in (C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d$  such that,  $a_n$  is uniformly bounded on the support of  $dm \otimes \mu$  and

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a_n(m, y) \cdot p_1(dm, dy) \rightarrow +\infty.$$

But then we would have  $g^*(p_1, p_2) = +\infty$ . So,

$$g^*(p_1, p_2) = \sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)^d} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left( a(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a(m, y), m) \right) m(dy) \mu(dm)$$

We now want to prove that

$$\begin{aligned} &\sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)^d} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left( a(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a(m, y), m) \right) m(dy) \mu(dm) \\ &= \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^* \left( y, \frac{dp_1}{dm \otimes \mu}(m, y), m \right) m(dy) \mu(dm). \end{aligned}$$

We have one inequality by definition of Fenchel's conjugate. Indeed,

$$\begin{aligned} & \sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left( a(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a(m, y), m) \right) m(dy) \mu(dm) \\ & \leq \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a^*(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a^*(m, y), m) m(dy) \mu(dm) \\ & = \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^* \left( y, \frac{dp_1}{dm \otimes \mu}(m, y), m \right) m(dy) \mu(dm), \end{aligned}$$

where

$$a^*(m, y) = D_a \mathcal{H}^* \left( y, \frac{dp_1}{dm \otimes \mu}(m, y), m \right).$$

For the opposite inequality we use a density argument. The function  $a^*$  could be not continuous but yet it must be measurable. Moreover, the growth of  $\mathcal{H}$  ensures that  $a^* \in L^2_\mu(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d; \mathbb{R}^d)$ . As  $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$  is a compact Hausdorff space, the set of continuous functions is dense in  $L^2_\mu(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)$ . Let  $a_n \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d; \mathbb{R}^d)$  be such that  $a_n \rightarrow a^*$  in  $L^2_\mu$ . Then,

$$\begin{aligned} & \sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a(m, y), m) m(dy) \mu(dm) \geq \\ & \lim_{n \rightarrow +\infty} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a_n(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a_n(m, y), m) m(dy) \mu(dm) = \\ & \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a^*(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a^*(m, y), m) m(dy) \mu(dm) = \\ & \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^* \left( y, \frac{dp_1}{dm \otimes \mu}(m, y), m \right) m(dy) \mu(dm). \end{aligned}$$

Therefore, we can conclude that

$$I = - \min_{(\mu, p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^* \left( y, \frac{dp_1}{dm \otimes \mu}, m \right) m(dy) \mu(dm),$$

where the minimum is taken over  $(\mu, p_1)$  satisfying condition (2.16).  $\square$

## 2.3 The $N$ -particle problem.

In the previous section, we introduced two problems in duality. These problems have a common value called  $I$  and we have checked that  $\lambda \leq I$ . The aim of this section is to show that there is actually an equality:  $\lambda = I$ . In the standard setting, this equality is proved by smoothing correctors by a convolution; by the convexity of the Hamiltonian, the smoothed corrector is a subsolution to the corrector equation (up to a small error term), thus providing a competitor for problem (2.13). In our framework, there is no exact equivalent of the convolution. We overcome this difficulty by considering the projection of the problem onto the set of empirical measures of size  $N$  (thus on  $(\mathbb{T}^d)^N$ ). For the  $N$ -particle problem, the corrector is smooth. We explain here that a suitable extension of this finite dimensional corrector to the set  $\mathcal{P}(\mathbb{T}^d)$  provides a smooth sub-corrector for the problem in  $\mathcal{P}(\mathbb{T}^d)$  when  $N$  is large. This shows the claimed equality and, in addition, the fact that the ergodic constant associated with the  $N$ -particle problem converges to  $\lambda$ .

More precisely, we consider  $v^N : (\mathbb{T}^d)^N \rightarrow \mathbb{R}$  the solution of:

$$-\sum_{i=1}^N \Delta_{x_i} v^N(\mathbf{x}) + \frac{1}{N} \sum_{i=1}^N \mathcal{H}(x_i, ND_{x_i} v^N(\mathbf{x}), m_{\mathbf{x}}^N) = \lambda^N, \quad (2.18)$$

where  $\mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{T}^d)^N$  and  $m_{\mathbf{x}}^N = N^{-1} \sum_{i=1}^N \delta_{x_i}$  and where  $\mathcal{H}$  is defined in (2.10). Let us recall that such a corrector exists (it is unique up to additive constants) and is smooth. To fix the ideas, we choose the solution  $v^N$  such that

$$\int_{(\mathbb{T}^d)^N} v^N = 0. \quad (2.19)$$

**Proposition 2.3.1.** *We have*

$$I \leq \liminf_{N \rightarrow +\infty} \lambda^N.$$

*Proof.* We define

$$W^N(m) := \int_{(\mathbb{T}^d)^N} v^N(x_1, \dots, x_N) \prod_{i=1}^N m(dx_i).$$

As  $v^N$  is smooth, it is clear that  $W^N$  is also smooth on  $\mathcal{P}(\mathbb{T}^d)$  and we have

$$D_m W^N(m, y) = \sum_{k=1}^N \int_{(\mathbb{T}^d)^{N-1}} D_{x_k} v^N(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_N) \prod_{i \neq k} m(dx_i)$$

and

$$\operatorname{div}_y D_m W^N(m, y) = \sum_{k=1}^N \int_{(\mathbb{T}^d)^{N-1}} \Delta_{x_k} v^N(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_N) \prod_{i \neq k} m(dx_i).$$

In view of the convexity of  $\mathcal{H}$  with respect to  $p$ , we obtain, for any  $m \in \mathcal{P}(\mathbb{T}^d)$ ,

$$\begin{aligned} & - \int_{\mathbb{T}^d} \operatorname{div}_y D_m W^N(m, y) m(dy) + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m W^N(m, y), m) m(dy) \\ &= - \sum_{k=1}^N \int_{\mathbb{T}^d} \int_{(\mathbb{T}^d)^{N-1}} \Delta_{x_k} v^N(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_N) \prod_{i \neq k} m(dx_i) m(dy) \\ & \quad + \int_{\mathbb{T}^d} \mathcal{H}\left(y, \sum_{k=1}^N \int_{(\mathbb{T}^d)^{N-1}} D_{x_k} v^N(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_N) \prod_{i \neq k} m(dx_i), m\right) m(dy) \\ & \leq \int_{(\mathbb{T}^d)^N} \left( - \sum_{k=1}^N \Delta_{x_k} v^N(x_1, \dots, x_N) + \frac{1}{N} \sum_{k=1}^N \mathcal{H}(x_k, ND_{x_k} v^N(x_1, \dots, x_N), m) \right) \prod_i m(dx_i). \end{aligned}$$

Following [31], the Glivenko-Cantelli Theorem states that

$$\int_{(\mathbb{T}^d)^N} \mathbf{d}_2(m_{\mathbf{x}}^N, m) \prod_i m(dx_i) \leq \epsilon_N := \begin{cases} N^{-1/2} & \text{if } d < 4, \\ N^{-1/2} \ln(N) & \text{if } d = 4, \\ N^{-2/d} & \text{otherwise.} \end{cases}$$

As  $\mathcal{H}$  has a separate form:  $\mathcal{H}(x, p, m) = H(x, p) - \mathcal{F}(m)$  where  $\mathcal{F}$  is Lipschitz continuous with respect to  $m$ , we infer that

$$\begin{aligned} & \int_{(\mathbb{T}^d)^N} \frac{1}{N} \sum_{k=1}^N \mathcal{H}(x_k, ND_{x_k} v^N(x_1, \dots, x_N), m) \prod_i m(dx_i) \\ & \leq \int_{(\mathbb{T}^d)^N} \frac{1}{N} \sum_{k=1}^N \mathcal{H}(x_k, ND_{x_k} v^N(x_1, \dots, x_N), m_{\mathbf{x}}^N) \prod_i m(dx_i) + C\epsilon_N. \end{aligned}$$

Recalling the equation satisfied by  $v^N$ , we conclude that

$$\begin{aligned} & - \int_{\mathbb{T}^d} \operatorname{div}_y D_m W^N(m, y) m(dy) + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m W^N(m, y), m) m(dy) \\ & \leq \int_{(\mathbb{T}^d)^N} \left( - \sum_{k=1}^N \Delta_{x_k} v^N(x_1, \dots, x_N) + \frac{1}{N} \sum_{k=1}^N \mathcal{H}(x_k, ND_{x_k} v^N(x_1, \dots, x_N), m_{\mathbf{x}}^N) \right) \prod_i m(dx_i) + C\epsilon_N \\ & \leq \lambda^N + C\epsilon_N. \end{aligned}$$

As  $W^N$  is smooth, this shows that  $I \leq \liminf_N \lambda^N$ .  $\square$

According to the above proposition and Proposition 2.2.1 we have that  $\lambda \leq I \leq \liminf_N \lambda^N$ . Therefore, to have that  $I = \lambda$  we need to show that  $\lim_N \lambda^N = \lambda$ . Before proving this result we introduce some estimates on  $v^N$ .

**Lemma 2.3.2.** *There exists a constant  $\bar{C}_1 > 0$ , independent of  $N$ , such that*

$$N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 \leq \bar{C}_1 \quad \forall \mathbf{x} \in (\mathbb{T}^d)^N. \quad (2.20)$$

*Proof.* The proof is an application of the Bernstein's type estimates used in [67]. First we show that the sequence  $\lambda^N$  is bounded. This is a direct consequence of the maximum principle. Indeed, if  $\mathbf{y}$  is a minimum point of  $v^N$ , then

$$\frac{1}{N} \sum_{i=1}^N \mathcal{H}(x_i, 0, m_{\mathbf{y}}^N) \geq - \sum_{i=1}^N \Delta_{x_i} v^N(\mathbf{y}) + \frac{1}{N} \sum_{i=1}^N \mathcal{H}(x_i, 0, m_{\mathbf{y}}^N) = \lambda^N.$$

As the left hand side is uniformly bounded with respect to  $N$  we get that  $\lambda^N$  is bounded from above. If instead of the minimum we consider a maximum we get the lower bound.

From (2.18) and the quadratic growth of the Hamiltonian we have

$$\begin{aligned} \bar{C}N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 - \bar{C} & \leq \sum_{i=1}^N \Delta_{x_i} v^N(\mathbf{x}) + \|\mathcal{F}\|_{\infty} + \lambda^N \leq \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})| + \|\mathcal{F}\|_{\infty} + \lambda^N \\ & \leq N^{\frac{1}{2}} \left( \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})|^2 \right)^{\frac{1}{2}} + \|\mathcal{F}\|_{\infty} + \lambda^N. \end{aligned}$$

As  $\lambda^N$  is bounded, there exists a constant  $C$  such that

$$\bar{C}N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 \leq N^{\frac{1}{2}} \left( \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})|^2 \right)^{\frac{1}{2}} + C \quad (2.21)$$

Now we define

$$w(\mathbf{x}) = \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2,$$

then

$$D_{x_j} w(\mathbf{x}) = 2 \sum_{i=1}^N D_{x_i} v^N(\mathbf{x}) D_{x_j, x_j}^2 v^N(\mathbf{x}).$$

By a direct computation we get

$$\begin{aligned} - \sum_{i=1}^N \Delta_{x_i} w(\mathbf{x}) &= -2 \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})|^2 + \sum_{i=1}^N D_{x_i} v^N(\mathbf{x}) D_{x_i} \left( \sum_{j=1}^N \Delta_{x_j} v^N(\mathbf{x}) \right) \\ &= -2 \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})|^2 + \sum_{i=1}^N D_{x_i} v^N(\mathbf{x}) D_{x_i} \left( N^{-1} \left( \sum_{j=1}^N H(x_j, N D_{x_j} v^N(\mathbf{x})) - \mathcal{F}(m_{\mathbf{x}}^N) \right) - \lambda^N \right) \\ &= -2 \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})|^2 + N^{-1} \sum_{i=1}^N D_{x_i} v^N(\mathbf{x}) D_x H(x_i, N D_{x_i} v^N(\mathbf{x})) \\ &\quad + \sum_{i,j=1}^N D_p H(x_j, N \Delta_{x_j} v^N(\mathbf{x})) D_{x_i, x_j}^2 v^N(\mathbf{x}) D_{x_i} v^N(\mathbf{x}) - N^{-1} \sum_{i=1}^N D_x F(m_{x_i}^N, x_i) D_{x_i} v^N(\mathbf{x}) \\ &= -2 \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})|^2 + N^{-1} \sum_{i=1}^N D_{x_i} v^N(\mathbf{x}) D_x H(x_i, N D_{x_i} v^N(\mathbf{x})) \\ &\quad + \sum_{j=1}^N D_p H(x_j, N \Delta_{x_j} v^N(\mathbf{x})) D_{x_j} w(\mathbf{x}) - N^{-1} \sum_{i=1}^N D_x F(m_{\mathbf{x}}^N, x_i) D_{x_i} v^N(\mathbf{x}) \end{aligned}$$

If  $\mathbf{x}$  is a maximum point of  $w$ , the above computations lead to

$$\begin{aligned} 2 \sum_{i,j=1}^N |D_{x_i, x_j}^2 v^N(\mathbf{x})|^2 &\leq N^{-1} \left[ \sum_{i=1}^N D_{x_i} v^N(\mathbf{x}) D_x H(x_i, N D_{x_i} v^N(\mathbf{x})) - \sum_{i=1}^N D_x F(m_{\mathbf{x}}^N, x_i) D_{x_i} v^N(\mathbf{x}) \right] \\ &\leq N^{-1} \left[ \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})| |D_x H(x_i, N D_{x_i} v^N(\mathbf{x}))| + \|D_x F\|_{\infty} \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})| \right] \\ &\leq N^{-1} \left[ CN \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 + (\|D_x F\|_{\infty} + C) \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})| \right], \end{aligned}$$

where for the last line we used the bounds on  $D_x H$  (2.9).

As  $D_x F$  is bounded, plugging the above inequality into (2.21) we get (for a possible different

constant  $C$ , independent of  $N$ , that might change from line to line)

$$\begin{aligned} N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 &\leq C \left[ N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 + \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})| \right]^{\frac{1}{2}} + C \\ &\leq C \left( N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 \right)^{\frac{1}{2}} + C \left( N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 \right)^{\frac{1}{4}} + C. \end{aligned}$$

The above inequalities ensure that there exists a  $\bar{C}_1 > 0$ , independent of  $N$ , such that

$$N \sum_{i=1}^N |D_{x_i} v^N(\mathbf{x})|^2 \leq \bar{C}_1.$$

□

**Lemma 2.3.3.** *Let  $v^N$  be the solution to (2.18) satisfying condition (2.19). There exists a Lipschitz continuous map  $V : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  and a subsequence  $\{N_k\}_k$  such that*

$$\lim_{k \rightarrow +\infty} \sup_{\mathbf{x} \in (\mathbb{T}^d)^{N_k}} |v^{N_k}(\mathbf{x}) - V(m_{\mathbf{x}}^{N_k})| = 0$$

*Proof.* The proof is a direct application of a compactness result due to Lions (see [17, Theorem 2.1]), for which we just need to check the hypotheses. Let us note that, by uniqueness (up to a constant) of the solution to (2.18),  $v^N$  is symmetrical. To apply [17, Theorem 2.1] we have to prove that  $v^N$  is uniformly bounded and that there exists a constant  $C$ , independent of  $N$ , such that, for any  $\mathbf{x}, \mathbf{y} \in (\mathbb{T}^d)^N$

$$|v^N(\mathbf{x}) - v^N(\mathbf{y})| \leq C \mathbf{d}(m_{\mathbf{x}}^N, m_{\mathbf{y}}^N). \quad (2.22)$$

The fact that (2.22) and the normalization condition (2.19) hold implies that the  $v^N$  are uniformly bounded. To prove (2.22), we fix two points  $\mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{T}^d)^N$  and  $\mathbf{y} = (y_1, \dots, y_N) \in (\mathbb{T}^d)^N$ . As  $v^N$  is symmetrical, we can assume without loss of generality that  $\mathbf{d}_2(m_{\mathbf{x}}^N, m_{\mathbf{y}}^N) = (N^{-1} \sum_{i=1}^N \mathbf{d}_{\mathbb{T}^d}^2(x_i, y_i))^{1/2}$ , where  $\mathbf{d}_{\mathbb{T}^d}$  is the distance on  $\mathbb{T}^d$ . Then, using again (2.20), we have

$$\begin{aligned} |v^N(\mathbf{x}) - v^N(\mathbf{y})| &\leq \sup_{\mathbf{z} \in (\mathbb{T}^d)^N} \sum_{i=1}^N |D_{x_i} v^N(\mathbf{z})| \mathbf{d}_{\mathbb{T}^d}(x_i, y_i) \\ &\leq \sup_{\mathbf{z} \in (\mathbb{T}^d)^N} \left( \sum_{i=1}^N |D_{x_i} v^N(\mathbf{z})|^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^N \mathbf{d}_{\mathbb{T}^d}^2(x_i, y_i) \right)^{\frac{1}{2}} \\ &\leq \bar{C}_1^{\frac{1}{2}} N^{-\frac{1}{2}} \left( \sum_{i=1}^N \mathbf{d}_{\mathbb{T}^d}^2(x_i, y_i) \right)^{\frac{1}{2}} = \bar{C}_1^{\frac{1}{2}} \mathbf{d}_2(m_{\mathbf{x}}^N, m_{\mathbf{y}}^N) \leq C \mathbf{d}(m_{\mathbf{x}}^N, m_{\mathbf{y}}^N). \end{aligned}$$

□

We now adapt the argument of [59] to prove that the function  $V$  is a corrector.



**Proposition 2.3.4.** *Let  $\{\lambda^{N_k}\}_k$  be the subsequence such that  $(v^{N_k})$  converges to  $V$  as in Lemma 2.3.3 and let  $\lambda^*$  be an accumulation point of  $\{\lambda^{N_k}\}_k$ . Then, for any  $T > 0$  and any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ , the function  $V : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  verifies*

$$V(m_0) = \inf_{(m, \alpha)} \left( \int_0^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + V(m(T)) \right) + \lambda^* T,$$

where  $(m, \alpha)$  solves in the sense of distributions  $-\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0$  with initial condition  $m_0$ .

Consequently,  $\lambda^* = \lambda$  and  $\lambda^N \rightarrow \lambda$ .

*Proof.* The Proposition is a consequence of a nice result due to Lacker [59], which roughly says that the mean field limit of an optimal stochastic control problem involving symmetric controllers is an optimal control problem of a McKean-Vlasov equation. As the adaptation to our framework, if relatively easy, requires cumbersome details, we refer the interested reader to the appendix.

The fact that  $\lambda^* = \lambda$  is a consequence of the uniqueness of the ergodic constant as stated in Proposition 2.1.3. As any accumulation point of the bounded sequence  $\lambda^N$  is equal to  $\lambda$ , we finally conclude that the whole sequence converges to  $\lambda$ .  $\square$

As we mentioned before, the immediate consequence of the above proposition is the following result.

**Proposition 2.3.5.** *We have*

$$I = \lambda = \lim_{N \rightarrow +\infty} \lambda^N.$$

## 2.4 On the support of the Mather measures

In this section we take a closer look at Mather measures and the properties of their support points.

**Definition 2.4.1.** *We say that the closed measure  $(\mu, p_1)$  is smooth if there exists a constant  $C > 0$  such that, for  $\mu$ -a.e.  $m \in \mathcal{P}(\mathbb{T}^d)$ ,  $m$  has a positive density and*

$$\|D \ln(m)\|_{L_m^2(\mathbb{T}^d)} \leq C.$$

The aim of this section is to prove the following property of smooth Mather measures.

**Proposition 2.4.2.** *Let  $(\mu, p_1)$  be a smooth Mather measure. Let us set*

$$q_1(y, m) := D_a \mathcal{H}^* \left( y, \frac{dp_1}{dm \otimes \mu}(y, m), m \right).$$

Then we have, for  $\mu$ -a.e.  $m \in \mathcal{P}(\mathbb{T}^d)$ ,

$$\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} \mathcal{H}(y, q_1(y, m), m) m(dy) = \lambda.$$

In order to prove the proposition, let us start with a preliminary step. Let  $(\mu, p_1)$  be optimal in problem (2.15) (where we recall that  $I = \lambda$ ) and let  $\Phi^N$  be any minimizing sequence in (2.13).

