UNIVERSITÀ DI ROMA TOR VERGATA
Dipartimento di Matematica Pura e Applicata

UNIVERSITÉ PARIS DAUPHINE
CEREMADE : Centre de Recherche en Mathématiques de la Décision

DOCTORAL SCHOOL IN MATHEMATICS

PH.D. THESIS

Some Advances in Mean Field Games Theory

Supervisors

Prof. Alessio Porretta
Prof. Pierre Cardaliaguet

Coordinator

Prof. Andrea Braides

Author

Michele Ricciardi

XXXII Cycle - Academic Year 2019/2020

May 26th, 2020

Alla mia famiglia

Resumé

La théorie des Jeux à Champ Moyen (MFG) a été développé par J.-M. Lasry et P.-L. Lions en 2006. Cette théorie veut décrire le comportement d'un système de N jouers, qui choisissent un contrôle et agissent pour minimiser une fonction de coût. La dynamique de chaque joueur est modélisé avec une équation différentielle stochastique.

Le système associé aux équilibres de Nash, sous certain hypothèses et quand N tend vers l'infini, converge à la solution du système MFG. Il s'agit d'une équation de Hamilton-Jacobi, pour la fonction valeur du système, couplé avec une equation de Fokker-Planck pour la densité du processus de chaque joueur.

Il y a une vaste littérature en ce qui concerne les Jeux à Champ Moyen, et beaucoup d'aspects ont été étudiés, comme par exemple existence, régularité et unicité des solutions, comportement en temps long etc., en utilisant soit des méthodes analytiques soit probabilistes.

Cependant, la plupart de la littérature considère seulement le cas que la dynamique des joueurs est confinée dans l'espace $\mathbb{R}^d$, ou, surtout dans la littérature analytique, dans le tore $\mathbb{T}^d$ (solutions périodiques). Plus, dans la plupart des cas les joueurs contrôlent seulement le drift de l'équation stochastique.

Mais dans beaucoup d'applications c'est très important travailler avec un processus qui reste dans un certain domaine d'existence.

Cette condition peut être obtenue, par exemple, en prescrivant des conditions de Neumann sur le système MFG, lesquelles correspondent à une reflexion dans l'équation différentielle stochastique du processus.

Alternativement, on peut confiner la dynamique dans un domaine borné en choisissant le contrôle, ou en prenant les termes de drift et de diffusion, afin de satisfaire la restriction requise. Dans ce cas on parle des MFG avec condition d'invariance ou contraintes sur l'état.

Dans ma thése je serai focusé sur les deux aspects.

Dans le premier chapitre j'étudierai le problème MFG avec contraintes sur l'état, en obtenant existence, unicité et résultats de régularité. Dans le deuxième chapitre j'étudierai le problème de convergence avec des conditions de Neumann sur la frontière du domaine. Finalement, dans le troisième chapitre je retournerai au système MFG, en étudiant un modèle avec un contrôle sur la diffusion.

Table des matières

Chapitre 1

Introduction

Game theory is a branch of mathematics which aims to study the behaviour of a group of players, called *agents*. The number of players will be denoted by N.

The dynamic of each agent depends on the interactions with the other players, in a noncooperative framework : i.e., the strategies of each player are made in order to pursue his own interest.

In this context, a generic agent chooses his own strategy, i.e. a suitable control, and set up his dynamic, typically modeled by a Stochastic Differential Equation. The control is chosen in order to minimize a certain cost functional, which also depends on the strategies and the dynamics of the other players.

A fundamental tool here is the notion of *Nash equilibrium*, introduced by Nash in [86]. A certain choice of strategies is called a Nash equilibrium if each agent is playing the optimal strategy in relation to the others. In other words, each player is not interested to be the only one who changes strategy.

In this framework, a crucial role is played by the *value function*. For each agent, the value function is defined as the cost functional of the player computed at the Nash equilibrium.

We will be more specific about the stochastic formulation and the value function later.

We will focus on *differential games*, i.e. games in continuous time and state space, introduced for the first time by Isaacs in [67] and, at the same time, by Pontryagin in [88]. In particular, Isaacs computed formally the link between the value function in these differential games and the Hamilton-Jacobi equations. In zero-sum differential games, this link became more rigorous thanks to the work of Crandall, Ishii and Lions on viscosity solutions, see

[45] and [68].

In non-zero sum differential games, the work becomes harder. In this case, the value functions v_i^N, $i = 1, \dots, N$, solve a coupled system of Hamilton-Jacobi equations, called *Nash system*.

We give an example about the structure of the Nash system in the case of uncontrolled diffusion. If the control of the agents acts only on the drift of the SDE and not on the diffusion, we can write the dynamic of the generic player i in this way :

$$\begin{cases} dX_t^i = b(X_t^i, \alpha_t^i)\, dt + \sqrt{2}\sigma(X_t^i) dB_t^i\,, \\ X_{t_0}^i = x_0^i\,. \end{cases} \tag{1.0.1}$$

Here, $x_0^i \in \mathbb{R}^d$, α_t^i is the control, chosen from a certain set A, b and σ are respectively the *drift* term and the *diffusion* matrix and $(B_t)^i$, $1 \le i \le N$, are independent d-dimensional Brownian motions.

From now on, we will use the notation $\mathbf{v}$ to indicate a vector of $\mathbb{R}^{Nd}$ defined by $\mathbf{v} = (v^1, \dots, v^N)$, where v^i is an already defined vector of $\mathbb{R}^d$.

Assume that the cost for the player i is given by the following functional :

$$J_i^N(t_0, \boldsymbol{x}_0, \boldsymbol{\alpha}_\cdot) = \mathbb{E}\left[\int_{t_0}^T \left(L(s, X_s^i, \alpha_s^i) + F_i^N(s, \boldsymbol{X}_s)\right) ds + G_i^N(\boldsymbol{X}_T)\right]\,,$$

where F_i^N and G_i^N are the *cost functions* of the player i and L is the *Langrangian* cost for the control, which we assume in this example not depending on i.

With these notations, a control $\boldsymbol{\alpha}_\cdot^*$ provides a Nash equilibrium if, for all controls $\boldsymbol{\alpha}_\cdot$ and for all i we have

$$J_i^N(t_0, \boldsymbol{x}_0, \boldsymbol{\alpha}_\cdot^*) \le J_i^N(t_0, \boldsymbol{x}_0, \alpha_i, (\alpha_j^*)_{j\neq i})\,,$$

i.e., each player chooses his optimal strategy, if we “freeze” the other players’ strategies.

Hence, the value function for the generic player i corresponds to the cost functional evalutated at the optimal control :

$$v_i^N(t_0, \boldsymbol{x}_0) = J_i^N(t_0, \boldsymbol{x}_0, \boldsymbol{\alpha}^*)\,.$$

Using Ito’s formula and the dynamic programming principle, one can prove that v_i^N solves the so-called Nash system :

$$
\begin{cases}
-\partial_t v_i^N(t,\boldsymbol{x}) - \sum_j \mathrm{tr}(a(x_j) D^2_{x_j x_j} v_i^N(t,\boldsymbol{x})) & + H(x_i, D_{x_i} v_i^N(t,\boldsymbol{x})) \\
\qquad + \sum_{j\neq i} H_p(x_j, D_{x_j} v_j^N(\boldsymbol{x}))\cdot D_{x_j} v_i^N(t,\boldsymbol{x}) & = F_i^N(\boldsymbol{x})\,, \\
v_i^N(T,\boldsymbol{x}) = G_i^N(\boldsymbol{x})\,,
\end{cases}
\tag{1.0.2}
$$

for $(t,\boldsymbol{x}) \in [0,T]\times\mathbb{R}^{Nd}$. Here H is the *Hamiltonian* of the system, i.e. the Fenchel conjugate of the Lagrangian, $a = \sigma\sigma^*$ and $H_p(x,p)$ denotes $\frac{\partial H(x,p)}{\partial p}$, for $p \in \mathbb{R}^N$.

Existence of solutions for this system is well known under some hypotheses of regularity and growth of the coefficients, see [17] and [74].

However, the structure of the N-players game becomes really intricate when $N \gg 1$, and in that case we are naturally interested in an asymptotic behaviour of (1.0.2) as $N \to \infty$, in order to simplify the configuration of the Nash system.

The system born to describe Nash equilibria in differential games with infinitely many (small and undistinguishable) agents is called ***Mean Field Games system***. It was introduced by J.M. Lasry and P.L. Lions ([75], [76], [77]), using tools from mean-field theories. A similar notion of Nash equilibria was also developed in the same years by P. Caines, M. Huang and R. Malhamé [63].

The macroscopic description used in mean field game theory leads to study coupled systems of PDEs, where the Hamilton-Jacobi-Bellman equation satisfied by the single agent's value function u is coupled with the Kolmogorov Fokker-Planck equation satisfied by the distribution law of the population m. The simplest form of this system is the following

$$
\begin{cases}
-\partial_t u - \mathrm{tr}(a(x)D^2 u) + H(x,Du) = F(x,m)\,, \\
\partial_t m - \sum_{i,j} \partial^2_{ij}(a_{ij}(x)m) - \mathrm{div}(mH_p(x,Du)) = 0\,, \\
m(0) = m_0 \qquad\qquad u(T) = G(x,m(T))\,,
\end{cases}
\tag{1.0.3}
$$

where $\partial^2_{ij}(\cdot) = \frac{\partial^2(\cdot)}{\partial x_i \partial x_j}$ denotes a second order partial differentiation and $\mathrm{div}(\cdot)$ is the usual divergence operator.

Unfortunately, there is no hope to obtain a convergence of the Nash system in absence of a symmetrical structure of (1.0.2). In other words, the agents and their dynamics have to be symmetric and *indistinguishable*. So, we suppose that the cost functions F_i^N and G_i^N

are of this form :

$$F_i^N(\boldsymbol{x}) = F(x_i, m_{\boldsymbol{x}}^{N,i}), \qquad G_i^N(t,\boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i}),$$

where

$$m_{\boldsymbol{x}}^{N,i} = \frac{1}{N-1}\sum_{j\neq i}\delta_{x_j}, \quad \text{where } \delta_x \text{ is the Dirac function at } x\,.$$

If we want to describe, at least heuristically, the structure of this limit problem when $N \to \infty$, we find a differential game with infinitely many players, where the dynamic of a generic player is driven by a stochastic differential equation of this type :

$$\begin{cases} dX_t = b(X_t,\alpha_t)\,dt + \sqrt{2}\sigma(X_t)dB_t\,, \\ X_{t_0} = x_0\,, \end{cases}$$

and each player chooses his own strategy in order to minimize

$$J(t_0,x_0,\alpha.) = \mathbb{E}\left[\int_{t_0}^T \left(L(s,X_s,\alpha_s) + F(X_s,m(s))\right)ds + G(X_T,m(T))\right],$$

where $m(\cdot)$ is the density of the population, obtained by the convergence of $m_{\boldsymbol{x}}^{N,i}$.

Since the early works of Lasry and Lions, the study of Mean Field Games Theory with infinitely many players is, in the literature, roughly divided into three main areas :

(i) Works on the Mean Field Games (MFG) system (1.0.3) : study of existence, uniqueness and qualitative properties of solutions, in all possible frameworks.

(ii) Works on the convergence problem of the system with N players towards the Mean Field Games system. In other words, show that (1.0.3) is a good approximation of (1.0.2).

(iii) Works that directly study the stochastic control problem and stochastic trajectories with probabilistic approach.

There is by now an extensive literature concerning mean field game systems of this kind (1.0.3), and many fundamental issues have been discussed so far such as existence or non-existence, regularity and uniqueness of solutions, long time behavior etc., using both an analytic and probabilistic approach. For example, in the analytic literature existence and uniqueness of smooth solutions was proved if F and G are non-decreasing operators, see [77] and [78]. A very general result of existence and uniqueness was subsequently proved by

Porretta in [89]. Other results, concerning also applications and numerical methods, can be found in [18], [60] and [61].

In the probabilistic literature, some pioneering results about the well-posedness of the MFG system were given in [16], [34], [64], [65], [66] and [69].

So far, most of the literature considers the case that the state variable x belongs to the flat torus (i.e. periodic solutions), or, especially in the probabilistic literature, in the whole space $\mathbb{R}^d$. But in many applications it is useful to work with a process that remains in a certain domain of existence.

This can be obtained, for instance, by prescribing Neumann boundary conditions at the equation (1.0.3), which correspond to a reflection term in the stochastic differential equation of the process. Some results of existence and uniqueness for MFG with Neumann (and Dirichlet) boundary conditions can be found in [89] and [43].

Another way to confine the dinamics into a bounded domain is to choose the control, or to build the drift-diffusion term, in order to satisfy the required restriction. In this case we talk about MFG with *invariance condition* or state constraint.

In this thesis I will be focused on both aspects.

In the first chapter I will study a MFG problem with state invariance, obtaining existence, uniqueness and regularity results. In the second chapter I will deal with the convergence problem in a framework of Neumann boundary conditions. We will be more specific about the convergence problem later. Finally, in the third chapter I come back to the Mean Field Games system, analyzing a model where also the diffusion is controlled.

1.1 Mean Field Games System under State Invariance : Chapter 1

As already sad, in many applied models boundary conditions turn out to be a crucial issue. A significant case occurs when the dynamical state needs to remain inside some given domain of existence, say if some natural restriction needs to be preserved. For instance, in many models appearing in economics, a scalar state variable needs to remain above or below given thresholds (e.g. if x denotes a stock quantity, or the reserve of a fossil fuel, or a wealth level, see models described in [2], [62]).

There are two typical ways in which the proposed models handle this kind of situation : either one considers the *state constraint* control problem, in which case one uses the control in order to satisfy the required restriction, or alternatively the drift-diffusion terms are built in the model so that the state does not leave the domain, regardless of the control. This latter situation is what we are going to study in Chapter 1 of my thesis. Namely, we assume that the state variable x belongs to a bounded domain $\Omega \subset \mathbb{R}^d$ and we will assume structure conditions, on the diffusion and the Hamiltonian terms, which imply that the domain Ω is an invariant set for the underlying controlled dynamics, and this invariance occurs *for any choice of the control*. In the control community, sometimes this property is referred to as the *invariance of the state space.*

Let us stress that considering the domain to be invariant for all controls is different from considering the state constraint control problem ; in very rough words, one can say that the viability of the state space plays like a regularizing condition of the underlying dynamics, whereas the state constraint problem leads to formation of singularities at the boundary, because of the forced action of the control.

In the case of uncontrolled SDEs

$$dX_t = b(t, X_t)\, dt + \sqrt{2}\, \sigma(X_t)\, dW_t\,, \quad X_0 = x \in \Omega \tag{1.1.1}$$

the conditions on the coefficients b and σ which let Ω be an invariant set are extensively discussed in the literature, at least in the case that σ and b are globally Lipschitz. We refer the reader to [27], [46], [47] and the literature therein. The case that the diffusion is controlled, so $b = b(t, x, \alpha)$, $\sigma = \sigma(t, x, \alpha)$ for α in a set of controls, was recently discussed in [9], [10] and [39] in terms of (viscosity solutions) of the associated Bellman operators. In our study, we let aside by now the possibility that the diffusion part is controlled ; on the other hand, we aim at giving general conditions on the diffusion matrix and drift terms which may apply to both degenerate and non degenerate operators, to possibly unbounded drifts and possibly non Lipschitz matrix σ. For the dynamics (1.1.1) in a C^2 domain Ω, we formulate this invariance condition by requiring that the following inequality holds in a neighborhood of the boundary :

$$\mathrm{tr}(a(x)D^2 d(x)) + b(t, x) \cdot Dd(x) \geq \frac{a(x)Dd(x) \cdot Dd(x)}{d(x)} - C\, d(x) \tag{1.1.2}$$

for some constant $C > 0$, where $d(x)$ is the distance function to the boundary, $a(x) = (\sigma\, \sigma^*)(x)$ is the diffusion matrix and $\mathrm{tr}(\cdot)$ is the trace operator. This condition reduces to the well known necessary and sufficient condition ([27]) if σ is Lipschitz continuous and $\sigma^*(x)Dd(x) = 0$ on the boundary. However, it also includes more general cases, like σ being only 1/2-Hölder continuous (to this respect, it generalizes the condition used in [9] for Hölder coefficients) and, last but not least, it also applies to the case of non degeneracy ($a(x)$ coercive up to the boundary) if the drift $b(x)$ is allowed to be unbounded (as in [79]).

In case of Bellman operators, say if the dynamics is controlled, the viability of the state space is ensured provided the same condition as (1.1.2) is required to hold for all controls. More specific examples will be given in the next Chapter. We point out that this kind of condition is also related to the notion of characteristic points of the boundary, namely to the question whether boundary conditions should be prescribed or not for the corresponding Bellman operator, see e.g. [12], [13] and [55] for the linear case.

We conclude this section giving a summary of the results obtained :

(i) First we obtain existence and uniqueness of bounded solutions to the Hamilton-Jacobi equation with bounded terminal conditions, with suitable hypotheses on the Hamiltonian H ;

(ii) Then we prove existence and uniqueness of weak solutions in $C([0,T];L^1(\Omega))$ for the Fokker-Planck equation, with initial condition in $L^1(\Omega)$. We will be more precise about these definitions of solutions in the next Chapter.

(iii) Collecting these results we are able to prove, under classical monotonicity assumptions on the couplings F and G, existence and uniqueness of solutions for the Mean Field Games system with invariance conditions.

(iv) With further hypotheses on the coefficients and the data of the system, we can improve the regularity of the solutions. In particular, the value function u is globally Lipschitz continuous and semi-concave in Ω, while the density function m is globally bounded.

We point out that these results are not restricted to smooth domains, but at the end of the chapter a generalization on domains with corners, e.g. rectangles, will be given.

1.2 The Master Equation and the Convergence Problem : Chapter 2

Once proved the results on the Mean Field Games problem, one naturally asks if this system can be a good approximation of the N-players system.

In this context, two kind of results can be shown :

(i) The optimal strategies in the Mean Field Games system provide approximated Nash equilibria (called ε-Nash equilibria) in the N-player game.

(ii) A Nash equilibrium in the N-player game converges, when $N \to +\infty$, towards an optimal strategy in the Mean Field Games.

The first question has been widely studied, using specific tools of the Mean Field theory. See, for instance, [34], [64], [69].

Conversely, many difficulties arise in the second question, due to the lack of compactness properties of the problem.

So far it has become clear that the Mean Field Games system cannot be sufficient to take into account the complexity of the problem with N players.

In order to overcome this problem, Lasry and Lions in [80] introduced a new infinite dimensional equation, the so-called ***Master Equation***, which summarizes the whole Mean Field Games system in a unique equation and is clearly connected with the Nash system.

The Master Equation is defined from its trajectories, which are solutions of the Mean Field Games system. To be more precise, we consider the solution (u, m) of the system (1.0.3) with initial condition $m(t_0) = m_0$ and we define the function

$$U : [0,T] \times \Omega \times \mathcal{P}(\Omega) \to \mathbb{R}\,, \qquad U(t_0, x, m_0) = u(t_0, x)\,, \tag{1.2.1}$$

where $\Omega \subseteq \mathbb{R}^d$ and $\mathcal{P}(\Omega)$ is the set of Borel probability measures on Ω.

If we compute, at least formally, the equation satisfied by U, we obtain a Hamilton-Jacobi equation in the space of measures :

$$\begin{cases} -\partial_t U(t,x,m) - \mathrm{tr}\left(a(x) D_x^2 U(t,x,m)\right) + H\left(x, D_x U(t,x,m)\right) \\ -\displaystyle\int_\Omega \mathrm{tr}\left(a(y) D_y D_m U(t,x,m,y)\right) dm(y) \\ +\displaystyle\int_\Omega D_m U(t,x,m,y) \cdot H_p(y, D_x U(t,y,m)) dm(y) = F(x,m)\,, \\ U(T,x,m) = G(x,m)\,. \end{cases} \tag{1.2.2}$$

Here, $D_m U$ is a suitable derivative of U with respect to the measure m. We will deal about this derivation in Chapter 2. This definition, however, is strictly related to the one given by Ambrosio, Gigli and Savaré in [4] and by Lions in [80].

The Master Equation plays a crucial role in order to prove the convergence of the N-player Nash equilibria, and his relevance was recognized in different papers. For example, in [19] and [20] Bensoussan, Frehse and Yam reformulated this equation as a PDE set on an L^2 space, and in [36] Carmona and Delarue interpreted it as a decoupling field of forward-backward stochastic differential equations in infinite dimension.

Once defined the Master Equation, there are two main steps which must be handled :

(i) Prove the well-posedness of the Master Equation : existence, uniqueness and regularity of solutions.

(ii) Prove that any solution of the Nash system (1.0.2) converge towards a solution of the Master Equation (1.2.2).

Some preliminary results about the first problem were given by Lions in [80] and a first exhaustive result of existence and uniqueness was proved, with a probabilistic approach, by Chassagneux, Crisan and Delarue in [42].

But the most important work in this direction is given by Cardaliaguet, Delarue, Lasry and Lions in [32], who give a complete proof for the well-posedness of (1.2.2) and for the convergence result.

These results are obtained in two different contexts : the first case is the so-called *First order Master Equation*, when the control of the generic player has the form (1.1.1) and the Master Equation is (1.2.2), and the *Second order Master Equation*, or *Master Equation with common noise*. In this case, the dynamic (1.1.1) has also an additional Brownian term dW_t, not depending on i (which justifies the adjective *common*). This leads to a different and more difficult type of Master Equation, with some additional terms depending also on the second derivative $D_{mm}U$. It is relevant to say that Mean Field Games with common noise were already studied by Carmona, Delarue and Lacker in [38].

Another convergence result using the Master Equation with common noise was given by Delarue, Lacker and Ramanan in [48], using large deviations results. Finally, a convergence result of Nash equilibria without using the Master Equation was given in [72].

Anyway, all these results were given in the periodic case, i.e.$\Omega = \mathbb{T}^d$, or possibly in the whole space $\mathbb{R}^N$. But, as already pointed out, in many economic and financial applications it is necessary to require some boundary conditions, which correspond to reflecting or absorbing processes for the N-player game.

In the second chapter of my thesis I study the well-posedness of the Master Equation and the convergence problem in case of reflecting processes, or equivalently Neumann boundary conditions, for a bounded domain Ω.

This part follows the main ideas of [32] : the function U is defined as in (1.2.1) and some estimates like global bounds and global Lipschitz regularity are proved. The main issue in order to obtain that U solves (1.2.2) is to prove the $\mathcal{C}^1$ character of U with respect to m. This proof passes through some estimations on a linearized system of the Mean Field Games one, (1.0.3).

However, these estimates requires strong regularities of U, and so of the Mean Field Games system, in the space and in the measure variable.

Concerning the measure variable, a suitable distance between measures has to be defined. This is called the *Wasserstein distance* and its definition will be given in the apposite Chapter.

The space regularity, besides, is obtained in [32] by differentiating the equation with respect to x.

But in the Neumann case, and in general in any boundary conditions case, these methods obviously cannot be applied, and these bounds are obtained using some space-time estimates, which require more effort to gain the same regularity.

Moreover, regularity estimates for Neumann parabolic equation require compatibility conditions between the final data and the value at the boundary (except for Hölder estimates).

Unfortunately, these compatibility conditions will be not always guaranteed in this context, especially in some Hamilton-Jacobi equations arising in the study of linearized systems, (see Proposition **??**).

This forces us to generalize the estimates obtained in [32], by defining generalized Wasserstein distances to take care of the boundary reflection and sometimes avoiding compatibility conditions, which allows us to obtain at least Hölder estimates.

Once proved existence and uniqueness of solutions, we are able to prove, following the same lines of [32], the convergence of Nash equilibria. We only need to pay attention about the stochastic process, which contains also a reflection term at the boundary.

The convergence is obtained defining the auxiliary functions

$$u_i^N(t,\boldsymbol{x}) = U(t,x_i,m_{\boldsymbol{x}}^{N,i})$$

and the stochastic processes related to this functions. Then, using a probabilistic approach, we can prove that $|u_i^N - v_i^N| \to 0$ in two different norms. See Theorem 3.7.7 for further details.

1.3 Mean Field Games with Controlled Diffusion : Chapter 3

Finally, we come back to the Mean Field Games problem, analyzing another kind of system.

In the model example we gave previously, the control α acts only on the drift. Some results with controlled diffusion are available, see for instance the elliptic case studied in [53], or the probabilistic approach developed in [15].

But, as far as we know, Mean Field Games system were mostly studied in case of uncontrolled diffusion. This leads to the study of a system of linear or possibly quasilinear PDE.

However, in many applied models it is interesting to study a framework where the agents can play their own control also on the diffusion term. See, for example, the financial articles of Avellaneda et al., [5], [6], [7].

This leads to the study of a fully nonlinear Hamilton-Jacobi equation, called in the model case *parabolic Bellman equation*. This equation was widely studied in the literature, both in elliptic and parabolic settings. See for instance [11], [44], [70], [71], [81], [97].

In this chapter, we study a particular case of nonlinear Mean Field Games, namely :

$$\begin{cases} \partial_t u + H^1(t,x,\nabla u) + H^2(t,x,\Delta u) + F(t,x,m) = 0\,, \\ \partial_t m - \Delta(mH^2_q(t,x,\Delta u)) + \mathrm{div}(mH^1_p(t,x,\nabla u)) = 0\,, \\ u(T,x) = G(x,m(T))\,, \qquad m(0) = m_0\,. \end{cases}$$

This system comes from a dynamic for the generic player of this type :

$$\begin{cases} dX_s = \alpha_s ds + \sigma_s dB_s\,, \\ X_t = x_0\,, \end{cases} \tag{1.3.1}$$

where two bounded different controls α and σ act on the drift and on the diffusion term. In order to ensure a strong ellipticity condition on the diffusion, we require σ bounded from above and below by two strictly positive constants.

Once defined a suitable definition of viscosity solution for the first equation (no weak solutions can be defined if there is nonlinearity in the second order term), existence and uniqueness of solutions for that one is immediately obtained.

But, since the diffusion of the second equation depends on Δu, we need to improve the regularity of u in order to study the FP equation.

These further regularities are widely studied in the elliptic case, see [24], [25], and in particular cases of fully nonlinear parabolic equations, see [85].

In our case, we start proving a Lipschitz bound for u, following the same ideas of [91]. Then we improve the regularity of u, adapting in a parabolic setting the results of [44], and proving a semiconcave bound for the value function.

Then, using a regularity result obtained by Krylov, see [70] and [71], we are able to prove that u is a classical solution of the system, at least in a regular case. This allows us to linearize the equation and apply the classical regularity results we need.

Finally, we can classically prove existence and uniqueness for m, since the equation of m is quasilinear, and then existence and uniqueness for Mean Field Games system using a standard fixed point argument.

Chapitre 2

Mean Field Games under Invariance Conditions for the State Space

2.1 Introduction

In this chapter, as already said, we investigate mean field game systems under invariance conditions for the state space, otherwise called *viability conditions* for the controlled dynamics. First we analyze separately the Hamilton-Jacobi and the Fokker-Planck equations, showing how the invariance condition on the underlying dynamics yields the existence and uniqueness, respectively in L^∞ and in L^1. Then we apply this analysis to mean field games. We investigate further the regularity of solutions proving, under some extra conditions, that the value function is (globally) Lipschitz and semiconcave. This latter regularity eventually leads the distribution density to be bounded, under suitable conditions. The results are not restricted to smooth domains.

The macroscopic description used in mean field game theory leads to study coupled systems of PDEs of the form (1.0.3), where the Hamilton-Jacobi-Bellman equation satisfied by the single agent's value function is coupled with the Kolmogorov Fokker-Planck equation satisfied by the distribution law of the population. We allow in this case a dependence on t

for the Hamiltonian H and the cost function F. The system assumes the following form :

$$\begin{cases} -\partial_t u - \sum_{i,j} a_{ij}(x)\partial^2_{ij} u + H(t,x,Du) = F(t,x,m)\,, & (t,x) \in (0,T)\times\Omega \\ \partial_t m - \sum_{i,j} \partial^2_{ij}(a_{ij}(x)m) - \operatorname{div}(mH_p(t,x,Du)) = 0\,, & (t,x) \in (0,T)\times\Omega \\ m(0) = m_0 \qquad u(T) = G(x, m(T)) \end{cases} \tag{2.1.1}$$

Here $\Omega \subseteq \mathbb{R}^N$ is an open and bounded set and $x \in \Omega$ represents the dynamical state of the generic agent, $m(t)$ is the distribution law of the agents at time t (and $m(t,x)$ denotes its density, if $m(t) \in L^1$), $F(t,x,m)$, $G(x,m(T))$ are respectively a running cost and a final pay-off and $H(t,x,Du)$ is the Hamiltonian function associated to the cost of dynamic control of the individuals. More details on the interpretation of solutions in terms of stochastic control will be given later.

We said in the introduction that for a stochastic dynamic (1.1.1), the invariance condition assumes the form (1.1.2). But for a general PDE approach to the Hamilton-Jacobi equation

$$-\partial_t u - \sum_{i,j} a_{ij}(x)\partial^2_{ij} u + H(t,x,Du) = 0\,, \qquad (t,x) \in (0,T)\times\Omega \tag{2.1.2}$$

we replace the invariance condition on the dynamics with a structure condition formulated directly on the Hamiltonian function. Namely, in the same spirit as above, we require that the diffusion matrix $a(x)$ and the Hamiltonian $H(t,x,p)$ satisfy the inequality

$$\operatorname{Tr}(a(x)D^2 d(x)) - H_p(t,x,p)\cdot Dd(x) \geq \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)} - C\, d(x) \qquad \forall p \in \mathbb{R}^N \tag{2.1.3}$$

for some constant $C > 0$, for x in a neighborhood of the boundary and $t \in (0,T)$.

As it is intrinsic to mean field games, we are going to study not only the properties of the HJB equation under condition (2.1.3) but also the properties of the Kolmogorov equation which appear, roughly speaking, in a dual form. The key point, as we will see in our results, is that the *invariance condition* ensures that, on one hand, the uniqueness holds just in the class of bounded solutions for HJB, on another hand a global L^1- stability holds for the KFP equation.

In the end, the contribution of this chapter will be the analysis of HJB equations, Fokker-Planck equations and, eventually, mean field games under the invariance structure conditions formulated above. In order to focus on the boundary behavior, we assume throughout this chapter that the matrix $a(x)$ is Lipschitz continuous in Ω and is elliptic in the interior of Ω, namely

$$a(x) \geq 0 \quad \forall\, x \in \overline{\Omega}, \quad \text{and} \;\; a(x) > 0 \text{ if } x \in \Omega. \tag{2.1.4}$$

This latter condition avoids the superposition of interior and boundary degeneracy which would make the analysis more complicated. Besides, the interior ellipticity will guarantee local compactness of the solutions which allows us to use a standard framework of weak (distributional) solutions. While we defer a precise statement of our results to the next sections, we list here a short summary of what we prove in this chapter :

(a) Assuming the structure conditions (2.1.3) and (2.1.4), and requiring the Hamiltonian $H(t, x, p)$ to be convex in p, locally bounded in x with at most quadratic growth with respect to p and such that $H(t, x, 0)$ is globally bounded, we show existence and uniqueness of bounded solutions to the HJB equation (2.1.2) with bounded terminal pay-off.

(b) Assuming that the drift $b(t, x)$ is locally bounded and the structure conditions (1.1.2) and (2.1.4), we prove that, for any initial probability density $m_0 \in L^1(\Omega)$ the Fokker-Planck equation

$$\begin{cases} \partial_t m - \sum\limits_{i,j} \partial^2_{ij}(a_{ij}(x)m) + \mathrm{div}(m\, b(t,x)) = 0\,, & (t,x) \in (0,T) \times \Omega \\ m(0) = m_0 \end{cases}$$

admits a unique weak solution $m \in C^0([0,T]; L^1(\Omega))$. Here by a weak solution we mean that

$$\int_0^T \int_\Omega m\, \mathcal{L}(\phi) dx dt = \int_\Omega m_0 \phi(0) dx$$

for every $\phi \in C^0([0,T]; L^1(\Omega)) \cap L^\infty$ such that $\mathcal{L}(\phi) \in L^\infty$, where $\mathcal{L}(\phi) = -\partial_t \phi - \sum\limits_{i,j} a_{ij}(x) \partial^2_{ij}\phi - b(t,x) \cdot D\phi$.

(c) Under the conditions on $a(x)$, H assumed in item (a) (in particular, under the invariance condition (2.1.3)), and assuming that the coupling terms $F(t, x, m)$ and $G(x, m)$ satisfy global bounds in L^∞ and suitable continuity conditions with respect to m, we prove that the mean field game system (2.1.1) admits a weak solution, where the two equations are formulated in the sense specified, respectively, by previous results in (a) and (b). This kind of solution of (2.1.1) is also unique under usual monotonicity conditions upon F and G (with respect to m).

(d) Assuming in addition that $a(x) = (\sigma\sigma^*)(x)$ with σ Lipschitz, plus a few natural structure conditions on the Hamiltonian H and further regularity of F and G, we prove additional regularity for the solution (u, m) of (2.1.1) ; namely, that u is (globally) Lipschitz continuous and semi concave in Ω. Moreover, in this case m is (globally) bounded provided $\sum_{i,j}[\frac{\partial a_{ij}}{\partial x_i} + H_{p_j}(t, x, Du)]\,\nu_j \geq 0$ on the boundary, where ν is the outward unit normal.

The spirit of the above results is that the invariance condition plays like a (transparent) boundary condition for the two equations. In fact, the existence of solutions will be provided by limit of (penalized) standard Neumann problems. The stochastic interpretation of the invariance condition easily explains that a kind of (transparent and soft) reflection naturally occurs near the boundary. The regularity results mentioned in item (d) show how the invariance condition may prevent the formation of singularities which, conversely, would occur in case of the state constraint problem. Indeed, global semi concavity may be lost in that case, see e.g. the recent paper [26].

Last but not least, we generalize our results to possibly non smooth domains. This generalization includes in particular the case that $\Omega = \prod_{i=1}^{N}(a_i, b_i)$ is a N-dimensional rectangle, which is often the case in applications.

We conclude by summarizing the organization of the chapter. In Section 2 we list the main notation and the standing assumptions which hold throughout the chapter ; then we give a few examples of control problems which fit our conditions and we properly state the main existence and uniqueness results which are proved. Section 3 is devoted to the study of the single HJB equation (2.1.2) under the invariance condition. Section 4 is devoted to the analysis of the single Fokker-Planck equation in the same context. The mean field game system (2.1.1) is studied and characterized in Section 5. Section 6 contains the additional regularity results on the solutions and specifically the Lipschitz and semi concavity regularity ; at this stage we need to make additional assumptions on the nonlinearity and this is why those results are not mentioned earlier in Section 2. Finally, Section 7 contains the generalization to non smooth domains. We leave to the Appendix the proof of a couple of technical results.

2.2 Preliminaries : assumptions and examples

We assume throughout the chapter that Ω is a bounded open subset of $\mathbb{R}^N$, $N \geq 1$. We denote $Q_T := (0,T) \times \Omega$. We recall that the *oriented distance* from $\partial\Omega$, denoted by d_Ω, is the function defined by

$$d_\Omega(x) = \begin{cases} d(x, \partial\Omega) & \text{if } x \in \Omega \\ -d(x, \partial\Omega) & \text{if } x \notin \Omega \end{cases}$$

where, as usual, $d(x, \partial\Omega) = \inf\limits_{y \in \partial\Omega} |x-y|$. We write d instead of d_Ω when there is no possible mistake for Ω. It is well-known that d_Ω is a 1-Lipschitz function which coincides with the unique viscosity solution of the eikonal equation

$$\begin{cases} |Du| = 1 & x \in \Omega \\ u = 0 & x \in \partial\Omega\,. \end{cases}$$

Moreover, if we require some regularity for the set Ω, we obtain further regularity for d_Ω.

Definition 2.2.1. *Let $K \subseteq \mathbb{R}^N$. We say that K is a compact domain of class $\mathcal{C}^2$ if K is a compact connected set and $\exists M \in \mathbb{N}$ such that $\forall 1 \leq i \leq M$ $\exists B_{r_i}(x_i)$, $x_i \in \partial K$ and a function $\phi_i : B_{r_i}(x_i) \to \mathbb{R}$ such that*
(i) $\partial K \subseteq \bigcup_{i=1}^M B_{r_i}(x_i)$
(ii) $\partial K \cap B_{r_i}(x_i) = \{\phi_i = 0\}$
(iii) ϕ_i is of class $\mathcal{C}^2$ with $D^2\phi_i$ bounded in $B_{r_i}(x_i)$.

In the following, we assume that Ω is an open set such that $\overline{\Omega}$ is a compact domain of class $\mathcal{C}^2$. We set

$$R_\varepsilon := \{x \in \mathbb{R}^N \,:\, |d_\Omega(x)| < \varepsilon\}$$
$$\Gamma_\varepsilon := \{x \in \Omega \,:\, d_\Omega(x) < \varepsilon\} = R_\varepsilon \cap \Omega\,.$$

We recall (see e.g. [27] and [50]) that

$$\overline{\Omega} \quad \text{is a compact domain of class } \mathcal{C}^2 \iff \exists \varepsilon_0 > 0 : \quad d_\Omega \in \mathcal{C}^2(R_{\varepsilon_0})$$

and

$$\begin{gathered} \forall x \in \Gamma_{\varepsilon_0}\ \exists!\ \overline{x} \in \partial\Omega \text{ s.t. } d_\Omega(x) = |x - \overline{x}| \\ \text{and } Dd_\Omega(x) = Dd_\Omega(\overline{x}) = -\nu(\overline{x}) \end{gathered} \tag{2.2.1}$$

where ν stands for the outward unit normal to $\partial\Omega$.

Remark : Actually, we will use the function d_Ω only near $\partial\Omega$, where this is a regular function. So, from now on, when we will write d_Ω (or d when there is no possible mistake) we will mean a $C^2(\overline{\Omega})$ function $\tilde{d}$ such that $\exists\varepsilon_0 > 0$ with $\tilde{d} = d$ in Γ_{ε_0}.

For every $r \in \mathbb{R}$ we set

$$D_r = \left\{x \in \Omega \text{ s.t. } d(x) \geq \frac{1}{r}\right\} = \Omega \setminus \Gamma_{\frac{1}{r}} \tag{2.2.2}$$

so that we build a sequence $\{D_n\}_{n\in\mathbb{N}}$ of compact domains of class $\mathcal{C}^2$ such that

$$D_n \subseteq \overset{\circ}{D}_{n+1} \qquad \text{and} \qquad \bigcup_{n=1}^{\infty} D_n = \Omega\,.$$

For each couple of vectors $(v, w) \in \mathbb{R}^n \times \mathbb{R}^m$, the tensor product $v \otimes w$ denotes the $n \times m$-matrix $v^T w$. Finally, throughout the proofs we use the notation C to denote a generic constant which may vary from line to line.

2.2.1 Standing assumptions.

Let us now make precise the assumptions on the coefficients a_{ij} and on the nonlinearities H, F, G of the system (2.1.1).

We assume that $a(x) = (a_{ij}(x))_{ij}$ is a $N \times N$-matrix which belongs to $W^{1,\infty}(\Omega)^{N\times N}$ and satisfies

$$a(x)\xi \cdot \xi > 0 \quad \forall\, x \in \Omega, \forall \xi \in \mathbb{R}^N\,. \tag{2.2.3}$$

We call $\lambda_r \in \mathbb{R}_{>0}$ the constant of uniform ellipticity of a in D_r, i.e.

$$a(x)\xi \cdot \xi \geq \lambda_r |\xi|^2 \qquad \forall x \in D_r\,, \quad \forall \xi \in \mathbb{R}^N\,. \tag{2.2.4}$$

Obviously we have $\lambda_r \leq \lambda_s$ if $r \geq s$. Moreover, by continuity the matrix $a(x)$ will be nonnegative on $\overline{\Omega}$, but it is allowed to vanish at the boundary, in which case $\lambda_r \searrow 0$.

We assume that $H(t, x, p)$ is a function such that $(t, x) \mapsto H(t, x, p)$ is measurable for any given $p \in \mathbb{R}^N$ and $p \mapsto H(t, x, p)$ is of class C^1 for almost every $(t, x) \in Q_T$. We assume in addition that

$$p \mapsto H(t,x,p) \quad \text{is convex.} \tag{2.2.5}$$

Concerning the growth of the Hamiltonian, we work assuming that it has at most quadratic growth with respect to p and is locally bounded with respect to x, with $H(t,x,0)$ globally bounded. Precisely, we assume that

$$H(t,x,0) \in L^\infty(Q_T) \tag{2.2.6}$$

and

$$\begin{gathered} \forall \text{ compact set } K \subset \Omega, \quad \exists C_K > 0 : \\ |H_p(t,x,p)| \le C_K(1+|p|) \qquad \forall p \in \mathbb{R}^N, \text{ and a.e. } x \in K, t \in [0,T]\,. \end{gathered} \tag{2.2.7}$$

Of course, (2.2.6)–(2.2.7) imply, by integration, that

$$|H(t,x,p)| \le C_K(1+|p|^2) \qquad \forall p \in \mathbb{R}^N, \text{ and a.e. } x \in K, t \in [0,T] \tag{2.2.8}$$

for a possibly different constant C_K.
The invariance condition will be formulated in terms of the diffusion matrix $a(x)$ and the Hamiltonian function $H(t,x,p)$. Namely, we assume that there exist $\delta > 0$ and $C > 0$ such that the following inequality holds :

$$\begin{gathered} \mathrm{tr}(a(x)D^2d(x)) - H_p(t,x,p)Dd(x) \ge \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)} - Cd(x) \\ \forall\ p \in \mathbb{R}^N \text{ and a.e. } x \in \Gamma_\delta,\ t \in [0,T], \end{gathered} \tag{2.2.9}$$

where, we recall, Γ_δ is the subset of Ω with $d(x) < \delta$.

A typical case when assumption (2.2.9) is satisfied occurs if there exists a $N \times N$-matrix $\sigma \in W^{1,\infty}$ such that $a = \sigma\sigma^*$,

$$\sigma^*(x)Dd(x) = 0 \qquad\qquad \forall x \in \partial\Omega\,, \tag{2.2.10}$$

and

$$\mathrm{tr}(a(x)D^2d(x)) - H_p(t,x,p)Dd(x) \ge 0$$

for all $p \in \mathbb{R}^N$ and all x in a neighborhood of $\partial\Omega$. Moreover, in the case that $H_p(t,x,p)$ is Lipschitz with respect to x (uniformly in t and p), the inequality can be required to hold only for $x \in \partial\Omega$, since (2.2.9) will be equally satisfied for x in some Γ_δ provided the constant

C is large enough. This is the typical condition which is given in the literature for linear operators, i.e. if $H(t,x,p) = b(x) \cdot p$, see e.g. [27].

However, we stress that assumption (2.2.9) is meant to include more general examples. On one hand, this condition includes the case of $a = \sigma\sigma^*$ with σ being only 1/2-Hölder continuous. On another hand, even the uniformly elliptic case is included in our setting, indeed $a(x)$ could be non degenerate at the boundary provided the drift part is sufficiently coercive in the (inward) normal direction. Situations of this kind were considered, for instance, in [79].

Finally, the assumptions on the coupling costs F, G. Here we assume that F is a map from $Q_T \times C^0([0,T]; L^1(\Omega))$ into $\mathbb{R}$. In particular, for any given $m \in C^0([0,T]; L^1(\Omega))$, $F(\cdot,\cdot,m)$ defines a function on Q_T. We assume that

$$\begin{gathered} m \mapsto F(\cdot,\cdot,m) \text{ maps bounded sets of } C^0([0,T]; L^1(\Omega)) \text{ into bounded sets of } L^\infty(Q_T), \\ \text{and is continuous in the } L^1(Q_T)\text{- topology.} \end{gathered} \tag{2.2.11}$$

We wish to include two model examples in the previous conditions. The simplest case is when F acts locally on the density $m(t,x)$: this means, for instance, that F is given through a real function $f : Q_T \times \mathbb{R} \to \mathbb{R}$ so that

$$F(t,x,m) := f(t,x,m(t,x))\,.$$

In this case the condition is satisfied whenever f is continuous with respect to m and is uniformly bounded. A second class of examples is given by nonlocal functions F, as, for instance, $F = K \star m$ for some bounded convolution kernel K.

A similar condition is assumed for G, although here we need to strengthen the requirements in order to ensure that Du is bounded up to $t = T$ (unless H is Lipschitz continuous, see also Remak 2.5.4). Namely, we assume that G is a map from $\Omega \times L^1(\Omega)$ into $\mathbb{R}$ such that

$$\begin{gathered} m \mapsto G(\cdot,m) \text{ is a continuous map from } L^1(\Omega) \text{ into } L^1(\Omega) \\ \text{which maps bounded sets of } L^1(\Omega) \text{ into bounded sets of } W^{1,\infty}(\Omega). \end{gathered} \tag{2.2.12}$$

As is customary in mean field game systems, we will require in addition some monotonicity of F, G in order to have uniqueness of solutions.

2.2.2 Short statement of the main results.

We list here the three main results that we prove in the chapter, standing on the assumptions previously introduced. The first one is just concerned with the Hamilton-Jacobi-Bellman equation. The notion of weak solution is a standard one and will be precisely given in Definition 2.3.1. Under the invariance condition, it turns out that the problem is well-posed in the class of (globally) bounded solutions, with no need of prescription of the boundary condition.

Theorem 2.2.2. *Assume that $a(x)$ and $H(t,x,p)$ satisfy assumptions (2.2.3), (2.2.5)-(2.2.7) and the invariance condition (2.2.9), and that $G \in L^\infty(\Omega)$.*

Then there is one and only one bounded weak solution of the problem

$$\begin{cases} -\partial_t u - \sum_{i,j} a_{ij}(x)\partial^2_{ij}u + H(t,x,Du) = 0\,, & (t,x) \in (0,T)\times\Omega \\ u(T) = G(x)\,, \qquad x \in \Omega\,. \end{cases}$$

The second result gives, somehow, a counterpart for the Fokker-Planck equation. Indeed, under the invariance condition the problem turns out to be well-posed in $L^1(\Omega)$. Here the notion of weak solution is defined in a dual way, see Definition 2.4.1, and incorporates somehow a transparent Neumann condition at the boundary.

With a slight abuse of notation, we denote by $L^\infty([0,T]; L^\infty_{loc}(\Omega))$ the space of measurable functions in Q_T which are bounded on $(0,T)\times K$ for every compact subset $K \subset \Omega$.

Theorem 2.2.3. *Let $m_0 \in L^1(\Omega)$, $m_0 \geq 0$. Let $a \in W^{1,\infty}(\Omega)$ satisfy (2.2.3). Assume that $b \in L^\infty([0,T]; L^\infty_{loc}(\Omega))$ and that there exist $\delta_0, C > 0$ such that the following inequality holds :*

$$\mathrm{tr}(a(x)D^2 d(x)) - b(t,x)\cdot Dd(x) \geq \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)} - C\,d(x) \tag{2.2.13}$$

for almost every $t \in (0,T)$ and $x \in \Gamma_{\delta_0}$.

Then there is one and only one weak solution (in the sense of Definition 2.4.1) of the problem

$$\begin{cases} \partial_t m - \sum_{i,j} \partial^2_{ij}(a_{ij}(x)m) - \mathrm{div}(m\,b(t,x)) = 0\,, & (t,x) \in (0,T)\times\Omega \\ m(0) = m_0(x)\,, \qquad x \in \Omega\,. \end{cases}$$

Our third main result is concerned with the mean field game system, where we join the two previous results, using the conditions on the coupling terms and the viability assumption on the Hamiltonian.

Theorem 2.2.4. *Assume that hypotheses (2.2.3), (2.2.5)-(2.2.7), (2.2.9), (2.2.11) and (2.2.12) hold true. Then there exists one solution (u, m) of (2.1.1), in the sense of Definition 2.5.1.*

If, in addition, F and G are monotone with respect to m, then the solution is unique.

The further results that we prove are concerned with the regularity of solutions, but in this case we postpone the proper statements to Section 7.

2.2.3 Probabilistic interpretation and examples.

Now we give the probabilistic interpretation of the system. Given a probability space $(\tilde{\Omega}, (\mathcal{F}_t)_t, \mathbb{P})$ (we use $\tilde{\Omega}$ instead of the classical Ω to avoid confusion with the state space Ω previously defined) and a Brownian motion $(B_t)_t$ adapted to the filtration $(\mathcal{F}_t)_t$, we consider for $s > t$ the solution $(X_s)_s$ of the following stochastic differential equation :

$$\begin{cases} dX_s = b(s, X_s, \alpha_s)ds + \sqrt{2}\sigma(X_s)dB_s \\ X_t = x \end{cases} \tag{2.2.14}$$

where b and $\sqrt{2}\sigma$ are, as usual, the *drift* and the *diffusion* coefficients of the process $X_\cdot$, and where the control α_s is a progressively measurable process adapted to the filtration $\mathcal{F}_t$ and taking values in $A \subseteq \mathbb{R}^N$.

We recall the *Ito's formula* : if $\phi \in \mathcal{C}^{1,2}[0,T] \times \mathbb{R}^N)$, then we have

$$\begin{aligned} d\phi(s, X_s) = \big(\phi_t(s, X_s) + \mathrm{tr}(a(X_s)D^2\phi(s, X_s)) + b(s, X_s, \alpha_s) \cdot \nabla\phi(s, X_s)\big)\, ds\, + \\ + \sqrt{2}(\nabla\phi(s, X_s))^*\sigma(X_s)dB_s\,. \end{aligned}$$

In many applications, it is required that the process $(X_t)_t$ remains in Ω for every $t \geq 0$ and *for all available controls*. This leads to the terminology of *invariance condition* for assumption (2.2.9), which is justified in view of the following result.

Proposition 2.2.5. *Let $\sigma \in W^{1,\infty}(\Omega)$ and $b(s, x, \alpha)$ be (locally) Lipschitz with respect to the time and space variables, with a Lipschitz constant (locally) uniform in α, and suppose* *that, for some* $\delta > 0$ *and* $C > 0$ *:*

$$\begin{aligned} &\operatorname{tr}(a(x)D^2 d(x)) + b(s,x,\alpha)\cdot Dd(x) \geq \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)} - Cd(x) \\ &\forall (s,x) \in [t,T]\times\Gamma_\delta\,, \quad \forall \alpha \in A\,, \end{aligned} \tag{2.2.15}$$

where $a = \sigma\sigma^*$*. Then, if* $(X_s)_s$ *is the solution of* (2.2.14) *with starting point* $x \in \Omega$*, we have*

$$\mathbb{P}(\{X_s \in \Omega\, \forall s > t\}) = 1\,. \tag{2.2.16}$$

Démonstration. The proof is classical (see e.g. [9], [27] for similar results) but we include it for the reader's convenience and because condition (2.2.15) applies to a more general setting than usual.

Let $(X_s)_s$ be the process solving (2.2.14). The existence and uniqueness of X_t is ensured by the local Lipschitz character of b, σ. For a bounded set $E \in \mathbb{R}^N$ we call τ_E the exit time from E of the process X_s : for $\omega \in \tilde{\Omega}$

$$\tau_E(\omega) = \inf\{s \geq t \mid X_s(\omega) \notin E\}\;.$$

So, proving (2.2.16) is equivalent to prove that $\mathbb{P}(\tau_\Omega < +\infty) = 0$. To this purpose, we will show that, for all $s > t$, we have

$$\mathbb{P}(\tau_\Omega \leq s) = 0\,. \tag{2.2.17}$$

Indeed, since

$$\begin{aligned} &\{\tau_\Omega \leq s\} \subseteq \{\tau_\Omega \leq r\} \qquad \text{for } s \leq r \\ \text{and} \qquad &\bigcup_s \{\tau_\Omega \leq s\} = \{\tau_\Omega < +\infty\}\;, \end{aligned}$$

then the assertion follows thanks to the monotone convergence theorem. To prove (2.2.17), we will show that

$$V(x) := -\log(d(x))$$

is, roughly speaking, a super solution up to a constant. Indeed, according to (2.2.15) we obtain

$$\begin{aligned} \operatorname{tr}(a(x)D^2V) + b(s,x,\alpha)\cdot DV = &-\frac{\operatorname{tr}(a(x)D^2 d(x)) + b(s,x,\alpha)\cdot Dd(x)}{d(x)} \\ &+ \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)^2} \leq C\,, \end{aligned}$$

for each $(s,x) \in [t,T] \times \Gamma_\delta$ and for each $\alpha \in A$. Now, by a standard localization argument, we obtain a non-negative $\mathcal{C}^2$ function U such that

$$\begin{cases} U(x) = V(x) & \text{for } x \in \Gamma_{\frac{\delta}{2}}\,, \\ \operatorname{tr}(a(x)D^2U(x)) + b(s,x,\alpha)\cdot DU(x) \le C & \text{for } x \in \Omega\,,\ \ s \in [t,T]\,,\ \ \alpha \in A\,; \end{cases}$$

we recall that the constant C can change from line to line. To conclude, we consider a sequence of compact domains $\{D_n\}_n$ converging to Ω, and the associated stopping times τ_{D_n}. Applying Ito's Formula to U and taking the expectation, we have

$$\mathbb{E}\left[U(X_{s\wedge\tau_{D_n}})\right] = U(x) + \mathbb{E}\left[\int_t^{s\wedge\tau_{D_n}} \big(\operatorname{tr}(a(X_r)D^2U(X_r) + b(r,X_r,\alpha_r)\cdot DU(X_r)\big)\,dr\right]$$

hence

$$\mathbb{E}\left[U(X_{s\wedge\tau_{D_n}})\right] \le U(x) + C(s-t) < +\infty$$

since $x \in \Omega$. Using Fatou's Lemma we get

$$\mathbb{E}\left[U(X_{s\wedge\tau_\Omega})\right] \le \liminf_{n\to\infty} \mathbb{E}\left[U(X_{s\wedge\tau_{D_n}})\right] \le U(x) + C(s-t) < +\infty\,.$$

Since $U(x)$ blows-up if $x \in \partial\Omega$, this implies

$$\mathbb{P}(\tau_\Omega \le s) = 0\,.$$

□

For $t \in [0,T]$ and $x \in \Omega$, we define now the value function

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

where, for every s, $m(s)$ is a probability density function.
Here F and G are the *cost functions* and the *Lagrangian* L satisfies standard conditions. Typically, we require the strict convexity of the function L. In a usual way, one can apply the Ito's formula and the dynamic programming principle to obtain a Hamilton-Jacobi-Bellman equation for the function u. So defining the Hamiltonian H as

$$H(t,x,p) := \sup_{\alpha\in A} \left(-b(t,x,\alpha)\cdot p - L(t,x,\alpha)\right)\,,$$

the value function u turns out to be the solution of the following equation :

$$\begin{cases} -\partial_t u - \mathrm{tr}(a(x)D^2 u) + H(t,x,Du) = F(t,x,m) \\ u(T) = G(x, m(T)) \,. \end{cases} .$$

Moreover, we obtain a feedback optimal control $b = -H_p(s,x,Du(s,x))$. Plugging this control in (2.2.14) and solving the SDE gives the optimal trajectories $(\tilde{X}_s)_s$. We say that the system is in equilibrium if the law of X_s coincides with $m(s)$ for every $s \in [0,T]$. Since the law of a solution of (2.2.14) solves a Fokker-Planck equation, we obtain a couple (u,m) solution of (2.1.1).

We give here two typical examples in which the hypotheses made so far upon H are satisfied. In particular, the invariance condition may be satisfied by using controlled perturbations of linear invariant processes.

Example 1 *(bounded controls).*

We consider a set of controls $A \subset \mathbb{R}^N$ which is compact and a positive number M. We take

$$b(x,\alpha) := MDd(x) + \alpha \,.$$

Then, for any $a(x)$ such that $a(x)Dd(x) \cdot Dd(x) = 0$ on $\partial\Omega$, the invariance condition (2.2.9) is satisfied for M sufficiently large.
Indeed, since the set of controls is bounded and a is Lipschitz, we have

$$\frac{a(x)Dd(x)\cdot Dd(x)}{d(x)} - \mathrm{tr}(a(x)D^2 d(x)) - \alpha \cdot Dd(x) \le C$$

for some constant $C > 0$ independent of α. So (2.2.9) holds provided M is large enough. Let us now check the other conditions assumed upon H. In this situation, the Hamiltonian H takes the form

$$H(x,p) = \sup_{\alpha \in A} \left(-\alpha \cdot p - MDd(x)\cdot p - L(x,\alpha)\right) \,.$$

Of course H is a convex function with respect to the variable p. Assume further that L is strictly convex with respect to the last variable. Then the supremum that arises in the definition of $H(x,p)$ is attained at a unique point, say $\alpha_{p,x}$, and the mapping $(x,p) \mapsto \alpha_{p,x}$

is continuous. In particular, $H(x,p)$ is a continuous function and one can further check that it is differentiable with respect to p with

$$H_p(x,p) = -b(x,\alpha_{p,x}) = -MDd(x) - \alpha_{p,x}\,, \tag{2.2.18}$$

which is continuous in both arguments and bounded uniformly with respect to (x,p). Thus, H satisfies (2.2.5)–(2.2.7).

In addition, we notice that, if L is Lipschitz with respect to x, uniformly in α, then H is Lipschitz with respect to x and

$$H_x = -M\,D^2d(x)p - L_x(x,\alpha_{p,x})\,,$$

which satisfies the growth condition $|H_x(x,p)| \le C(1+|p|)$. If we have further that L is C^1 with respect to x with L_x continuous with respect to α, then H is also $\mathcal{C}^1$ in both variables.

Finally, let us stress that a similar example can be adapted to the case that $a(x)$ does not degenerate at the boundary, namely if $a(x) > 0$ in $\overline{\Omega}$. In that case, it is enough to take $b(x,\alpha) = M\,\frac{Dd(x)}{d(x)} + \alpha$ in order to build a similar example.

Example 2 *(unbounded controls and coercive Hamiltonian).*

Here we consider a case in which the set of controls is unbounded. This gives us an Hamiltonian with super-linear growth in p.

We take $A = \{\alpha \in \mathbb{R}^N \text{ s.t. } \alpha_i \ge 0\ \forall i\}$ and we set

$$b(x,\alpha) := MDd(x) + B(x)\alpha\,,$$

where $\forall x\ B(x)$ is a $N\times N$-real valued matrix. Let $c_0 \ge 0$ be such that

$$\operatorname{tr}(a(x)D^2d(x)) \ge -c_0\,.$$

Then we have

$$\operatorname{tr}(a(x)D^2d(x)) + b(x,\alpha)Dd(x) \ge M - c_0 + B(x)\alpha\cdot Dd(x) \ge M - c_0\,,$$

if we choose B such that

$$B(x)\alpha\cdot Dd(x) \ge 0\,, \qquad \forall \alpha \in A\,. \tag{2.2.19}$$

For example, we can take $B(x)_{ij} = Dd(x)_i\delta_{ij}$. If (2.2.19) holds, then the invariance condition (2.2.9) is satisfied provided M is sufficiently large.

Let us suppose further that the matrix $B(x)$ is bounded and continuous. As in the previous example, let the function $L(x,\alpha)$ be continuous in both arguments and strictly convex in α, and assume the following coercivity condition : $\exists \eta > 0$, $q > 1$, $c_0 > 0$ such that

$$L(x,\alpha) \geq \eta|\alpha|^q - c_0 \qquad\qquad \forall \alpha \in A\,,\; x \in \overline{\Omega}\,.$$

Then one can readily check the properties of H as before. In particular, the supremum that arises in the definition of $H(x,p)$ is attained at a unique point $\alpha_{p,x}$, which is continuous with respect to (x,p) and now satisfies the estimate :

$$|\alpha_{p,x}| \leq C(1 + |p|^{q'-1}) \qquad \forall (x,p)\,,$$

where q' is the conjugate exponent of q, i.e. $q' = \frac{q}{q-1}$. As a consequence, there exists a constant $C > 0$ such that

$$-C(1+|p|) \leq H(x,p) \leq C(1 + |p|^{q'})\,.$$

Similarly, the differentiability of H with respect to p and formula (2.2.18) imply that

$$|H_p(x,p)| \leq C(1 + |p|^{q'-1})\,,$$

so that H has at most quadratic growth for $q \geq 2$ and satisfies (2.2.5)–(2.2.7).

Moreover, if L and B are $\mathcal{C}^1$ with respect to x, and the derivative of L is continuous in α, then H is $\mathcal{C}^1$ in both variables. Finally, we notice that, if L_x has a linear growth in α, then $|H_x(x,p)| \leq C(1 + |p|^{q'})$.

2.3 The Hamilton-Jacobi-Bellman equation

In this Section we study the Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{cases} -u_t - \mathrm{tr}(a(x)D^2u) + H(t,x,Du) = 0 & \text{in } (0,T)\times\Omega, \\ u(T) = G(x) & \text{in } \Omega \end{cases} . \tag{2.3.1}$$

under conditions of invariance of the domain (see (2.2.9)). In particular, we observe that no boundary condition is prescribed. Here, the boundedness of u will be enough to characterize the solution.

Note that in this formulation we do not have a function F, since, for m fixed, it can be included in the Hamiltonian H. Since F does not depend on Du, the derivative H_p does not change, and the conditions of invariance (2.2.9) are not affected from this inclusion.

We assume that G is a bounded function and the Hamiltonian H satisfies (2.2.6) and (2.2.8).

Different notions of solutions could be used for problem (2.3.1). Since we are assuming a to be Lipschitz continuous, the problem can be formulated in divergence form as well, namely

$$\begin{cases} -u_t - \operatorname{div}(a(x)Du) + \tilde{b}(x) \cdot Du + H(t,x,Du) = 0 \\ u(T) = G(x) \end{cases}$$

where $\tilde{b}$ is defined as

$$\tilde{b}_j(x) = \sum_{i=1}^{N} \frac{\partial a_{ij}}{\partial x_i}(x)\,, \qquad j = 1, \dots N\,. \tag{2.3.2}$$

This fact allows us to use a weak (distributional) formulation which avoids any continuity requirement on the solution as well as on F, H with respect to t and x. The natural growth condition (2.2.8) also leads us to consider local H^1 solutions as defined below.

Definition 2.3.1. *We say that u is a weak solution (resp. subsolution, supersolution) of the problem* (2.3.1) *if*

(i) $u \in L^\infty([0,T] \times \Omega)$ *;*
(ii) $u \in L^2([0,T]; W^{1,2}(K))$ *for each* $K \subset\subset \Omega$ *;*
(iii) $\forall \phi \in C_c^\infty((0,T] \times \Omega)$ *(resp.* ≥ 0*) the weak formulation holds :*

$$\int_0^T \int_\Omega u\phi_t \, dxdt \; + \int_0^T \int_\Omega a(x)Du \cdot D\phi \, dxdt \; + \\ + \int_0^T \int_\Omega (H(t,x,Du) + \tilde{b}(x)Du)\phi \, dxdt = \int_\Omega G(x)\phi(T)dx\,,$$

(resp. $\leq$, $\geq$*), where* $\tilde{b}$ *is defined above.*

Remark 2.3.2. We observe that, if u is a weak solution of (2.3.1), then $u \in \mathcal{C}([0,T]; L^p(\Omega))$ for each $p \geq 1$.
Indeed, from $(ii)-(iii)$ we have $u_t \in L^2([0,T]; W^{-1,2}(K)) + L^1((0,T) \times K)$ for each $K \subset\subset \Omega$, and so $u \in C([0,T]; L^1(K))$ (see [89, Theorem 2.2]). Since $u \in L^\infty([0,T] \times \Omega)$, one can actually conclude that $u \in C([0,T]; L^p(\Omega))$ for each $p \geq 1$.

2.3.1 Existence of solutions

We start by proving the existence of at least one weak solution. This is achieved without using the invariance condition and actually follows by a standard use of global bounds and local compactness. The next lemma is by now standard, following the arguments in [22]. For the reader's convenience, since the statement may not be found exactly in this form in previous references, we will give a short proof in the Appendix.

Lemma 2.3.3. *Let $\{\Omega_\varepsilon\}$ be a sequence of domains such that $\Omega_\varepsilon \subseteq \Omega_\eta \subseteq \Omega$ for $\varepsilon > \eta$, and $\bigcup_\varepsilon \Omega_\varepsilon = \Omega$. Let H_ε be a sequence of Carathéodory functions such that*

$$|H_\varepsilon(t,x,p)| \leq C_\varepsilon(1+|p|^2) \quad \text{for a.e. } (t,x) \in (0,T) \times \Omega_\varepsilon \text{ and } \forall p \in \mathbb{R}^N,$$

where C_ε is bounded (independently of ε) in $(0,T) \times K$, for every compact set $K \subset \Omega$.

Assume that a_ε is a sequence of matrices which is uniformly bounded and, locally, uniformly coercive. Let u_ε be a sequence of solutions of

$$-(u_\varepsilon)_t - \operatorname{div}(a_\varepsilon(x) Du_\varepsilon) + H_\varepsilon(t,x,Du_\varepsilon) = 0\,, \quad (t,x) \in [0,T] \times \Omega_\varepsilon\,, \tag{2.3.3}$$

such that $\|u_\varepsilon\|_\infty$ is uniformly bounded.

Then there exists a function $u \in L^\infty(Q_T) \cap L^2(0,T; W^{1,2}_{loc}(\Omega))$ and a subsequence u_ε converging to u weakly in $L^2(0,T;W^{1,2}(K))$ and strongly in $L^p((0,T)\times K)$, for all compact sets $K \subset\subset \Omega$ and all $p < \infty$.

In addition, if $a_\varepsilon(x)$ converges almost everywhere in Ω to some matrix $a(x)$, then u_ε converges to u strongly in $L^2(0,t;W^{1,2}(K))$ for all $t < T$, and the convergence holds up to $t = T$ if $u_\varepsilon(T)$ converges almost everywhere in Ω.

Démonstration. See the Appendix. □

From the above Lemma we deduce the following stability result.

Proposition 2.3.4. *Let $\{\Omega_\varepsilon\}$ be a sequence of domains such that $\Omega_\varepsilon \subseteq \Omega_\eta \subseteq \Omega$ for $\varepsilon > \eta$, and $\bigcup_\varepsilon \Omega_\varepsilon = \Omega$. Assume that $H_\varepsilon(t,x,p)$ is a sequence of Carathéodory functions such that :*

$$\begin{aligned} &H_\varepsilon(t,x,p) \to H(t,x,p) \qquad \text{for a.e. } (t,x) \in (0,T) \times \Omega \text{ and every } p \in \mathbb{R}^N. \\ &\|H_\varepsilon(t,x,0)\|_\infty \leq C\,, \\ &|H_\varepsilon(t,x,p)| \leq C_\varepsilon(1+|p|^2) \qquad \text{for a.e. } (t,x) \in (0,T) \times \Omega_\varepsilon \text{ and every } p \in \mathbb{R}^N, \end{aligned} \tag{2.3.4}$$

where C is a constant independent of ε and C_ε is a constant which is bounded (independently of ε) in $(0,T) \times K$, for every compact set $K \subset \Omega$.

Assume that a_ε is a sequence of matrices which is uniformly bounded and, locally, uniformly coercive and, moreover, there exists a matrix $a(x)$ such that

$$a_\varepsilon(x) \to a(x) \qquad \text{for a.e. } x \in \Omega\,.$$

Finally, assume that $\{G_\varepsilon\}_\varepsilon$ is a sequence of uniformly bounded functions such that $\exists G \in L^\infty(\Omega)$ with

$$G_\varepsilon(x) \to G(x) \qquad \text{for a.e. } x \in \Omega\,.$$

Let $u_\varepsilon \in L^2([0,T]; W^{1,2}(\Omega_\varepsilon))$ be the unique solution of the approximating system

$$\begin{cases} -(u_\varepsilon)_t - \operatorname{div}(a_\varepsilon(x)Du_\varepsilon) + H_\varepsilon(t,x,Du_\varepsilon) = 0 & (t,x) \in (0,T) \times \Omega_\varepsilon \\ u_\varepsilon(T) = G_\varepsilon(x) \\ a_\varepsilon(x)Du_\varepsilon \cdot \nu_{|\partial\Omega_\varepsilon} = 0\,. \end{cases} \tag{2.3.5}$$

Then we have that there exists $u \in L^\infty(Q_T) \cap L^2(0,T; W^{1,2}_{loc}(\Omega))$ and a subsequence u_ε such that

$$u_\varepsilon \to u \qquad \text{in } L^2(0,T; W^{1,2}(K)) \cap C^0([0,T]; L^p(\Omega)) \text{ for any } p < \infty$$

for all compact sets $K \subset \Omega$, and u is a weak solution of

$$\begin{cases} -u_t - \operatorname{div}(a(x)Du) + H(t,x,Du) = 0\,, & (t,x) \in (0,T) \times \Omega \\ u(T) = G(x)\,. \end{cases}$$

Démonstration. Since $\|H_\varepsilon(t,x,0)\|_\infty$ is uniformly bounded, by maximum principle we have that $\|u_\varepsilon\|_\infty$ is uniformly bounded. From Lemma 2.3.3, we deduce that there exists a function $u \in L^\infty(Q_T) \cap L^2(0,T; W^{1,2}_{loc}(\Omega))$ and a subsequence u_ε converging to u strongly in $L^p(Q_T)$ $\forall p$ and in $L^2(0,T; W^{1,2}(K))$, for all compact sets $K \subset \Omega$. In particular, we have that $Du_\varepsilon \to Du$ in $L^2([0,T] \times K)$ for any compact subset K. Thanks to the pointwise convergence of H_ε towards H, and due to the growth assumptions, we infer by Lebesgue theorem that

$$H_\varepsilon(t,x,Du_\varepsilon) \to H(t,x,Du) \qquad\qquad \text{in } L^1([0,T] \times K)$$

for all compact sets $K \subset \Omega$. Since the equation implies that $(u_\varepsilon)_t$ strongly converges in $L^2(0,T; W^{-1,2}(K)) + L^1([0,T] \times K)$, from the embedding result in [89, Theorem 2.2], we deduce that

$$u_\varepsilon \to u \qquad \text{in } C^0([0,T]; L^1(K))$$

and due to the L^∞ bound the convergence actually holds in $C^0([0,T]; L^p(\Omega))$ for every $p < \infty$. Now we can pass to the limit in the weak formulation of (2.3.5) to show that u is a weak solution. □

We finally deduce the existence of a solution for our problem.

Theorem 2.3.5. *Suppose $G \in L^\infty(\Omega)$. Assume that a satisfies (2.2.3) and that $H(t,x,p)$ satisfies (2.2.6) and (2.2.8). Then the problem (2.3.1) has at least one solution.*

Démonstration. We first notice that, replacing $H(t,x,p)$ by $\tilde{H}(t,x,p) = H(t,x,p) + \tilde{b}(x)\cdot p$, where $\tilde{b}$ is defined in (2.3.2), the function $\tilde{H}$ satisfies the same hypotheses as H. So, without loss of generality, it is enough to prove the existence of solutions for the following equation :

$$\begin{cases} -u_t - \operatorname{div}(a(x)Du) + H(t,x,Du) = 0 \\ u(T) = G(x)\,. \end{cases} \tag{2.3.6}$$

Here we define the truncation of H at levels $\pm\frac{1}{\varepsilon}$:

$$H_\varepsilon(t,x,p) := \min\left\{\max\left\{H(t,x,p), -\frac{1}{\varepsilon}\right\}, \frac{1}{\varepsilon}\right\}.$$

and we consider $u_\varepsilon \in L^2([0,T];W^{1,2}(\Omega))$ as the unique solution of the penalized Neumann problem

$$\begin{cases} -(u_\varepsilon)_t - \operatorname{div}((a(x)+\varepsilon I)Du_\varepsilon) + H_\varepsilon(t,x,Du_\varepsilon) = 0\,, & (t,x) \in (0,T)\times\Omega\,, \\ u_\varepsilon(T) = G(x) \\ Du_\varepsilon \cdot \nu_{|\partial\Omega} = 0\,. \end{cases} \tag{2.3.7}$$

Applying Proposition 2.3.4 with $\Omega_\varepsilon = \Omega$, we conclude. □

We stress that an alternative way to construct a solution of (2.3.1) would be to use Neumann problems on a sequence of domains converging to Ω ; the local ellipticity of $a(x)$ and the local boundedness of H would avoid to approximate the nonlinearities, in this case. The proof is again a straightforward consequence of Lemma 2.3.3 and Propositon 2.3.4.

Proposition 2.3.6. *Let us set $\Omega^\varepsilon := \{x \in \Omega \,:\, d(x) > \varepsilon\}$ and let u_ε be the unique solution of the Neumann problem*

$$\begin{cases} -(u_\varepsilon)_t - \operatorname{tr}(a(x)D^2u_\varepsilon) + H(t,x,Du_\varepsilon) = 0 & \text{in } (0,T)\times\Omega^\varepsilon, \\ u_\varepsilon(T) = G(x) & \text{in } \Omega^\varepsilon, \\ a(x)Du_\varepsilon \cdot \nu_{|\partial\Omega^\varepsilon} = 0\,. \end{cases} \tag{2.3.8}$$

Then, up to subsequences, u^ε converges (in $L^2(0,T;W^{1,2}(K))$, for all compact sets $K \subset \Omega$) to a weak solution u of problem (2.3.1).

Remark 2.3.7. The above compactness, and so the existence results, do not need that the matrix $a(x)$ be globally Lipschitz in Ω, since a local Lipschitz condition is enough.

Moreover, we point out that similar results could be proved for other kind of equations, including for instance fully nonlinear equations (Bellman operators, etc...). In fact, it is clear that two only ingredients are required in the above construction : a global L^∞ bound (typically ensured by the bound on $\|H(t,x,0)\|_\infty$) and a local compactness and stability for equi-bounded solutions. For instance, in the fully nonlinear case this may be achieved in the topology of uniform convergence in order to build a viscosity solution inside.

2.3.2 Uniqueness of solutions

Now we prove that the HJB equation (2.3.1) has a unique solution if the invariance condition holds. The strategy is classical and relies on the existence of a blow-up supersolution and the convexity of H ; a similar principle can be found e.g. in [79, Lemma 6].

Theorem 2.3.8. *Suppose $G \in L^\infty(\Omega)$. Assume that $a(x)$ satisfies (2.2.3), that $H(t,x,p)$ satisfies (2.2.5)-(2.2.7) and that the invariance condition (2.2.9) holds true.*

Then there is at most one bounded weak solution of the problem (2.3.1).

Démonstration. Let u, v be two bounded solutions of (2.3.1). For $\varepsilon > 0$, we set

$$v_\varepsilon = v + \varepsilon^2(M - \log d(x)) + \varepsilon\sqrt{T-t}\,.$$

A straightforward computation implies

$$\begin{aligned} &-(v_\varepsilon)_t - \operatorname{tr}(a(x)D^2 v_\varepsilon) + H(t,x,Dv_\varepsilon) = -v_t - \operatorname{tr}(a(x)D^2 v) + H(t,x,Dv_\varepsilon) \\ &\quad + \varepsilon^2\left(\frac{\operatorname{tr}(a(x)D^2 d(x)))}{d(x)} - \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)^2}\right) + \frac{\varepsilon}{2\sqrt{T-t}}\,. \end{aligned}$$

By convexity of H, we have

$$\begin{aligned} H(t,x,Dv_\varepsilon) &\geq H(t,x,Dv) + H_p(t,x,Dv)\cdot(Dv_\varepsilon - Dv) \\ &= H(t,x,Dv) - \varepsilon^2\,\frac{H_p(t,x,Dv)\cdot Dd(x)}{d(x)} \end{aligned}$$

so we deduce

$$\begin{aligned} &-(v_\varepsilon)_t - \operatorname{tr}(a(x)D^2 v_\varepsilon) + H(t,x,Dv_\varepsilon) \geq -v_t - \operatorname{tr}(a(x)D^2 v) + H(t,x,Dv) \\ &+ \varepsilon^2\left(\frac{\operatorname{tr}(a(x)D^2 d(x))) - H_p(t,x,Dv)\cdot Dd(x)}{d(x)} - \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)^2}\right) + \frac{\varepsilon}{2\sqrt{T-t}}\,. \end{aligned} \tag{2.3.9}$$

Using assumption (2.2.9), we have that there exists $\delta > 0$ such that

$$\frac{\mathrm{tr}(a(x)D^2d(x))) - H_p(t,x,Dv)\cdot Dd(x)}{d(x)} - \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)^2} \geq -C \qquad (2.3.10)$$

for all $x \in \Gamma_\delta$. However, due to parabolic regularity (see e.g. *Theorem V.3.1* of [74]), we know that each solution of (2.3.1) is locally Lipschitz in the space variable, and in particular, thanks to the global L^∞ bound of the solutions, we have

$$|Dv(t,x)| \leq \frac{C_\delta}{\sqrt{T-t}} \qquad \forall (t,x) : \, t \in (0,T), \, d(x) \geq \delta \, .$$

Because of (2.2.7), we deduce that

$$|H_p(t,x,Dv)| \leq C_\delta \left(1 + \frac{1}{\sqrt{T-t}}\right) \forall (t,x) : \, t \in (0,T), \, d(x) \geq \delta \, .$$

Therefore, since $d(x)$ is a smooth extension of the distance function, the inequality (2.3.10) extends to the whole domain as follows :

$$\frac{\mathrm{tr}(a(x)D^2d(x))) - H_p(t,x,Dv)\cdot Dd(x)}{d(x)} - \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)^2} \geq -\frac{C}{\sqrt{T-t}} \qquad (2.3.11)$$

for any $(t,x) \in Q_T$, for a possibly different constant C. Finally, using (2.3.11) and the fact that v is a super solution, we deduce from (2.3.9)

$$-(v_\varepsilon)_t - \mathrm{tr}(a(x)D^2 v_\varepsilon) + H(t,x,Dv_\varepsilon) \geq -\frac{C\,\varepsilon^2}{\sqrt{T-t}} + \frac{\varepsilon}{2\sqrt{T-t}} \geq 0$$

provided ε is sufficiently small.

Thus, v_ε is a super solution and clearly $u - v_\varepsilon < 0$ in a neighborhood of $\partial\Omega$ since u, v are bounded while $\log d(x) \to -\infty$. For a convenient choice of M, we also have that $v_\varepsilon(T) \geq v(T) \geq u(T)$. Therefore, u and v_ε are a pair of sub and super solution in any subset $\Omega \setminus \Gamma_\eta$ and $u \leq v_\varepsilon$ on $\partial(\Omega \setminus \Gamma_\eta)$ if η is sufficiently small. We can apply e.g. Proposition 2.1 in [92] [1] to conclude that $u \leq v_\varepsilon$ in $\Omega \setminus \Gamma_\eta$. Letting $\eta \to 0$, we get $u \leq v_\varepsilon$ in $(0,T) \times \Omega$. As $\varepsilon \to 0$, we obtain $u \leq v$. Reversing the roles of u, v, we conclude with the uniqueness of solutions. □

The above uniqueness result yields important consequences in terms of stability of solutions. Namely, under the invariance condition all different approximations, as those suggested in the previous subsection, converge towards the same solution.

1. even if [92, Proposition 2.1] is written with $a(x) = I$, the same proof applies without any change to the case of a bounded coercive matrix $a(x)$.

Corollary 2.3.9. *Let the assumptions of Theorem 2.3.8 hold true. Then, given any sequences $a_\varepsilon, H_\varepsilon, G_\varepsilon$ satisfying the hypotheses of Proposition 2.3.4, the (whole) sequence u_ε of solutions of (2.3.5) converges to the unique weak solution u of (2.3.1). In addition, u is also the limit of the (whole) sequence of solutions of problem (2.3.8).*

2.4 The Fokker-Planck equation

In this section we turn the attention to the Fokker-Planck equation under the invariance conditions. So we consider the following equation

$$\begin{cases} m_t - \sum_{i,j} \partial^2_{ij}(a_{ij}(x)m) - \operatorname{div}(m\, b(t,x)) = 0 & \text{in } Q_T, \\ m(0) = m_0 & \text{in } \Omega, \end{cases} \tag{2.4.1}$$

where $b : [0,T] \times \overline{\Omega} \to \mathbb{R}^N$ is a vector field which is locally bounded in $[0,T] \times \Omega$ (i.e. it is bounded in $(0,T) \times K$, for any compact set $K \subset \Omega$).

Definition 2.4.1. *Let $m \in L^1([0,T] \times \Omega)$. We say that m is a weak solution of* (2.4.1) *if*

(i) $m \in C([0,T]; L^1(\Omega))$, $m \geq 0$;

(ii) For each $\phi \in C([0,T]; L^1(\Omega)) \cap L^\infty([0,T] \times \Omega)$ such that ϕ satisfies

$$\begin{cases} -\phi_t - \operatorname{tr}(a(x)D^2\phi) + b \cdot D\phi \in L^\infty([0,T] \times \Omega) \\ \phi(T) = 0 \end{cases}$$

in the sense of Definition 2.3.1, the weak formulation holds :

$$\int_0^T \int_\Omega m \left(-\phi_t - \operatorname{tr}(a(x)D^2\phi)) + b \cdot D\phi\right) dxdt = \int_\Omega m_0 \phi(0) dx\,.$$

Since a is assumed to be Lipschitz, here we also have

$$\sum_{i,j} \partial^2_{ij}(a_{ij}(x)m) = \operatorname{div}(a^*(x)\, Dm) + \operatorname{div}(m\tilde{b}(x))\,,$$

where $\tilde{b}(x)$ is defined by (2.3.2). So, there is no loss of generality in considering only divergence form operators :

$$\begin{cases} m_t - \operatorname{div}(a^*(x)Dm) - \operatorname{div}(m\, b(t,x)) = 0 \\ m(0) = m_0 \end{cases} \tag{2.4.2}$$

Weak solutions of (2.4.2) are defined by duality exactly as in Definition 2.4.1. In other words, m is a weak solution of (2.4.1) if and only if it is a weak solution of (2.4.2) with b replaced by $b - \tilde{b}$. Hereafter, we will deal with problem (2.4.2), where the adjoint matrix a^* appears in the divergence operator, in order to keep consistency with the dual HJB equation considered before.

2.4.1 Existence of solutions

As for the existence of solutions of Hamilton-Jacobi equation, we reason through approximation and use compactness arguments. We start with the following lemma, whose proof is postponed to the Appendix.

Lemma 2.4.2. *Let $\{\Omega_\varepsilon\}$ be a sequence of domains such that $\Omega_\varepsilon \subseteq \Omega_\eta \subseteq \Omega$ for $\varepsilon > \eta$, and $\bigcup_\varepsilon \Omega_\varepsilon = \Omega$. Let $b_\varepsilon \in L^\infty(0,T; L^\infty_{loc}(\Omega))$ be a sequence such that, for every compact set $K \subset \Omega$, b_ε is bounded in $(0,T) \times K$ uniformly in ε. Assume that a_ε is a sequence of matrices which is uniformly bounded and, locally, uniformly coercive. Let m_ε be a solution of*

$$(m_\varepsilon)_t - \operatorname{div}(a^*_\varepsilon(x) Dm_\varepsilon + b_\varepsilon\, m_\varepsilon) = 0 \quad \text{in } (0,T) \times \Omega_\varepsilon\,, \tag{2.4.3}$$

such that $\|m_\varepsilon(t)\|_{L^1(\Omega)}$ is uniformly bounded with respect to ε and $t \in [0,T]$.

Then there exists a function $m \in L^\infty((0,T); L^1(\Omega))$ and a subsequence m_ε such that

$$m_\varepsilon \to m \qquad \text{in } L^1((0,T) \times K)$$

for all compact sets $K \subset \Omega$. Moreover, if, for some m_0, $m_\varepsilon(0) \to m_0$ in $L^1_{loc}(\Omega)$ and, for some $a(x)$, $b(t,x)$, we have $a_\varepsilon(x) \to a(x)$ and $b_\varepsilon(t,x) \to b(t,x)$ almost everywhere in Q_T, then we also have

$$m_\varepsilon \to m \qquad \text{in } C^0([0,T]; L^1(K))$$

for all compact sets $K \subset \Omega$.

Démonstration. See the Appendix. □

The next step says that, under the invariance condition, there is a *global* L^1 stability.

Proposition 2.4.3. *Let $\{\Omega_\varepsilon\}$ be a sequence of domains such that $\Omega_\varepsilon \subseteq \Omega_\eta \subseteq \Omega$ for $\varepsilon > \eta$, and $\bigcup_\varepsilon \Omega_\varepsilon = \Omega$. Let $m_0 \in L^1(\Omega)$, $m_0 \geq 0$. Assume that $b_\varepsilon \in L^\infty([0,T] \times \Omega_\varepsilon)$ is such that*

$$\begin{cases} \forall \text{ compact } K \subset \Omega,\ b_\varepsilon \text{ is uniformly bounded in } (0,T) \times K, \\ b_\varepsilon(t,x) \to b(t,x) \quad \text{a.e. in } Q_T. \end{cases} \tag{2.4.4}$$

Let $a_\varepsilon(x)$ be a sequence of locally uniformly coercive and Lipschitz matrices such that $a_\varepsilon(x) \to a(x)$ almost everywhere in Ω.

We call $d_\varepsilon(x) = d_{\Omega_\varepsilon}(x)$ and we assume that b_ε satisfies the following condition in Ω_ε : there exist $\delta_0, C > 0$ and two sequences $r_\varepsilon, h_\varepsilon \to 0$ such that

$$\operatorname{div}(a_\varepsilon(x)Dd_\varepsilon(x)) - b_\varepsilon(t,x) \cdot Dd_\varepsilon(x) \geq \frac{a_\varepsilon(x)Dd_\varepsilon(x) \cdot Dd_\varepsilon(x)}{d_\varepsilon(x) + r_\varepsilon} - C(d_\varepsilon(x) + h_\varepsilon) \qquad (2.4.5)$$

for all $\varepsilon > 0$ and a.e. $t \in (0,T)$, $x \in \Gamma_{\delta_0} \cap \Omega_\varepsilon$.

Let m_ε be the solution, in Ω_ε, of the Neumann problem

$$\begin{cases} (m_\varepsilon)_t - \operatorname{div}(a_\varepsilon^*(x)Dm_\varepsilon) - \operatorname{div}(m_\varepsilon\, b_\varepsilon) = 0 & (t,x) \in (0,T) \times \Omega_\varepsilon\,, \\ m_\varepsilon(0) = m_0 & \\ (a_\varepsilon^*(x)Dm_\varepsilon + b_\varepsilon\, m_\varepsilon) \cdot \nu_{|\partial\Omega_\varepsilon} = 0\,. & \end{cases} . \qquad (2.4.6)$$

Then there exists $m \in L^1(\Omega)$ such that (defining $m_\varepsilon = 0$ in $\Omega \setminus \Omega_\varepsilon$) we have, up to a subsequence,

$$m_\varepsilon \to m \qquad \text{in } C^0([0,T]; L^1(\Omega))$$

and m is a weak solution of problem (2.4.2).

Remark 2.4.4. Despite the above statement is given in a general version, the reader should keep in mind at least two typical examples of approximations. The first one occurs if $a(x)$ degenerates on the boundary in the normal direction, i.e. if $a(x)Dd(x) \cdot Dd(x) = 0$ on $\partial\Omega$, and if $b(t,x)$ is bounded in Q_T and the invariance property (1.1.2) holds true. In this case we can use this result with $\Omega_\varepsilon = \Omega$, $b_\varepsilon = b$ and $a_\varepsilon(x) = a(x) + \varepsilon I$. Then (2.4.5) is satisfied provided $\varepsilon = o(r_\varepsilon)$ and with $h_\varepsilon \simeq \frac{\varepsilon}{r_\varepsilon}$. A second example occurs if the drift b is unbounded near the boundary, which is certainly the case whenever $a(x)$ does not degenerate and the invariance condition (1.1.2) holds. In this case one needs to work on internal domains and the above result can be used with $\Omega_\varepsilon = \{x : d(x) > \varepsilon\}$, $b_\varepsilon = b$, $a_\varepsilon = a$ and $r_\varepsilon = h_\varepsilon = \varepsilon$ (see Theorem 2.4.6 below).

Démonstration. For simplicity, we divide the proof into steps.

Step 1 : local convergence. Hereafter, we extend m_ε to Ω defining $m_\varepsilon = 0$ in $\Omega \setminus \Omega_\varepsilon$. Integrating the equation in (2.4.6), one has immediately, for each $t \in [0,T]$,

$$\int_\Omega m_\varepsilon(t)dx = \int_{\Omega_\varepsilon} m_0\, dx \to \int_\Omega m_0\, dx\,. \qquad (2.4.7)$$

Moreover, by the maximum principle, we have $m_\varepsilon \geq 0$ in Q_T.

We use Lemma 2.4.2 to deduce that m_ε is relatively compact and, up to a subsequence, converges to some $m \in L^1(\Omega)$; the convergence holds almost everywhere in Q_T and, in addition, in $C^0([0,T]; L^1(K)$ for every compact subset $K \subset \Omega$.

Step 2 : global L^1 convergence. Now we want to prove that, for every $t \in [0,T]$,

$$m_\varepsilon(t) \to m(t) \qquad \text{strongly in } L^1(\Omega)\,.$$

Since $m_\varepsilon \geq 0$, it suffices to prove that, for every $t \in [0,T]$,

$$m_\varepsilon(t) \to m(t) \quad \text{a.e. in } \Omega \quad \text{and} \quad \int_\Omega m_\varepsilon(t)dx \to \int_\Omega m(t)dx\,. \tag{2.4.8}$$

The almost everywhere convergence of $m_\varepsilon(t)$ to $m(t)$ is already given by Lemma 2.4.2 (up to a subsequence, which is not relabeled). For the convergence of the integrals, by (2.4.7) we have to prove that

$$\int_\Omega m(t)dx = \int_\Omega m_0\, dx\,. \tag{2.4.9}$$

By Fatou's lemma and (2.4.7) we have

$$\int_\Omega m(t)dx \leq \int_\Omega m_0\, dx\,, \tag{2.4.10}$$

which in particular implies that $m \in L^1(\Omega)$. To prove the reverse inequality, for $\delta > 0$ we consider the the auxiliary function

$$\phi_\varepsilon = \log(d_\varepsilon(x) + \delta) - \log\delta\,. \tag{2.4.11}$$

Of course we have that $\phi_\varepsilon \to \phi := \log(d(x)+\delta) - \log\delta$. We use ϕ_ε as a test function in the equation of (2.4.6). Using the Neumann condition for m_ε and that $a_\varepsilon D\phi_\varepsilon \cdot \nu \leq 0$ on $\partial\Omega_\varepsilon$, integrating by parts twice we obtain

$$\int_{\Omega_\varepsilon} m_\varepsilon(t)\phi_\varepsilon(t)\, dx + \int_0^t \int_{\Omega_\varepsilon} m_\varepsilon(-\mathrm{div}(a_\varepsilon(x)D\phi_\varepsilon) + b_\varepsilon\, D\phi_\varepsilon)\, dxds \geq \int_{\Omega_\varepsilon} m_0\phi_\varepsilon(0)\, dx\,. \tag{2.4.12}$$

Computing the gradient of ϕ_ε we get, for ε sufficiently small,

$$\begin{aligned}
&\int_0^t \int_{\Omega_\varepsilon} m_\varepsilon(\mathrm{div}(a_\varepsilon(x)D\phi_\varepsilon) - b_\varepsilon\, D\phi_\varepsilon)\, dxds \\
&= \int_0^t \int_{\Omega_\varepsilon} \frac{m_\varepsilon}{d_\varepsilon(x)+\delta} \left\{ (\mathrm{div}(a_\varepsilon(x)Dd_\varepsilon) - b_\varepsilon\, Dd_\varepsilon) - \frac{a_\varepsilon(x)Dd_\varepsilon \cdot Dd_\varepsilon}{d_\varepsilon(x)+\delta} \right\} dxds \\
&\geq \int_0^t \int_{\Omega_\varepsilon} \frac{m_\varepsilon}{d_\varepsilon(x)+\delta} \left\{ \mathrm{div}(a_\varepsilon(x)Dd_\varepsilon) - b_\varepsilon\, Dd_\varepsilon - \frac{a_\varepsilon(x)Dd_\varepsilon \cdot Dd_\varepsilon}{d_\varepsilon(x)+r_\varepsilon} \right\} dxds\,.
\end{aligned}$$

and thanks to assumption (2.4.5) we deduce

$$\int_0^t \int_{\Omega_\varepsilon} m_\varepsilon(\operatorname{div}(a_\varepsilon(x)D\phi_\varepsilon) - b_\varepsilon\, D\phi_\varepsilon)\, dxds \geq -C \int\limits_{\Omega_\varepsilon \cap \Gamma_{\delta_0}} m_\varepsilon \frac{d_\varepsilon(x) + h_\varepsilon}{d_\varepsilon(x) + \delta}\, dxds$$

$$+ \int_0^t \int\limits_{\Omega_\varepsilon \setminus \Gamma_{\delta_0}} \frac{m_\varepsilon}{d_\varepsilon(x) + \delta} \left\{ \operatorname{div}(a_\varepsilon(x)Dd_\varepsilon) - b_\varepsilon\, Dd_\varepsilon - \frac{a_\varepsilon(x)Dd_\varepsilon \cdot Dd_\varepsilon}{d_\varepsilon(x) + r_\varepsilon} \right\} dxds\,.$$

The first integral in the right-hand side is uniformly bounded because $h_\varepsilon \leq \delta$ for small ε. Since b_ε is locally uniformly bounded in Q_T, a_ε is Lipschitz and, for ε sufficiently small, $d_\varepsilon(x) \geq \frac{\delta_0}{2}$ in $\Omega \setminus \Gamma_{\delta_0}$, the second integral is also bounded uniformly. So we conclude that

$$\int_0^t \int_{\Omega_\varepsilon} m_\varepsilon(\operatorname{div}(a_\varepsilon(x)D\phi_\varepsilon) - b_\varepsilon\, D\phi_\varepsilon)\, dxds \geq -C$$

for some $C > 0$ independent of ε, δ. Plugging this estimate into (2.4.12), we get

$$\int_\Omega m_\varepsilon(t)\phi_\varepsilon(t)\, dx \geq \int_{\Omega_\varepsilon} m_0 \phi_\varepsilon(0)\, dx - C\,. \tag{2.4.13}$$

Now we observe that the integral in the left side converges : indeed, for any $\eta > 0$ we have (we call $\Gamma_\eta^\varepsilon = \{x \,:\, d_\varepsilon(x) < \eta\}$)

$$\int_\Omega |m_\varepsilon(t)\phi_\varepsilon(t) - m(t)\phi(t)|\, dx \leq \int_{\Gamma_\eta} |m_\varepsilon(t)\phi_\varepsilon(t) - m(t)\phi(t)|\, dx$$

$$+ \int_{\Omega \setminus \Gamma_\eta} |m_\varepsilon(t)\phi_\varepsilon(t) - m(t)\phi(t)| dx$$

$$\leq C \log\left(\frac{\eta + \delta}{\delta}\right) + \int_{\Omega \setminus \Gamma_\eta} |m_\varepsilon(t)\phi_\varepsilon(t) - m(t)\phi(t)| dx\,.$$

We let first $\varepsilon \to 0$, using the L^1_{loc} convergence of m_ε, and then we let $\eta \to 0$, so that last two terms will vanish. Hence we deduce

$$\int_\Omega m_\varepsilon(t)\phi_\varepsilon(t)\, dx \to \int_\Omega m(t)\, \phi(t)\, dx\,.$$

Therefore we obtain from (2.4.13), letting $\varepsilon \to 0$,

$$\int_\Omega m(t)\phi(t)\, dx \geq \int_\Omega m_0 \phi(0)\, dx - C\,.$$

Since $\phi = \log(d(x) + \delta) - \log \delta \leq |\log \delta| + c$ we deduce that

$$\int_\Omega m(t)\, dx \geq \int_\Omega m_0 \frac{\log(d(x) + \delta) - \log \delta}{|\log \delta|}\, dx - \frac{C}{|\log \delta|}\,.$$

We now let $\delta \to 0$; using Lebesgue's theorem and since $m_0 \in L^1(\Omega)$, we get

$$\int_\Omega m(t)\, dx \geq \int_\Omega m_0\, dx\,.$$

Thus (2.4.9) is proved, we have (2.4.8) and with that we conclude that $m_\varepsilon(t) \to m(t)$ in $L^1(\Omega)$, for all $t > 0$. By Lebesgue theorem, we also deduce the convergence of m_ε to m in $L^1(Q_T)$.

Step 3 : convergence in $C^0([0,T]; L^1(\Omega))$.

First we observe that $m \in C([0,T]; L^1(\Omega))$, with a similar argument as used above. Indeed, let $t_n \to t$; since $m \in C([0,T]; L^1_{loc}(\Omega))$ we have that $m(t_n)$ always admits a subsequence converging to $m(t)$ a.e. in Ω. Since $\int_\Omega m(t_n)\, dx = \int_\Omega m_0\, dx = \int_\Omega m(t)\, dx$, we deduce again that $m(t_n) \to m(t)$ in $L^1(\Omega)$.

Since $m \in C^0([0,T]; L^1(\Omega))$, a compactness argument implies that

$$\lim_{|E| \to 0} \sup_{t \in [0,T]} \int_E m(t)\, dx = 0\,. \tag{2.4.14}$$

Using positive and negative parts, i.e. $s = s^+ - s^-$, $|s| = s^+ + s^-$, we split

$$\begin{aligned} \int_\Omega |m_\varepsilon(t) - m(t)|\, dx &= \int_\Omega (m_\varepsilon(t) - m(t))\, dx + 2 \int_\Omega (m_\varepsilon(t) - m(t))^-\, dx \\ &= - \int_{\Omega \setminus \Omega_\varepsilon} m_0\, dx + 2 \int_\Omega (m_\varepsilon(t) - m(t))^-\, dx \end{aligned} \tag{2.4.15}$$

because of mass conservation. Last integral is restricted to where $m_\varepsilon(t) \leq m(t)$. So, we split once more

$$\begin{aligned} \int_\Omega (m_\varepsilon(t) - m(t))^-\, dx &\leq \int_{\Omega \setminus \Gamma_\eta} (m_\varepsilon(t) - m(t))^-\, dx + 2 \int_{\Gamma_\eta} m(t)\, dx \\ &\leq \int_{\Omega \setminus \Gamma_\eta} |m_\varepsilon(t) - m(t)|\, dx + 2 \int_{\Gamma_\eta} m(t)\, dx \end{aligned}$$

which yields

$$\sup_{t \in [0,T]} \int_\Omega (m_\varepsilon(t) - m(t))^-\, dx \leq \sup_{t \in [0,T]} \int_{\Omega \setminus \Gamma_\eta} |m_\varepsilon(t) - m(t)|\, dx + 2 \sup_{t \in [0,T]} \int_{\Gamma_\eta} m(t)\, dx\,.$$

Now recall that $m_\varepsilon \to m$ in $C^0([0,T]; L^1(K))$, for any compact subset K. So, when we let $\varepsilon \to 0$ the first term in the right-hand side vanishes and we get

$$\lim_{\varepsilon \to 0} \sup_{t \in [0,T]} \int_\Omega (m_\varepsilon(t) - m(t))^-\, dx \leq 2 \sup_{t \in [0,T]} \int_{\Gamma_\eta} m(t)\, dx\,.$$

Finally, we let $\eta \to 0$ and we use (2.4.14) in the last term, and we conclude that

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \int_\Omega (m_\varepsilon(t) - m(t))^- \, dx = 0 \, .$$

Then from (2.4.15) we deduce that

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \int_\Omega |m_\varepsilon(t) - m(t)| \, dx = 0$$

that is $m_\varepsilon \to m$ in $C^0([0,T];L^1(\Omega))$.

Step 4 : Conclusion. We take $\phi \in C([0,T];L^1(\Omega)) \cap L^\infty([0,T]\times\Omega)$ such that ϕ satisfies

$$\begin{cases} -\phi_t - \mathrm{div}(a(x)D\phi) + b\cdot D\phi \in L^\infty([0,T]\times\Omega) \\ \phi(T) = 0 \end{cases}$$

in the weak sense. Let us consider the solution ϕ_ε of the following problem

$$\begin{cases} -(\phi_\varepsilon)_t - \mathrm{div}(a_\varepsilon(x)D\phi_\varepsilon) + b_\varepsilon \cdot D\phi_\varepsilon = f & \text{in } (0,T)\times\Omega_\varepsilon \\ a_\varepsilon(x)D\phi_\varepsilon\cdot\nu = 0 & \text{in } (0,T)\times\partial\Omega_\varepsilon \\ \phi_\varepsilon(T) = 0 \end{cases}$$

$$\text{where} \quad f := -\phi_t - \mathrm{div}(a(x)D\phi) + bD\phi \, .$$

As always, we set $\phi_\varepsilon := 0$ on $\Omega\setminus\Omega_\varepsilon$.

Applying Corollary 2.3.9, we have that ϕ_ε converges to ϕ in $L^2(0,T;W^{1,2}(K))$ and in $C^0([0,T];L^p(\Omega))$ for all $p<\infty$. Now, taking ϕ_ε as test function in (2.4.6) we get

$$-\int_\Omega m_0\phi_\varepsilon(0)\,dx + \int_0^T\int_\Omega m_\varepsilon(-\phi_t - \mathrm{div}(a(x)D\phi) + b\cdot D\phi)\,dxdt = 0\,.$$

Since $\phi_\varepsilon(0) \to \phi(0)$ a.e. in Ω, we can pass to the limit in the first term with Lebesgue's theorem. In the second term, we use the L^1 convergence of m_ε towards m. Finally, we obtain that m satisfies

$$\int_0^T\int_\Omega m(-\phi_t - \mathrm{div}(a(x)D\phi) + bD\phi)\,dxdt = \int_\Omega m_0\phi(0)\,dx\,.$$

□

Finally, we conclude with the existence part.

Theorem 2.4.5. *Let $m_0 \in L^1(\Omega)$, $m_0 \geq 0$. Let $a \in W^{1,\infty}(\Omega)$ satisfy (2.2.3). Assume that $b \in L^\infty(0,T;L^\infty_{loc}(\Omega))$ and that there exist $\delta_0, C > 0$ such that the following inequality holds :*

$$\operatorname{div}(a(x)Dd(x)) - b(t,x) \cdot Dd(x) \geq \frac{a(x)Dd(x) \cdot Dd(x)}{d(x)} - C\, d(x) \tag{2.4.16}$$

for almost every $t \in (0,T)$ and $x \in \Gamma_{\delta_0}$. Then there exists a solution of problem (2.4.2)*.*

Démonstration. For each ε we consider $a_\varepsilon = a$, $b_\varepsilon = b$ and $\Omega_\varepsilon = \{d(x) > \varepsilon\}$. It is immediate to check that the assumptions of Proposition 2.4.3 are satisfied : indeed, since $d_\varepsilon(x) = d(x) - \varepsilon$ near $\partial\Omega$, we have in Ω_ε

$$\begin{aligned}\operatorname{div}(a_\varepsilon(x)Dd_\varepsilon(x)) - b_\varepsilon(t,x) \cdot Dd_\varepsilon(x) &= \operatorname{div}(a(x)Dd(x)) - b(t,x) \cdot Dd(x) \\ \geq \frac{a(x)Dd(x) \cdot Dd(x)}{d(x)} - C\, d(x) &= \frac{a_\varepsilon(x)Dd_\varepsilon(x) \cdot Dd_\varepsilon(x)}{d_\varepsilon(x) + \varepsilon} - C\, d_\varepsilon(x) - C\varepsilon\end{aligned}$$

which gives (2.4.5). So by solving the approximating problems (2.4.6) and passing to the limit, thanks to Proposition 2.4.3, we obtain the existence of a solution. □

2.4.2 Uniqueness of solutions

The uniqueness of solutions for the Fokker-Planck equation comes easily from the existence of solutions of the Hamilton-Jacobi-Bellman equation.

Theorem 2.4.6. *Let $m_0 \in L^1(\Omega)$, $m_0 \geq 0$. Assume that $a(\cdot) \in W^{1,\infty}(\Omega)$ satisfies* (2.2.3)*. Assume also that $b \in L^\infty(0,T;L^\infty_{loc}(\Omega))$ and that (2.4.16) holds true. Then there exists a unique weak solution of the Fokker-Planck equation* (2.4.2)*.*

Démonstration. Let m_1 and m_2 be two solutions of (2.4.2). Then $m := m_1 - m_2$ solves in the weak sense

$$\begin{cases} m_t - \operatorname{div}(a^*(x)Dm) - \operatorname{div}(m\, b) = 0 \\ m(0) = 0\,. \end{cases}$$

Now, we take ϕ as the solution of the Hamilton-Jacobi equation

$$\begin{cases} -\phi_t - \operatorname{div}(a(x)D\phi) + bD\phi = \operatorname{sgn}(m) \\ \phi(T) = 0\,, \end{cases}$$

where $\operatorname{sgn}(m) = m/|m|\, \mathbb{1}_{\{m \neq 0\}}$. We use ϕ as test function in the weak formulation of m, obtaining

$$\int_0^T \int_\Omega |m|\, dxdt = \int_0^T \int_\Omega m \operatorname{sgn}(m)\, dxdt = 0\,.$$

So $m \equiv 0$ and the proof is concluded. □

In the end, from the equivalence between the two problems (2.4.1) and (2.4.2), we have proved Theorem 2.2.3.

2.5 The Mean Field Game system

We are now ready to study the mean field game system (2.1.1) under the invariance conditions. For convenience, we rewrite here the system, which reads as

$$\begin{cases} -\partial_t u - \sum\limits_{i,j} a_{ij}(x)\partial^2_{ij} u + H(t,x,Du) = F(t,x,m)\,, & (t,x) \in (0,T)\times\Omega \\ u(T) = G(x,m(T)) \end{cases} \tag{2.5.1}$$

$$\begin{cases} \partial_t m - \sum\limits_{i,j} \partial^2_{ij}(a_{ij}(x)m) - \operatorname{div}(mH_p(t,x,Du)) = 0\,, & (t,x) \in (0,T)\times\Omega \\ m(0) = m_0\,. \end{cases} \tag{2.5.2}$$

Let us recall that the matrix $a(x)$ satisfies (2.2.3), the Hamiltonian H satisfies assumptions (2.2.5)-(2.2.7) and that the invariance condition (2.2.9) holds true. The nonlinearities F, G satisfy conditions (2.2.11), (2.2.12). This implies that, for any given $m \in C([0,T]\times L^1(\Omega))$, the HJB equation has a unique solution thanks to Theorem 2.3.5 and Theorem 2.3.8. Conversely, for every u which is locally Lipschitz, the growth condition (2.2.7) guarantees that $H_p(t,x,Du)$ is a locally bounded vector field, so the FP equation has a unique solution given by Theorem 2.4.6. This justifies our definition below.

Definition 2.5.1. *We say that a couple $(u,m) \in L^\infty([0,T]\times\Omega)\times C([0,T]\times L^1(\Omega))$ is a weak solution of the system* (2.1.1) *if u is a solution of the Hamilton-Jacobi-Bellman equation* (2.5.1) *in the sense of Definition 2.3.1 and m is a solution of the Fokker-Planck equation* (2.5.2) *in the sense of Definition 2.4.1.*

2.5.1 Existence of solutions

Here we prove the existence of a solution to the mean field game system.

Theorem 2.5.2. *Assume that hypotheses (2.2.3), (2.2.5)-(2.2.7), (2.2.9), (2.2.11) and (2.2.12) hold true. Then there exists at least one solution (u,m) of* (2.5.1)*-(2.5.2), in the sense of Definition 2.5.1.*

Démonstration. For each $\varepsilon > 0$, we define $\Omega_\varepsilon = \{d(x) > \varepsilon\}$ and $(u_\varepsilon, m_\varepsilon)$ as the solution, in $[0,T] \times \Omega_\varepsilon$, of the mean-field game system

$$\begin{cases} -\partial_t u_\varepsilon - \mathrm{tr}(a(x)D^2 u_\varepsilon) + H(t,x,Du_\varepsilon) = F(x,m_\varepsilon)\,, & (t,x) \in (0,T) \times \Omega_\varepsilon \\ \partial_t m_\varepsilon - \mathrm{div}(a^*(x)Dm_\varepsilon) - \mathrm{div}(m_\varepsilon(H_p(t,x,Du_\varepsilon) + \tilde{b}(x)) = 0 & (t,x) \in (0,T) \times \Omega_\varepsilon \\ m_\varepsilon(0) = m_0\,, \qquad u_\varepsilon(T) = G(x, m_\varepsilon(T)) \\ a(x)Du_\varepsilon \cdot \nu_{|\partial\Omega_\varepsilon} = 0 \qquad \Big[a(x)Dm_\varepsilon + m_\varepsilon(\tilde{b}(x) + H_p(t,x,Du_\varepsilon))\Big] \cdot \nu_{|\partial\Omega_\varepsilon} = 0\,. \end{cases} \tag{2.5.3}$$

As before, we extend the solutions to the whole of Ω by setting $u_\varepsilon = m_\varepsilon = 0$ in $\Omega \setminus \Omega_\varepsilon$.

By conservation of mass, we have that $\int_{\Omega_\varepsilon} m_\varepsilon(t) = \int_\Omega m_0\,dx$ for all $t \in (0,T)$, and additionally $m_\varepsilon \geq 0$. Then, assumptions (2.2.11), (2.2.12) imply that $F(x,m_\varepsilon)$ and $G(x,m_\varepsilon(T))$ are uniformly bounded. By maximum principle, we deduce that $\|u_\varepsilon\|_\infty$ is uniformly bounded. Applying Lemma 2.3.3, we deduce that there exists a function $u \in L^\infty(Q_T) \cap L^2(0,T;W^{1,2}_{loc}(\Omega))$ and a subsequence u_ε converging to u weakly in $L^2(0,T;W^{1,2}(K))$, for all compact sets $K \subset \Omega$. Moreover, the assumption upon G, the global bound on u_ε and the natural growth conditions ensure that the sequence Du_ε is also bounded in any set $(0,T) \times K$, for K compact. This allows us to use Lemma 2.4.2 for m_ε. In fact, since $H_p(t,x,Du_\varepsilon)$ is locally bounded and converges a.e. to $H_p(t,x,Du)$, and since the invariance condition (2.2.9) holds, we are in the position to apply the stability result of Proposition 2.4.3 as well. Therefore, we conclude that

$$m_\varepsilon \to m \qquad \text{in } C^0([0,T];L^1(\Omega))$$

and m is a solution of (2.5.2). Finally, the continuity assumptions upon F and G now imply that $F(t,x,m_\varepsilon)$ converges almost everywhere to $F(t,x,m)$ in Q_T and $G(x,m_\varepsilon(T))$ converges to $G(x,m(T))$ a.e. and therefore in L^p for all $p < \infty$. We have now access to the stability result of Proposition 2.3.4 and we deduce that the limit function u is a solution of (2.5.1). □

2.5.2 Uniqueness of solutions

Theorem 2.5.3. *Suppose the hypotheses of Theorem 2.5.2 are satisfied and, in addition, F and G are nondecreasing with respect to m, in the sense of operators. If at least one of* *the two following conditions holds :*

$$(i)\quad \begin{cases} \int_\Omega (F(t,x,m_0)-F(t,x,m_1))d(m_0-m_1)=0 \;\Rightarrow\; F(t,x,m_0)=F(t,x,m_1) \\ \int_\Omega (G(x,m_0)-G(x,m_1))d(m_0-m_1)=0 \;\Rightarrow\; G(x,m_0)=G(x,m_1) \end{cases}$$

$$(ii)\quad H(t,x,p_1)-H(t,x,p_2)-H_p(t,x,p_2)(p_1-p_2)=0 \;\Rightarrow\; H_p(t,x,p_1)=H_p(t,x,p_2)\,. \tag{2.5.4}$$

then the solution of (2.1.1) *is unique.*

Démonstration. Let (u,m) and (v,μ) be two solutions of the mean field game system. We want to prove that $v=u$, $\mu=m$.
To do this, we reason as always through approximation. Having defined Ω_ε as in Theorem 2.5.2, we consider $(u_\varepsilon,m_\varepsilon)$ solution of the problem

$$\begin{cases} -\partial_t u_\varepsilon - \mathrm{tr}(a(x)D^2u_\varepsilon)+H(t,x,Du_\varepsilon)=F(t,x,m) & (t,x)\in(0,T)\times\Omega_\varepsilon \\ \partial_t m_\varepsilon - \mathrm{div}(a^*(x)Dm_\varepsilon)-\mathrm{div}(m_\varepsilon(H_p(t,x,Du_\varepsilon)+\tilde b(x))=0 & (t,x)\in(0,T)\times\Omega_\varepsilon \\ m_\varepsilon(0)=m_0 \qquad u_\varepsilon(T)=G(x,m(T)) \\ Du_\varepsilon\cdot\nu_{|\partial\Omega_\varepsilon}=0 \qquad \Big[\varepsilon Dm_\varepsilon+m_\varepsilon(\tilde b(x)+H_p(t,x,Du_\varepsilon))\Big]\cdot\nu_{|\partial\Omega_\varepsilon}=0 \end{cases}. \tag{2.5.5}$$

Similarly, $(v_\varepsilon,\mu_\varepsilon)$ will be the solution of the problem

$$\begin{cases} -\partial_t v_\varepsilon - \mathrm{tr}(a(x)D^2v_\varepsilon)+H(t,x,Dv_\varepsilon)=F(t,x,\mu) & (t,x)\in(0,T)\times\Omega_\varepsilon \\ \partial_t \mu_\varepsilon - \mathrm{div}(a^*(x)D\mu_\varepsilon)-\mathrm{div}(\mu_\varepsilon(H_p(t,x,Dv_\varepsilon)+\tilde b(x))=0 & (t,x)\in(0,T)\times\Omega_\varepsilon \\ \mu_\varepsilon(0)=m_0 \qquad v_\varepsilon(T)=G(x,\mu(T)) \\ Dv_\varepsilon\cdot\nu_{|\partial\Omega_\varepsilon}=0 \qquad \Big[\varepsilon D\mu_\varepsilon+\mu_\varepsilon(\tilde b(x)+H_p(t,x,Dv_\varepsilon))\Big]\cdot\nu_{|\partial\Omega_\varepsilon}=0\,. \end{cases} \tag{2.5.6}$$

Notice that the equations of u_ε and v_ε are *decoupled* from the system, since they do not depend, respectively, upon m_ε and μ_ε. Using Corollary 2.3.9, we know that $u_\varepsilon\to u$ and $v_\varepsilon\to v$ in $C([0,T];L^p(K))$ and in $L^2([0,T];H^1(K))$, for each compact $K\subset\subset\Omega$.

Using this information, and the local gradient bounds, we know that $H_p(x,Du_\varepsilon)$ and $H_p(x,Dv_\varepsilon)$ are locally bounded sequences which converge, respectively, to $H_p(x,Du)$ and $H_p(x,Dv)$. From Proposition 2.4.3 we deduce that $m_\varepsilon\to m$ and $\mu_\varepsilon\to\mu$ in $C([0,T];L^1(\Omega))$.

Now we use the classical monotonicity argument in mean field game systems. We estimate in two different ways the quantity

$$\int_0^T\int_{\Omega_\varepsilon}((u_\varepsilon-v_\varepsilon)(m_\varepsilon-\mu_\varepsilon))_t\;dxdt\,.$$

First, computing directly the time integral we find

$$\int_0^T \int_{\Omega_\varepsilon} ((u_\varepsilon - v_\varepsilon)(m_\varepsilon - \mu_\varepsilon))_t \; dxdt = \int_{\Omega_\varepsilon} (G(x, m(T)) - G(x, \mu(T)))(m_\varepsilon(T) - \mu_\varepsilon(T)) \, dx \, .$$

Besides, if we use the weak formulations of u_ε, v_ε, m_ε, μ_ε, we obtain

$$\int_0^T \int_{\Omega_\varepsilon} ((u_\varepsilon - v_\varepsilon)(m_\varepsilon - \mu_\varepsilon))_t \; dxdt = -\int_0^T \int_{\Omega_\varepsilon} (F(x, m) - F(x, \mu))(m_\varepsilon - \mu_\varepsilon) \, dxdt -$$
$$- \int_0^T \int_{\Omega_\varepsilon} m_\varepsilon (H(t, x, Dv_\varepsilon) - H(t, x, Du_\varepsilon) + H_p(x, Du_\varepsilon)(Dv_\varepsilon - Du_\varepsilon)) \, dxdt -$$
$$- \int_0^T \int_{\Omega_\varepsilon} \mu_\varepsilon (H(t, x, Du_\varepsilon) - H(t, x, Dv_\varepsilon) + H_p(x, Dv_\varepsilon)(Du_\varepsilon - Dv_\varepsilon)) \, dxdt$$

which yields

$$\int_{\Omega_\varepsilon} [G(x, m(T)) - G(x, \mu(T))] \, (m_\varepsilon(T) - \mu_\varepsilon(T)) \, dx + \int_0^T \int_{\Omega_\varepsilon} (F(x, m) - F(x, \mu))(m_\varepsilon - \mu_\varepsilon) \, dxdt$$
$$+ \int_0^T \int_{\Omega_\varepsilon} m_\varepsilon (H(t, x, Dv_\varepsilon) - H(t, x, Du_\varepsilon) + H_p(x, Du_\varepsilon)(Dv_\varepsilon - Du_\varepsilon)) \, dxdt$$
$$+ \int_0^T \int_{\Omega_\varepsilon} \mu_\varepsilon (H(t, x, Du_\varepsilon) - H(t, x, Dv_\varepsilon) + H_p(x, Dv_\varepsilon)(Du_\varepsilon - Dv_\varepsilon)) \, dxdt \leq 0 \, .$$

Since H is convex, we can apply Fatou's lemma in the last two integrals. Moreover, using that (extending the functions to zero outside Ω_ε) $m_\varepsilon(T) - \mu_\varepsilon(T) \to m(T) - \mu(T)$ in $L^1(\Omega)$, we can pass to the limit in the remaining two integrals. We obtain

$$\int_{\Omega} [G(x, m(T)) - G(x, \mu(T))] \, (m(T) - \mu(T)) \, dx + \int_0^T \int_{\Omega} (F(x, m) - F(x, \mu))(m - \mu) \, dxdt$$
$$+ \int_0^T \int_{\Omega} m (H(t, x, Dv) - H(t, x, Du) + H_p(x, Du)(Dv - Du)) \, dxdt$$
$$+ \int_0^T \int_{\Omega} \mu (H(t, x, Du) - H(t, x, Dv) + H_p(x, Dv)(Du - Dv)) \, dxdt \leq 0 \, ,$$

and therefore all integrals must vanish. Now we conclude with assumption (2.5.4). Indeed, if (i) holds we deduce that

$$F(t, x, m) = F(t, x, \mu) \, , \qquad\qquad G(x, m(T)) = G(x, \mu(T)) \, .$$

This means that u and v solve the same Hamilton-Jacobi-Bellman equation. From Theorem 2.3.8, we know that they coincide. So $v = u$, hence $H_p(x, Du) = H_p(x, Dv)$. Coming back to the Fokker-Planck equation, we deduce that $\mu = m$ from Theorem 2.4.6. Otherwise, if (ii) holds, we proceed in the opposite way : first we deduce that $H_p(x, Du) = H_p(x, Dv)$, and then $\mu = m$, which in turn implies that $u = v$ by uniqueness of the HJB equation. □

Remark 2.5.4. We stress that the assumption on the final pay-off G can be relaxed in case that $H(t, x, p)$ is globally Lipschitz continuous with respect to p. Indeed, in this case the drift term $H_p(t, x, Du)$ is always bounded and it is not needed that the range of G be bounded in $W^{1,\infty}(\Omega)$. It would be enough, in this case, to require that the range of G is bounded in $L^\infty(\Omega)$, similar as it is done for the internal coupling F. In particular, this condition would include local couplings, i.e. $G = G(x, r)$ is a bounded real function.

2.6 Further regularity of solutions

In this Section we get an improvement of regularity for u or m with a suitable strengthening of hypotheses.

2.6.1 Lipschitz regularity of the value function

We follow the classical Bernstein method in order to get gradient bounds for the solution of u. The approach is borrowed from [79] and yields the global Lipschitz character of the value function.

Theorem 2.6.1. *Assume that $a(x)$ satisfies (2.2.3) and there exists a matrix $\sigma \in W^{1,\infty}(\Omega)$ such that $a(x) = \sigma(x)\sigma(x)^*$. Let $H \in C^1(Q_T \times \mathbb{R}^N)$ satisfy conditions (2.2.5)-(2.2.7) and, in addition, the following assumption :*

$$H_x(t, x, p) \cdot p \geq -C\,(1 + |p|^2) \qquad \forall (t, x) \in Q_T\,, p \in \mathbb{R}^N \tag{2.6.1}$$

for some constant $C > 0$. Moreover, assume that the invariance condition (2.2.9) holds true. Let F, G satisfy (2.2.11), (2.2.12) and assume that $m \mapsto F(\cdot, m)$ has bounded range in $L^\infty((0,T); W^{1,\infty}(\Omega))$. Then $u \in L^\infty((0,T); W^{1,\infty}(\Omega))$.

Démonstration. Let $\Omega^\varepsilon = \{x : d(x) > \varepsilon\}$. We consider u_ε solution of the problem

$$\begin{cases} -(u_\varepsilon)_t - \operatorname{tr}(a(x)D^2 u_\varepsilon) + H(t, x, Du_\varepsilon) = F(x, m)\,, & (t, x) \in (0,T) \times \Omega^\varepsilon \\ u_\varepsilon(T) = G(x, m(T)) \\ Du_\varepsilon \cdot \nu_{|\partial\Omega^\varepsilon} = 0\,. \end{cases} \tag{2.6.2}$$

We know from Proposition 2.3.6 that $u_\varepsilon \to u$ when $\varepsilon \to 0$, with $Du_\varepsilon \to Du$ a.e. in Q_T.

Let us set $w_\varepsilon := |Du_\varepsilon|^2 e^{\theta(d(x))}$, where $\theta \in C^2(0,1)$ is a bounded function to be defined later. Computing the derivatives of w_ε, we find :

$$\begin{aligned}
Dw_\varepsilon &= e^{\theta(d)}\left(2Du_\varepsilon D^2u_\varepsilon + |Du_\varepsilon|^2\theta'(d)Dd\right) ;\\
D^2w_\varepsilon &= e^{\theta(d)}\left(2D^2u_\varepsilon D^2u_\varepsilon + 2Du_\varepsilon D^3u_\varepsilon + |Du_\varepsilon|^2\theta'(d)D^2d + 4\theta'(d)D^2u_\varepsilon Du_\varepsilon \otimes Dd\right) + \\
&\quad + e^{\theta(d)}\,|Du_\varepsilon|^2[\theta''(d) + (\theta'(d))^2]Dd \otimes Dd
\end{aligned}$$

where $Du_\varepsilon D^3u_\varepsilon = \sum_k (u_\varepsilon)_{x_k}(u_\varepsilon)_{x_i x_j x_k}$.

Thus when we form the equation for w_ε we obtain (see also [79, Lemma 8])

$$-(w_\varepsilon)_t - \mathrm{tr}(a(x)D^2w_\varepsilon) + b_\varepsilon(t,x)Dw_\varepsilon + c_\varepsilon(t,x)w_\varepsilon = r_\varepsilon(t,x)\,,$$

where

$$\begin{aligned}
b_\varepsilon(t,x) &= H_p(x, Du_\varepsilon) + 2\theta'(d)\,(a(x)Dd)\,;\\
c_\varepsilon(t,x) &= \theta'(d)\left(-H_p(x,Du_\varepsilon)\cdot Dd + \mathrm{tr}(a(x)D^2d)\right) + \left(\theta''(d) - \left(\theta'(d)\right)^2\right)\,a(x)Dd\cdot Dd;\\
r_\varepsilon(t,x) &= 2e^{\theta(d)}\left(\mathrm{tr}(\tilde a(u_\varepsilon)D^2u_\varepsilon) - \mathrm{tr}(a(x)D^2u_\varepsilon D^2u_\varepsilon) - H_x(x,Du_\varepsilon)\cdot Du_\varepsilon + DF\,\cdot Du_\varepsilon\right)
\end{aligned}$$

and we denoted $\tilde a(u_\varepsilon)_{i,j} = \sum_k (a_{i,j}(x))_{x_k}(u_\varepsilon)_{x_k}$.