**Lemma 2.4.3.** *The sequence  $(D_p \mathcal{H}(y, D_m \Phi^N(m, y)))$  converges to  $\frac{dp_1}{dm \otimes \mu}$  in  $L^2(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, dm \otimes \mu)$ . Moreover,*

$$\lim_{N \rightarrow +\infty} - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) = \lambda. \quad (2.23)$$

*Proof.* Recall that  $\Phi^N$  satisfies

$$- \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \leq \lambda + o_N(1). \quad (2.24)$$

We integrate equation (2.24) against  $\mu$  and add the problem for  $(\mu, p_1)$  to find:

$$\begin{aligned} & - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) \\ & + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}, m\right) m(dy) \mu(dm) \leq o_N(1). \end{aligned}$$

Using the uniform convexity of  $\mathcal{H}$ , this implies that

$$\begin{aligned} & - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \frac{dp_1}{dm \otimes \mu} \cdot D_m \Phi^N(m, y) m(dy) \mu(dm) \\ & + C^{-1} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left| \frac{dp_1}{dm \otimes \mu} - D_p \mathcal{H}(y, D_m \Phi^N(m, y), m) \right|^2 m(dy) \mu(dm) \leq o_N(1). \end{aligned}$$

Then, (2.16) implies that the first line vanishes and so

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left| \frac{dp_1}{dm \otimes \mu} - D_p \mathcal{H}(y, D_m \Phi^N(m, y), m) \right|^2 m(dy) \mu(dm) \leq o_N(1),$$

which proves the first statement of the lemma. We now turn to (2.23). First of all, as  $\Phi^N$  is a minimizing sequence for (2.13), we have that

$$\limsup_{N \rightarrow +\infty} - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) \leq \lambda.$$

To prove the other inequality we start with

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm).$$

We add and subtract the same quantity to get

$$\begin{aligned} & - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \frac{dp_1}{dm \otimes \mu} \cdot D_m \Phi^N(m, y) m(dy) \mu(dm) + \\ & \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \frac{dp_1}{dm \otimes \mu} \cdot D_m \Phi^N(m, y) m(dy) \mu(dm). \end{aligned}$$

As  $(\mu, p_1)$  verifies (2.16), the first line above vanishes. Then, using the Fenchel's inequality, we find that

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) \geq$$

$$- \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}, m\right) m(dy) \mu(dm).$$

By hypothesis,  $(\mu, p_1)$  is a minimizer for (2.15), so, for any  $N \in \mathbb{N}$ ,

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) \geq \lambda.$$

Therefore,

$$\liminf_{N \rightarrow +\infty} - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) \geq \lambda.$$

and the result follows.  $\square$

*Proof of Proposition 2.4.2.* From our assumption on  $(\mu, p_1)$ , we have

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) = - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi^N(m, y) \cdot Dm(y) dy \mu(dm).$$

As, by Lemma 2.4.3, the sequence  $(D_p \mathcal{H}(y, D_m \Phi^N(m, y)), m)$  converges to  $\frac{dp_1}{dm \otimes \mu}$  in  $L^2(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, dm \otimes \mu)$ , we also have that  $(D_m \Phi^N(m, y))$  converges to  $q_1$  in  $L^2(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, dm \otimes \mu)$  by regularity and invertibility of  $D_p \mathcal{H}$  and  $D_a \mathcal{H}^*$ . On the other hand,  $\frac{Dm(y)}{m(y)}$  is bounded in  $L^2_m(\mathbb{T}^d)$  for  $\mu$ -a.e.  $m \in \mathcal{P}(\mathbb{T}^d)$ . Therefore

$$\lim_{N \rightarrow +\infty} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) = - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} q_1(m, y) \cdot Dm(y) dy \mu(dm).$$

We conclude thanks to (2.23) that

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} q_1(m, y) \cdot Dm(y) dy \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, q_1(m, y), m) m(dy) \mu(dm) = \lambda.$$

On the other hand, extracting a subsequence if necessary, the sequence  $\{D_m \Phi^N(m, y)\}_N$  converges  $\mu$ -a.e. to  $q_1$ . So, by (2.24), we have, for  $m$ -a.e.  $m \in \mathcal{P}(\mathbb{T}^d)$ ,

$$\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} \mathcal{H}(y, q_1(y, m), m) m(dy) \leq \lambda.$$

Putting together the previous inequality with the previous equality gives the result.  $\square$

## 2.5 The long time behavior of potential MFG

In this section, we prove the two main results of the paper: the first one is the convergence, as  $T \rightarrow +\infty$ , of  $\mathcal{U}^T(0, \cdot) + \lambda T$ . The second one states that limits of time-dependent minimizing mean field games equilibria, as the horizon tends to infinity, are calibrated curves.

### 2.5.1 Convergence of $\mathcal{U}^T(0, \cdot) + \lambda T$

We recall that  $\mathcal{U}^T(t, m_0)$  is defined by

$$\mathcal{U}^T(t, m_0) = \inf_{(m, \alpha)} \int_t^T \left( \int_{\mathbb{T}^d} H^*(y, \alpha(s, y)) m(s, y) dy + \mathcal{F}(m(s)) \right) ds,$$

where  $(m, \alpha)$  verifies the usual constraint

$$\partial_t m - \Delta m - \operatorname{div}(m\alpha) = 0 \text{ in } (t, T) \times \mathbb{T}^d, \quad m(t) = m_0.$$

Let  $(m^T, \alpha^T)$  be a minimizer of the problem, then,  $\alpha^T(s, x) = D_p H(x, Du^T(s, x))$ , where  $(u^T, m^T)$  solves

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [t, T] \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [t, T] \\ m(t) = m_0, u(T, x) = 0 & \text{in } \mathbb{T}^d, \end{cases} \quad (2.25)$$

(see for instance [11] for details).

We also take from [68, Lemma 1.3] some uniform estimates on the solutions of (2.25) which will be useful in the next propositions.

**Lemma 2.5.1.** *There exists  $C > 0$  independent of  $m_0$  and  $T$  such that, if  $(u, m)$  is a classical solution of (2.25), then*

- $\|Du\|_{L^\infty([0, T] \times \mathbb{T}^d)} + \|D^2 u\|_{L^\infty([0, T] \times \mathbb{T}^d)} \leq C$
- $\mathbf{d}(m(s), m(l)) \leq C|l - s|^{1/2}$  for any  $l, s \in [0, T]$

Consequently, we also have that  $|\partial_t u(s, \cdot)| \leq C$  for any  $s \in [0, T]$ .

**Lemma 2.5.2.** *For any  $(u, m)$  solution of the MFG system (2.25), there exists  $c(u, m) \in \mathbb{R}$  such that, for any  $t \in [0, T]$ ,*

$$\int_{\mathbb{T}^d} \left( H(x, Du^T(t, x)) - \Delta u^T(t, x) \right) m^T(t, dx) - \mathcal{F}(m^T(t)) = c(m, u).$$

*Proof.* As for any  $t > 0$  both  $m^T$  and  $u^T$  are smooth in time and space, the integral

$$\int_{\mathbb{T}^d} H(x, Du(t, x)) m^T(t, x) + Du(t, x) \cdot Dm(t, x) dx - \mathcal{F}(m(t))$$

is well defined and we can derive it in time. Then,

$$\begin{aligned} & \int_{\mathbb{T}^d} D_p H(x, Du(t, x)) \partial_t Du(t, x) m(t, x) + H(x, Du(t, x)) \partial_t m(t, x) + \\ & \int_{\mathbb{T}^d} \partial_t Du(t, x) \cdot Dm(t, x) dx + Du(t, x) \cdot \partial_t Dm(t, x) - F(x, m(t)) \partial_t m(t, x). \end{aligned}$$

Integrating by parts and rearranging the terms we get that the above expression is equal to

$$\int_{\mathbb{T}^d} (-\Delta m(t, x) - \operatorname{div}(m(t, x) D_p H(x, Du^T(t, x)))) \partial_t u^T(t, x) dx +$$

$$(-\Delta u^T(t, x) + H(x, Du^T(t, x) - F(x, m^T(t, x)))\partial_t m^T(t, x)dx.$$

If we plug into the last equality the equations verified by  $u^T$  and  $m^T$ , we get

$$\begin{aligned} \frac{d}{dt} \left( \int_{\mathbb{T}^d} H(x, Du^T(t, x))m^T(t, x) + Du^T(t, x) \cdot Dm^T(t, x)dx - \mathcal{F}(m^T(t)) \right) = \\ \int_{\mathbb{T}^d} -\partial_t m^T(t, x)\partial_t u^T(t, x) + \partial_t m^T(t, x)\partial_t u^T(t, x) = 0. \end{aligned}$$

Integrating by parts the term  $Du^T(t, x) \cdot Dm^T(t, x)$  and using the continuity of  $t \rightarrow u(t, \cdot)$  in  $C^2(\mathbb{T}^d)$  and the continuity of  $t \rightarrow m(t)$  in  $\mathcal{P}(\mathbb{T}^d)$  we conclude that the result holds.  $\square$

**Proposition 2.5.3.** *Let  $(m^T, \alpha^T)$  be optimal for  $\mathcal{U}(T, m_0)$  and  $(u^T, m^T)$  be a solution of (2.25) associated to  $(m^T, \alpha^T)$ . Then,  $c(u^T, m^T) \rightarrow \lambda$  as  $T \rightarrow +\infty$ . Moreover, this limit is uniform with respect to the initial condition  $m_0$  and the choice of the minimizer  $(m^T, \alpha^T)$ .*

*Proof.* We prove it by contradiction. Let us suppose that there exist a sequence  $T_i \rightarrow +\infty$  and a sequence  $(m^i, \alpha^i)$ , minimizing  $\mathcal{U}^{T_i}(0, m_0^i)$ , such that, for any  $i \in \mathbb{N}$  and a some  $\varepsilon > 0$ ,

$$|c(u^i, m^i) - \lambda| \geq \varepsilon, \quad (2.26)$$

where, as usual,  $\alpha^i(s, x) = D_p H(x, Du^i(s, x))$  and  $(u^i, m^i)$  solves (2.25). Thanks to Lemma 2.5.1 we know that there exists  $C > 0$ , independent of  $i$  such that

$$\sup_{T_i > 0} \sup_{t \in [0, T_i]} \|\alpha^i(t)\|_\infty + \|D\alpha^i(t)\|_\infty \leq C.$$

Let  $E$  be the set

$$E := \{\alpha \in W^{1, \infty}(\mathbb{T}^d, \mathbb{R}^d), \|\alpha\|_\infty + \|D\alpha\|_\infty \leq C\}.$$

Then  $E$ , endowed with the topology of the uniform convergence, is compact. Moreover,  $\alpha^i(t) \in E$  for any  $t \in [0, T_i]$ . Let us define the probability measure  $\nu^i$  on  $\mathcal{P}(\mathbb{T}^d) \times E$  by

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} f(m, \alpha) \nu^i(dm, d\alpha) = \frac{1}{T_i - 1} \int_1^{T_i} f(m^i(t), \alpha^i(t)) dt \quad \forall f \in C(\mathcal{P}(\mathbb{T}^d) \times E).$$

As  $\mathcal{P}(\mathcal{P}(\mathbb{T}^d) \times E)$  is compact,  $\nu^i$  converges, up to a subsequence denoted in the same way, to some probability measure  $\nu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d) \times E)$ . Note that

$$\begin{aligned} \frac{1}{T_i} \mathcal{U}(0, m_0^i) &= \frac{1}{T_i} \int_0^1 \left( \int_{\mathbb{T}^d} H^*(y, \alpha^i(s, y)) m^i(s, y) dy \right) + \mathcal{F}(m^i(s)) ds \\ &\quad + \frac{T_i - 1}{T_i} \int_{\mathcal{P}(\mathbb{T}^d) \times E} \left( \int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right) \nu^i(dm, d\alpha) \end{aligned}$$

Hence, as the left-hand side converges, uniformly with respect to  $m_0^i$ , to  $-\lambda$  (see [68]), we obtain

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} \left( \int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right) \nu(dm, d\alpha) = -\lambda. \quad (2.27)$$

Now we make the link between  $\nu$  and the measure  $(\mu, p_1)$  of Section 2.4. Let  $\mu$  be the first marginal of  $\nu$  and let us define the vector measure  $p_1$  on  $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$  as

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \phi(m, y) \cdot p_1(dm, dy) = \int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \phi(m, y) \cdot \alpha(y) m(dy) \nu(dm, d\alpha)$$

for any test function  $\phi \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, \mathbb{R}^d)$ . We note that  $p_1$  is absolutely continuous with respect to  $\mu$ , since, if we disintegrate  $\nu$  with respect to  $\mu$ :  $\nu = \nu_m(d\alpha)\mu(dm)$ , then

$$p_1(dm, dy) = \int_E \alpha(y)m(dy)\nu_m(d\alpha)\mu(dm).$$

Therefore,

$$\frac{dp_1}{dm \otimes \mu}(m, y) = \int_E \alpha(y)\nu_m(d\alpha).$$

Let us check that  $(\mu, p_1)$  is closed. Indeed, for any map  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ , we have

$$\begin{aligned} \frac{d}{dt}\Phi(m^i(t)) &= \int_{\mathbb{T}^d} \operatorname{div}(D_m\Phi(m^i(t), y))m^i(t, dy) - \int_{\mathbb{T}^d} D_m\Phi(m^i(t), y) \cdot H_p(y, Du^i(t, y))m^i(dy) \\ &= \int_{\mathbb{T}^d} \operatorname{div}(D_m\Phi(m^i(t), y))m^i(t, dy) + \int_{\mathbb{T}^d} D_m\Phi(m^i(t), y) \cdot \alpha^i(t, y)m^i(dy). \end{aligned}$$

So,

$$\begin{aligned} &\int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \operatorname{div}(D_m\Phi(m, y)) + D_m\Phi(m, y) \cdot \alpha(y)m(dy) \nu^i(dm, d\alpha) \\ &= \frac{1}{T_i - 1} \int_1^{T_i} \int_{\mathbb{T}^d} \operatorname{div}(D_m\Phi(m^i(t), y))m^i(t, dy) + \int_{\mathbb{T}^d} D_m\Phi(m^i(t), y) \cdot \alpha^i(t, y)m^i(dy) dt \\ &= \frac{1}{T_i - 1} [\Phi(m^i(T_i)) - \Phi(m^i(1))]. \end{aligned}$$

Letting  $i \rightarrow +\infty$  gives

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \operatorname{div}(D_m\Phi(m, y)) + D_m\Phi(m, y) \cdot \alpha(y)m(dy) \nu(dm, d\alpha) = 0,$$

which can be rewritten, in view of the definition of  $p_1$ , as

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \operatorname{div}(D_m\Phi(m, y))m(dy)\mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m\Phi(m, y) \cdot p_1(dm, dy) = 0.$$

This proves that  $(\mu, p_1)$  is closed. Next we come back to (2.27): using the convexity of  $H^*$ , we also have

$$\begin{aligned} -\lambda &= \int_{\mathcal{P}(\mathbb{T}^d) \times E} \left[ \int_{\mathbb{T}^d} H^*(y, \alpha(y))m(dy) + \mathcal{F}(m) \right] \nu(dm, d\alpha) = \\ &\int_{\mathcal{P}(\mathbb{T}^d)} \int_E \left[ \int_{\mathbb{T}^d} H^*(y, \alpha(y))m(dy) + \mathcal{F}(m) \right] \nu_m(d\alpha)\mu(dm) \geq \\ &\int_{\mathcal{P}(\mathbb{T}^d)} \left[ \int_{\mathbb{T}^d} H^* \left( y, \int_E \alpha(y)\nu_m(d\alpha) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm) = \\ &\int_{\mathcal{P}(\mathbb{T}^d)} \left[ \int_{\mathbb{T}^d} H^* \left( y, \frac{dp_1}{dm \otimes \mu}(m, y) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm). \end{aligned}$$

Therefore,

$$\int_{\mathcal{P}(\mathbb{T}^d)} \left[ \int_{\mathbb{T}^d} H^* \left( y, \frac{dp_1}{dm \otimes \mu}(m, y) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm) \leq -\lambda, \quad (2.28)$$

which proves the minimality of  $(\mu, p_1)$ . By the uniform convexity of  $H$ , relation (2.28) shows also that, for  $\mu$ -a.e.  $m \in \mathcal{P}(\mathbb{T}^d)$  and for  $\nu_m$ -a.e.  $\alpha$ , one has

$$\frac{dp_1}{dm \otimes \mu}(m, y) = \alpha(y). \quad (2.29)$$

Note also that,  $m^i(t)$  has a positive density for any  $t \in [1, T_i]$  and there exists a constant  $C > 0$  independent of  $i$  such that

$$\sup_{t \in [1, T_i]} \|1/m^i(t, \cdot)\|_\infty + \|Dm^i(t, \cdot)\|_\infty \leq C. \quad (2.30)$$

The bounds on  $Dm^i$  are standard and we refer to [60, Ch4, Theorem 5.1]. While, for the estimates on  $1/m^i$ , we used the Harnack's inequality in [10, Theorem 8.1.3]. In our setting, this theorem states that, for any  $x, y \in \mathbb{T}^d$  and for any  $0 < s < t < T_i$ , there exists a constant  $C_{t-s}$ , depending only on  $|t - s|$ , such that

$$m^i(t, x) \geq C_{t-s} m^i(s, y).$$

As we work on the torus, for any  $s > 0$ , there exists a point  $y_s \in \mathbb{T}^d$  such that  $m^i(s, y_s) \geq 1$ . Then, we can chose  $s = t - 1$  and we get that for any  $t > 1$

$$m^i(t, x) \geq C_1 m^i(t - 1, y_{t-1}) \geq C_1,$$

which proves (2.30).

The estimates in (2.30) ensure that the pair  $(\mu, p_1)$  is smooth in the sense of Definition 2.4.1. In particular, we know by Proposition 2.4.2 that, for  $\mu$ -a.e.  $m \in \mathcal{P}(\mathbb{T}^d)$ ,

$$-\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} H(y, q_1(y, m)) m(dy) - \mathcal{F}(m) = \lambda,$$

where

$$q_1(y, m) := D_a H^* \left( y, \frac{dp_1}{dm \otimes \mu}(y, m) \right).$$

By the convergence of  $\nu^i$  to  $\nu$ , there exists (up to a subsequence again)  $t_i \in [1, T_i]$  such that  $(m^i(t_i), \alpha^i(t_i))$  converges to an element  $(m, \alpha) \in \mathcal{P}(\mathbb{T}^d) \times E$  which belongs to the support of  $\mu$ . Then by (2.29),  $\alpha = \frac{dp_1}{dm \otimes \mu}$ . Thus  $Du^i(t_i) = D_a H^*(y, \alpha^i(t_i))$  converges uniformly to  $q_1(\cdot, m)$ . This shows that

$$\begin{aligned} & \lim_{i \rightarrow +\infty} c(u^i, m^i) \\ &= \lim_{i \rightarrow +\infty} - \int_{\mathbb{T}^d} Du^i(t_i, y) \cdot Dm^i(t_i, y) dy + \int_{\mathbb{T}^d} H(y, Du^i(t_i, y)) m^T(t_i, dy) - \mathcal{F}(m^i(t_i)) \\ &= - \int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} H(y, q_1(y, m)) m(dy) - \mathcal{F}(m) = \lambda, \end{aligned}$$

which is in contradiction with (2.26). □

The next step towards Theorem 2.5.7 is to prove that the map  $(s, m) \rightarrow \mathcal{U}^T(s, m) + \lambda(T - s)$  has a limit. In the next proposition we prove that  $(s, m) \rightarrow \mathcal{U}^T(s, m) + \lambda(T - s)$  is bounded and equicontinuous on  $[0, T] \times \mathcal{P}(\mathbb{T}^d)$  and so that there exists a subsequence  $(\mathcal{U}^{T_n} + \lambda(T_n - \cdot))$  which, locally in time, converges uniformly to a continuous function  $\xi$ .

**Proposition 2.5.4.** *The maps  $(s, m) \rightarrow \mathcal{U}^T(s, m) + \lambda(T - s)$  are uniformly bounded and uniformly continuous.*

*Proof.* We first prove that  $(s, m) \rightarrow \mathcal{U}^T(s, m) + \lambda(T - s)$  is bounded, uniformly in  $T$ . Let  $\chi$  be a corrector function. As  $\chi$  is a continuous function on the compact set  $\mathcal{P}(\mathbb{T}^d)$ , there exists a constant  $C > 0$  such that  $0 \leq \chi(m) + C$  for any  $m \in \mathcal{P}(\mathbb{T}^d)$ . If  $(m(t), w(t))$  is an admissible trajectory for the minimization problem of  $\mathcal{U}^T(s, m)$ , then

$$\begin{aligned} & \int_s^T \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) dm(t) + \mathcal{F}(m(t)) dt + \lambda(T - s) \\ & \leq \int_s^T \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) dm(t) + \mathcal{F}(m(t)) dt + \chi(m(T)) + \lambda(T - s) + C. \end{aligned}$$

Taking the infimum over all the possible  $(m, \alpha)$ , the definition of  $\mathcal{U}^T(s, m)$  and the dynamic programming principle verified by  $\chi$  lead to

$$\mathcal{U}^T(s, m) + \lambda(T - s) \leq \chi(m) + C.$$

As  $\chi$  is bounded, we get an upper bound independent of  $T$ ,  $m$  and  $s$ . The lower bound is analogous.

We turn to the equicontinuity. For what concern the continuity in the  $m$  variable, one can adapt the proof of [68, Theorem 1.5] with minor adjustments, to show that, if  $T - s \geq \varepsilon > 0$  for given  $\varepsilon > 0$ , then there exists a constant  $K$  independent of  $T$  and  $s$  such that  $\mathcal{U}^T(s, \cdot)$  is  $K$ -Lipschitz continuous.