First we estimate the quantity $\mathrm{tr}(\tilde a(u_\varepsilon)D^2u_\varepsilon) - \mathrm{tr}(a(x)D^2u_\varepsilon D^2u_\varepsilon)$. Here, since $a(\cdot) = \sigma(\cdot)\sigma(\cdot)^*$, using Young's inequality we get

$$\begin{aligned}
&\mathrm{tr}(\tilde a(u_\varepsilon)D^2u_\varepsilon) - \mathrm{tr}(a(x)D^2u_\varepsilon D^2u_\varepsilon) =\\
=&\sum_{i,j,k}(a_{ij})_{x_k}(u_\varepsilon)_{x_k}(u_\varepsilon)_{x_i x_j} - \sum_{i,j,k} a_{ij}(u_\varepsilon)_{x_j x_k}(u_\varepsilon)_{x_i x_k} =\\
=&2\sum_{i,j,k,l}(\sigma_\varepsilon)_{jl}((\sigma_\varepsilon)_{il})_{x_k}(u_\varepsilon)_{x_k}(u_\varepsilon)_{x_i x_j} - \sum_k |\sigma_\varepsilon^* D(u_\varepsilon)_{x_k}|^2 \le C|Du_\varepsilon|^2\,.
\end{aligned}$$

Using also (2.6.1) and the condition on F, we estimate

$$r_\varepsilon \le C\,e^{\theta(d)}\left(1 + |Du_\varepsilon|^2\right)\,.$$

Therefore, we have that w_ε satisfies

$$-(w_\varepsilon)_t - \mathrm{tr}(a(x)D^2w_\varepsilon) + b_\varepsilon(t,x)Dw_\varepsilon + (c_\varepsilon(t,x) - C)w_\varepsilon \le C,$$

for a suitable $C > 0$.

Now we estimate c_ε thanks to the invariance condition (2.2.9). Indeed, if $d(x) < \delta_0$, we get

$$c_\varepsilon \geq \left(\frac{\theta'(d)}{d} + \theta''(d) - \big(\theta'(d)\big)^2\right) a(x)Dd \cdot Dd - C\, d\,\theta'(d)\,.$$

Choosing $\theta(d) = d^\gamma$, with $\gamma \in (0,1)$, we get

$$c_\varepsilon \geq \gamma\, d^{\gamma-1}\left(\frac{1}{d} + (\gamma-1)d^{-1} - \gamma\, d^{\gamma-1}\right) a(x)Dd \cdot Dd - C\,\gamma\, d^\gamma$$

hence c_ε is uniformly bounded below. If $d(x) \geq \delta_0$, we recall that u is Lipschitz by elliptic regularity, so c_ε is also bounded from below by a constant possibly dependent on δ_0, but independent from ε. We conclude that w_ε satisfies

$$-(w_\varepsilon)_t - \mathrm{tr}(a(x)D^2 w_\varepsilon) + b_\varepsilon(t,x)Dw_\varepsilon - C\, w_\varepsilon \leq C\,.$$

Since the maximum of w_ε cannot be taken on the boundary due to the Neumann condition (see e.g. [79, Lemma 4]), and since at $t = T$ we use the Lipschitz bound on $G(\cdot, m(T))$, we conclude applying the maximum principle that the maximum of Dw_ε is uniformly bounded in ε. As $\varepsilon \to 0$, this implies that $Du \in L^\infty(Q_T)$. □

Remark 2.6.2. We stress that the above proof may admit some variants which possibly apply to other interesting cases. For instance, assume that the invariance condition is strengthened as follows :

$$\mathrm{tr}(a(x)D^2 d(x)) - H_p(x,p)Dd(x) \geq \frac{a(x)Dd(x)\cdot Dd(x)}{d(x)^{1+\rho}} - Cd(x)\,, \tag{2.6.3}$$

for some $\rho > 0$, and in addition that $a(\cdot) = \sigma(\cdot)\sigma(\cdot)^*$ with

$$|D\sigma(x)|^2 \leq c_0 + c_1|\sigma(x)Dd(x)|^2\, d(x)^{\gamma-2-\rho} \tag{2.6.4}$$

for some $\gamma > 0$. Then the conclusion of Theorem 2.6.1 remains true, and the proof can be easily modified accordingly.

In particular, whenever (2.6.3) is satisfied, this generalization includes the case that $\sigma(x)$ is $\frac{1}{2}$-Hölder continuous with $|\sigma(x)Dd(x)| \geq c\, d(x)^{\frac{1}{2}}$ and $|D\sigma(x)| \leq C\, d(x)^{-\frac{1}{2}}$, in which case (2.6.4) holds for any $\gamma < \rho$. Otherwise, (2.6.3)–(2.6.4) are satisfied if $\sigma(x)$ is β-Hölder continuous, $\beta > \frac{1}{2}$, $|\sigma(x)Dd(x)| \geq c\, d(x)^\beta$ and $|D\sigma(x)| \leq C\, d(x)^{\beta-1}$ and the Hamiltonian satisfies, in a neighborhood of the boundary, that

$$\mathrm{tr}(a(x)D^2 d(x)) - H_p(x,p)Dd(x) \geq c\, d(x)^\eta$$

for some $\eta < 2\beta - 1$. An assumption of this kind appears for instance in [9], [39].

Remark 2.6.3. An assumption as (2.6.1) may not allow the application to Hamiltonians with super linear growth and inhomogeneous coefficients. However, we point out that more general conditions on the growth of the Hamiltonian could still lead to the Lipschitz bound, exactly as it is done in [79, Theorem 4]. The strategy in that case is to use a change of unknown (typically of exponential type) in addition to the usual Bernstein method. However this leads to an increase of technicalities which we decided to omit here, for the sake of brevity.

2.6.2 Semiconcavity of the value function

If we require stronger assumptions, we can also prove a semi-concavity bound on u. This will be helpful to improve the regularity of m under suitable assumptions.

We recall that a function f is said to be *semiconcave* in Ω if $\exists\, C > 0$ such that

$$f(x+h) + f(x-h) - 2f(x) \leq C|h|^2\,, \tag{2.6.5}$$

for each $x \in \Omega$, $h \in \mathbb{R}^N$ such that $x+h, x-h \in \Omega$.

In order to prove that u is semi concave, we will follow the tripling variable method used in [44]. To this purpose, we define the following function, that will play a crucial role :

$$\psi(x,y,z) = |x-z|^4 + |y-z|^4 + 2|x+y-2z|^2\,.$$

Then the semi-concavity of f is true if the following relation holds :

$$f(x) + f(y) - 2f(z) \leq C\sqrt{\psi(x,y,z)}\,, \qquad \forall x,y,z \in \Omega\,.$$

Indeed, it suffices to take $x = x' + h$, $y = x' - h$, $z = x'$ to obtain (2.6.5). We also recall that an equivalent formulation of this latter condition is the following : there exists a constant $C > 0$ such that

$$f(x) + f(y) - 2f(z) \leq C\left(\delta + \frac{\psi(x,y,z)}{\delta}\right) \quad \forall \delta > 0\ \forall x,y,z \in \Omega\,.$$

Moreover, it is well-known that a function f is in $W^{2,\infty}(\Omega)$ if and only if $\exists C > 0$ such that

$$|f(x) + f(y) - 2f(z)| \leq C\sqrt{\psi}\,,$$

or, equivalently, $\forall \delta > 0$

$$|f(x) + f(y) - 2f(z)| \leq C\left(\delta + \frac{\psi(x,y,z)}{\delta}\right)\,. \tag{2.6.6}$$

Theorem 2.6.4. *Suppose the hypotheses of Theorem 2.6.1 are satisfied. Moreover, suppose that $a(x) = \sigma(x)\sigma(x)^*$, with $\sigma \in W^{2,\infty}(\Omega)$ and that $F(\cdot, m) \in W^{2,\infty}(\Omega)$, $G(\cdot, m) \in W^{2,\infty}(\Omega)$ uniformly with respect to m. Finally, we require the following hypothesis on H : there exist constants C_0, C_1 such that*

$$H(t,x,p) + H(t,y,q) - 2H\left(t, z, \frac{p+q}{2}\right) \geq \\ - C_0(|x-z|^2 + |y-z|^2 + |x+y-2z|)(1+|p+q|) - C_1 |x-y| \, |p-q| \tag{2.6.7}$$

for any $(x, y, z) \in \Omega$, $(p, q) \in \mathbb{R}^N$, $t \in (0, T)$.

Then $u(t, \cdot)$ is a semiconcave function for all $t \in [0, T]$, with a semiconcavity constant bounded uniformly for $t \in (0, T)$. Namely, we have

$$D^2 u(t) \leq M \qquad \forall t \in (0, T)\,,$$

where M depends on T, $\|D^2\sigma\|_\infty, \|D^2 F\|_\infty, \|D^2 G\|_\infty$ and on H (through the constants appearing in the growth conditions).

Démonstration. We closely follow the proof given in [44, Theorem VII.3] with two main novelties : the boundary contribution, which will be handled through the invariance condition, and the structure condition (2.6.7) rather than the case of pure Bellman operators.

In the end, we wish to prove that there exist, $M > 0$ such that

$$u(t,x) + u(t,y) - 2u(t,z) \leq M\left(\delta + \frac{1}{\delta}\psi(x,y,z)\right)\,, \tag{2.6.8}$$

for every $t \in [0, T]$, every $x, y, z \in \Omega$ and for any $\delta > 0$ sufficiently small.
For a given $k > 0$ we consider the function $v(t, x) = e^{-k(T-t)} u(t, x)$. It satisfies the parabolic equation

$$\begin{cases} -v_t - \mathrm{tr}(a(x)D^2 v) + \tilde{H}(t,x,Dv) + kv = \tilde{F}(t,x,m) \\ v(T) = G(x, m(T))\,, \end{cases} \tag{2.6.9}$$

with

$$\tilde{H}(t,x,p) = e^{-k(T-t)} H(t,x,e^{k(T-t)}p)\,, \qquad \tilde{F}(t,x,m) = e^{-k(T-t)} F(x,m)\,.$$

We note that $\tilde{H}$ satisfies (2.6.7) and (2.2.9) uniformly in t, k, and that $\tilde{F}(t, \cdot, m) \in W^{2,\infty}(\Omega)$ uniformly in t. For $(\gamma, \delta, M) \in (0, +\infty)^3$ we take the following function :

$$\varphi(x,y,z) = M\left(\delta + \frac{\psi(x,y,z)}{\delta}\right) - \gamma \log(d(x)d(y)d(z))$$

where, without loss of generality, we assume that $d(x) \leq 1$ in Ω. Hence $\log(d(x)d(y)d(z)) \leq 0$. As usual, we assume that

$$\sup_{[0,T]\times\Omega^3} \left(v(t,x) + v(t,y) - 2v(t,z) - \varphi(x,y,z)\right) > 0\,,$$

and we will reach a contradiction if k, M are sufficiently large and γ sufficiently small, independently of the choice of δ.

Since $\phi(x,y,z) = +\infty$ if one of x, y, z lies in $\partial\Omega$, the sup is attained at a point $(\bar t, \bar x, \bar y, \bar z) = (\bar t_{\delta,M,\gamma,k}, \bar x_{\delta,M,\gamma,k}, \bar y_{\delta,M,\gamma,k}, \bar z_{\delta,M,\gamma,k}) \in [0,T] \times \Omega^3$. We drop the indexes for simplicity of notation. We observe further that, if $\bar t = T$, then we have

$$\begin{aligned} v(t,x) + v(t,y) - 2v(t,z) - \varphi(x,y,z) &= \\ &= G(x,m(T)) + G(y,m(T)) - 2G(z,m(T)) - \varphi(x,y,z)\,, \end{aligned}$$

which implies, thanks to the regularity of G,

$$v(t,x) + v(t,y) - 2v(t,z) \leq (C - M)\left(\delta + \frac{\psi(x,y,z)}{\delta}\right) \leq 0$$

provided $M \geq C$. So the sup must be attained at $\bar t < T$.

Now we proceed with typical viscosity solutions' arguments. Indeed, standing on the uniqueness result, it is easy to see that v is also a viscosity solution. It may actually be the case that v is smooth inside the domain, but we prefer to keep the argument in viscosity sense for a possibly wider generality. By Jensen's lemma, there exists matrices X, Y, Z and scalars a, b, c such that

$$\begin{pmatrix} X & 0 & 0 \\ 0 & Y & 0 \\ 0 & 0 & Z \end{pmatrix} \leq D^2\varphi(\bar x, \bar y, \bar z) \tag{2.6.10}$$

and

$$\begin{aligned} -a - \mathrm{tr}(a(\bar x)X) + \tilde H(\bar t, \bar x, D_x\varphi) + kv(\bar t, \bar x) &\leq \tilde F(\bar t, \bar x, m(\bar t)) \\ -b - \mathrm{tr}(a(\bar y)Y) + \tilde H(\bar t, \bar y, D_y\varphi) + kv(\bar t, \bar y) &\leq \tilde F(\bar t, \bar y, m(\bar t)) \\ -c - \mathrm{tr}\left(a(\bar z)\left(-\frac{1}{2}Z\right)\right) + \tilde H(\bar t, \bar z, -\frac{1}{2}D_z\varphi) + kv(\bar t, \bar z) &\geq \tilde F(\bar t, \bar z, m(\bar t)) \end{aligned}$$

where φ is computed at $(\bar t, \bar x, \bar y, \bar z)$ and where a, b, c (the *time derivatives* in viscosity sense) are real numbers such that $a + b \leq 2c$. We multiply by 2 the latter inequality, we sum and we get

$$
\begin{aligned}
&-\operatorname{tr}(a(\bar{x})X + a(\bar{y})Y + a(\bar{z})Z) + k\{v(\bar{t},\bar{x}) + v(\bar{t},\bar{y}) - 2v(\bar{t},\bar{z})\} \\
&\quad + \hat{H}(\bar{t},\bar{x},\bar{y},\bar{z}) \le \hat{F}(\bar{t},\bar{x},\bar{y},\bar{z})
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{H}(t,x,y,z) &= \tilde{H}(t,x,D_x\varphi(t,x)) + \tilde{H}(t,y,D_y\varphi(t,y)) - 2\tilde{H}(t,z,-\frac{1}{2}D_z\varphi(t,z))\,, \\
\hat{F}(t,x,y,z) &= \tilde{F}(t,x,m(t)) + \tilde{F}(t,y,m(t)) - 2\tilde{F}(t,z,m(t))\,.
\end{aligned}
$$

We multiply inequality (2.6.10) by the matrix $\Sigma = \Sigma(x,y,z)$, which is defined (in blocks) as

$$
\Sigma(x,y,z) = \begin{pmatrix} \sigma(x)\sigma^*(x) & \sigma(x)\sigma^*(y) & \sigma(x)\sigma^*(z) \\ \sigma(y)\sigma^*(x) & \sigma(y)\sigma^*(y) & \sigma(y)\sigma^*(z) \\ \sigma(z)\sigma^*(x) & \sigma(z)\sigma^*(y) & \sigma(z)\sigma^*(z) \end{pmatrix}\,,
$$

so we estimate

$$
\operatorname{tr}(a(\bar{x})X + a(\bar{y})Y + a(\bar{z})Z) \le \operatorname{tr}\left(\Sigma(\bar{x},\bar{y},\bar{z})D^2\varphi(\bar{x},\bar{y},\bar{z})\right)\,.
$$

We also estimate, since $\bar{t},\bar{x},\bar{y},\bar{z}$ is a maximum point and the maximum is positive

$$
v(\bar{t},\bar{x}) + v(\bar{t},\bar{y}) - 2v(\bar{t},\bar{z}) \ge \varphi(\bar{x},\bar{y},\bar{z})\,.
$$

Finally, we deduce that φ satisfies

$$
k\varphi(\bar{x},\bar{y},\bar{z}) \le \hat{F}(t,x,y,z) + \operatorname{tr}\left(\Sigma(\bar{x},\bar{y},\bar{z})D^2\varphi(\bar{x},\bar{y},\bar{z})\right) - \hat{H}(\bar{t},\bar{x},\bar{y},\bar{z})\,.
$$

Using the $W^{2,\infty}$ regularity of F, which therefore satisfies (2.6.6), we get

$$
k\varphi(\bar{x},\bar{y},\bar{z}) \le C\left(\delta + \frac{\psi(\bar{x},\bar{y},\bar{z})}{\delta}\right) + \operatorname{tr}\left(\Sigma(\bar{x},\bar{y},\bar{z})D^2\varphi(\bar{x},\bar{y},\bar{z})\right) - \hat{H}(\bar{t},\bar{x},\bar{y},\bar{z})\,. \tag{2.6.11}
$$

We analyse the two latter terms. From now on, we will omit the dependences from $(\bar{t},\bar{x},\bar{y},\bar{z})$ when there will be no possible mistake. We have

$$
\operatorname{tr}(\Sigma D^2\varphi) = \frac{M}{\delta}\operatorname{tr}(\Sigma D^2\psi) - \gamma\operatorname{tr}\left(\Sigma D^2\left(\log(d(x)d(y)d(z))\right)_{|(x,y,z)=(\bar{x},\bar{y},\bar{z})}\right)\,.
$$

So, we start computing the Hessian matrix of the function ψ. We get

$$
D_x\psi = 4|\bar{x}-\bar{z}|^2(\bar{x}-\bar{z}) + 4(\bar{x}+\bar{y}-2\bar{z})\,, \tag{2.6.12}
$$

$$
D_y\psi = 4|\bar{y}-\bar{z}|^2(\bar{y}-\bar{z}) + 4(\bar{x}+\bar{y}-2\bar{z})\,, \tag{2.6.13}
$$

$$
D_z\psi = -4|\bar{x}-\bar{z}|^2(\bar{x}-\bar{z}) - 4|\bar{y}-\bar{z}|^2(\bar{y}-\bar{z}) - 8(\bar{x}+\bar{y}-2\bar{z})\,, \tag{2.6.14}
$$

and so

$$
\begin{aligned}
&D^2_{xx}\psi = 8(\bar{x}-\bar{z})\otimes(\bar{x}-\bar{z}) + 4|\bar{x}-\bar{z}|^2 I + 4I\,, \qquad D^2_{xy}\psi = 4I\,,\\
&D^2_{xz}\psi = -8(\bar{x}-\bar{z})\otimes(\bar{x}-\bar{z}) - 4|\bar{x}-\bar{z}|^2 I - 8I\,,\\
&D^2_{yy}\psi = 8(\bar{y}-\bar{z})\otimes(\bar{y}-\bar{z}) + 4|\bar{y}-\bar{z}|^2 I + 4I\,,\\
&D^2_{yz}\psi = -8(\bar{y}-\bar{z})\otimes(\bar{y}-\bar{z}) - 4|\bar{y}-\bar{z}|^2 I - 8I\,,\\
&D^2_{zz}\psi = 8(\bar{x}-\bar{z})\otimes(\bar{x}-\bar{z}) + 4|\bar{x}-\bar{z}|^2 I + 8(\bar{y}-\bar{z})\otimes(\bar{y}-\bar{z}) + 4|\bar{y}-\bar{z}|^2 I + 16I\,,
\end{aligned}
$$

Therefore, computing the first trace we found

$$
\begin{aligned}
\frac{M}{\delta}\mathrm{tr}(\Sigma D^2\psi) = {}& 4\frac{M}{\delta}\mathrm{tr}((\sigma(\bar{x})+\sigma(\bar{y})-2\sigma(\bar{z}))(\sigma^*(\bar{x})+\sigma^*(\bar{y})-2^*\sigma(\bar{z}))+\\
&+4\frac{M}{\delta}|\bar{x}-\bar{z}|^2\,\mathrm{tr}\left((\sigma(\bar{x})-\sigma(\bar{z}))(\sigma^*(\bar{x})-\sigma^*(\bar{z}))\right)+\\
&+4\frac{M}{\delta}|\bar{y}-\bar{z}|^2\,\mathrm{tr}\left((\sigma(\bar{y})-\sigma(\bar{z}))(\sigma^*(\bar{y})-\sigma^*(\bar{z}))\right)+\\
&+8\frac{M}{\delta}|(\sigma^*(\bar{x})-\sigma^*(\bar{z}))(\bar{x}-\bar{z})|^2 + 8\frac{M}{\delta}|(\sigma^*(\bar{y})-\sigma^*(\bar{z}))(\bar{y}-\bar{z})|^2\,.
\end{aligned}
$$

Using the $W^{2,\infty}$ continuity of the function σ we easily obtain

$$
\frac{M}{\delta}\mathrm{tr}(\Sigma D^2\psi) \le c\,M\frac{\psi(\bar{x},\bar{y},\bar{z})}{\delta}\,.
$$

A straightforward computation shows us that

$$
\begin{aligned}
&D^2\left(\log(d(x)d(y)d(z))\right)_{|(x,y,z)=(\bar{x},\bar{y},\bar{z})} =\\
&=\begin{pmatrix}
\frac{D^2d(\bar{x})}{d(\bar{x})}-\frac{Dd(\bar{x})\otimes Dd(\bar{x})}{d^2(\bar{x})} & 0 & 0\\
0 & \frac{D^2d(\bar{y})}{d(\bar{y})}-\frac{Dd(\bar{y})\otimes Dd(\bar{y})}{d^2(\bar{y})} & 0\\
0 & 0 & \frac{D^2d(\bar{z})}{d(\bar{z})}-\frac{Dd(\bar{z})\otimes Dd(\bar{z})}{d^2(\bar{z})}
\end{pmatrix}.
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
&-\gamma\mathrm{tr}\left(\Sigma D^2\left(\log(d(x)d(y)d(z))\right)_{|(x,y,z)=(\bar{x},\bar{y},\bar{z})}\right) =\\
&= -\frac{\gamma}{d(\bar{x})}\left(\mathrm{tr}(a(\bar{x})D^2d(\bar{x})) - \frac{a(\bar{x})Dd(\bar{x})\cdot Dd(\bar{x})}{d(\bar{x})}\right)\\
&\quad -\frac{\gamma}{d(\bar{y})}\left(\mathrm{tr}(a(\bar{y})D^2d(\bar{y})) - \frac{a(\bar{y})Dd(\bar{y})\cdot Dd(\bar{y})}{d(\bar{y})}\right)\\
&\quad -\frac{\gamma}{d(\bar{z})}\left(\mathrm{tr}(a(\bar{z})D^2d(\bar{z})) - \frac{a(\bar{z})Dd(\bar{z})\cdot Dd(\bar{z})}{d(\bar{z})}\right)\,.
\end{aligned}
$$

Now we have to analyze the Hamiltonian term. As before, we need to split the computation in two parts, the first one including only the ψ function and the last one involving the logarithmic term. First of all, we recall that

$$\hat{H}(\bar{t},\bar{x},\bar{y},\bar{z}) = \tilde{H}(\bar{t},\bar{x},D_x\varphi(\bar{x},\bar{y},\bar{z})) + \tilde{H}(\bar{t},\bar{y},D_y\varphi(\bar{x},\bar{y},\bar{z})) - 2\tilde{H}\left(\bar{t},\bar{z},-\frac{1}{2}D_z\varphi(\bar{x},\bar{y},\bar{z})\right).$$

Since

$$D_x\varphi(\bar{x},\bar{y},\bar{z}) = \frac{M}{\delta}D_x\psi(\bar{x},\bar{y},\bar{z}) - \gamma D\log d(\bar{x})$$

and the same holds for $D_y\varphi$, $D_z\varphi$, we can write

$$\begin{aligned}\hat{H}(\bar{t},\bar{x},\bar{y},\bar{z}) = {}& \tilde{H}\left(\bar{t},\bar{x},\frac{M}{\delta}D_x\psi\right) - \gamma\int_0^1 \tilde{H}_p(\bar{t},\bar{x},p_1(\lambda))\frac{Dd(\bar{x})}{d(\bar{x})}d\lambda + \\ & + \tilde{H}\left(\bar{t},\bar{y},\frac{M}{\delta}D_y\psi\right) - \gamma\int_0^1 \tilde{H}_p(\bar{t},\bar{y},p_2(\lambda))\frac{Dd(\bar{y})}{d(\bar{y})}d\lambda - \\ & -2\tilde{H}\left(\bar{t},\bar{z},-\frac{M}{2\delta}D_z\psi\right) - \gamma\int_0^1 \tilde{H}_p(\bar{t},\bar{z},p_3(\lambda))\frac{Dd(\bar{z})}{d(\bar{z})}d\lambda\,,\end{aligned}$$

where

$$\begin{aligned} p_1(\lambda) &= \frac{M}{\delta}D_x\psi(\bar{x},\bar{y},\bar{z}) - \lambda\gamma D\log d(\bar{x})\,, \\ p_2(\lambda) &= \frac{M}{\delta}D_y\psi(\bar{x},\bar{y},\bar{z}) - \lambda\gamma D\log d(\bar{y})\,, \\ p_3(\lambda) &= -\frac{M}{2\delta}D_z\psi(\bar{x},\bar{y},\bar{z}) + \frac{1}{2}\lambda\gamma D\log d(\bar{z})\,. \end{aligned}$$

Putting these estimates in (2.6.11), one finds

$$\begin{aligned} k\varphi(\bar{x},\bar{y},\bar{z}) \le {}& C(1+M)\left(\delta + \frac{\psi(\bar{x},\bar{y},\bar{z})}{\delta}\right) \\ & -\tilde{H}\left(\bar{t},\bar{x},\frac{M}{\delta}D_x\psi\right) - \tilde{H}\left(\bar{t},\bar{y},\frac{M}{\delta}D_y\psi\right) + 2\tilde{H}\left(\bar{t},\bar{z},-\frac{M}{2\delta}D_z\psi\right) \\ & -\frac{\gamma}{d(\bar{x})}\int_0^1\left(\mathrm{tr}(a(\bar{x})D^2d(\bar{x})) - \frac{a(\bar{x})Dd(\bar{x})\cdot Dd(\bar{x})}{d(\bar{x})} - \tilde{H}_p(\bar{t},\bar{x},p_1(\lambda))Dd(\bar{x})\right)d\lambda \\ & -\frac{\gamma}{d(\bar{y})}\int_0^1\left(\mathrm{tr}(a(\bar{y})D^2d(\bar{y})) - \frac{a(\bar{y})Dd(\bar{y})\cdot Dd(\bar{y})}{d(\bar{y})} - \tilde{H}_p(\bar{t},\bar{y},p_2(\lambda))Dd(\bar{y})\right)d\lambda \\ & -\frac{\gamma}{d(\bar{z})}\int_0^1\left(\mathrm{tr}(a(\bar{z})D^2d(\bar{z})) - \frac{a(\bar{z})Dd(\bar{z})\cdot Dd(\bar{z})}{d(\bar{z})} - \tilde{H}_p(\bar{t},\bar{z},p_3(\lambda))Dd(\bar{z})\right)d\lambda\,. \end{aligned}$$

We use the invariance condition (2.2.9) to get rid of the latter terms. So the inequality becomes

$$\begin{aligned} k\varphi(\bar x,\bar y,\bar z) \leq\ & C(1+M)\left(\delta+\frac{\psi(\bar x,\bar y,\bar z)}{\delta}\right) \\ -\tilde H\left(\bar t,\bar x,\frac{M}{\delta}D_x\psi\right) &- \tilde H\left(\bar t,\bar y,\frac{M}{\delta}D_y\psi\right)+2\tilde H\left(\bar t,\bar z,-\frac{M}{2\delta}D_z\psi\right)+c\,\gamma\,. \end{aligned} \tag{2.6.15}$$

Finally, since $-D_z\psi = D_x\psi + D_y\psi$, we use (2.6.7) to estimate the last terms involving $\tilde H$:

$$\begin{aligned} -\tilde H\left(\bar t,\bar x,\frac{M}{\delta}D_x\psi\right) - \tilde H\left(\bar t,\bar y,\frac{M}{\delta}D_y\psi\right)+2\tilde H\left(\bar t,\bar z,-\frac{M}{2\delta}D_z\psi\right) \leq \\ \leq C_0(|x-z|^2+|y-z|^2+|x+y-2z|)(1+\frac{M}{\delta}\,|D_x\psi+D_y\psi|) \\ + C_1|x-y|\frac{M}{\delta}\,|D_x\psi-D_y\psi|\ . \end{aligned}$$

Using the precise values of $D_x\psi$ and $D_y\psi$, we estimate thanks to Young's inequality

$$\begin{aligned} C_1\,|x-y|\frac{M}{\delta}\,|D_x\psi-D_y\psi| &\leq C\,|x-y|\frac{M}{\delta}\left(|x-z|^3+|y-z|^3\right) \\ &\leq C\,\frac{M}{\delta}\psi \end{aligned}$$

and similarly

$$\begin{aligned} & C_0(|x-z|^2+|y-z|^2+|x+y-2z|)(1+\frac{M}{\delta}\,|D_x\psi+D_y\psi|) \leq C\left(\sqrt{\psi}+\frac{M}{\delta}\psi\right) \\ & \leq C\left(\delta+\frac{M}{\delta}\psi\right)\ . \end{aligned}$$

Eventually, we end up with

$$-\tilde H\left(\bar t,\bar x,\frac{M}{\delta}D_x\psi\right) - \tilde H\left(\bar t,\bar y,\frac{M}{\delta}D_y\psi\right)+2\tilde H\left(\bar t,\bar z,-\frac{M}{2\delta}D_z\psi\right) \leq C\left(\delta+\frac{M}{\delta}\psi\right)\ . \tag{2.6.16}$$

Since

$$\begin{aligned} k\varphi(\bar x,\bar y,\bar z) =& kM\left(\delta+\frac{\psi(\bar x,\bar y,\bar z)}{\delta}\right) - k\gamma\log(d(\bar x)d(\bar y)d(\bar z)) \geq \\ \geq& kM\left(\delta+\frac{\psi(\bar x,\bar y,\bar z)}{\delta}\right)\ , \end{aligned}$$

we obtain from (2.6.15)–(2.6.16)

$$kM\left(\delta+\frac{\psi(\bar x,\bar y,\bar z)}{\delta}\right) \leq C(1+M)\left(\delta+\frac{\psi(\bar x,\bar y,\bar z)}{\delta}\right)+C\,\gamma\,.$$

We choose k such that $kM - C(1+M) \geq 1$ to have

$$\delta \leq \left(\delta + \frac{\psi(\bar{x},\bar{y},\bar{z})}{\delta}\right) \leq C\,\gamma\,.$$

Choosing γ sufficiently small we obtain a contradiction.

So, we have found that for k sufficiently large and γ sufficiently small, we have

$$v(t,x) + v(t,y) - 2v(t,z) \leq M\left(\delta + \frac{\psi(x,y,z)}{\delta}\right) - \gamma \log(d(x)d(y)d(z))\,,$$

for every $x, y, z \in \Omega$, and any $\delta > 0$. Letting $\gamma \to 0$, we obtain

$$u(t,x) + u(t,y) - 2u(t,z) \leq e^{kT} M\left(\delta + \frac{\psi(\bar{x},\bar{y},\bar{z})}{\delta}\right)\,.$$

So, (2.6.8) is proved and the proof is concluded. □

Remark 2.6.5. We note that the hypothesis (2.6.7) is satisfied at least for classical Bellman equations where

$$H(x,p) = \sup_{\alpha\in A}(-b(x,\alpha)\cdot p - L(x,\alpha))$$

assuming that $L(\cdot,\alpha)$ and $b(\cdot,\alpha)$ are $W^{2,\infty}$, and both conditions hold uniformly with respect to $\alpha \in A$. Indeed, we have

$$\begin{aligned}
2&\left(-b(z,\alpha)\cdot\left(\frac{p+q}{2}\right) - L(z,\alpha)\right) = (-b(z,\alpha)\cdot p - L(x,\alpha)) \\
&\quad + (-b(z,\alpha)\cdot q - L(y,\alpha)) + (L(x,\alpha) + L(y,\alpha) - 2L(z,\alpha)) \\
&= (-b(x,\alpha)\cdot p - L(x,\alpha)) + (-b(y,\alpha)\cdot q - L(y,\alpha)) \\
&\quad + \frac{1}{2}(b(x,\alpha) - b(y,\alpha))\cdot(p-q) + \frac{1}{2}(b(x,\alpha) + b(y,\alpha) - 2b(z,\alpha))(p+q) \\
&\quad + (L(x,\alpha) + L(y,\alpha) - 2L(z,\alpha)) \\
&\leq H(x,p) + H(y,q) + C\,|x-y|\,|p-q| \\
&\quad\quad + C\left(|x-z|^2 + |y-z|^2 + |x+y-2z|\right)|p+q| \\
&\quad\quad + (L(x,\alpha) + L(y,\alpha) - 2L(z,\alpha))\,.
\end{aligned}$$

Hence, taking the sup_α in the left-hand side, and using the regularity of L, implies (2.6.7).

Finally, we show with the next result that the semiconcavity of u leads to the boundedness of the function m.

Proposition 2.6.6. *Assume that the hypotheses of Theorem 2.6.4 are satisfied, and, in addition, that $H_p(t,x,p) \in W^{1,\infty}(Q_T \times K)$ for all compact sets $K \subset \mathbb{R}^N$. Suppose that the following condition holds near the boundary : there exists δ_0 such that*

$$\Big(\tilde{b}(x) + H_p(t,x,p)\Big) \cdot Dd(x) \le 0\,, \qquad \forall x \in \Gamma_{\delta_0},\ \forall p \in \mathbb{R}^N\,. \tag{2.6.17}$$

Then, if (u,m) is a solution of (2.1.1), we have $m \in L^\infty([0,T]\times\Omega)$.

Démonstration. We are going to apply the comparison principle in the equation of m. To do so, we call μ_ε the solution of the following problem :

$$\begin{cases} (\mu_\varepsilon)_t - \operatorname{div}(a(x)D\mu_\varepsilon) - \operatorname{div}((\tilde{b}(x) + H_p(t,x,Du))\mu_\varepsilon) = 0\,, & (t,x) \in (0,T)\times\Omega_\varepsilon \\ \mu_\varepsilon(0) = m_0 \\ [a(x)D\mu_\varepsilon + (\tilde{b}(x) + H_p(t,x,Du))\mu_\varepsilon]\cdot\nu_{|\partial\Omega_\varepsilon} = 0\,, \end{cases}$$

where, as before, $\Omega_\varepsilon = \{x : d(x) > \varepsilon\}$. With the same arguments used previously we obtain

$$\mu_\varepsilon \to m \qquad a.e. \text{ in } [0,T]\times\Omega\,.$$

Since u is a semiconcave function, we can split the last divergence term in order to get

$$\begin{cases} (\mu_\varepsilon)_t - \operatorname{div}(a(x)D\mu_\varepsilon) - (\tilde{b} + H_p(t,x,Du))D\mu_\varepsilon + c(t,x)\mu_\varepsilon = 0\,, & (t,x) \in (0,T)\times\Omega_\varepsilon \\ \mu_\varepsilon(0) = m_0 \\ [a(x)D\mu_\varepsilon + (\tilde{b}(x) + H_p(t,x,Du))\mu_\varepsilon]\cdot\nu_{|\partial\Omega_\varepsilon} = 0\,, \end{cases}$$

where c is defined as follows :

$$c(t,x) = -\operatorname{div}(\tilde{b}) - \operatorname{tr}(H_{px}(t,x,Du)) - \operatorname{tr}(H_{pp}(t,x,Du)D^2u)\,.$$

Recall that in Ω_ε the matrix $a(x)$ is elliptic, so u enjoys the standard parabolic regularity and $c(t,x)$ is well defined (at least in Lebesgue spaces). We now estimate the function c. Since $H_{px}(t,x,Du)$ is in $L^\infty(Q_T)$ (because u is globally Lipschitz) and $D^2u \le CI$, we get, up to changing C,

$$c(t,x) \ge -C - \operatorname{tr}(H_{pp}(x,Du)(D^2u - CI)) - C\operatorname{tr}(H_{pp}(x,Du)) \ge -k\,,$$

for a certain $k > 0$ and since $H_{pp}(D^2u - CI)$ is a negative semi-definite matrix.
Calling $\mu_\varepsilon^k = e^{-kt}\mu_\varepsilon$, we have that μ_ε^k is the solution of the following equation

$$\begin{cases} (\mu_\varepsilon^k)_t - \operatorname{div}((a(x)+\varepsilon I)D\mu_\varepsilon^k) - (\tilde{b} + H_p(x,Du))D\mu_\varepsilon^k + (k + c(t,x))\mu_\varepsilon^k = 0 \\ \mu_\varepsilon^k(0) = m_0 \\ [\varepsilon D\mu_\varepsilon^k + (\tilde{b}(x) + H_p(x,Du))\mu_\varepsilon^k]\cdot\nu_{|\partial\Omega_\varepsilon} = 0 \end{cases}\,.$$

We choose k such that $k + c(t,x) \geq 0$ for each $(t,x) \in [0,T] \times \Omega_\varepsilon$. Then, thanks to (2.6.17), it is immediate to prove that M is a super-solution of the equation of μ_ε^k, for $M \geq \|m_0\|_\infty$. Now we can easily conclude the proof : thanks to the comparison principle, we have

$$\mu_\varepsilon^k(t,x) \leq M \implies \mu_\varepsilon(t,x) \leq e^{kT} M \overset{\varepsilon \to 0}{\implies} m(t,x) \leq e^{kT} M \, ,$$

where the estimates are true almost everywhere in (t,x). Since $m \geq 0$, the proof is concluded.

□

2.7 Non-smooth domains

Unfortunately, in many applications one needs to consider that the state variable does not belong to a $\mathcal{C}^2$ domain. This implies that the distance function from the boundary of Ω turns out not to be a $\mathcal{C}^2$ function, and the invariance condition (2.2.9) becomes meaningless. However, a generalization of the results obtained so far is possible, in the following setting.

Theorem 2.7.1. *Suppose that* $\exists \psi \in \mathcal{C}^2(\Omega)$ *such that* $\psi > 0$ *in* Ω, $\psi = 0$ *in* $\partial\Omega$ *and the following inequality holds in a neighborhood* V *of* $\partial\Omega$ *:*

$$\begin{aligned} &\operatorname{tr}(a(x)D^2\psi(x)) - H_p(t,x,p)D\psi(x) \geq \frac{a(x)D\psi(x)\cdot D\psi(x)}{\psi(x)} - C\psi(x) \\ &\qquad \forall\, p \in \mathbb{R}^N \ , \forall t \in [0,T] \ \text{ and a.e. } x \in V \, . \end{aligned} \tag{2.7.1}$$

Then, all the results of the previous sections remain true replacing (2.2.9) with (2.7.1).

All the proofs can be done in the same way, replacing d by ψ and the set Γ_ε by $\{\psi < \varepsilon\} \cap \Omega$.

This generalization plays a crucial role in order that the smoothness assumption for the domain Ω be weakened. As an example, we define a class of non-smooth domains and we prove that hypothesis (2.7.1) is satisfied for those ones.

Definition 2.7.2. *Let* $\Omega \subseteq \mathbb{R}^N$. *We say that* Ω *is a generalized* $\mathcal{C}^2$ *domain if* $\exists n \in \mathbb{N}$ *and a collection of sets* $\{\Omega^i\}_{1 \leq i \leq n}$ *such that* $\overline{\Omega}^i$ *is a compact domain of class* $\mathcal{C}^2$ *and*

$$\Omega = \bigcap_{i=1}^{n} \Omega^i \, .$$

From now on, when we write $d(x)$ and $d_i(x)$, in the case of a generalized $\mathcal{C}^2$ domain, we mean respectively $d_\Omega(x)$ and $d_{\Omega^i}(x)$. Moreover, we will use the following notation :

$$\Omega^i_\varepsilon = \{d_i(x) > \varepsilon\} \cap \Omega\,, \qquad \Gamma^i_\varepsilon = \{d_i(x) < \varepsilon\} \cap \Omega\,.$$

We now show that Theorem 2.7.1 applies to generalized $\mathcal{C}^2$ domains.

Proposition 2.7.3. *Let* $\Omega = \bigcap_{i=1}^{n} \Omega^i$ *be a generalized* $\mathcal{C}^2$ *domain. Suppose that* $\exists\ \delta, C_0 > 0$ *s.t.* $\forall\ p \in \mathbb{R}^N$, $\forall 1 \le i \le n$ *and* $\forall\ x \in \Gamma^i_\delta$ *the following inequality holds :*

$$\operatorname{tr}(a(x)D^2 d_i(x)) - H_p(x,p)Dd_i(x) \ge \frac{a(x)Dd_i(x)Dd_i(x)}{d_i(x)} - C_0 d_i(x)\,. \tag{2.7.2}$$

Then all the results of the previous sections remain true for the non-smooth domain Ω.

Démonstration. We have to prove that condition (2.7.1) is satisfied for a certain $\mathcal{C}^2$ function ψ.
To do that, we consider $\phi : [0, +\infty) \to \mathbb{R}$ a C^2 function such that $\phi(s) = s$ when $s \le \frac{\delta}{2}$, $\phi \equiv 1$ for $s \ge \delta$ and $\phi'(s) \ge 0$. Moreover, we require that $\phi(x) \ge x$ in $[0, \delta]$. This can be done for δ sufficiently small.

We take $\psi(x) = \prod_i \phi(d_i(x))$ and we prove that (2.7.1) is satisfied in $V = \bigcup_\delta \Gamma^i_\delta$ for a certain $\delta > 0$.

From now on, we will write ϕ_i, ϕ'_i, ϕ''_i instead of $\phi(d_i(x))$, $\phi'(d_i(x))$, $\phi''(d_i(x))$ to simplify the notation. Computing the derivative $D\psi$ and $D^2\psi$ we find

$$D\psi = \sum_i \prod_{j\neq i} \phi_j\, \phi'_i Dd_i\,,$$

$$D^2\psi = \sum_i \prod_{j\neq i} \phi_j\, \phi'_i D^2 d_i + \sum_i \prod_{j\neq i} \phi_j\, \phi''_i Dd_i \otimes Dd_i + \sum_{i,k\neq i} \prod_{j\neq i,k} \phi_j\, \phi'_i \phi'_k Dd_i \otimes Dd_k\,.$$

Plugging these computations in (2.7.1) we find

$$\begin{aligned} &\operatorname{tr}(a(x)D^2\psi(x)) - H_p(t,x,p)D\psi(x) - \frac{a(x)D\psi(x)\cdot D\psi(x)}{\psi(x)} + C\psi(x) = \\ =&\sum_i \prod_{j\neq i} \phi_j\, \phi'_i \left(\operatorname{tr}(a(x)D^2 d_i) - H_p(x,p)Dd_i\right) + \sum_i \prod_{j\neq i} \phi_j\, \phi''_i\, a(x)Dd_i \cdot Dd_i + \\ &+\sum_{i,k\neq i} \prod_{j\neq i,k} \phi_j\, \phi'_i \phi'_k\, a(x)Dd_i \cdot Dd_k - \frac{\sum_{i,k} \prod_{j\neq i} \phi_j \prod_{l\neq k} \phi_l\, \phi'_i \phi'_k\, a(x)Dd_i \cdot Dd_k}{\prod_i \phi_i} + C\psi(x). \end{aligned}$$

We start analyzing the first term. Because of the presence of ϕ_i', we can study each term of the sum only in Γ_δ^i. So, choosing δ such that (2.7.2) holds true, we get

$$\sum_i \prod_{j\neq i} \phi_j\, \phi_i' \left(\operatorname{tr}(a(x)D^2 d_i) - H_p(x,p)Dd_i\right) \geq \sum_i \prod_{j\neq i} \phi_j\, \phi_i' \left(\frac{a(x)Dd_i\cdot Dd_i}{d_i} - C_0 d_i\right).$$

Since $d_i \leq \phi_i$ in Γ_δ^i, one has

$$-C_0 \sum_i \prod_{j\neq i} \phi_j\, \phi_i' d_i \geq -C_1 \sum_i \prod_j \phi_j \geq -C_1 \psi(x)\,,$$

where C_1 is a constant depending on C_0 and n that can change from line to line.

Then we look at the third and the fourth terms. Since, for $i \neq k$,

$$\frac{\prod_{j\neq i}\phi_j \prod_{l\neq k}\phi_l}{\prod_i \phi_i} = \prod_{j\neq i,k} \phi_j\,,$$

then we have

$$\begin{aligned}
&\sum_{i,k\neq i} \prod_{j\neq i,k} \phi_j\, \phi_i'\phi_k'\, a(x)Dd_i\cdot Dd_k - \frac{\sum_{i,k}\prod_{j\neq i}\phi_j \prod_{l\neq k}\phi_l\, \phi_i'\phi_k'\, a(x)Dd_i\cdot Dd_k}{\prod_i \phi_i} \\
&= -\frac{\sum_i \left(\prod_{j\neq i}\phi_j\right)^2 (\phi_i')^2 a(x)Dd_i\cdot Dd_i}{\prod_i \phi_i} = -\sum_i \prod_{j\neq i} \phi_j\, \frac{(\phi_i')^2}{\phi_i} a(x)Dd_i\cdot Dd_i\,.
\end{aligned}$$

Using these estimates, we obtain

$$\begin{aligned}
&\operatorname{tr}(a(x)D^2\psi(x)) - H_p(t,x,p)D\psi(x) - \frac{a(x)D\psi(x)\cdot D\psi(x)}{\psi(x)} + C\psi(x) \geq \\
&\geq \sum_i \prod_{j\neq i} \phi_j\, a(x)Dd_i\cdot Dd_i \left(\frac{\phi_i'}{d_i} - \frac{(\phi_i')^2}{\phi_i} + \phi_i''\right) + (C - C_1)\psi(x)\,.
\end{aligned} \tag{2.7.3}$$

To conclude, we want to prove that

$$\frac{\phi_i'}{d_i} - \frac{(\phi_i')^2}{\phi_i} + \phi_i'' \geq -C_2 \phi_i$$

for a certain constant C_2. This is equivalent to prove that, for all $x \in \mathbb{R}^+$,

$$\phi''(x)\phi(x)x - (\phi'(x))^2 x + \phi'(x)\phi(x) \geq -C_2 \phi^2(x)x\,.$$

Since $\phi(s) = s$ in $[0, \frac{\delta}{2}]$, we obtain immediately that the left hand side term vanishes in this interval, and so the relation is verified. A similar computation occurs for $s \geq \delta$. Finally, for $s \in [\frac{\delta}{2}, \delta]$ the relation is certainly satisfied for a constant C_2 depending on δ. Therefore we obtain from (2.7.3)

$$\operatorname{tr}(a(x)D^2\psi(x)) - H_p(t,x,p)D\psi(x) - \frac{a(x)D\psi(x)\cdot D\psi(x)}{\psi(x)} + C\psi(x) \geq (C - C_1 - C_2)\psi(x)\,,$$

which, for C sufficiently large, proves that condition (2.7.1) is satisfied. This concludes the proof. □

2.8 Appendix

In this Appendix, we give the proof of two technical results.

Proof of Lemma 2.3.3. We consider the sequence of compact sets $\{D_k\}_{k\in\mathbb{N}}$ defined in (2.2.2). For each $k \in \mathbb{N}$ we take a cut-off function ξ_k such that

$$\begin{cases} \xi_k \in C_c^\infty(\Omega)\,, \quad 0 \leq \xi_k \leq 1 \\ \xi_k(x) \equiv 1 \quad \text{for } x \in D_k \\ \xi_k(x) \equiv 0 \quad \text{for } x \in D_{k+1}. \end{cases} \tag{2.8.1}$$

For ε small enough and $\lambda > 0$, we multiply the equation (2.3.7) by $e^{\lambda u_\varepsilon}\xi_k^2$ and we integrate in $[t,T]\times\Omega_\varepsilon$:

$$\begin{aligned} &\frac{1}{\lambda}\int_{D_{k+1}} e^{\lambda u_\varepsilon}(t)\xi_k^2\,dx + \lambda\int_t^T\int_{D_{k+1}} e^{\lambda u_\varepsilon}a_\varepsilon(x)Du_\varepsilon\cdot Du_\varepsilon\,\xi_k^2\,dxdt \\ &+\int_t^T\int_{D_{k+1}} H_\varepsilon(t,x,Du_\varepsilon)e^{\lambda u_\varepsilon}\xi_k^2\,dxdt + \int_t^T\int_{D_{k+1}} a_\varepsilon(x)Du_\varepsilon\cdot D\xi_k\,2e^{\lambda u_\varepsilon}\xi_k\,dxdt \\ &\qquad = \frac{1}{\lambda}\int_{D_{k+1}} e^{\lambda u_\varepsilon}(T)\xi_k^2\,dx \leq C \end{aligned}$$

since u_ε is uniformly bounded. From the local uniform coercivity of a_ε, we have $a_\varepsilon(x) \geq \lambda_{k+1}\,I$ for $x \in D_{k+1}$. Using also the local uniform natural growth assumed upon H_ε, and a local bound on a_ε, we deduce that

$$\lambda\,\lambda_{k+1}\int_t^T\int_{D_{k+1}} e^{\lambda u_\varepsilon}\,|Du_\varepsilon|^2\,\xi_k^2\,dxdt \leq C_k\int_0^T\int_{D_{k+1}} e^{\lambda u_\varepsilon}(1+|Du_\varepsilon|^2)\xi_k^2\,dxdt$$

for some constant C_k only depending on k. Choosing λ sufficiently large (depending on k), we can bound the gradient of u_ε in D_k. Hence, together with the L^∞ bound, we deduce that u_ε is bounded in $L^2([0,T];W^{1,2}(D_k))$ for each $k \in \mathbb{N}$.

From (2.3.7) now we get that $(u_\varepsilon)_t$ is bounded in $L^2([0,T];W^{-1,2}(D_k))$. So, by [93, Corollary 4], we deduce that u_ε is relatively compact in $L^2(D_k)$. By a standard diagonal argument, we can therefore extract a subsequence, which we still denote by u_ε, such that

$$u_\varepsilon \to u \text{ weakly in } L^2([0,T];W^{1,2}(K)) \text{ and strongly in } L^p([0,T]\times K) \text{ for every } p<\infty,$$

for any compact subset $K \subset \Omega$.

We now aim at getting the strong convergence. To this purpose we assume that the matrix $a_\varepsilon(x)$ converges (up to subsequences) almost everywhere in Ω towards some matrix $a(x)$. We further suppose by now that $u_\varepsilon(T)$ converges almost everywhere in Ω to some function $g(x)$, in order to get the full convergence up to $t=T$. We notice that, since u_ε is uniformly bounded, this implies that $u_\varepsilon(T) \to g$ strongly in $L^p(\Omega)$ for all $p<\infty$. Moreover, since $(u_\varepsilon)_t$ converges to u_t weakly in $L^2([0,T];W^{-1,2}(D_k))$ for all D_k, one has that $u_t \in L^2([0,T];W^{-1,2}(D_k))$, so $u \in C^0([0,T];L^2(D_k))$ and actually $u(T)=g(x)$ in Ω.

Now we multiply (2.3.7) by $\psi(u_\varepsilon-u)\xi_k^2$, for a convenient increasing function ψ to be chosen later. We proceed in a similar way as above obtaining

$$\begin{aligned}
&\lambda_{k+1}\int_0^T\int_{D_{k+1}} \psi'(u_\varepsilon-u)\,|Du_\varepsilon-Du|^2\,\xi_k^2\,dxdt\\
&\quad\le -\int_0^T\int_{D_{k+1}} \psi'(u_\varepsilon-u)a_\varepsilon(x)Du\cdot(Du_\varepsilon-Du)\,\xi_k^2\\
&\quad+C_k\int_0^T\int_{D_{k+1}} |\psi(u_\varepsilon-u)|(1+|Du_\varepsilon|^2)\,\xi_k^2\,dxdt-\int_0^T\langle\partial_t u_\varepsilon,\psi(u_\varepsilon-u)\xi_k^2\rangle\,,
\end{aligned}$$

which yields, using $|Du_\varepsilon|^2 \le 2(|Du_\varepsilon-Du|^2+|Du|^2)$:

$$\begin{aligned}
&\int_0^T\int_{D_{k+1}} [\lambda_{k+1}\psi'(u_\varepsilon-u)-2C_k|\psi(u_\varepsilon-u)|]\,|Du_\varepsilon-Du|^2\,\xi_k^2\,dxdt\le\\
&\le -\int_0^T\int_{D_{k+1}} \psi'(u_\varepsilon-u)a_\varepsilon(x)Du\cdot(Du_\varepsilon-Du)\,\xi_k^2\\
&\quad+C_k\int_0^T\int_{D_{k+1}} |\psi(u_\varepsilon-u)|(1+2|Du|^2)\,\xi_k^2\,dxdt+\int_0^T\langle\partial_t u_\varepsilon,\psi(u_\varepsilon-u)\xi_k^2\rangle\,.
\end{aligned}$$

Now we choose ψ such that $s\psi(s) \geq 0$ and $\lambda_{k+1}\psi'(s) - 2C_k|\psi(s)| > 0$ for all $s \in \mathbb{R}$. A typical choice is e.g. $\psi(s) = se^{bs^2}$ with $b = \frac{C_k^2}{\lambda_{k+1}^2}$. So we get

$$
\begin{aligned}
&\int_0^T \int_{D_{k+1}} |Du_\varepsilon - Du|^2 \, \xi_k^2 \, dxdt \leq \\
&\leq -C_k \int_0^T \int_{D_{k+1}} \psi'(u_\varepsilon - u) a_\varepsilon(x) Du \cdot (Du_\varepsilon - Du) \, \xi_k^2 \qquad (2.8.2) \\
&\quad + C_k \int_0^T \int_{D_{k+1}} |\psi(u_\varepsilon - u)|(1 + 2|Du|^2) \, \xi_k^2 \, dxdt + C_k \int_0^T \langle \partial_t u_\varepsilon, \psi(u_\varepsilon - u)\xi_k^2 \rangle \,,
\end{aligned}
$$

where $\langle \cdot, \cdot \rangle$ is the duality between $W^{1,2}(\Omega_\varepsilon)$ and its dual. We conclude by showing that all terms in the right-hand side go to zero as $\varepsilon \to 0$. Indeed, if $\Psi(s) = \int_0^s \psi(r)dr$, we have

$$
\begin{aligned}
\int_0^T \langle \partial_t u_\varepsilon, \psi(u_\varepsilon - u)\xi_k^2 \rangle &= \int_\Omega \Psi(u_\varepsilon(t) - u(t))dx \, \overset{T}{\underset{0}{I}} + \int_0^T \langle \partial_t u, \psi(u_\varepsilon - u)\xi_k^2 \rangle \\
&\leq \int_\Omega \Psi(u_\varepsilon(T) - u(T))dx + \int_0^T \langle \partial_t u, \psi(u_\varepsilon - u)\xi_k^2 \rangle \, dt
\end{aligned}
$$

and the last two terms converge to zero because $\psi(u_\varepsilon - u)\xi_k^2$ weakly converges to zero in $L^2([0,T]; W^{1,2}(D_k))$ and $u_\varepsilon(T) \to g(x) = u(T)$ almost everywhere (hence $\Psi(u_\varepsilon(T) - u(T)) \to 0$ in L^1 by dominated convergence). Still using Lebesgue's dominated convergence theorem, we have that $|\psi(u_\varepsilon - u)|(1 + 2|Du|^2) \to 0$ in $L^1((0,T) \times D_{k+1})$. Finally, using that $a_\varepsilon(x)$ converges almost everywhere, we can deduce that $\psi'(u_\varepsilon - u)a_\varepsilon(x)Du$ converges strongly in $L^2((0,T) \times D_{k+1})$, and since Du_ε converges weakly to Du, the first integral in the right-hand side of (2.8.2) converges to zero as well. Therefore, (2.8.2) implies that

$$
\int_0^T \int_{D_{k+1}} |Du_\varepsilon - Du|^2 \, \xi_k^2 \, dxdt \overset{\varepsilon \to 0}{\to} 0
$$

which implies the strong convergence of u_ε to u in $L^2([0,T]; W^{1,2}(D_k))$ for all D_k.

We finish by noticing that, in case $u_\varepsilon(T)$ is not assume to converge strongly, then the above argument needs to be localized, which means using the test function $\psi(u_\varepsilon - u)\xi_k^2 \, (T - t)$. With the same arguments as above, in that case one concludes the strong convergence in $L^2([0,t]; W^{1,2}(D_k))$ for all $t < T$, but not up to $t = T$. □

We now provide the proof of the analog compactness result for the Fokker-Planck equation.

Proof of Lemma 2.4.2. We only sketch the proof since this is just a local version of compactness results which are well-known in the case of boundary value problems.

We first obtain local estimates as in [90, Lemma 3.3] : namely, for any $q < \frac{N+2}{N+1}$,

$$\int_0^T \int_{D_k} |Dm_\varepsilon|^q \, dxdt \le C_{q,k} \tag{2.8.3}$$

for every D_k defined in (2.2.2). This estimate in particular implies that m_ε is bounded in $L^q(0,T;W^{1,q}(K))$ for $q < \frac{N+2}{N+1}$, for any compact subset K. From the equation and the local boundedness of a_ε, b_ε, we deduce that $(m_\varepsilon)_t$ is bounded in $L^q(0,T;W^{-1,q}(K))$, so applying standard compactness results (see [93, Corollary 4]) we get that m_ε is relatively compact in $L^1(0,T;L^1(K))$. Through a diagonal procedure, we can extract a subsequence, which is not relabeled, such that m_ε converges almost everywhere in Q_T, and in $L^1(0,T;L^1(K))$ for every compact subset $K \subset \Omega$, towards some function m which actually belongs to $L^\infty((0,T);L^1(\Omega))$ because of Fatou's lemma and the fact that $\|m_\varepsilon(t)\|_{L^1(\Omega)}$ is uniformly bounded.

In order to obtain a strong convergence in $C^0([0,T];L^1_{loc}(\Omega))$, we use some renormalization argument similar as in [89, Theorem 6.1]. To this purpose, we use the auxiliary function

$$S_n(r) = n\,S\left(\frac{r}{n}\right)\,, \text{ where } S(r) = \textstyle\int_0^r S'(r)dr, \quad S'(r) = \begin{cases} 1 & \text{if } |s| \le 1 \\ 2-|s| & \text{if } 1 < |s| \le 2 \\ 0 & \text{if } |s| > 2 \end{cases} \tag{2.8.4}$$

so that S_n is a sequence of bounded functions which converges to the identity locally uniformly as $n \to \infty$. Let ξ_k be the cut-off function defined in (2.8.1). By choosing $(1 - S_n'(m_\varepsilon))\xi_k^2$ as test function in (2.4.3) one obtains

$$\begin{aligned} &\int_\Omega (m_\varepsilon - S_n(m_\varepsilon))(t)\xi_k^2\,dx + \lambda_{k+1}\frac{1}{n}\int_0^t\int_{D_k} |Dm_\varepsilon|^2\,\mathbb{1}_{\{n<m_\varepsilon<2n\}}dxds \\ &\quad \le C_k\left\{\int_0^t\int_{D_{k+1}} m_\varepsilon\,\mathbb{1}_{\{n<m_\varepsilon\}}dxds + \int_0^t\int_{D_{k+1}} |Dm_\varepsilon|\mathbb{1}_{\{n<m_\varepsilon\}}dxds\right\} \\ &\qquad + \int_\Omega (m_\varepsilon(0) - S_n(m_\varepsilon(0)))\,\xi_k^2\,dx\,, \end{aligned}$$

where we used the local ellipticity of a_ε, and the local boundedness of a_ε and b_ε. Since $0 \le r - S_n(r) \le r\,\mathbb{1}_{\{r>n\}}$, and since $m_\varepsilon(0)$ converges in $L^1(K)$ for all compact subsets K, last term is small as $n \to \infty$, uniformly with respect to ε. The same holds for the previous terms in the right-hand side due to the local bounds on m_ε, Dm_ε. Hence it holds

$$\lim_{n\to\infty} \sup_\varepsilon \frac{1}{n} \int_0^T \int_\Omega |Dm_\varepsilon|^2 \mathbb{1}_{\{n<m_\varepsilon<2n\}} \xi^2 \, dxdt = 0 \tag{2.8.5}$$

and

$$\lim_{n\to\infty} \sup_{\{\varepsilon>0, t\in[0,T]\}} \int_\Omega (m_\varepsilon - S_n(m_\varepsilon))(t)\xi^2 \, dx = 0 \tag{2.8.6}$$

for any cut-off function ξ.