We now need to estimate  $|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_1, m_0)|$ . We suppose  $t_2 > t_1$  and we fix  $(\bar{m}(s), \bar{\alpha}(s))$  an optimal trajectory for  $\mathcal{U}^T(t_1, m_0)$ , then

$$|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_1, m_0)| \leq |\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_2, \bar{m}(t_2))| + |\mathcal{U}^T(t_2, \bar{m}(t_2)) - \mathcal{U}^T(t_1, m_0)|. \quad (2.31)$$

We can estimate the first term on the right hand-side using at first the uniform Lipschitz continuity we discussed before and then the estimates on the solution of the MFG system in Lemma 2.5.1. So,

$$|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_2, \bar{m}(t_2))| \leq K \mathbf{d}(m_0, \bar{m}(t_2)) \leq KC|t_1 - t_2|^{\frac{1}{2}}. \quad (2.32)$$

To estimate the second term in the right hand-side of (2.31), we just need to use that

$$\mathcal{U}^T(t_1, m_1) = \inf_{(m, \alpha)} \left\{ \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) dm(t) + \mathcal{F}(m(t)) dt + \mathcal{U}^T(t_2, m(t_2)) \right\}.$$

As  $(\bar{m}, \bar{\alpha})$  is optimal for  $\mathcal{U}^T(t_1, m_1)$ , we get

$$\mathcal{U}^T(t_1, m_1) - \mathcal{U}^T(t_2, \bar{m}(t_2)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t, x)) d\bar{m}(t) + \mathcal{F}(\bar{m}(t)) dt.$$

Note that,  $\bar{\alpha}(t, x) = D_p H(x, D\bar{u}(x, t))$  where  $(\bar{m}, \bar{u})$  solves the MFG system (2.25) and, according to Lemma 2.5.1, we have uniform estimates on  $\bar{u}$ . Therefore,

$$|\mathcal{U}^T(t_2, \bar{m}(t_2)) - \mathcal{U}^T(t_1, m_0)| \leq \int_{t_1}^{t_2} \left| \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t, x)) d\bar{m}(t) + \mathcal{F}(\bar{m}(t)) \right| dt \leq C(t_2 - t_1).$$

Putting together the last inequality with (2.32) we have that, for a possibly different constant  $C > 0$ , independent of  $T$ ,  $m_0$  and  $m_1$ ,

$$|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_1, m_0)| \leq C(|t_1 - t_2|^{\frac{1}{2}} + |t_1 - t_2|),$$

which, in turn, implies the uniform continuity in time.  $\square$



From now on we fix a continuous map  $\xi : [0, +\infty) \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ , limit of a subsequence (denoted in the same way) of the sequence  $(\mathcal{U}^T + \lambda(T - \cdot))$  as  $T \rightarrow +\infty$ .

**Proposition 2.5.5.** *The map  $\xi_0(\cdot) := \xi(0, \cdot)$  satisfies*

$$\xi_0(m) \leq \inf_{(m, \alpha)} \left\{ \int_0^t \left( \int_{\mathbb{T}^d} H^*(y, \alpha(s, y)) m(s, y) dy + \mathcal{F}(m(s)) \right) ds + \xi_0(m(t)) \right\} + \lambda t.$$

*Proof.* We first claim that  $\xi$  is a viscosity solution to

$$-\partial_t \xi \geq 0 \quad \text{in } [0, +\infty) \times \mathcal{P}(\mathbb{T}^d). \quad (2.33)$$

Let  $\Phi = \Phi(t, m)$  be a smooth test function such that  $\xi \geq \Phi$  with an equality only at  $(t_0, m_0)$ . Then there exists a subsequence  $(t_n, \bar{m}_n)$  converging to  $(t_0, m_0)$  and such that  $\mathcal{U}^{T_n} + \lambda(T_n - s) - \Phi$  has a minimum at  $(t_n, \bar{m}_n)$ . Let  $(m_n, \alpha_n)$  be a minimizer for  $\mathcal{U}^{T_n}(t_n, \bar{m}_n)$ . We consider  $u_n \in C^{1,2}([t_n, T_n] \times \mathbb{T}^d)$  such that  $(u_n, m_n)$  is a solution to the MFG system (2.25) and  $\alpha_n = D_p H(x, Du_n)$ . Then, by Lemma 2.6.1, we have

$$-\partial_t \Phi(t_n, \bar{m}_n) - \lambda + \int_{\mathbb{T}^d} (H(y, Du_n(t_n, y)) - \Delta u_n(t_n, y)) \bar{m}_n(dy) - \mathcal{F}(\bar{m}_n) \geq 0.$$

By Lemma 2.5.2 and Proposition 2.5.3, we have, given  $\epsilon > 0$ ,

$$\int_{\mathbb{T}^d} (H(y, Du_n(t_n, y)) - \Delta u_n(t_n, y)) \bar{m}_n(dy) - \mathcal{F}(\bar{m}_n) \leq \lambda + \epsilon,$$

for  $n$  large enough. So

$$-\partial_t \Phi(t_n, \bar{m}_n) \geq -\epsilon.$$

We obtain therefore, after letting  $n \rightarrow +\infty$  and then  $\epsilon \rightarrow 0$ ,

$$-\partial_t \Phi(t_0, m_0) \geq 0.$$

This shows that  $\xi$  satisfies (2.33) holds in the viscosity solution sense.

We now prove that (2.33) implies that  $\xi$  is nonincreasing in time. Fix  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and assume on the contrary that there exists  $0 \leq t_1 < t_2$  such that  $\xi(t_1, m_0) < \xi(t_2, m_0)$ . Let  $\Psi = \Phi(m)$  be a smooth test function such that  $\Psi > 0$  on  $\mathcal{P}(\mathbb{T}^d) \setminus \{m_0\}$  with  $\Phi(m_0) = 0$ . Then, we can find  $\eta > 0$  small such that, if  $m_1$  and  $m_2$  are such that  $\xi(t_1, \cdot) + \eta^{-1}\Psi$  has a minimum at  $m_1$  and  $\xi(t_2, \cdot) + \eta^{-1}\Psi$  has a minimum at  $m_2$ , then  $\xi(t_1, m_1) < \xi(t_2, m_2)$ . Note that this implies that

$$\begin{aligned} \min_{m \in \mathcal{P}(\mathbb{T}^d)} \xi(t_1, m) + \eta^{-1}\Psi(m) &= \xi(t_1, m_1) + \eta^{-1}\Psi(m_1) \\ &< \xi(t_2, m_2) + \eta^{-1}\Psi(m_2) = \min_{m \in \mathcal{P}(\mathbb{T}^d)} \xi(t_2, m) + \eta^{-1}\Psi(m). \end{aligned}$$

Recalling that  $\xi$  is bounded, this implies that we can find  $\epsilon > 0$  small such that the map  $(t, m) \rightarrow \xi(t, m) + \eta^{-1}\Psi(m) + \epsilon t$  has an interior minimum on  $[t_1, +\infty) \times \mathcal{P}(\mathbb{T}^d)$  at some point  $(t_3, m_3) \in (t_1, +\infty) \times \mathcal{P}(\mathbb{T}^d)$ . This contradicts (2.33).

Now that we have proved the monotonicity in time of  $\xi$  we can finally show the statement of the proposition. We have that  $\mathcal{U}^T$  verifies the following dynamic programming principle

$$\mathcal{U}^T(0, m_0) = \inf_{(m, \alpha)} \left\{ \int_0^t \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) dm(t) + \mathcal{F}(m(t)) dt + \mathcal{U}^T(t, m(t)) \right\}. \quad (2.34)$$

Then, adding on both sides  $\lambda T$  and passing to the limit  $T \rightarrow +\infty$ , one easily checks that  $\xi$  satisfies

$$\xi(0, m) = \inf_{(m, \alpha)} \left\{ \int_0^t \left( \int_{\mathbb{T}^d} H^*(y, \alpha(s, y)) m(s, y) dy + \mathcal{F}(m(s)) \right) ds + \xi(t, m(t)) \right\} + \lambda t.$$

Using the fact that  $\xi$  is nonincreasing in time we get the desired result.  $\square$

Before we can prove that  $\xi_0$  is a corrector function we need to state some standard properties of  $\tau_h : C^0(\mathcal{P}(\mathbb{T}^d)) \rightarrow C^0(\mathcal{P}(\mathbb{T}^d))$  which is defined as follows. For  $h > 0$  and  $\Phi \in C^0(\mathbb{T}^d)$  we set

$$\tau_h \Phi(m_0) = \inf_{(m, \alpha)} \left\{ \int_0^h \left( \int_{\mathbb{T}^d} H^*(y, \alpha(s, y)) m(s, y) dy + \mathcal{F}(m(s)) \right) ds + \Phi(m(h)) \right\} + \lambda h,$$

where  $(m, \alpha)$  solves in the sense of distribution

$$\begin{cases} -\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0 & \text{in } [0, h] \times \mathbb{T}^d \\ m(0) = m_0. \end{cases}$$

**Lemma 2.5.6.** *The function  $\tau_h$  verifies the following properties*

1. For any  $h_1, h_2 > 0$ ,  $\tau_{h_1} \circ \tau_{h_2} = \tau_{h_1+h_2}$
2. For any  $h > 0$ ,  $\tau_h$  is not expansive, i.e. for any  $\Phi, \Psi \in C^0(\mathcal{P}(\mathbb{T}^d))$

$$\|\tau_h \Phi - \tau_h \Psi\|_\infty \leq \|\Phi - \Psi\|_\infty.$$

3. For any  $h > 0$ ,  $\tau_h$  is order preserving, i.e. for any  $\Phi, \Psi \in C^0(\mathcal{P}(\mathbb{T}^d))$  such that  $\Phi \leq \Psi$ ,

$$\tau_h \Phi \leq \tau_h \Psi.$$

4. Let  $\Phi \in C^0(\mathcal{P}(\mathbb{T}^d))$  be such that, for any  $h > 0$ ,  $\Phi \leq \tau_h \Phi$ . Then, for any  $0 < h_1 < h_2$ ,

$$\Phi \leq \tau_{h_1} \Phi \leq \tau_{h_2} \Phi.$$

*Proof.* The proof is standard, see for instance, in a closely related context, [42].  $\square$

**Theorem 2.5.7.**  $\xi_0$  is a corrector and  $\mathcal{U}^T(0) + \lambda T$  converges uniformly to  $\xi_0$  on  $\mathcal{P}(\mathbb{T}^d)$ .

*Proof.* The proof follows closely the one of [42, Theorem 6.3.1]. We define  $\tilde{\mathcal{U}}^T(t, m) = \mathcal{U}^T(t, m) + \lambda(T - t)$ . Let  $T_n \rightarrow +\infty$  be a sequence such that  $\tilde{\mathcal{U}}^{T_n}$  converges locally uniformly to  $\xi$  on  $[0, +\infty) \times \mathcal{P}(\mathbb{T}^d)$ . We can suppose that, if we define  $s_n = T_{n+1} - T_n$ , then  $s_n \rightarrow +\infty$ . Note that  $\tilde{\mathcal{U}}^{T_{n+1}}(s_n, m) = \tilde{\mathcal{U}}^{T_n}(0, m)$ . Then, using (2.34), we get

$$\tilde{\mathcal{U}}^{T_{n+1}}(0, m) = \tau_{s_n} \tilde{\mathcal{U}}^{T_{n+1}}(s_n, m) = \tau_{s_n} \tilde{\mathcal{U}}^{T_n}(0, m).$$

We also know from Lemma 2.5.6 that  $\tau_h$  is a contraction and that it verifies the semigroup property. Therefore,

$$\begin{aligned} \|\tau_{s_n} \xi_0 - \xi_0\|_\infty &\leq \|\tau_{s_n} \xi_0 - \tau_{s_n} \tilde{\mathcal{U}}^{T_n}(0)\|_\infty + \|\tau_{s_n} \tilde{\mathcal{U}}^{T_n}(0) - \xi_0\|_\infty \\ &\leq \|\xi_0 - \tilde{\mathcal{U}}^{T_n}(0)\|_\infty + \|\tilde{\mathcal{U}}^{T_{n+1}}(0) - \xi_0\|_\infty \rightarrow 0, \end{aligned}$$

Moreover, Proposition 2.5.5 and Lemma 2.5.6 prove that  $\tau_s \xi_0$  is monotone in  $s$ . Then, for any  $s > 0$ , we have that, for a sufficiently large  $n \in \mathbb{N}$ ,

$$\xi_0 \leq \tau_s \xi_0 \leq \tau_{s_n} \xi_0 \rightarrow \xi_0,$$

which proves that  $\tau_t \xi_0$  is constant in  $t$  and so  $\xi_0$  is corrector function. It remains to check that the whole sequence  $\tilde{\mathcal{U}}^T(0)$  converges to  $\xi_0$ . Let  $T > T_n$ , then

$$\|\tilde{\mathcal{U}}^T(0) - \xi_0\|_\infty = \|\tau_{T-T_n} \tilde{\mathcal{U}}^{T_n}(0) - \tau_{T-T_n} \xi_0\|_\infty \leq \|\tilde{\mathcal{U}}^{T_n}(0) - \xi_0\|_\infty \rightarrow 0$$

and the result follows.  $\square$

## 2.5.2 Convergence of optimal trajectories

Now that we have proved the convergence of  $\mathcal{U}^T(0, \cdot) + \lambda T$  to a corrector function  $\chi$ , we can properly define the limit trajectories for time dependent MFG and the set where these trajectories lay. We will show that this set is a subset of the projected Mather set  $\mathcal{M}$  as was suggested in [68]. We recall the definition of  $\mathcal{M}$ .

**Definition 2.5.8.** *We say that  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  belongs to the projected Mather set  $\mathcal{M} \subset \mathcal{P}(\mathbb{T}^d)$  if there exists a calibrated curve  $(m(t), \alpha(t))$  such that  $m(0) = m_0$ .*

**Remark 2.5.1.** *Note that the notion of projected Mather set that we use here is consistent with the one that was already introduced in [68]. On the other hand, Definition 2.5.8 is not the transposition of the definition of projected Mather set that is generally used in standard Weak KAM theory. In this latter case the projected Mather set is the union of the projection on the torus of the supports of Mather measures. What we call here projected Mather set would be rather the projected Aubry set or the projected Mané set (we refer to [43] and [42] for these definitions). We decided to use Definition 2.5.8 mostly to be consistent with the terminology in [68]. Moreover, it is worthwhile to mention that in the standard theory the Mather set and the Aubry set are deeply connected while in this framework such a relation is no longer clear. In particular, in standard Weak KAM theory the Mather set is contained in the Aubry set (where the latter is defined as the intersection of graphs of calibrated curves). One can check this inclusion defining a calibrated curve, starting from any point of the Mather set, through the Lagrangian flow. In the MFG setting, the lack of uniqueness of solutions and the forward/backward structure of the system prevent from defining any sensible notion of flow. Moreover, an other important difference, that highlights how the connection between Mather set and Aubry set is not clear in the MFG framework, is that, on the one hand, we know that calibrated curves lay on smooth probability measures but, on the other, we know nothing about the regularity of Mather measures' support points (reason why we introduced the notion of "smooth" Mather measure).*

We recall also that a couple  $(\bar{m}, \bar{\alpha})$ , which satisfies  $-\partial_t \bar{m}(t) + \Delta \bar{m}(t) + \operatorname{div}(\bar{\alpha}(t) \bar{m}(t)) = 0$  for any  $t \in \mathbb{R}$ , is a calibrated curve, if there exists a corrector function  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  such that, for any  $t_1 < t_2 \in \mathbb{R}$ ,

$$\chi(\bar{m}(t_1)) = \lambda(t_2 - t_1) + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(s)) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + \chi(\bar{m}(t_2)).$$

We fix  $(m^T, \alpha^T)$  a minimizer for  $\mathcal{U}^T(-T, m_0)$ . As usual  $\alpha^T = D_p H(x, D\bar{u}^T)$  where  $(\bar{u}^T, m^T)$  solves (2.25) on  $[-T, T] \times \mathbb{T}^d$ . We define  $u^T(t, x) = \bar{u}^T(t, x) - u^T(0, \bar{x})$ , for a fixed  $\bar{x} \in \mathbb{T}^d$ . We know from Lemma 2.5.1 that  $D^2 \bar{u}^T$  and  $\partial_t \bar{u}^T$  are uniformly bounded. This means that  $u^T$  and

$Du^T$  are uniformly bounded and uniformly continuous on any compact set of  $\mathbb{R} \times \mathbb{T}^d$ . Therefore, we have that up to subsequence  $u^T$  converges to a function  $u \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$ . The convergence, up to subsequence, of  $m^T$  to a function  $m \in C^0(\mathbb{R}, \mathcal{P}(\mathbb{T}^d))$  is ensured again by Lemma 2.5.1 and the uniform  $C^{\frac{1}{2}}([0, T], \mathcal{P}(\mathbb{T}^d))$  bounds on  $m^T$  therein. It is standard that the couple  $(u, m)$  solves in classical sense

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

As by product we have that  $(m^T, \alpha^T)$  uniformly converges on compact sets to the couple  $(m, D_p H(x, Du))$ . Moreover, if we define  $\alpha = D_p H(x, Du)$ , then  $(m, \alpha)$  solves  $-\partial_t m + \Delta m + \operatorname{div}(m \alpha) = 0$ . We can now prove that  $(m, \alpha)$  is a calibrated curve and therefore that  $\mathcal{M}$  contains the uniform limits of optimal trajectories.

**Theorem 2.5.9.** *Let  $(m^T, \alpha^T)$  be an optimal trajectory for  $\mathcal{U}^T(-T, m_0)$ . Then,  $(m^T, \alpha^T)$  converges, up to subsequence, to a calibrated curve  $(m, \alpha)$ . Consequently,  $m(t) \in \mathcal{M}$  for any  $t \in \mathbb{R}$ .*

*Proof.* As we have already discussed the convergence of  $(m^T, \alpha^T)$  to  $(m, \alpha)$  we just need to check that  $(m, \alpha)$  is a calibrated curve. We fix  $t_1 < t_2 \in \mathbb{R}$ , then, by dynamic programming principle,

$$\mathcal{U}^T(t_1, m_0) + \lambda(T - t_1) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \alpha^T(s)) dm^T(s) + \mathcal{F}(m^T(s)) ds + \mathcal{U}^T(t_2, m^T(t_2)) + \lambda(T - t_1).$$

We recall that  $\mathcal{U}^T(t, m_0) = \mathcal{U}^{T-t}(0, m_0)$ . Given the continuity of  $\mathcal{U}^T(0, \cdot)$ , the uniform convergence of  $(m^T, \alpha^T)$  on compact subsets and the uniform convergence of  $\mathcal{U}^T(0, \cdot) + \lambda T$  to  $\chi(\cdot)$ , we can pass to the limit in  $T$  and we get that, for any interval  $[t_1, t_2]$ , the couple  $(m, \alpha)$  verifies

$$\chi(m(t_1)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t_2)) + \lambda(t_2 - t_1).$$

So  $(m, \alpha)$  is a calibrated curve. □

## 2.6 Appendix

### 2.6.1 Proof of Proposition 2.3.4

*Proof of Proposition 2.3.4.* To simplify the notation, we argue as if the convergence of  $v^N$  to  $V$  in Lemma 2.3.3 holds for the full sequence (i.e.,  $N_k = N$ ). We denote with  $\mathcal{P}_2(\mathbb{R}^d)$  the set of Borel probability measures  $m$  on  $\mathbb{R}^d$  such that

$$\int_{\mathbb{R}^d} |x|^2 m(dx) < +\infty$$

We denote by  $\pi : \mathbb{R}^d \rightarrow \mathbb{T}^d$  the usual projection (and, by abuse of notation, we set  $\pi(\mathbf{x}) = (\pi(x^1), \dots, \pi(x^N))$  for any  $\mathbf{x} = (x^1, \dots, x^N) \in (\mathbb{R}^d)^N$ ). Let  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$  with compact support and  $(\xi^i)_{i \in \mathbb{N}}$  be an i.i.d. sequence of random variables on  $\mathbb{R}^d$  with law  $m_0$ . Fix also  $(B^i)$  independent Brownian motions on  $\mathbb{R}^d$  which are also independent of  $(\xi^i)$  and set  $\xi^N = (\xi^1, \dots, \xi^N)$ . When we look at  $v^N$  as a  $\mathbb{Z}^d$ -periodic map on  $(\mathbb{R}^d)^N$ , we have by classical representation formula that

$$\mathbb{E} [v^N(\xi^N)] = \inf_{\alpha} \mathbb{E} \left[ N^{-1} \sum_{i=1}^N \int_0^T H^*(X_t^i, \alpha_t^i) dt + \int_0^T \mathcal{F}(\pi \# m_{\mathbf{X}_t}^N) dt + v^N(\mathbf{X}_T) \right] + \lambda^N T, \quad (2.35)$$

where the infimum is taken over progressively measurable controls  $\alpha = (\alpha^1, \dots, \alpha^N)$  (adapted to the filtration generated by the  $(B^i)$  and the  $(\xi^i)$ ) and where  $\mathbf{X} = (X^1, \dots, X^N)$  solves

$$dX_t^i = \alpha_t^i dt + \sqrt{2} dB_t^i, \quad t \in [0, T], \quad X_0^i = \xi^i.$$

The optimal feedback in (2.35) is well-known: it is given by  $\alpha^{*i,N}(t, \mathbf{x}) := D_p H(x_i, D_{x_i} v^N(\mathbf{x}))$ . We denote by  $\alpha^{*N} = (\alpha^{*1,N}, \dots, \alpha^{*N,N})$  and  $\mathbf{X}^{*N} = (X^{*1,N}, \dots, X^{*N,N})$  the corresponding optimal solution:

$$dX_t^{*i,N} = \alpha_t^{*i,N} dt + \sqrt{2} dB_t^i, \quad t \in [0, T], \quad X_0^{*i,N} = \xi^i.$$

By Lemma 2.3.3, we have

$$\lim_{N \rightarrow +\infty} \epsilon_N = 0 \quad \text{where} \quad \epsilon_N := \sup_{\mathbf{x} \in (\mathbb{R}^d)^N} \left| V(\pi \# m_{\mathbf{x}}^N) - v^N(\mathbf{x}) \right|, \quad (2.36)$$

because  $v^N(\mathbf{x}) = v^N((\pi(x^1), \dots, \pi(x^N)))$  while  $\pi \# m_{\mathbf{x}}^N = m_{\pi(\mathbf{x})}^N$ .

Note that (2.36) implies, on the one hand, that

$$\left| \mathbb{E} \left[ N^{-1} \sum_{i=1}^N \int_0^T H^*(X_t^{*i,N}, \alpha_t^{*i,N}) dt + \int_0^T \mathcal{F}(\pi \# m_{\mathbf{X}_t^{*N}}^N) dt + V(\pi \# m_{\mathbf{X}_T^{*N}}^N) + \lambda^N T \right] - \mathbb{E} \left[ v^N(\xi^N) \right] \right| \leq \epsilon_N, \quad (2.37)$$

and, on the other hand, that  $\alpha^{*N}$  is  $2\epsilon_N$  optimal for the problem if one replaces  $v^N$  by  $V$  in the optimal control problem:

$$\left| \mathbb{E} \left[ N^{-1} \sum_{i=1}^N \int_0^T H^*(X_t^{*i,N}, \alpha_t^{*i,N}) dt + \int_0^T \mathcal{F}(\pi \# m_{\mathbf{X}_t^{*N}}^N) dt + V(\pi \# m_{\mathbf{X}_T^{*N}}^N) \right] - \inf_{\alpha} \mathbb{E} \left[ N^{-1} \sum_{i=1}^N \int_0^T H^*(X_t^i, \alpha_t^i) dt + \int_0^T \mathcal{F}(\pi \# m_{\mathbf{X}_t}^N) dt + V(\pi \# m_{\mathbf{X}_T}^N) \right] \right| \leq 2\epsilon_N. \quad (2.38)$$

We aim at letting  $N \rightarrow +\infty$  in the above inequalities. By the law of large numbers we know that  $(m_{\xi^N}^N)$  converges a.s. and in expectation in  $\mathcal{P}_2$  to  $m_0$ . Therefore, by the Lipschitz continuity of  $V$ , we obtain

$$\begin{aligned} & \limsup \left| \mathbb{E} \left[ v^N(\xi^N) \right] - V(\pi \# m_0) \right| \\ & \leq \limsup \mathbb{E} \left[ \left| v^N(\xi^N) - V(\pi \# m_{\xi^N}^N) \right| \right] + \limsup \mathbb{E} \left[ \left| V(\pi \# m_{\xi^N}^N) - V(\pi \# m_0) \right| \right] \\ & \leq \limsup \epsilon_N + C \mathbb{E} \left[ \mathbf{d}_2(m_{\xi^N}^N, m_0) \right] = 0. \end{aligned} \quad (2.39)$$

In order to pass to the limit in (2.38), we use several results of [59]. The first one (Corollary 2.13) states that  $m_{\mathbf{X}^{*N}}^N$  is precompact in  $\mathcal{P}_2([0, T], \mathcal{P}_2(\mathbb{R}^d))$  and that every weak limit has a support in the set of relaxed minimizers of the McKean Vlasov optimal control problem in weak Markovian formulation (expressed here with—almost—the notation of [59], see Proposition 2.5):

$$\kappa := \inf_{\mathbb{P}} \mathbb{E}^{\mathbb{P}} \left[ \int_0^T H^*(X_t, \hat{\alpha}(t, X_t)) + \mathcal{F}(\pi \# (\mathbb{P} \circ X_t^{-1})) dt + V(\pi \# (\mathbb{P} \circ X_T^{-1})) \right] + \lambda^* T, \quad (2.40)$$

where the infimum is taken over the family of probability spaces  $(\Omega, (\mathcal{F}_t), \mathbb{P})$  supporting a  $d$ -dimensional process  $X$  and a  $d$ -dimensional Brownian motion  $B$ , such that  $\mathbb{P} \circ X_0^{-1} = m_0$ ,  $\hat{\alpha} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is Borel measurable, and the following holds:

$$dX_t = \hat{\alpha}(t, X_t)dt + \sqrt{2}dB_t, \quad \text{with} \quad \mathbb{E}^{\mathbb{P}} \left[ \int_0^T |X_t|^2 + |\hat{\alpha}(t, X_t)|^2 dt \right] < +\infty. \quad (2.41)$$

The above problem can be reformulated in PDE terms as

$$\kappa = \inf_{(m, \hat{\alpha})} \int_0^T \int_{\mathbb{R}^d} H^*(x, \hat{\alpha}(t, x))m(t, dx)dt + \int_0^T \mathcal{F}(\pi\#m(t))dt + V(\pi\#m(T)) + \lambda^*T, \quad (2.42)$$

where the infimum is taken over the pairs  $(m, \hat{\alpha})$  with  $m \in C^0([0, T], \mathcal{P}_2(\mathbb{R}^d))$ ,  $\hat{\alpha} \in L^2([0, T], L^2_{m(t)}(\mathbb{R}^d))$  and

$$\partial_t m - \Delta m - \text{div}(m\hat{\alpha}) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad m(0) = m_0. \quad (2.43)$$

Indeed, if  $\mathbb{P}$  is admissible in (2.40), we just need to set  $m(t) = \mathbb{P} \circ X_t^{-1}$  and then  $(m, \hat{\alpha})$  is admissible in (2.42). Conversely, if  $(m, \hat{\alpha})$  is admissible in (2.42), then there exists a weak solution to the SDE (2.41): this precisely means that there exists a stochastic basis which is admissible for (2.40).

Next, we note that the proof of Theorem 2.11 in [59] (and more precisely inequality (6.1)) shows that

$$\lim_N \mathbb{E} \left[ N^{-1} \sum_{i=1}^N \int_0^T H^*(X_t^{*i, N}, \alpha_t^{*i, N})dt + \int_0^T \mathcal{F}(\pi\#m_{\mathbf{X}_t^{*N}}^N) dt + V(\pi\#m_{\mathbf{X}_T^{*N}}^N) \right] + \lambda^*T = \kappa.$$

Putting together (2.37), (2.39), (2.42) and the above equality shows that

$$\begin{aligned} V(\pi\#m_0) &= \lim_N \mathbb{E} \left[ v^N(\xi^N) \right] \\ &= \lim_N \mathbb{E} \left[ N^{-1} \sum_{i=1}^N \int_0^T H^*(X_t^{*i, N}, \alpha_t^{*i, N})dt + \int_0^T \mathcal{F}(\pi\#m_{\mathbf{X}_t^{*N}}^N) dt + V(\pi\#m_{\mathbf{X}_T^{*N}}^N) \right] + \lambda^*T \\ &= \kappa = \inf_{(m, \hat{\alpha})} \int_0^T \int_{\mathbb{R}^d} H^*(x, \hat{\alpha}(t, x))m(t, dx)dt + \int_0^T \mathcal{F}(\pi\#m(t))dt + V(\pi\#m(T)) + \lambda^*T, \end{aligned} \quad (2.44)$$

where the infimum is computed as above.

It remains to explain why we can work in  $\mathbb{T}^d$  instead of  $\mathbb{R}^d$ . For this, let us define, for any  $m_0 \in \mathcal{P}(\mathbb{R}^d)$  and any  $\tilde{m}_0 \in \mathcal{P}_2(\mathbb{T}^d)$ ,  $J_{\mathbb{R}^d}(m_0)$  and  $J_{\mathbb{T}^d}(\tilde{m}_0)$  by

$$J_{\mathbb{R}^d}(m_0) := \inf_{(m, \hat{\alpha})} \int_0^T \int_{\mathbb{R}^d} H^*(x, \hat{\alpha}(t, x))m(t, dx)dt + \int_0^T \mathcal{F}(\pi\#m(t))dt + V(\pi\#m(T)),$$

where the infimum is taken over the pairs  $(m, \hat{\alpha})$  as above, and

$$J_{\mathbb{T}^d}(\tilde{m}_0) := \inf_{(\tilde{m}, \tilde{w})} \int_0^T \int_{\mathbb{T}^d} H^*(x, -\frac{d\tilde{w}}{d\tilde{m}}(t, x))\tilde{m}(t, dx)dt + \int_0^T \mathcal{F}(\tilde{m}(t))dt + V(\tilde{m}(T)),$$

where the infimum is computed (as usual) over the pairs  $(\tilde{m}, w)$  such that  $\tilde{m} \in C^0([0, T], \mathcal{P}(\mathbb{T}^d))$ ,  $\tilde{w}$  is a vector measure on  $[0, T] \times \mathbb{T}^d$  with values in  $\mathbb{R}^d$  with first marginal  $dt$  and which is absolutely continuous with respect to  $\tilde{m}$  with

$$\int_0^T \int_{\mathbb{T}^d} \left| \frac{d\tilde{w}}{d\tilde{m}}(t, x) \right|^2 \tilde{m}(t, dx)dt < +\infty$$

and the continuity equation

$$\partial_t \tilde{m} - \Delta \tilde{m} + \operatorname{div}(\tilde{w}) = 0 \text{ in } (0, T) \times \mathbb{T}^d, \quad \tilde{m}(0) = \tilde{m}_0$$

holds. From the Lipschitz continuity of  $V$  and the convexity of  $H^*$  one can easily prove that  $J_{\mathbb{R}^d}$  is continuous on  $\mathcal{P}_2(\mathbb{R}^d)$ . Our aim is to show that  $J_{\mathbb{R}^d}(m_0) = J_{\mathbb{T}^d}(\pi\sharp m_0)$  and  $J_{\mathbb{T}^d}(\tilde{m}_0) = J_{\mathbb{R}^d}(\tilde{m}_0 \mathbf{1}_{Q_1})$  for any  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$  and any  $\tilde{m}_0 \in \mathcal{P}(\mathbb{T}^d)$ , where  $Q_1 = [-1/2, 1/2)^d$ .

Let  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$  and  $(m, \hat{\alpha})$  be admissible for  $J_{\mathbb{R}^d}(m_0)$ . We see  $w := -\hat{\alpha}m$  as a vector measure on  $[0, T] \times \mathbb{R}^d$  with values in  $\mathbb{R}^d$  and finite mass (since  $\hat{\alpha} \in L^2([0, T], L^2_{m(t)}(\mathbb{R}^d))$ ). Let us set  $\tilde{m}(t) = \pi\sharp m(t)$  and  $\tilde{w}(t) := \pi\sharp w(t)$ . Then, as  $(m, \alpha)$  satisfies (2.43),  $(\tilde{m}, \tilde{w})$  solves the continuity equation

$$\partial_t \tilde{m} - \Delta \tilde{m} + \operatorname{div}(\tilde{w}) = 0 \text{ in } (0, T) \times \mathbb{T}^d, \quad \tilde{m}(0) = \pi\sharp m_0.$$

Indeed, if  $\varphi$  is a smooth function with compact support on  $[0, T] \times \mathbb{T}^d$  we have

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}^d} (\partial_t \varphi(t, x) + \Delta \varphi(t, x)) \tilde{m}(t, dx) + D\varphi(t, x) \tilde{w}(t, dx) + \int_{\mathbb{T}^d} \phi(0) m_0 \\ &= \int_0^T \int_{\mathbb{R}^d} (\partial_t \varphi(t, \pi(y)) + \Delta \varphi(t, \pi(y))) m(t, dy) + D\varphi(t, \pi(y)) w(t, dy) + \int_{\mathbb{T}^d} \phi(0) m_0 = 0, \end{aligned}$$

where the last equality holds because the map  $(t, y) \mapsto \varphi(t, \pi(y))$  is smooth and bounded with bounded derivatives on  $[0, T] \times \mathbb{R}^d$  and  $(m, w)$  verifies  $-\partial_t m + \Delta m - \operatorname{div} w = 0$ . As

$$\begin{aligned} \int_0^T \int_{\mathbb{T}^d} H^*(x, \frac{d\tilde{w}}{d\tilde{m}}(t, x)) \tilde{m}(t, dx) dt &= \int_0^T \int_{\mathbb{R}^d} H^*(\pi(x), \frac{dw}{dm}(t, x)) m(t, dx) dt \\ &= \int_0^T \int_{\mathbb{R}^d} H^*(x, \hat{\alpha}(t, x)) m(t, dx) dt, \end{aligned}$$

(since  $L$  is  $\mathbb{Z}^d$  periodic in the first variable), we easily derive that  $J_{\mathbb{R}^d}(m_0) \geq J_{\mathbb{T}^d}(\pi\sharp m_0)$ .

To prove the opposite inequality let  $(\tilde{m}, \tilde{w})$  be  $\theta$ -optimal for  $J_{\mathbb{T}^d}(\tilde{m}_0)$ . We define  $(\tilde{m}_\varepsilon, \tilde{w}_\varepsilon) := (\tilde{m} * \xi_\varepsilon, \tilde{w} * \xi_\varepsilon)$  where  $\xi_\varepsilon$  is a standard mollification kernel in the space variable. The couple  $(\tilde{m}_\varepsilon, \tilde{w}_\varepsilon)$  solves the usual continuity equation with initial condition  $\tilde{m}_{0,\varepsilon} = \tilde{m}_0 * \xi_\varepsilon$ . Then

$$\int_0^T \int_{\mathbb{T}^d} H^*(x, -\frac{d\tilde{w}_\varepsilon}{d\tilde{m}_\varepsilon}(t, x)) \tilde{m}_\varepsilon(t, dx) dt + \int_0^T \mathcal{F}(\tilde{m}_\varepsilon(t)) dt + V(\tilde{m}_\varepsilon(T)) \leq J_{\mathbb{T}^d}(\tilde{m}_0) + \theta + O_\varepsilon(1).$$

Set  $\tilde{\alpha}_\varepsilon(t, x) = -(\tilde{w}_\varepsilon / \tilde{m}_\varepsilon)(t, x)$ . Let  $m_\varepsilon$  be the solution to

$$\partial_t m_\varepsilon - \Delta m_\varepsilon - \operatorname{div}(\tilde{\alpha}_\varepsilon(t, \pi(x)) m_\varepsilon) = 0 \text{ in } (0, T) \times \mathbb{R}^d, \quad m_\varepsilon(0, y) = \tilde{m}_{0,\varepsilon}(\pi(y)) \mathbf{1}_{Q_1}(y) \quad y \in \mathbb{R}^d.$$

Then, by periodicity of  $\tilde{\alpha}_\varepsilon$ ,  $\tilde{\mu}_\varepsilon(t) := \pi\sharp m_\varepsilon(t)$  solves

$$\partial_t \tilde{\mu}_\varepsilon - \Delta \tilde{\mu}_\varepsilon - \operatorname{div}(\tilde{\alpha}_\varepsilon(t, x) \tilde{\mu}_\varepsilon) = 0 \text{ in } (0, T) \times \mathbb{T}^d, \quad \tilde{\mu}_\varepsilon(0) = \pi\sharp m_\varepsilon(0) = \tilde{m}_{0,\varepsilon},$$

which has  $\tilde{m}_\varepsilon$  as unique solution since  $\tilde{\alpha}_\varepsilon$  is smooth in space. This shows that  $\pi\sharp m_\varepsilon(t) = \tilde{m}_\varepsilon(t)$ . Therefore, as  $(m_\varepsilon(\cdot, \cdot), \tilde{\alpha}_\varepsilon(\cdot, \pi(\cdot)))$  is an admissible competitor for  $J_{\mathbb{R}^d}(m_\varepsilon(0))$ , we get

$$\begin{aligned} J_{\mathbb{R}^d}(m_\varepsilon(0)) &\leq \int_0^T \int_{\mathbb{R}^d} H^*(x, \tilde{\alpha}_\varepsilon(t, \pi(x))) m_\varepsilon(t, dx) dt + \int_0^T \mathcal{F}(\pi\sharp m_\varepsilon(t)) dt + V(\pi\sharp m_\varepsilon(T)) \\ &= \int_0^T \int_{\mathbb{T}^d} H^*(x, \tilde{\alpha}_\varepsilon(t, x)) \tilde{m}_\varepsilon(t, dx) dt + \int_0^T \mathcal{F}(\tilde{m}_\varepsilon(t)) dt + V(\tilde{m}_\varepsilon(T)) \leq J_{\mathbb{T}^d}(\tilde{m}_0) + \theta + O_\varepsilon(1). \end{aligned}$$

As  $J_{\mathbb{R}^d}$  is continuous, we can pass to the limit as  $\varepsilon$  and then  $\theta$  tend to 0. Then we find that  $J_{\mathbb{R}^d}(\tilde{m}_0 \mathbf{1}_{Q_1}) \leq J_{\mathbb{T}^d}(\tilde{m}_0)$  and, therefore, that  $J_{\mathbb{R}^d}(\tilde{m}_0 \mathbf{1}_{Q_1}) = J_{\mathbb{T}^d}(\tilde{m}_0)$ . In view of (2.44), this completes the proof of the proposition.  $\square$

### 2.6.2 A viscosity solution property

**Lemma 2.6.1.** *Let  $\Phi = \Phi(t, m)$  be a smooth test function such that  $\mathcal{U}^T - \Phi$  has a minimum at a point  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$ . Let  $(u, m)$  be a solution of the MFG system (2.25) starting from  $(t_0, m_0)$  and such that  $(m, D_p H(x, Du(t, x)))$  is optimal for  $\mathcal{U}(t_0, m_0)$ . Then*

$$D_m \Phi(t_0, m_0, x) = Du(t_0, x) \text{ for } m_0\text{-a.e. } x \in \mathbb{T}^d \quad (2.45)$$

and

$$-\partial_t \Phi(t_0, m_0) + \int_{\mathbb{T}^d} (H(x, Du(t_0, x)) - \Delta u(t_0, x)) m_0(dx) - \mathcal{F}(m_0) \geq 0. \quad (2.46)$$

*Proof.* Without loss of generality we assume that  $\Phi(t_0, m_0) = \mathcal{U}^T(t_0, m_0)$ . Let  $m'(t_0) \in \mathcal{P}(\mathbb{T}^d)$  and  $m' = m'(t)$  be the solution to

$$\partial_t m' - \Delta m' - \operatorname{div}(m' D_p H(x, Du(t, x))) = 0.$$

We set  $\mu(t) = m'(t) - m(t)$  and, for  $h \in (0, 1]$  and note that the pair  $(m + h\mu, D_p H(x, Du(t, x)))$  is a solution to  $-\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0$  with initial condition  $m_0^h := (1 - h)m_0 + hm'(t_0)$ . Hence, by the definition of  $\Phi$  and  $\mathcal{U}^T$  we have

$$\begin{aligned} \Phi(t_0, m_0^h) &\leq \mathcal{U}^T(t_0, m_0^h) \\ &\leq \int_{t_0}^T \left( \int_{\mathbb{T}^d} H^*(x, D_p H(x, Du(t, x))) (m + h\mu)(t, dx) + \mathcal{F}((m + h\mu)(t)) \right) dt \\ &\leq \Phi(t_0, m_0) + h \left( \int_{t_0}^T \int_{\mathbb{T}^d} (H^*(x, D_p H(x, Du(t, x))) + F(x, m(t))) \mu(t, dx) dt + o_h(1) \right). \end{aligned}$$

Next we use the equation for  $u$  and then for  $\mu$ :

$$\begin{aligned} &\int_{t_0}^T \int_{\mathbb{T}^d} (H^*(x, D_p H(x, Du(t, x))) + F(x, m(t))) \mu(t, dx) dt \\ &= \int_{t_0}^T \int_{\mathbb{T}^d} (H^*(x, D_p H(x, Du(t, x))) - \partial_t u - \Delta u + H(x, Du(t, x))) \mu(t, dx) dt \\ &= \int_{\mathbb{T}^d} u(t_0, x) \mu(t_0, dx) = \int_{\mathbb{T}^d} u(t_0, x) (m'(t_0) - m_0)(dx). \end{aligned}$$

Plugging this into the estimate of  $\Phi(t_0, m_0^h)$  above, we obtain, dividing by  $h$  and letting  $h \rightarrow 0$ ,

$$\int_{\mathbb{T}^d} \frac{\delta \Phi(t_0, m_0, x)}{\delta m} (m'(t_0) - m_0)(dx) \leq \int_{\mathbb{T}^d} u(t_0, x) (m'(t_0) - m_0)(dx).$$

Recalling the convention (2.8) on the derivative and the arbitrariness of  $m'(t_0)$ , we infer, by choosing Dirac masses for  $m'(t_0)$ , that

$$\frac{\delta \Phi(t_0, m_0, x)}{\delta m} \leq u(t_0, x) - \int_{\mathbb{T}^d} u(t_0, y) m_0(dy) \quad \forall x \in \mathbb{T}^d,$$

while

$$\int_{\mathbb{T}^d} \frac{\delta \Phi(t_0, m_0, x)}{\delta m} m_0(dx) = \int_{\mathbb{T}^d} \left( u(t_0, x) - \int_{\mathbb{T}^d} u(t_0, y) m_0(dy) \right) m_0(dx) = 0.$$

Therefore

$$\frac{\delta \Phi(t_0, m_0, x)}{\delta m} = u(t_0, x) - \int_{\mathbb{T}^d} u(t_0, y) m_0(dy), \quad m_0\text{-a.e. } x \in \mathbb{T}^d.$$



This shows that the map  $x \rightarrow \frac{\delta\Phi(t_0, m_0, x)}{\delta m} - u(t_0, x)$  has a maximum on  $\mathbb{T}^d$  at  $m_0$ -a.e.  $x \in \mathbb{T}^d$  and thus (2.45) holds.