Now one can renormalize the equation for m_ε. Indeed, thanks to (2.8.5) the function $S_n(m_\varepsilon)$ satisfies

$$(S_n(m_\varepsilon))_t - \mathrm{div}(a_\varepsilon^*(x)DS_n(m_\varepsilon) + b_\varepsilon \, m_\varepsilon \, S_n'(m_\varepsilon)) = R_{\varepsilon,n} \quad \text{in } (0,T)\times\Omega_\varepsilon \tag{2.8.7}$$

where $R_{\varepsilon,n}$ is such that

$$\lim_{n\to\infty} \sup_\varepsilon \int_0^T \int_\Omega |R_{\varepsilon,n}| \, \xi^2 \, dxdt = 0 \, . \tag{2.8.8}$$

Consider now a sequence $\{m_j^n\}$ in $L^2(0,T;W^{1,2}_{loc}(\Omega))$ approximating the function $S_n(m)$ with the following properties :

$$\begin{cases} \partial_t m_j^n = -j(m_j^n - S_n(m)) \, , & \|m_j^n\|_\infty \le n \\ m_j^n \overset{j\to\infty}{\to} S_n(m) \quad \text{in } L^2(0,T;W^{1,2}(K)), & m_j^n(0) \overset{j\to\infty}{\to} S_n(m_0) \quad \text{in } L^1(K), \end{cases}$$

for any compact $K \subset \Omega$. We take $T_1(S_n(m_\varepsilon) - m_j^n)\,\xi^2$ as test function in (2.8.7) and we obtain (we denote $\Theta_1(r) = \int_0^r T_1(s)ds$) :

$$\begin{aligned} &\int_\Omega \Theta_1\left(S_n(m_\varepsilon) - m_j^n\right)(t)\,\xi^2 dx + C\int_0^t \int_\Omega |DT_1(S_n(m_\varepsilon) - m_j^n)|^2\,\xi^2 \\ &\le \int_\Omega \Theta_1\left(S_n(m_\varepsilon(0)) - m_j^n(0)\right)\,\xi^2 dx - \int_0^t \int_\Omega b_\varepsilon \, m_\varepsilon \, S_n'(m_\varepsilon) DT_1(S_n(m_\varepsilon) - m_j^n)\,\xi^2 \\ &\quad - \int_0^t \int_\Omega a_\varepsilon^*(x)\, Dm_j^n \, DT_1(S_n(m_\varepsilon) - m_j^n)\,\xi^2 \, dxds - \int_0^t \int_\Omega (m_j^n)_t T_1(S_n(m_\varepsilon) - m_j^n)\,\xi^2 dxds \\ &\quad 2\int_0^t \int_\Omega [a_\varepsilon^*(x)\, DS_n(m_\varepsilon) + b_\varepsilon \, m_\varepsilon \, S_n'(m_\varepsilon)] D\xi \, T_1(S_n(m_\varepsilon) - m_j^n)\,\xi \, dxds \\ &\quad + \sup_\varepsilon \int_0^T \int_\Omega |R_{\varepsilon,n}|\,\xi^2 \, dxdt \, . \end{aligned}$$

We will let first $\varepsilon \to 0$ and then $j \to \infty$. Since all the integrals are localized in a compact subset of Ω, we use the almost everywhere convergence of m_ε (hence $S_n(m_\varepsilon)$ converges in L^1) and the weak convergence of $DS_n(m_\varepsilon)$ in L^2, as well as the convergences of $a_\varepsilon(x)$ and $b_\varepsilon(t,x)$. Then we find, as $\varepsilon \to 0$:

$$
\begin{aligned}
&\limsup_{\varepsilon\to 0} \int_\Omega \Theta_1\left(S_n(m_\varepsilon) - m_j^n\right)(t)\,\xi^2 dx \\
&\le \int_\Omega \Theta_1\left(S_n(m_0) - m_j^n(0)\right)\,\xi^2 dx - \int_0^t \int_\Omega b\, m\, S_n'(m) DT_1(S_n(m) - m_j^n)\,\xi^2 \\
&\quad - \int_0^t \int_\Omega a^*(x)\, Dm_j^n\, DT_1(S_n(m) - m_j^n)\,\xi^2\, dxds - \int_0^t \int_\Omega (m_j^n)_t T_1(S_n(m) - m_j^n)\,\xi^2 dxds \\
&\quad 2\int_0^t \int_\Omega [a^*(x)\, DS_n(m) + b\, m\, S_n'(m)] D\xi\, T_1(S_n(m) - m_j^n)\,\xi\, dxds \\
&\quad + \sup_\varepsilon \int_0^T \int_\Omega |R_{\varepsilon,n}|\,\xi^2\, dxdt\,.
\end{aligned}
$$

Thanks to the properties of m_j^n we have $(m_j^n)_t T_1(S_n(m) - m_j^n) \ge 0$. Hence that term can be dropped. For all other terms, we let $j \to \infty$; using that m_j^n converges to $S_n(m)$ and $m_j^n(0)$ to $S_n(m_0)$ we conclude that

$$
\limsup_{j\to\infty} \limsup_{\varepsilon\to 0} \int_\Omega \Theta_1\left(S_n(m_\varepsilon) - m_j^n\right)(t)\,\xi^2 dx \le \omega(n)\,, \qquad \omega(n) := \sup_\varepsilon \int_0^T \int_\Omega |R_{\varepsilon,n}|\,\xi^2\, dxdt\,. \tag{2.8.9}
$$

Last term will vanish as $n \to \infty$ due to (2.8.8). Now we estimate

$$
\int_\Omega |S_n(m_\varepsilon) - S_n(m)|(t)\,\xi^2\, dx \le \int_\Omega |S_n(m_\varepsilon) - m_j^n|(t)\,\xi^2 dx + \int_\Omega |S_n(m) - m_j^n|(t)\,\xi^2 dx\,.
$$

Using that $|s| \le C\,\max(\Theta_1(s), \sqrt{\Theta_1(s)})$, due to (**??**) and to the convergence of m_j^n towards $S_n(m)$, we get after letting $\varepsilon \to 0$ and $j \to \infty$:

$$
\limsup_{\varepsilon\to 0} \int_\Omega |S_n(m_\varepsilon) - S_n(m)|(t)\,\xi^2\, dx \le C\omega(n) \overset{n\to\infty}{\to} 0\,,
$$

and the above holds uniformly with respect to $t \in [0,T]$. Putting together this estimate with (2.8.6), we conclude that

$$
\lim_{\varepsilon\to 0} \sup_{[0,T]} \int_\Omega |m_\varepsilon(t) - m|(t)\,\xi^2\, dx = 0\,.
$$

Hence $m \in C^0([0,T]; L^1_{loc}(\Omega))$ and the uniform convergence holds. □

Chapitre 3

The Master Equation in a Bounded Domain with Neumann Conditions

3.1 Introduction

In this chapter we analyze the asymptotic behaviour of an N-players differential game, where each player chooses his own control and play his dynamic in a bounded domain $\Omega \subseteq \mathbb{R}^d$, with a reflecting process at the boundary $\partial\Omega$.

As said in the Introduction of my thesis, this convergence problem is studied using the so-called *Master Equation*, an equation in the space of measures introduced by P.-L. Lions in his courses at College de France, who gives also a formal link with the Nash system. See [80].

There are many papers about the well-posedness of the Master Equation. Buckhdan, Li, Peng and Rainer in [23] proved the existence of a classical solution using probabilistic arguments, when there is no coupling and no common noise. We point out here also the work of Chassagneux, Crisan and Delarue in [42], who gave the first existence result of solution on the Master Equation with diffusion and without common noise. Furthermore, Gangbo and Swiech proved a short time existence for the Master Equation with common noise, see [59].

But the most important result in this framework was achieved by Cardaliaguet, Delarue, Lasry and Lions in [32], who proved existence and uniqueness of solutions for the Master Equation with and without common noise, and also the convergence problem, in a periodic setting ($\Omega = \mathbb{T}^d$).

Many other results about the convergence problem are given in the literature. The convergence in the whole space, under weaker condition than in [32], was given by Carmona and Delarue in [35]. In [31], Cardaliaguet, Cirant and Porretta studied the convergence for the major-minor problem. Very important are the works of Delarue, Lacker and Ramanan, who used the Master Equation for the analysis of the large deviation problem and the central limit theorem, see [48, 49]. As regards finite state problems, some recents developments were studied by Bayraktar and Cohen in [14] and by Cecchin and Pelino in [41].

There also convergence result obtained without using Master Equation. See, for example, the work by Lacker in [73]. Other important papers about Master Equation and convergence problem are [30, 37, 40, 52, 56, 72, 87].

However, most of these papers proves the convergence in a periodic framework ($\Omega = \mathbb{T}^d$) or in the whole space $\mathbb{R}^d$. But in many applications it is useful to work in a framework with boundary conditions, in particular reflections at the boundary ; see for instance the models analyzed by Achdou, Bardi and Cirant in [1].

Here we want to extend the convergence results in our context. The chapter is clearly inspired by the ideas of [32], but, as already said in the introduction, many issues appear in the further technical estimates, and more effort has to be done in order to gain the same results.

The chapter is divided as follows. In section 3.2 we define some useful tools and we state the main assumptions we will need in order to prove the next results.

The rest of the chapter is divided into two parts. In the first part (section 3.3 to 3.6), we analyze the well-posedness of the Master Equation with a Neumann condition on $\partial\Omega$ with respect to x and m :

$$\begin{cases} -\partial_t U(t,x,m) - \operatorname{tr}\left(a(x)D_x^2 U(t,x,m)\right) + H\left(x, D_x U(t,x,m)\right) \\ -\displaystyle\int_\Omega \operatorname{tr}\left(a(y)D_y D_m U(t,x,m,y)\right) dm(y) + \\ \displaystyle\int_\Omega D_m U(t,x,m,y)\cdot H_p(y, D_x U(t,y,m)) dm(y) = F(x,m) \\ \text{in } (0,T)\times\Omega\times\mathcal{P}(\Omega)\,, \\ \\ U(T,x,m) = G(x,m) \qquad \text{in } \Omega\times\mathcal{P}(\Omega)\,, \\ a(x)D_x U(t,x,m)\cdot\nu(x) = 0 \qquad \text{for } (t,x,m)\in(0,T)\times\partial\Omega\times\mathcal{P}(\Omega)\,, \\ a(y)D_m U(t,x,m,y)\cdot\nu(y) = 0 \qquad \text{for } (t,x,m,y)\in(0,T)\times\Omega\times\mathcal{P}(\Omega)\times\partial\Omega\,, \end{cases} \tag{3.1.1}$$

where ν is the normal outward at $\partial\Omega$ and $D_m U$ is a derivation of U with respect to the measure, whose precise definition will be given later.

We stress the fact that the Neumann condition $aD_m U\cdot\nu = 0$ is completely new in the literature. Actually, we have to pay attention to the space where U is defined, i.e. $[0,T]\times\Omega\times\mathcal{P}(\Omega)$. Then, together with final data and Neumann condition with respect to x, there is another boundary condition caused by the boundary of $\mathcal{P}(\Omega)$.

The idea is quite classical : for each (t_0, m_0) we consider the MFG system in $[t_0,T]\times\Omega$ with a Neumann condition :

$$\begin{cases} -u_t - \operatorname{tr}(a(x)D^2u) + H(x,Du) = F(x,m(t))\,, \\ m_t - \operatorname{div}(a(x)Dm) - \operatorname{div}(m(\tilde{b} + H_p(x,Du))) = 0\,, \\ m(t_0) = m_0\,, \qquad u(x,T) = G(x,m(T))\,, \\ a(x)Du\cdot\nu(x)_{|\partial\Omega} = 0 \qquad \left(a(x)Dm + (\tilde{b} + H_p(x,Du))m\right)\nu_{|\partial\Omega} = 0\,, \end{cases} \tag{3.1.2}$$

where $\tilde{b}$ is a vector field defined as follows :

$$\tilde{b}_i(x) = \sum_{j=1}^d \frac{\partial a_{ji}}{\partial x_j}(x)\,, \qquad i = 1,\dots,d\,.$$

Then we define

$$U(t_0,x,m_0) = u(t_0,x) \tag{3.1.3}$$

and we prove that this U is a solution of the Master Equation.

Since the function U depends also on the measure, a suitable distance between these measures has to be defined : this is called the *Wasserstein distance*, which definition will be given in section 3.2. For a more detailed description about this kind of distance, see [29].

The main issue is to prove the $\mathcal{C}^1$ character of U with respect to m, and Section $3.3-3.5$ are completely devoted to prove technical results to ensure this kind of differentability.

In section 3.3 we prove a first estimate of a solution (u,m) of the Mean Field Games system, namely

$$\|u\|_{1+\frac{\alpha}{2},2+\alpha} \le C\,, \qquad \mathbf{d}_1(m(t),m(s)) \le C|t-s|^{\frac{1}{2}}\,,$$

where $\mathbf{d}_1$ is the Wasserstein distance defined in Section 3.2. This implies

$$\|U(t,\cdot,m)\|_{2+\alpha} \le C\,.$$

In section 3.4 we use the definition of U from the Mean Field Games system in order to prove a Lipschitz character of U with respect to m :

$$\|U(t,\cdot,m_1) - U(t,\cdot,m_2)\|_{2+\alpha} \le C\mathbf{d}_1(m_1,m_2)\,.$$

In section 3.5 we prove the $\mathcal{C}^1$ character of U with respect to m. This passes through different estimates on linearized MFG systems.

Once proved the $\mathcal{C}^1$ character of U, we can prove that U is actually the unique solution of the Master Equation (3.1.1). This will be done in Section 3.6.

In the second part we analyze the asymptotic behavior of the N-players game and the convergence of the Nash system towards a solution of the Master Equation. To do that, we need to give a precise stochastic interpretation of the differential game in this framework.

We start considering a finite control problem with N players, where each player chooses his own strategy in order to minimize a certain cost functional.

The dynamic of the player i, with $1 \le i \le N$, is given by the following stochastic differential equation :

$$\begin{cases} dX_t^i = b(X_t^i,\alpha_t^i)\,dt + \sqrt{2}\sigma(X_t^i)dB_t^i - dk_t^i\,, \\ X_{t_0}^i = x_0^i\,. \end{cases} \tag{3.1.4}$$

Here, α_t^i is the control, chosen from a certain set A, $b : [0,T] \times \Omega \times A \to \mathbb{R}^d$ is the *drift* function, $\sigma : \Omega \to \mathbb{R}^{d\times d}$ is the *diffusion* matrix, $a = \sigma\sigma^*$ and ν is the outward normal at $\partial\Omega$.

Moreover, $(B_t)^i$, $1 \le i \le N$ are independent d-dimensional Brownian motions, $x_0^i \in \Omega$ and k_t^i is a *reflected process along the co-normal.* According to [83], this reflected process satisfies the following properties :

$$k_t^i = \int_0^t a(X_s^i)\nu(X_s^i)\, d|k|_s^i\,, \qquad |k|_t^i = \int_0^t \mathbb{1}_{\{X_s^i \in \partial\Omega\}}\, d|k|_s^i\,,$$

where ν is the outward normal at $\partial\Omega$. This reflection along the co-normal forces the process to stay into Ω for all $t \ge 0$.

Existence results for stochastic differential equations with reflection were already obtained in [21], [54], [82], [83], [94], [95], [96], so we will not discuss about it here.

Throughout the chapter we use the notation $\mathbf{v}$ to indicate a vector of $\mathbb{R}^{Nd}$ defined by $\mathbf{v} = (v^1, \ldots, v^N)$, where v^i is an already defined vector of $\mathbb{R}^d$.

We consider a control $\boldsymbol{\alpha}.$ and a starting position $\boldsymbol{x}_0$ at time t_0. The cost for the player i is given by the following functional :

$$J_i^N(t_0, \boldsymbol{x}_0, \boldsymbol{\alpha}.) = \mathbb{E}\left[\int_{t_0}^T \left(L(X_s^i, \alpha_s^i) + F^{N,i}(\boldsymbol{X}_s)\right) ds + G^{N,i}(\boldsymbol{X}_T)\right]\,,$$

where L is called *Lagrangian* of the system and $F^{N,i}$ and $G^{N,i}$ are the *cost functions* for the player i.

In order to obtain a convergence when N is large, we need to assume that the players are *indistinguishable* : i.e., there has to be some symmetry in the cost functional.

So we assume that, for certain functions F and G,

$$F^{N,i}(\boldsymbol{x}) = F(x_i, m_{\boldsymbol{x}}^{N,i})\,, \qquad G^{N,i}(\boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i})\,,$$

with

$$m_{\boldsymbol{x}}^{N,i} = \frac{1}{N-1}\sum_{j\ne i}\delta_{x_j}\,, \quad \text{where } \delta_x \text{ is the Dirac function at } x\,. \tag{3.1.5}$$

The optimal control strategy for each players appears in the study of *Nash equilibria.* Namely, a control $\boldsymbol{\alpha}^*_.$ provides a Nash equilibrium if, for all controls $\boldsymbol{\alpha}.$ and for all i we have

$$J_i^N(t_0, \boldsymbol{x}_0, \boldsymbol{\alpha}^*_.) \le J_i^N(t_0, \boldsymbol{x}_0, \alpha_i, (\alpha_j^*)_{j\ne i})\,,$$

i.e., each player chooses his optimal strategy, if we "freeze" the other players' strategies.

Now we define the value function for the generic player i as the cost functional at the optimal control :

$$v_i^N(t_0,\boldsymbol{x}_0)=J_i^N(t_0,\boldsymbol{x}_0,\boldsymbol{\alpha}^*)\,.$$

A well-known computation, using Ito's formula and dynamic programming principle (see [57]), proves that, if v_i^N solves the so-called *Nash system*,

$$\begin{cases}-\partial_t v_i^N-\sum_j \mathrm{tr}(a(x_j)D^2_{x_jx_j}v_i^N)+H(x_i,D_{x_i}v_i^N)+\sum_{j\neq i}H_p(x_j,D_{x_j}v_j^N)\cdot D_{x_j}v_i^N=F(x_i,m_{\boldsymbol{x}}^{N,i})\,,\\ v_i^N(T,\boldsymbol{x})=G(x_i,m_{\boldsymbol{x}}^{N,i})\,,\\ a(x_j)D_{x_j}v_i^N\cdot\nu(x_j)_{|x_j\in\partial\Omega}=0\,,\qquad j=1,\cdots,N\,,\end{cases}\tag{3.1.6}$$

then v_i^N is the value function of a Nash equilibrium.

In Section 3.7 we analyze this convergence problem. This is done by defining the following functions u_i^N :

$$u_i^N(t,\boldsymbol{x})=U(t,x_i,m_{\boldsymbol{x}}^{N,i})\,,$$

and proving that u_i^N solves "almost" the Nash system, with an error of order $\frac{1}{N}$.

Then we define the following related process for v_i^N :

$$\begin{cases}dY_t^i=-H_p(Y_t^i,D_{x_i}v_i^N(t,\boldsymbol{Y}_t))\,dt+\sqrt{2}\sigma(Y_t^i)dB_t^i-dk_t^i\,,\\ Y_{t_0}^i=Z^i\,,\end{cases}$$

where $\boldsymbol{Z}=(Z_i)_i$ are i.i.d. random variables of law m_0, and we prove that

$$|u_i^N(t_0,\boldsymbol{Z})|-v_i^N(t_0,\boldsymbol{Z})|\le\frac{C}{N}\qquad\mathbb{P}-a.s.\,,$$

$$\mathbb{E}\left[\int_{t_0}^T\left|D_{x_i}v_i^N(t,\boldsymbol{Y}_t)-D_{x_i}u_i^N(t,\boldsymbol{Y}_t)\right|^2dt\right]\le\frac{C}{N^2}\,.$$

With these asymptotic estimates, we are able to prove the two main convergence results. Actually, if we define

$$m_{\boldsymbol{x}}^N:=\frac{1}{N}\sum_i\delta_{x_i}\,,$$

we have

$$\sup_i |v_i^N(t_0, \boldsymbol{x}) - U(t_0, x_i, m_{\boldsymbol{x}}^N)| \le \frac{C}{N}\,.$$

Moreover, if we set

$$w_i^N(t_0, x_i, m_0) := \int_{\Omega^{N-1}} v_i^N(t_0, \boldsymbol{x}) \prod_{j\neq i} m_0(dx_j)\,,$$

then

$$\left\| w_i^N(t_0, \cdot, m_0) - U(t_0, \cdot, m_0) \right\|_{L^1(m_0)} \le C\omega_N\,, \tag{3.1.7}$$

where

$$\omega_N = \begin{cases} CN^{-\frac{1}{d}} & \text{if} \quad d \ge 3\,, \\ CN^{-\frac{1}{2}}\log(N) & \text{if} \quad d = 2\,, \\ CN^{-\frac{1}{2}} & \text{if} \quad d = 1\,. \end{cases}$$

3.2 Notation and assumptions

Throughout this chapter, we fix a time $T > 0$. $\Omega \subset \mathbb{R}^d$ will be the closure of an open bounded set, with boundary of class $\mathcal{C}^{2+\alpha}$, and we define $Q_T := [0,T] \times \Omega$.

First, we need to define some useful tools.

For $n \ge 0$ and $\alpha \in (0,1)$ we denote with $\mathcal{C}^{n+\alpha}(\Omega)$, or simply $\mathcal{C}^{n+\alpha}$, the space of functions $\phi \in \mathcal{C}^n(\Omega)$ with, for each $\ell \in \mathbb{N}^r$, $1 \le r \le n$, the derivative $D^\ell \phi$ is Hölder continuous with Hölder constant α. The norm is defined in the following way :

$$\|\phi\|_{n+\alpha} := \sum_{|\ell|\le n} \left\| D^l \phi \right\|_\infty + \sum_{|\ell|=n} \sup_{x\neq y} \frac{|D^\ell \phi(x) - D^\ell \phi(y)|}{|x-y|^\alpha}\,.$$

Sometimes, in order to deal with Neumann boundary conditions, we will need to work with a suitable subspace of $\mathcal{C}^{n+\alpha}(\Omega)$.

So we will call $\mathcal{C}^{n+\alpha,N}(\Omega)$, or simply $\mathcal{C}^{n+\alpha,N}$, the set of functions $\phi \in \mathcal{C}^{n+\alpha}$ such that $aD\phi \cdot \nu_{|\partial\Omega} = 0$, endowed with the same norm $\|\phi\|_{n+\alpha}$.

Then, we define several parabolic spaces we will need to work with during the chapter, starting from $\mathcal{C}^{\frac{n+\alpha}{2}, n+\alpha}([0,T] \times \Omega)$.

We say that $\phi : [0,T] \times \Omega \to \mathbb{R}$ is in $\mathcal{C}^{\frac{n+\alpha}{2},n+\alpha}([0,T] \times \Omega)$ if ϕ is continuous in both variables, together with all derivatives $D_t^r D_x^s \phi$, with $2r + s \leq n$. Moreover, $\|\phi\|_{\frac{n+\alpha}{2},n+\alpha}$ is bounded, where

$$\|\phi\|_{\frac{n+\alpha}{2},n+\alpha} := \sum_{2r+s\leq n} \|D_t^r D_x^s \phi\|_\infty + \sum_{2r+s=n} \sup_t \|D_t^r D_x^s \phi(t,\cdot)\|_\alpha + \sum_{0<n+\alpha-2r-s<2} \sup_x \|D_t^r D_x^s \phi(\cdot,x)\|_{\frac{n+\alpha-2r-s}{2}} .$$

The space of continuous space-time functions which satisfy a Hölder condition in x will be denoted by $\mathcal{C}^{0,\alpha}([0,T] \times \Omega)$. It is endowed with the norm

$$\|\phi\|_{0,\alpha} = \sup_{t\in[0,T]} \|\phi(t,\cdot)\|_\alpha .$$

The same definition can be given for the space $\mathcal{C}^{\alpha,0}$. Finally, we define the space $\mathcal{C}^{1,2+\alpha}$ of functions differentiable in time and twice differentiable in space, with all derivatives in $\mathcal{C}^{0,\alpha}(\overline{Q_T})$. The natural norm for this space is

$$\|\phi\|_{1,2+\alpha} := \|\phi\|_\infty + \|\phi_t\|_{0,\alpha} + \|D_x\phi\|_\infty + \left\| D_x^2 \phi \right\|_{0,\alpha} .$$

We note that, thanks to *Lemma 5.1.1* of [85], the first order derivatives of $\phi \in \mathcal{C}^{1,2}$ satisfy also a Hölder condition in time. Namely

$$\|D_x\phi\|_{\frac{1}{2},\alpha} \leq C \|f\|_{1,2+\alpha} . \tag{3.2.1}$$

In order to study distributional solutions for the Fokker-Planck equation, we also need to define a structure for the dual spaces of regular functions.

We define, for $n \geq 0$ and $\alpha \in (0,1)$, the space $\mathcal{C}^{-(n+\alpha)}(\Omega)$, called for simplicity $\mathcal{C}^{-(n+\alpha)}$ in this article, as the dual space of $\mathcal{C}^{n+\alpha}$, endowed with the norm

$$\|\rho\|_{-(n+\alpha)} = \sup_{\|\phi\|_{n+\alpha}\leq 1} \langle \rho, \phi \rangle .$$

With the same notations we define the space $\mathcal{C}^{-(n+\alpha),N}$ as the dual space of $\mathcal{C}^{n+\alpha,N}$ endowed with the same norm :

$$\|\rho\|_{-(n+\alpha),N} = \sup_{\substack{\|\phi\|_{n+\alpha}\leq 1 \\ aD\phi\cdot\nu_{|\partial\Omega}=0}} \langle \rho, \phi \rangle .$$

Finally, for $k \geq 1$ and $1 \leq p \leq +\infty$, we can also define the space $W^{-k,p}(\Omega)$, called for simplicity $W^{-k,p}$, as the dual space of $W^{k,p}(\Omega)$, endowed with the norm

$$\|\rho\|_{W^{-k,p}} = \sup_{\|\phi\|_{W^{k,p}} \leq 1} \langle \rho, \phi \rangle \,.$$

Definition 3.2.1. *Let $m_1, m_2 \in \mathcal{P}(\Omega)$ two Borel probability measures on Ω.*
We call the Wasserstein distance *between m_1 and m_2, and we write $\mathbf{d}_1(m_1, m_2)$ the quantity*

$$\mathbf{d}_1(m_1, m_2) := \sup_{Lip(\phi) \leq 1} \int_\Omega \phi(x) d(m_1 - m_2)(x) \,. \tag{3.2.2}$$

We note that we can also write (3.2.2) as

$$\mathbf{d}_1(m_1, m_2) := \sup_{\substack{\|\phi\|_{W^{1,\infty}} \leq C \\ Lip(\phi) \leq 1}} \int_\Omega \phi(x) d(m_1 - m_2)(x) \,, \tag{3.2.3}$$

for a certain $C > 0$. Actually, for a fixed $x_0 \in \Omega$, we can restrict ourselves to the functions ϕ such that $\phi(x_0) = 0$, since

$$\int_\Omega \phi(x) d(m_1 - m_2)(x) = \int_\Omega (\phi(x) - \phi(x_0)) d(m_1 - m_2)(x) \,,$$

and these functions obviously satisfies $\|\phi\|_{W^{1,\infty}} \leq C$ for a certain $C > 0$.

We will always work with (3.2.3), where the restriction in $W^{1,\infty}$ allows us to obtain some desired estimates with respect to $\mathbf{d}_1$.

In order to give a sense to equation (3.1.1), we need to define a suitable derivation of U with respect to the measure m.

Definition 3.2.2. *Let $U : \mathcal{P}(\Omega) \to \mathbb{R}$. We say that U is of class $\mathcal{C}^1$ if there exists a continuous map $K : \mathcal{P}(\Omega) \times \Omega \to \mathbb{R}$ such that, for all m_1, $m_2 \in \mathcal{P}(\Omega)$ we have*

$$\lim_{t \to 0} \frac{U(m_1 + s(m_2 - m_1)) - U(m_1)}{s} = \int_\Omega K(m_1, x)(m_2(dx) - m_1(dx)) \,. \tag{3.2.4}$$

Note that the definition of K is up to additive constants. Then, we define the derivative $\frac{\delta U}{\delta m}$ as the unique map K satisfying (3.2.4) and the normalization convention

$$\int_\Omega K(m, x) dm(x) = 0 \,.$$

As an immediate consequence, we obtain the following fundamental equality, that we will use very often in the rest of the chapter : for each m_1, $m_2 \in \mathcal{P}(\Omega)$ we have

$$U(m_2) - U(m_1) = \int_0^1 \int_\Omega \frac{\delta U}{\delta m}((m_1) + s(m_2 - m_1), x)(m_2(dx) - m_1(dx)) \,.$$

Finally, we can define the *intrinsic derivative* of U with respect to m.

Definition 3.2.3. *Let $U : \mathcal{P}(\Omega) \to \mathbb{R}$. If U is of class $\mathcal{C}^1$ and $\frac{\delta U}{\delta m}$ is of class $\mathcal{C}^1$ with respect to the last variable, we define the intrinsic derivative $D_m U : \mathcal{P}(\Omega) \times \Omega \to \mathbb{R}^d$ as*

$$D_m U(m,x) := D_x \frac{\delta U}{\delta m}(m,x)\,.$$

We need the following assumptions

Hypotheses 3.2.4. *Assume that*

(i) (Uniform ellipticity) $\|a(\cdot)\|_{1+\alpha} < \infty$ *and* $\exists \lambda > 0$ *s.t.* $\forall \xi \in \mathbb{R}^d$ $\langle a(x)\xi, \xi\rangle \geq \lambda |\xi|^2$ *;*

(ii) $H : \Omega \times \mathbb{R}^d \to \mathbb{R}$, $G : \Omega \times \mathcal{P}(\Omega) \to \mathbb{R}$ *and* $F : \Omega \times \mathcal{P}(\Omega) \to \mathbb{R}$ *are smooth functions with* H *Lipschitz with respect to the last variable;*

(iii) $\exists C > 0$ *s.t.*

$$0 < H_{pp}(x,p) \leq C I_{d\times d}\,;$$

(iv) F *satisfies, for some* $0 < \alpha < 1$ *and* $C_F > 0$*,*

$$\int_\Omega \big(F(x,m) - F(x,m')\big)\, d(m-m')(x) \geq 0$$

and

$$\sup_{m\in\mathcal{P}(\Omega)} \left(\|F(\cdot,m)\|_\alpha + \left\| \frac{\delta F}{\delta m}(\cdot,m,\cdot) \right\|_{\alpha,2+\alpha} \right) + \operatorname{Lip}\left(\frac{\delta F}{\delta m}\right) \leq C_F\,,$$

with

$$\operatorname{Lip}\left(\frac{\delta F}{\delta m}\right) := \sup_{m_1\neq m_2} \left(\mathbf{d}_1(m_1,m_2)^{-1} \left\| \frac{\delta F}{\delta m}(\cdot,m_1,\cdot) - \frac{\delta F}{\delta m}(\cdot,m_2,\cdot) \right\|_{\alpha,1+\alpha} \right)\,;$$

(v) G *satisfies the same estimates as* F *with* α *and* $1+\alpha$ *replaced by* $2+\alpha$*, i.e.*

$$\sup_{m\in\mathcal{P}(\Omega)} \left(\|G(\cdot,m)\|_{2+\alpha} + \left\| \frac{\delta G}{\delta m}(\cdot,m,\cdot) \right\|_{2+\alpha,2+\alpha} \right) + \operatorname{Lip}\left(\frac{\delta G}{\delta m}\right) \leq C_G\,,$$

with

$$\operatorname{Lip}\left(\frac{\delta G}{\delta m}\right) := \sup_{m_1\neq m_2} \left(\mathbf{d}_1(m_1,m_2)^{-1} \left\| \frac{\delta G}{\delta m}(\cdot,m_1,\cdot) - \frac{\delta G}{\delta m}(\cdot,m_2,\cdot) \right\|_{2+\alpha,2+\alpha} \right)\,;$$

(vi) The following Neumann boundary conditions are satisfied :

$$\left\langle a(y) D_y \frac{\delta F}{\delta m}(x,m,y), \nu(y) \right\rangle_{|\partial\Omega} = 0\,, \qquad \left\langle a(y) D_y \frac{\delta G}{\delta m}(x,m,y), \nu(y) \right\rangle_{|\partial\Omega} = 0\,,$$

$$\langle a(x) D_x G(x,m), \nu(x)\rangle_{|\partial\Omega} = 0\,,$$

for all $m \in \mathcal{P}(\Omega)$*.*

Some comments about the previous hypotheses : the first five are standard hypotheses in order to obtain existence and uniqueness of solutions for the Mean Field Games system. The hypotheses about the derivative of F and G with respect to the measure will be essential in order to obtain some estimates on a linearized MFG system.

As regards hypotheses (vi), the second and the third boundary conditions are natural compatibility conditions, essential to obtain a classical solution for the MFG and the linearized MFG system. The first boundary condition will be essential in order to prove the Neumann boundary condition of $D_m U$, see Corollary 3.5.13.

With these hypotheses we are able to prove existence and uniqueness of a classical solution for the Master Equation (3.1.1).

But first, we have to prove some preliminary estimates about the Mean Field Games system and some other estimates on a linearyzed Mean Field Games system, which will be essential in order to ensure the $\mathcal{C}^1$ character of U with respect to m.

3.3 Preliminary estimates and Mean Field Games system

In this section we start giving some technical results for linear parabolic equations, which will be useful in the rest of the Chapter.

Then we will obtain some preliminary estimates for the Master Equation, obtained by a deep analysis of the Mean Field Games related system.

We start with this technical Lemma.

Lemma 3.3.1. *Suppose a satisfies (i) of Hypotheses 3.2.4, $b, f \in L^\infty(Q_T)$. Furthermore, let $\psi \in \mathcal{C}^{1+\alpha,N}(\Omega)$, with $0 \le \alpha < 1$. Then the unique solution z of the problem*

$$\begin{cases} -z_t - \operatorname{tr}(a(x)D^2 z) + b(t,x)\cdot Dz = f(t,x)\,, \\ z(T) = \psi\,, \\ aDz\cdot \nu_{|\partial\Omega} = 0 \end{cases}$$

satisfies

$$\|z\|_{\frac{1+\alpha}{2},1+\alpha} \le C\left(\|f\|_\infty + \|\psi\|_{1+\alpha}\right)\,. \tag{3.3.1}$$

Démonstration. Note that, if f and b are continuous bounded functions, with b depending only on x, this result is simply *Theorem 5.1.18* of [85]. In the general case, we argue as follows.

We can write $z = z_1 + z_2$, where z_1 satisfies

$$\begin{cases} -(z_1)_t - \mathrm{tr}(a(x)D^2 z_1) = 0\,, \\ z_1(T) = \psi\,, \\ aDz_1 \cdot \nu_{|\partial\Omega} = 0\,. \end{cases} \tag{3.3.2}$$

and z_2 satisfies

$$\begin{cases} -(z_2)_t - \mathrm{tr}(a(x)D^2 z_2) + b(t,x)\cdot Dz_2 = f(t,x) - b(t,x)\cdot Dz_1\,, \\ z_2(T) = 0\,, \\ aDz_2 \cdot \nu_{|\partial\Omega} = 0\,, \end{cases} \tag{3.3.3}$$

Since in the equation (3.3.2) of z_1 we do not have a drift term depending on time, we can apply *Theorem 5.1.18* of [85] and obtain

$$\|z_1\|_{\frac{1+\alpha}{2},1+\alpha} \le C\,\|\psi\|_{1+\alpha}\,.$$

As regards (3.3.3), obviously $z_2(T) \in W^{2,p}(\Omega)$ $\forall p$, and from the estimate of z_1 we know that $f - bDz \in L^\infty$. So we can apply the Corollary of *Theorem IV.9.1* of [74] to obtain that, that $\forall r \ge \frac{d+2}{2}$,

$$\|z_2\|_{1-\frac{d+2}{2r},2-\frac{d+2}{r}} \le C\,\|f - bDz\|_\infty \le C\left(\|f\|_\infty + \|\psi\|_{1+\alpha}\right)\,.$$

Choosing $r = \frac{d+2}{1-\alpha}$, one has $2 - \frac{d+2}{r} = 1+\alpha$, and so (3.3.1) is satisfied for z_2.

Since $z = z_1 + z_2$, estimate (3.3.1) holds also for z. This concludes the proof. □

If the data $f = 0$, we can generalize the result of Lemma 3.3.1 if ψ is only a Lipschitz function.

This result is well-known if $a \in \mathcal{C}^2(\Omega)$, by applying a classical Bernstein method. In our framework, we have the following result.

Lemma 3.3.2. *Suppose a and b be bounded continuous functions, and $\psi \in W^{1,\infty}(\Omega)$. Then the unique solution z of the problem*

$$\begin{cases} -z_t - \mathrm{tr}(a(x)D^2 z) + b(t,x)\cdot Dz = 0\,, \\ z(T) = \psi\,, \\ aDz \cdot \nu_{|\partial\Omega} = 0 \end{cases} \tag{3.3.4}$$

satisfies a Hölder condition in t and a Lipschitz condition in x, namely $\exists C$ such that

$$|z(t,x)-z(s,x)| \le C\,\|\psi\|_{W^{1,\infty}}\,|t-s|^{\frac{1}{2}}\,, \qquad |z(t,x)-z(t,y)| \le C\,\|\psi\|_{W^{1,\infty}}\,|x-y|\,. \tag{3.3.5}$$

Démonstration. If $\psi \in \mathcal{C}^{1,N}$, estimates (3.3.5) is guaranteed by (3.3.1) of Lemma 3.3.1.

In the general case, we take $\psi^n \in \mathcal{C}^1$ such that $\psi^n \to \psi$ in $\mathcal{C}([0,T]\times\Omega)$ and $\|\psi^n\|_1 \le C\,\|\psi\|_{W^{1,\infty}}$, and we want to make a suitable approximation of it in order to obtain a function $\tilde{\psi}^n \in \mathcal{C}^{1,N}$, also converging to ψ.

In order to do that, we first need some useful tools.

For $\delta > 0$, $d(\cdot)$ the distance function from $\partial\Omega$, $\Omega_\delta = \{x \in \Omega \,|\, d(x) \ge \delta\}$ and $x \in \Omega \setminus \Omega_\delta$, we consider the following ODE in $\mathbb{R}^d$:

$$\begin{cases} \xi'(t;x) = -a(\xi(t;x))\nu(\xi(t;x))\,, \\ \xi(0;x) = x\,, \end{cases} \tag{3.3.6}$$

where ν is an extension of the outward unit normal in $\Omega \setminus \Omega_\delta$. Actually, we know from [50] that

$$Dd(x)_{|\partial\Omega} = -\nu(x)\,,$$

so a suitable extension can be $\nu(x) = -Dd(x)$.

Then we consider the corresponding hitting time of $\partial\Omega_\delta$:

$$T(x) := \inf\left\{t \ge 0 \,|\, \xi(t;x) \notin \Omega \setminus \Omega_\delta\right\}\,.$$

We have that $T(x) < +\infty$ for each $x \in \Omega \setminus \Omega_\delta$. To prove that, we consider the auxiliary function

$$\Phi(t,x) = \delta - d(\xi(t;x)).$$

So, the function $T(x)$ can be rewritten as

$$T(x) = \inf\left\{t \ge 0 \,|\, \Phi(t,x) = 0\right\}\,,$$

and his finiteness is an obvious consequence of the decreasing character of Φ in time :

$$\partial_t\Phi(t,x) = -Dd(\xi(t;x))\cdot\xi'(t;x) = -\langle a(\xi(t;x))\nu(\xi(t;x)), \nu(\xi(t;x))\rangle \le -\lambda < 0\,.$$

Moreover, thanks to Dini's theorem we obtain that $T(x)$ is a $\mathcal{C}^1$ function and his gradient is given by

$$\nabla T(x) = -\frac{\nabla_x\Phi(T(x),x)}{\partial_t\Phi(T(x),x)} = \frac{\nu(\xi(T(x);x))\mathrm{Jac}_x\xi(T(x);x)}{\langle a\nu,\nu\rangle(\xi(T(x);x))}\,.$$

Actually, thanks to the regularity of a and Ω, we can differentiate w.r.t. x the ODE (3.3.6) and obtain that $\xi(t;\cdot) \in \mathcal{C}^1$.

Now we define the approximating functions $\tilde{\psi}^n$ in the following way :

$$\tilde{\psi}^n(x) = \begin{cases} \psi^n(x) & \text{if} \quad x \in \Omega_\delta \,, \\ \psi^n(\xi(T(x);x)) & \text{if} \quad x \in \Omega \setminus \Omega_\delta \,, \end{cases} \tag{3.3.7}$$

eventually considering a $\mathcal{C}^1$ regularization in $\Omega_\delta \setminus \Omega_{2\delta}$.

From the definition of $\tilde{\psi}^n$ and the $\mathcal{C}^1$ regularity of ξ and T we have $\tilde{\psi}^n \in \mathcal{C}^1$ and

$$\|\tilde{\psi}^n\|_1 \le C\,\|\psi^n\|_1 \le C\,\|\psi\|_{W^{1,\infty}} \ .$$

Moreover, since near the boundary $\tilde{\psi}^n$ is constant along the trajectories $a(\cdot)\nu(\cdot)$, we have that on $\partial\Omega$

$$a(x)D\tilde{\psi}^n(x)\cdot\nu(x)_{|\partial\Omega} = \frac{\partial\tilde{\psi}^n}{\partial(a\nu(x))}(x)_{|\partial\Omega} = 0\,,$$

so $\tilde{\psi}^n \in \mathcal{C}^{1,N}$.

Now we consider z^n as the solution of (3.3.4) with ψ replaced by $\tilde{\psi}^n$. Then Lemma 3.3.1 implies that $\tilde{\psi}^n$ satisfies

$$\|z^n\|_{\frac12,1} \le C\|\tilde{\psi}\|_1 \le C\,\|\psi\|_{W^{1,\infty}} \ .$$

Then, Ascoli-Arzelà's Theorem tells us that $\exists z$ such that $z^n \to z$ in $\mathcal{C}([0,T]\times\Omega)$. Passing to the limit in the weak formulation of z^n, we obtain that z is the unique solution of (3.3.4).

Finally, since z^n satisfies (3.3.5), we can pass to the pointwise limit when $n \to +\infty$ and obtain the estimate (3.3.5) for z. This concludes the Lemma. □

Now we start with the first estimates for the Master Equation.

The first result is obtained by the study of some regularity properties of the MFG system, uniformly in m_0.

Proposition 3.3.3. *The system* (3.1.2) *has a unique classical solution* $(u,m) \in \mathcal{C}^{1+\frac{\alpha}{2},2+\alpha} \times \mathcal{C}([0,T];\mathcal{P}(\Omega))$*, and this solution satisfies*

$$\sup_{t_1 \neq t_2} \frac{\mathbf{d}_1(m(t_1),m(t_2))}{|t_1-t_2|^{\frac12}} + \|u\|_{1+\frac{\alpha}{2},2+\alpha} \le C\,, \tag{3.3.8}$$

where C does not depend on (t_0,m_0).
Furthermore, $m(t)$ has a positive density for each $t>0$ and, if $m_0 \in \mathcal{C}^{2+\alpha}$ and satisfies the Neumann boundary condition

$$\Big(a(x)Dm_0 + (\tilde{b}(0,x) + H_p(x,Du(0,x)))m_0\Big)\cdot\nu_{|\partial\Omega} = 0\,, \tag{3.3.9}$$

then $m \in \mathcal{C}^{1+\frac{\alpha}{2},2+\alpha}$.
Finally, the solution is stable : if $m_{0n} \to m_0$ *in* $\mathcal{P}(\Omega)$*, then there is convergence of the corresponding solutions of* (3.1.2) *:* $(u_n, m_n) \to (u, m)$ *in* $\mathcal{C}^{1,2} \times \mathcal{C}([0,T];\mathcal{P}(\Omega))$.

Démonstration. We use a Schauder fixed point argument.
Let $X \subset \mathcal{C}([t_0,T];\mathcal{P}(\Omega))$ be the set

$$X := \left\{ m \in \mathcal{C}([t_0,T];\mathcal{P}(\Omega)) \text{ s.t. } \mathbf{d}_1(m(t),m(s)) \le L|t-s|^{\frac{1}{2}} \ \forall s,t \in [t_0,T] \right\} ,$$

where L is a constant that will be chosen later.
It is easy to prove that X is a convex compact set for the uniform distance.
We define a map $\Phi : X \to X$ as follows.
Given $\beta \in X$, we consider the solution of the following Hamilton-Jacobi equation

$$\begin{cases} -u_t - \mathrm{tr}(a(x)D^2u) + H(x,Du) = F(x,\beta(t)) \\ u(T) = G(x,\beta(T)) \\ a(x)Du \cdot \nu(x)_{|\partial\Omega} = 0 \end{cases} . \tag{3.3.10}$$

Thanks to hypothesis (iv) of 3.2.4, we have $F(\cdot,\beta(\cdot)) \in \mathcal{C}^{\frac{\alpha}{2},\alpha}$ and its norm is bounded by a constant independent of β. For the same reason $G(\cdot,\beta(T)) \in \mathcal{C}^{2+\alpha}$.
It is well known that these hypotheses guarantee the existence and uniqueness of a classical solution. A proof can be found in [74], *Theorem V.7.4*.
So, we can expand with Taylor formula the gradient term and obtain a linear equation satisfied by u :

$$\begin{cases} -u_t - \mathrm{tr}(a(x)D^2u) + H(x,0) + V(t,x)\cdot Du = F(x,\beta(t)) \\ u(T) = G(x,\alpha(T)) \\ a(x)Du \cdot \nu_{\partial\Omega} = 0 \end{cases} ,$$

with

$$V(t,x) := \int_0^1 H_p(x,\lambda Du(t,x))\, d\lambda \,.$$

Thanks to the Lipschitz hypothesis on H, (ii) of 3.2.4, we know that $V \in L^\infty$. So, we can use the Corollary of *Theorem IV.9.1* of [74] to obtain

$$Du \in \mathcal{C}^{\frac{\alpha}{2},\alpha} \implies V \in \mathcal{C}^{\frac{\alpha}{2},\alpha} \,.$$

So, we can apply *Theorem IV.5.3* of [74] and get

$$\|u\|_{1+\frac{\alpha}{2},2+\alpha} \leq C\left(\|F\|_{\frac{\alpha}{2},\alpha} + \|G\|_{2+\alpha}\right),$$

where the constant C does not depend on β, t_0, m_0.
Now, we define $\Phi(\beta) = m$, where $m \in \mathcal{C}([t_0,T];\mathcal{P}(\Omega))$ is the solution of the Fokker-Planck equation

$$\begin{cases} m_t - \operatorname{div}(a(x)Dm) - \operatorname{div}(m(\tilde{b}(x) + H_p(x,Du))) = 0 \\ m(t_0) = m_0 \\ \Big(a(x)Dm + (\tilde{b} + H_p(x,Du))m\Big) \cdot \nu_{|\partial\Omega} = 0 \end{cases}. \tag{3.3.11}$$

It is easy to prove that the above equation has a unique solution in the sense of distribution. A proof in a more general case will be given in the next section, in Proposition 3.5.3. We want to check that $m \in X$.
Thanks to the distributional formulation, we have

$$\begin{aligned} &\int_\Omega \phi(t,x)m(t,dx) - \int_\Omega \phi(s,x)m(s,dx) + \\ &\int_s^t \int_\Omega (-\phi_t - \operatorname{tr}(a(x)D^2\phi) + H_p(x,Du)\cdot D\phi)m(r,dx)dr = 0\,, \end{aligned} \tag{3.3.12}$$

for each $\phi \in L^\infty$ satisfying in the weak sense

$$\begin{cases} -\phi_t - \operatorname{tr}(a(x)D^2\phi) + H_p(x,Du)\cdot D\phi \in L^\infty(Q_T) \\ aD\phi\cdot\nu_{|\partial\Omega} = 0 \end{cases}.$$

Take $\psi(\cdot)$ a 1-Lipschitz function in Ω. So, we choose ϕ in the weak formulation as the solution in $[t,T]$ of the following linear equation

$$\begin{cases} -\phi_t - \operatorname{tr}(a(x)D^2\phi) + H_p(x,Du)\cdot D\phi = 0\,, \\ \phi(t) = \psi\,, \\ a(x)D\phi\cdot\nu_{|\partial\Omega} = 0\,. \end{cases} \tag{3.3.13}$$

Thanks to Lemma 3.3.2, we know that $\phi(\cdot,x) \in \mathcal{C}^{\frac{1}{2}}([0,T])$ and its Hölder norm in time is bounded uniformly if ψ is 1-Lipschitz.
Coming back to (3.3.12), we obtain

$$\int_\Omega \psi(x)(m(t,dx) - m(s,dx)) = \int_\Omega (\phi(t,x) - \phi(s,x))m(s,dx) \leq C|t-s|^{\frac{1}{2}}\,,$$

and taking the sup over the ψ 1-Lipschitz,

$$\mathbf{d}_1(m(t), m(s)) \le C|t-s|^{\frac{1}{2}}\,.$$

Choosing $L = C$, we have proved that $m \in X$.
Since X is convex and compact, to apply Schauder's theorem we only need to show the continuity of Φ.
Let $\beta_n \to \beta$, and let u_n and m_n the solutions of (3.3.10) and (3.3.11) related to β_n. Since $\{u_n\}_n$ is uniformly bounded in $\mathcal{C}^{1+\frac{\alpha}{2},2+\alpha}$, from Ascoli-Arzelà's Theorem we have $u_n \to u$ in $\mathcal{C}^{1,2}$.
To prove the convergence of $\{m_n\}_n$, we take ϕ_n as the solution of (3.3.13) with Du replaced by Du_n. Then, as before, $\{\phi_n\}_n$ is a Cauchy sequence in $\mathcal{C}^1$. Actually, the difference $\phi_{n,m} := \phi_n - \phi_m$ satisfies

$$\begin{cases} -(\phi_{n,m})_t - \mathrm{tr}(a(x)D^2\phi_{n,m}) + H_p(x, Du_n)\cdot D\phi_{n,m} = (H_p(x,Du_m) - H_p(x,Du_n))\cdot D\phi_m\,, \\ \phi_{n,m}(t) = 0\,, \\ a(x)D\phi_{n,m}\cdot \nu_{|\partial\Omega} = 0\,, \end{cases}$$

and so Lemma 3.3.1 implies

$$\|\phi_{n,m}\|_{\frac{1+\alpha}{2},1+\alpha} \le C\,\|(H_p(x,Du_m) - H_p(x,Du_n))\cdot D\phi_m\|_\infty \le C\,\|Du_m - Du_n\|_\infty \le \omega(n,k)\,,$$

where $\omega(n,k) \to 0$ when $n, k \to \infty$, and where we use Lemma 3.3.2 in order to bound $D\phi_m$ in L^∞, without compatibility conditions.
Using (3.3.12) with (m_n, ϕ_n) and (m_k, ϕ_k), for $n, k \in \mathbb{N}$, $s = 0$, and subtracting the two equalities, we get

$$\int_\Omega \psi(x)(m_n(t,dx) - m_k(t,dx)) = \int_\Omega (\phi_n(0,x) - \phi_k(0,x))m_0(dx) \le \omega(n,k)\,.$$

Taking the sup over the ψ 1-Lipschitz and over $t \in [0,T]$, we obtain

$$\sup_{t\in[0,T]} \mathbf{d}_1(m_n(t)), m_k(t)) \le \omega(n,k)\,,$$

which proves that $\{m_n\}_n$ is a Cauchy sequence in X. Then, $\exists m$ such that $m_n \to m$ in X. Passing to the limit in (3.3.11), we immediately obtain $m = \Phi(\beta)$, which conclude the proof of continuity.
So we can apply Schauder's theorem and obtain a classical solution of the problem (3.1.2). The estimate (3.3.8) follows from the above estimates for (3.3.10) and (3.3.11).

To prove the uniqueness, let (u_1, m_1), (u_2, m_2) be two solutions of (3.1.2).
We use inequality (3.3.15), whose proof will be given in the next lemma, with $m_{01}(t_0) = m_{02}(t_0) = m_0$:

$$\int_{t_0}^{T} \int_{\Omega} (H(x, Du_2) - H(x, Du_1) - H_p(x, Du_1)(Du_2 - Du_1)) m_1(t, dx) dt +$$
$$+ \int_{t_0}^{T} \int_{\Omega} (H(x, Du_1) - H(x, Du_2) - H_p(x, Du_2)(Du_1 - Du_2)) m_2(t, dx) dt \leq 0$$

Since H is strictly convex, the above inequality gives us $Du_1 = Du_2$ in the set $\{m_1 > 0\} \cup \{m_2 > 0\}$. Then m_1 and m_2 solve the same Fokker-Planck equation, and for uniqueness we have $m_1 = m_2$.
So $F(x, m_1(t)) = F(x, m_2(t))$, $G(x, m_1(T)) = G(x, m_2(T))$ and u_1 and u_2 solve the same Hamilton-Jacobi equation, which implies $u_1 = u_2$. The proof of uniqueness is complete.
Finally, if $m_0 \in \mathcal{C}^{2+\alpha}$ satisfies (3.3.9), then, splitting the divergence terms in (3.3.11), we have

$$\begin{cases} m_t - \operatorname{tr}(a(x)D^2 m) - m \operatorname{div}\left(\tilde{b}(x) + H_p(x, Du)\right) - \left(2\tilde{b}(x) + H_p(x, Du)\right) Dm = 0 \\ m(t_0) = m_0 \\ \left(a(x)Dm + (\tilde{b} + H_p(x, Du))m\right) \cdot \nu_{|\partial\Omega} = 0 \end{cases} .$$

Then, thanks to *Theorem IV.5.3* of [74], m is of class $\mathcal{C}^{1+\frac{\alpha}{2}, 2+\alpha}$.
The stability of solutions is obtained in the same way we used for the continuity of Φ. This concludes the proof. □

With this proposition, we have obtained that

$$\sup_{t \in [0,T]} \sup_{m \in \mathcal{P}(\Omega)} \|U(t, \cdot, m)\|_{2+\alpha} \leq C \,, \tag{3.3.14}$$

which gives us an initial regularity result for the function U.
To complete the previous proposition, we need the following lemma, based on the so-called *Lasry-Lions monotonicity argument*.

Lemma 3.3.4. *Let* (u_1, m_1) *and* (u_2, m_2) *be two solutions of System* (3.1.2)*, with* $m_1(t_0) =$

m_{01}, $m_2(t_0) = m_{02}$. *Then*

$$\begin{aligned} &\int_{t_0}^{T}\int_{\Omega}(H(x,Du_2)-H(x,Du_1)-H_p(x,Du_1)(Du_2-Du_1))m_1(t,dx)dt \\ &+\int_{t_0}^{T}\int_{\Omega}(H(x,Du_1)-H(x,Du_2)-H_p(x,Du_2)(Du_1-Du_2))m_2(t,dx)dt \\ \leq& -\int_{\Omega}(u_1(t_0,x)-u_2(t_0,x))(m_{01}(dx)-m_{02}(dx))\,. \end{aligned} \tag{3.3.15}$$

Démonstration. The idea is to estimate the quantity

$$\frac{d}{dt}\int_{\Omega}(u_1-u_2)(m_1(t,dx)-m_2(t,dx))\,.$$

Using (3.1.2) and integrating by parts, we find

$$\begin{aligned} &\frac{d}{dt}\int_{\Omega}(u_1-u_2)(m_1(t,dx)-m_2(t,dx)) = \\ =&-\int_{\Omega}\left(F(x,m_1(t))-F(x,m_2(t))\right)\left(m_1(t,dx)-m_2(t,dx)\right)+ \\ &-\int_{\Omega}(H(x,Du_2)-H(x,Du_1)-H_p(x,Du_1)(Du_2-Du_1))m_1(t,dx)+ \\ &-\int_{\Omega}(H(x,Du_1)-H(x,Du_2)-H_p(x,Du_2)(Du_1-Du_2))m_2(t,dx)\,. \end{aligned}$$

Integrating the above equality for $t \in [t_0, T]$ we find

$$\begin{aligned} &\int_{t_0}^{T}\int_{\Omega}(H(x,Du_2)-H(x,Du_1)-H_p(x,Du_1)(Du_2-Du_1))m_1(t,dx)dt \\ &+\int_{t_0}^{T}\int_{\Omega}(H(x,Du_1)-H(x,Du_2)-H_p(x,Du_2)(Du_1-Du_2))m_2(t,dx)dt \\ \leq&-\int_{t_0}^{T}\int_{\Omega}\left(F(x,m_1(t))-F(x,m_2(t))\right)\left(m_1(t,dx)-m_2(t,dx)\right)dt \\ &-\int_{\Omega}(G(x,m_1(T))-G(x,m_2(T)))(m_1(T,dx)-m_2(T,dx)) \\ &-\int_{\Omega}(u_1(t_0,x)-u_2(t_0,x))(m_{01}(dx)-m_{02}(dx))\,. \end{aligned}$$

Using the monotonicity of the couplings F and G we obtain (3.3.15) and we conclude. □

3.4 Lipschitz continuity of U

Proposition 3.4.1. *Let* (u_1, m_1) *and* (u_2, m_2) *be two solutions of system* (3.1.2)*, with* $m_1(t_0) = m_{01}$*,* $m_2(t_0) = m_{02}$*. Then*

$$\begin{aligned} \|u_1 - u_2\|_{1,2+\alpha} &\le C\mathbf{d}_1(m_{01}, m_{02}), \\ \sup_{t\in[t_0,T]} \mathbf{d}_1(m_1(t), m_2(t)) &\le C\mathbf{d}_1(m_{01}, m_{02}), \end{aligned} \tag{3.4.1}$$

where C *does not depend on* t_0*,* m_{01}*,* m_{02}*. In particular*

$$\sup_{t\in[0,T]} \sup_{m_1\neq m_2} \left[(\mathbf{d}_1(m_1, m_2))^{-1} \|U(t,\cdot,m_1) - U(t,\cdot,m_2)\|_{2+\alpha}\right] \le C\,.$$

So, the solution of the Master Equation is Lipschitz continuous in the measure variable. This will be essential in order to prove the $\mathcal{C}^1$ character of U with respect to m.

Démonstration. For simplicity, we show the result for $t_0 = 0$.
First step : An initial estimate. Thanks to the hypotheses on H and the Lipschitz bound of u_1 and u_2, (3.3.15) implies

$$\begin{aligned} &\int_0^T \int_\Omega |Du_1 - Du_2|^2 (m_1(t, dx) + m_2(t, dx))dt \le \\ &\le C \int_\Omega (u_1(0, x) - u_2(0, x))(m_{01}(dx) - m_{02}(dx)) \le C \|u_1 - u_2\|_{\frac{1+\alpha}{2},1+\alpha} \mathbf{d}_1(m_{01}, m_{02}). \end{aligned}$$

Second step : An estimate on $m_1 - m_2$. We call $m := m_1 - m_2$. We take ϕ a sufficiently regular function satisfying $aD\phi\cdot\nu = 0$, which will be chosen later. By subtracting the weak formulations (3.3.12) of m_1 and m_2 for $s = 0$ and for ϕ as test function, we obtain

$$\begin{aligned} &\int_\Omega \phi(t, x)m(t, dx) + \int_0^t \int_\Omega \left(-\phi_t - \mathrm{tr}(a(x)D^2\phi) + H_p(x, Du_1)D\phi\right) m(s, dx)ds + \\ &+ \int_0^t \int_\Omega (H_p(x, Du_1) - H_p(x, Du_2))D\phi\, m_2(s, dx)ds = \int_\Omega \phi(0, x)(m_{01}(dx) - m_{02}(dx))\,. \end{aligned} \tag{3.4.2}$$

We choose ϕ as the solution of (3.3.13) related to u_1, with terminal condition $\psi \in W^{1,\infty}$. Using the Lipschitz continuity of H_p with respect to p, we get

$$\int_\Omega \psi(x)m(t, dx) \le C \int_0^t \int_\Omega |Du_1 - Du_2|m_2(s, dx)ds + C\mathbf{d}_1(m_{01}, m_{02}),$$

since, for Lemma 3.3.2, ϕ is Lipschitz continuous with a constant bounded uniformly if ψ is 1-Lipschitz.
Now we use the Young's inequality and the first step to obtain

$$\int_\Omega \psi(x)m(t,dx) \le C\left(\int_0^t\int_\Omega |Du_1-Du_2|^2 m_2(s,dx)\right)^{\frac12} + C\mathbf{d}_1(m_{01},m_{02}) \le$$
$$\le C\left(\|u_1-u_2\|^{\frac12}_{\frac{1+\alpha}{2},1+\alpha}\mathbf{d}_1(m_{01},m_{02})^{\frac12} + \mathbf{d}_1(m_{01},m_{02})\right)\,,$$

and finally, taking the sup over the ψ 1-Lipschitz and the over $t\in[0,T]$,

$$\sup_{t\in[0,T]}\mathbf{d}_1(m_1(t),m_2(t)) \le C\left(\|u_1-u_2\|^{\frac12}_{\frac{1+\alpha}{2},1+\alpha}\mathbf{d}_1(m_{01},m_{02})^{\frac12} + \mathbf{d}_1(m_{01},m_{02})\right)\,. \quad (3.4.3)$$

Third step : Estimate on u_1-u_2 and conclusion. We call $u := u_1-u_2$. Then u solves the following equation

$$\begin{cases} -u_t - \mathrm{tr}(a(x)D^2u) + V(t,x)Du = f(t,x)\\ u(T) = g(x)\\ a(x)Du\cdot\nu_{|\partial\Omega} = 0\end{cases},$$

where

$$V(t,x) = \int_0^1 H_p(x,\lambda Du_1(t,x) + (1-\lambda)Du_2(t,x)d\lambda\,;$$
$$f(t,x) = \int_0^1\int_\Omega \frac{\delta F}{\delta m}(x,\lambda m_1(t)+(1-\lambda)m_2(t),y)(m_1(t,dy)-m_2(t,dy))d\lambda\,;$$
$$g(x) = \int_0^1\int_\Omega \frac{\delta G}{\delta m}(x,\lambda m_1(T)+(1-\lambda)m_2(T),y)(m_1(T,dy)-m_2(T,dy))d\lambda\,.$$

From the regularity of u_1 and u_2, we have $V(t,\cdot)$ bounded in $\mathcal{C}^{\frac{\alpha}{2},\alpha}$.
We want to apply *Theorem 5.1.21* of [85]. To do this, we have to estimate $\sup_t\|f(t,\cdot)\|_\alpha$

First, we call

$$m_\lambda(\cdot) := \lambda m_1(\cdot) + (1-\lambda)m_2(\cdot)\,.$$

We get

$$\sup_{t\in[0,T]}\|f(t,\cdot)\|_\alpha \le \sup_{t\in[0,T]}\int_0^1\left\|D_y\frac{\delta F}{\delta m}(\cdot,m_\lambda(t),\cdot)\right\|_{\alpha,\infty}d\lambda\,\mathbf{d}_1(m_1(t),m_2(t))$$
$$\le C\sup_{t\in[0,T]}\mathbf{d}_1(m_1(t),m_2(t))\,,$$

where C depends on the constant C_F in hypotheses 3.2.4.
In the same way

$$\|g(\cdot)\|_{2+\alpha} \leq C \sup_{r\in[0,T]} \mathbf{d}_1(m_1(r), m_2(r))\,. \tag{3.4.4}$$

So we can apply *Theorem 5.1.21* of [85] and obtain

$$\|u_1 - u_2\|_{1,2+\alpha} \leq C \sup_{r\in[0,T]} \mathbf{d}_1(m_1(r), m_2(r))\,. \tag{3.4.5}$$

Coming back to (3.4.3), this implies

$$\sup_{t\in[0,T]} \mathbf{d}_1(m_1(t), m_2(t)) \leq$$
$$\leq C\left(\left(\sup_{r\in[0,T]} \mathbf{d}_1(m_1(r), m_2(r))\right)^{\frac{1}{2}} \mathbf{d}_1(m_{01}, m_{02})^{\frac{1}{2}} + \mathbf{d}_1(m_{01}, m_{02})\right)\,,$$

and, using a generalized Young's inequality, this allows us to conclude :

$$\sup_{t\in[0,T]} \mathbf{d}_1(m_1(t), m_2(t)) \leq C\mathbf{d}_1(m_{01}, m_{02})\,. \tag{3.4.6}$$

Plugging this estimate in (3.4.5), we finally obtain

$$\|u_1 - u_2\|_{1,2+\alpha} \leq C\mathbf{d}_1(m_{01}, m_{02})\,.$$

□

3.5 Linearized system and differentiability of U with respect to the measure

The proof of existence and uniqueness of solutions for the Master Equation strongly relies on the $\mathcal{C}^1$ character of U with respect to m.
The definition of the derivative $\frac{\delta U}{\delta m}$ is strictly related to the solution (v, μ) of the following

linearized system :

$$\begin{cases} -v_t - \mathrm{tr}(a(x)D^2v) + H_p(x,Du)\cdot Dv = \dfrac{\delta F}{\delta m}(x,m(t))(\mu(t))\,, \\ \mu_t - \mathrm{div}(a(x)D\mu) - \mathrm{div}(\mu(H_p(x,Du)+\tilde{b})) - \mathrm{div}(mH_{pp}(x,Du)Dv) = 0\,, \\ v(T,x) = \dfrac{\delta G}{\delta m}(x,m(T))(\mu(T))\,, \qquad \mu(t_0) = \mu_0\,, \\ a(x)Dv\cdot\nu_{|\partial\Omega} = 0\,, \qquad \Big(a(x)D\mu + \mu(H_p(x,Du)+\tilde{b}) + mH_{pp}(x,Du)Dv\Big)\cdot\nu_{|\partial\Omega} = 0\,, \end{cases} \tag{3.5.1}$$

where we use the notation

$$\frac{\delta F}{\delta m}(x,m(t))(\rho(t)) := \left\langle \frac{\delta F}{\delta m}(x,m(t),\cdot),\rho(t)\right\rangle$$

and the same for G.