As  $\mathcal{U}^T$  satisfies a dynamic programming principle and  $\mathcal{U}^T - \Phi$  has a minimum at  $(t_0, m_0)$ , it is standard that  $\Phi$  also satisfies

$$-\partial_t \Phi(t_0, m_0) + \int_{\mathbb{T}^d} H(x, D_m \Phi(t_0, m_0, x)) m_0(dx) - \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi(t_0, m_0, x) m_0(dx) - \mathcal{F}(m_0) \geq 0.$$

Using (2.45) one then infers that (2.46) holds.  $\square$

### 2.6.3 Smooth test functions

Here we fix a corrector  $\chi$  and construct a smooth function that touches  $\chi$  from above. We fix  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and  $\tau > 0$ . We know from [68, Appendix] that, if  $(\bar{m}, \bar{\alpha})$  verifies

$$\chi(m_0) = \int_0^{2\tau} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + 2\lambda\tau + \chi(\bar{m}_{2\tau}), \quad (2.47)$$

then there exists a couple  $(\bar{u}, \bar{m})$  which solves the MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, 2\tau], \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, 2\tau], \\ m(0) = m_0, \end{cases} \quad (2.48)$$

such that  $\bar{\alpha} = D_p H(x, D\bar{u})$ .

For any  $m_1 \in \mathcal{P}(\mathbb{T}^d)$  we define  $m(t)$  and  $\alpha(t)$  as follows. We first consider  $\tilde{m}$  solution of

$$\begin{cases} -\partial_t \tilde{m} + \Delta \tilde{m} + \operatorname{div}(\tilde{m} D_p H(x, D\bar{u})) = 0 & \text{in } [0, \tau] \times \mathbb{T}^d \\ \tilde{m}(0) = m_1 \end{cases} \quad (2.49)$$

and then we set

$$m(s, x) = \begin{cases} \tilde{m}(s, x), & \text{if } s \in (0, \tau] \\ \frac{2\tau-s}{\tau} \tilde{m}(s, x) + \frac{s-\tau}{\tau} \bar{m}(s, x), & \text{if } s \in [\tau, 2\tau]. \end{cases}$$

Let  $\zeta : [0, 2\tau] \times \mathbb{T}^d \rightarrow \mathbb{R}$  be the solution to  $\Delta \zeta(t) = \bar{m}(t) - \tilde{m}(t)$  so that  $\int_{\mathbb{T}^d} \zeta(s, x) = 0$ . Then, the drift  $\alpha$  will be  $\alpha(t, x) = D_p H(x, D\bar{u}(t, x)) + \frac{D\zeta(t, x)}{\tau m(t, x)}$  in  $[\tau, 2\tau]$  and  $\alpha(t, x) = D_p H(x, D\bar{u}(t, x))$  elsewhere.

We define the function  $\Psi(m_1)$  as

$$\Psi(m_1) = \int_0^{2\tau} \int_{\mathbb{T}^d} H^*(x, \alpha(t)) dm(t) + \mathcal{F}(m(t)) dt + 2\lambda\tau + \chi(\bar{m}(2\tau)). \quad (2.50)$$

**Proposition 2.6.2.** *The function  $\Psi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  defined in (2.50) is twice differentiable with respect to  $m$  with  $C^2$  continuous derivatives in space and with derivatives bounded independently of  $\chi$ .*

*Proof.* We first introduce  $\Gamma : \mathbb{R}^+ \times \mathbb{T}^d \times \mathbb{R}^+ \times \mathbb{T}^d \rightarrow \mathbb{R}$ , the fundamental solution of (2.49), i.e.  $\Gamma(\cdot, \cdot; s, x)$  is the solution of (2.49) starting at time  $s$  with initial condition  $\Gamma(s, y; s, x) = \delta_x(y)$ . Then, by superposition, the solution  $\tilde{m}(t)$  of (2.49) is given by  $\tilde{m}(t, x) = \int_{\mathbb{T}^d} \Gamma(t, x; 0, y) m_1(dy)$  (for  $t > 0$ ).

We consider separately the following two integrals:

$$I_1(m_1) = \int_0^\tau \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t)) d\bar{m}(t) + \mathcal{F}(\bar{m}(t)) dt \quad (2.51)$$

and

$$I_2(m_1) = \int_\tau^{2\tau} \int_{\mathbb{T}^d} H^*(x, \alpha(t)) dm(t) + \mathcal{F}(m(t)) dt.$$

Note that  $\Psi = I_1 + I_2 + 2\lambda\tau + \chi(\bar{m}(2\tau))$ .

If we plug  $\Gamma$  into (2.51), then

$$I_1(m_1) = \int_0^\tau \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t)) \Gamma(t, x, 0, y) m_1(dy) dx dt + \mathcal{F} \left( \int_{\mathbb{T}^d} \Gamma(t, \cdot, 0, y) m_1(dy) \right) dt$$

We can now derive  $I_1$  with respect to  $m_1$  and we get

$$\frac{\delta I_1}{\delta m}(m_1, y) = \int_0^\tau \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t)) \Gamma(t, x, 0, y) dx dt + \int_0^\tau \int_{\mathbb{T}^d} F(x, \bar{m}(t)) \Gamma(t, x, 0, y) dx dt.$$

As the functions  $\bar{\alpha}$  and  $\bar{m}$  are smooth, standar results in parabolic equation ensures that  $\delta I_1 / \delta m(m_1, \cdot)$  is smooth (see for instance Chapter 4§14 in [60]).

We now focus on  $I_2$ . We fix  $G : \mathbb{T}^d \rightarrow \mathbb{R}$  the kernel associated to the integral representation of the solution of the Poisson equation (see for instance [4, Theorem 4.13]). Explicitly, if  $\Delta\zeta = f$ , then

$$\zeta(x) = \int_{\mathbb{T}^d} G(x - y) f(y) dy$$

We first analyse the integral  $\int_\tau^{2\tau} \int_{\mathbb{T}^d} H^*(x, \alpha(t)) dm(t)$ , which explicitly becomes

$$\int_\tau^{2\tau} \int_{\mathbb{T}^d} H^* \left( x, D_p H(x, D\bar{u}) + \frac{D\zeta}{(2\tau - s)\bar{m} + (s - \tau)\bar{m}} \right) \left( \frac{2\tau - s}{\tau} \tilde{m}(s, x) + \frac{s - \tau}{\tau} \bar{m}(s, x) \right).$$

If we derive the above expresion we get

$$\begin{aligned} \frac{1}{\tau} \int_\tau^{2\tau} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} -D_p H^*(\alpha) \frac{G(x - z) D_z \Gamma_0(t, z; y) m(t, x) + (2\tau - s) \Gamma_0(t, x; y) D\zeta(t, x)}{m(t, x)} dx dz \\ + \frac{1}{\tau} \int_\tau^{2\tau} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} H^*(\alpha) (2\tau - s) \Gamma_0(t, x; y) dx dz. \end{aligned} \quad (2.52)$$

Therefore,  $\frac{\delta I_2}{\delta m}(m_1, y)$  is equal to

$$\begin{aligned} \frac{1}{\tau} \int_\tau^{2\tau} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} -D_p H^*(\alpha) \frac{G(x - z) D_z \Gamma_0(t, z; y) m(t, x) + (2\tau - s) \Gamma_0(t, x; y) D\zeta(t, x)}{m(t, x)} dx dz dt + \\ \frac{1}{\tau} \int_\tau^{2\tau} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} H^*(x, \alpha) (2\tau - s) \Gamma_0(t, x; y) dx dz + \frac{1}{\tau} \int_\tau^{2\tau} \int_{\mathbb{T}^d} (2\tau - s) F(x, m(s, x)) \Gamma_0(t, x; y) dx dt \end{aligned}$$

and

$$\frac{\delta \Psi}{\delta m}(m_1, y) = \frac{\delta I_1}{\delta m}(m_1, y) + \frac{\delta I_2}{\delta m}(m_1, y)$$

Note that, as in the above expression we are looking at a time interval bounded away from zero, the parabolic regularity ensures that all the functions therein are smooth with respect to the state variable and that  $m(t, x)$  is bounded away from zero. This implies that also  $D_m \Psi$  is well defined.

We omit the proof for second order derivatives. It does not present any further difficulties. Indeed, the parabolic regularity, enjoyed by the solutions of the MFG system at any time  $t > 0$ , ensures that we can deploy the same kind of computations that we used in (2.52) and so that both  $D_{mm}^2 \Psi$  and  $D_{my}^2 \Psi$  are well defined and bounded.  $\square$

**Lemma 2.6.3.** *For any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ , there exists a function  $\bar{\Psi} \in C^2(\mathcal{P}(\mathbb{T}^d))$  such that  $\bar{\Psi}(m) > \chi(m)$  for any  $m \neq m_0$  and  $\bar{\Psi}(m_0) = \chi(m_0)$ .*

*Moreover, we can choose  $\bar{\Psi}$  such that  $D_{mm}^2 \bar{\Psi}$  and  $D_{ym}^2 \bar{\Psi}$  are bounded independently of  $\chi$  and with  $D_m \bar{\Psi}(m_0, x) = D\bar{u}(0, x)$  where  $\bar{u}$  is defined in (2.47) and (2.48).*

*Proof.* Let  $\{\phi_n\}_n$  be a countable collection of  $C^\infty(\mathbb{T}^d)$  maps such that  $\{\phi_n\}_n$  is dense in the set of  $\text{Lip}_1(\mathbb{T}^d)$ , which is the set 1-Lipschitz function on  $\mathbb{T}^d$ . Then,

$$\mathbf{d}(m, m_0) = \sup_{f \in \text{Lip}_1(\mathbb{T}^d)} \int_{\mathbb{T}^d} f(x)(m - m_0)(dx) = \sup_{n \in \mathbb{N}} \int_{\mathbb{T}^d} \phi_n(x)(m - m_0)(dx).$$

We define  $Q : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  as follows

$$Q(m) = \sum_{n \in \mathbb{N}} \frac{(\int_{\mathbb{T}^d} \phi_n(x)(m - m_0)(dx))^2}{(n+1)^2(\|\phi_n\|_\infty + \|D^2 \phi_n\|_\infty + 1)}.$$

The denominator in the above fraction ensures that  $Q$  is well defined for any  $m \in \mathcal{P}(\mathbb{T}^d)$ . Note that  $Q(m) = 0$  if and only if, for any  $n \in \mathbb{N}$ ,  $\int_{\mathbb{T}^d} \phi_n(x)(m - m_0)(dx) = 0$ . In this case,  $\mathbf{d}(m, m_0) = 0$  and so  $m = m_0$ . One easily checks that  $Q$  is smooth and that its derivatives

$$\begin{aligned} D_m Q(m, y) &= 2 \sum_{n \in \mathbb{N}} \frac{(\int_{\mathbb{T}^d} \phi_n(x)(m - m_0)(dx)) D\phi_n(y)}{(n+1)^2(\|\phi_n\|_\infty + \|D^2 \phi_n\|_\infty + 1)}, \\ D_{my}^2 Q(m, y) &= 2 \sum_{n \in \mathbb{N}} \frac{(\int_{\mathbb{T}^d} \phi_n(x)(m - m_0)(dx)) D^2 \phi_n(y)}{(n+1)^2(\|\phi_n\|_\infty + \|D^2 \phi_n\|_\infty + 1)}, \end{aligned}$$

and

$$D_{mm}^2 Q(m, y, z) = \sum_{n \in \mathbb{N}} \frac{D\phi_n(y) \otimes D\phi_n(z)}{(n+1)^2(\|\phi_n\|_\infty + \|D^2 \phi_n\|_\infty + 1)}$$

are bounded. Note also that  $D_m Q(m_0, y) = 0$  for any  $y \in \mathcal{P}(\mathbb{T}^d)$ . We can now define

$$\bar{\Psi}(m) = \Psi(m) + Q(m),$$

where  $\Psi$  is the function defined in (2.50). By construction,  $\bar{\Psi}$  is such that  $\bar{\Psi}(m) > \chi(m)$  for any  $m \neq m_0$  and  $\bar{\Psi}(m_0) = \chi(m_0)$ . Moreover,

$$D_m \bar{\Psi}(m_0, y) = D_m \Psi(m_0, y) + D_m Q(m_0, y) = D_m \Psi(m_0, y) = D\bar{u}(0, y).$$

The boundedness of the derivatives comes from Proposition 2.6.2 and the properties of  $Q$  that we discussed above.  $\square$

## Chapter 3

# Convergence of the MFG discounted HJ equation

We consider the solution  $\mathcal{V}_\delta$  of the discounted Hamilton-Jacobi equation in the Wasserstein space arising from potential MFG and we prove its full convergence to a corrector function  $\chi_0$ . We follow the structure of the proof of the analogue result in the finite dimensional setting provided by Davini, Fathi, Iturriaga, Zavidovique in 2016. We characterize the limit  $\chi_0$  through a particular set of smooth Mather measures. A major point that distinguishes the techniques deployed in the standard setting from the ones that we use here is the lack of mollification in the Wasserstein space.

This chapter was submitted to *ESAIM: Control, Optimisation and Calculus of Variations* (ESAIM: COCV).

### 3.1 Introduction

The Mean Field Games theory (briefly MFG) is a branch of the broader field of dynamic games which is devoted to the analysis of those models where a large number of *small* players interact with each others. This theory was introduced simultaneously and independently by Lasry and Lions [61, 62] and Huang, Caines and Malhamé [57]. Under appropriate assumptions, the Nash equilibria of this models can be analyzed through the solutions of the, so called, MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{R}^d \times [0, T] \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{R}^d \times [0, T] \\ m(0) = m_0, u(T, x) = G(x, m(T)) & \text{in } \mathbb{R}^d \end{cases}$$

with unknown the couple  $(u, m)$ . The value  $u(t, x)$  is the best that a player can get starting from  $x$  at time  $t$  while  $m(t)$  is a probability measure that represents the distribution of players at time  $t$ .

In this paper we will consider a specific class of MFG which is the class of *potential MFG*. When the functions  $F$  and  $G$  are respectively the derivatives of the potentials  $\mathcal{F}$  and  $\mathcal{G}$ , the MFG system can be derived as optimality condition of the following minimization problem

$$\mathcal{U}^T(t, m_0) = \inf_{(m, \alpha)} \int_t^T \int_{\mathbb{R}^d} H^*(x, \alpha(s, x)) dm(s) + \mathcal{F}(m(s)) ds + \mathcal{G}(m(T)),$$

where  $(m, \alpha)$  verifies the Fokker Plank equation  $-\partial_t m + \Delta m + \operatorname{div}(\alpha m) = 0$  with  $m(t) = m_0$  (the precise assumptions on  $(m, \alpha)$  can be found in Section 3.2.2) and  $H^*$  is the Fenchel conjugate of  $H$  with respect to the second variable.

Starting from [62], where this class of MFG was introduced, several papers have been focusing on this setting. The variational structure of these models often allows to push the analysis further than in the standard setting. See for instance [21, 70, 20] for existence results or [63, 72, 27] for regularity. An other reason to exploit the variational structure of potential MFG is to understand how the solutions of the MFG system behave when the time horizon goes to infinity. The problem of the long time convergence has been addressed in different papers starting from [64] and the Mexican wave model in [56] to more recent contributions in [18, 22, 23, 50]. In [68, 26] the authors tackled the problem using techniques from weak KAM theory. Following Fathi's seminal papers [39, 40, 41] and his book [42], they adapted the main arguments of the weak KAM theory into the infinite dimensional framework of MFG. This approach allows to disregard the convexity of the variational problem  $\mathcal{U}^T$  so that a new range of models can be analyzed. Note that, infinite dimensional weak KAM theorems were already known in the context of Wasserstein spaces (see [45, 46, 47, 48]).

This papers is meant to answer a natural question that arises in this theory and more specifically in the context of *ergodic approximation*. In [26], it was proved that there exists a critical value  $\lambda \in \mathbb{R}$  such that  $\mathcal{U}^T(0, \cdot) + \lambda T$  uniformly converges to a corrector function  $\chi$  when  $T \rightarrow +\infty$ . We say that  $\chi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is a corrector function if, for any  $t \in \mathbb{R}$ ,  $\chi$  verifies the following dynamical programming principle

$$\chi(m_0) = \inf_{(m, \alpha)} \left( \int_0^t H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t, \quad (3.1)$$

where  $(m, \alpha)$  solves in the sense of distributions  $-\partial_t m + \Delta m + \operatorname{div}(m \alpha) = 0$  with initial condition  $m(0) = m_0$ . A fundamental result that was needed to get the above convergence was to know a priori that the set of corrector functions was not empty. In [68] this was proven through the

so called ergodic approximation. As in [66], the idea is to define the infinite horizon discounted problem

$$\mathcal{V}_\delta(m_0) = \inf_{(m,\alpha)} \int_0^\infty e^{-\delta t} \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) dm(t) + \mathcal{F}(m(t)) dt.$$

Letting  $\delta \rightarrow 0$ , one gets that  $\delta \mathcal{V}_\delta(\cdot) \rightarrow -\lambda$  and that, up to subsequence,  $\mathcal{V}_\delta(\cdot) - V_\delta(m_0)$  uniformly converges to a corrector function  $\chi$ .

It was proved by Davini, Fathi, Iturriaga and Zavidovique [35] that, in the standard finite dimensional setting, the solution of the discounted equation converges to a solution of the critical Hamilton-Jacobi equation, i.e. to a corrector function. In this paper we prove the analogous result in the context of potential MFG. If we define

$$\bar{\mathcal{V}}_\delta(m) = \mathcal{V}_\delta(m) + \frac{\lambda}{\delta},$$

then the main result is the following. The whole family  $\bar{\mathcal{V}}_\delta$  converges, as  $\delta$  tends to zero, to a corrector function  $\chi_0$ . Moreover, we can characterize this limit as

$$\chi_0(m) = \sup_{\chi \in \mathcal{S}^-} \chi(m).$$

where  $\mathcal{S}^-$  is the set of subsolutions  $\chi$  of

$$- \int_{\mathbb{T}^d} \operatorname{div}_y D_m \chi(m, y) m(dy) + \int_{\mathbb{T}^d} H(y, D_m \chi(m, y)) m(dy) - \mathcal{F}(m) = \lambda, \quad m \in \mathcal{P}(\mathbb{T}^d) \quad (3.2)$$

such that  $\int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu(dm) \leq 0$  for any  $\nu \in \mathcal{M}_\nu$ . We say that  $\nu \in \mathcal{M}_\nu$  if it is induced by an optimal trajectory for  $\mathcal{V}_\delta$ : if  $(m_\delta, \alpha_\delta)$  is an optimal trajectory for  $\bar{\mathcal{V}}_\delta(m_0)$  we define  $\nu_\delta^{m_0} \in \mathcal{P}(\mathbb{T}^d \times C^1(\mathbb{T}^d, \mathbb{R}^d))$  as follows

$$\int_{\mathcal{P}(\mathbb{T}^d) \times C^1(\mathbb{T}^d, \mathbb{R}^d)} f(m, \alpha) \nu_\delta^{m_0}(dm, d\alpha) = \delta \int_0^\infty e^{-\delta s} f(m^\delta(s), \alpha^\delta(s)) ds.$$

If  $\nu^{m_0}$  is the limit of  $\nu_\delta^{m_0}$  and  $m_0$  has smooth density, then  $\nu^{m_0} \in \mathcal{M}_\nu$ . Moreover, we will also prove that if  $\nu^{m_0} \in \mathcal{M}_\nu$  then  $\nu^{m_0}$  is a *Mather measure* (Definition 3.2.4). This selection principle for Mather measures was firstly introduced in the finite dimensional setting in [53].

The structure of the paper itself is inspired by the one in [35]. Even though the steps to get to the result are mostly the same, the techniques deployed to prove the main points are quite different. In [35], a major ingredient is that one can approximate a viscosity solution with a smooth function and this approximation is also an approximated solution of the equation. While this is quite standard in the finite dimensional setting and it is generally proved through mollification, in the space of functions over  $\mathcal{P}(\mathbb{T}^d)$  one cannot expect the same. In [71] it was recently proved that one can uniformly approximate a continuous function on  $\mathcal{P}(\mathbb{T}^d)$  with a sequence of smooth ones (similar approximation were also introduced in [64]). The problem is that this convergence is not as strong as the one given by mollification in the finite dimensional setting. More specifically, we know that if  $\chi$  is a corrector function then it is also a viscosity solution of the critical equation (3.2). The result in [71] allows us to approximate  $\chi$  uniformly but one cannot expect that this approximation is also an approximated solution of (3.2). This lack of regularity prevented us to work at the level of solutions of the critical equation and it forced us to rely only on the dynamical programming principle (3.1) and the properties of optimal trajectories. Moreover, the lack of regularity led to an other difference. The characterization of

the limit  $\chi_0$  used by Davini et al. is slightly different. In their definition of  $\mathcal{S}^-$  the subsolutions are tested against any Mather measure and not only against the subset  $\mathcal{M}_V$ .

We now briefly discuss the structure of the paper. In Section 3.2 we set the problem in term of assumptions and notation and we collect the main results previously proved in [68, 26] that are used in the paper.

Section 3.3 is devoted to the convergence of  $\bar{V}_\delta$ . In Subsection 3.3.1 we introduce the notion of probability measures induced by optimal trajectories of  $\mathcal{V}_\delta$  and we prove that the limit of these probability measures are Mather measures.