We want to prove that this system admits a solution and that the following equality holds :

$$v(t_0,x) = \left\langle \frac{\delta U}{\delta m}(t_0,x,m_0,\cdot),\mu_0\right\rangle\,. \tag{3.5.2}$$

First, we have to analyze separately the well-posedness of the Fokker-Planck equation in distribution sense :

$$\begin{cases} \mu_t - \mathrm{div}(a(x)D\mu) - \mathrm{div}(\mu b) = f\,, \\ \mu(0) = \mu_0\,, \\ (a(x)D\mu + \mu b)\cdot\nu_{|\partial\Omega} = 0\,, \end{cases} \tag{3.5.3}$$

where $f\in L^1(W^{-1,\infty})$, $\mu_0\in\mathcal{C}^{-(1+\alpha)}$, $b\in L^\infty$.

A suitable distributional definition of solution is the following :

Definition 3.5.1. *Let* $f\in L^1(W^{-1,\infty})$*,* $\mu_0\in\mathcal{C}^{-(1+\alpha)}$*,* $b\in L^\infty$*. We say that a function* $\mu\in\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),N})\cap L^1(Q_T)$ *is a weak solution of* (3.5.3) *if, for all* $\psi\in L^\infty(\Omega)$*,* $\xi\in\mathcal{C}^{1+\alpha,N}$ *and* ϕ *solution in* $[0,t]\times\Omega$ *of the following linear equation*

$$\begin{cases} -\phi_t - \mathrm{div}(aD\phi) + bD\phi = \psi\,, \\ \phi(t) = \xi\,, \\ aD\phi\cdot\nu_{|\partial\Omega} = 0\,, \end{cases} \tag{3.5.4}$$

the following formulation holds :

$$\langle \mu(t), \xi \rangle + \int_0^t \int_\Omega \mu(s,x)\psi(s,x)\, dxds = \langle \mu_0, \phi(0,\cdot) \rangle + \int_0^t \langle f(s), \phi(s,\cdot) \rangle \, ds \,, \tag{3.5.5}$$

where $\langle \cdot, \cdot \rangle$ *denotes the duality between* $\mathcal{C}^{-(1+\alpha),N}$ *and* $\mathcal{C}^{1+\alpha,N}$ *in the first case, between* $\mathcal{C}^{-(1+\alpha)}$ *and* $\mathcal{C}^{1+\alpha}$ *in the second case and between* $W^{-1,\infty}$ *and* $W^{1,\infty}$ *in the last case.*

We note that the definition is well-posed. Actually, $\phi(s,\cdot)$ is in $\mathcal{C}^{1+\alpha}$ $\forall s$ thanks to Lemma (3.3.1), so $\langle \mu_0, \phi(0,\cdot) \rangle$ and $\langle f(s), \phi(s,\cdot) \rangle$ are well defined. Moreover, we have

$$\|\phi(s,\cdot)\|_{W^{1,\infty}} \le C \,.$$

Hence, since $f \in L^1(W^{-1,\infty})$, the last integral is well defined too.

Remark 3.5.2. We are mainly interested in a particular case of distribution f. If there exists an integrable function $c : [0,T] \times \Omega \to \mathbb{R}^n$ such that $\forall \phi \in W^{1,\infty}$

$$\langle f(t), \phi \rangle = \int_\Omega c(t,x) \cdot D\phi(x)\, dx \,,$$

then we can write the problem (3.5.3) in this way :

$$\begin{cases} \mu_t - \operatorname{div}(a(x)D\mu) - \operatorname{div}(\mu b) = \operatorname{div}(c) \,, \\ \mu(0) = \mu_0 \,, \\ (a(x)D\mu + \mu b + c) \cdot \nu_{|\partial\Omega} = 0 \,, \end{cases}$$

writing f like a divergence and adjusting the Neumann condition, in order to make sense out of the integration by parts in the regular case.

In this case, in order to ensure the condition $f \in L^1(W^{-1,\infty})$, we can simply require $c \in L^1(Q_T)$. Actually we have, using Jensen's inequality,

$$\|f\|_{L^1(W^{-1,\infty})} = \int_0^T \sup_{\|\phi\|_{W^{1,\infty}} \le 1} \left(\int_\Omega c(t,x) \cdot D\phi(x)\, dx \right) dt \le C \int_0^T \int_\Omega |c(t,x)|\, dxdt = \|c\|_{L^1} \,,$$

where $|\cdot|$ is any equivalent norm in $\mathbb{R}^d$.

The next Proposition gives us an exhaustive existence and uniqueness result for (3.5.3).

Proposition 3.5.3. *Let $f \in L^1(W^{-1,\infty})$, $\mu_0 \in \mathcal{C}^{-(1+\alpha)}$, $b \in L^\infty$. Then there exists a unique solution of the Fokker-Planck equation* (3.5.3).

This solution satisfies

$$\sup_t \|\mu(t)\|_{-(1+\alpha),N} + \|\mu\|_{L^p} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|f\|_{L^1(W^{-1,\infty})}\right), \tag{3.5.6}$$

where $p = \frac{d+2}{d+1+\alpha}$.

Finally, the solution is stable : if $\mu_0^n \to \mu_0$ in $\mathcal{C}^{-(1+\alpha)}$, $\{b^n\}_n$ uniformly bounded and $b^n \to b$ in L^p $\forall p$, $f^n \to f$ in $L^1(W^{-1,\infty})$, then, calling μ^n and μ the solutions related, respectively, to (μ_0^n, b^n, f^n) and (μ_0, b, f), we have $\mu^n \to \mu$ in $\mathcal{C}([0,T]; \mathcal{C}^{-(1+\alpha),N}) \cap L^p(Q_T)$.

Démonstration. For the existence part, we start assuming that f, b, μ_0 are smooth functions, and that μ_0 satisfies

$$(a(x)D\mu_0 + \mu_0 b) \cdot \nu_{|\partial\Omega} = 0\,. \tag{3.5.7}$$

In this case, we can split the divergence terms in (3.5.3) and obtain that μ is a solution of a linear equation with smooth coefficients. So the existence of solutions in this case is a straightforward consequence of the classical results in [74], [85].

We consider the unique solution ϕ of (3.5.4) with $\psi = 0$ and $\xi \in \mathcal{C}^{1+\alpha,N}$. Multiplying the equation of μ for ϕ and integrating by parts in $[0,t] \times \Omega$ we obtain

$$\langle \mu(t), \xi \rangle = \langle \mu_0, \phi(0,\cdot) \rangle + \int_0^t \langle f(s), \phi(s,\cdot) \rangle \, ds\,. \tag{3.5.8}$$

Thanks to Lemma 3.3.1, we know that

$$\|\phi\|_{\frac{1+\alpha}{2},1+\alpha} \le C\,\|\xi\|_{1+\alpha}\,. \tag{3.5.9}$$

Then the right hand side term of (3.5.8) is bounded in this way :

$$\langle \mu_0, \phi(0,\cdot) \rangle + \int_0^t \langle f(s), \phi(s,\cdot) \rangle \, ds \le C\,\|\xi\|_{1+\alpha}\left(\|\mu_0\|_{-(1+\alpha)} + \int_0^t \|f(s)\|_{W^{1,\infty}}\right).$$

Coming back to (3.5.8) and passing to the *sup* when $\xi \in \mathcal{C}^{1+\alpha,N}$, $\|\xi\|_{1+\alpha} \le 1$, we obtain

$$\sup_t \|\mu(t)\|_{-(1+\alpha),N} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|f\|_{L^1(W^{-1,\infty})}\right). \tag{3.5.10}$$

Now we have to prove the L^p estimate. We consider the solution of (3.5.4) with $t = T$, $\xi = 0$ and $\psi \in L^r$, with $r > d+2$ (we recall that in this chapter we call d the dimension of the space.).

Then the Corollary of *Theorem IV.9.1* of [74] tells us that

$$\|\phi\|_{1-\frac{d+2}{2r},2-\frac{d+2}{r}} \le C\,\|\psi\|_{L^r}\,. \tag{3.5.11}$$

Choosing $r=\frac{d+2}{1-\alpha}$, one has $2-\frac{d+2}{r}=1+\alpha$. Integrating in $[0,T]\times\Omega$ the equation of μ one has

$$\int_0^T\int_\Omega \mu\psi\,dxds = \langle \mu_0, \phi(0,\cdot)\rangle + \int_0^T \langle f(s), \phi(s,\cdot)\rangle\,ds\,.$$

Thanks to (3.5.11) we can estimate the terms on the right-hand side and obtain

$$\int_0^T\int_\Omega \mu\psi\,dxds \le C\,\|\psi\|_{L^r}\left(\|\mu_0\|_{-(1+\alpha)} + \|f\|_{L^1(W^{-1,\infty})}\right)\,. \tag{3.5.12}$$

Passing to the *sup* for $\|\psi\|_{L^r}\le 1$, we finally get

$$\|\mu\|_{L^p} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|f\|_{L^1(W^{-1,\infty})}\right)\,,$$

with p defined as the conjugate exponent of r, i.e. $p=\frac{d+2}{d+1+\alpha}$.
This proves estimates (3.5.6) in the regular case.

In the general case, we consider suitable smooth approximations μ_0^k, f^k, b^k converging to μ_0, f, b respectively in $\mathcal{C}^{-(1+\alpha),N}$, $L^1(W^{-1,\infty})$ and $L^q(Q_T)$ $\forall q\ge 1$, with b_k bounded uniformly in k and with μ_0^k satisying (3.5.7).

We call μ^k the related solution of (3.5.3). The above convergences tells us that, for a certain C,

$$\|\mu_0^k\|_{-(1+\alpha)} \le C\,\|\mu_0\|_{-(1+\alpha)}\,, \qquad \|b_k\|_\infty \le C\,\|b\|_\infty$$
$$\|f^k\|_{L^1(W^{-1,\infty})} \le C\,\|f\|_{L^1(W^{-1,\infty})}\,.$$

Then we apply (3.5.6), to obtain, uniformly in k,

$$\sup_t \|\mu^k(t)\|_{-(1+\alpha),N} + \|\mu^k\|_{L^p} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|f\|_{L^1(W^{-1,\infty})}\right)\,, \tag{3.5.13}$$

where C actually depends on b^k, but since $b^k\to b$ it is bounded uniformly in k.

Moreover, the function $\mu^{k,h}:=\mu^k-\mu^h$ also satisfies (3.5.3) with data $b=b^k$, $f=f^k-f^h+\mathrm{div}(\mu^h(b^k-b^h))$, $\mu^0=\mu_0^k-\mu_0^h$. Then estimates (3.5.6) tell us that

$$\begin{aligned}&\sup_t \|\mu^{k,h}(t)\|_{-(1+\alpha),N} + \|\mu^{k,h}\|_{L^p}\\ &\le C\left(\|\mu_0^k-\mu_0^h\|_{-(1+\alpha)} + \|f^k-f^h\|_{L^1(W^{-1,\infty})} + \|\mathrm{div}(\mu^h(b^k-b^h))\|_{L^1(W^{-1,\infty})}\right)\,,\end{aligned}$$

The first two terms in the right-hand side easily go to 0 when $h, k \to +\infty$, since μ_0^k and f^k are Cauchy sequences. As regards the last term, calling p' the conjugate exponent of p, we have

$$\|\mathrm{div}(\mu^h(b^k - b^h))\|_{L^1(W^{-1,\infty})} \leq C \int_0^T \int_\Omega \left|\mu^h(b^k - b^h)\right| \, dxdt \leq C\|b^k - b^h\|_{L^{p'}} \,, \qquad (3.5.14)$$

since μ^k is bounded in L^p by (3.5.13) (here C depends also on μ_0 and f). So, also the last term goes to 0 since b^k is a Cauchy sequence in L^q $\forall q \geq 1$.

Hence, $\{\mu^k\}_k$ is a Cauchy sequence, and so there exists $\mu \in \mathcal{C}([0,T]; \mathcal{C}^{-(1+\alpha),N}) \cap L^p(Q_T)$ such that

$$\mu^k \to \mu \qquad \text{strongly in } \mathcal{C}([0,T]; \mathcal{C}^{-(1+\alpha),N}) \,, \text{ strongly in } L^p(Q_T) \,.$$

Furthermore, μ satisfies (3.5.6).

To conclude, we have to prove that μ is actually a solution of (3.5.3) in the sense of Definition 3.5.1.

We take ϕ and ϕ^k as the solutions of (3.5.4) related to b and b^k. The weak formulation for μ^k implies that

$$\langle \mu^k(t), \xi \rangle + \int_0^t \int_\Omega \mu^k(s,x)\psi(s,x) \, dxds = \langle \mu_0^k, \phi^k(0,\cdot) \rangle + \int_0^t \langle f^k(s), \phi^k(s,\cdot) \rangle \, ds \,,$$

We can immediately pass to the limit in the left-hand side, using the convergence of μ^k previously obtained.

For the right-hand side, we first need to prove the convergence of ϕ^k towards ϕ. This is immediate : actually, the function $\tilde{\phi}^k := \phi^k - \phi$ satisfies

$$\begin{cases} -\tilde{\phi}_t^k - \mathrm{div}(aD\tilde{\phi}^k) + b^k D\tilde{\phi}^k = (b^k - b)D\phi \,, \\ \tilde{\phi}^k(t) = 0 \,, \\ aD\phi \cdot \nu_{|\partial\Omega} = 0 \,. \end{cases}$$

Then, the Corollary of *Theorem IV.9.1* of [74] implies, for a certain $q > d + 2$ and depending on α,

$$\|\tilde{\phi}^k\|_{\frac{1+\alpha}{2}, 1+\alpha} \leq C\|(b^k - b)D\phi\|_{L^q} \to 0 \,,$$

since $D\phi$ is bounded in L^∞ using Lemma 3.3.1.

Hence, $\phi^k \to \phi$ in $\mathcal{C}^{\frac{1+\alpha}{2}, 1+\alpha}$. This allows us to pass to the limit in the right-hand side too and prove that (3.5.5) holds true, and so that μ is a weak solution of (3.5.3). This concludes the existence part.

For the uniqueness part, we consider μ_1 and μ_2 two weak solutions of the system. Then, by linearity, the function $\mu := \mu_1 - \mu_2$ is a weak solution of

$$\begin{cases} \mu_t - \operatorname{div}(a(x)D\mu) - \operatorname{div}(\mu b) = 0\,, \\ \mu(0) = 0\,, \\ (a(x)D\mu + \mu b)\cdot \nu_{|\partial\Omega} = 0\,. \end{cases}$$

Hence, the weak estimation (3.5.5) implies, $\forall \psi \in L^\infty$ and $\forall \xi \in \mathcal{C}^{1+\alpha,N}$,

$$\langle \mu(t), \xi\rangle + \int_0^t \int_\Omega \mu(s,x)\psi(s,x)\,dxds = 0\,,$$

which implies

$$\|\mu\|_{L^1} = \sup_{t\in[0,T]} \|\mu(t)\|_{-(1+\alpha),N} = 0$$

and concludes the uniquess part.

Finally, the stability part is an easy consequence of the estimates obtained previously. Let $f^n \to f$, $\mu_0^n \to \mu_0$ and $b^n \to b$. Then the function $\tilde{\mu}^n := \mu^n - \mu$ satisfies (3.5.3) with $b\,, \mu_0$ and f replaced by b^n, $\mu_0^n - \mu_0$, $f^n - f + \operatorname{div}(\mu(b^n - b))$. Then we use (3.5.6) to obtain

$$\begin{aligned} &\sup_t \|\tilde{\mu}^n\|_{-(1+\alpha),N} + \|\tilde{\mu}^n\|_{L^p} \\ &\le C\left(\|\mu_0^n - \mu_0\|_{-(1+\alpha)} + \|f^n - f\|_{L^1(W^{-1,\infty})} + \|\operatorname{div}(\mu(b^n - b))\|_{L^1(W^{-1,\infty})}\right)\,, \end{aligned}$$

The first two terms in the right-hand side go to 0. For the last term, the same computations of (3.5.14) imply

$$\|\operatorname{div}(\mu(b^n - b))\|_{L^1(W^{-1,\infty})} \le C\|b^n - b\|_{L^{p'}} \to 0\,.$$

Then $\mu^n \to \mu$ in $\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),N}) \cap L^p(Q_T)$, which concludes the Proposition. □

The last proposition allows us to get another regularity result of μ, when the data b is more regular. This result will be essential in order to improve the regularity of $\frac{\delta U}{\delta m}$ with respect to y.

Corollary 3.5.4. *Let* $\mu_0 \in \mathcal{C}^{-(1+\alpha)}$, $f \in L^1(W^{-1,\infty})$, $b \in \mathcal{C}^{\frac{\alpha}{2},\alpha}$. *Then the unique solution* μ *of* (3.5.3) *satisfies*

$$\sup_{t\in[0,T]} \|\mu(t)\|_{-(2+\alpha),N} \le C\left(\|\mu_0\|_{-(2+\alpha)} + \|f\|_{L^1(W^{-1,\infty})}\right)\,. \tag{3.5.15}$$

Démonstration. We take ϕ as the solution of (3.5.4), with $\xi \in C^{2+\alpha,N}(\Omega)$ and $\psi = 0$. Then we know from the classical results of [74], [85] (it is important here that $b \in \mathcal{C}^{\frac{\alpha}{2},\alpha}$), that

$$\|\phi\|_{1+\frac{\alpha}{2},2+\alpha} \le C\,\|\xi\|_{2+\alpha}\,.$$

The weak formulation of μ (3.5.5) tells us that

$$\langle \mu(t),\xi\rangle = \langle \mu_0,\phi(0,\cdot)\rangle + \int_0^T \langle f(s),\phi(s,\cdot)\rangle\,ds \le C\left(\|\mu_0\|_{-(2+\alpha)} + \|f\|_{L^1(W^{-1,\infty})}\right)\|\xi\|_{2+\alpha}\,.$$

Hence, we can pass to the *sup* for $\xi \in \mathcal{C}^{2+\alpha,N}$ with $\|\xi\|_{2+\alpha} \le 1$ and obtain (3.5.15). □

Remark 3.5.5. We stress the fact that *we shall not formulate problem* (3.5.3) *directly with* $\mu_0 \in \mathcal{C}^{-(2+\alpha)}$. Actually, the core of the existence theorem is the L^p bound in space-time of μ, and this is obtained by duality, considering test functions ϕ with data $\psi \in L^r$. For this function it is not guaranteed that $\phi(0,\cdot) \in \mathcal{C}^{2+\alpha}(\Omega)$, and an estimation like (3.5.12) is no longer possible.

We can also obtain some useful estimates for the density function m, as stated in the next result.

Corollary 3.5.6. *Let* (u,m) *be the solution of the MFG system defined in* (3.1.2). *Then we have* $m \in L^p(Q_T)$ *for* $p = \frac{d+2}{d+1+\alpha}$, *with*

$$\|m\|_{L^p} \le C\,\|m_0\|_{-(1+\alpha)}\,. \tag{3.5.16}$$

Furthermore, if (u_1,m_1) *and* (u_2,m_2) *are two solutions of* (3.1.2) *with initial conditions* m_{01} *and* m_{02}, *then we have*

$$\|m_1 - m_2\|_{L^p(Q_T)} \le C\mathbf{d}_1(m_{01},m_{02})\,. \tag{3.5.17}$$

Démonstration. Since m satisfies (3.5.3) with $\mu = m_0 \in \mathcal{P}(\Omega) \subset \mathcal{C}^{-(1+\alpha)}$, $b = H_p(x,Du) + \tilde{b} \in L^\infty$ and $f = 0$, inequality (3.5.16) comes from Proposition 3.5.3.

For the second inequality, we consider $m := m_1 - m_2$. Then m solves the equation

$$\begin{cases} m_t - \mathrm{div}(aDm) - \mathrm{div}(m(H_p(x,Du_1)+\tilde{b})) = \mathrm{div}(m_2(H_p(x,Du_2) - H_p(x,Du_1)))\,,\\ m(t_0) = m_{01} - m_{02}\,,\\ \Big[aDm + m\tilde{b} + m_1 H_p(x,Du_1) - m_2 H_p(x,Du_2)\Big]\cdot \nu_{|\partial\Omega} = 0\,, \end{cases}$$

i.e. m is a solution of (3.5.3) with $f = \mathrm{div}(m_2(H_p(x, Du_2) - H_p(x, Du_1)))$, $\mu_0 = m_{01} - m_{02}$, $b = H_p(x, Du_1)$. Then estimations (3.5.6) imply

$$\|m_1 - m_2\|_{L^p(Q_T)} \leq C \left(\|\mu_0\|_{-(1+\alpha)} + \|f\|_{L^1(W^{-1,\infty})} \right) .$$

We estimate the right-hand side term. As regards μ_0 we have

$$\|\mu_0\|_{-(1+\alpha)} = \sup_{\|\phi\|_{1+\alpha} \leq 1} \int_\Omega \phi(x)(m_{01} - m_{02})(dx) \leq C\mathbf{d}_1(m_{01}, m_{02}) .$$

For the f term we argue in the following way :

$$\begin{aligned} \|f\|_{L^1(W^{-1,\infty})} &= \int_0^T \sup_{\|\phi\|_{W^{1,\infty}} \leq 1} \left(\int_\Omega H_p(x, Du_2) - H_p(x, Du_1) D\phi \, m_2(t, dx) \right) dt \\ &\leq C \, \|u_1 - u_2\|_{\frac{1+\alpha}{2}, 1+\alpha} \leq C\mathbf{d}_1(m_{01}, m_{02}) , \end{aligned}$$

which allows us to conclude. □

In order to prove the representation formula (3.5.2), we need to obtain some estimates for a more general linearized system of the form

$$\begin{cases} -z_t - \mathrm{tr}(a(x)D^2 z) + H_p(x, Du)Dz = \dfrac{\delta F}{\delta m}(x, m(t))(\rho(t)) + h(t, x) , \\ \rho_t - \mathrm{div}(a(x)D\rho) - \mathrm{div}(\rho(H_p(x, Du) + \tilde{b})) - \mathrm{div}(mH_{pp}(x, Du)Dz + c) = 0 , \\ z(T, x) = \dfrac{\delta G}{\delta m}(x, m(T))(\rho(T)) + z_T(x) , \qquad \rho(t_0) = \rho_0 , \\ a(x)Dz \cdot \nu_{|\partial\Omega} = 0 , \quad \left(a(x)D\rho + \rho(H_p(x, Du) + \tilde{b}) + mH_{pp}(x, Du)Dz + c \right) \cdot \nu_{|\partial\Omega} = 0 , \end{cases} \tag{3.5.18}$$

where we require

$$z_T \in \mathcal{C}^{2+\alpha}, \quad \rho_0 \in \mathcal{C}^{-(1+\alpha)}, \quad h \in \mathcal{C}^{0,\alpha}([t_0, T] \times \Omega), \quad c \in L^1([t_0, T] \times \Omega) .$$

Moreover, z_T satisfies

$$aDz_T \cdot \nu_{|\partial\Omega} = 0 . \tag{3.5.19}$$

A suitable definition of solution for this system is the following :

Definition 3.5.7. *We say that a couple* $(z, \rho) \in \mathcal{C}^{1,2+\alpha} \times \left(\mathcal{C}([0, T]; \mathcal{C}^{-(1+\alpha),N}(\Omega)) \cap L^1(Q_T) \right)$ *is a solution of the equation* (3.5.18) *if*

— z is a classical solution of the linear equation ;

— *ρ is a distributional solution of the Fokker-Planck equation in the sense of Definition 3.5.1.*

We start with the following existence result.

Proposition 3.5.8. *Let hypotheses 3.2.4 hold for $0<\alpha<1$. Then there exists a unique solution $(z,\rho)\in\mathcal{C}^{1,2+\alpha}\times\big(\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),N}(\Omega))\cap L^1(Q_T)\big)$ of system (3.5.18). This solution satisfies, for a certain $p>1$,*

$$\|z\|_{1,2+\alpha}+\sup_t\|\rho(t)\|_{-(1+\alpha),N}+\|\rho\|_{L^p}\le CM\,, \tag{3.5.20}$$

where C depends on H and where M is given by

$$M:=\|z_T\|_{2+\alpha}+\|\rho_0\|_{-(1+\alpha)}+\|h\|_{0,\alpha}+\|c\|_{L^1}\,. \tag{3.5.21}$$

Démonstration. As always, we can assume $t_0=0$ without loss of generality.

The main idea is to apply Schaefer's Theorem.

Step 1 : Definition of the map $\mathbf{\Phi}$ *satisfying Schaefer's Theorem.* We set $X:=\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),N})$, endowed with the norm

$$\|\phi\|_X:=\sup_{t\in[0,T]}\|\phi(t)\|_{-(1+\alpha),N}\,.$$

For $\rho\in X$, we consider the classical solution z of the following equation

$$\begin{cases}-z_t-\operatorname{tr}(a(x)D^2z)+H_p(x,Du)Dz=\dfrac{\delta F}{\delta m}(x,m(t))(\rho(t))+h(t,x)\,,\\ z(T)=\dfrac{\delta G}{\delta m}(x,m(T))(\rho(T))+z_T\,,\\ a(x)Dz\cdot\nu_{|\partial\Omega}=0\,.\end{cases} \tag{3.5.22}$$

We note that, from Hypotheses 3.2.4, we have

$$\langle a(x)D_xG(x,m),\nu(x)\rangle_{|\partial\Omega}=0\quad\forall m\in\mathcal{P}(\Omega)\implies\left\langle a(x)D_x\frac{\delta G}{\delta m}(x,m(T))(\mu(T)),\nu(x)\right\rangle_{|\partial\Omega}=0\,.$$

Hence, compatibility conditions are satisfied for equation (3.5.22) and, from *Theorem 5.1.21* of [85], z satisfies

$$\begin{aligned}\|z\|_{1,2+\alpha}&\le C\left(\|z_T\|_{2+\alpha}+\sup_{t\in[0,T]}\|\rho(t)\|_{-(2+\alpha),N}+\|h\|_{0,\alpha}\right)\\&\le C\left(M+\sup_{t\in[0,T]}\|\rho(t)\|_{-(1+\alpha),N}\right)\,,\end{aligned} \tag{3.5.23}$$

where we also use hypothesis (vi) of 3.2.4, for the boundary condition of $\frac{\delta F}{\delta m}$.

Then we define $\mathbf{\Phi}(\rho) := \tilde{\rho}$, where $\tilde{\rho}$ is the solution in the sense of Definition 3.5.1 to :

$$\begin{cases} \tilde{\rho}_t - \mathrm{div}(a(x)D\tilde{\rho}) - \mathrm{div}(\tilde{\rho}(H_p(x,Du)+\tilde{b})) - \mathrm{div}(mH_{pp}(x,Du)Dz + c) = 0 \\ \tilde{\rho}(0) = \rho_0 \\ \Big(a(x)D\tilde{\rho} + \tilde{\rho}(H_p(x,Du)+\tilde{b}) + mH_{pp}(x,Du)Dz + c\Big)\cdot \nu_{|\partial\Omega} = 0 \end{cases} . \quad (3.5.24)$$

Thanks to Proposition 3.5.3, we have $\tilde{\rho} \in X$. We want to prove that the map $\mathbf{\Phi}$ is continuous and compact.

For the compactness, let $\{\rho_n\}_n \subset X$ be a subsequence with $\|\rho_n\|_X \le C$ for a certain $C > 0$. We consider for each n the solutions z_n and $\tilde{\rho}_n$ of (3.5.22) and (3.5.24) associated to $\tilde{\rho}_n$. Using (3.5.23), we have $\|z_n\|_{1,2+\alpha} \le C_1$, where C_1 depends on C. Then, thanks to Ascoli-Arzelà's Theorem, and using also (3.2.1), $\exists z$ s.t. $z_n \to z$ up to subsequences at least in $\mathcal{C}([0,T];\mathcal{C}^1(\Omega))$.

Using the pointwise convergence of Dz_n and the L^p boundedness of m stated in (3.5.16), we immediately obtain

$$mH_{pp}(x,Du)Dz_n + c \to mH_{pp}(x,Du)Dz + c \qquad \text{in } L^1(Q_T)\,,$$

which immediately implies

$$\mathrm{div}(mH_{pp}(x,Du)Dz_n + c) \to \mathrm{div}(mH_{pp}(x,Du)Dz + c) \qquad \text{in } L^1(W^{-1,\infty})\,.$$

Hence, stability results proved in Proposition 3.5.3 proves that $\tilde{\rho}_n \to \tilde{\rho}$ in X, where $\tilde{\rho}$ is the solution related to Dz. This proves the compactness result.

The continuity of Φ can be proved used the same computations of the compactness.

Finally, in order to apply Schaefer's theorem, we have to prove that

$$\exists M > 0 \text{ s.t. } \rho = \sigma\mathbf{\Phi}(\rho) \;\text{ and } \sigma \in [0,1] \implies \|\rho\|_X \le M\,.$$

We will prove in the next step that, if $\rho = \sigma\mathbf{\Phi}(\rho)$, then the couple (z,ρ) satisfies (3.5.20). This allows us to apply Schaefer's theorem and also gives us the desired estimate (3.5.20), since each solution (z,ρ) of the system satisfies $\rho = \sigma\mathbf{\Phi}(\rho)$ with $\sigma = 1$.

Step 2 : Estimate of ρ *and* z. Let $(\rho,\sigma) \in X \times [0,1]$ such that $\rho = \sigma\mathbf{\Phi}(\rho)$. Then the couple

(z, ρ) satisfies

$$\begin{cases} -z_t - \mathrm{tr}(a(x)D^2 z) + H_p(x, Du)Dz = \dfrac{\delta F}{\delta m}(x, m(t))(\rho(t)) + h(t,x) \\ \rho_t - \mathrm{div}(a(x)D\rho) - \mathrm{div}(\rho(H_p(x,Du) + \tilde{b})) - \sigma \mathrm{div}(mH_{pp}(x,Du)Dz + c) = 0 \\ z(T,x) = \dfrac{\delta G}{\delta m}(x, m(T))(\rho(T)) + z_T(x) \qquad\qquad \rho(0) = \sigma\rho_0 \\ a(x)Dz \cdot \nu_{|\partial\Omega} = 0 \qquad \Big(a(x)D\rho + \rho(H_p(x,Du) + \tilde{b}) + \sigma(mH_{pp}(x,Du)Dz + c)\Big) \cdot \nu_{|\partial\Omega} = 0 \end{cases} .$$

We want to use z as test function for the equation of ρ. This is allowed since z satisfies (3.5.4) with

$$\psi = \frac{\delta F}{\delta m}(x, m(t))(\rho(t)) + h(t,x) \in L^\infty(\Omega)\,, \qquad \xi = \frac{\delta G}{\delta m}(x, m(T))(\rho(T)) + z_T(x) \in \mathcal{C}^{1+\alpha,N}$$

We obtain from the weak formulation of ρ :

$$\int_\Omega (\rho(T,x)z(T,x) - \sigma\rho_0(x)z(0,x))\,dx = -\sigma\int_0^T\int_\Omega \langle c, Dz\rangle dxdt + \\ - \int_0^T\int_\Omega \rho(t,x)\left(\frac{\delta F}{\delta m}(x,m(t))(\rho(t)) + h\right)dxdt - \sigma\int_0^T\int_\Omega m\langle H_{pp}(x,Du)Dz, Dz\rangle\,dxdt\,.$$

Using the terminal condition of z and the monotonicity of F and G, we get a first estimate :

$$\begin{aligned} \sigma\int_0^T\int_\Omega m\langle H_{pp}(x,Du)Dz, Dz\rangle\,dxdt &\le \sup_{t\in[0,T]} \|\rho(t)\|_{-(2+\alpha),N}\,\|z_T\|_{2+\alpha} + \|\rho\|_{L^p}\,\|h\|_\infty \\ &\quad + \|z\|_{1,2+\alpha}\left(\|\rho_0\|_{-(2+\alpha),N} + \|c\|_{L^1}\right) \\ &\le M\left(\sup_{t\in[0,T]} \|\rho(t)\|_{-(1+\alpha),N} + \|\rho\|_{L^1} + \|z\|_{1,2+\alpha}\right). \end{aligned} \tag{3.5.25}$$

We already know an initial estimate on z in (3.5.23). Now we need to estimate ρ.

Using (3.5.6) we obtain

$$\sup_{t\in[0,T]} \|\rho\|_{-(1+\alpha),N} + \|\rho\|_{L^p} \le C\left(\|\sigma m H_{pp}(x,Du)Dz\|_{L^1} + \|c\|_{L^1} + \|\rho_0\|_{-(1+\alpha)}\right) \tag{3.5.26}$$

As regards the first term in the right hand side, we can use Hölder's inequality and

(3.5.25) to obtain

$$
\begin{aligned}
\|mH_{pp}(x,Du)Dz\|_{L^1} &= \sigma \sup_{\substack{\|\phi\|_\infty\le 1\\ \phi\in L^\infty(Q_T;\mathbb{R}^d)}} \int_0^T\int_\Omega m\langle H_{pp}(x,Du)Dz,\phi\rangle\,dxdt\\
&\le \sigma\left(\int_0^T\int_\Omega m\langle H_{pp}(x,Du)Dz,Dz\rangle\,dxdt\right)^{\frac12}\left(\int_0^T\int_\Omega m\langle H_{pp}(x,Du)\phi,\phi\rangle\,dxdt\right)^{\frac12}\\
&\le M^{\frac12}\left(\sup_{t\in[0,T]}\|\rho(t)\|_{-(1+\alpha),N}^{\frac12}+\|\rho\|_{L^1}^{\frac12}+\|z\|_{1,2+\alpha}^{\frac12}\right)
\end{aligned}
$$

Putting these estimates into (3.5.26) we obtain

$$
\sup_{t\in[0,T]}\|\rho\|_{-(1+\alpha),N}+\|\rho\|_{L^p}\le C\left(M+M^{\frac12}\left(\sup_{t\in[0,T]}\|\rho(t)\|_{-(1+\alpha),N}^{\frac12}+\|\rho\|_{L^1}^{\frac12}+\|z\|_{1,2+\alpha}^{\frac12}\right)\right).
$$

Using a generalized Young's inequality with suitable coefficients, we get

$$
\sup_{t\in[0,T]}\|\rho\|_{-(1+\alpha),N}+\|\rho\|_{L^p}\le C\left(M+M^{\frac12}\,\|z\|_{1,2+\alpha}^{\frac12}\right). \tag{3.5.27}
$$

This gives us an initial estimate for ρ, depending on the estimate of z.

Coming back to (3.5.23), (3.5.27) implies

$$
\|z\|_{1,2+\alpha}\le C\left(M+M^{\frac12}\,\|z\|_{1,2+\alpha}^{\frac12}\right).
$$

Using a generalized Young's inequality with suitable coefficients, this implies

$$
\|z\|_{1,2+\alpha}\le Cm\,.
$$

Plugging this estimate in (3.5.27), we finally obtain

$$
\|z\|_{1,2+\alpha}+\sup_{t\in[0,T]}\|\rho\|_{-(1+\alpha),N}+\|\rho\|_{L^p}\le CM\,.
$$

This concludes the existence result.

Step 3. Uniqueness. Let (z_1,ρ_1) and (z_2,ρ_2) be two solutions of (3.5.18). Then the couple $(z,\rho):=(z_1-z_2,\rho_1-\rho_2)$ satisfies the following linear system :

$$
\begin{cases}
-z_t-\operatorname{tr}(a(x)D^2z)+H_p(x,Du)Dz=\dfrac{\delta F}{\delta m}(x,m(t))(\rho(t))=0\,,\\
\rho_t-\operatorname{div}(a(x)D\rho)-\operatorname{div}(\rho(H_p(x,Du)+\tilde b))-\operatorname{div}(mH_{pp}(x,Du)Dz)=0\,,\\
z(T,x)=\dfrac{\delta G}{\delta m}(x,m(T))(\rho(T))\,,\qquad \rho(t_0)=0\,,\\
a(x)Dz\cdot\nu_{|\partial\Omega}=0\,,\quad \Big(a(x)D\rho+\rho(H_p(x,Du)+\tilde b)+mH_{pp}(x,Du)Dz\Big)\cdot\nu_{|\partial\Omega}=0\,,
\end{cases}
$$

i.e., a system of the form (3.5.18) with $h = c = z_T = \rho_0 = 0$. Then estimation (3.5.20) tells us that

$$\|z\|_{1,2+\alpha} + \sup_{t\in[0,T]} \|\rho\|_{-(1+\alpha),N} + \|\rho\|_{L^p} \leq 0\,.$$

and so $z = 0$, $\rho = 0$. This concludes the Proposition. □

We are ready to prove that (3.5.1) has a fundamental solution. This solution will be the desired derivative $\frac{\delta U}{\delta m}$.

Proposition 3.5.9. *Equation* (3.5.1) *has a fundamental solution, i.e. there exists a function* $K : [0, T] \times \Omega \times \mathcal{P}(\Omega) \times \Omega \to \mathbb{R}$ *such that, for any* (t_0, m_0, μ_0) *we have*

$$v(t_0, x) = {}_{-(1+\alpha)}\langle \mu_0, K(t_0, x, m_0, \cdot)\rangle_{1+\alpha} \tag{3.5.28}$$

Moreover, $K(t_0, \cdot, m_0, \cdot) \in \mathcal{C}^{2+\alpha}(\Omega) \times \mathcal{C}^{1+\alpha}(\Omega)$ *with*

$$\sup_{(t,m)\in[0,T]\times\mathcal{P}(\Omega)} \|K(t, \cdot, m, \cdot)\|_{2+\alpha,1+\alpha} \leq C\,, \tag{3.5.29}$$

and the second derivatives w.r.t. x *and the first derivatives w.r.t.* y *are continuous in all variables.*

Démonstration. From now on, we indicate with $v(t, x; \mu_0)$ the solution of the first equation of (3.5.1) related to μ_0. We start considering, for $y \in \Omega$, $\mu_0 = \delta_y$, the Dirac function at y. We define

$$K(t_0, x, m_0, y) = v(t_0, x; \delta_y)$$

Thanks to (3.5.20), one immediately knows that K is twice differentiable w.r.t. x and

$$\|K(t_0, \cdot, m_0, y)\|_{2+\alpha} \leq C\, \|\delta_y\|_{-(1+\alpha)} = C$$

Moreover, we can use the linearity of the system (3.5.18) to obtain

$$\frac{K(t_0, x, m_0, y + he_j) - K(t_0, x, m_0, y)}{h} = v(t_0, x; \Delta_{h,j}\delta_y)\,,$$

where $\Delta_{h,j}\delta_y = \frac{1}{h}(\delta_{y+he_j} - \delta_y)$. Using stability results for (3.5.1), proved previously, we can pass to the limit and find that

$$\frac{\partial K}{\partial y_j}(t_0, x, m_0, y) = v(t_0, x; -\partial_{y_j}\delta_y)\,,$$

where the derivative of the Dirac delta function is in the sense of distribution. Since $\partial_{y_i}\delta_y$ is bounded in $\mathcal{C}^{-(1+\alpha)}$ for all i, j, from (3.5.20) we deduce that the second derivatives of K with respect to x are well defined and bounded.

The representation formula (3.5.28) is an immediate consequence of the linear character of the equation and of the density of the set generated by the Dirac functions. This concludes the proof. □

Now we are ready to prove the differentiability of the function U with respect to the measure m.

In particular, we want to prove that this fundamental solution K is actually the derivative of U with respect to the measure.

Theorem 3.5.10. *Let* (u_1, m_1) *and* (u_2, m_2) *be two solutions of the Mean Field Games system* (3.1.2)*, associated with the starting initial conditions* (t_0, m_0^1) *and* (t_0, m_0^2)*. Let* (v, μ) *be the solution of the linearized system* (3.5.1) *related to* (u_2, m_2)*, with initial condition* $(t_0, m_0^1 - m_0^2)$*. Then we have*

$$\|u_1 - u_2 - v\|_{1,2+\alpha} + \sup_{t\in[0,T]} \|m_1(t) - m_2(t) - \mu(t)\|_{-(1+\alpha),N} \le C\mathbf{d}_1(m_0^1, m_0^2)^2 \,, \qquad (3.5.30)$$

Consequently, the function U *defined in* (3.1.3) *is differentiable with respect to* m*.*

Démonstration. We call $(z, \rho) = (u_1 - u_2 - v, m_1 - m_2 - \mu)$. Then (z, ρ) satisfies

$$\begin{cases} -z_t - \operatorname{tr}(a(x)D^2 z) + H_p(x, Du_2)Dz = \dfrac{\delta F}{\delta m}(x, m_2(t))(\rho(t)) + h(t,x)\,, \\ \rho_t - \operatorname{div}(a(x)D\rho) - \operatorname{div}(\rho(H_p(x, Du_2) + \tilde{b})) - \operatorname{div}(mH_{pp}(x, Du_2)Dz + c) = 0\,, \\ z(T,x) = \dfrac{\delta G}{\delta m}(x, m_2(T))(\rho(T)) + z_T(x)\,, \qquad \rho(t_0) = 0,, \\ a(x)Dz\cdot\nu_{|\partial\Omega} = 0\,, \quad \Big(a(x)D\rho + \rho(H_p(x, Du) + \tilde{b}) + mH_{pp}(x, Du)Dz + c\Big)\cdot\nu_{|\partial\Omega} = 0\,, \end{cases}$$

$$h(t,x) = h_1(t,x) + h_2(t,x)\,,$$
$$h_1 = -\int_0^1 (H_p(x, sDu_1 + (1-s)Du_2) - H_p(x, Du_2)) \cdot D(u_1 - u_2)\, ds\,,$$
$$h_2 = \int_0^1 \int_\Omega \left(\frac{\delta F}{\delta m}(x, sm_1(t) + (1-s)m_2(t), y) - \frac{\delta F}{\delta m}(x, m_2(t), y) \right) (m_1(t) - m_2(t))(dy)ds,$$
$$c(t) = c_1(t) + c_2(t)\,,$$
$$c_1(t) = (m_1(t) - m_2(t)) H_{pp}(x, Du_2)(Du_1 - Du_2)\,,$$
$$c_2(t) = m_1 \int_0^1 (H_{pp}(x, sDu_1 + (1-s)Du_2) - H_{pp}(x, Du_2))\,(Du_1 - Du_2)\, ds\,,$$
$$z_T = \int_0^1 \int_\Omega \left(\frac{\delta G}{\delta m}(x, sm_1(T) + (1-s)m_2(T), y) - \frac{\delta G}{\delta m}(x, m_2(T), y) \right) (m_1(T) - m_2(T))(dy)ds\,.$$

So, (3.5.20) implies that

$$\|u_1 - u_2 - v\|_{1,2+\alpha} + \sup_{t\in[0,T]} \|m_1(t) - m_2(t) - \mu(t)\|_{-(1+\alpha),N} \le C \left(\|h\|_{0,\alpha} + \|c\|_{L^1} + \|z_T\|_{2+\alpha} \right). \tag{3.5.31}$$

Now we bound the right-hand side term in order to obtain (3.5.30).

We start with the term $h = h_1 + h_2$. We can write

$$h_1 = -\int_0^1 \int_0^1 s \left\langle H_{pp}(x, rsDu_1 + (1-rs)Du_2)\,(Du_1 - Du_2)\,,\,(Du_1 - Du_2) \right\rangle drds\,.$$

Using the properties of Hölder norm and (3.4.1), it is immediate to obtain

$$\|h_1\|_{0,\alpha} \le C\, \|D(u_1 - u_2)\|_{0,\alpha}^2 \le C\mathbf{d}_1(m_0^1, m_0^2)^2\,.$$

As regards the h_2 term, we can immediately bound the quantity

$$|h_2(t,x) - h_2(t,y)|$$

by

$$|x-y|^\alpha \mathbf{d}_1(m_1(t), m_2(t)) \int_0^1 \|D_m F(\cdot, sm_1(t) + (1-s)m_2(t), \cdot) - D_m F(\cdot, m_2(t), \cdot)\|_{\alpha,\infty}\, ds\,.$$

Using the regularity of F and (3.4.1), we get

$$\|h_2\|_{0,\alpha} = \sup_{t\in[0,T]} \|h_2(t,\cdot)\|_\alpha \le C\mathbf{d}_1(m_0^1, m_0^2)^2\,.$$

A similar estimate holds for the function z_T. As regards the function c, we have

$$\begin{aligned}\|c_1\|_{L^1} &= \int_0^T \int_\Omega H_{pp}(x, Du_2)(Du_1 - Du_2)(m_1(t,dx) - m_2(t,dx))\, dt \\ &\le C\, \|u_1 - u_2\|_{1,2+\alpha}\, \mathbf{d}_1(m_1(t), m_2(t)) \le C\mathbf{d}_1(m_0^1, m_0^2)^2\,,\end{aligned}$$

and, using the notation $u_{1+s} := sDu_1 + (1-s)Du_2$,

$$\begin{aligned}\|c_2\|_{L^1} &= \int_0^1 \int_0^T \int_\Omega \left(H_{pp}(x, Du_{1+s}) - H_{pp}(x, Du_2)\right)(Du_1 - Du_2) m_1(t,dx)\, dtds \\ &\le C\, \|Du_1 - Du_2\|_\infty^2 \le C\mathbf{d}_1(m_0^1, m_0^2)^2\,.\end{aligned}$$

Substituting these estimates in (3.5.31), we obtain (3.5.30) and we conclude the proof. □

Since

$$v(t_0, x) = \int_\Omega K(t_0, x, m_{02}, y)(m_{01}(dy) - m_{02}(dy))\,,$$

equation (3.5.30) implies

$$\left\| U(t_0, \cdot, m_{01}) - U(t_0, \cdot, m_{02}) - \int_\Omega K(t_0, \cdot, m_{02}, y)(m_{01} - m_{02})(dy) \right\|_\infty \le C\mathbf{d}_1(m_{01}, m_{02})^2.$$

As a straightforward consequence, we have that U is differentiable with respect to m and

$$\frac{\delta U}{\delta m}(t, x, m, y) = K(t, x, m, y)\,.$$

Consequently, using (3.5.29) we obtain

$$\sup_t \left\| \frac{\delta U}{\delta m}(t, \cdot, m, \cdot) \right\|_{2+\alpha, 1+\alpha} \le C\,. \tag{3.5.32}$$

But, in order to make sense to equation (3.1.1), we need at least that $\frac{\delta U}{\delta m}$ is almost everywhere twice differentiable with respect to y.

To do that, we need to improve the estimates (3.5.20) for a couple (v, μ) solution of (3.5.1).

Proposition 3.5.11. *Let $\mu_0 \in \mathcal{C}^{-(1+\alpha)}$. Then the unique solution (v, μ) satisfies*

$$\|v\|_{1,2+\alpha} + \sup_{t \in [0,T]} \|\mu(t)\|_{-(2+\alpha), N} \le C\, \|\mu_0\|_{-(2+\alpha)}\,. \tag{3.5.33}$$

Démonstration. We consider the solution (v, μ) obtained in Proposition 3.5.8. Since μ satisfies $\mu = \sigma\Phi(\mu)$ with $\sigma = 1$, we can use (3.5.23) with $z_T = h = 0$ and obtain

$$\|v\|_{1,2+\alpha} \le C \sup_{t\in[0,T]} \|\mu(t)\|_{-(2+\alpha),N} . \tag{3.5.34}$$

We want to estimate the right-hand side. Using (3.5.15) we have

$$\sup_{t\in[0,T]} \|\mu(t)\|_{-(2+\alpha),N} \le C \left(\|\mu_0\|_{-(2+\alpha)} + \|mH_{pp}(x, Du)Dv\|_{L^1} \right) . \tag{3.5.35}$$

The last term is estimated, as in Proposition 3.5.8, by

$$\|\sigma m H_{pp}(x, Du)Dv\|_{L^1} \le C \left(\int_0^T \int_\Omega m\langle H_{pp}(x, Du)Dv, Dv\rangle \, dxdt \right)^{\frac{1}{2}} . \tag{3.5.36}$$

The right-hand side term can be bounded using (3.5.25) with $h = z_T = c = 0$:

$$\int_0^T \int_\Omega m\langle H_{pp}(x, Du)Dv, Dv\rangle \, dxdt \le \|v\|_{1,2+\alpha} \|\mu_0\|_{-(2+\alpha)} . \tag{3.5.37}$$

Hence, plugging estimates (3.5.36) and (3.5.37) into (3.5.35) we obtain

$$\sup_{t\in[0,T]} \|\mu(t)\|_{-(2+\alpha),N} \le C \left(\|\mu_0\|_{-(2+\alpha)} + \|v\|^{\frac{1}{2}}_{1,2+\alpha} \|\mu_0\|^{\frac{1}{2}}_{-(2+\alpha),N} \right) . \tag{3.5.38}$$

Coming back to (3.5.34) and using a generalized Young's inequality, we get

$$\|v\|_{1,2+\alpha} \le C \|\mu_0\|_{-(2+\alpha)} ,$$

and finally, substituting the last estimate into (3.5.38), we obtain (3.5.33) and we conclude. □

As an immediate Corollary, we get the desired estimate for $\frac{\delta U}{\delta m}$.

Corollary 3.5.12. *Suppose hypotheses 3.2.4 satisfied. Then the derivative $\frac{\delta U}{\delta m}$ is twice differentiable with respect to y, together with its first and second derivatives with respect to x, and the following estimate hold :*

$$\left\| \frac{\delta U}{\delta m}(t, \cdot, m, \cdot) \right\|_{2+\alpha,2+\alpha} \le C . \tag{3.5.39}$$

Démonstration. We want to prove that, $\forall\, i, j$, the incremental ratio

$$R^h_{i,j}(x,y) := \frac{\partial_{y_i}\frac{\delta U}{\delta m}(t_0,x,m_0,y+he_j) - \partial_{y_i}\frac{\delta U}{\delta m}(t_0,x,m_0,y)}{h} \tag{3.5.40}$$

is a Cauchy sequence for $h \to 0$ together with its first and second derivatives with respect to x. Then we have to estimate, for $h,k>0$, the quantity $\left|D^l_x R^h_{i,j}(x,y) - D^l_x R^k_{i,j}(x,y)\right|$, for $|l| \leq 2$.

We already know that

$$\partial_{y_i}\frac{\delta U}{\delta m}(t_0,x,m_0,y) = v(t_0,x;-\partial_{y_i}\delta_y)\,.$$

Using the linearity of the system (3.5.1), we obtain that

$$\left|D^l_x R^h_{i,j}(x,y) - D^l_x R^k_{i,j}(x,y)\right| = D^l_x v\left(t_0,x;\Delta^j_h(-\partial_{y_i}\delta_y) - \Delta^j_k(-\partial_{y_i}\delta_y)\right)\,,$$

where $\Delta^j_h(-\partial_{y_i}\delta_y) = -\frac{1}{h}(\partial_{y_i}\delta_{y+he_j} - \partial_{y_i}\delta_y)\,.$

Hence, estimate (3.5.33) and Lagrange's Theorem implies

$$\begin{aligned}
&\left|D^l_x R^h_{i,j}(x,y) - D^l_x R^k_{i,j}(x,y)\right| \leq C\left\|\Delta^j_h(-\partial_{y_i}\delta_y) - \Delta^j_k(-\partial_{y_i}\delta_y)\right\|_{-(2+\alpha)} \\
&= \sup_{\|\phi\|_{2+\alpha}\leq 1}\left(\frac{\partial_{y_i}\phi(y+he_j) - \partial_{y_i}\phi(y)}{h} - \frac{\partial_{y_i}\phi(y+ke_j) - \partial_{y_i}\phi(y)}{k}\right) \\
&= \sup_{\|\phi\|_{2+\alpha}\leq 1}\left(\partial^2_{y_iy_j}\phi(y_{\phi,h}) - \partial^2_{y_iy_j}\phi(y_{\phi,k})\right) \leq \sup_{\|\phi\|_{2+\alpha}\leq 1}|y_{\phi,h} - y_{\phi,k}|^\alpha \leq |h|^\alpha + |k|^\alpha\,,
\end{aligned}$$

for a certain $y_{\phi,h}$ in the line segment between y and $y+he_j$ and $y_{\phi,k}$ in the line segment between y and $y+ke_j$.

Since the last term goes to 0 when $h,k\to 0$, we have proved that the incremental ratio (3.5.40) and its first and second derivative w.r.t x are Cauchy sequences in h, and so converging when $h\to 0$. This proves that $D^l_x\frac{\delta U}{\delta m}$ is twice differentiable with respect to y, for all $0\leq |l|\leq 2\,.$

In order to show the Hölder bound for $\frac{\delta U}{\delta m}$ w.r.t. y, we consider $y, y' \in \Omega$ and we consider the function

$$R^h_{i,j}(x,y) - R^h_{i,j}(x,y')\,.$$

Then we know from the linearity of (3.5.1)

$$R^h_{i,j}(x,y) - R^h_{i,j}(x,y') = v(t_0,x;\Delta^j_h(-\partial_{y_i}\delta_y) - \Delta^j_h(-\partial_{y_i}\delta_{y'}))\,,$$

and so, using (3.5.33) and

$$\left\| R^h_{i,j}(\cdot,y) - R^h_{i,j}(\cdot,y') \right\|_{2+\alpha} \le C \left\| \Delta^j_h(-\partial_{y_i}\delta_y) - \Delta^j_h(-\partial_{y_i}\delta_{y'}) \right\|_{-(2+\alpha)} .$$

Now we pass to the limit when $h \to 0$. It is immediate to prove that

$$\Delta^j_h(-\partial_{y_i}\delta_y) - \Delta^j_h(-\partial_{y_i}\delta_{y'}) \overset{h\to 0}{\longrightarrow} \partial_{y_j}\partial_{y_i}\delta_y - \partial_{y_j}\partial_{y_i}\delta_{y'} \qquad \text{in } \mathcal{C}^{-(2+\alpha)} .$$

Since $D^l_x R^h_{i,j}(x,y) \to \partial^2_{y_i y_j} D^l_x \frac{\delta U}{\delta m}(x,y)$ for all $|l| \le 2$, we can use Ascoli-Arzelà to obtain that

$$\left\| \partial^2_{y_i y_j}\frac{\delta U}{\delta m}(t,\cdot,m,y) - \partial^2_{y_i y_j}\frac{\delta U}{\delta m}(t,\cdot,m,y') \right\|_{2+\alpha} \le C \left\| \partial_{y_j}\partial_{y_i}\delta_y - \partial_{y_j}\partial_{y_i}\delta_{y'} \right\|_{-(2+\alpha)} \le C|y-y'|^\alpha ,$$

which proves (3.5.39) and concludes the proof. □

We conclude this part with a last property on the derivative $D_m U$, which will be essential in order to prove the uniqueness of solutions for the Master Equation and the convergence problem.

Corollary 3.5.13. *The function U satisfies the following Neumann boundary conditions :*

$$\begin{aligned} a(x)D_x\frac{\delta U}{\delta m}(t,x,m,y)\cdot\nu(x) &= 0 , \qquad \forall x\in\partial\Omega, y\in\Omega, t\in[0,T], m\in\mathcal{P}(\Omega) ,\\ a(y)D_m U(t,x,m,y) \quad \cdot\nu(y) &= 0 , \qquad \forall x\in\Omega, y\in\partial\Omega, t\in[0,T], m\in\mathcal{P}(\Omega) . \end{aligned}$$

Démonstration. Since $\frac{\delta U}{\delta m}(t_0,x,m_0,y) = v(t_0,x)$, where (v,μ) is the solution of (3.5.1) with $\mu_0 = \delta_y$, the first condition is immediate because of the Neumann condition of (3.5.1).

For the second condition, we consider $y \in \partial\Omega$ and we take

$$\mu_0 = -\partial_w(\delta_y) , \qquad \text{with } w = a(y)\nu(y) .$$

We want to prove that $(v,\mu) = (0,\mu)$ is a solution of (3.5.1) with $\mu_0 = -\partial_w\delta_y$, where μ is the unique solution in the sense of Definition 3.5.1 of

$$\begin{cases} \mu_t - \operatorname{div}(a(x)D\mu) - \operatorname{div}(\mu(H_p(x,Du)+\tilde b)) = 0 ,\\ \mu(t_0) = \mu_0 ,\\ \left(a(x)D\mu + \mu(H_p(x,Du)+\tilde b)\right)\cdot\nu_{|\partial\Omega} = 0 . \end{cases}$$

We only have to check that, if μ is a solution of this equation, then $v = 0$ solves

$$\begin{cases} -v_t - \mathrm{tr}(a(x)D^2v) + H_p(x, Du) \cdot Dv = \dfrac{\delta F}{\delta m}(x, m(t))(\mu(t))\,, \\ v(T, x) = \dfrac{\delta G}{\delta m}(x, m(T))(\mu(T))\,, \\ a(x)Dv \cdot \nu_{|\partial\Omega} = 0\,, \end{cases} \tag{3.5.41}$$

which reduces to prove that

$$\frac{\delta F}{\delta m}(x, m(t))(\mu(t)) = \frac{\delta G}{\delta m}(x, m(T))(\mu(T)) = 0\,.$$

We will give a direct proof.

Choosing a test function $\phi(t, y)$ satisfying (3.5.4), with $\psi(t, y) = 0$ and $\xi(y) = \frac{\delta F}{\delta m}(x, m(t), y)$, we have from boundary conditions of $\frac{\delta F}{\delta m}$ that ϕ is a $\mathcal{C}^{\frac{1+\alpha}{2}, 1+\alpha}$ function satisfying Neumann boundary conditions.

It follows from the weak formulation of μ that

$$\frac{\delta F}{\delta m}(x, m(t))(\mu(t)) = \langle \mu(t), \frac{\delta F}{\delta m}(x, m(t), \cdot)\rangle = \langle \mu_0, \phi(0, \cdot)\rangle = 0\,,$$

since $aD\phi \cdot \nu_{|\partial\Omega} = 0$ and

$$\langle \mu_0, \phi(0, \cdot)\rangle = \langle -\partial_w \delta_y, \phi(0, \cdot)\rangle = a(y)D\phi(0, y) \cdot \nu(y) = 0\,.$$

Same computations hold for $\frac{\delta G}{\delta m}$, proving that $v = 0$ satisfies (3.5.41).

Then we can easily conclude :

$$\begin{aligned} a(y)D_mU(t_0, x, m_0, y) \cdot \nu(y) &= D_y \frac{\delta U}{\delta m}(t_0, x, m_0, y) \cdot w \\ &= \left\langle \frac{\delta U}{\delta m}(t_0, x, m_0, \cdot), \mu_0 \right\rangle = v(t_0, x) = 0\,. \end{aligned}$$

□

3.6 Solvability of the first-order Master Equation

The $\mathcal{C}^1$ character of U with respect to m is crucial in order to prove the main theorem of this chapter.

Theorem 3.6.1. *Suppose hypotheses 3.2.4 are satisfied. Then there exists a unique classical solution U of the Master Equation* (3.1.1)*.*

Démonstration. We start from the existence part.

Existence. We start assuming that m_0 is a smooth and positive function satisfying (3.3.9), and we consider (u, m) the solution of MFG system starting from m_0 at time t_0. Then

$$\partial_t U(t_0, x, m_0)$$

can be computed as the sum of the two limits :

$$\lim_{h\to 0} \frac{U(t_0 + h, x, m_0) - U(t_0 + h, x, m(t_0 + h))}{h}$$

and

$$\lim_{h\to 0} \frac{U(t_0 + h, x, m(t_0 + h)) - U(t_0, x, m_0)}{h}.$$

The second limit, using the very definition of U, is equal to

$$\lim_{h\to 0} \frac{u(t_0 + h, x) - u(t_0, x)}{h} = u_t(t_0, x) = -\mathrm{tr}(a(x)D^2u(t_0, x)) + H(x, Du(t_0, x))$$
$$-F(x, m(t_0)) = -\mathrm{tr}(a(x)D_x^2U(t_0, x, m_0)) + H(x, D_xU(t_0, x, m_0)) - F(x, m_0)\,.$$

As regards the first limit, defining $m_s := (1-s)m(t_0)+sm(t_0+h)$ and using the $\mathcal{C}^1$ regularity of U with respect to m, we can write it as

$$-\lim_{h\to 0} \int_0^1 \int_\Omega \frac{\delta U}{\delta m}(t_0 + h, x, m_s, y)\frac{(m(t_0 + h, y) - m(t_0, y))}{h}\, dyds$$
$$= -\int_0^1 \int_\Omega \frac{\delta U}{\delta m}(t_0, x, m_0, y)m_t(t_0, y)\, dyds = \int_\Omega \frac{\delta U}{\delta m}(t_0, x, m_0, y)m_t(t_0, y)\, dy$$
$$= -\int_\Omega \frac{\delta U}{\delta m}(t_0, x, m_0, y)\,\mathrm{div}\Big(a(y)Dm(t_0, y) + m(t_0, y)(\tilde{b} + H_p(y, Du(t_0, y)))\Big)\, dy\,.$$

Taking into account the representation formula (3.5.2) for $\frac{\delta U}{\delta m}$, we integrate by parts and use the boundary condition of $\frac{\delta U}{\delta m}$ and m to obtain

$$\int_\Omega [H_p(y, D_xU(t_0, y, m_0))D_mU(t_0, x, m_0, y) - \mathrm{tr}\,(a(y)D_yD_mU(t_0, x, m_0, y))]\, dm_0(y)$$

So with the computation of the two limits we obtain

$$\partial_t U(t, x, m) = -\mathrm{tr}\left(a(x)D_x^2U(t, x, m)\right) + H\left(x, D_xU(t, x, m)\right)$$
$$-\int_\Omega \mathrm{tr}\,(a(y)D_yD_mU(t, x, m, y))\, dm(y)+$$
$$\int_\Omega D_mU(t, x, m, y) \cdot H_p(y, D_xU(t, y, m))dm(y) - F(x, m)\,.$$

So the equation is satisfied for all $m_0 \in \mathcal{C}^\infty$ satisfying (3.3.9), and so, with a density argument, for all $m_0 \in \mathcal{P}(\Omega)$.

The boundary conditions are easily verified thanks to Corollary 3.5.13. This concludes the existence part.

Uniqueness. Let V be another solution of the Master Equation (3.1.1) with Neumann boundary conditions. We consider, for fixed t_0 and m_0, with m_0 smooth satisfying (3.3.9), the solution $\tilde{m}$ of the Fokker-Planck equation :

$$\begin{cases} \tilde{m}_t - \operatorname{div}(a(x)D\tilde{m}) - \operatorname{div}\Big(\tilde{m}\,\Big(H_p(x, D_xV(t,x,\tilde{m})) + \tilde{b}\Big)\Big) = 0\,, \\ \tilde{m}(t_0) = m_0\,, \\ \Big[a(x)D\tilde{m} + (\tilde{b} + D_xV(t,x,\tilde{m}))\Big]\cdot \nu(x)_{|\partial\Omega} = 0\,. \end{cases}$$

This solution is well defined since D_xV is Lipschitz continuous with respect to the measure variable.

Then we define $\tilde{u}(t,x) = V(t,x,\tilde{m}(t))$. Using the equations of V and $\tilde{m}$, we obtain

$$\begin{aligned} \tilde{u}_t(t,x) =& V_t(t,x,\tilde{m}(t)) + \int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\,\tilde{m}_t(t,y)\,dy \\ =& V_t(t,x,\tilde{m}(t)) + \int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\operatorname{div}(a(y)D\tilde{m}(t,y))\;dy \\ &+ \int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\operatorname{div}\Big(\tilde{m}\,\Big(H_p(x,D_xV(t,x,\tilde{m})) + \tilde{b}\Big)\Big)\;dy\,. \end{aligned}$$

We compute the two integrals by parts. As regards the first, we have

$$\begin{aligned} &\int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\operatorname{div}(a(y)D\tilde{m}(t,y))\;dy \\ =& -\int_\Omega a(y)D\tilde{m}(t,y)\,D_mV(t,x,\tilde{m}(t),y)\,dy + \int_{\partial\Omega}\frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\,a(y)D\tilde{m}(t,y)\cdot\nu(t,y)\,dy \\ =& \int_\Omega \operatorname{div}(a(y)D_yD_mV(t,x,\tilde{m}(t),y))\,\tilde{m}(t,y)\,dy - \int_\Omega a(y)D_mV(t,x,\tilde{m}(t),y)\cdot\nu(y)\tilde{m}(t,y)dy \\ &+ \int_{\partial\Omega}\frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\,a(y)D\tilde{m}(t,y)\cdot\nu(t,y)\,dy\,, \end{aligned}$$

while for the second

$$\begin{aligned} &\int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\operatorname{div}\Big(\tilde{m}\,\Big(H_p(x,D_xV(t,x,\tilde{m})) + \tilde{b}\Big)\Big)\;dy \\ =& -\int_\Omega \Big(H_p(x,D_xV(t,x,\tilde{m})) + \tilde{b}\Big)\,D_mV(t,x,\tilde{m},y)\tilde{m}(t,y)dy \\ &+ \int_{\partial\Omega}\frac{\delta V}{\delta m}(t,x,\tilde{m}(t),y)\,\Big(H_p(x,D_xV(t,x,\tilde{m})) + \tilde{b}\Big)\cdot\nu(y)\,\tilde{m}(t,y)dy\,. \end{aligned}$$

Putting together these estimates and taking into account the boundary conditions on V and m :

$$\Big[a(x)D\tilde{m} + (\tilde{b} + D_x V(t,x,\tilde{m}))\Big] \cdot \nu(x)_{|x\in\partial\Omega} = 0\,, \qquad a(y)D_m V(t,x,m,y)\cdot\nu(y)_{|y\in\partial\Omega} = 0\,,$$

and the relation between the divergence and the trace term

$$\operatorname{div}(a(x)D\phi(x)) = \operatorname{tr}(a(x)D^2\phi(x)) + \tilde{b}(x)D\phi(x)\,, \qquad \forall\phi\in W^{2,\infty}(\Omega)\,,$$

we find

$$\begin{aligned}
\tilde{u}_t(t,x) &= V_t(t,x,\tilde{m}(t)) + \int_\Omega \operatorname{tr}(a(y)D_yD_mV(t,x,\tilde{m},y))\,d\tilde{m}(y) \\
&\quad - \int_\Omega -H_p(y,D_xV(t,y,\tilde{m}))D_mV(t,x,\tilde{m},y)\,d\tilde{m}(y) \\
&= -\operatorname{tr}(a(x)D_x^2V(t,x,\tilde{m}(t))) + H(x,D_xV(t,x,\tilde{m}(t))) - F(x,\tilde{m}(t)) \\
&= -\operatorname{tr}(a(x)D^2\tilde{u}(t,x)) + H(x,D\tilde{u}(t,x)) - F(x,\tilde{m}(t))\,.
\end{aligned}$$

This means that $(\tilde{u},\tilde{m})$ is a solution of the MFG system (3.1.2). Since the solution of the Mean Field Games system is unique, we get $(\tilde{u},\tilde{m}) = (u,m)$ and so $V(t_0,x,m_0) = U(t_0,x,m_0)$ whenever m_0 is smooth.
Then, using a density argument, the uniqueness is proved. □

3.7 The convergence problem

In this section we analyze the convergence of the Nash system of N players towards a solution of the Master Equation.

We consider, for an integer $N \geq 2$, a classical solution v_i^N of the *Nash System* :

$$\begin{cases}
-\partial_t v_i^N - \sum\limits_j \operatorname{tr}(a(x_j)D^2_{x_jx_j}v_i^N) + H(x_i,D_{x_i}v_i^N) + \sum\limits_{j\neq i} H_p(x_j,D_{x_j}v_j^N)\cdot D_{x_j}v_i^N = F(t,x_i,m_{\boldsymbol{x}}^{N,i})\,, \\
v_i^N(T,\boldsymbol{x}) = G(x_i,m_{\boldsymbol{x}}^{N,i})\,, \\
a(x_j)D_{x_j}v_i^N\cdot\nu(x_j)_{|x_j\in\partial\Omega} = 0\,, \qquad j=1,\cdots,N\,.
\end{cases} \tag{3.7.1}$$

where $1\leq i\leq N$ and $m_{\boldsymbol{x}}^{N,i}$ is defined in (3.1.5).

If hypotheses 3.2.4 are satisfied, then there exists a unique solution of the Master Equation (3.1.1).

In order to prove the convergence of v_i^N towards U, the main idea is to work with suitable *finite dimensional projections* of U, proving that these projections are nearly solutions to the Nash-system.

So, for $N \geq 2$ and $1 \leq i \leq N$, we define the following functions u_i^N :

$$u_i^N(t, \boldsymbol{x}) = U(t, x_i, m_{\boldsymbol{x}}^{N,i}) . \tag{3.7.2}$$

Thanks to the regularity of U, we already know that

$$u_i^N \in \mathcal{C}^{1+\frac{\alpha}{2}, 2+\alpha} \qquad \text{with respect to the couple } (t, x_i) . \tag{3.7.3}$$

In order to prove a regularity result for u_i^N with respect to the other variables $(x_j)_{j \neq i}$, we need to prove two technical results about the solution U of the Master Equation (3.1.1).

The first theorem is the following :

Theorem 3.7.1. *Suppose hypotheses 3.2.4 are satisfied. Then the derivative of the Master Equation* $\frac{\delta U}{\delta m}$ *is Lipschitz continuous with respect to the measure :*

$$\sup_{t \in [0,T]} \sup_{m_1 \neq m_2} (\mathbf{d}_1(m_1, m_2))^{-1} \left\| \frac{\delta U}{\delta m}(t, \cdot, m_1, \cdot) - \frac{\delta U}{\delta m}(t, \cdot, m_2, \cdot) \right\|_{2+\alpha, 1+\alpha} \leq C . \tag{3.7.4}$$

Démonstration. We consider, for $i = 1, 2$, the solution (v_i, μ_i) of the linearized system (3.5.1) related to (u_i, m_i).

To avoid too heavy notations, we take $t_0 = 0$ and we define

$$H_i'(t,x) := H_p(x, Du_i(t,x)) , \qquad H_i''(t,x) = H_{pp}(x, Du_i(t,x)) ,$$
$$F'(x, m, \mu) = \int_\Omega \frac{\delta F}{\delta m}(x, m, y)\mu(dy) , \qquad G'(x, m, \mu) = \int_\Omega \frac{\delta G}{\delta m}(x, m, y)\mu(dy) .$$

Then the couple $(z, \rho) := (v_1 - v_2, \mu_1 - \mu_2)$ satisfies the following linear system :

$$\begin{cases} -z_t - \operatorname{tr}(a(x)D^2 z) + H_1' \cdot Dz = F'(x, m_1(t), \rho(t)) + h , \\ \rho_t - \operatorname{div}(a(x)D\rho) - \operatorname{div}(\rho(H_1' + \tilde{b})) - \operatorname{div}(m_1 H_1'' Dz + c) = 0 , \\ z(T, x) = G'(x, m_1(T), \rho(T)) + z_T , \qquad \rho(t_0) = 0 , \\ a(x)Dz \cdot \nu_{|\partial\Omega} = 0 , \quad \Big(a(x)D\rho + \rho(H_1' + \tilde{b}) + mH_1''Dz + c\Big) \cdot \nu_{|\partial\Omega} = 0 , \end{cases}$$

where

$$\begin{aligned} h(t,x) &= h_1(t,x) + h_2(t,x)\,, \\ h_1(t,x) &= F'(x, m_1(t), \mu_2(t)) - F'(x, m_2(t), \mu_2(t))\,, \\ h_2(t,x) &= (H_1'(t,x) - H_2'(t,x)) \cdot Dv_2(t,x)\,, \\ c(t,x) &= \mu_2(t)(H_1' - H_2')(t,x) + [(m_1 H_1'' - m_2 H_2'')Dv_2](t,x)\,, \\ z_T(x) &= G'(x, m_1(T), \mu_2(T)) - G'(x, m_2(T), \mu_2(T))\,. \end{aligned}$$

Applying (3.5.20) we obtain this estimate on z :

$$\|z\|_{1,2+\alpha} \le C\left(\|z_T\|_{2+\alpha} + \|h\|_{0,\alpha} + \|c\|_{L^1}\right)\,.$$

Now we estimate the terms in the right-hand side.

The term with z_T, thanks to (3.5.20), is immediately estimated :

$$\begin{aligned} \|z_T\|_{2+\alpha} &\le \left\| \frac{\delta G}{\delta m}(\cdot, m_1(T), \cdot) - \frac{\delta G}{\delta m}(\cdot, m_2(T), \cdot) \right\|_{2+\alpha,1+\alpha} \|\mu_2(T)\|_{-(1+\alpha),N} \\ &\le C\mathbf{d}_1(m_{01}, m_{02})\,\|\mu_0\|_{-(1+\alpha)}\,. \end{aligned}$$

As regards the space estimate for h, we have

$$\|h(t,\cdot)\|_\alpha \le \big\| F'(\cdot, m_1(t), \mu_2(t)) - F'(\cdot, m_2(t), \mu_2(t)) \big\|_\alpha + \big\| (H_1' - H_2')(t,\cdot) Dv_2(t,\cdot) \big\|_\alpha\,.$$

The first term is bounded as z_T :

$$\big\| F'(\cdot, m_1(t), \mu_2(t)) - F'(\cdot, m_2(t), \mu_2(t)) \big\|_\alpha \le C\mathbf{d}_1(m_{01}, m_{02})\,\|\mu_0\|_{-(1+\alpha)}\,.$$

The second term, using (3.4.1) and (3.5.20) can be estimated in this way :

$$\big\| (H_1' - H_2')(t,\cdot) Dv_2(t,\cdot) \big\|_\alpha \le C\,\|(u_1 - u_2)(t)\|_{1+\alpha}\,\|v_2(t)\|_{1+\alpha} \le C\mathbf{d}_1(m_{01}, m_{02})\,\|\mu_0\|_{-(1+\alpha)}\,.$$

In summary,

$$\|h\|_{0,\alpha} = \sup_{t\in[0,T]} \|b(t,\cdot)\|_\alpha \le C\mathbf{d}_1(m_{01}, m_{02})\,\|\mu_0\|_{-(1+\alpha)}\,.$$

Finally, we estimate $\|c\|_{L^1}$. We have

$$\begin{aligned} \|c\|_{L^1} &= \int_0^T\!\!\int_\Omega (H_1' - H_2')(t,x)\,\mu_2(t,dx)\,dt + \int_0^T\!\!\int_\Omega H_1''(t,x) Dv_2(t,x)\,(m_1(t) - m_2(t))(dx)\,dt \\ &+ \int_0^T\!\!\int_\Omega (H_1'' - H_2'')(t,x)\,Dv_2(t,x)\,m_2(t,dx)\,dt \le C\,\|u_1 - u_2\|_{\frac{1+\alpha}{2},1+\alpha}\,\|\mu_2\|_{L^1} \\ &+ C\,\|u_1\|_{\frac{1+\alpha}{2},1+\alpha}\,\|v_2\|_{\frac{1+\alpha}{2},1+\alpha}\,\|m_1 - m_2\|_{L^1} + C\,\|u_1 - u_2\|_{\frac{1+\alpha}{2},1+\alpha}\,\|v_2\|_{\frac{1+\alpha}{2},1+\alpha} \end{aligned}$$

The first term in the right-hand side, thanks to (3.4.1) and (3.5.20), is bounded by

$$C\,\|u_1-u_2\|_{\frac{1+\alpha}{2},1+\alpha}\,\|\mu_2\|_{L^1}\le C\,\mathbf{d}_1(m_{01},m_{02})\,\|\mu_0\|_{-(1+\alpha)}\ ,$$

while the second term is estimated from above, using (3.5.17) and (3.5.20), by

$$C\,\|u_1\|_{\frac{1+\alpha}{2},1+\alpha}\,\|v_2\|_{\frac{1+\alpha}{2},1+\alpha}\,\|m_1(t)-m_2(t)\|_{L^1(\Omega)}\le C\,\mathbf{d}_1(m_{01},m_{02})\,\|\mu_0\|_{-(1+\alpha)}\ .$$

Finally, for the third term we use again (3.4.1) and (3.5.20) in order to obtain this bound :

$$C\,\|u_1-u_2\|_{\frac{1+\alpha}{2},1+\alpha}\,\|v_2\|_{\frac{1+\alpha}{2},1+\alpha}\le C\,\mathbf{d}_1(m_{01},m_{02})\,\|\mu_0\|_{-(1+\alpha)}\ .$$

Then

$$\|c\|_{L^1}\le C\mathbf{d}_1(m_{01},m_{02})\,\|\mu_0\|_{-(1+\alpha)}\ .$$

Putting together all these estimates, we finally obtain :

$$\|z\|_{1,2+\alpha}\le C\mathbf{d}_1(m_{01},m_{02})\,\|\mu_0\|_{-(1+\alpha)}\ .$$

Since

$$z(t_0,x)=\int_\Omega\left(\frac{\delta U}{\delta m}(t_0,x,m_1,y)-\frac{\delta U}{\delta m}(t_0,x,m_2,y)\right)\mu_0(dy)\,,$$

we have proved (3.7.4). □

This theorem is a fundamental step in order to prove this technical lemma.

Lemma 3.7.2. *Suppose hypotheses of the previous theorem are satisfied. Then, if* $m\in\mathcal{P}(\Omega)$ *and* $\phi\in L^2(m,\mathbb{R}^d)$ *is a bounded vector field,*

$$\left\|U(t,\cdot,(id+\phi)\sharp m)-U(t,\cdot,m)-\int_\Omega D_mU(t,\cdot,m,y)\cdot\phi(y)\,dm(y)\right\|_{1+\alpha}\le C\,\|\phi\|^2_{L^2(m)}\ . \tag{3.7.5}$$

Démonstration. For simplicity, we divide the proof in two steps.

Step 1. We start proving the following inequality for all m_0, $m_1\in\mathcal{P}(\Omega)$:

$$\left\|U(t,\cdot,m_1)-U(t,\cdot,m_0)-\int_\Omega\frac{\delta U}{\delta m}(t,\cdot,m_0,y)d(m_1-m_0)(y)\right\|_{1+\alpha}\le C\mathbf{d}_1(m_1,m_0)^2\,. \tag{3.7.6}$$

Since

$$U(t,x,m_1)-U(t,x,m_0)=\int_0^1\int_\Omega\frac{\delta U}{\delta m}(t,x,m_s,y)(m_1(dy)-m_0(dy))\,ds\,,$$

with $m_s = sm_1 + (1-s)m_0$, we have that the left-hand side of (3.7.6) is equal to

$$\left\| \int_0^1 \int_\Omega \left(\frac{\delta U}{\delta m}(t,\cdot,m_s,y) - \frac{\delta U}{\delta m}(t,\cdot,m_0,y) \right) (m_1 - m_0)(dy)\, ds \right\|_{1+\alpha} . \tag{3.7.7}$$

Using (3.7.4), (3.7.7) is estimated from above by

$$\left\| \frac{\delta U}{\delta m}(t,\cdot,m_s,\cdot) - \frac{\delta U}{\delta m}(t,\cdot,m_0,\cdot) \right\|_{1+\alpha,1+\alpha} \mathbf{d}_1(m_s,m_0) \le C\mathbf{d}_1(m_1,m_0)^2 ,$$

which concludes the first step.

Step 2. In this step we prove (3.7.5).

Thanks to (3.7.6), we know that

$$\begin{aligned} &\left\| U(t,\cdot,(id+\phi)\sharp m) - U(t,\cdot,m) - \int_\Omega \frac{\delta U}{\delta m}(t,\cdot,m,y) d((id+\phi)\sharp m - m)(y) \right\|_{1+\alpha} \\ &\le C\mathbf{d}_1((id+\phi)\sharp m, m)^2 \le C \left\|\phi\right\|^2_{L^2(m)} . \end{aligned}$$

So, it is sufficient to prove that

$$\left\| \int_\Omega \frac{\delta U}{\delta m}(t,\cdot,m,y) d((id+\phi)\sharp m - m)(y) - \int_\Omega D_m U(t,\cdot,m,y)\cdot\phi(y)\, dm(y) \right\|_{1+\alpha} \le C \left\|\phi\right\|^2_{L^2(m)} .$$

Since $D_m U = D_y \frac{\delta U}{\delta m}$, the left-hand side can be rewritten as

$$\begin{aligned} &\left\| \int_\Omega \left(\frac{\delta U}{\delta m}(t,\cdot,m,y+\phi(y)) - \frac{\delta U}{\delta m}(t,\cdot,m,y) - D_m U(t,\cdot,m,y)\cdot\phi(y) \right) dm(y) \right\|_{1+\alpha} \\ =& \left\| \int_0^1 \int_\Omega (D_m U(t,\cdot,m,y+t\phi(y)) - D_m U(t,\cdot,m,y))\cdot\phi(y)\, dm(y) dt \right\|_{1+\alpha} \\ =& \left\| \int_0^1 \int_0^1 \int_\Omega s D_y D_m U(t,\cdot,m,st\phi(y))\phi(y)\phi(y)\, dm(y)\, dt\, ds \right\|_{1+\alpha} \\ \le& C \left\| \frac{\delta U}{\delta m} \right\|_{1+\alpha,2} \left\|\phi\right\|^2_{L^2(m)} \le C \left\|\phi\right\|^2_{L^2(m)} , \end{aligned}$$

where we used (3.5.39) in the last passage. This concludes the proof. □

Now we are ready to prove a regularity result for the u_i^N functions defined in (3.7.2).

Proposition 3.7.3. *For all $j \neq i$, the following formulas for the derivatives of u_i^N hold true :*

$$D_{x_j} u_i^N(t,\boldsymbol{x}) = \frac{1}{N-1} D_m U(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j) , \tag{3.7.8}$$

$$D^2_{x_i,x_j} u_i^N(t,\boldsymbol{x}) = \frac{1}{N-1} D_x D_m U(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j) , \tag{3.7.9}$$

$$\left| D^2_{x_j,x_j} u_i^N(t,\boldsymbol{x}) - \frac{1}{N-1} D_y D_m U(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j) \right| \le \frac{C}{N^2} , \tag{3.7.10}$$

where the last inequality holds a.e. $\boldsymbol{x} \in \Omega^N$*.*

Démonstration. Thanks to the regularity of $D_m U$, the second equality is an obvious consequence of the first one. So, we restrict ourselves to the proof of the first formula.

We consider $\boldsymbol{x} = (x_1, \dots, x_N)$ with $x_j \neq x_k$ when $j \neq k$ and $\varepsilon := \min_{j \neq k} |x_j - x_k|$.

We fix a vector $\boldsymbol{v} = (v_1, \dots, v_N)$ with $v_i = 0$ and we consider a smooth and bounded vector field such that, for all $x \in B_{\frac{\varepsilon}{3}}(x_j)$,

$$\phi(x) = v_j\,.$$

We note that

$$u_i^N(t, \boldsymbol{x} + \boldsymbol{v}) = U(t, x_i, (id + \phi)\sharp m_{\boldsymbol{x}}^{N,i})\,, \qquad u_i^N(t, \boldsymbol{x}) = U(t, x_i, m_{\boldsymbol{x}}^{N,i})\,.$$

Then, (3.7.5) implies that

$$u_i^N(t, \boldsymbol{x} + \boldsymbol{v}) = u_i^N(t, \boldsymbol{x}) + \int_\Omega D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, y) \cdot \phi(y)\, dm_{\boldsymbol{x}}^{N,i}(y) + o\left(\|v\|_{L^2\left(m_{\boldsymbol{x}}^{N,i}\right)}\right)$$

So, computing the integral and the norm in the right-hand side, we find

$$u_i^N(t, \boldsymbol{x} + \boldsymbol{v}) = u_i^N(t, \boldsymbol{x}) + \frac{1}{N-1} \sum_{j \neq i} D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j) \cdot v_j + o(|v|)\,.$$

This Taylor expansion proves the first formula for all points $\boldsymbol{x}$ with $x_j \neq x_k$ when $j \neq k$. Since this subset is dense in Ω^N, the theorem is concluded thanks to the continuity of $D_m U$.

As regards the last inequality, we start showing that $D_{x_j} u_i^N$ is a Lipschitz function in the space variable. Actually

$$|D_{x_j} u_i^N(t, \boldsymbol{x}) - D_{x_j} u_i^N(t, \boldsymbol{y})| \le \frac{C}{N} \left(|D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j) - D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, y_j)|\right)$$
$$+ \frac{C}{N} \left(|D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, y_j) - D_m U(t, x_i, m_{\boldsymbol{y}}^{N,i}, y_j)|\right)\,.$$

The first term in the right-hand side is immediately controlled by

$$\frac{C}{N} |\boldsymbol{x} - \boldsymbol{y}|\,.$$

As regards the second term, we have

$$\frac{C}{N} \left(|D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, y_j) - D_m U(t, x_i, m_{\boldsymbol{y}}^{N,i}, y_j)|\right) \le \frac{C}{N} \mathbf{d}_1(m_{\boldsymbol{x}}^{N,i}, m_{\boldsymbol{y}}^{N,i}) \le \frac{C}{N} |\boldsymbol{x} - \boldsymbol{y}|\,.$$

This means that $D^2_{x_j,x_j}u_i^N$ exists almost everywhere, and

$$\left\| D^2_{x_j,x_j}u_i^N \right\|_\infty \le \frac{C}{N}\,.$$

To prove (3.7.10), we estimate the quantity

$$\left| \frac{D_{x_j}u^{N,i}(t,\boldsymbol{x}+h\boldsymbol{e_j^k}) - D_{x_j}u^{N,i}(t,\boldsymbol{x})}{h} - \frac{1}{N-1}\partial_{y_k}D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\right|\,,$$

where $\boldsymbol{e_{jk}} = (e^1_{jk},\dots,e^N_{jk})$, with $e^l_{jk}=0$ if $l\neq j$ and $e^j_{jk}=e_k\in\mathbb{R}^d$.

Using (3.7.8), we can bound the quantity above by

$$\frac{C}{N}\left|\frac{D_mU(t,x_i,m^{N,i}_{\boldsymbol{x}+h\boldsymbol{e_j^k}},x_j+he_k) - D_mU(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j+he_k)}{h}\right|$$
$$+\frac{C}{N}\left|\frac{D_mU(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j+he_k) - D_mU(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j)}{h} - \partial_{y_k}D_mU(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j)\right|\,.$$

The first term is estimated from above, using (3.7.4), by

$$\frac{C}{Nh}\mathbf{d}_1(m^{N,i}_{\boldsymbol{x}+h\boldsymbol{e_j^k}},m^{N,i}_{\boldsymbol{x}}) \le \frac{C}{N^2}\,,$$

while the second term, using Lagrange's Theorem and (3.5.39), is equal to

$$\frac{C}{N}|\partial_{y_k}D_mU(t,x_i,m^{N,i}_{\boldsymbol{x}},x_h) - \partial_{y_k}D_mU(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j)| \le \frac{C}{N}h^\alpha\,,$$

for a certain x_h in the line segment between x_j and x_j+he_k. Then, for h sufficiently small, we obtain

$$\left| \frac{D_{x_j}u^{N,i}(t,\boldsymbol{x}+h\boldsymbol{e_j^k}) - D_{x_j}u^{N,i}(t,\boldsymbol{x})}{h} - \frac{1}{N-1}\partial_{y_k}D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\right| \le \frac{C}{N^2}\,,$$

and, passing to the limit as $h\to 0$ for all $1\le k\le d$,

$$\left|D^2_{x_j,x_j}u_i^N(t,\boldsymbol{x}) - \frac{1}{N-1}D_yD_mU(t,x_i,m^{N,i}_{\boldsymbol{x}},x_j)\right| \le \frac{C}{N^2}\,,$$

which is exactly (3.7.10). □

Now we are ready to prove a first result, showing that $(u_i^N)_{1\le i\le N}$ is "almost" a solution to the Nash system (3.7.1).

Theorem 3.7.4. *Let hypotheses 3.2.4 satisfied. Then* $u_i^N \in \mathcal{C}^1([0,T]\times\Omega^N)$ *,* $u_i^N(t,\cdot)\in W^{2,\infty}(\Omega^N)$ *and* u_i^N *solves almost everywhere the following equation :*

$$\begin{cases} -\partial_t u_i^N - \sum_j \mathrm{tr}(a(x_j)D^2_{x_jx_j}u_i^N) + H(x_i, D_{x_i}u_i^N) + \sum_{j\neq i} H_p(x_j, D_{x_j}u_j^N)\cdot D_{x_j}u_i^N \\ = F(t,x_i,m_{\boldsymbol{x}}^{N,i}) + r_i^N(t,\boldsymbol{x})\,, \\ u_i^N(T,\boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i})\,, \\ a(x_j)D_{x_j}u_i^N\cdot\nu(x_j)_{|x_j\in\partial\Omega} = 0\,, \qquad j=1,\cdots,N\,. \end{cases} \tag{3.7.11}$$

where $r_i^N \in L^\infty$ *with*

$$\|r_i^N\|_\infty \leq \frac{C}{N}\,.$$

Démonstration. The regularity of u_i^N follows from (3.7.3) and Proposition 3.7.3.

The boundary condition of u_i^N is an immediate consequence of the representation formula for the derivatives of u_i^N and the boundary conditions of (3.1.1).

Actually for $j=i$ we have

$$a(x_i)D_{x_i}u_i^N\cdot\nu(x_i)_{|x_i\in\partial\Omega} = a(x_i)D_xU(t,x_i,m_{\boldsymbol{x}}^{N,i})\cdot\nu(x_i)_{|x_i\in\partial\Omega} = 0$$

for the first boundary condition of (3.1.1). On the other hand, for $j\neq i$ we have

$$a(x_j)D_{x_j}u_i^N\cdot\nu(x_j)_{|x_j\in\partial\Omega} = \frac{1}{N-1}a(x_j)D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot\nu(x_j)_{|x_j\in\partial\Omega} = 0$$

for the second boundary condition of (3.1.1).

Since U is a solution of the Master Equation, we find, evaluating (3.1.1) at $(t,x_i,m_{\boldsymbol{x}}^{N,i})$,

$$\begin{aligned} &-\partial_t U - \mathrm{tr}\left(a(x_i)D_x^2U\right) + H(x_i, D_xU(t,x,m)) - \int_\Omega \mathrm{tr}\left(a(y)D_yD_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},y)\right)dm_{\boldsymbol{x}}^{N,i}(y) \\ &+ \int_\Omega D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},y)\cdot H_p(y, D_xU(t,y,m_{\boldsymbol{x}}^{N,i}))dm_{\boldsymbol{x}}^{N,i}(y) = F(t,x_i,m_{\boldsymbol{x}}^{N,i})\,. \end{aligned}$$

So, as $\partial_t U(t,x_i,m_{\boldsymbol{x}}^{N,i}) = \partial_t u_i^N(t,\boldsymbol{x})$ and $D_xU(t,x_i,m_{\boldsymbol{x}}^{N,i}) = D_{x_i}u_i^N(t,\boldsymbol{x})$, we obtain initially this equation for u_i^N :

$$\begin{aligned} &-\partial_t u_i^N - \mathrm{tr}(a(x_i)D^2_{x_ix_i}u_i^N) + H(x_i, D_{x_i}u_i^N) - \int_\Omega \mathrm{tr}\left(a(y)D_yD_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},y)\right)dm_{\boldsymbol{x}}^{N,i}(y) \\ &+ \frac{1}{N-1}\sum_{j\neq i} H_p(x_j, D_xU(t,x_j,m_{\boldsymbol{x}}^{N,i}))\cdot D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j) = F(t,x_i,m_{\boldsymbol{x}}^{N,i})\,. \end{aligned}$$

Thanks to the derivative formulas of u_N^i, we know that

$$\frac{1}{N-1} D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j) = D_{x_j} u_i^N(t, \boldsymbol{x}) \implies \| D_{x_j} u_i^N \|_\infty \le \frac{C}{N}\,.$$

Moreover, using the Lipschitz continuity of $D_x U$ with respect to m, stated in (3.7.4), we have

$$|H_p(x_j, D_x U(t, x_j, m_{\boldsymbol{x}}^{N,i})) - H_p(x_j, D_x U(t, x_j, m_{\boldsymbol{x}}^{N,j}))| \le C\mathbf{d}_1(m_{\boldsymbol{x}}^{N,i}, m_{\boldsymbol{x}}^{N,j}) \le \frac{C}{N}\,.$$

Hence, we get

$$\begin{aligned}\frac{1}{N-1}\sum_{j\neq i} H_p(x_j, D_x U(t, x_j, m_{\boldsymbol{x}}^{N,i}))\cdot D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j)& \\ = \sum_{j\neq i} H_p(x_j, D_{x_j} u_j^N)\cdot D_{x_j} u_i^N + O\left(\frac{1}{N}\right)&\,.\end{aligned}$$

We conclude analyzing the last integral term. We have

$$\begin{aligned}\int_\Omega \operatorname{tr}\left(a(y) D_y D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, y)\right) dm_{\boldsymbol{x}}^{N,i}(y) &= \frac{1}{N-1}\sum_{j\neq i} \operatorname{tr}(a(x_j) D_y D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j)) \\ &= \sum_{j\neq i} D^2_{x_j, x_j} u_i^N(t, \boldsymbol{x}) + O\left(\frac{1}{N}\right)\,.\end{aligned}$$

Collecting all these estimations, we obtain (3.7.11), which concludes the theorem. □

Now we turn to the main convergence result. To do that, we consider the two functions $(u_i^N)_i$ and $(v_i^N)_i$, where u_i^N is defined in (3.7.2) and $(v_i^N)_i$ are solutions of the system (3.7.1).

We note that these solutions are *symmetrical*. This means that there exist two functions V^N and $U^N : \Omega \times \Omega^{N-1} \to \mathbb{R}$ such that, for all $x \in \Omega$, the functions $(y_1, \dots, y_{N-1}) \to V^N(x, (y_1, \dots, y_{N-1}))$ and $(y_1, \dots, y_{N-1}) \to U^N(x, (y_1, \dots, y_{N-1}))$ are invariant under permutations and

$$\begin{aligned} v_i^N(t, \boldsymbol{x}) &= V^N(x_i, (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N))\,, \\ u_i^N(t, \boldsymbol{x}) &= U^N(x_i, (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N))\,. \end{aligned}$$

We fix $t_0 \in [0,T)$, $m_0 \in \mathcal{P}(\Omega)$ and $\boldsymbol{Z} = (Z^i)_i$ a family of i.i.d random variables of law m_0. We consider the process $\boldsymbol{Y}_t = (Y_t^i)_i$ solution of the following system :

$$\begin{cases} dY_t^i = -H_p(Y_t^i, D_{x_i} v_i^N(t, \boldsymbol{Y}_t))\, dt + \sqrt{2}\sigma(Y_t^i) dB_t^i - dk_t^i\,, \\ Y_{t_0}^i = Z^i\,, \end{cases} \tag{3.7.12}$$

where k_i^t is a *reflected process along the co-normal.*

We need to use an extension of the Ito's formula, stated in the following Lemma.

Lemma 3.7.5. *Let $\phi : [0,T] \times \Omega \to \mathbb{R}$ a $W^{2,\infty}$ function with respect to x and a C^1 function with respect to t, such that*

$$a(x) D\phi(x) \cdot \nu(x)_{|\partial\Omega} = 0\,.$$

Let $m_0 \in \mathcal{P}(\Omega)$ and let X_t be a process in the probability space $(\tilde{\Omega}, (\mathcal{F}_t)_t, \mathbb{P})$, with initial density m_0, satisfying

$$dX_t = b(t, X_t)\, dt + \sigma(X_t)\, dB_t - dk_t^i\,,$$

where $(B_t)_t$ is a Brownian motion, b and σ are bounded functions respectively in L^∞ and $\mathcal{C}^{1+\alpha}$, with σ a uniformly elliptic matrix, and k_i^t is a reflected process along the co-normal. Then the following formula holds $\forall t$ and a.s. in $\omega \in \tilde{\Omega}$:

$$\begin{aligned} \phi(t, X_t) = \phi(0, X_0) &+ \int_0^t \left(\phi_t(s, X_s) + \frac{1}{2} \mathrm{tr}(a(X_s) D^2\phi(s, X_s)) + b(s, X_s) \cdot D\phi(s, X_s) \right) ds \\ &+ \int_0^t \sigma(X_s) D\phi(s, X_s)\, dB_s\,. \end{aligned}$$

Démonstration. We consider ϕ^n a sequence of $\mathcal{C}^{1,2,N}$ functions, bounded uniformly in n together with their derivatives, such that $\phi^n \to \phi$ pointwise together with its first order derivatives in space and time, and almost everywhere for the second order derivatives in space.

We define $a = \sigma\sigma^*$. The classical Ito's formula for ϕ^n tells us that

$$\begin{aligned} \phi^n(t, X_t) = \phi^n(0, X_0) &+ \int_0^t \left(\phi_t^n(s, X_s) + \frac{1}{2} \mathrm{tr}(a(X_s) D^2\phi^n(s, X_s)) + b(s, X_s) \cdot D\phi^n(s, X_s) \right) ds \\ &+ \int_0^t \sigma(X_s) D\phi^n(s, X_s) dB_s - \int_0^t a(X_s) D\phi^n(s, X_s) \nu(X_s) d|k|_s\,, \end{aligned}$$

and so, since ϕ^n satisfies $a(x) D\phi^n \cdot \nu_{|\partial\Omega} = 0\,,$

$$\begin{aligned} \phi^n(t, X_t) = \phi^n(0, X_0) &+ \int_0^t \left(\phi_t^n(s, X_s) + \frac{1}{2} \mathrm{tr}(a(X_s) D^2\phi^n(s, X_s)) + b(s, X_s) \cdot D\phi^n(s, X_s) \right) ds \\ &+ \int_0^t \sigma(X_s) D\phi^n(s, X_s) dB_s\,. \end{aligned}$$

Since $\phi^n \to \phi$ pointwise, we can pass to the limit for the terms outside the integrals.

For the term in the deterministic integral, we note that the law $m(t)$ of the process X_t satisfies the following Fokker-Planck equation :

$$\begin{cases} m_t - \operatorname{div}(a(x)Dm) - \operatorname{div}(m(b+\tilde{b})) = 0\,, \\ m(0) = m_0\,, \\ \left[aDm + m(b+\tilde{b})\right] \cdot \nu_{|\partial\Omega} = 0\,, \end{cases}$$

and so, with the same strategies of Corollary (3.5.6), we have that m is globally bounded in $L^p(Q_T)$ for some $p > 1$.

Hence, we have

$$\begin{aligned} &\mathbb{E}\left[\int_0^t \left|\phi_t^n(s,X_s) + \frac{1}{2}\operatorname{tr}(a(X_s)D^2\phi^n(s,X_s)) + b(s,X_s)\cdot D\phi^n(s,X_s)\right.\right. \\ &\left.\left. -\phi_t(s,X_s) - \frac{1}{2}\operatorname{tr}(a(X_s)D^2\phi(s,X_s)) - b(s,X_s)\cdot D\phi(s,X_s)\right| ds\right] \\ \le &\int_0^t \int_\Omega \left(|\phi_t^n - \phi_t| + \frac{1}{2}|\operatorname{tr}(aD^2(\phi^n - \phi))| + |b\cdot(D\phi^n - D\phi)|\right) m(s,x)\,dxds \to 0\,, \end{aligned}$$

where the dominated convergence is guaranteed by the *a.e.* convergence of ϕ_t^n, $D\phi^n$, $D^2\phi^n$ and the global boundedness of m in L^p and of ϕ_t^n, $D\phi^n$ and $D^2\phi^n$ in L^∞.

As regards the last term, the *a.s.* convergence is guaranteed by the property of the stochastic integral. Actually we have

$$\begin{aligned} \mathbb{E}\left[\left(\int_0^t \sigma_s D\phi^n(s,X_s)dB_s - \int_0^t \sigma_s D\phi(s,X_s)dB_s\right)^2\right] &= \mathbb{E}\left[\int_0^t |\sigma_s(D\phi^n - D\phi)(s,X_s)|^2\,ds\right] \\ &\le C\,\|D\phi^n - D\phi\|_\infty^2 \to 0\,. \end{aligned}$$

This concludes the Lemma. □

We note that, if there are no reflection terms, then the boundary condition for ϕ can be removed.

The last theorem before the main result is the following.

Theorem 3.7.6. *Assume hypotheses 3.2.4 hold. Then, for any $1 \le i \le N$, we have*

$$\mathbb{E}\left[\int_{t_0}^T \left|D_{x_i}v_i^N(t,\boldsymbol{Y}_t) - D_{x_i}u_i^N(t,\boldsymbol{Y}_t)\right|^2 dt\right] \le \frac{C}{N^2}\,. \tag{3.7.13}$$

Moreover, $\mathbb{P}$-a.s.,

$$|u_i^N(t_0,\boldsymbol{Z}) - v_i^N(t_0,\boldsymbol{Z})| \le \frac{C}{N}\,. \tag{3.7.14}$$

Démonstration. The proof is almost exactly the same of *Theorem 6.2.1* of [32].

Without loss of generality, we work with $t_0 = 0$ and we start proving (3.7.13).

To simplify the rest of the proof, we will use the following notations :

$$\begin{aligned} U_t^{N,i} &= u_i^N(t, \boldsymbol{Y}_t)\,, \qquad & DU_t^{N,i,j} &= D_{x_j} u_i^N(t, \boldsymbol{Y}_t)\,, \\ V_t^{N,i} &= v_i^N(t, \boldsymbol{Y}_t)\,, \qquad & DV_t^{N,i,j} &= D_{x_j} v_i^N(t, \boldsymbol{Y}_t)\,. \end{aligned}$$

Using Lemma 3.7.5 and the equation (3.7.1) satisfied by v_i^N, we obtain

$$\begin{aligned} dV_t^{N,i} = &\Big[H(Y_t^i, DV_t^{N,i,i}) - DV_t^{N,i,i} \cdot H_p(Y_t^i, DV_t^{N,i,i}) - F(Y_t^i, m_{\boldsymbol{Y}_t}^{N,i})\Big]\, dt \\ &+ \sqrt{2} \sum_j \sigma(Y_t^i) DV_t^{N,i,j} dB_t^j \end{aligned}$$

Similarly, using equation (3.7.11) satisfied by u_i^N, we obtain

$$\begin{aligned} dU_t^{N,i} = &\Big[H(Y_t^i, DU_t^{N,i,i}) - DU_t^{N,i,i} \cdot H_p(Y_t^i, DU_t^{N,i,i}) - F(Y_t^i, m_{\boldsymbol{Y}_t}^{N,i}) - r_i^N(t, \boldsymbol{Y}_t)\Big]\, dt \\ &- \sum_j DU_t^{N,i,j} \Big(H_p(Y_t^j, DV_t^{N,j,j}) - H_p(Y_t^j, DU_t^{N,j,j})\Big)\, dt + \sqrt{2} \sum_j \sigma(Y_t^i) DU_t^{N,i,j} dB_t^j \end{aligned}$$

Both reflection terms are null, because of the boundary conditions on u_i^N and v_i^N, and Lemma 3.7.5.

Now we apply the Ito's formula to the process

$$\Big(U_t^{N,i} - V_t^{N,i}\Big)^2 \,,$$

obtaining

$$d\Big(U_t^{N,i} - V_t^{N,i}\Big)^2 = (A_t + B_t)dt + 2\sqrt{2}\Big(U_t^{N,i} - V_t^{N,i}\Big) \sum_j \Big[\sigma(Y_t^i)(DV_t^{N,i,j} - DU_t^{N,i,j})\Big]\, dB_t^j \,,$$

where

$$\begin{aligned} A_t = &\,2(U_t^{N,i} - V_t^{N,i}) \Big(H(Y_t^i, DU_t^{N,i,i}) - H(Y_t^i, DV_t^{N,i,i})\Big) \\ &- 2(U_t^{N,i} - V_t^{N,i}) \Big(DU_t^{N,i,i}(H_p(Y_t^i, DU_t^{N,i,i}) - H_p(Y_t^i, DV_t^{N,i,i}))\Big) \\ &- 2(U_t^{N,i} - V_t^{N,i}) \Big((DU_t^{N,i,i} - DV_t^{N,i,i}) H_p(Y_t^i, DV_t^{N,i,i}) - r_i^N(t, \boldsymbol{Y}_t)\Big) \end{aligned}$$

and

$$
\begin{aligned}
B_t &= 2\sum_j \langle a(Y_i^t)(DU_t^{N,i,j} - DV_t^{N,i,j}), DU_t^{N,i,j} - DV_t^{N,i,j}\rangle \\
&\quad - 2(U_t^{N,i} - V_t^{N,i}) \sum_j DU_t^{N,i,j} \left(H_p(Y_t^j, DV_t^{N,j,j}) - H_p(Y_t^j, DU_t^{N,j,j}) \right)
\end{aligned}
$$

Now we want to integrate from t to T the above formula and take the conditional expectation given $\boldsymbol{Z}$.

We recall that a is a uniformly elliptic matrix, and $\exists \nu > 0$ such that

$$
\langle a(x)v, v\rangle \geq \nu |v|^2 , \quad \text{for all } v \in \Omega .
$$

Moreover, thanks to the previous results,

$$
|DU_t^{N,i,j}| \leq \frac{C}{N} , \qquad |r_i^N| \leq \frac{C}{N} \qquad \text{for a certain } C > 0 .
$$

We obtain :

$$
\begin{aligned}
&\mathbb{E}^{\boldsymbol{Z}} \left[|U_t^{N,i} - V_t^{N,i}|^2 \right] + 2\nu \sum_j \mathbb{E}^{\boldsymbol{Z}} \left[\int_t^T |DU_s^{N,i,j} - DV_s^{N,i,j}|^2 ds \right] \\
&\leq \frac{C}{N} \int_t^T \mathbb{E}^{\boldsymbol{Z}} \left[|U_s^{N,i} - V_s^{N,i}| \right] ds + C \int_t^T \mathbb{E}^{\boldsymbol{Z}} \left[|U_s^{N,i} - V_s^{N,i}| \cdot |DU_s^{N,i,i} - DV_s^{N,i,i}| \right] ds \\
&+ \frac{C}{N} \sum_{j \neq i} \int_t^T \mathbb{E}^{\boldsymbol{Z}} \left[|U_s^{N,i} - V_s^{N,i}| \cdot |DU_s^{N,j,j} - DV_s^{N,j,j}| \right] ds .
\end{aligned}
$$

By a standard convexity argument, with the use of a generalized Young's inequality, we get

$$
\begin{aligned}
&\mathbb{E}^{\boldsymbol{Z}} \left[|U_t^{N,i} - V_t^{N,i}|^2 \right] + \nu \mathbb{E}^{\boldsymbol{Z}} \left[\int_t^T |DU_s^{N,i,i} - DV_s^{N,i,i}|^2 ds \right] \\
&\leq \frac{C}{N^2} + C \int_t^T \mathbb{E}^{\boldsymbol{Z}} \left[|U_s^{N,i} - V_s^{N,i}|^2 \right] ds + \frac{\nu}{2N} \sum_j \mathbb{E}^{\boldsymbol{Z}} \left[\int_t^T |DU_s^{N,j,j} - DV_s^{N,j,j}|^2 ds \right] .
\end{aligned}
\tag{3.7.15}
$$

The last term in the right-hand side can be removed by taking the mean of the inequalities over $i \in 1, \ldots, N$. So we get

$$\begin{aligned}&\frac{1}{N}\sum_i \mathbb{E}^{\boldsymbol{Z}}\left[|U_t^{N,i}-V_t^{N,i}|^2\right]+\frac{\nu}{2N}\sum_i \mathbb{E}^{\boldsymbol{Z}}\left[\int_t^T |DU_s^{N,i,i}-DV_s^{N,i,i}|^2 ds\right]\\&\le \frac{C}{N^2}+\frac{C}{N}\sum_i\int_t^T \mathbb{E}^{\boldsymbol{Z}}\left[|U_s^{N,i}-V_s^{N,i}|^2\right]ds\,.\end{aligned}\tag{3.7.16}$$

Then we use Gronwall's Lemma and the terminal condition on u_i^N and v_i^N in order to obtain

$$\sup_{0\le t\le T}\left[\frac{1}{N}\sum_i \mathbb{E}^{\boldsymbol{Z}}[|U_t^{N,i}-V_t^{N,i}|^2]\right]\le\frac{C}{N^2}\,.$$

Plugging the last bound in (3.7.16) one has

$$\frac{1}{N}\sum_i \mathbb{E}^{\boldsymbol{Z}}\left[\int_0^T |DU_s^{N,i,i}-DV_s^{N,i,i}|^2 ds\right]\le\frac{C}{N^2}\,.$$

Finally, using this estimation in (3.7.15) and applying again Gronwall's Lemma, we conclude the first step :

$$\sup_{0\le t\le T}\mathbb{E}^{\boldsymbol{Z}}[|U_t^{N,i}-V_t^{N,i}|^2]+\mathbb{E}^{\boldsymbol{Z}}\left[\int_t^T |DU_s^{N,i,i}-DV_s^{N,i,i}|^2 ds\right]\le\frac{C}{N^2}\,.\tag{3.7.17}$$

This proves (3.7.13).

From (3.7.17), evaluated in $t=0$, we obtain $\mathbb{P}$-almost surely

$$|u_i^N(0,\boldsymbol{Z})-v_i^N(0,\boldsymbol{Z})|^2=\mathbb{E}^{\boldsymbol{Z}}[|U_t^{N,i}-V_t^{N,i}|^2]_{|t=0}\le\frac{C}{N^2}\,.$$

So, (3.7.14) is proved. □

Now we are ready to the last theorem of this chapter, which proves the main convergence result of the Nash systems towards the Master Equation.

Theorem 3.7.7. *Suppose hypotheses* (3.2.4) *hold true. Then, if we define*

$$m_{\boldsymbol{x}}^N := \frac{1}{N}\sum_i \delta_{x_i}\,,$$

we have

$$\sup_i |v_i^N(t_0,\boldsymbol{x})-U(t_0,x_i,m_{\boldsymbol{x}}^N)|\le\frac{C}{N}\,.\tag{3.7.18}$$

Moreover, if we set

$$w_i^N(t_0, x_i, m_0) := \int_{\Omega^{N-1}} v_i^N(t_0, \boldsymbol{x}) \prod_{j\neq i} m_0(dx_j)\,,$$

then

$$\left\| w_i^N(t_0, \cdot, m_0) - U(t_0, \cdot, m_0) \right\|_{L^1(m_0)} \leq C\omega_N\,, \tag{3.7.19}$$

where

$$\omega_N = \begin{cases} CN^{-\frac{1}{d}} & \text{if} \quad d \geq 3\,, \\ CN^{-\frac{1}{2}}\log(N) & \text{if} \quad d = 2\,, \\ CN^{-\frac{1}{2}} & \text{if} \quad d = 1\,. \end{cases}$$

Démonstration. We start choosing $m_0 = 1$. Then, (3.7.14) implies

$$|U(t_0, Z_i, m_{\boldsymbol{Z}}^{N,i}) - v_i^N(t_0, \boldsymbol{Z})| \leq \frac{C}{N} \qquad \mathbb{P} - a.s.\,.$$

Since the support of m_0 is Ω, this means, thanks to the continuity of v_i^N and U, that

$$|U(t_0, x_i, m_{\boldsymbol{x}}^{N,i}) - v_i^N(t_0, \boldsymbol{x})| \leq \frac{C}{N} \qquad \forall \boldsymbol{x} \in \Omega^N\,.$$

By the Lipschitz continuity of U, we have

$$|U(t_0, x_i, m_{\boldsymbol{x}}^{N,i}) - U(t_0, x_i, m_{\boldsymbol{x}}^N)| \leq C\mathbf{d}_1(m_{\boldsymbol{x}}^{N,i}, m_{\boldsymbol{x}}^N) \leq \frac{C}{N}\,.$$

Putting together the last two inequalities, we obtain (3.7.18).
To prove (3.7.19), we use the results of [3], [51], [58] in order to obtain

$$\int_{\Omega^{N-1}} |u_i^N(t_0, \boldsymbol{x}) - U(t_0, x_i, m_0)| \prod_{j\neq i} m_0(dx_j) \leq C \int_{\Omega^{N-1}} \mathbf{d}_1(m_{\boldsymbol{x}}^{N,i}, m_0) \prod_{j\neq i} m_0(dx_j) \leq C\omega_N\,,$$

where ω_N is defined exactly as in the right-hand side of (3.7.19).
With this inequality, we can conclude in this way :

$$\begin{aligned}
&\left\| w_i^N(t_0, \cdot, m_0) - U(t_0, \cdot, m_0) \right\|_{L^1(m_0)} \\
&= \int_\Omega \left| \int_{\Omega^{N-1}} \left(v_i^N(t_0, \boldsymbol{x}) - U(t_0, x_i, m_0)\right) \prod_{j\neq i} m_0(dx_j) \right| m_0(dx_i) \\
&\leq \mathbb{E}\left[|v_i^N(t, \boldsymbol{Z}) - u_i^N(t, \boldsymbol{Z})|\right] + \int_{\Omega^N} |u_i^N(t_0, \boldsymbol{x}) - U(t, x_i, m_0)| \prod_j m_0(dx_j) \\
&\leq \frac{C}{N} + C\omega_N \leq C\omega_N\,.
\end{aligned}$$

This proves (3.7.19) and concludes the theorem. □

Chapitre 4

Mean Field Games PDE with Controlled Diffusion

4.1 Introduction

In the last chapter of my thesis we study a general class of Mean Field Games of this form :

$$\begin{cases} \partial_t u + H^1(t,x,\nabla u) + H^2(t,x,\Delta u) = -F(t,x,m)\,, & (t,x) \in (0,T)\times\mathbb{R}^n\,, \\ \partial_t m - \Delta(m H^2_q(t,x,\Delta u)) + \mathrm{div}(m H^1_p(t,x,\nabla u)) = 0\,, & (t,x) \in (0,T)\times\mathbb{R}^n\,, \\ u(T,x) = G(x,m(T))\,, \qquad m(0,x) = m_0(x)\,, & x \in \mathbb{R}^n\,. \end{cases} \tag{4.1.1}$$

Here, $H^1(t,x,p)$ and $H^2(t,x,q)$ denotes the two Hamiltonians of the system, F is the running cost and G the final pay-off ; with the notation H^1_p and H^2_q we mean respectively the gradient and the derivative of H^1 and H^2 with respect to the last variable.

The most important difference with the other Mean Field Games systems we studied before is that here the nonlinearity of the Hamilton-Jacobi equation regards not only the first order term, but the second order too.

This class of MFG systems appears when, in the stochastic modelization, the control of the generic agent involves not only the drift, but also the diffusion term.

In the next Section we will give a significative modelization of this situation. For now, we stress the fact that models with controlled diffusion are widely required in financial applications, see for instance the works of Avellaneda et al. in [5], [6], [7].

Some cases of fully nonlinear Mean Field Games were studied in the literature. In [53] there are results for an ergodic Mean Field Games (so elliptic) with nonlinear second order term, and in [15] we find another class of fully nonlinear Mean Field Games systems, studied with a probabilistic approach.

But, as far as we know, there are no results for the well-posedness of a Mean Field Games system like (4.1.1). This is the aim of this chapter.

We give a short summary of the results.

In Section 2, as already said, we give a stochastic interpretation of the system, with a process controlled by two different controls, α for the drift term and σ for the diffusion term. Then we prove, in a simpler case, that the value function of this game is actually a viscosity solution of a fully nonlinear Hamilton-Jacobi equation.

In Section 3 we start studying the Hamilton-Jacobi equation of the system (4.1.1). Equations of this type has been studied in the literature, see for instance the works [11], [44], [81] and [97].

We start giving a suitable definition of viscosity solution. Then we prove existence and uniqueness of bounded solutions. Further we obtain, with a strengthening of the hypotheses, Lipschitz and semiconcave estimates for u.

In Section 4 we restrict ourselves in a regular case, with stronger hypotheses on H^1 and H^2; applying a regularity result of Krylov,see [70] and [71], we obtain that $u \in \mathcal{C}^{1+\frac{\gamma}{2},2+\gamma}$ for a certain $0 < \gamma < 1$, and so u is actually a solution of a linear PDE. Then, using the regularity results of [74], we obtain first an estimate for u in $\mathcal{C}^{1+\frac{\alpha}{2},2+\alpha}$, and then a $C^{3+\alpha}$ regularity for u in the space.

These regularity results will be essential in order to obtain, by approximation, a $\mathcal{C}^{1+\frac{\alpha}{2},3+\alpha}$ solution for the value function u in our case of Bellman operators. This will be done in Section 5.

This regularity will be essential in Section 6, where we finally prove existence and uniqueness of solutions for the problem (4.1.1), with a classical fixed point argument.

4.2 From stochastic model to deterministic PDEs

Now we give a stochastic interpretation of the system (4.1.1).

In this framework, the generic player choose two controls, $\alpha.$ and $\sigma.$, and plays his dynamic, modelized by the following process :

$$\begin{cases} dX_s^{\alpha,\sigma} = b(X_s^{\alpha,\sigma}, \alpha_s)ds + \Gamma(X_s^{\alpha,\sigma}, \sigma_s)dB_s \\ X_t^{\alpha,\sigma} = x, , \end{cases} \tag{4.2.1}$$

where $x \in \mathbb{R}^n$, $t \in [0, T)$ and also the random variable X_s takes values in the whole space $\mathbb{R}^n$.

The time dependent control variables α_s, σ_s live in the space of bounded controls $\mathcal{U} \times \mathcal{S}$, denoting the compact sets $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{S} \subset \mathbb{R}$, whereas for the continuous functions

$$b : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}\,, \qquad \Gamma : \mathbb{R}^n \times \mathcal{S} \to \mathbb{R}$$

we require that, for some constants $M > 0$, $\lambda_2 > \lambda_1 > 0$,

$$\begin{aligned} &|b(x, \alpha)| \le M && \forall x \in \mathbb{R}^n,\ \alpha \in \mathcal{U}\,, \\ &\lambda_1 < \Gamma(x, \sigma) < \lambda_2 && \forall x \in \mathbb{R}^n\ \sigma \in \mathcal{S}\,. \end{aligned} \tag{4.2.2}$$

In particular, Γ is bounded from below (and from above) by a positive constant, in order to ensure the uniform ellipticity of the process. We denote the spaces of admissible controls where α and σ live respectively as $\mathcal{A}_t^{\mathcal{U}}$ and $\mathcal{A}_t^{\mathcal{S}}$. From now on we omit the superscripts on the process to simplify the notations.

The cost function for the player is defined as

$$\mathcal{J}(x, t, \alpha., \sigma.) := \mathbb{E}\left[\int_t^T (L_1(s, X_s, \alpha_s) + L_2(s, X_s, \sigma_s) + F(s, X_s, m(s)))\, ds + G(X_T, m(T))\right], \tag{4.2.3}$$

where $m(s)$ is a density function, which is fixed at the moment and which will denote the density of the generic player, with initial condition $X_0 = m_0$.

Our aim is to find a Hamilton-Jacobi equation, which is solved by the function

$$u(x, t) := \inf_{\substack{\alpha. \in \mathcal{A}_t^{\mathcal{U}} \\ \sigma. \in \mathcal{A}_t^{\mathcal{S}}}} \mathcal{J}(x, t, \alpha., \sigma.)\,. \tag{4.2.4}$$

For this we use the dynamic programming principle written in the following form :

$$\begin{aligned} u(x, t) = \inf_{\substack{\alpha. \in \mathcal{A}_t^{\mathcal{U}} \\ \sigma. \in \mathcal{A}_t^{\mathcal{S}}}} \mathbb{E}\left[\int_t^{t+h} (L_1(s, X_s, \alpha_s) + L_2(s, X_s, \sigma_s)\right. \\ \left. + F(s, X_s, m(s))ds + v(X_{t+h}, t + h)\right]. \end{aligned} \tag{4.2.5}$$

The proof of the verification theorem in this general case is quite technical and can be found in [57]. We will give here a direct proof in a simpler case, where the process of the generic player follows the equation

$$\begin{cases} dX_s^{\alpha,\sigma} = \alpha_s ds + \sigma_s dB_s \\ X_t^{\alpha,\sigma} = x\,, \end{cases} \tag{4.2.6}$$

where the time dependent control variables α_s, σ_s live in the space of bounded controls $\mathcal{U} \times \mathcal{S}$, and $\mathcal{S}$ is bounded from below (and from above) by a positive constant.

The following lemma will come in handy in the proof of the main result :

Lemma 4.2.1. *Let $t, r \in [0,T]$, $x, y \in \mathbb{R}^d$. We consider the two processes*

$$\begin{cases} dX_s = \alpha_s ds + \sigma_s dB_s\,, \\ X_t = x \end{cases} \tag{4.2.7}$$

and

$$\begin{cases} dY_s = \beta_s ds + \eta_s dB_s\,, \\ Y_r = y\,, \end{cases} \tag{4.2.8}$$

with $X_s = x$ for $s < t$ and $Y_s = y$ for $s < r$, and where $\alpha., \beta. \in \mathcal{A}_t^{\mathcal{U}}$, $\sigma., \eta. \in \mathcal{A}_t^{\mathcal{S}}$ are bounded processes in $[0,T]$.

Let $h > 0$. Then there exists a constant C (not depending on h) such that

$$\mathbb{E}\left[\sup_{t \le s \le t+h} |X_s - Y_s|\right] \le C\left(|x-y| + \sqrt{|t-r|} + M\sqrt{h}\right)\,, \tag{4.2.9}$$

where C depends on $\|\alpha\|_\infty$, $\|\beta\|_\infty$, $\|\sigma\|_\infty$, $\|\eta\|_\infty$ and where $M := \|\alpha-\beta\|_\infty + \|\eta-\sigma\|_\infty$.

In particular, we have

$$\mathbb{E}\left[\sup_{t \le s \le t+h} |X_s - x|\right] \le C\sqrt{h}\,. \tag{4.2.10}$$

Démonstration. Without loss of generality, we suppose $t > r$.

We start noticing that (4.2.10) is a directly consequence of (4.2.9).

Actually, taking a process $(X_s)_s$ satisfying (4.2.7), the constant process $Y_s = x$ satisfies (4.2.8) with $r = t$, $y = x$, $\beta_s = 0$, $\eta_s = 0$. Then (4.2.9) tells us that

$$\mathbb{E}\left[\sup_{t \le s \le t+h} |X_s - x|\right] \le C\sqrt{h}\,.$$

Hence, we only have to prove (4.2.9).

Since $t > r$, we can write

$$X_s - Y_s = x + \int_t^s (\alpha_u - \beta_u)du + (\sigma_u - \eta_u)dB_u - y - \int_r^t \beta_u du + \eta_u dB_u \,.$$

Using the boundedness of α and β, we get for $s \in [t, t+h]$

$$|X_s - Y_s| \le |x - y| + Mh + C(t-r) + \left| \int_t^s (\sigma_u - \eta_u)dB_u \right| + \left| \int_r^t \eta_u dB_u \right| ,$$

and so

$$\begin{aligned} \mathbb{E}\left[\sup_{t \le s \le t+h} |X_s - x| \right] &\le |x-y| + C(t-r) + Ch \\ + \mathbb{E}\left[\sup_{s\in[t,t+h]} \left| \int_t^s (\sigma_u - \eta_u)dB_u \right| \right] &+ \mathbb{E}\left[\left| \int_r^t \eta_u dB_u \right| \right] . \end{aligned} \tag{4.2.11}$$

The last two terms in the right-hand side are estimated using (7.23) and (7.17) of [8] :

$$\begin{aligned} \mathbb{E}\left[\sup_{s\in[t,t+h]} \left| \int_t^s (\sigma_u - \eta_u)dB_u \right| \right] &\le \sqrt{\mathbb{E}\left[\sup_{s\in[t,t+h]} \left(\int_t^s (\sigma_u - \eta_u)dB_u \right)^2 \right]} \\ \le C\sqrt{\mathbb{E}\left[\int_t^{t+h} (\sigma_u - \eta_u)^2 du \right]} &\le CM\sqrt{h} \,. \end{aligned}$$

In an easier way we estimate the last term :

$$\mathbb{E}\left[\left| \int_r^t \eta_u \, dB_u \right| \right] \le \sqrt{\mathbb{E}\left[\left(\int_r^t \eta_u \, dB_u \right)^2 \right]} = \sqrt{\mathbb{E}\left[\int_r^t \eta_u^2 \, du \right]} \le C\sqrt{t-r} \,.$$

Plugging these estimates in (4.2.11), we obtain (4.2.9). □

Now we are able to prove the following theorem.

Theorem 4.2.2. *Suppose that L^1, L^2, F and G are Lipschitz continuous in the space variable and globally bounded. Furthermore, suppose that L^1 and L^2 are continuous in time and space, and that the function*

$$(t, x) \to F(t, x, m(t))$$

is continuous in time and space.

Then u is a viscosity solution of the Hamilton-Jacobi equation

$$\begin{cases} u_t(t,x) + H^1(t,x,\nabla u(t,x)) + H^2(t,x,u(t,x)) + F(t,x,m(t)) = 0\,, \\ u(T,x) = G(x,m(T))\,, \end{cases} \tag{4.2.12}$$

where

$$H^1(t,x,p) := \inf_{\alpha\in\mathcal{U}} \{\langle p,\alpha\rangle + L_1(t,x,\alpha)\}$$

and

$$H^2(t,x,q) := \inf_{\sigma\in\mathcal{S}} \left\{\frac{\sigma^2}{2}q + L_2(t,x,\sigma)\right\} \tag{4.2.13}$$

are the so called Hamiltonians.