In Subsection 3.3.2 we prove that, if  $\chi$  is a subsolution of (3.2) and  $\nu_\delta^{m_0}$  is a probability measure induced by the optimal trajectory  $(m_\delta, \alpha_\delta)$ , then

$$\bar{V}_\delta(m_0) \geq \chi(m_0) - \int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu_\delta^{m_0}(dm).$$

The proof is done comparing the dynamic programming principle verified by  $\bar{V}_\delta$  and the one verified by the subsolution  $\chi$ .

In Subsection 3.3.3 we introduce the notion of *smooth Mather measure* (Definition 3.3.4) which are Mather measures which have smooth densities. The smoothness of these measures allows to overcome the lack of regularity of  $\bar{V}_\delta$  so that we can prove that if  $(\mu, p_1)$  is a smooth Mather measure then

$$\delta \int_{\mathcal{P}(\mathbb{T}^d)} \bar{V}_\delta(\bar{m}) \mu(d\bar{m}) \leq 0.$$

Moreover, we prove that if  $\nu^{m_0} \in \mathcal{M}_V$  then  $\nu^{m_0}$  a smooth Mather measure.

In Subsection 3.3.4 we collect all the previous results and we finally prove that the limit  $\chi_0$  of  $\bar{V}_\delta$  is uniquely defined by

$$\chi_0(m) = \sup_{\chi \in \mathcal{S}^-} \chi(m).$$

## 3.2 Assumptions and preliminary results

### 3.2.1 Notation and assumptions

As it was mentioned in the introduction, this paper is meant to expand the weak KAM theory in the context of MFG introduced in [26] and [68]. Therefore, we will suppose that the very same assumptions are in place.

**Remark 3.2.1.** *Note that some of these hypothesis will not be explicitly used in this paper; especially the ones on the growth of the Hamiltonian and its derivatives. Nonetheless, as we give for granted several results of [26, 68] we still need to impose them.*

We will use as state space the  $d$ -dimensional flat torus  $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ . This domain is chosen to avoid boundary conditions and to set the problem on a compact domain. We denote respectively by  $\mathcal{M}(\mathbb{T}^d)$  and  $\mathcal{P}(\mathbb{T}^d)$  the set of Borel measures and probability measures on  $\mathbb{T}^d$ . The set  $\mathcal{P}(\mathbb{T}^d)$  is a compact, complete and separable metric space when endowed with the 1-Wasserstein distance  $\mathbf{d}(\cdot, \cdot)$ . If  $m \in C^0([t, T], \mathcal{P}(\mathbb{T}^d))$  then we set  $L_m^2([t, T] \times \mathbb{T}^d)$  the set of  $m$ -measurable functions  $f$  such that the integral of  $|f|^2 dm(s)$  over  $[t, T] \times \mathbb{T}^d$  is finite.

We use the notion of derivative on the metric space  $\mathcal{P}(\mathbb{T}^d)$  introduced in [19]. We say that  $\Phi : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is  $C^1$  if there exists a continuous function  $\frac{\delta\Phi}{\delta m} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}$  such that

$$\Phi(m_1) - \Phi(m_2) = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta\Phi}{\delta m}((1-t)m_1 + tm_2, x)(m_1 - m_2)(dx)dt, \quad \forall m_1, m_2 \in \mathcal{P}(\mathbb{T}^d).$$

As this derivative is defined up to an additive constant, we use the standard normalization

$$\int_{\mathbb{T}^d} \frac{\delta\Phi}{\delta m}(m, x)m(dx) = 0.$$

If  $\frac{\delta\Phi}{\delta m}$  is derivable in the second variable, we define the *intrinsic derivative*

$$D_m\Phi(m, x) = D_x \frac{\delta\Phi}{\delta m}(m, x).$$

Moreover, we say that a function  $\Phi$  belongs to  $C^{1,1}(\mathbb{T}^d)$  if  $\Phi \in C^1(\mathbb{T}^d)$ ,  $D_m\Phi$  is well defined and  $D_m\Phi$  is smooth in the second variable.

We recall that, if  $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$ , the 1-Wasserstein distance is defined by

$$\mathbf{d}(\mu, \nu) = \sup \left\{ \int_{\mathbb{T}^d} \phi(x) d(\mu - \nu)(x) \mid \text{continuous } \phi : \mathbb{T}^d \rightarrow \mathbb{R}, \text{Lip}(\phi) \leq 1 \right\}.$$

**Assumptions:** Throughout the paper we will suppose the following conditions:

1.  $H : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is of class  $C^2$ ,  $p \mapsto D_{pp}H(x, p)$  is Lipschitz continuous, uniformly with respect to  $x$ . Moreover, there exists  $\bar{C} > 0$  that verifies

$$\bar{C}^{-1}I_d \leq D_{pp}H(x, p) \leq \bar{C}I_d, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d$$

and  $\theta \in (0, 1)$ ,  $C > 0$  such that the following conditions hold true:

$$|D_x H(x, p)| \leq C + C|p| \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d$$

and

$$|D_{xx}H(x, p)| \leq C(1 + |p|)^{1+\theta}, \quad |D_{x,p}H(x, p)| \leq C(1 + |p|)^\theta, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d.$$

2.  $\mathcal{F} : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is of class  $C^2$ . Its derivative  $F : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is twice differentiable in  $x$  and  $D_{xx}^2 F$  is bounded.

### 3.2.2 Definitions and preliminary results

We define  $\mathcal{V}_\delta : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  as

$$\mathcal{V}_\delta(m_0) = \inf_{(m, \alpha)} \int_0^\infty e^{-\delta t} \int_{\mathbb{T}^d} H^*(x, \alpha) dm(t) + \mathcal{F}(m(t)) dt, \quad m_0 \in \mathcal{P}(\mathbb{T}^d) \quad (3.3)$$

where  $\delta > 0$ ,  $m \in C^0([0, +\infty), \mathcal{P}(\mathbb{T}^d))$ ,  $\alpha \in L_{m, \delta}^2([0, +\infty) \times \mathbb{T}^d, \mathbb{R}^d)$ , that is  $L_m^2$  with weight  $e^{-\delta t}$ , and  $(m, \alpha)$  verifies on  $[0, +\infty) \times \mathbb{T}^d$  the Fokker-Plank equation

$$\begin{cases} -\partial_t m + \Delta m + \text{div}(m\alpha) = 0 \\ m(0) = m_0. \end{cases} \quad (3.4)$$



Standard arguments in optimal control ensures that  $\mathcal{V}_\delta$  solves the following dynamic programming principle: for any  $t > 0$

$$\mathcal{V}_\delta(m_0) = \inf_{(m,\alpha)} \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta t} \mathcal{V}_\delta(m(t)). \quad (3.5)$$

where  $(m, \alpha)$  are defined as before. Note that one can prove through standard arguments that, at least formally,  $\mathcal{V}_\delta$  is a solution of discounted Hamilton-Jacobi equation

$$\delta \mathcal{V}_\delta(m) - \int_{\mathbb{T}^d} \operatorname{div}_y D_m \mathcal{V}_\delta(m, y) m(dy) + \int_{\mathbb{T}^d} H(y, D_m \mathcal{V}_\delta(m, y)) m(dy) - \mathcal{F}(m) = 0 \quad \forall m \in \mathcal{P}(\mathbb{T}^d).$$

As we have anticipated in the introduction we will characterize the limit of  $\mathcal{V}_\delta$  through a special class of subsolutions of the critical equation

$$- \int_{\mathbb{T}^d} \operatorname{div}_y D_m \chi(m, y) m(dy) + \int_{\mathbb{T}^d} H(y, D_m \chi(m, y)) m(dy) - \mathcal{F}(m) = \lambda, \quad m \in \mathcal{P}(\mathbb{T}^d), \quad (3.6)$$

where the value  $\lambda$

$$-\lambda = \lim_{\delta \rightarrow 0} \delta \mathcal{V}_\delta(\cdot). \quad (3.7)$$

In [68, Proposition 1.6] it is proven that the above limit is well defined and uniform with respect to the argument of  $\mathcal{V}_\delta$ . Hereafter, anytime the constant  $\lambda$  appears we implicitly refer to the one defined by (3.7)

We will most often work with functions that do not enjoy enough regularity to solve in classical sense the above equation. Therefore, we need to introduce a weaker definition of a subsolution and, accordingly, of the notion of corrector function.

**Definition 3.2.2.** *We say that a continuous real function  $\chi$  on  $\mathcal{P}(\mathbb{T}^d)$  is a subsolution of the critical equation (3.6) if, for any  $h > 0$  and  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,*

$$\chi(m_0) \leq \inf_{(m,\alpha)} \int_0^h \int_{\mathbb{T}^d} H^*(x, \alpha(t)) dm(t) + \mathcal{F}(m(t)) dt + \chi(m(h)) + \lambda h, \quad (3.8)$$

where  $m \in C^0([0, h], \mathcal{P}(\mathbb{T}^d))$ ,  $\alpha \in L_m^2([0, h] \times \mathbb{T}^d, \mathbb{R}^d)$  and  $(m, \alpha)$  verifies (3.4) in  $[0, h]$  with initial condition  $m(0) = m_0$ .

If, otherwise,  $\chi$  verifies (3.8), for any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and any  $h > 0$ , as an equality, we say that  $\chi$  is corrector function.

We recall that  $\mathcal{V}_\delta$  is uniformly Lipschitz continuous with respect to  $\delta$  and that  $\delta \mathcal{V}_\delta \rightarrow -\lambda$  ([68, Proposition 1.6]). We define

$$\bar{\mathcal{V}}_\delta(m) = \mathcal{V}_\delta(m) + \frac{\lambda}{\delta}.$$

It is clear that  $\delta \bar{\mathcal{V}}_\delta \rightarrow 0$ . Moreover we claim that if  $c \in \mathbb{R}$  is such that  $\mathcal{V}_\delta + c/\delta$  is uniformly bounded, then  $c = \lambda$ . We also claim that  $\bar{\mathcal{V}}_\delta$  converges, at least up to subsequence, to a corrector  $\chi_0$  (Definition 3.2.2). The proof of these claims is postponed in Lemma 3.3.3 for which we need some preliminary results.

We will use more than once the fact that, for any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and any  $\delta > 0$ , the minimization problem (3.3) admits minimizer  $(m_\delta, \alpha_\delta)$  and, as proved in [68, Proposition 1.1], that  $\alpha_\delta(t, x) = D_p H(x, Du_\delta(t, x))$  where,  $u_\delta \in C^{1,2}([0, +\infty) \times \mathbb{T}^d)$  and  $(u_\delta, m_\delta)$  solves

$$\begin{cases} -\partial_t u - \Delta u + \delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, +\infty) \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, +\infty) \\ m(0) = m_0, u \in L^\infty([0, +\infty) \times \mathbb{T}^d). \end{cases} \quad (3.9)$$

Moreover,  $(m_\delta, \alpha_\delta)$  enjoys the following estimates (for the proof see [68, Lemma 1.3]):

**Lemma 3.2.3.** *There exists  $C_1 > 0$  independent of  $m_0, \delta$  such that, if  $(u_\delta, m_\delta)$  is a classical solution of (3.9), then*

- $\|Du_\delta\|_{L^\infty([0, +\infty) \times \mathbb{T}^d)} \leq C_1$
- $\|D^2 u_\delta\|_{L^\infty([0, +\infty) \times \mathbb{T}^d)} \leq C_1$ .
- $\mathbf{d}(m_\delta(s), m_\delta(l)) \leq C_1 |l - s|^{1/2}$  for any  $l, s \in [0, +\infty)$

Consequently, we also have that  $\|\partial_t u_\delta\|_{L^\infty([0, +\infty) \times \mathbb{T}^d)} \leq C_1$  for any  $s \in [0, +\infty)$ .

**Definition 3.2.4.** *We call Mather measure any minimizer of the following minimization problem*

$$\inf_{(\mu, p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} H^*\left(y, \frac{dp_1}{dm \otimes \mu}\right) m(dy) + \mathcal{F}(m) \mu(dm), \quad (3.10)$$

where the minimum is taken over  $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$  and  $p_1$ , Borel vector measure on  $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$  with value in  $\mathbb{R}^d$ , such that  $p_1$  is absolutely continuous with respect to the measure  $dm \otimes \mu := m(dy)\mu(dm)$  and such that  $(\mu, p_1)$  is closed, in the sense that, for any  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ ,

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi(m, y) m(dy) \mu(dm) = 0. \quad (3.11)$$

As in the standard Aubry-Mather theory, in [26] it was proven that, if  $(\mu, p_1)$  is a Mather measure, then

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} H^*\left(y, \frac{dp_1}{dm \otimes \mu}\right) m(dy) + \mathcal{F}(m) \mu(dm) = -\lambda.$$

### 3.3 Convergence of $\bar{\mathcal{V}}_\delta$

#### 3.3.1 Measures induced by optimal trajectories

Let  $(m_\delta, \alpha_\delta)$  be an optimal trajectory for  $\bar{\mathcal{V}}_\delta(m_0)$ , then if  $C_1 > 0$  is the constant that appears in Lemma (3.2.3), we have  $\|\alpha_\delta\|_{W^{1,\infty}(\mathbb{T}^d \times [0, +\infty))} \leq C_1$ .

We define  $\nu_\delta^{m_0} \in \mathcal{P}(\mathbb{T}^d \times E)$  as follows

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} f(m, \alpha) \nu_\delta^{m_0}(dm, d\alpha) = \delta \int_0^\infty e^{-\delta s} f(m^\delta(s), \alpha^\delta(s)) ds, \quad (3.12)$$

where

$$E := \{\alpha \in W^{1,\infty}(\mathbb{T}^d, \mathbb{R}^d), \|\alpha\|_\infty + \|D\alpha\|_\infty \leq C\}.$$

Note that, we know from Lemma 3.2.3 that we can chose  $C$  independent of  $\delta$ . As  $\mathcal{P}(\mathbb{T}^d) \times E$  is compact when endowed with the uniform convergence, we can suppose that  $\nu_\delta^{m_0}$  weakly

converges, up to subsequence, to a probability measure  $\bar{\nu}$  on  $E$ . Let  $\mu$  be the first marginal of  $\bar{\nu}$  and let us define the vector measure  $p_1$  on  $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$  as

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \phi(m, y) \cdot p_1(dm, dy) = \int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \phi(m, y) \cdot \alpha(y) m(dy) \bar{\nu}(dm, d\alpha)$$

for any test function  $\phi \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, \mathbb{R}^d)$ . Note that  $p_1$  is absolutely continuous with respect to  $\mu$ , since, if we disintegrate  $\bar{\nu}$  with respect to  $\mu$ :  $\bar{\nu} = \bar{\nu}_m(d\alpha)\mu(dm)$ , then

$$p_1(dm, dy) = \int_E \alpha(y) m(dy) \bar{\nu}_m(d\alpha) \mu(dm) \quad (3.13)$$

and so

$$\frac{dp_1}{dm \otimes \mu}(m, y) = \int_E \alpha(y) \bar{\nu}_m(d\alpha).$$

In the next proposition we prove that the couple  $(\mu, p_1)$  defined above is a Mather measure.

**Proposition 3.3.1.** *Let  $(m_\delta, \alpha_\delta)$  be an optimal trajectory for  $\bar{\mathcal{V}}_\delta(m_0)$  and  $\nu_\delta^{m_0}$  be the probability measure defined by (3.12). If  $\bar{\nu} \in \mathcal{P}(\mathbb{T}^d)$  is a weak limit of  $\nu_\delta^{m_0}$  (possibly up to subsequence),  $\mu$  is the first marginal of  $\bar{\nu}$  and  $p_1$  is defined as (3.13), then  $(\mu, p_1)$  is a Mather measure.*

*Proof.* We first need to check that  $(\mu, p_1)$  is closed in the sense of (3.11). Let  $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ , then

$$\frac{d}{dt} \Phi(m_\delta(t)) = \int_{\mathbb{T}^d} \operatorname{div}(D_m \Phi(m_\delta(t), y)) m_\delta(t, dy) + \int_{\mathbb{T}^d} D_m \Phi(m_\delta(t), y) \cdot \alpha_\delta(t, y) m_\delta(t, dy).$$

So,

$$\int_0^t \int_{\mathbb{T}^d} \operatorname{div}(D_m \Phi(m_\delta(t), y)) m_\delta(t, dy) + \int_{\mathbb{T}^d} D_m \Phi(m_\delta(t), y) \cdot \alpha_\delta(t, y) m_\delta(t, dy) dt = \Phi(m_\delta(t)) - \Phi(m_0).$$

Using the above relation and integrating by parts, we get

$$\begin{aligned} & \left| \int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \operatorname{div}(D_m \Phi(m, y)) + D_m \Phi(m, y) \cdot \alpha(y) m(dy) \nu_\delta^{m_0}(dm, d\alpha) \right| \\ &= \delta \left| \int_0^{+\infty} e^{-\delta t} \int_{\mathbb{T}^d} \operatorname{div}(D_m \Phi(m_\delta(t), y)) m_\delta(t, dy) + \int_{\mathbb{T}^d} D_m \Phi(m_\delta(t), y) \cdot \alpha_\delta(t, y) m_\delta(t, dy) dt \right| \\ &\leq \delta \left| \lim_{T \rightarrow +\infty} e^{-\delta T} (\Phi(m_\delta(T)) - \Phi(m_\delta(m_0))) \right| + \delta^2 \left| \int_0^{+\infty} e^{-\delta t} (\Phi(m_\delta(t)) - \Phi(m_\delta(m_0))) dt \right| \\ &\leq \delta^2 K \int_0^{+\infty} e^{-\delta t} dt = \delta K, \end{aligned}$$

where  $K = \sup_{m, n \in \mathcal{P}(\mathbb{T}^d)} |\Phi(m) - \Phi(n)|$ . Letting  $\delta \rightarrow 0$ , we find

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \operatorname{div}(D_m \Phi(m, y)) + D_m \Phi(m, y) \cdot \alpha(y) m(dy) \bar{\nu}(dm, d\alpha) = 0.$$

According to the definition of  $p_1$  we can read the last equality as

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \operatorname{div}(D_m \Phi(m, y)) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi(m, y) \cdot p_1(dm, dy) = 0,$$

which proves that  $(\mu, p_1)$  is closed. The last step is to prove that  $(\mu, p_1)$  is a minimizer of (3.10). Indeed, by convexity of  $H^*$ , we have

$$\begin{aligned}
& \int_{\mathcal{P}(\mathbb{T}^d)} \left[ \int_{\mathbb{T}^d} H^* \left( y, \frac{dp_1}{dm \otimes \mu}(m, y) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm) \\
&= \int_{\mathcal{P}(\mathbb{T}^d)} \left[ \int_{\mathbb{T}^d} H^* \left( y, \int_E \alpha(y) \bar{\nu}_m(d\alpha) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm) \\
&\leq \int_{\mathcal{P}(\mathbb{T}^d)} \int_E \left[ \int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right] \bar{\nu}_m(d\alpha) \mu(dm) \\
&= \int_{\mathcal{P}(\mathbb{T}^d) \times E} \left[ \int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right] \bar{\nu}(dm, d\alpha) \\
&= \lim_{\delta \rightarrow 0} \int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \nu_\delta^{m_0}(dm, d\alpha) = \lim_{\delta \rightarrow 0} \delta \mathcal{V}_\delta(m_0) = -\lambda.
\end{aligned}$$

□

### 3.3.2 Lower bound

In this section we will first prove the analogue of [35, Lemma 3.5] and then the boundedness of  $\bar{V}_\delta$ . Note that, the lack of a proper mollification for functions defined on  $\mathcal{P}(\mathbb{T}^d)$ , forced us to find a proof that differs from the one in [35]. While [35] used approximation of solutions of the critical equation, we work here at the level of optimal trajectories. We also mention that analogous bounds were similarly proved in [53] in the finite dimensional setting.