We observe that the condition on F is satisfied, for instance, if

$$F : [0,T]\times\mathbb{R}^n\times\mathcal{P}(\Omega)\to\mathbb{R}\,,$$

where $\mathcal{P}(\Omega)$ is the space of Borel probability measures with finite first order moment, equipped with the Wasserstein distance $\mathbf{d}_1$ defined in (3.2.3), is continuous in all variables, and if

$$m : [0,T]\to\mathcal{P}(\Omega)$$

is continuous.

Démonstration. Step 1 : Continuity of u. The first step consists to check the continuity of the value function u in both variables.

Let $(t_n,x_n)\to(t,x)$. Using the definition of u, for each $\varepsilon>0$ we considers controls $\alpha_s^{\varepsilon,n}$ and $\sigma_s^{\varepsilon,n}$ such that

$$\mathbb{E}\left[\int_{t_n}^T (L_1(s,X_s^{\varepsilon,n},\alpha_s^{\varepsilon,n}) + L_2(s,X_s^{\varepsilon,n},\sigma_s^{\varepsilon,n}) + F(s,X_s^{\varepsilon,n},m(s)))\,ds + G(X_T^{\varepsilon,n},m(T))\right] \le u(t_n,x_n)+\varepsilon\,,$$

where $X_s^{\varepsilon,n}$ is the process related to the controls $\alpha_s^{\varepsilon,n}$ and $\sigma_s^{\varepsilon,n}$, with $X_s^{\varepsilon,n}=x_n$ for $s<t_n$.

We take the process $(X_s)_s$ defined $X_s=x$ for $s<t$ and satisfying for $s\ge t$

$$\begin{cases} dX_s = \alpha_s^{\varepsilon,n}ds + \sigma_s^{\varepsilon,n}dB_s\,, \\ X_t = x\,. \end{cases}$$

Then we have, using the hypotheses on the cost functions and the Lagrangians,

$$u(t,x) \leq \mathbb{E}\left[\int_t^T \left(L_1(s,X_s,\alpha_s^{\varepsilon,n}) + L_2(s,X_s,\sigma_s^{\varepsilon,n}) + F(s,X_s,m(s))\right) ds + G(X_T,m(T))\right]$$
$$\leq \mathbb{E}\left[\int_{t_n}^T \left(L_1(s,X_s,\alpha_s^{\varepsilon,n}) + L_2(s,X_s,\sigma_s^{\varepsilon,n}) + F(s,X_s,m(s))\right) ds + G(X_T,m(T))\right] + C|t_n - t|$$
$$\leq C\mathbb{E}\left[\sup_{t\in[0,T]} |X_s - X_s^{\varepsilon,n}|\right] + C|t_n-t| + u(t_n,x_n) + \varepsilon \leq C\left(|x_n - x| + \sqrt{|t_n-t|}\right) + u(t_n,x_n) + \varepsilon\,,$$

where we used in the last step (4.2.9) with $M = 0$.

So, passing to the lim inf when $(t_n, x_n) \to (t, x)$, we obtain

$$u(t,x) \leq \liminf_{(t_n,x_n)\to(t,x)} u(t_n,x_n) + \varepsilon\,,$$

which gives, for the arbitrariness of ε,

$$u(t,x) \leq \liminf_{(t_n,x_n)\to(t,x)} u(t_n,x_n)\,.$$

In the same way, we prove

$$u(t,x) \geq \limsup_{(t_n,x_n)\to(t,x)} u(t_n,x_n)\,,$$

which finally gives the continuity of u.

Step 2 : Subsolution argument. Let $\varphi \in C^\infty([0,T]\times\mathbb{R}^n)$, with all derivatives bounded, a test function with

$$\varphi(\bar{t},\bar{x}) \geq u(\bar{t},\bar{x}), \;\; \forall(\bar{t},\bar{x}) \in [0,T]\times\mathbb{R}^n$$

and a minimum value at the point (t, x), where

$$\varphi(t,x) = u(t,x).$$

The dynamical programming principle (4.2.5) gives

$$0 = \inf_{\substack{\alpha.\in\mathcal{A}_t^{\mathcal{U}}\\ \sigma.\in\mathcal{A}_t^{\mathcal{S}}}} \mathbb{E}\left[\int_t^{t+h} (L_1(s,X_s,\alpha_s) + L_2(s,X_s,\sigma_s) + F(s,X_s,m(s)))ds + u(t+h,X_{t+h}) - u(t,x)\right]$$
$$\leq \inf_{\substack{\alpha.\in\mathcal{A}_t^{\mathcal{U}}\\ \sigma.\in\mathcal{A}_t^{\mathcal{S}}}} \mathbb{E}\left[\int_t^{t+h} (L_1(s,X_s,\alpha_s) + L_2(s,X_s,\sigma_s) + F(s,X_s,m(s)))ds + \varphi(t+h,X_{t+h}) - \varphi(t,x)\right]. \tag{4.2.14}$$

We now use Ito's formula on φ and obtain, for any admissible control (α_s, σ_s),

$$\mathbb{E}[\varphi(t+h, X_{t+h}) - \varphi(t,x)] = \mathbb{E}\left[\int_t^{t+h} \left\{[\varphi_t(s,X_s) + \langle \nabla\varphi(s,X_s), \alpha_s\rangle + \frac{\sigma_s^2}{2}\Delta\varphi(s,X_s)\right\} ds\right]$$

This expression, combined with (4.2.14), gives us

$$\begin{aligned} 0 \leq \mathbb{E}&\left[\int_t^{t+h} \left\{[\varphi_t(s,X_s) + \langle \nabla\varphi(s,X_s), \alpha_s\rangle + \frac{\sigma_s^2}{2}\Delta\varphi(s,X_s)\right\} ds\right] \\ &+ \mathbb{E}\left[\int_t^{t+h} \{L_1(s,X_s,\alpha_s) + L_2(s,X_s,\sigma_s) + F(s,X_s,m(s))\}\, ds\right], \end{aligned}$$

for each control $\alpha.$ and $\sigma.$.

Now it is necessary to write down our PDE as a purely deterministic expression. For this, we consider for now only a subset of the control spaces $\mathcal{A}_t^{\mathcal{U}}$ and $\mathcal{A}_t^{\mathcal{S}}$, namely the set of constant controls. This controls are denoted by α and σ instead of α_s and σ_s, taking values in $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{S} \subset \mathbb{R}$. Our inequality now reads

$$\begin{aligned} 0 \leq \mathbb{E}&\left[\int_t^{t+h} \left\{[\varphi_t(s,X_s) + \langle \nabla\varphi(s,X_s), \alpha\rangle + \frac{\sigma^2}{2}\Delta\varphi(s,X_s)\right\} ds\right] \\ &+ \mathbb{E}\left[\int_t^{t+h} \{L_1(s,X_s,\alpha) + L_2(s,X_s,\sigma) + F(s,X_s,m(s))\}\, ds\right], \end{aligned}$$

Now we divide by h and we let $h \to 0$. Using the *a.s.* time continuity of the trajectories $X.(\omega)$ and the continuity of the cost functions and the Lagrangians, we obtain

$$0 \leq \phi_t(t,x) + \langle \nabla\phi(t,x), \alpha\rangle + \frac{\sigma^2}{2}\Delta\phi(t,x) + L_1(t,x,\alpha) + L_2(t,x,\sigma) + F(t,x,m(t))\,.$$

Passing to the inf when $\alpha \in \mathcal{U}$ and $\sigma \in \mathcal{S}$, we get

$$0 \geq -\varphi_t(t,x) - H^1(t,x,\nabla\varphi(t,x)) - H^2(t,x,\Delta\varphi(t,x)) - F(t,x,m(t,x)) \qquad (4.2.15)$$

for all test functions φ, which proves that $u(t,x)$ is indeed a viscosity sub solution for our PDE.

Step 3 : Supersolution argument. Now we take $\varphi \in C_b^\infty([0,T]\times\mathbb{R}^n)$, a test function with

$$\varphi(\bar{t},\bar{x}) \leq u(\bar{t},\bar{x}),\ \ \forall(\bar{t},\bar{x}) \in [0,T]\times\mathbb{R}^n$$

and a maximum value at the point (t, x), where

$$\varphi(t, x) = u(t, x).$$

As before, the dynamical programming principle (4.2.5) gives

$$0 \geq \inf_{\substack{\alpha_\cdot \in \mathcal{A}_t^{\mathcal{U}} \\ \sigma_\cdot \in \mathcal{A}_t^{\mathcal{S}}}} \mathbb{E}\left[\int_t^{t+h} (L_1(s, X_s, \alpha_s) + L_2(s, X_s, \sigma_s) + F(s, X_s, m(s)))ds + \varphi(t+h, X_{t+h}) - \varphi(t, x)\right].$$

Using the definition of ϕ, we take controls α_s^h, σ_s^h and the related process X_s^h such that

$$\mathbb{E}\left[\int_t^{t+h} (L_1(s, X_s^h, \sigma_s^h) + L_2(s, X_s^h, \alpha_s^h) + F(s, X_s^h, m(s)))ds + \varphi(t+h, X_{t+h}^h) - \varphi(t, x)\right] \leq h^2 .$$

Applying Ito's formula we obtain

$$\begin{aligned} &\mathbb{E}\left[\int_t^{t+h} \left\{\varphi_t(s, X_s^h) + \langle \nabla\varphi(s, X_s^h), \alpha_s^h\rangle + \frac{(\sigma_s^h)^2}{2}\Delta\varphi(s, X_s^h)\right\} ds\right] \\ &+ \mathbb{E}\left[\int_t^{t+h} \left\{L_1(s, X_s^h, \alpha_s^h) + L_2(s, X_s^h, \sigma_s^h) + F(s, X_s^h, m(s))\right\} ds\right] \leq h^2 . \end{aligned}$$

We estimate all the terms in the same way, using the Lipschitz continuity of the cost functions with respect to x and Lemma 4.2.1. For instance, for the L^1 term we get

$$\begin{aligned} &\mathbb{E}\left[\int_t^{t+h} L_1(s, X_s^h, \alpha_s^h)\, ds\right] \geq \mathbb{E}\left[\int_t^{t+h} \Big(L_1(s, x, \alpha_s^h) - |X_s^h - x|\Big)\, ds\right] \\ &\geq \mathbb{E}\left[\int_t^{t+h} L_1(s, x, \alpha_s^h)\, ds\right] - h\mathbb{E}\left[\sup_{s\in[t,t+h]} |X_s^h - x|\right] \geq \mathbb{E}\left[\int_t^{t+h} L_1(s, x, \alpha_s^h)\, ds\right] - Ch\sqrt{h}\, . \end{aligned}$$

Hence, with all these estimates we get

$$\begin{aligned} &\mathbb{E}\left[\int_t^{t+h} \left\{\varphi_t(s, x) + \langle \nabla\varphi(s, x), \alpha_s^h\rangle + \frac{(\sigma_s^h)^2}{2}\Delta\varphi(s, x)\right\} ds\right] \\ &+ \mathbb{E}\left[\int_t^{t+h} \left\{L_1(s, x, \alpha_s^h) + L_2(s, x, \sigma_s^h) + F(s, x, m(s))\right\} ds\right] \leq h^2 + Ch\sqrt{h} \leq Ch\sqrt{h}\, . \end{aligned}$$

Using the definition of H^1 and H^2, we obtain

$$\mathbb{E}\left[\int_t^{t+h} \left\{\varphi_t(s, x) + H^1(s, x, \nabla\phi(s, x)) + H^2(s, x, \Delta\phi(s, x)) + F(s, x, m(s))\right\} ds\right] \leq Ch\sqrt{h}\, .$$

Now we divide by h and we use the continuity of H^1 and H^2 in time variable, which is an immediate consequence of the time continuity of L^1 and L^2. We finally obtain

$$-\varphi_t(t,x) - H^1(t,x,\nabla\varphi(t,x)) - H^2(t,x,\Delta\varphi(t,x)) - F(t,x,m(t,x)) \geq 0\,, \tag{4.2.16}$$

which proves that u is a supersolution of the HJB equation and concludes the Theorem. $\square$

Remark 4.2.3. We observe that, up to defining the set $\mathcal{S}' := \left\{\frac{\sigma^2}{2} \mid \sigma \in \mathcal{S}\right\}$, the function H^2 can be rewritten as

$$H^2(t,x,q) := \inf_{\eta\in\mathcal{S}'} \{\eta q + L_3(t,x,\eta)\}\ ,$$

where $L_3(t,x,\eta) = L_2(t,x,\sqrt{2\eta})$.

Henceforth, when we will talk about this simplified model case, we will refer to

$$\begin{aligned} H^1(t,x,p) &:= \inf_{\alpha\in\mathcal{U}} \{\langle p,\alpha\rangle + L_1(t,x,\alpha)\}\ , \\ H^2(t,x,q) &:= \inf_{\eta\in\mathcal{S}'} \{\eta q + L_3(t,x,\eta)\}\ . \end{aligned} \tag{4.2.17}$$

4.3 The Hamilton-Jacobi Equation

This Section is completely devoted to the study of the following equation :

$$\begin{cases} u_t + H^2(t,x,\Delta u) + H^1(t,x,\nabla u) = -F(t,x)\,, & (t,x)\in[0,T]\times R^N\,, \\ u(T,x) = G(x)\,, \end{cases}$$

where

$$H^1(t,x,p) := \inf_{\alpha\in\mathcal{U}} \{\langle p, b(x,\alpha)\rangle + L_1(t,x,\alpha)\}$$

and

$$H^2(t,x,q) := \inf_{\sigma\in\mathcal{S}} \left\{\frac{\Gamma(x,\sigma)^2}{2} q + L_2(t,x,\sigma)\right\}$$

Since there is no dependence on m in this step, we can omit F in (4.3.1), including it into H^1 (up to changing the Lagrangian L^1). Hence, the equation we want to study is the following

$$\begin{cases} u_t + H^2(t,x,\Delta u) + H^1(t,x,\nabla u) = 0\,, & (t,x)\in[0,T]\times R^N\,, \\ u(T,x) = G(x)\,, \end{cases} \tag{4.3.1}$$

We note that, with the change of variable $v(t,x) = e^{-\lambda(T-t)}u(t,x)$, the system (4.3.1) is equivalent to the following one :

$$\begin{cases} -u_t - H^2_\lambda(t,x,\Delta u) - H^1_\lambda(t,x,\nabla u) + \lambda u = 0\,, \\ u(T,x) = G(x)\,, \end{cases} \tag{4.3.2}$$

where

$$H^2_\lambda(t,x,q) = e^{-\lambda(T-t)}H^2(t,x,e^{\lambda(T-t)}q) = \inf_{\sigma\in\mathcal{S}}\left\{\frac{\Gamma(x,\sigma)^2}{2}q + e^{-\lambda(T-t)}L_2(t,x,\sigma)\right\}\,,$$
$$H^1_\lambda(t,x,p) = e^{-\lambda(T-t)}H^1(t,x,e^{\lambda(T-t)}p) = \inf_{\alpha\in\mathcal{U}}\left\{\langle p, b(x,\alpha)\rangle + e^{-\lambda(T-t)}L_1(t,x,\alpha)\right\}\,.$$

for each $\lambda > 0$.

So, in order to prove existence, uniqueness and regularity of solutions, we will work with (4.3.2). For simplicity, we will write H^1 and H^2 instead of H^1_λ and H^2_λ, up to changing again the Lagrangians L^1 and L^2.

Due to the non-linearity of the second-order term, we need to work in viscosity sense.
First, for each $u : (0,T)\times\mathbb{R}^n \to \mathbb{R}$ we define the following sets :

Definition 4.3.1. *We denote with* $\mathcal{P}^{2,+}u(t_0,x_0)$ *the set of all points* $(a,p,X) \in \mathbb{R}\times\mathbb{R}^N\times Sym(n,n)$ *such that for* $(t,x)\to(t_0,x_0)$ *we have*

$$u(t,x) \le u(t_0,x_0) + a(t-t_0) + \langle p, x-x_0\rangle + \frac{1}{2}\langle X(x-x_0), x-x_0\rangle + o(|t-t_0| + |x-x_0|^2)\,.$$

Similarly, we define $\mathcal{P}^{2,-}u(t_0,x_0) = -\mathcal{P}^{2,+}(-u)(t_0,x_0)$*, i.e. the set of all points* $(a,p,X) \in \mathbb{R}\times\mathbb{R}^N\times Sym(n,n)$ *such that for* $(t,x)\to(t_0,x_0)$ *we have*

$$u(t,x) \ge u(t_0,x_0) + a(t-t_0) + \langle p, x-x_0\rangle + \frac{1}{2}\langle X(x-x_0), x-x_0\rangle + o(|t-t_0| + |x-x_0|^2)\,.$$

Now we give a suitable definition of solution.

Definition 4.3.2. *We say that a function* $u \in USC((0,T]\times\mathbb{R}^n)$ *and bounded from above is a subsolution of* (4.3.2) *if* $\forall(t,x)\in(0,T)\times\mathbb{R}^N$ *and* $\forall(a,p,X)\in\mathcal{P}^{2,+}u(t,x)$ *we have*

$$-a - H^2(t,x,\mathrm{tr}(X)) - H^1(t,x,p) + \lambda u(t,x) \le 0\,;$$

moreover, $\forall x \in \mathbb{R}^n$ *we require* $u(T,x) \leq G(x)$*.*

Similarly, we say that a function $u \in LSC((0,T] \times \mathbb{R}^n)$ *and bounded from below is a supersolution of* (4.3.2) *if* $\forall (t,x) \in (0,T) \times \mathbb{R}^N$ *and* $\forall (a,p,X) \in \mathcal{P}^{2,-}u(t,x)$ *we have*

$$-a - H^2(t,x,\mathrm{tr}(X)) - H^1(t,x,p) + \lambda u(t,x) \geq 0\,;$$

moreover, $\forall x \in \mathbb{R}^n$ *we require* $u(T,x) \geq G(x)$*.*

Finally, we say that a bounded continuous function u is a solution of (4.3.2) *if it is both a subsolution and supersolution.*

The new term λu and the boundedness hypotheses play an essential role in order to prove the comparison principle.

In order to prove it, we also need the following proposition (Theorem 8.3 of [45]) , which will be also useful in the rest of the section.

Proposition 4.3.3. *Let* $u_1, \ldots, u_N$ *subsolutions of* (4.3.2)*.*
Consider a function $\phi : (0,T) \times \mathbb{R}^{Nn} \to \mathbb{R}$*, once continuously differentiable in* t *and twice continuously differentiable in* $(x_1, \ldots, x_N)$*.*
Suppose that

$$u_1(t,x_1) + \cdots + u_N(t,x_N) - \phi(t,x_1,\ldots,x_N)$$

achieves his maximum in $(\tilde{t}, \tilde{x}_1, \ldots, \tilde{x}_N)$*.*
Then there exist $a_i \in \mathbb{R}$*,* $X_i \in Sym(n,n)$*,* $i = 1, \cdots, N$ *such that, defining* $A = (D^2_x\phi)(\tilde{t}, \tilde{x}_1, \ldots, \tilde{x}_N)$*, one has*

$$(a_i, \nabla_{x_i}\phi(\tilde{t}, \tilde{x}_1, \ldots, \tilde{x}_N), X_i) \in \overline{\mathcal{P}^{2,+}u_i(\tilde{t}, \tilde{x}_i)}\,;$$

$$\begin{pmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_N \end{pmatrix} \leq A\,;$$

$$\sum_i a_i = \phi_t(\tilde{t}, \tilde{x}_1, \ldots, \tilde{x}_N)\,.$$

Now we are ready to prove the comparison principle, by adapting the proof given by Crandall, Ishii and Lions in [44].

Theorem 4.3.4. *Suppose* (4.2.2) *is satisfied. Moreover, suppose that*

— L^1 *continuous in all variables and Hölder in* x*, uniformly in* $t \in [0,T]$ *and locally uniformly in* p*, i.e.* $\forall |p| \le L\ \exists K_L$ *such that*

$$|L^1(t,x,p) - L^1(t,y,p)| \le K_L|x-y|^\beta\,, \qquad \textit{for a certain } 0 < \beta \le 1\,, \quad \forall t \in [0,T]\,.$$

The same hypotheses hold for L^2*.*

— Γ*,* b *and* G *are Lipschitz in the space variable.*

Then the comparison principle holds for equation (4.3.2)*, i.e., if* u *and* v *are respectively a subsolution and a supersolution of* (4.3.2)*, then* $u(t,x) \le v(t,x)$ *for all* $(t,x) \in (0,T] \times \mathbb{R}^n$*.*

Démonstration. Suppose, by contradiction, that $\exists (s,z) \in (0,T] \times \mathbb{R}^n$ such that $u(s,z) - v(s,z) = \delta > 0$.

We consider, for $\alpha, \nu > 0$, the following quantity

$$u(t,x) - v(t,y) - \frac{\alpha}{2}|x-y|^2 - \frac{1}{\alpha}|x|^2 - \frac{\nu}{t}\,. \tag{4.3.3}$$

Due to the coercive term $\frac{1}{\alpha}|x|^2$ and the boundedness of u of v, we know that (4.3.3) achieves a maximum.

We denote the maximum by M and one of its maximum points with $(\bar t, \bar x, \bar y) \in (0,T] \times \mathbb{R}^{2n}$.

We must have, for ν sufficiently small and α sufficiently large,

$$u(\bar t,\bar x) - v(\bar t,\bar y) - \frac{\alpha}{2}|\bar x-\bar y|^2 - \frac{1}{\alpha}|\bar x|^2 - \frac{\nu}{\bar t} \ge u(s,z) - v(s,z) - \frac{1}{\alpha}|z|^2 - \frac{\nu}{s} \ge \delta - \frac{\delta}{2} = \frac{\delta}{2}\,. \tag{4.3.4}$$

This implies, thanks to the boundedness of u and v,

$$\frac{1}{\alpha}|\bar x|^2 + \frac{\alpha}{2}|\bar x-\bar y|^2 \le u(\bar t,\bar x) - v(\bar t,\bar y) \le C \implies \lim_{\alpha\to\infty} |\bar x - \bar y| = 0\,.$$

This cannot happen if $\bar t = T$. In this case we have

$$\frac{\delta}{2} \le u(s,z) - v(s,z) - \frac{1}{\alpha}|z|^2 - \frac{\nu}{s} \le u_T(\bar x) - v_T(\bar y) - \frac{\alpha}{2}|\bar x-\bar y|^2 - \frac{1}{\alpha}|\bar x|^2\,.$$

Since G is Lipschitz, $u_T \le v_T$ and $|\bar x - \bar y| \to 0$, we have for α sufficiently large

$$u_T(\bar x) - v_T(\bar y) \le \frac{\delta}{3} \implies u_T(\bar x) - v_T(\bar y) - \frac{\alpha}{2}|\bar x-\bar y|^2 - \frac{1}{\alpha}|\bar x|^2 \le \frac{\delta}{3}\,,$$

which gives a contradiction.

Then we must have $\bar t \in (0,T)$. Applying Proposition 4.3.3 with

$$u_1(t,x) = u(t,x)\,, \quad u_2(t,y) = -v(t,y)\,,$$
$$\phi(t,x,y) = \frac{\alpha}{2}|x-y|^2 + \frac{1}{\alpha}|x|^2 + \frac{\nu}{t}\,,$$

we obtain that there exist a, b, X, Y such that

$$(a, \nabla_x\phi(\bar{t}, \bar{x}, \bar{y}), X) \in \overline{\mathcal{P}^{2,+}u(\bar{t}, \bar{x})}\,, \qquad (-b, -\nabla_y\phi(\bar{t}, \bar{x}, \bar{y}), -Y) \in \overline{\mathcal{P}^{2,-}v(\bar{t}, \bar{y})}\,,$$

and

$$a + b = -\frac{\nu}{\bar{t}^2}\,, \qquad \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq D^2\phi(\bar{t}, \bar{x}, \bar{y})\,. \tag{4.3.5}$$

From now on, we will omit for the function ϕ his dependence on $(\bar{t}, \bar{x}, \bar{y})$.

Since u is a subsolution and v a supersolution, one has

$$\begin{aligned} -a - H^2(\bar{t}, \bar{x}, \mathrm{tr}(X)) - H^1(\bar{t}, \bar{x}, \nabla_x\phi) + \lambda u(\bar{t}, \bar{x}) \leq 0\,, \\ b - H^2(\bar{t}, \bar{y}, -\mathrm{tr}(Y)) - H^1(\bar{t}, \bar{y}, -\nabla_y\phi) + \lambda v(\bar{t}, \bar{y}) \geq 0\,. \end{aligned}$$

Subtracting the two inequalities we obtain

$$\frac{\nu}{\bar{t}^2} + \lambda(u(\bar{t}, \bar{x}) - v(\bar{t}, \bar{y})) \leq H^2(\bar{t}, \bar{x}, \mathrm{tr}(X)) - H^2(\bar{t}, \bar{y}, -\mathrm{tr}(Y)) + \tag{4.3.6}$$

$$+ H^1(\bar{t}, \bar{x}, \nabla_x\phi) - H^1(\bar{t}, \bar{y}, -\nabla_y\phi)\,. \tag{4.3.7}$$

The first term in the left-hand side is non-negative, so we can ignore it. For the second term, we use (4.3.4) to get

$$\begin{aligned} \frac{\delta}{2} + \lambda\frac{\alpha}{2}|\bar{x} - \bar{y}|^2 \leq\, & H^2(\bar{t}, \bar{x}, \mathrm{tr}(X)) - H^2(\bar{t}, \bar{y}, -\mathrm{tr}(Y)) + \\ & + H^1(\bar{t}, \bar{x}, \nabla_x\phi) - H^1(\bar{t}, \bar{y}, -\nabla_y\phi)\,. \end{aligned}$$

In order to estimate the last two terms of the right-hand side term, we first compute the derivatives of ϕ. We have

$$\begin{aligned} &\nabla_x\phi = \alpha(\bar{x} - \bar{y}) + \frac{2}{\alpha}\bar{x}\,, \qquad \nabla_y\phi = -\alpha(\bar{x} - \bar{y})\,, \\ &D^2\phi = \alpha\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \frac{2}{\alpha}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\,. \end{aligned}$$

Using the very definition of H^2, we obtain, calling σ the optimal control for $H^2(\bar{t}, \bar{y}, -\mathrm{tr}(Y))$,

$$\begin{aligned} & H^2(\bar{t}, \bar{x}, \mathrm{tr}(X)) - H^2(\bar{t}, \bar{y}, -\mathrm{tr}(Y)) \\ \leq\, & \frac{\Gamma(\bar{x}, \sigma)^2}{2}\mathrm{tr}(X) + \frac{\Gamma(\bar{y}, \sigma)^2}{2}\mathrm{tr}(Y) + e^{-\lambda(T-t))}\left(L^2(\bar{t}, \bar{x}, \sigma) - L^2(\bar{t}, \bar{y}, \sigma)\right) \\ \leq\, & C|\bar{x} - \bar{y}|^\beta + \mathrm{tr}\left(\frac{\Gamma(\bar{x}, \sigma)^2}{2}X + \frac{\Gamma(\bar{y}, \sigma)^2}{2}Y\right) \leq \omega(\alpha) + \mathrm{tr}\left(\frac{\Gamma(\bar{x}, \sigma)^2}{2}X + \frac{\Gamma(\bar{y}, \sigma)^2}{2}Y\right)\,. \end{aligned}$$

The last term is estimated as follows.

We have

$$\mathrm{tr}\left(\frac{\Gamma(\bar{x},\sigma)^2}{2}X+\frac{\Gamma(\bar{y},\sigma)^2}{2}Y\right)=\frac{1}{2}\mathrm{tr}\left(B\begin{pmatrix}X & 0\\ 0 & Y\end{pmatrix}\right),$$

where B is equal to

$$\begin{pmatrix}\Gamma(\bar{x},\sigma)^2 & \Gamma(\bar{x},\sigma)\Gamma(\bar{y},\sigma)\\ \Gamma(\bar{x},\sigma)\Gamma(\bar{y},\sigma) & \Gamma(\bar{y},\sigma)^2\end{pmatrix}.$$

Since B is non-negative definite, we can use (4.3.5) to obtain

$$\begin{aligned}\mathrm{tr}\left(\frac{\Gamma(\bar{x},\sigma)^2}{2}X+\frac{\Gamma(\bar{y},\sigma)^2}{2}Y\right)&\le \mathrm{tr}\left(BD^2\phi\right)=\\ =\alpha n|\Gamma(\bar{x},\sigma)-\Gamma(\bar{y},\sigma)|^2+\frac{2n}{\alpha}\Gamma(\bar{x},\sigma)^2&\le C\alpha|\bar{x}-\bar{y}|^2+\omega(\alpha),\end{aligned}$$

where $\omega(\alpha)$ is a quantity depending on α such that $\omega(\alpha)\to 0$ when $\alpha\to\infty$.

We argue in a similar way in order to bound the H^1 term. We have, calling α_0 the optimal control for $H^1(\bar{t},\bar{y},-\nabla_y\phi)$,

$$\begin{aligned}&H^1(\bar{t},\bar{x},\nabla_x\phi)-H^1(\bar{t},\bar{y},-\nabla_y\phi)\\ &\le\langle b(\bar{x},\alpha_0),\nabla_x\phi\rangle+\langle b(\bar{y},\alpha_0),\nabla_y\phi\rangle+e^{-\lambda(T-t)}\left(L^1(\bar{t},\bar{x},\alpha_0)-L^1(\bar{t},\bar{y},\alpha_0)\right)\\ &\le|\bar{x}-\bar{y}|^\beta+\langle b(\bar{x},\alpha_0)-b(\bar{y},\alpha_0),\alpha(\bar{x}-\bar{y})\rangle+\frac{2}{\alpha}\langle b(\bar{x},\alpha_0),\bar{x}\rangle\le C\alpha|\bar{x}-\bar{y}|^2+\omega(\alpha),\end{aligned}$$

where in the last passage we used the Lipschitz bound on b and the boundedness of b.

Putting together all the estimates in (4.3.6), we obtain

$$\frac{\delta}{2}+\lambda\frac{\alpha}{2}|\bar{x}-\bar{y}|^2\le C\alpha|\bar{x}-\bar{y}|^2+\omega(\alpha),$$

which gives a contradiction for α sufficiently small and λ sufficiently large, and proves the Theorem. □

The existence of solutions of (4.3.2) is a natural adaptation of the elliptic case, whose proof can be found in [45]. Following the ideas of that article, existence of solutions is guaranteed if there exists a continuous subsolution $\underline{u}$ and a continuous supersolution $\overline{u}$ such that (in the case of Dirichlet conditions)

$$\underline{u}(T,x)=\overline{u}(T,x)=G(x),\tag{4.3.8}$$

The solution turns to be the following one :

$$u(t,x) = \sup\left\{ w(t,x) : \underline{u} \le w \le \overline{u} \text{ and } w \text{ is a subsolution of (4.3.2)} \right\}.$$

We look for a function $\underline{u}$ of this type :

$$\underline{u}(t,x) = G(x) + \xi(t), \qquad \text{with } \xi(T) = 0\,.$$

Obviously (4.3.8) is satisfied. In order to have $\underline{u}$ subsolution of (4.3.2), it has to be

$$-\xi'(t) + \lambda\xi(t) \le H^2(t,x,\Delta G(x)) + H^1(t,x,\nabla G(x)) - \lambda G(x)\,, \qquad (4.3.9)$$

for each $(t,x) \in (0,T) \times \mathbb{R}^n$. If we require $G \in \mathcal{C}^2(\mathbb{R}^n)$ with bounded derivatives, we obtain that the right-hand side of (4.3.9) is bounded from below. So it is sufficient to take ξ as the solution of

$$\begin{cases} -\xi'(t) + \lambda\xi(t) = -M\,, \\ \xi(T) = 0\,, \end{cases}$$

with M sufficiently large.

Reasoning in the same way, a good supersolution $\overline{u}$ will be :

$$\overline{u}(t,x) = G(x) + \xi(t)\,,$$

with ξ solution, for M sufficiently large, of the *ODE*

$$\begin{cases} -\xi'(t) + \lambda\xi(t) = M\,, \\ \xi(T) = 0\,. \end{cases}$$

As regards uniqueness, it is an obvious consequence of the comparison principle.

So, we have proved the following

Theorem 4.3.5. *Suppose hypotheses of Theorem 4.3.4 and* (4.2.2) *are satisfied, and suppose that* $G \in \mathcal{C}^2(\mathbb{R}^n)$ *with bounded derivatives.*
Then there exists a unique viscosity solution for equation (4.3.2)*, and, consequently,* (4.3.1)*.*

Now, we prove some regularity results for the solution u. These results will be essential in order to work with the Fokker-Planck equation.

Theorem 4.3.6. *Suppose* (4.2.2) *hold true and hypotheses of theorem 4.3.4 with* $\beta = 1$ *are satisfied. Moreover, suppose that* G *is a globally Lipschitz function.*

Then, every solution of (4.3.2) *(and consequently of* (4.3.1)*) is globally Lipschitz in the space variable, with a Lipschitz constant bounded uniformly in* t.

Démonstration. We have to prove that

$$|u(t,x) - u(t,y)| \leq L|x-y|\,,$$

for a certain constant L. Using Young's inequality, the last inequality is equivalent

$$|u(t,x) - u(t,y)| \leq M\left(\delta + \frac{|x-y|^2}{\delta}\right)\,, \qquad \forall \delta > 0\,, \tag{4.3.10}$$

for a certain constant $M > 0$. So, we will prove (4.3.10).
We consider, for $\delta, \gamma, \nu > 0$, the following quantity

$$u(t,x) - u(t,y) - M\left(\delta + \frac{|x-y|^2}{\delta}\right) - \gamma|x|^2 - \frac{\nu}{t}\,. \tag{4.3.11}$$

Due to the coercive term $\gamma|x|^2$ and the boundedness of u, we know that the function in (4.3.11) achieves a maximum.
We denote one of its maximum points with $(\bar{t}, \bar{x}, \bar{y}) \in (0,T] \times \mathbb{R}^{2n}$.
Suppose $\bar{t} = T$. Then

$$u(t,x) - u(t,y) - M\left(\delta + \frac{|x-y|^2}{\delta}\right) - \gamma|x|^2 - \frac{\nu}{t} \leq G(\bar{x}) - G(\bar{y}) - M\left(\delta + \frac{|\bar{x}-\bar{y}|^2}{\delta}\right) - \gamma|\bar{x}|^2 - \frac{\nu}{T}\,,$$

which implies for M sufficiently large, since G is Lipschitz,

$$u(t,x) - u(t,y) \leq M\left(\delta + \frac{|x-y|^2}{\delta}\right) + \gamma|x|^2 + \frac{\nu}{t}\,. \tag{4.3.12}$$

Suppose now $\bar{t} \in (0,T)$. If the maximum achieved by the function in (4.3.11) is non-positive, then (4.3.12) remains true.

Now, suppose by contradiction that (4.3.11) attains a strictly positive maximum. This implies

$$u(\bar{t}, \bar{x}) - u(\bar{t}, \bar{y}) \geq M\left(\delta + \frac{|\bar{x}-\bar{y}|^2}{\delta}\right) \tag{4.3.13}$$

and

$$\gamma|\bar{x}|^2 \leq u(\bar{t}, \bar{x}) - u(\bar{t}, \bar{y}) \leq C\,, \tag{4.3.14}$$

for a certain $C > 0$.

Applying Proposition 4.3.3 with

$$u_1(t,x) = u(t,x)\,, \quad u_2(t,y) = -u(t,y)\,,$$
$$\phi(t,x,y) = M\left(\delta + \frac{|x-y|^2}{\delta}\right) + \gamma|x|^2 + \frac{\nu}{t}\,,$$

we obtain that there exist a, b, X, Y such that

$$(a, \nabla_x\phi(\bar t,\bar x,\bar y), X) \in \overline{\mathcal{P}^{2,+}u(\bar t,\bar x)}\,, \qquad (-b, -\nabla_y\phi(\bar t,\bar x,\bar y), -Y) \in \overline{\mathcal{P}^{2,-}u(\bar t,\bar y)}\,,$$

and

$$a + b = -\frac{\nu}{\bar t^2}\,, \qquad \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \le D^2\phi(\bar t,\bar x,\bar y)\,. \tag{4.3.15}$$

From now on, we will omit for the function ϕ his dependence on $(\bar t,\bar x,\bar y)$.

Since u is both a subsolution and a supersolution, one has

$$-a - H^2(\bar t,\bar x,\mathrm{tr}(X)) - H^1(\bar t,\bar x,\nabla_x\phi) + \lambda u(\bar t,\bar x) \le 0\,,$$
$$b - H^2(\bar t,\bar y,-\mathrm{tr}(Y)) - H^1(\bar t,\bar y,-\nabla_y\phi) + \lambda u(\bar t,\bar y) \ge 0\,.$$

Subtracting the two inequalities we obtain

$$\frac{\nu}{\bar t^2} + \lambda(u(\bar t,\bar x) - u(\bar t,\bar y)) \le H^2(\bar t,\bar x,\mathrm{tr}(X)) - H^2(\bar t,\bar y,-\mathrm{tr}(Y)) + \\ + H^1(\bar t,\bar x,\nabla_x\phi) - H^1(\bar t,\bar y,-\nabla_y\phi)\,.$$

The first term in the left-hand side is non-negative, so we can ignore it. For the second term, we use (4.3.13) to get

$$\lambda M\left(\delta + \frac{|\bar x-\bar y|^2}{\delta}\right) \le H^2(\bar t,\bar x,\mathrm{tr}(X)) - H^2(\bar t,\bar y,-\mathrm{tr}(Y)) + \\ + H^1(\bar t,\bar x,\nabla_x\phi) - H^1(\bar t,\bar y,-\nabla_y\phi)\,.$$

In order to estimate the last two terms of the right-hand side term, we first compute the derivatives of ϕ. We have

$$\nabla_x\phi = \frac{2M}{\delta}(\bar x-\bar y) + 2\gamma\bar x\,, \qquad \nabla_y\phi = -\frac{2M}{\delta}(\bar x-\bar y)\,,$$
$$D^2\phi = \frac{2M}{\delta}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + 2\gamma\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\,.$$

Using the very definition of H^2, we obtain, calling σ the optimal control for $H^2(\bar{t}, \bar{y}, -\mathrm{tr}(Y))$,

$$
\begin{aligned}
&H^2(\bar{t}, \bar{x}, \mathrm{tr}(X)) - H^2(\bar{t}, \bar{y}, -\mathrm{tr}(Y)) \\
\leq &\frac{\Gamma(\bar{x}, \sigma)^2}{2}\mathrm{tr}(X) + \frac{\Gamma(\bar{y}, \sigma)^2}{2}\mathrm{tr}(Y) + e^{-\lambda(T-t)}\left(L^2(\bar{t}, \bar{x}, \sigma) - L^2(\bar{t}, \bar{y}, \sigma)\right) \\
\leq &C|\bar{x} - \bar{y}| + \mathrm{tr}\left(\frac{\Gamma(\bar{x}, \sigma)^2}{2}X + \frac{\Gamma(\bar{y}, \sigma)^2}{2}Y\right).
\end{aligned}
$$

The last term is estimated as in the comparison Theorem.

We have

$$
\mathrm{tr}\left(\frac{\Gamma(\bar{x}, \sigma)^2}{2}X + \frac{\Gamma(\bar{y}, \sigma)^2}{2}Y\right) = \frac{1}{2}\mathrm{tr}\left(B\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}\right),
$$

where B is equal to

$$
\begin{pmatrix} \Gamma(\bar{x}, \sigma)^2 & \Gamma(\bar{x}, \sigma)\Gamma(\bar{y}, \sigma) \\ \Gamma(\bar{x}, \sigma)\Gamma(\bar{y}, \sigma) & \Gamma(\bar{y}, \sigma)^2 \end{pmatrix}.
$$

Since B is non-negative definite, we can use (4.3.15) to obtain

$$
\begin{aligned}
&\mathrm{tr}\left(\frac{\Gamma(\bar{x}, \sigma)^2}{2}X + \frac{\Gamma(\bar{y}, \sigma)^2}{2}Y\right) \leq \mathrm{tr}\left(BD^2\phi\right) = \frac{2M}{\delta}n|\Gamma(\bar{x}, \sigma) - \Gamma(\bar{y}, \sigma)|^2 + 2\gamma n\Gamma(\bar{x}, \sigma)^2 \\
&\leq C\frac{2M}{\delta}n|\bar{x} - \bar{y}|^2 + 2\gamma n\Gamma(\bar{x}, \sigma)^2 \leq CM\left(\delta + \frac{|\bar{x} - \bar{y}|^2}{\delta}\right) + C\gamma.
\end{aligned}
$$

We argue in a similar way in order to bound the H^1 term. We have, calling α the optimal control for $H^1(\bar{t}, \bar{y}, -\nabla_y\phi)$,

$$
\begin{aligned}
&H^1(\bar{t}, \bar{x}, \nabla_x\phi) - H^1(\bar{t}, \bar{y}, -\nabla_y\phi) \\
\leq &\langle b(\bar{x}, \alpha), \nabla_x\phi\rangle + \langle b(\bar{y}, \alpha), \nabla_y\phi\rangle + e^{-\lambda(T-t)}\left(L^1(\bar{t}, \bar{x}, \alpha) - L^1(\bar{t}, \bar{y}, \alpha)\right) \\
\leq &C|\bar{x} - \bar{y}| + \frac{2M}{\delta}\langle b(\bar{x}, \alpha) - b(\bar{y}, \alpha), (\bar{x} - \bar{y})\rangle + \frac{2}{\gamma}\alpha\langle b(\bar{x}, \alpha), \bar{x}\rangle \leq CM\left(\delta + \frac{|\bar{x} - \bar{y}|^2}{\delta}\right) + C\gamma|\bar{x}|,
\end{aligned}
$$

where in the last passage we used the Lipschitz bound on b and the boundedness of b.

Because of (4.3.14), we have $\gamma|\bar{x}| \to 0$ when $\gamma \to 0$.

Putting together all the estimates, we obtain, for M large enough,

$$
\lambda M\left(\delta + \frac{|\bar{x} - \bar{y}|^2}{\delta}\right) \leq CM\left(\delta + \frac{|\bar{x} - \bar{y}|^2}{\delta}\right) + \omega(\gamma),
$$

where $\omega(\gamma)$ is a quantity that tends to 0 when $\gamma \to 0$.

So, for γ small enough and $\lambda > 1$, we obtain a contradiction.

So in each case (4.3.12) remains true and, letting γ and ν go to 0, we get

$$u(t,x) - u(t,y) \leq M\left(\delta + \frac{|x-y|^2}{\delta}\right)$$

for all $\delta > 0$. Finally, choosing $\delta = |x-y|$ and reversing the role of x and y, we conclude :

$$|u(t,x) - u(t,y)| \leq 2M|x-y|\,.$$

□

Our next goal is to show that, with a strengthening of hypotheses, u satisfies a stronger estimate, which includes, as a direct consequence, a twice differentiability almost everywhere with respect to the x variable.

For this reason, we introduce the notion of *semiconcavity* :

Definition 4.3.7. *We say that a function $f : \Omega \to \mathbb{R}$ is semiconcave if for all $x \in \Omega$ and for all h such that $x+h$, $x-h \in \Omega$ we have*

$$f(x+h) + f(x-h) - 2f(x) \leq C|h|^2 \tag{4.3.16}$$

for a certain constant C.

A semiconcave function is twice differentiable almost everywhere. Actually, the following result holds (see [28]) :

Remark 4.3.8. *The following properties are equivalent :*

— *f is a semiconcave function ;*

— *There exists a constant C such that*

$$f(x) - C|x|^2$$

is a concave function ;

— *f is a.e. twice differentiable and there exists a constant C such that*

$$D^2 f \leq CI\,,$$

where I is the $n \times n$-identity matrix.

From this remark we immediately obtain :

$$f \in W^{2,\infty}(\Omega) \implies f \text{ semiconcave in } \Omega\,,$$

since each $W^{2,\infty}$ function satisfies $\left\| D^2 f \right\|_\infty \leq C$ for a certain C, and so the third condition of the remark is satisfied.

More generally, if f is in $W^{2,\infty}$ then f and $-f$ are semiconcave functions.

In order to prove the semiconcave regularity for the value function, we need the following technical lemma on semiconcave and $W^{2,\infty}$ functions :

Lemma 4.3.9. *Let $f : \Omega \to \mathbb{R}$. Suppose that the following inequality holds for every $x, y, z \in \Omega$, $\delta > 0$ and for a certain $C > 0$:*

$$f(x) + f(y) - 2f(z) \leq C\left(\delta + \frac{1}{\delta}\left(|x-z|^4 + |y-z|^4 + |x+y-2z|^2\right)\right)\,. \tag{4.3.17}$$

Then f is locally Lipschitz and semiconcave.

Furthermore, if f is Lipschitz and semiconcave then (4.3.17) *holds for all $x, y, z \in \Omega$, $\delta > 0$ and for a certain $C > 0$.*

Finally, if $f \in W^{2,\infty}(\Omega)$ then there exists a $C > 0$ such that, for every $x, y, z \in \Omega$ and $\delta > 0$

$$|f(x) + f(y) - 2f(z)| \leq C\left(\delta + \frac{1}{\delta}\left(|x-z|^4 + |y-z|^4 + |x+y-2z|^2\right)\right)\,. \tag{4.3.18}$$

Démonstration. If (4.3.17) is true, then, taking $x = z + h$, $y = z - h$, $\delta = |h|^2$, we obtain (4.3.16) and this proves that f is semiconcave.

Furthermore, if we wake $z = y$ and $\delta = |x - y|$ from (4.3.17) we obtain

$$f(x) - f(y) \leq C\left(1 + |x-y|^2\right)|x-y|\,,$$

proving that f is locally Lipschitz.

On the other hand, suppose that f is Lipschitz and semiconcave. Then we know from (4.3.16) that

$$f(\tilde{x} + h) + f(\tilde{x} - h) - 2f(\tilde{x}) \leq C|h|^2\,,$$

for each $\tilde{x} \in \Omega$, $h \in \mathbb{R}^n$ such that $\tilde{x} + h, \tilde{x} - h \in \Omega$.

We choose $\tilde{x} = \frac{x+y}{2}$, $h = \frac{x-y}{2}$, obtaining (up to changing C)

$$f(x) + f(y) - 2f\left(\frac{x+y}{2}\right) \leq C|x-y|^2\,.$$

Then we have, using the Lipschitz bound on f,

$$f(x)+f(y)-2f(z)\le C|x-y|^2+2f\left(\frac{x+y}{2}\right)-2f(z)\le C\left(|x-y|^2+|x+y-2z|\right)$$

Writing $x-y=(x-z)+(y-z)$ and using a generalized Young's inequality, we get

$$f(x)+f(y)-2f(z)\le C\left(\delta+\frac{1}{\delta}\left(|x-z|^4+|y-z|^4+|x+y-2z|^2\right)\right)\,,$$

and so (4.3.17) holds.

Finally, if f is $W^{2,\infty}$ then f and $-f$ are Lipschitz and semiconcave functions. Then (4.3.17) holds for f and $-f$, proving that (4.3.18) is true. □

Proposition 4.3.10. *Suppose hypotheses of theorem 4.3.6 are satisfied. Furthermore, suppose that G, L^1, and L^2 are Lipschitz and semiconcave functions in the space variable, and b and Γ are $W^{2,\infty}$ functions in the space variable .*

Then, every solution of (4.3.2) *is semiconcave in the space variable.*

Démonstration. We want to prove that

$$u(t,x)+u(t,y)-2u(t,z)\le C\left(\delta+\frac{1}{\delta}\left(|x-z|^4+|y-z|^4+|x+y-2z|^2\right)\right)\,,$$

for a certain $C>0$ and for all $\delta>0$, $t\in(0,T]$, $x,y,z\in\Omega$.
In fact, choosing $x=z+h$, $y=z-h$, $\delta=|h|^2$, we obtain (4.3.16).

To do that, we argue as in the Lipschitz case, following the ideas of *Theorem VII.3* of [68].

We consider the following auxiliary function :

$$\phi(t,x,y,z)=M\left(\delta+\frac{1}{\delta}\left(|x-z|^4+|y-z|^4+|x+y-2z|^2\right)\right)+\gamma|x|^2+\frac{\nu}{t}\,, \qquad (4.3.19)$$

Due to the coercive term $\gamma|x|^2$ and the boundedness of u, we know that the quantity

$$u(t,x)+u(t,y)-2u(t,z)-\phi(t,x,y,z) \qquad (4.3.20)$$

achieves a maximum.
We denote one of its maximum points with $(\bar t,\bar x,\bar y,\bar z)\in(0,T]\times\mathbb{R}^{3n}$.
Suppose $\bar t=T$. Then

$$u(t,x)+u(t,y)-2u(t,z)-\phi(t,x,y,z)\le G(\bar x)+G(\bar y)-2G(\bar z)-\phi(T,\bar x,\bar y,\bar z)\,.$$

Since G is Lipschitz and semiconcave, the right-hand side term is non-positive for M sufficiently large, using (4.3.17). This implies

$$u(t,x)+u(t,y)-2u(t,z)\le \phi(t,x,y,z)\,. \tag{4.3.21}$$

Suppose now $\bar t\in(0,T)$. If the maximum of (4.3.20) is non-positive, then (4.3.21) remains true.

Now, suppose by contradiction that (4.3.20) attains a strictly positive maximum. This implies

$$u(\bar t,\bar x)+u(\bar t,\bar y)-2u(\bar t,\bar z)\ge M\left(\delta+\frac{1}{\delta}\left(|\bar x-\bar z|^4+|\bar y-\bar z|^4+|\bar x+\bar y-2\bar z|^2\right)\right) \tag{4.3.22}$$

and, since u is Lipschitz,

$$\gamma|\bar x|^2\le u(\bar t,\bar x)+u(\bar t,\bar y)-2u(\bar t,\bar z)\le C(|\bar x-\bar z|+|\bar y-\bar z|)\,, \tag{4.3.23}$$

for a certain $C>0$.

Since $-2u(t,x)$ is a subsolution of

$$-z_t-\tilde H^2(t,x,\Delta z)-\tilde H^1(t,x,\nabla z)=-2F(t,x)\,,$$

with $\tilde H^2(t,x,q)=-2H^2(t,x,-\frac12 q)$ and $\tilde H^1(t,x,p)=-2H^1(t,x,-\frac12 p)$, we can apply Proposition 4.3.3 with

$$u_1(t,x)=u(t,x)\,,\quad u_2(t,y)=u(t,y)\,,\quad u_3(t,z)=-2u(t,z)\,,$$
$$\phi(t,x,y,z)\text{ as defined before}\,,$$

obtaining that there exist a,b,c,X,Y,Z such that

$$(a,\nabla_x\phi(\bar t,\bar x,\bar y,\bar z),X)\in\overline{\mathcal P^{2,+}u(\bar t,\bar x)}\,,\qquad (b,\nabla_y\phi(\bar t,\bar x,\bar y,\bar z),Y)\in\overline{\mathcal P^{2,+}u(\bar t,\bar y)}\,,$$
$$(c,\nabla_z\phi,Z)\in\overline{\mathcal P^{2,+}(-2u)(\bar t,\bar z)}\in\overline{\mathcal P^{2,+}(-2u)(\bar t,\bar z)}\,,$$

and

$$a+b+c=-\frac{\nu}{\bar t^2}\,,\qquad \begin{pmatrix}X&0&0\\0&Y&0\\0&0&Z\end{pmatrix}\le D^2\phi(\bar t,\bar x,\bar y,\bar z)\,. \tag{4.3.24}$$

From now on, we will omit for the function ϕ its dependence on $(\bar t,\bar x,\bar y,\bar z)$.

It is immediate to prove that, if $(c, \nabla_z\phi, Z) \in \overline{\mathcal{P}^{2,+}(-2u)(\bar{t},\bar{z})}$, then

$$\left(-\frac{c}{2}, -\frac{1}{2}\nabla_z\phi(\bar{t},\bar{x},\bar{y},\bar{z}), -\frac{1}{2}Z\right) \in \overline{\mathcal{P}^{2,-}u(\bar{t},\bar{z})}$$

.

Hence, since u is both a subsolution and a supersolution, one has

$$\begin{aligned}
&-a - H^2(\bar{t},\bar{x},\mathrm{tr}(X)) - H^1(\bar{t},\bar{x},\nabla_x\phi) + \lambda u(\bar{t},\bar{x}) \le 0\,,\\
&-b - H^2(\bar{t},\bar{y},\mathrm{tr}(Y)) - H^1(\bar{t},\bar{y},\nabla_y\phi) + \lambda u(\bar{t},\bar{y}) \le 0\,,\\
&\frac{c}{2} - H^2\left(\bar{t},\bar{z},-\frac{1}{2}\mathrm{tr}(Z)\right) - H^1\left(\bar{t},\bar{z},-\frac{1}{2}\nabla_z\phi\right) + \lambda u(\bar{t},\bar{z}) \ge 0\,.
\end{aligned}$$

Adding the first two equalities and subtracting twice the third we obtain

$$\begin{aligned}
&\frac{\nu}{\bar{t}^2} + \lambda(u(\bar{t},\bar{x}) + u(\bar{t},\bar{y}) - 2u(\bar{t},\bar{z}))\\
\le\,& H^2(\bar{t},\bar{x},\mathrm{tr}(X)) + H^2(\bar{t},\bar{y},\mathrm{tr}(Y)) - 2H^2\left(\bar{t},\bar{z},-\frac{1}{2}\mathrm{tr}(Z)\right)\\
+\,& H^1(\bar{t},\bar{x},\nabla_x\phi) + H^1(\bar{t},\bar{y},\nabla_y\phi) - 2H^1\left(\bar{t},\bar{z},-\frac{1}{2}\nabla_z\phi\right)\,.
\end{aligned}$$

The first term in the left-hand side is non-negative, so we can ignore it. For the second term, we use (4.3.22) to get

$$\begin{aligned}
&\lambda M\left(\delta + \frac{1}{\delta}\left(|x-z|^4 + |y-z|^4 + |x+y-2z|^2\right)\right)\\
\le\,& H^2(\bar{t},\bar{x},\mathrm{tr}(X)) + H^2(\bar{t},\bar{y},\mathrm{tr}(Y)) - 2H^2\left(\bar{t},\bar{z},-\frac{1}{2}\mathrm{tr}(Z)\right)\\
+\,& H^1(\bar{t},\bar{x},\nabla_x\phi) + H^1(\bar{t},\bar{y},\nabla_y\phi) - 2H^1\left(\bar{t},\bar{z},-\frac{1}{2}\nabla_z\phi\right)\,.
\end{aligned} \tag{4.3.25}$$

In order to estimate the right-hand side terms, we first compute the derivatives of ϕ. We have

$$\begin{aligned}
\nabla_x\phi &= \frac{4M}{\delta}\left(|\bar{x}-\bar{z}|^2(\bar{x}-\bar{z}) + (\bar{x}+\bar{y}-2\bar{z})\right) + 2\gamma\bar{x}\,,\\
\nabla_y\phi &= \frac{4M}{\delta}\left(|\bar{y}-\bar{z}|^2(\bar{y}-\bar{z}) + (\bar{x}+\bar{y}-2\bar{z})\right)\,,\\
\nabla_z\phi &= -\frac{4M}{\delta}\left(|\bar{x}-\bar{z}|^2(\bar{x}-\bar{z}) - |\bar{y}-\bar{z}|^2(\bar{y}-\bar{z}) - 2(\bar{x}+\bar{y}-2\bar{z})\right)\,,
\end{aligned}$$

and so

$$D^2_{xx}\phi = \frac{4M}{\delta}\big(2(\bar x-\bar z)\otimes(\bar x-\bar z)+|\bar x-\bar z|^2 I+I\big)+2\gamma I\,, \qquad D^2_{xy}\phi=\frac{4M}{\delta}I\,,$$

$$D^2_{xz}\phi = -\frac{4M}{\delta}\left(2(\bar x-\bar z)\otimes(\bar x-\bar z)+|\bar x-\bar z|^2 I+2I\right)\,,$$

$$D^2_{yy}\phi = \frac{4M}{\delta}\left(2(\bar y-\bar z)\otimes(\bar y-\bar z)+|\bar y-\bar z|^2 I+I\right)\,,$$

$$D^2_{yz}\phi = -\frac{4M}{\delta}\left(2(\bar y-\bar z)\otimes(\bar y-\bar z)+|\bar y-\bar z|^2 I+2I\right)\,,$$

$$D^2_{zz}\phi = \frac{4M}{\delta}\left(2(\bar x-\bar z)\otimes(\bar x-\bar z)+|\bar x-\bar z|^2 I+2(\bar y-\bar z)\otimes(\bar y-\bar z)+|\bar y-\bar z|^2 I+4I\right)\,.$$

We estimate the H^2 terms in (4.3.25). Choosing σ as the optimal control for $H^2\left(\bar t,\bar z,-\frac12\mathrm{tr}(Z)\right)$, we obtain

$$H^2(\bar t,\bar x,\mathrm{tr}(X))+H^2(\bar t,\bar y,\mathrm{tr}(Y))-2H^2\left(\bar t,\bar z,-\frac12\mathrm{tr}(Z)\right)\le \frac{\Gamma(\bar x,\sigma)^2}{2}\mathrm{tr}(X)+\frac{\Gamma(\bar y,\sigma)^2}{2}\mathrm{tr}(Y)$$
$$+\frac{\Gamma(\bar z,\sigma)^2}{2}\mathrm{tr}(Z)+e^{-\lambda(T-t)}\left(L^2(\bar t,\bar x,\sigma)+L^2(\bar t,\bar y,\sigma)-2L^2(\bar t,\bar z,\sigma)\right)$$
$$\le C\phi+\frac12\mathrm{tr}(\Gamma(\bar x,\sigma)^2X+\Gamma(\bar y,\sigma)^2Y+\Gamma(\bar z,\sigma)^2Z)\,.$$

We estimate the trace term in the following way :

$$\mathrm{tr}(\Gamma(\bar x,\sigma)^2X+\Gamma(\bar y,\sigma)^2Y+\Gamma(\bar z,\sigma)^2Z)=$$
$$=\mathrm{tr}\left(\begin{pmatrix}\Gamma^2(\bar x,\sigma)I & \Gamma(\bar x,\sigma)\Gamma(\bar y,\sigma)I & \Gamma(\bar x,\sigma)\Gamma(\bar z,\sigma)I\\ \Gamma(\bar x,\sigma)\Gamma(\bar y,\sigma)I & \Gamma^2(\bar y,\sigma)I & \Gamma(\bar y,\sigma)\Gamma(\bar z,\sigma)I\\ \Gamma(\bar x,\sigma)\Gamma(\bar z,\sigma)I & \Gamma(\bar y,\sigma)\Gamma(\bar z,\sigma)I & \Gamma^2(\bar z,\sigma)I\end{pmatrix}\begin{pmatrix}X&0&0\\0&Y&0\\0&0&Z\end{pmatrix}\right)\le$$
$$\le\mathrm{tr}\left(\begin{pmatrix}\Gamma^2(\bar x,\sigma)I & \Gamma(\bar x,\sigma)\Gamma(\bar y,\sigma)I & \Gamma(\bar x,\sigma)\Gamma(\bar z,\sigma)I\\ \Gamma(\bar x,\sigma)\Gamma(\bar y,\sigma)I & \Gamma^2(\bar y,\sigma)I & \Gamma(\bar y,\sigma)\Gamma(\bar z,\sigma)I\\ \Gamma(\bar x,\sigma)\Gamma(\bar z,\sigma)I & \Gamma(\bar y,\sigma)\Gamma(\bar z,\sigma)I & \Gamma^2(\bar z,\sigma)I\end{pmatrix}D^2\phi\right)\,,$$

since the matrix on the left is non-negative definite and thanks to (4.3.24).

So we obtain, with standard computations,

$$\mathrm{tr}(\Gamma(\bar x,\sigma)^2X+\Gamma(\bar y,\sigma)^2Y+\Gamma(\bar z,\sigma)^2Z)\le$$
$$\le\frac{4M}{\delta}\left[(2+n)|\bar x-\bar z|^2(\Gamma(\bar x,\sigma)-\Gamma(\bar z,\sigma))^2+(2+n)|\bar y-\bar z|^2(\Gamma(\bar y,\sigma)-\Gamma(\bar z,\sigma))^2+\right.$$
$$\left.+n\left(\Gamma(\bar x,\sigma)+\Gamma(\bar y,\sigma)-2\Gamma(\bar z,\sigma)\right)^2\right]+2\gamma n\Gamma^2(\bar x,\sigma)\,.$$

Using the hypotheses on Γ, we finally obtain this bound for the H^2 term :

$$H^2(\bar{t},\bar{x},\mathrm{tr}(X)) + H^2(\bar{t},\bar{y},\mathrm{tr}(Y)) - 2H^2\left(\bar{t},\bar{z},-\frac{1}{2}\mathrm{tr}(Z)\right) \leq C\phi + C\gamma\,,$$

where C depends also on n.

We reason in a similar way in order to bound the H^1 term. We have, choosing α as the optimal control for $H^1(\bar{t},\bar{x},-\frac{1}{2}\nabla_z\phi)$, and since $\nabla_z\phi = -(\nabla_x\phi + \nabla_y\phi) + 2\gamma\bar{x}$,

$$\begin{aligned}
&H^1(\bar{t},\bar{x},\nabla_x\phi) + H^1(\bar{t},\bar{y},\nabla_y\phi) - 2H^1\left(\bar{t},\bar{z},-\frac{1}{2}\nabla_z\phi\right) \leq \langle b(\bar{x},\alpha),\nabla_x\phi\rangle + \langle b(\bar{y},\alpha),\nabla_y\phi\rangle \\
&+\langle b(\bar{z},\alpha),\nabla_z\phi\rangle + e^{-\lambda(T-t)}\left(L^1(\bar{t},\bar{x},\alpha) + L^1(\bar{t},\bar{y},\alpha) - 2L^1(\bar{t},\bar{z},\alpha)\right) \\
\leq{}& C\phi + \frac{4M}{\delta}|\bar{x}-\bar{z}|^2\langle b(\bar{x},\sigma) - b(\bar{z},\sigma),\bar{x}-\bar{z}\rangle + \frac{4M}{\delta}|\bar{y}-\bar{z}|^2\langle b(\bar{y},\sigma) - b(\bar{z},\sigma),\bar{y}-\bar{z}\rangle \\
+{}& \frac{2M}{\delta}\langle b(\bar{x},\sigma) + b(\bar{y},\sigma) - 2b(\bar{z},\sigma),\bar{x}+\bar{y}-2\bar{z}\rangle + 2\gamma\langle b(\bar{x},\sigma),\bar{x}\rangle \leq C\phi + C\gamma|\bar{x}|\,.
\end{aligned}$$

Because of (4.3.23), we have $\gamma|\bar{x}| \to 0$ when $\gamma \to 0$.

Putting together all the estimates, we obtain, for M large enough,

$$\lambda\phi \leq C\phi + \omega(\gamma)\,,$$

where $\omega(\gamma)$ is a quantity that tends to 0 when $\gamma \to 0$.

So, for γ small enough and $\lambda \gg 1$, we obtain a contradiction.

So in each case (4.3.21) remains true and, letting γ and ν go to 0, we get

$$u(t,x) + u(t,y) - 2u(t,z) \leq M\left(\delta + \frac{1}{\delta}\left(|x-z|^4 + |y-z|^4 + |x+y-2z|^2\right)\right)\,,$$

which concludes the proof. □

4.4 Classical solutions in a regular case

In this section we want to prove that, for a specified class of $\mathcal{C}^2$ Hamiltonians, the solution u of (4.3.1) is actually in the space $\mathcal{C}^{1+\frac{\gamma}{2},2+\gamma}$, for some $\gamma > 0$. This estimate allows us to linearize problem (4.3.1) and apply the classical regularity results of linear parabolic equations in order to obtain a $\mathcal{C}^{1+\frac{\alpha}{2},3+\alpha}$, depending only on the coefficients H^1, H^2 and the data G. This last estimation will be crucial in order to obtain a $\mathcal{C}^{1+\frac{\alpha}{2},3+\alpha}$ solution in our framework.

To do that, we need to use the regularity results obtained by Krylov in 1983. According to [70], we define a class of functions for which the $\mathcal{C}^{1+\frac{\gamma}{2},2+\gamma}$ regularity will hold.

Definition 4.4.1. *Consider a function $M : [0,T] \times \mathbb{R} \times \mathbb{R}_{>0} \times \mathbb{R}^{n^2} \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ of variables $(t, x, \beta, B, \underline{p}, s)$, with $B = (b_{ij})_{ij}$ and $\underline{p} = (p_i)_i$.*

We say that $M \in \mathscr{M}$ if the following conditions are satisfied :

(i) M is positive homogeneous of first order, with respect to the variables $(\beta, B, \underline{p}, s)$;

(ii) M is twice continuously differentiable in the variables $(x, \beta, B, \underline{p}, s)$, for all $t \in [0,T]$;

(iii) $\exists \nu > 0$ such that

$$\sum_{i,j} \frac{\partial M}{\partial b_{ij}} \xi_i \xi_j \geq \nu |\xi|^2 , \qquad \forall \xi \in \mathbb{R}^n ;$$

(iv) M is concave with respect to $(b_{ij})_{ij}$;

(v) The second order directional derivative of M with respect to $(B, \underline{p}, s)$ along a vector $(B_0, \underline{p_0}, s_0)$ is bounded from above by

$$C\beta^{-1} \left(|\underline{p_0}|^2 + |s_0|^2 \right) ,$$

for a certain $C > 0$;

(vi) There exists $C > 0$ such that, $\forall i, j$,

$$|M_{b_{ij}}| + |M_{p_i}| + |M_\beta| + |M_{b_{ij}x}| + |M_{sx}| + |M_{\beta x}| \leq C ; \tag{4.4.1}$$

(vii) M continuous in all variables and differentiable with respect to t, with

$$|M_t(t, x, \beta, B, \underline{p}, s)| + |M_{x_i x_j}(t, x, \beta, B, \underline{p}, s)| \leq C \sqrt{\beta^2 + s^2 + |\underline{p}|^2 + |B|^2} ;$$

With these hypotheses, we can state the main result we will use in order to prove the classical regularity of u. This is *Theorem 1.1* of [70].

Theorem 4.4.2. *Let, for $r \in \mathbb{N}$, $(M_r)_r$ a sequence of function such that $M_r \in \mathscr{M}$ $\forall r$. Let morever $\phi \in \mathcal{C}^{2+\alpha}$, for a certain $0 < \alpha < 1$. Then, if we define $M = \inf_r M_r$, we have that the problem*

$$\begin{cases} u_t - M(t, x, 1, D^2 u, Du, u) = 0 , & (t, x) \in (0,T) \times \mathbb{R}^n \\ u(0, x) = \phi(x) , \quad x \in \mathbb{R}^n \end{cases}$$

admits a unique solution $u \in \mathcal{C}^{2+\gamma}([0,T] \times \mathbb{R}^n)$, where $\gamma \in (0,1)$.

Now we are ready to prove the $\mathcal{C}^{1+\frac{\alpha}{2},2+\alpha}$ regularity of u.

Theorem 4.4.3. *Let H^1 and H^2 be differentiable with respect to t and twice continuously differentiable with respect to the other variables. Let, moreover, $G \in \mathcal{C}^{2+\alpha}(\mathbb{R}^n)$, for a certain $0 < \alpha < 1$.*

Furthermore, we suppose that H^2 and H^1 are concave with respect to the last variable, and

$$H^2_q(t,x,q) \geq \nu \qquad \forall (t,x,q) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \text{ and for a certain } \nu > 0\,. \tag{4.4.2}$$

Finally, we require the following assumptions : $\exists\, C > 0$ such that

$$\left\| H^1(\cdot,\cdot,0) \right\|_\infty + \left\| H^2(\cdot,\cdot,0) \right\|_\infty \leq C\,, \tag{4.4.3}$$

$$\left\| H^2_q(\cdot,\cdot,\cdot) \right\|_\infty + \left\| H^1_p(\cdot,\cdot,\cdot) \right\|_\infty \leq C\,, \tag{4.4.4}$$

$$\left\| H^2_{qx}(\cdot,\cdot,\cdot) \right\|_\infty \leq C\,, \tag{4.4.5}$$

$$H^1_p(t,x,p)\cdot p - H^1(t,x,p) \geq -C\,, \qquad H^2_q(t,x,q)\cdot q - H^2(t,x,q) \geq -C\,, \tag{4.4.6}$$

$$|H^1_{px_i}(t,x,p)\cdot p - H^1_{x_i}(t,x,p)| \leq C\,, \qquad |H^2_{qx_i}(t,x,q)\cdot q - H^2_{qx_i}(t,x,q)| \leq C\,, \tag{4.4.7}$$

$$\left\| H^2_{xx}(\cdot,\cdot,q) \right\|_\infty + \left\| H^2_t(\cdot,\cdot,q) \right\|_\infty \leq C(1+|q|)\,, \tag{4.4.8}$$

$$\left\| H^1_{xx}(\cdot,\cdot,p) \right\|_\infty + \left\| H^1_t(\cdot,\cdot,p) \right\|_\infty \leq C(1+|p|)\,, \tag{4.4.9}$$

Then, if u is the unique solution of (4.3.1)*, we have $u \in \mathcal{C}^{1+\frac{\gamma}{2},2+\gamma}([0,T]\times\mathbb{R}^n)$, for a certain $0 < \gamma < 1$.*

Démonstration. We define $\forall\, r$ the function $M_r = M$, where

$$M(t,x,\beta,B,\underline{p},s) = \beta H^2\left(t,x,\beta^{-1}\mathrm{tr}(B)\right) + \beta H^1\left(t,x,\beta^{-1}\underline{p}\right)\,.$$

We note that this definition is well-posed, since $\beta \in \mathbb{R}_{>0}$, and we have

$$M(t,x,1,D^2u,Du,u) = H^2(t,x,\Delta u) + H^1(t,x,Du)\,.$$

So, if we can apply Theorem 4.4.2, we will obtain a $\mathcal{C}^{1+\frac{\gamma}{2},2+\gamma}$ solution for (4.3.1), and we will have finished.

We only have to check that $M \in \mathscr{M}$. We check all the conditions required by Definition 4.4.1.

(i), (ii). The positive homogeneity with respect to $(\beta, B, \underline{p})$ and the twice continuously differentiability with respect to $(x, \beta, B, \underline{p})$ are immediate consequences of the definition of M and the hypotheses of regularity for H^1, H^2.

(iii). Since M depends only on b_{ij} when $i=j$, we have, $\forall\, \xi \in \mathbb{R}^n$,

$$\sum_{i,j} \frac{\partial M}{\partial b_{ij}} \xi_i \xi_j = \sum_i \beta\beta^{-1} H_q^2(t,x,\beta^{-1}\mathrm{tr}(B))\xi_i^2 \geq \nu|\xi|^2 \,.$$

(iv) Since $\beta > 0$, the concavity of M with respect to (b_{ij}) is a direct consequence of the concavity of H^2 with respect to q. Actually, we have

$$\frac{\partial^2 M}{\partial b_{ij}^2} = \beta\beta^{-2} H_{qq}^2(t,x,\beta^{-1}\mathrm{tr}(B)) \leq 0 \,.$$

(v) Take a vector $(B_0, \underline{p_0})$ and define the function $\phi(r) = M(t,x,\beta,B+rB_0,\underline{p}+r\underline{p_0})$. We have to prove that

$$\phi''(0) \leq C\beta^{-1}|p_0|^2 \,. \tag{4.4.10}$$

Computing the first derivative, we find

$$\phi'(r) = H_q^2(t,x,\beta^{-1}\mathrm{tr}(B+rB_0))\mathrm{tr}(B_0) + H_p^1(t,x,\beta^{-1}(\underline{p}+r\underline{p_0}))\cdot \underline{p_0}\,,$$

and so

$$\phi''(0) = \beta^{-1}\left[H_{qq}^2(t,x,\beta^{-1}\mathrm{tr}(B))\mathrm{tr}(B_0)^2 + \langle H_{pp}^1(t,x,\beta^{-1}\underline{p})\,\underline{p_0}\,,\,\underline{p_0}\rangle\right] \,.$$

Using the concavity of H^2 and the semiconcavity of H^1 with respect to the last variable, (4.4.10) is proved.

(vi) We analyze each term of (4.4.1). As regards $M_{b_{ij}}$ and M_{p_i}, we have for (4.4.4)

$$|M_{b_{ij}}| + |M_{p_i}| = |H_q^2(t,x,\beta^{-1}\mathrm{tr}(B))| + |H_p^1(t,x,\beta^{-1}\underline{p})| \leq C \,.$$

For M_β, we compute the derivative and we obtain

$$\begin{aligned} M_\beta = {} & H^2(t,x,\beta^{-1}\mathrm{tr}(B)) - \beta^{-1}\mathrm{tr}(B)H_q^2(t,x,\beta^{-1}\mathrm{tr}(B)) \\ & + H^1(t,x,\beta^{-1}\underline{p}) - \beta^{-1}\underline{p}H_q^2(t,x,\beta^{-1}\underline{p}) \,. \end{aligned}$$

Hence, the concavity of H^2 and H^1, with conditions (4.4.3) and (4.4.6), easily allow to obtain a bound for $|M_\beta|$. Similar and easier computations are made in order to bound $|M_{b_{ij}x}|$ and $|M_{\beta x}|$, using (4.4.5) and (4.4.7).

(vii) The continuity of M follows from the continuity of H^1 and H^2. We prove only the estimate for $|M_{x_i x_j}|$, since the $|M_t|$ goes along the same computations. We have

$$\begin{aligned} |M_{x_i x_j}| & \leq |\beta H_{x_i x_j}^2(t,x,\beta^{-1}\mathrm{tr}(B))| + |\beta H_{x_i x_j}^1(t,x,\beta^{-1}\underline{p})| \\ & \leq C\left(|\mathrm{tr}(B)| + |\underline{p}| + \beta\right) \leq C\sqrt{\beta^2 + |\underline{p}|^2 + |B|^2}\,, \end{aligned}$$

where we used (4.4.8), (4.4.9).

Then $M \in \mathscr{M}$ and we are allowed to apply Theorem 4.4.2. This concludes the proof. $\square$

Since u is a classical function, we are allowed to linearize problem (4.3.1). Actually, we get that u satisfies

$$\begin{cases} -u_t - V(t,x)\Delta u - Z(t,x)Du = b(t,x)\,, \\ u(T,x) = G(x)\,, \end{cases} \tag{4.4.11}$$

where

$$V(t,x) = \int_0^1 H_q^2\left(t,x,\lambda\Delta u(t,x)\right)d\lambda\,, \qquad Z(t,x) = \int_0^1 H_p^1\left(t,x,\lambda Du(t,x)\right)d\lambda\,,$$
$$b(t,x) = H^2(t,x,0) + H^1(t,x,0)\,.$$

Now we can use the linear character of the equation (4.4.11) in order to improve the regularity of u. The theorem is the following :

Theorem 4.4.4. *Suppose the hypotheses of Theorem 4.4.3 are satisfied. Furthermore, assume that, for some $k > 0$ and $C > 0$,*

$$|H_x^2(t,x,q)| + |H_x^1(t,x,p)| \le C(1+|p|) \qquad \forall t \in [0,T], x \in \mathbb{R}^n\,. \tag{4.4.12}$$

Morever, we require, for a certain $\alpha \in (0,1)$,

$$\left\| H_q^2(\cdot,\cdot,q) \right\|_{\frac{\alpha}{2},\alpha} + \left\| H_p^1(\cdot,\cdot,p) \right\|_{\frac{\alpha}{2},\alpha} + \left\| H_x^2(\cdot,\cdot,q) \right\|_{\frac{\alpha}{2},\alpha} + \left\| H_x^1(\cdot,\cdot,p) \right\|_{\frac{\alpha}{2},\alpha} \le C_L \quad \forall |p| \le L, |q| \le L\,, \tag{4.4.13}$$

and all the derivatives H_q^2, H_p^1, H_x^2, H_x^1 be Lipschitz in the last variable.

Finally, suppose that $G \in \mathcal{C}^{3+\alpha}$, with

$$\|G\|_{3+\alpha} \le C\,. \tag{4.4.14}$$

Then the solution u of (4.3.1) *satisfies the following estimate :*

$$\|u\|_{1+\frac{\alpha}{2},3+\alpha} \le C\,, \tag{4.4.15}$$

where C depends on H^2, H^1, G, T, n.

Démonstration. We consider, for $1 \le i \le N$, the function $v := u_{x_i}$.

Differentiating the equation (4.3.1) with respect to x_i, we obtain for the function v

$$\begin{cases} -v_t - H_q^2(t,x,\Delta u)\Delta v - H_p^1(t,x,Du)\cdot Dv = H_{x_i}^2(t,x,\Delta u) + H_{x_i}^1(t,x,Du)\,, \\ v(T,x) = G_{x_i}(x)\,. \end{cases} \tag{4.4.16}$$

Thanks to the hypotheses, the coefficients H^2_q, H^1_p are in L^∞ and $G_{x_i} \in \mathcal{C}^{2+\alpha}$. Moreover, with the same ideas of Theorem 4.3.6 (see also [91]), we have that $|Du|$ is globally bounded in L^∞, which implies, thanks to (4.4.12),

$$H^2_{x_i}(t, x, \Delta u) + H^1_{x_i}(t, x, Du) \in L^\infty \,.$$

With these hypotheses we know from standard regularity results (see e.g. [74]) that

$$\|v\|_{\frac{1+\alpha}{2}, 1+\alpha} \le C \,,$$

where C depends on G, H^2, H^1.

Coming back to the linear equation (4.4.11) satisfied by u, we get that

$$V, Z, b \in C^{\frac{\alpha}{2}, \alpha} \,, \qquad G \in C^{2+\alpha} \,.$$

Hence, *Theorem IV.5.1* of [74] implies that

$$u \in C^{1+\frac{\alpha}{2}, 2+\alpha} \quad \text{ and } \|u\|_{1+\frac{\alpha}{2}, 2+\alpha} \le C \,.$$

Now, the coefficients of the equation (4.4.16) are in $C^{\frac{\alpha}{2}, \alpha}$ and $G_{x_i} \in C^{2+\alpha}$. Applying again *Theorem IV.5.1* of [74] we obtain

$$\|v\|_{1+\frac{\alpha}{2}, 2+\alpha} \le C \implies \|u\|_{1+\frac{\alpha}{2}, 3+\alpha} \le C \,,$$

which concludes the proof. □

4.5 Regular solutions for the Bellman operator

It is not obvious at all to have in our examples $\mathcal{C}^2$ Hamiltonian in all variables. In order to handle this problem, we introduce the following Proposition.

Proposition 4.5.1. *Let H^2 and H^1 be almost everywhere twice differentiable in the space and in the last variable, and almost everywhere differentiable in the time variable, and suppose H^1 and H^2 be concave in the last variable and satisfying estimates from* (4.4.2) *to* (4.4.9) *and from* (4.4.12) *to* (4.4.14)*, where pointwise estimates have to be intended almost everywhere.*

Then the unique solution u of (4.3.1) *satisfies*

$$\|u\|_{1+\frac{\alpha}{2}, 3+\alpha} \le C \,. \tag{4.5.1}$$

Démonstration. We consider a sequence $H^{2,k}$ and $H^{1,k}$ of smooth functions converging to H^2 and H^1, taking, for $\delta > 0$, the convolutions

$$H^{2,k} = H^2 * \rho_\delta \,, \qquad H^{1,k} = H^1 * \rho_\delta \,,$$

where $\rho_\delta \geq 0$ is a non-negative function with compact support in $B(0, \delta)$. Then we take the related solutions u^k of (4.3.1).

It is immediate to prove that $H^{2,k}$ and $H^{1,k}$ satisfy conditions from (4.4.2) to (4.4.9) and from (4.4.12) to (4.4.14), with all bounds independent on k. Just to give an example, we show that

$$H_p^{1,k}(t,x,p) \cdot p - H^{1,k}(t,x,p) \geq -C \,.$$

We have

$$\begin{aligned}
&H_p^{1,k}(t,x,p) \cdot p - H^{1,k}(t,x,p) \\
&= \int_0^T \int_{\mathbb{R}^n} \int_{-\infty}^{+\infty} \left(H_p^1(t-s, x-y, p-r) \cdot p - H^1(t-s, x-y, p-r) \right) \rho_\delta(s,y,r) \, dsdydr \\
&= \int_0^T \int_{\mathbb{R}^n} \int_{-\infty}^{+\infty} \left(H_p^1(t-s, x-y, p-r) \cdot (p-r) - H^1(t-s, x-y, p-r) \right) \rho_\delta(s,y,r) \, dsdydr \\
&+ \int_0^T \int_{\mathbb{R}^n} \int_{-\infty}^{+\infty} H_p^1(t-s, x-y, p-r) \cdot r \rho_\delta(s,y,r) \, dsdydr \geq -C \,,
\end{aligned}$$

where we use in the last passage hypotheses (4.4.4) and (4.4.6) for H^1.

Then, from the results of the previous sections,

$$\|u^k\|_{1+\frac{\alpha}{2}, 3+\alpha} \leq C \,.$$

Hence, we can use Ascoli-Arzelà Theorem on any compact set $K \subset \mathbb{R}^n$ and obtain that $\exists\, u$ such that, up to subsequences, $u^k \to u$ pointwise with all the derivatives. Moreover, u satisfies (4.5.1).

Passing to the limit in the equation of u^k, we obtain that u satisfies (4.3.1). This concludes the Proposition.

□

Remark 4.5.2. We stress the fact that condition (4.4.6) is satisfied at least in the model case (4.2.17), where

$$\begin{aligned}
H^1(t,x,p) &:= \inf_{\alpha \in \mathcal{U}} \left\{ \langle p, \alpha \rangle + L_1(t,x,\alpha) \right\} \,, \\
H^2(t,x,q) &:= \inf_{\eta \in \mathcal{S}'} \left\{ \eta q + L_3(t,x,\eta) \right\} \,,
\end{aligned}$$

provided L_1 and L_3 are uniformly bounded and strictly convex with respect to the last variable.

Actually, the linear character of $\langle p, \alpha\rangle$ and ηq with respect to α and η and the strict convexity of the Lagrangian functions immediately imply that the inf in H^1 and H^2 is attained at a unique point $\alpha_{t,x,p}$, $\eta_{t,x,q}$.

The uniqueness of the *argmin* plays a crucial role in order to obtain the $\mathcal{C}^1$ character of H^1 and H^2 with respect to the last variable. Actually, we have

$$H^1_p(t,x,p) = \alpha_{t,x,p}\,, \qquad H^2_q(t,x,q) = \eta_{t,x,q}\,,$$

and so

$$\begin{aligned} |H^1_p(t,x,p)\cdot p - H^1(t,x,p)| &= |L_1(t,x,\alpha_{t,x,p})| \le C\,,\\ |H^2_q(t,x,q)\cdot q - H^2(t,x,q)| &= |L_3(t,x,\eta_{t,x,p})| \le C\,. \end{aligned}$$

Moreover, if we strengthen the regularity hypotheses on L_1 and L_3 with respect to all variables, we obtain that also conditions (4.4.2),. . .,(4.4.9) of Theorem 4.4.3 are satisfied.

Finally, we note that the convexity condition of L_3 is satisfied if we require, in (4.2.13), that L_2 is convex and non-increasing with respect to the last variable. Actually, from Remark 4.2.3

$$L_3(t,x,\eta) = L_2(t,x,\sqrt{2\eta})\,,$$

and so, if L_3 and L_2 are twice differentiable in the last variable,

$$\partial^2_{\eta\eta}L_3(t,x,\eta) = \partial^2_{\sigma\sigma}L_x(t,x,\sqrt{2\eta})(2\eta)^{-\frac{1}{2}} - \partial_\sigma L_2(t,x,\sqrt{2\eta})(2\eta)^{-\frac{3}{2}} > 0\,.$$

4.6 The Fokker-Planck Equation and The Mean Field Games System

The existence and uniqueness for the Fokker-Planck equation follow from well-known arguments, thanks to its linearity character.

Before starting the study of the Mean-Field Games system, we need to prove the following Proposition about some regularity estimates of m.

Proposition 4.6.1. *Let m be the unique solution in $\mathcal{C}([0,T];\mathcal{P}(\mathbb{R}^n))$ of*

$$\begin{cases} m_t - \Delta(a(t,x)m) + \operatorname{div}(mb(t,x)) = 0\,, & (t,x) \in (0,T)\times\mathbb{R}^n\,, \\ m(0,x) = m_0(x)\,, & x\in\mathbb{R}^n\,, \end{cases} \tag{4.6.1}$$

where a is uniformly elliptic, a, b bounded in L^∞, and $m_0 \in \mathcal{P}(\mathbb{R}^n)$. Then m satisfies, for a certain $p_0 > 1$ and $C > 0$ both independent of m_0,

$$\|m\|_{L^p([0,T]\times\mathbb{R}^n)} + \sup_{s\neq t}\frac{\mathbf{d}_1(m(t),m(s))}{|t-s|^{\frac12}} + \sup_{t\in[0,T]}\int_{\mathbb{R}^n}|x|\,m(t,dx) \le C\,, \tag{4.6.2}$$

for all $p < p_0$.

Démonstration. We start assuming m_0 smooth, and we consider the process satisfying

$$\begin{cases} dX_t = b(t,X_t)dt + \sigma(t,X_t)dB_t\,, \\ X_0 = Z\,, \end{cases}$$

where Z is a random variable of law m_0, $(B_t)_t$ is a standard Brownian motion and $\sigma\sigma^* = a$. Then, assuming without loss of generality that $t > s$, we have

$$\begin{aligned} \mathbb{E}[|X_t - X_s|] &\le \mathbb{E}\left[\int_s^t |b(r,X_r)|\,dr\right] + \mathbb{E}\left[\left|\int_s^t \sigma(r,X_r)\,dB_r\right|\right] \\ &\le C|t-s| + \mathbb{E}\left[\int_s^t |\sigma(r,X_r)|^2\,dr\right]^{\frac12} \le C|t-s|^{\frac12}\,. \end{aligned}$$

Since X_t has law $m(t)$ and X_s has law $m(s)$, we obtain

$$\sup_{s\neq t}\frac{\mathbf{d}_1(m(t),m(s))}{|t-s|^{\frac12}} \le C\,.$$

To prove the L^p bound for m, we consider ϕ as the solution in $[0,T]\times\mathbb{R}^n$ of

$$\begin{cases} -\phi_t - a(t,x)\Delta\phi + b(t,x)\cdot D\phi = \psi\,, \\ \phi(T) = 0\,, \end{cases}$$

where $\psi \in L^q(\Omega)$ for a certain q which will be specified later. Then, *Theorem IV.9.1* of [74] tells us that

$$\|\phi\|_{W^{1,q}([0,T]\times\mathbb{R}^n)} \le C\,\|\psi\|_{L^q}\,.$$

Hence, for $q > n+1$, ϕ satisfies a Hölder estimate in space and time. Multiplying the equation of m by ϕ and integrating by parts, we obtain

$$\int_0^T \int_{\mathbb{R}^n} \psi(t,x)m(t,x)\,dxdt = \int_{\mathbb{R}^n} \phi(0,x)m_0(dx) \le C \left\|\psi\right\|_{L^q} \,,$$

which implies that

$$\left\|m\right\|_{L^p} \le C \,,$$

for all $p < p_0$, where p_0 is the conjugate exponent of $n+1$.

Finally, we consider ϕ as the solution in $[0,t] \times \mathbb{R}^n$ of

$$\begin{cases} -\phi_t - a(s,x)\Delta\phi + b(s,x) \cdot D\phi = 0 \,, \\ \phi(s) = |x| \,. \end{cases}$$

From standard regularity results (see e.g. [91]), we know that

$$\sup_{s\in[0,t]} \left\|\phi(s,\cdot)\right\|_\infty \le C(1+|x|) \,.$$

Hence, multiplying the equation of m by ϕ and integrating by parts, we obtain

$$\int_{\mathbb{R}^n} |x|m(t,x)\,dxdt = \int_{\mathbb{R}^n} \phi(0,x)m_0(dx) \le C\left(1 + \int_{\mathbb{R}^n} |x|m_0(dx)\right) \le C \,,$$

since $m_0 \in \mathcal{P}(\Omega)$.

Since these estimates do not depend on the smoothness of m_0, with a standard approximation we can obtain (4.6.2) $\forall m_0 \in \mathcal{P}(\mathbb{R}^n)$, which concludes the Proposition. □

Now we are ready to study the full Mean-Field Games system.

We handle the case where our couplings F and G are *non-local*.

Hypotheses 4.6.2. *$F : [0,T] \times \mathbb{R}^N \times \mathcal{P}(\mathbb{R}^N) \to \mathbb{R}$ is a $\mathcal{C}^{1+\alpha}$ function in the space variable and $\mathcal{C}^{\frac{\alpha}{2}}$ in the time variable, for a certain $\alpha \in (0,1)$, satisfying*

$$\sup_{m\in\mathcal{P}(\mathbb{R}^n)} \left\|F(\cdot,\cdot,m)\right\|_{\frac{\alpha}{2},1+\alpha} \le C \,.$$

Furthermore, F and satisfies a Lipschitz estimate with respect to the measure :

$$\left\|F(t,\cdot,m_1) - F(t,\cdot,m_2)\right\|_{1+\alpha} \le C\mathbf{d}_1(m_1,m_2) \,, \qquad \forall m_1,m_2 \in \mathcal{P}(\mathbb{R}^N) \,. \tag{4.6.3}$$

For G, we required that is a $\mathcal{C}^{3+\alpha}$ function in the space variable satisfying (4.6.3)*, with a $\mathcal{C}^{3+\alpha}$ norm bounded uniformly in m .*

We begin with the existence part.

Theorem 4.6.3. *Suppose the hypotheses of Proposition 4.5.1 are satisfied, and suppose* $m_0 \in \mathcal{P}(\mathbb{R}^N)$*. Then there exists at least one solution* $(u, m) \in \mathcal{C}^{\frac{3+\alpha}{2}, 3+\alpha} \times \mathcal{C}([0,T]; \mathcal{P}(\mathbb{R}^N))$ *for the Mean-Field Games system* (4.1.1)*.*

Furthermore, if one of these conditions is satisfied :

(i) H^1 *and* H^2 *strictly concave with respect to the last variable and* F *and* G *non-decreasing with respect to* m *:*

$$\begin{aligned} &\int_{\mathbb{R}^n} (F(t,x,m_1) - F(t,x,m_2))\,(m_1(dx) - m_2(dx)) \geq 0\,, \\ &\int_{\mathbb{R}^n} (G(x,m_1) - G(x,m_2))\,(m_1(dx) - m_2(dx)) \geq 0\,. \end{aligned} \tag{4.6.4}$$

(ii) H^1 *and* H^2 *concave with respect to the last variable and* F *and* G *strictly increasing with respect to* m*, i.e.* F *and* G *satisfy* (4.6.4) *and in addition :*

$$\int_{\mathbb{R}^n} (F(t,x,m_1) - F(t,x,m_2))\,(m_1(dx) - m_2(dx)) = 0 \implies F(t,x,m_1) = F(t,x,m_2)\,,$$
$$\int_{\mathbb{R}^n} (G(x,m_1) - G(x,m_2))\,(m_1(dx) - m_2(dx)) = 0 \implies G(x,m_1) = G(x,m_2)\,,$$

then the solution is unique.

Démonstration. The existence relies on Schauder fixed point Theorem.

We consider the following metrix space :

$$X = \left\{ \gamma \in \mathcal{C}([0,T]; \mathcal{P}(\mathbb{R}^N)) : \mathbf{d}_1(\gamma(t), \gamma(s)) \leq C|t-s|^{\frac{1}{2}} \right\}\,,$$

where C will be defined later.

It is immediate to note that X is a convex closed space.

We want to apply the Schauder fixed point Theorem. First, we define a suitable functional Φ.

For $\gamma \in X$ we consider u as the solution of the Hamilton-Jacobi equation

$$\begin{cases} u_t + H^2(t,x,\Delta u) + H^1(t,x,\nabla u) + F(t,x,\gamma(t)) = 0\,, & (t,x) \in [0,T] \times R^N\,, \\ u(T,x) = G(x,\gamma(T))\,, & x \in \mathbb{R}^n\,. \end{cases} \tag{4.6.5}$$

and then we define $\Phi(\gamma) = m$, where m is the solution of

$$\begin{cases} m_t - \Delta(m H^2_q(t,x,\Delta u)) + \mathrm{div}(m H^1_p(t,x,\nabla u)) = 0\,, & (t,x) \in (0,T) \times \mathbb{R}^n\,, \\ m(0,x) = m_0(x)\,, & x \in \mathbb{R}^n\,. \end{cases} \tag{4.6.6}$$

Thanks to the regularity results for u, we know that m is well-defined in the space $\mathcal{C}([0,T];\mathcal{P}(\Omega))$.

Proposition 4.6.1 implies that, if we choose wisely C in the definition of X, we have $m \in X$ and, in addition to that,

$$\|m\|_{L^p} \leq C\,,$$

with C not depending on γ.

In order to apply Schauder's Theorem we need to show the continuity of Φ and that $\Phi(X)$ is a relatively compact set.

We start by showing that $\Phi(X)$ is relatively compact.

Let $\{\gamma_n\}_n \subset X$, and let u_n and m_n be the solutions of (4.6.5) and (4.6.6) related to γ_n. From (4.4.15) we know that

$$\|u_n\|_{\frac{3+\alpha}{2},3+\alpha} \leq C\,,$$

where C does not depend on n. Hence, Ascoli-Arzelà and Banach-Alaoglu Theorems imply that $\exists\{u_{n_k}\}_k, u \in \mathcal{C}^{1+\frac{\alpha}{2},3+\alpha}$ such that $u_{n_k} \to u$ in $\mathcal{C}^{1,3}_{loc}$ and all the derivatives converge pointwise.

For simplicity, we call the subsequence u_{n_k} as u_n, forgetting the dependence on k.

To prove the convergence of $\{m_n\}_n$ (up to subsequences), we consider the processes X^n_s and X^k_s satisfying

$$\begin{aligned}
dX^n_s &= H^1_p(s, X^n_s, \nabla u_n(s, X^n_s))ds + \sqrt{2H^2_q(s, X^n_s, \Delta u_n(s, X^n_s))}dB_s\,,\\
dX^k_s &= H^1_p(s, X^k_s, \nabla u_k(s, X^k_s))ds + \sqrt{2H^2_q(s, X^k_s, \Delta u_k(s, X^k_s))}dB_s\,,
\end{aligned}$$

with $X^n_0 = X^k_0 = Z$, a process with density m_0. Then we have

$$\begin{aligned}
\mathbb{E}[|X^n_t - X^k_t|^2] \leq\ & \mathbb{E}\left[\int_0^t |H^1_p(s, X^n_s, \nabla u_n(s, X^n_s)) - H^1_p(s, X^k_s, \nabla u_k(s, X^k_s))|^2\, ds\right]\\
&+ \mathbb{E}\left[\left|\int_0^t \sqrt{2H^2_q(s, X^n_s, \Delta u_n(s, X^n_s))} - \sqrt{2H^2_q(s, X^k_s, \Delta u_k(s, X^k_s))}\, dB_s\right|^2\right].
\end{aligned} \tag{4.6.7}$$

We analyze each term. The last term is equal to

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t \left(\sqrt{2H_q^2(s,X_s^n,\Delta u_n(s,X_s^n))} - \sqrt{2H_q^2(s,X_s^k,\Delta u_k(s,X_s^k))}\right)^2 ds\right] \\
&\le \mathbb{E}\left[\int_0^t \left(\sqrt{2H_q^2(s,X_s^n,\Delta u_n(s,X_s^n))} - \sqrt{2H_q^2(s,X_s^n,\Delta u_k(s,X_s^n))}\right)^2 ds\right] \\
&+ \mathbb{E}\left[\int_0^t \left(\sqrt{2H_q^2(s,X_s^n,\Delta u_k(s,X_s^n))} - \sqrt{2H_q^2(s,X_s^k,\Delta u_k(s,X_s^k))}\right)^2 ds\right] \\
&\le C\int_0^t\int_{\mathbb{R}^n} |\Delta(u_n-u_k)|^2 m_n(s,x)\,dxds + C\int_0^t \mathbb{E}\left[|X_s^n - X_s^k|^2\right] \\
&= \omega(n,k) + C\int_0^t \mathbb{E}\left[|X_s^n - X_s^k|^2\right],
\end{aligned}$$

where we use the Lipschitz bound of H_q^2 and Δu_n.

In order to handle the first integral in the right-hand side, we consider a bounded domain $E \subseteq \mathbb{R}^n$ and we write

$$\begin{aligned}
&\int_0^t\int_{\mathbb{R}^n} |\Delta(u_n-u_k)|^2 m_n(s,x)\,dxds \\
&= \int_0^t\int_E |\Delta(u_n-u_k)|^2 m_n(s,x)\,dxds + \int_0^t\int_{E^c} |\Delta(u_n-u_k)|^2 m_n(s,x)\,dxds\,.
\end{aligned}$$

The first integral goes to 0 for any E bounded, thanks to the uniform convergence of Δu_n in bounded sets and the L^p bound of m_n. For the second integral, we note that $|\Delta(u_n-u_k)|^2$ is uniformly bounded and that m_n has a finite first order moment thanks to (4.6.2). Hence, $\forall\,\varepsilon > 0$ we can choose E sufficiently large such that

$$\int_0^t\int_{E^c} |\Delta(u_n-u_k)|^2 m_n(s,x)\,dxds \le \varepsilon\,.$$

For the arbitrariness of ε, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t \left(\sqrt{2H_q^2(s,X_s^n,\Delta u_n(s,X_s^n))} - \sqrt{2H_q^2(s,X_s^k,\Delta u_k(s,X_s^k))}\right)^2 ds\right] \\
&\le \omega(n,k) + C\int_0^t \mathbb{E}\left[|X_s^n - X_s^k|^2\right],
\end{aligned}$$

where $\omega(n,k)$ is a quantity converging to 0 when $n,k \to \infty$.

Similar estimates are made in order to bound the first term of (4.6.7). Hence we get

$$\mathbb{E}[|X_t^n - X_t^k|^2] \le \omega(n,k) + C\int_0^t \mathbb{E}\left[|X_s^n - X_s^k|^2\right],$$

and so, applying Gronwall's inequality,

$$\mathbb{E}[|X_t^n - X_t^k|] \leq \mathbb{E}[|X_t^n - X_t^k|^2] \leq C\omega(n,k) = \omega(n,k)\,,$$

which immediately implies

$$\sup_{t\in[0,T]} \mathbf{d}_1(m_n(t)), m_k(t)) \leq \omega(n,k)\,.$$

Hence, we have proved that $\{m_n\}_n$ is a Cauchy sequence in X. Then, up to subsequences, $\exists m$ such that $m_n \to m$. This proves the relatively compactness of $\Phi(X)$.

The continuity is an easy consequence of the compactness.

We consider $\gamma_n \to \gamma$. From the compactness, there exist subsequences $\{u_{n_k}\}_k$, $\{m_{n_k}\}_k$ converging to some u, m. Passing to the limit in the formulations of u_{n_k} and m_{n_k}, we obtain that u and m are the (unique) solutions of (3.3.10) and (3.3.11) related to γ and u.

Then, for each converging subsequence $\{m_{n_h}\}_h$, we must have $m_{n_h} \to m$. This means that the whole sequence $\Phi(\gamma_n) = m_n \to m = \Phi(\gamma)$. This proves the continuity of Φ.

So we can apply Schauder's theorem and obtain a classical solution of the problem (4.1.1).

To prove the uniqueness, let (u_1, m_1), (u_2, m_2) be two solutions of (4.1.1).

We want to estimate in two different ways the quantity

$$\int_0^T \int_{\mathbb{R}^n} ((u_1 - u_2)(m_1 - m_2))_t$$

First, computing directly the time integral we obtain

$$\int_{\mathbb{R}^n} (G(x, m_1(T)) - G(x, m_2(T)))(m_1(T) - m_2(T))\, dx \geq 0\,. \tag{4.6.8}$$

On the other hand, if we compute the derivative and use the weak formulation of (u_1, m_1) and (u_2, m_2), we obtain

$$
\begin{aligned}
&-\int_0^T \int_{\mathbb{R}^n} (F(x,m_1(t)) - F(x,m_2(t)))\,(m_1(t,dx) - m_2(t,dx))\,dt \\
&+\int_0^T \int_{\mathbb{R}^n} \left(H^2(t,x,\Delta u_2) - H^2(t,x,\Delta u_1) - H^2_q(t,x,\Delta u_1)\Delta(u_2-u_1)\right)\, m_1(t,dx)dt \\
&+\int_0^T \int_{\mathbb{R}^n} \left(H^2(t,x,\Delta u_1) - H^2(t,x,\Delta u_2) - H^2_q(t,x,\Delta u_1)\Delta(u_1-u_2)\right)\, m_2(t,dx)dt \\
&+\int_0^T \int_{\mathbb{R}^n} \left(H^1(t,x,Du_2) - H^1(t,x,Du_1) - H^1_p(t,x,Du_1)D(u_2-u_1)\right)\, m_1(t,dx)dt \\
&+\int_0^T \int_{\mathbb{R}^n} \left(H^1(t,x,Du_1) - H^1(t,x,Du_2) - H^1_p(t,x,Du_2)D(u_1-u_2)\right)\, m_2(t,dx)dt
\end{aligned}
$$

Since H^2 and H^1 are concave functions, all the above integrals are non-positive. Then, combining this result with (4.6.8), we obtain

$$
\int_0^T \int_{\mathbb{R}^n} ((u_1-u_2)(m_1-m_2))_t = 0\,,
$$

which means

$$
\begin{aligned}
\int_{\mathbb{R}^n} (G(x,m_1(T)) - G(x,m_2(T)))(m_1(T) - m_2(T))\,dx = 0\,, \\
\int_0^T \int_{\mathbb{R}^n} (F(x,m_1(t)) - F(x,m_2(t)))\,(m_1(t,dx) - m_2(t,dx)) = 0\,, \\
\int_0^T \int_{\mathbb{R}^n} \left(H^2(t,x,\Delta u_i) - H^2(t,x,\Delta u_j) - H^2_q(t,x,\Delta u_j)\Delta(u_i-u_j)\right)\, m_j(t,dx)dt = 0\,, \\
\int_0^T \int_{\mathbb{R}^n} \left(H^1(t,x,Du_i) - H^1(t,x,Du_j) - H^1_p(t,x,Du_j)D(u_i-u_j)\right)\, m_j(t,dx)dt = 0\,,
\end{aligned}
$$

for $(i,j) = (1,2)$ or $(2,1)$.

This allows us to conclude : actually, thanks to the hypotheses, there are two cases :

(i) if F and G are strictly increasing, then

$$
F(t,x,m_1(t)) = F(t,x,m_2(t))\,, \qquad\qquad G(x,m_1(T)) = G(x,m_2(T))\,.
$$

Hence, u_1 and u_2 solve the same HJB equation, which implies $u_1 = u_2$. Coming back to the FP equation, we have that $\Delta u_1 = \Delta u_2$ and $Du_1 = Du_2$. So m_1 and m_2 solve the same FP equation, which implies $m_1 = m_2$.

(ii) if H^1 and H^2 are strictly concave, we argue in a similar way. Actually, we have for example

$$H_p^1(t,x,Du_1) = H_p^1(t,x,Du_2)\,, \quad H_q^2(t,x,\Delta u_1) = H_q^2(t,x,\Delta u_2) \quad \text{for } x \in \operatorname{supp}(m_1(t))\,,$$

which implies

$$m_1 H_p^1(t,x,Du_1) = m_1 H_p^1(t,x,Du_2)\,, \quad m_1 H_q^2(t,x,\Delta u_1) = m_1 H_q^2(t,x,\Delta u_2)\,.$$

Hence, m_1 and m_2 solve the same FP equation, which implies $m_1 = m_2$. Coming back to the HJB equation, we have that $F(t,x,m_1(t)) = F(t,x,m_2(t))$ and $G(x,m_1(T)) = G(x,m_2(T))$. So u_1 and u_2 solve the same HJB equation, which implies $u_1 = u_2$.

Hence, the proof of uniqueness is completed.

□

Ringraziamenti

This thesis was a joint work between the Università di Roma Tor Vergata and the Université Paris Dauphine. So I spent my time in two different places, which meant a lot for me, both from a professional and a personal point of view. So, my thanks will be split in two different parts.

Pour la partie française, mon premier remerciement est pour mon directeur de thèse, M. Pierre Cardaliaguet. Merci pour tout l'aide que vous m'avez donné, et pour toute la patience que vous avez eu vers moi. Je ne l'oublierai jamais.

Je remercie aussi mon rapporteur, M. François Delarue, pour la précis attention qu'il a eu dans la lecture de ma thèse et pour toutes les corrections qui m'ont permis de la améliorer.

Nessuna frase è in grado di spiegare la gratitudine che ho nei confronti del mio relatore italiano, prof. Alessio Porretta. Grazie, grazie, grazie, per l'immenso aiuto datomi in questi tre anni, per avermi fatto crescere come matematico e come persona, per tutti i rimproveri nei momenti di inadempienza e per l'infinita pazienza nei miei confronti.

Grazie mille alla Prof. Annalisa Cesaroni, per essere stata referee di questa tesi e per tutti i consigli e le correzioni. Grazie al prof. Braides e a Simonetta De Nicola per essere sempre presenti e disponibili per ogni questione di carattere burocratico. Grazie poi a Vincenzo Ignazio, collaboratore insieme ad Ariel Neufeld per l'ultima parte della tesi.

È il momento di ringraziare i miei affetti più cari, che mi hanno sempre sostenuto in tutti i momenti di questi tre anni.

Il ringraziamento più grande è per la famiglia, i miei genitori e mia sorella, per essere sempre vicini a me con il loro amore e la loro comprensione. Un immenso grazie anche a mio nonno, ai suoi preziosi consigli, al suo affetto e alla sua saggezza sempre presenti nella mia vita. Grazie poi a mio zio Vincenzo e a tutti i parenti, vicini e lontani.

Grazie a MJ, per tutto ciò che abbiamo passato assieme e per esserci sempre stata nei momenti belli e brutti.

Infine, i ringraziamenti ai miei amici, romani, irpini e non solo, per la loro vicinanza e per tutti i bei momenti trascorsi insieme.

Fra gli amici romani non posso non incominciare con Andrea. Grazie, amico di sempre, per la tua costante presenza nella mia vita.

Un grazie speciale ai miei amici all'estero, lontani ma sempre vicini con il loro affetto, soprattutto nei momenti difficili : Elena, Alberto e Federica.

Grazie a Federico, Fix, Simone, Ersilia, e tutti gli amici del liceo, "Compagnia bella", per esserci sempre stati. Grazie agli amici d'infanzia, David e Giorgio, e grazie a Beatrice : un'amicizia nata ai tempi del dottorato ma subito divenuta fortissima e sincera.

Un grazie particolare va agli amici dell'università, per aver condiviso con me da vicino questo percorso. Prima fra tutti, grazie a Giulia, il mio "angelo custode" dal primo anno di triennale. Compagna di studi dal primo esame, coinquilina a Parigi e soprattutto, amica di sempre. Grazie mille per tutto.

Grazie a tutti gli amici di questo percorso accademico. Grazie al gruppo "Tarantino", Simone e Giovanna, compagni di studio dal primo anno e amici lontani ma vicini, ad Anna Paola e Giulio, ad Anna e Maurizia, per i momenti trascorsi a Roma e a Parigi, e agli amici dello studio 1225, Rossana, Gianluca, Duccio, Dario, Lorenzo, Benny, Guido.

Voglio poi ringraziare gli amici di Calitri, il bellissimo paese dell'Altirpinia che non mi ha dato i natali ma che chiamerò sempre *casa*.

Grazie a mio cugino Lorenzo, a Zuma e a Grazia per la vera amicizia e per la vicinanza dimostratami nei momenti più difficili.

Grazie ad Attilio, "amico, conterraneo e datore di lavoro sul fan club", per l'amicizia e i continui "sfavellamenti" che fanno iniziare e concludere ogni giornata con il sorriso.

Grazie poi a tutti quelli che rendono fantastici i miei soggiorni a Calitri. Grazie al mitico Tenny su tutti, grazie a CI, Eliana, i coniugi Sanacore, Michele, Maria Francesca, Aleo, Coach, Tateo, Marco, il Maestro Alfred, Fabio, Cicc, Gaetano, Tony, Antonio, Jhon. So per certo di aver dimenticato qualcuno, per questo concludo con un ringraziamento complessivo. Grazie a tutti voi, grazie Calitri.

Grazie poi a Gerardo, coinquilino Erasmus del lontano 2014, e da quel giorno amico sincero, "frato" a tutti gli effetti. E grazie agli altri amici Erasmus, ad Antonella, Gabriela, Ilenia.

Infine, un pensiero per le mie nonne, che non ci sono più ma il cui ricordo rimarrà sempre nel mio cuore.

Bibliographie

[1] Y. Achdou, M. Bardi, M. Cirant : *Mean Field Games models of segregation.* Math. Models Methods Appl. Sci., 2017.

[2] Y. Achdou, F.J. Buera, J-M. Lasry, P-L. Lions, B. Moll, *Partial differential equation models in macroeconomics*, Phil. Trans. R. Soc. A 372 (2014).

[3] M. Ajtai, J. Komlos, G. Tusnády : *On optimal matchings.* Combinatorica, Vol. 4(4), 259-264, 1984.

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

[5] M. Avellaneda, R. Buff : *Combinatorial Implications of Nonlinear Uncertain Volatility Models : the case of Barrier Options.* Applied Mathematical Finance, 1, 1-18, 1999.

[6] M. Avellaneda, A. Levy, A. Paras : *Pricing and hedging derivative securities in markets with uncertain volatilities.* Applied Mathematical Finance, 2, 73-88, 1995.

[7] M. Avellaneda, A. Paras : *Managing the Volatility Risk of Portfolios of Derivative Securities : the Lagrangian Uncertain Volatility Model.* Applied Mathematical Finance, 3, 21-52, 1996.

[8] P. Baldi : *Stochastic Calculus. An Introduction Through Theory and Exercises.* Springer International Publishing, 2017.

[9] M. Bardi, A. Cesaroni, L. Rossi : *Nonexistence of nonconstant solutions of some degenerate Bellman equations and applications to stochastic control*, ESAIM Control Optim. Calc. Var. 22 (2016), 842-861.

[10] M. Bardi, R. Jensen : *A Geometric Characterization of Viable Sets for Controlled Degenerate Diffusions.* Set-Valued Analysis, Vol. 10, 2-3, 129-141, 2002.

[11] G. Barles : *Fully nonlinear Neumann type boundary conditions for second-order elliptic and parabolic equations*, J. Differential Equations, 106, 90-106, 1993.

[12] G. Barles, J. Burdeau, *The Dirichlet problem for semilinear second-order degenerate elliptic equations and applications to stochastic exit time control problems*, Comm. Partial Differential Equations 20 (1995), 129-178.

[13] G. Barles, E. Rouy, *A strong comparison result for the Bellman equation arising in stochastic exit time control problems and its applications* Comm. Partial Differential Equations 23 (1998), 1995-2033.

[14] Bayraktar, E., Cohen, A., *Analysis of a finite state many player game using its master equation.* SIAM Journal on Control and Optimization, 56(5), 3538-3568, 2018.

[15] C. Benazzoli, L. Campi, G. Di Persio : *Mean field games with controlled jump-diffusion dynamics : Existence results and an illiquid interbank market model*, arXiv :1703.01919v2, 2018.

[16] Bensoussan, A., Frehse, J., *Control and Nash Games with Mean Field effect*, Chinese Annals of Mathematics, Series B, 34 (2), 161-192, 2012.

[17] A. Bensoussan, J. Frehse : *Smooth solutions of systems of quasilinear parabolic equations.* ESAIM : Control, Optimisation and Calculus of Variations, 8, 169-193, 2002.

[18] Bensoussan, A., Frehse J., Yam, S.C.P. : *Mean field games and mean field type control theory.* Briefs in Mathematics, 2013.

[19] A. Bensoussan, J. Frehse, S.C.P. Yam : *The Master Equation in mean field theory.* J. Math. Pures et Appliquées, 103, 1441-1474, 2015.

[20] A. Bensoussan, J. Frehse, S.C.P. Yam : *On the interpretation of the Master Equation.* Stoc. Proc. App., 127, 2093-2137, 2017.

[21] A. Bensoussan, P.-L. Lions : *Contrôle Impulsionnel et Inéquations Quasi-Variationnelles*, Dunod, Paris, 1982.

[22] L. Boccardo, F. Murat, J.-P. Puel, *Existence results for some quasilinear parabolic equations.* Nonlinear Anal. 13 (1989), 373-392.

[23] Buckdahn, R., Li, J., Peng, S., and Rainer, C., *Mean-field stochastic differential equations and associated PDEs.* Ann. Probab., 45, 824–878, 2017.

[24] X. Cabre, L.A. Caffarelli : *Fully Nonlinear Elliptic Equations*, AMS Colloquium Publications, Vol. 43, 1995.

[25] X. Cabre, L.A. Caffarelli : *Regularity for viscosity solutions of fully nonlinear equations* $F(D^2u) = 0$, Topol. Methods Nonlinear Anal. Vol. 6, N.1, 31-48, 1995.

[26] P. Cannarsa, R. Capuani, P. Cardaliaguet, *Mean Field Games with state constraints : from mild to pointwise solutions of the PDE system*, preprint : https ://hal.archives-ouvertes.fr/hal-01964755.

[27] P. Cannarsa, G. Da Prato, H. Frankowska : *Invariant measures associated to degenerate elliptic operators.* Indiana Univ. Math. J. 59 (2010), 53-78.

[28] P. Cannarsa, C. Sinestrari : *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control.* Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser Boston, 2004.

[29] P. Cardaliaguet : *Notes on Mean Field Games.* From P.-L. Lions' lectures at the Collège de France, 2010.

[30] Cardaliaguet, P., *The convergence problem in mean field games with local coupling. Applied Mathematics & Optimization,* 76(1), 177-215, 2017.

[31] P. Cardaliaguet, M. Cirant, and A. Porretta, *Remarks on nash equilibria in mean field game models with a major player*, arXiv preprint arXiv :1811.02811, 2018.

[32] P. Cardaliaguet, F. Delarue, J.M. Lasry, P.L. Lions : *The Master Equation and the Convergence Problem in Mean Field Games.* Annals of Mathematics Studies, Vol. 2, 2019.

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

[34] R. Carmona, F. Delarue, *Probabilist analysis of Mean Field Games.* SIAM Journal on Control and Optimization, 51(4), 2705-2734, 2013.

[35] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications*, Springer Verlag, 2017.

[36] R. Carmona, F. Delarue : *The Master Equation for large population equilibriums.* Stochastic Analysis and Applications 2014, Editors : D. Crisan, B. Hambly, T. Zariphopoulou. Springer, 2014.

[37] Carmona, R., Delarue, F., *The Master Field and the Master Equation.* Probabilistic Theory of Mean Field Games with Applications II (pp. 239-321). Springer, Cham, 2018.

[38] Carmona, R., Delarue, F. and Lacker, D. : *Probabilistic anal- ysis of mean field games with a common noise.* Ann. Probab, 44, 3740–3803, 2016.

[39] D. Castorina, A. Cesaroni, L. Rossi, *On a parabolic Hamilton-Jacobi-Bellman equation degenerating at the boundary*, Commun. Pure Appl. Anal. 15 (2016), 1251-1263.

[40] Cecchin, A., Pra, P.D., Fischer, M., Pelino, G., *On the convergence problem in mean field games : a two state model without uniqueness.* SIAM Journal on Control and Optimization, 57(4), 2443-2466, 2019.

[41] Cecchin, A., Pelino, G., *Convergence, fluctuations and large deviations for finite state mean field games via the master equation.* Stochastic Processes and their Applications, 129(11), 4510-4555, 2019.

[42] J.F. Chassagneux, D. Crisan, F. Delarue : *Classical solutions to the Master Equation for large population equilibria.* arXiv preprint arXiv :1411.3009, 2014.

[43] M. Cirant, *Multi-population Mean Field Games Systems with Neumann Boundary Conditions.* J. Math. Pures Appl., 2014

[44] M.G. Crandall, P.L. Lions : *Viscosity solutions of Hamilton-Jacobi equations.* Transactions of the American Mathematical Society, 277 (1), 1-42, 1983.

[45] M.G. Crandall, H. Ishii, P.-L. Lions : *User's guide to viscosity solutions of second order partial differential equations.* Bull. Amer. Soc., 27, 1-67, 1992.

[46] G. Da Prato, H. Frankowska : *Stochastic viability for compact sets in terms of the distance function*, Dynamic Systems and Applications, 10, 177-184, 2001.

[47] G. Da Prato, H. Frankowska : *Stochastic viability of convex sets.* J. Math. Analysis and Appl., 333, 151-163, 2007.

[48] F. Delarue, D. Lacker, K. Ramanan : *From the master equation to mean field game limit theory : a central limit theorem.* Electron. J. Probab. 24, no. 51, 1–54, 2019.

[49] Delarue, F., Lacker, D., Ramanan, K., *From the master equation to mean field game limit theory : Large deviations and concentration of measure.* arXiv preprint arXiv :1804.08550, 2018.

[50] M.C. Delfour, J.-P. Zolesio : *Shape analysis via oriented distance function.* J. Funct. Anal. 123 (1994), 129-201.

[51] S. Dereich, M. Scheutzow, R. Schottstedt : *Constructive quantization : approximation by empirical measures.* Annales de l'IHP, Probabilités et Statistiques, Vol. 49(4), 1183-1203, 2013.

[52] Doncel, J., Gast, N., Gaujal, B., *Discrete mean field games : Existence of equilibria and convergence.* arXiv preprint arXiv :1909.01209, 2019.

[53] F. Dragoni, E. Feleqi : *Ergodic Mean Field Games with Hörmander diffusions*, Calculus of Variations and Partial Differential Equations, Vol. 57, 116, 2018.

[54] El Karoui, N., and Chaleyat-Maurel, M., *Un problème de réflexion et ses applications au temps local et aux équations différentielles stochastiques sur* $\mathbb{R}$, Cas continu, Temps Locaux, Astérisque 52-53, pp. 117-144, 1978.

[55] G. Fichera, *Sulle equazioni differenziali lineari ellittico-paraboliche del secondo ordine*, Atti Accad. Naz. Lincei. Mem. Cl. Sci. Fis. Mat. Nat. Sez. I. 5 (1956), 1-30.

[56] Fischer, M., *On the connection between symmetric n-player games and mean field games.* The Annals of Applied Probability, 27(2), 757-810, 2017.

[57] W.H. Fleming, M. Soner : *Controlled Markov Processes and Viscosity Solutions*, Springer Science+Business Media, 2006.

[58] N. Fournier, A. Guillin : *On the rate of convergence in Wasserstein distance of the empirical measure.* Probability Theory and Related Fields, Vol. 162(3), 707-738, 2015.

[59] W. Gangbo and A. Swiech, *Existence of a solution to an equation arising from the theory of mean field games*, Journal of Differential Equations, 259, pp. 6573–6643, 2015.

[60] D. Gomes, J. Saude : *Mean Field Games Models - A Brief Survey.* Dynamic Games and Applications, 4(2), 110-154, 2014.

[61] D. Gomes, E. Pimentel, V. Voskanyan : *Regularity theory for Mean-Field Game Systems.* SpringerBriefs in Mathematics, Springer, 2016.

[62] O. Guéant, J-M. Lasry, P-L. Lions. *Mean field games and applications.* In Paris-Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Mathematics, vol. 2003 (2011), 205-266, Springer Berlin ed.

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

[64] Huang, M., Caines, P.E., Malhamé, R.P., *Large population Cost-Coupled LQG Problems With Nonuniform Agents : Individual-Mass Behavior and Decentralized ε-Nash Equilibria.* IEEE Transactions on Automatic Control, 52(9), 1560-1571, 2007.

[65] Huang, M., Caines, P.E., Malhamé, R.P. : *The Nash Certainty Equivalence Principle and McKean-Vlasov Systems : an Invariance Principle and Entry Adaptation.* 46th IEEE Conference on Decision and Control, 121–123, 2007.

[66] Huang, M., Caines, P.E., and Malhamé, R.P. : *An Invariance Principle in Large Population Stochastic Dynamic Games.* Journal of Systems Science & Complexity, 20 (2), 162–172, 2007.

[67] R. Isaacs : *Differential Games.* Wiley, New York, 1965.

[68] H. Ishii, P.-L. Lions : *Viscosity Solutions of Fully Nonlinear Second-Order Elliptic Partial Differential Equations.* J. Diff. Eq. 83 (1990), 26-78.

[69] V.N. Kolokoltsov, J. Li, W. Yang, *Mean Field Games and nonlinear Markov Processes.* Preprint arXiv :1112.3744, 2011.

[70] N.V. Krylov, *Boundedly Nonhomogeneous Elliptic and Parabolic Equations.* Math. USSR Izvestiya, Vol. 20, n. 3, 1983.

[71] N.V. Krylov, *Boundedly Nonhomogeneous Elliptic and Parabolic Equations in a Domain.* Math. USSR Izvestiya, Vol. 22, n. 1, 1984.

[72] Lacker, D. : *A general characterization of the mean field limit for stochastic differential games.* Probability Theory and Related Fields, 165, 581–648, 2016.

[73] Lacker, D., *On the convergence of closed-loop Nash equilibria to the mean field game limit.* arXiv preprint arXiv :1808.02745, 2018.

[74] O.A. Ladyženskaja, V.A. Solonnikov, N.N. Ural'ceva : *Linear and Quasi-linear Equations of Parabolic Type.* Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence R.I., 1967.

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

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

[77] Lasry, J.-M., Lions, P.-L., *Mean field games.* Jpn. J. Math. 2 (2007), no. 1, 229–260.

[78] Lasry, J.-M., Lions, P.-L., Guèant, O.,*Application of Mean Field Games to Growth Theory.* In : Paris-Princeton lectures on mathematical finance, 2010 ; Lecture notes in Mathematics. Springer, Berlin, 2011.

[79] T.Leonori, A. Porretta : *Gradient bounds for elliptic problems singular at the boundary,* Arch. Rat. Mech. Anal. 202 (2011), 663-705.

[80] P.-L. Lions,*Cours au Collège de France.* www.college-de-france.fr .

[81] P.-L. Lions : *On the Hamilton-Jacobi-Bellman equations*, Acta Appl. 1, 17-41, 1983.

[82] Lions, P.L., Menaldi. J.L., and Sznitman, A.S., *Construction de processus de diffusion réfléchis par pénalisation du domaine*, Comptes-Rendus Paris 292, pp. 559-562, 1981.

[83] P.L. Lions, A.S. Sznitman : *Stochastic Differential Equations with Reflecting Boundary Conditions.* Communications on Pure and Applied Mathematics, Vol. 37, 511-537, 1984.

[84] T.J. Lyons : *Uncertain volatility and the risk-free synthesis of derivatives.* Applied Mathematical Finance, Vol 2, 117133, 1995.

[85] A. Lunardi : *Analytic Semigroups and Optimal Regularity in Parabolic Problems.* Modern Birkhäuser Classics, 2012.

[86] J. Nash : *Non-cooperative games.* Annals of mathematics, 286-295, 1951.

[87] Nutz, M., Zhang, Y., *A mean field competition. Mathematics of Operations Research*, 44(4), 1245-1263, 2019.

[88] N.S. Pontryagin : *Linear Differential Games I and II.* Soviet Math. Doklady, Vol. 8, 3 & 4, 769-771 & 910-912, 1968.

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

[90] A. Porretta, *On the weak theory for mean field games systems*, Boll. Unione Mat. Ital. 10 (2017), 411-439.

[91] A. Porretta, E. Priola : *Global Lipschitz regularizing effects for linear and nonlinear parabolic equations*, J. Math. Pures Appl., 100, 633–686, 2013.

[92] A. Porretta, E. Zuazua, *Null controllability of viscous Hamilton-Jacobi equations*, Ann. I.H.P. Analyse Nonlinèaire 29 (2012), 301-333.

[93] Simon, J., *Compact sets in* $L^p(0,T;B)$, Ann. Mat. Pura Appl. 146 (4) (1987), 65-96.

[94] Skorokhod, A. V., *Stochastic equations for diflusion processes in a bounded region* 1, 2, Theor. Veroyatnost. i Primenen, 6, pp. 264-274, 7, 1962, pp. 3-23, 1961.

[95] Stroock, D. W., and Varadhan, S.R. S., *Diffusion Processes with boundary conditions*, Comm. Pure Appl. Math. 24, pp. 147-225, 1971.

[96] Tanaka, H., *Stochastic differential equations with reflecting boundary condition in convex regions*, Hiroshima Math. J. 9, pp. 163-177, 1979.

[97] D. Tataru, *Viscosity solutions of Hamilton-Jacobi equations with unbounded nonlinear terms*, J. Math. Anal. Appl. 163, 345–392, 1992.