**Proposition 3.3.2.** *Let  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,  $(m_\delta, \alpha_\delta)$  be an optimal trajectory for  $\bar{V}_\delta(m_0)$  and  $\nu_\delta^{m_0}$  be the probability measure defined by (3.12). Then, for any  $\chi$  subsolution of (3.6),*

$$\bar{V}_\delta(m_0) \geq \chi(m_0) - \int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu_\delta^{m_0}(dm).$$

*Proof.* The dynamic programming principle (3.5) says that, if  $(m_\delta, \alpha_\delta)$  is an optimal trajectory for  $\bar{V}_\delta(m_0)$ , then, for any  $t > 0$ ,  $(m_\delta, \alpha_\delta)$  is a minimizer of

$$\bar{V}_\delta(m_0) = \inf_{(m, \alpha)} \int_0^t e^{-\delta s} \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds + e^{-\delta t} \bar{V}_\delta(m(t)) + \frac{\lambda}{\delta}. \quad (3.14)$$

In order to keep the notation as simple as possible we define

$$\mathcal{L}(t, s) = \int_{\mathcal{P}(\mathbb{T}^d)} H^*(\alpha_\delta(t+s)) dm_\delta(t+s) + \mathcal{F}(m_\delta(t+s)) + \lambda$$

and

$$\mathbb{L}_\delta(t, h) = \int_0^h e^{-\delta s} \mathcal{L}(t, s) ds,$$

so that, according to the dynamic programming principle (3.14),

$$\bar{V}_\delta(m_0) = \bar{V}_\delta(m_\delta(t)) - e^{-\delta h} \bar{V}_\delta(m_\delta(t+h)). \quad (3.15)$$

We start with the following computation

$$\int_0^h \mathcal{L}(t, s) ds = \int_0^h e^{\delta s} e^{-\delta s} \mathcal{L}(t, s) ds = e^{\delta h} \mathbb{L}_\delta(t, h) - \delta \int_0^h e^{\delta s} \mathbb{L}_\delta(t, s) ds.$$

Plugging (3.15) into the last equality, we get

$$\int_0^h \mathcal{L}(t, s) ds = e^{\delta h} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \bar{\mathcal{V}}_\delta(m_\delta(t+h)) - \delta \int_0^h e^{\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \bar{\mathcal{V}}_\delta(m_\delta(t+s)) ds. \quad (3.16)$$

Now we can focus on  $\int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu_\delta^{m_0}(dm)$ . Using the above relations and the definitions of  $\nu_\delta^{m_0}$  and subsolution of (3.6), we get

$$\int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu_\delta^{m_0}(dm) = \delta \int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t)) dt \leq \delta \int_0^{+\infty} e^{-\delta t} \int_0^h \mathcal{L}(t, s) ds + \chi(m_\delta(t+h)) dt.$$

Arranging the terms and dividing by  $\delta$ , we find

$$\int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t)) dt - \int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t+h)) dt \leq \int_0^{+\infty} e^{-\delta t} \int_0^h \mathcal{L}(t, s) ds dt. \quad (3.17)$$

We first consider the left hand side. If we run a change of variable we get

$$\begin{aligned} & \int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t)) dt - \int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t+h)) dt \\ &= \int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t)) dt - e^{\delta h} \int_h^{+\infty} e^{-\delta t} \chi(m_\delta(t)) dt \\ &= (1 - e^{\delta h}) \int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t)) dt + e^{\delta h} \int_0^h e^{-\delta t} \chi(m_\delta(t)) dt. \end{aligned} \quad (3.18)$$

We now work on the right hand side of (3.17). Using again (3.16) we get

$$\begin{aligned} & \int_0^{+\infty} e^{-\delta t} \int_0^h \mathcal{L}(t, s) ds dt \\ &= \int_0^{+\infty} e^{-\delta t} \left( e^{\delta h} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \bar{\mathcal{V}}_\delta(m_\delta(t+h)) \right) dt - \delta \int_0^{+\infty} e^{-\delta t} \int_0^h e^{\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \bar{\mathcal{V}}_\delta(m_\delta(t+s)) ds dt. \end{aligned}$$

We now look at each addend of the above line. First we have that

$$\begin{aligned} & \int_0^{+\infty} e^{-\delta t} \left( e^{\delta h} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \bar{\mathcal{V}}_\delta(m_\delta(t+h)) \right) dt \\ &= \int_0^{+\infty} e^{-\delta(t-h)} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \int_h^{+\infty} e^{-\delta(t-h)} \bar{\mathcal{V}}_\delta(m_\delta(t)) dt \\ &= \int_0^h e^{-\delta(t-h)} \bar{\mathcal{V}}_\delta(m_\delta(t)) dt = e^{\delta h} \int_0^h e^{-\delta t} \bar{\mathcal{V}}_\delta(m_\delta(t)) dt, \end{aligned} \quad (3.19)$$

then,

$$\begin{aligned} & \left| \int_0^{+\infty} e^{-\delta t} \int_0^h e^{\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \bar{\mathcal{V}}_\delta(m_\delta(t+s)) ds dt \right| \\ & \leq \int_0^{+\infty} e^{-\delta t} \int_0^h e^{\delta s} \left| \bar{\mathcal{V}}_\delta(m_\delta(t)) - e^{-\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t+s)) \right| ds dt \\ & \leq e^{\delta h} \int_0^{+\infty} e^{-\delta t} \int_0^h \left| \bar{\mathcal{V}}_\delta(m_\delta(t)) - e^{-\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t+s)) \right| ds dt. \end{aligned} \quad (3.20)$$

Note that  $\bar{\mathcal{V}}_\delta$  is Lipschitz continuous. Moreover, we know from Lemma 3.2.3 that there exists a  $C > 0$  independent of  $m_0$  and  $t$  such that

$$\mathbf{d}(m_\delta(t+h), m_\delta(t)) \leq Ch^{\frac{1}{2}}.$$

Then, for small  $h > 0$ , if  $s \leq h$

$$\begin{aligned} \left| \bar{\mathcal{V}}_\delta(m_\delta(t)) - e^{-\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t+s)) \right| &= \left| (1 - e^{-\delta s}) \bar{\mathcal{V}}_\delta(m_\delta(t)) - e^{-\delta s} (\bar{\mathcal{V}}_\delta(m_\delta(t+s)) - \bar{\mathcal{V}}_\delta(m(t))) \right| \\ &\leq \left| \bar{\mathcal{V}}_\delta(m_\delta(t))(1 - e^{-\delta s}) \right| + \left| (\bar{\mathcal{V}}_\delta(m_\delta(t+s)) - \bar{\mathcal{V}}_\delta(m(t))) \right| \end{aligned}$$

Note that,

$$\left| \bar{\mathcal{V}}_\delta(m_\delta(t))(1 - e^{-\delta s}) \right| \leq |\bar{\mathcal{V}}_\delta(m(t))|(\delta h + o(\delta h)) = h |\bar{\mathcal{V}}_\delta(m(t))|(\delta + o(\delta)).$$

As  $\delta \bar{\mathcal{V}}_\delta(\cdot) \rightarrow 0$ , there exists a constant  $C_2 > 0$ , independent of  $\delta$ , such that  $|\bar{\mathcal{V}}_\delta(\cdot)|(\delta + o(\delta)) \leq C_2$ . Then,

$$\left| \bar{\mathcal{V}}_\delta(m_\delta(t))(1 - e^{-\delta s}) \right| \leq h C_2$$

Therefore, if  $K$  is the Lipschitz constant of  $\bar{\mathcal{V}}_\delta$ , which we know to be independent of  $\delta$  ([68, Theorem 1.5]), we have that there exists  $C_1 > 0$  such that

$$\begin{aligned} \left| \bar{\mathcal{V}}_\delta(m_\delta(t)) - e^{-\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t+s)) \right| &\leq |\bar{\mathcal{V}}_\delta(m(t))|(\delta h + o(\delta h)) + \left| (\bar{\mathcal{V}}_\delta(m_\delta(t+s)) - \bar{\mathcal{V}}_\delta(m(t))) \right| \\ &\leq C_2 h + K \mathbf{d}(m_\delta(t+s), m_\delta(t)) \leq C_1 h^{\frac{1}{2}} \end{aligned}$$

Coming back to (3.20), we get

$$\begin{aligned} \left| \int_0^{+\infty} e^{-\delta t} \int_0^h e^{\delta s} \bar{\mathcal{V}}_\delta(m_\delta(t)) - \bar{\mathcal{V}}_\delta(m_\delta(t+s)) ds dt \right| & \tag{3.21} \\ &\leq e^{\delta h} \int_0^{+\infty} e^{-\delta t} C_1 h h^{\frac{1}{2}} dt = \frac{e^{\delta h}}{\delta} C_1 h h^{\frac{1}{2}}. \end{aligned}$$

If we plug (3.18), (3.19) and (3.21) into (3.17), we get

$$\begin{aligned} (1 - e^{\delta h}) \int_0^{+\infty} e^{-\delta(t)} \chi(m_\delta(t)) dt + e^{\delta h} \int_0^h e^{-\delta t} \chi(m_\delta(t)) dt \\ \leq e^{\delta h} \int_0^h e^{-\delta t} \bar{\mathcal{V}}_\delta(m_\delta(t)) dt + \frac{e^{\delta h}}{\delta} C_1 h h^{\frac{1}{2}}. \end{aligned}$$

Now, we just need to divide by  $h$  and let  $h$  go to 0 to find

$$-\delta \int_0^{+\infty} e^{-\delta t} \chi(m_\delta(t)) dt + \chi(m_0) \leq \bar{\mathcal{V}}_\delta(m_0).$$

□

We can now prove that  $\bar{\mathcal{V}}$  does actually converge, up to subsequence, to a corrector  $\chi_0$ .

**Lemma 3.3.3.** *The family of function  $\{\bar{\mathcal{V}}_\delta\}_\delta$  is uniformly bounded, uniformly Lipschitz continuous and therefore admits, up to subsequence, a uniform limit  $\chi_0$ . Moreover,  $\chi_0$  is a corrector function.*

*Proof.* The only thing that we have to prove is that  $\bar{\mathcal{V}}_\delta$  is uniformly bounded. Indeed, the rest of the claim was proven in [68, Proposition 1.6] and [68, Lemma 1.7].

We start showing that  $\bar{\mathcal{V}}_\delta$  is bounded below. Let  $\chi$  be a corrector function and  $\bar{m} \in \mathcal{P}(\mathbb{T}^d)$  be such that  $\max_{m \in \mathcal{P}(\mathbb{T}^d)} \chi(m) = \chi(\bar{m})$  (the existence of a corrector function is guaranteed by [68, Proposition 1.6] and [68, Lemma 1.7]). From Proposition 3.3.2, if we fix  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  we get

$$\bar{\mathcal{V}}_\delta(m_0) \geq \chi(m_0) - \int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu_\delta^{m_0}(dm).$$

Then,

$$\begin{aligned} \bar{\mathcal{V}}_\delta(m_0) &\geq \chi(m_0) - \int_{\mathcal{P}(\mathbb{T}^d)} \max_{m \in \mathcal{P}(\mathbb{T}^d)} \chi(m) \nu_\delta^{m_0}(dm) \\ &= \chi(m_0) - \chi(\bar{m}) \geq -K \mathbf{d}(m_0, \bar{m}) \geq -K \text{diam}(\mathcal{P}(\mathbb{T}^d)) \end{aligned}$$

and the Lipschitz constant  $K$  does not depend on  $\chi$  (see [68, Proposition 1.8]).

We now focus on the upper bound. We fix  $(m, \alpha)$  an optimal trajectory for  $\chi(m_0)$  which means that, for any  $t > 0$ ,

$$\chi(m_0) - \chi(m(t)) = \int_0^t \int_{\mathbb{T}^d} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s)) ds + \lambda t. \quad (3.22)$$

To have a lighter notation we introduce

$$\mathcal{L}(s) = \int_{\mathcal{P}(\mathbb{T}^d)} H^*(x, \alpha(s)) dm(s) + \mathcal{F}(m(s))$$

and

$$\mathbb{L}(t) = \int_0^t \mathcal{L}(s) ds.$$

Note that according to (3.22) we have

$$\mathbb{L}(t) = \chi(m_0) - \chi(m(t)) - t\lambda. \quad (3.23)$$

Moreover, we recall that  $\bar{\mathcal{V}}_\delta$  verifies the dynamical programming principle (3.14). Therefore,

$$\bar{\mathcal{V}}_\delta(m_0) \leq \int_0^T e^{-\delta s} \mathcal{L}(s) ds + e^{-\delta T} \bar{\mathcal{V}}_\delta(m(T)) + \frac{\lambda}{\delta}.$$

Integrating by parts we find

$$\bar{\mathcal{V}}_\delta(m_0) \leq e^{-\delta T} \mathbb{L}(T) + \delta \int_0^T e^{-\delta s} \mathbb{L}(s) ds + e^{-\delta T} \bar{\mathcal{V}}_\delta(m(T)) + \frac{\lambda}{\delta}. \quad (3.24)$$

We now consider each addends on the right hand side to prove that it either converges to zero when  $T \rightarrow +\infty$  or it is uniformly bounded. From (3.23), we deduce that

$$e^{-\delta T} \mathbb{L}(T) = e^{-\delta T} (\chi(m_0) - \chi(m(T)) - T\lambda) \xrightarrow{T \rightarrow +\infty} 0.$$

Then, there exists a  $C > 0$  such that

$$\begin{aligned} \delta \int_0^T e^{-\delta s} \mathbb{L}(s) ds &= \delta \int_0^T e^{-\delta s} (-s\lambda + \chi(m_0) - \chi(m(s))) ds = \\ &= -(1 - e^{-\delta T}) \frac{\lambda}{\delta} + \delta \int_0^T e^{-\delta s} (\chi(m_0) - \chi(m(s))) ds \leq \\ &= -(1 - e^{-\delta T}) \frac{\lambda}{\delta} + (1 - e^{-\delta T}) C \xrightarrow{T \rightarrow +\infty} -\frac{\lambda}{\delta} + C, \end{aligned}$$

where in the last inequality we used the boundedness of  $\chi$ . Note that, for fixed  $\delta$ , also the term  $e^{-\delta T} \bar{V}_\delta(m(T))$  converges to zero when  $T \rightarrow +\infty$ . Then, plugging these computations into (3.24) and letting  $T \rightarrow +\infty$ , we get

$$\bar{V}_\delta(m_0) \leq -\frac{\lambda}{\delta} + C + \frac{\lambda}{\delta} = C.$$

□

### 3.3.3 Smooth Mather measures

In this section we will work with measures with a smooth density. In this case, we will often identify a measure  $m \in \mathcal{P}(\mathbb{T}^d)$  with its density.

**Definition 3.3.4.** *We say that  $(\mu, p_1) \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d)) \times \mathcal{M}(\mathcal{P}(\mathbb{T}^d) \times E, \mathbb{R}^d)$  is smooth if there exists  $C > 0$  such that, for any  $m \in \text{supp } \mu$ ,*

1.  $\|m\|_{C^2(\mathbb{T}^d)} \leq C$ .
2.  $\|dp_1/(dm \otimes \mu)(\cdot, m)\|_{C^1(\mathbb{T}^d)} \leq C$ .

**Lemma 3.3.5.** *Let  $(\mu, p_1)$  be a smooth (in the sense Definition 3.3.4) closed measure. Then, if  $\Phi \in C^0(\mathcal{P}(\mathbb{T}^d))$ ,*

$$\int_{\mathcal{P}(\mathbb{T}^d)} \left( \Delta m(x) + \text{div} \left( \frac{dp_1}{dm \otimes \mu}(x, m) m(x) \right) \right) \Phi(m) \mu(dm) = 0$$

*Proof.* First of all, note that the left hand side of the above relation is integrable because of the bounds on the space derivatives of  $m$  and  $p_1/dm \otimes \mu$ .

According to [71, Theorem 2.2], we can pick a sequence  $\Phi_n \in C^{1,1}$  which converges uniformly to  $\Phi$ . Then, using that  $(\mu, p_1)$  is closed and smooth we get

$$\begin{aligned} 0 &= - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi_n(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \text{div}_y D_m \Phi_n(m, y) m(dy) \mu(dm) \\ &= - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi_n(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(x, m) dm(x) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \text{div}_y D_m \Phi_n(m, y) m(dy) \mu(dm) \\ &= \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left( \Delta m(x) + \text{div} \left( \frac{dp_1}{dm \otimes \mu}(x, m) m(x) \right) \right) \Phi_n(m) \mu(dm), \end{aligned}$$

that is

$$0 = \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left( \Delta m(x) + \text{div} \left( \frac{dp_1}{dm \otimes \mu}(x, m) m(x) \right) \right) \Phi_n(m) \mu(dm)$$

As  $\Phi_n$  converges uniformly to  $\Phi$  we just need to pass to the limit  $n \rightarrow +\infty$ . □



**Proposition 3.3.6.** *For any  $\delta > 0$  and any smooth Mather measure  $(\mu, p_1)$ , we have*

$$\int_{\mathcal{P}(\mathbb{T}^d)} \bar{\mathcal{V}}_\delta(m) \mu(dm) \leq 0$$

*Proof.* For any  $\bar{m} \in \mathcal{P}(\mathbb{T}^d)$  there exists a smooth function  $\Phi^{\bar{m}} \in C^{1,1}$  such that  $\Phi^{\bar{m}} \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ ,  $\Phi^{\bar{m}}(m) \geq \bar{\mathcal{V}}_\delta(m)$  with an equality only for  $m = \bar{m}$  and such that

$$\delta \bar{\mathcal{V}}_\delta(\bar{m}) \leq \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi^{\bar{m}}(\bar{m}, y) \bar{m}(dy) - \int_{\mathbb{T}^d} H(y, D_m \Phi^{\bar{m}}(\bar{m}, y)) \bar{m}(dy) + \mathcal{F}(\bar{m}) + \lambda.$$

For the construction of such a function one can look at [26, Lemma 6.3].

By convexity of  $H$  with respect to the second variable, we get

$$\begin{aligned} \delta \bar{\mathcal{V}}_\delta(\bar{m}) &\leq \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi^{\bar{m}}(\bar{m}, y) \bar{m}(dy) - \int_{\mathbb{T}^d} D_m \Phi^{\bar{m}}(\bar{m}, y) \cdot \frac{dp_1}{dm \otimes \mu}(\bar{m}, y) \bar{m}(dy) \\ &\quad + \int_{\mathbb{T}^d} H^* \left( y, \frac{dp_1}{dm \otimes \mu}(\bar{m}, y) \right) \bar{m}(dy) + \mathcal{F}(\bar{m}) + \lambda. \end{aligned}$$

If  $\bar{m} \in \operatorname{supp} \mu$  then we can integrate by parts and we get

$$\begin{aligned} \delta \bar{\mathcal{V}}_\delta(\bar{m}) &\leq \int_{\mathbb{T}^d} \left( \Delta \bar{m}(x) + \operatorname{div} \left( \frac{dp_1}{dm \otimes \mu}(x, \bar{m}) \bar{m}(x) \right) \right) \Phi^{\bar{m}}(\bar{m}) \\ &\quad + \int_{\mathbb{T}^d} H^* \left( y, \frac{dp_1}{dm \otimes \mu}(\bar{m}, y) \right) \bar{m}(dy) + \mathcal{F}(\bar{m}) + \lambda \\ &= \int_{\mathbb{T}^d} \left( \Delta \bar{m}(x) + \operatorname{div} \left( \frac{dp_1}{dm \otimes \mu}(x, \bar{m}) \bar{m}(x) \right) \right) \bar{\mathcal{V}}_\delta(\bar{m}) \\ &\quad + \int_{\mathbb{T}^d} H^* \left( y, \frac{dp_1}{dm \otimes \mu}(\bar{m}, y) \right) \bar{m}(dy) + \mathcal{F}(\bar{m}) + \lambda \end{aligned}$$

If we integrate against  $\mu$  we get from Lemma 3.3.5 that

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \left( \Delta \bar{m}(x) + \operatorname{div} \left( \frac{dp_1}{dm \otimes \mu}(x, \bar{m}) \bar{m}(x) \right) \right) \bar{\mathcal{V}}_\delta(\bar{m}) \mu(d\bar{m}) = 0.$$

Moreover, as  $(\mu, p_1)$  is a Mather measure we also have

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} H^* \left( y, \frac{dp_1}{dm \otimes \mu}(\bar{m}, y) \right) \bar{m}(dy) + \mathcal{F}(\bar{m}) \mu(d\bar{m}) = -\lambda.$$

Therefore,

$$\delta \int_{\mathcal{P}(\mathbb{T}^d)} \bar{\mathcal{V}}_\delta(\bar{m}) \mu(d\bar{m}) \leq 0.$$

□

**Definition 3.3.7.** *We say that  $(\mu, p_1) \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d)) \times \mathcal{M}(\mathcal{P}(\mathbb{T}^d) \times E, \mathbb{R}^d)$  belongs to  $\mathcal{M}_\gamma$  if  $(\mu, p_1)$  is smooth in the sense of Definition 3.3.4 and if it is the limit of  $\nu_\delta^{m_0}$  (in the sense of Proposition 3.3.1) for a certain  $m_0$  and along a subsequence  $\delta \rightarrow 0$ .*

Note that, according to Proposition 3.3.1,  $\mathcal{M}_\gamma$  is a subset of the set of Mather measures. Moreover, we will say with an abuse of terminology that a measure  $\nu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d) \times E)$  belongs to  $\mathcal{M}_\gamma$  is the couple  $(\mu, p_1)$  defined as in Proposition 3.3.1 does.

**Lemma 3.3.8.** *For any  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  with a  $C^\infty(\mathbb{T}^d)$  density, there exists  $\delta_n \rightarrow 0$  such that  $\nu_{\delta_n}^{m_0} \rightarrow \nu^{m_0}$  and  $\nu^{m_0} \in \mathcal{M}_\mathcal{V}$ .*

*Proof.* We know from Lemma 3.2.3 that, if  $(m_\delta, \alpha_\delta)$  is a minimizer of  $\bar{\mathcal{V}}_\delta$  then  $\alpha_\delta$  is uniformly bounded, with respect to  $\delta$ , in  $C^{1,1}([0, +\infty] \times \mathbb{T}^d)$ . Therefore,  $m_\delta$  solves the Fokker-Plank equation (3.4) with a smooth drift. If, moreover, the initial condition  $m_0$  is smooth, by standard parabolic estimates, we have that  $m_\delta$  is uniformly bounded in  $C^{1,2}([0, +\infty] \times \mathbb{T}^d)$ . Then, as  $\text{supp } \nu_\delta^{m_0} = \{(m_\delta(t), \alpha_\delta(t))\}_{t \geq 0}$ , there exists a constant  $C > 0$  such that for any  $(m, \alpha) \in \text{supp } \nu_\delta^{m_0}$ ,  $\|(m, \alpha)\|_{C^2(\mathbb{T}^d) \times C^1(\mathbb{T}^d, \mathbb{R}^d)} \leq C$ .

Let now  $\delta_n$  be a sequence such that  $\nu_{\delta_n}^{m_0} \rightarrow \nu^{m_0}$ . Given that  $\text{supp } \nu^{m_0} \subset \limsup_{\delta \rightarrow 0} \text{supp } \nu_\delta^{m_0}$ , the bounds that we have discussed ensure that the points of the support of  $\nu^{m_0}$  are smooth in the sense of Definition 3.3.4.  $\square$

Note that the proof of the above lemma has as byproduct that the set of smooth Mather measures (Definition 3.3.4) is not empty.

### 3.3.4 Conclusion

We define  $\mathcal{S}^-$  the set of subsolution  $\Psi$  of (3.6) such that for any  $(\mu, p_1) \in \mathcal{M}_\mathcal{V}$  we have

$$\int_{\mathcal{P}(\mathbb{T}^d)} \Psi(m) \mu(dm) \leq 0.$$

We set

$$\bar{\chi}(m) = \sup_{\chi \in \mathcal{S}^-} \chi(m).$$

To give sense to the terms that appear in the above relation we need first to prove the following lemma.

**Lemma 3.3.9.** *The family of function  $\mathcal{S}^-$  is not empty and uniformly bounded from above.*

*Proof.* Let  $\chi$  be a corrector function and  $C > 0$  be such that  $\chi - C < 0$ . If we set  $\Psi = \chi - C$ , then  $\Psi$  is a corrector function and for any measure  $(\mu, p_1) \in \mathcal{M}_\mathcal{V}$

$$\int_{\mathcal{P}(\mathbb{T}^d)} \Psi(m) \mu(dm) \leq 0.$$

Therefore,  $\mathcal{S}^-$  is not empty. Following the structure of [68, Theorem 1.5], one can easily prove that the set of subsolution of (3.6) is uniformly Lipschitz continuous. Then, if we fix a  $(\mu, p_1) \in \mathcal{M}_\mathcal{V}$  and a subsolution  $\chi \in \mathcal{S}^-$ , we have

$$\min_{\nu \in \mathcal{P}(\mathbb{T}^d)} \chi(\nu) = \int_{\mathcal{P}(\mathbb{T}^d)} \min_{\nu \in \mathcal{P}(\mathbb{T}^d)} \chi(\nu) \mu(dm) \leq \int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \mu(dm) \leq 0.$$

If we use the Lipschitz continuity of  $\chi$  we get that, for any  $m \in \mathcal{P}(\mathbb{T}^d)$ ,

$$\chi(m) \leq \min_{\nu \in \mathcal{P}(\mathbb{T}^d)} \chi(\nu) + K \text{diam}(\mathcal{P}(\mathbb{T}^d)) \leq K \text{diam}(\mathcal{P}(\mathbb{T}^d)),$$

which proves the claim.  $\square$

**Theorem 3.3.10.** *The function  $\bar{\mathcal{V}}_\delta$  uniformly converges to the corrector  $\bar{\chi}$  defined by*

$$\chi_0(m) = \sup_{\chi \in \mathcal{S}^-} \chi(m).$$

*Proof.* In Lemma 3.3.3 we proved that there exists a sequence  $\delta_n \rightarrow 0$  and a corrector function  $\chi_0$  such that  $\bar{\mathcal{V}}_{\delta_n}$  uniformly converges to  $\chi_0$ . We know from Proposition 3.3.6 that, for any  $\delta_n > 0$  and any smooth Mather measure  $(\mu, p_1)$ ,

$$\int_{\mathcal{P}(\mathbb{T}^d)} \bar{\mathcal{V}}_{\delta_n}(m) \mu(dm) \leq 0.$$

In particular the above relation holds true for any  $(\mu, p_1) \in \mathcal{M}_\nu$ . Letting  $\delta_n \rightarrow 0$ , we get

$$\int_{\mathcal{P}(\mathbb{T}^d)} \chi_0(m) \mu(dm) \leq 0, \quad \forall (\mu, p_1) \in \mathcal{M}_\nu,$$

which proves that  $\chi_0 \in \mathcal{S}^-$  and, consequently, that  $\chi_0 \leq \bar{\chi}$ .

For the other inequality we fix  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  and a sequence  $m_\varepsilon$  of smooth measures such that  $m_\varepsilon \rightarrow m_0$  when  $\varepsilon \rightarrow 0$ . From Lemma 3.3.2 we know that, for any subsolution  $\chi$  of (3.6), we have

$$\bar{\mathcal{V}}_{\delta_n}(m_\varepsilon) \geq \chi(m_\varepsilon) - \int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu_{\delta_n}^{m_\varepsilon}(dm). \quad (3.25)$$

As  $m_\varepsilon$  is smooth, we know from Lemma 3.3.8 that there exists a further subsequence  $\delta_{n_k}$  such that  $\nu_{\delta_{n_k}}^{m_\varepsilon} \rightarrow \nu^{m_\varepsilon}$  and  $\nu^{m_\varepsilon} \in \mathcal{M}_\nu$ . Therefore, if we let in (3.25)  $\delta_{n_k} \rightarrow 0$ , we get

$$\chi_0(m_\varepsilon) \geq \chi(m_\varepsilon) - \int_{\mathcal{P}(\mathbb{T}^d)} \chi(m) \nu^{m_\varepsilon}(dm).$$

If now we suppose that  $\chi \in \mathcal{S}^-$ , the above inequality becomes

$$\chi_0(m_\varepsilon) \geq \chi(m_\varepsilon).$$

As  $\chi$  and  $\chi_0$  are continuous we can let  $\varepsilon \rightarrow 0$  to finally find that

$$\chi_0(m_0) \geq \chi(m_0).$$

By the arbitrariness of  $\chi$  and  $m_0$  we deduce that

$$\chi_0 \geq \sup_{\chi \in \mathcal{S}^-} \chi = \bar{\chi}.$$

Note that  $\bar{\chi}$  is uniquely defined and depends neither on  $\delta_n$  nor on  $\delta_{n_k}$ . This implies that also  $\chi_0$  is uniquely defined and, therefore, the full convergence of  $\bar{\mathcal{V}}_\delta$ .  $\square$

# Bibliography

- [1] Yves Achdou, Francisco J Buera, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, *Partial differential equation models in macroeconomics*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372** (2014), no. 2028, 20130397.
- [2] Yves Achdou, Fabio Camilli, and Italo Capuzzo-Dolcetta, *Mean field games: convergence of a finite difference method*, SIAM Journal on Numerical Analysis **51** (2013), no. 5, 2585–2612.
- [3] Yves Achdou and Italo Capuzzo-Dolcetta, *Mean field games: Numerical methods*, SIAM Journal on Numerical Analysis **48** (2010), no. 3, 1136–1162.
- [4] Thierry Aubin, *Some nonlinear problems in riemannian geometry*, Springer Science & Business Media, 2013.
- [5] G. Barles and P. E. Souganidis, *Space-time periodic solutions and long-time behavior of solutions to quasi-linear parabolic equations*, SIAM Journal on Mathematical Analysis **32** (2001), no. 6, 1311–1323.
- [6] Jean-David Benamou and Yann Brenier, *A computational fluid mechanics solution to the monge-kantorovich mass transfer problem*, Numerische Mathematik **84** (2000), no. 3, 375–393.
- [7] Jean-David Benamou, Guillaume Carlier, and Filippo Santambrogio, *Variational mean field games*, Active Particles, Volume 1, Springer, 2017, pp. 141–171.
- [8] Charles Bertucci, *Optimal stopping in mean field games, an obstacle problem approach*, Journal de Mathématiques Pures et Appliquées **120** (2018), 165–194.
- [9] V. I. Bogachev, N. V. Krylov, and M. Röckner, *Elliptic and parabolic equations for measures*, Russian Mathematical Surveys **64** (2009), no. 6, 973.
- [10] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, *Fokker-planck-kolmogorov equations*, vol. 207, American Mathematical Soc., 2015.
- [11] Ariela Briani and Pierre Cardaliaguet, *Stable solutions in potential mean field game systems*, Nonlinear Differential Equations and Applications NoDEA **25** (2018), no. 1, 1.
- [12] Simone Cacace, Fabio Camilli, and Claudio Marchi, *A numerical method for mean field games on networks*, ESAIM: Mathematical Modelling and Numerical Analysis **51** (2017), no. 1, 63–88.

- [13] Filippo Cagnetti, Diogo Gomes, Hiroyoshi Mitake, and Hung V Tran, *A new method for large time behavior of degenerate viscous hamilton–jacobi equations with convex hamiltonians*, **32** (2015), no. 1, 183–200.
- [14] Fabio Camilli and Francisco Silva, *A semi-discrete in time approximation for a model 1st order-finite horizon mean field game problem*, *Networks and heterogeneous media* **7** (2011), no. 2, 263–277.
- [15] P. Cannarsa and C. Sinestrari, *Semiconcave functions, hamilton-jacobi equations, and optimal control*, vol. 58, Springer Science & Business Media, 2004.
- [16] Piermarco Cannarsa and Rossana Capuani, *Existence and uniqueness for mean field games with state constraints*, *PDE Models for Multi-Agent Phenomena*, Springer, 2018, pp. 49–71.
- [17] P. Cardaliaguet, *Notes on mean field games*, Tech. report, Technical report, 2010.
- [18] ———, *Long time average of first order mean field games and weak kam theory*, *Dynamic Games and Applications* **3** (2013), no. 4, 473–488.
- [19] P. Cardaliaguet, François D., J.-M. Lasry, and P.-L. Lions, *The master equation and the convergence problem in mean field games:(ams-201)*, Princeton University Press **381** (2019).
- [20] P. Cardaliaguet and P. J. Graber, *Mean field games systems of first order*, *ESAIM: Control, Optimisation and Calculus of Variations* **21** (2015), no. 3, 690–722.
- [21] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon, *Second order mean field games with degenerate diffusion and local coupling*, *Nonlinear Differential Equations and Applications NoDEA* **22** (2015), no. 5, 1287–1317.
- [22] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta, *Long time average of mean field games with a nonlocal coupling*, *SIAM Journal on Control and Optimization* **51** (2013), no. 5, 3558–3591.
- [23] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, A. Porretta, et al., *Long time average of mean field games.*, *NHM* **7** (2012), no. 2, 279–301.
- [24] Pierre Cardaliaguet, *Weak solutions for first order mean field games with local coupling*, *Analysis and geometry in control theory and its applications*, Springer, 2015, pp. 111–158.
- [25] Pierre Cardaliaguet and Charles-Albert Lehalle, *Mean field game of controls and an application to trade crowding*, *Mathematics and Financial Economics* (2017), 1–29.
- [26] Pierre Cardaliaguet and Marco Masoero, *Weak kam theory for potential mfg*, *Journal of Differential Equations*, in press.
- [27] Pierre Cardaliaguet, Alpár R Mészáros, and Filippo Santambrogio, *First order mean field games with density constraints: pressure equals price*, *SIAM Journal on Control and Optimization* **54** (2016), no. 5, 2672–2709.
- [28] Pierre Cardaliaguet and Alessio Porretta, *Long time behavior of the master equation in mean field game theory*, *Analysis & PDE* **12** (2019), no. 6, 1397–1453.
- [29] Elisabetta Carlini and Francisco José Silva Álvarez, *A semi-lagrangian scheme for a degenerate second order mean field game system*, arXiv preprint arXiv:1404.5932 (2014).

- [30] René Carmona and François Delarue, *Probabilistic analysis of mean-field games*, SIAM Journal on Control and Optimization **51** (2013), no. 4, 2705–2734.
- [31] ———, *Probabilistic theory of mean field games with applications i-ii*, Springer, 2018.
- [32] René Carmona, François Delarue, Daniel Lacker, et al., *Mean field games with common noise*, The Annals of Probability **44** (2016), no. 6, 3740–3803.
- [33] M. Cirant and L. Nurbekyan, *The variational structure and time-periodic solutions for mean-field games systems*, arXiv preprint arXiv:1804.08943 (2018).
- [34] Marco Cirant, *On the existence of oscillating solutions in non-monotone mean-field games*, Journal of Differential Equations (2019).
- [35] Andrea Davini, Albert Fathi, Renato Iturriaga, and Maxime Zavidovique, *Convergence of the solutions of the discounted hamilton-jacobi equation*, Inventiones mathematicae **206** (2016), no. 1, 29–55.
- [36] J. De Vries, *Elements of topological dynamics*, vol. 257, Springer Science & Business Media, 2013.
- [37] Boualem Djehiche, Alain Tcheukam, and Hamidou Tembine, *Mean-field-type games in engineering*, arXiv preprint arXiv:1605.03281 (2016).
- [38] Lawrence C Evans, *Weak kam theory and partial differential equations*, Calculus of variations and nonlinear partial differential equations, Springer, 2008, pp. 123–154.
- [39] A. Fathi, *Solutions kam faibles conjuguées et barrières de peierls*, Comptes Rendus de l’Académie des Sciences-Series I-Mathematics **325** (1997), no. 6, 649–652.
- [40] ———, *Théoreme kam faible et théorie de mather sur les systemes lagrangiens*, Comptes Rendus de l’Académie des Sciences-Series I-Mathematics **324** (1997), no. 9, 1043–1046.
- [41] ———, *Sur la convergence du semi-groupe de lax-oleinik*, Comptes Rendus de l’Académie des Sciences-Series I-Mathematics **327** (1998), no. 3, 267–270.
- [42] ———, *Weak kam theorem in lagrangian dynamics preliminary version number 10*, by CUP (2008).
- [43] Albert Fathi, *Weak kam theory: the connection between aubry-mather theory and viscosity solutions of the hamilton-jacobi equation*, Proceedings of the International Congress of Mathematicians, vol. 3, 2014, pp. 597–621.
- [44] Rita Ferreira and Diogo A Gomes, *On the convergence of finite state mean-field games through  $\gamma$ -convergence*, Journal of Mathematical Analysis and Applications **418** (2014), no. 1, 211–230.
- [45] Wilfrid Gangbo and Adrian Tudorascu, *Lagrangian dynamics on an infinite-dimensional torus; a weak kam theorem*, Advances in Mathematics **224** (2010), no. 1, 260–292.
- [46] ———, *Weak kam theory on the wasserstein torus with multidimensional underlying space*, Communications on Pure and Applied Mathematics **67** (2014), no. 3, 408–463.

- [47] Diogo Gomes and Levon Nurbekyan, *On the minimizers of calculus of variations problems in hilbert spaces*, Calculus of Variations and Partial Differential Equations **52** (2015), no. 1-2, 65–93.
- [48] ———, *An infinite-dimensional weak kam theory via random variables*, Discrete & Continuous Dynamical Systems-A **36** (2016), no. 11, 6167–6185.
- [49] Diogo Gomes and Marc Sedjro, *One-dimensional, forward-forward mean-field games with congestion*, Discrete and Continuous Dynamical Systems - Series S **11** (2017), 901.
- [50] Diogo A Gomes, Joana Mohr, and Rafael Rigao Souza, *Discrete time, finite state space mean field games*, Journal de mathématiques pures et appliquées **93** (2010), no. 3, 308–328.
- [51] Diogo A Gomes and Enrico Valdinoci, *Duality theory, representation formulas and uniqueness results for viscosity solutions of hamilton–jacobi equations*, Dynamics, Games and Science II, Springer, 2011, pp. 361–386.
- [52] Diogo A Gomes and Vardan K Voskanyan, *Extended deterministic mean-field games*, SIAM Journal on Control and Optimization **54** (2016), no. 2, 1030–1055.
- [53] Diogo Aguiar Gomes, *Generalized mather problem and selection principles for viscosity solutions and mather measures*, Advances in Calculus of Variations **1** (2008), no. 3, 291–307.
- [54] P Jameson Graber, Alpár R Mészáros, Francisco J Silva, and Daniela Tonon, *The planning problem in mean field games as regularized mass transport*, arXiv preprint arXiv:1811.02706 (2018).
- [55] Olivier Guéant, *Mean field games equations with quadratic hamiltonian: a specific approach*, Mathematical Models and Methods in Applied Sciences **22** (2012), no. 09, 1250022.
- [56] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions, *Mean field games and applications*, Paris-Princeton lectures on mathematical finance 2010, Springer, 2011, pp. 205–266.
- [57] M. Huang, R. P. Malhamé, P. E. Caines, et al., *Large population stochastic dynamic games: closed-loop mckean–vlasov systems and the nash certainty equivalence principle*, Communications in Information & Systems **6** (2006), no. 3, 221–252.
- [58] Daniel Lacker, *A general characterization of the mean field limit for stochastic differential games*, Probability Theory and Related Fields **165** (2016), no. 3-4, 581–648.
- [59] ———, *Limit theory for controlled mckean–vlasov dynamics*, SIAM Journal on Control and Optimization **55** (2017), no. 3, 1641–1672.
- [60] O. A. Ladyzhenskaia, V. A. Solonnikov, and N. N Ural'tseva, *Linear and quasi-linear equations of parabolic type*, vol. 23, American Mathematical Soc., 1988.
- [61] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. i–le cas stationnaire*, Comptes Rendus Mathématique **343** (2006), no. 9, 619–625.
- [62] ———, *Jeux à champ moyen. ii–horizon fini et contrôle optimal*, Comptes Rendus Mathématique **343** (2006), no. 10, 679–684.

- [63] Hugo Lavenant and Filippo Santambrogio, *New estimates on the regularity of the pressure in density-constrained mean field games*, 2018.
- [64] P.-L. Lions, *Cours au collège de france*, [www.college-de-france.fr](http://www.college-de-france.fr).
- [65] ———, *Mean-field games. cours au collège de france (2007–2008)*.
- [66] P.-L. Lions, G. Papanicolaou, and S. R. S. Varadhan, *Homogenization of hamilton-jacobi equations*, Unpublished preprint (1987).
- [67] Pierre-Louis Lions and Panagiotis E Souganidis, *Homogenization of degenerate second-order pde in periodic and almost periodic environments and applications*, *Annales de l'IHP Analyse non linéaire*, vol. 22, 2005, pp. 667–677.
- [68] Marco Masoero, *On the long time convergence of potential mfg*, *Nonlinear Differential Equations and Applications NoDEA* **26** (2019), no. 2, 15.
- [69] John N. Mather, *Action minimizing invariant measures for positive definite lagrangian systems*, *Mathematische Zeitschrift* **207** (1991), no. 1, 169–207.
- [70] A. R. Mészáros and F. J. Silva, *On the variational formulation of some stationary second-order mean field games systems*, *SIAM Journal on Mathematical Analysis* **50** (2018), no. 1, 1255–1277.
- [71] Chenchen Mou and Jianfeng Zhang, *Weak solutions of mean field game master equations*, arXiv preprint [arXiv:1903.09907](https://arxiv.org/abs/1903.09907) (2019).
- [72] Adam Prosimski and Filippo Santambrogio, *Global-in-time regularity via duality for congestion-penalized mean field games*, 2017.







## RÉSUMÉ

---

Cette thèse porte sur l'étude du comportement en temps long des jeux à champ moyen (MFG) potentiels, indépendamment de la convexité du problème de minimisation associé. Pour le système hamiltonien de dimension finie, des problèmes de même nature ont été traités par la théorie KAM faible. Nous transposons de nombreux résultats de cette théorie dans le contexte des jeux à champ moyen potentiels. Tout d'abord, nous caractérisons par approximation ergodique la valeur limite associée aux systèmes MFG à horizon fini. Nous fournissons des exemples explicites dans lesquels cette valeur est strictement supérieure au niveau d'énergie des solutions stationnaires du système MFG ergodique. Cela implique que les trajectoires optimales des systèmes MFG à horizon fini ne peuvent pas converger vers des configurations stationnaires. Ensuite, nous prouvons la convergence du problème de minimisation associé à MFG à horizon fini vers une solution de l'équation Hamilton-Jacobi critique dans l'espace de mesures de probabilité. De plus, nous montrons une limite de champ moyen pour la constante ergodique associée à l'équation Hamilton-Jacobi de dimension finie correspondante. Dans la dernière partie, nous caractérisons la limite du problème de minimisation à horizon infini que nous avons utilisé pour l'approximation ergodique dans la première partie du manuscrit.

## MOTS CLÉS

---

Jeux à champ moyen, Contrôl optimal, Théorie KAM faible

## ABSTRACT

---

The purpose of this thesis is to shed some light on the long time behavior of potential Mean Field Games (MFG), regardless of the convexity of the minimization problem associated. For finite dimensional Hamiltonian systems, problems of the same nature have been addressed through the so-called weak KAM theory. We transpose many results of this theory in the infinite dimensional context of potential MFG. First, we characterize through an ergodic approximation the limit value associated to time dependent MFG systems. We provide explicit examples where this value is strictly greater than the energy level of stationary solutions of the ergodic MFG system. This implies that optimal trajectories of time dependent MFG systems cannot converge to stationary configurations. Then, we prove the convergence of the minimization problem associated to time dependent MFGs to a solution of the critical Hamilton-Jacobi equation in the space of probability measures. In addition, we show a mean field limit for the ergodic constant associated with the corresponding finite dimensional Hamilton-Jacobi equation. In the last part we characterize the limit of the infinite horizon discounted minimization problem that we use for the ergodic approximation in the first part of the manuscript.

## KEYWORDS

---

Mean Field Games, Optimal control, Weak KAM Theory

