Noise through an additional variable for mean field games master equation on finite state space

Charles Bertucci, Charles Meynard

Noise through an additional variable for mean field games master equation on finite state space

Charles Bertucci, Charles Meynard

February 7, 2024

Abstract

This paper provides a mathematical study of the well-posedness of master equation on finite state space involving terms modelling common noise. In this setting, the solution of the master equation depends on an additional variable modelling the value of a stochastic process impacting all players. Using technique from viscosity solutions, we give sufficient conditions for the existence of a Lipschitz continuous solution on any time interval. Under some structural assumptions, we are even able to treat cases in which the dynamics of this stochastic process depend on the state of the game.

Contents

1 Introduction

This paper is the first of a series devoted to the systematic study of mean field games (MFG) master equations associated to games in which a common source of randomness (or noise) is affecting all players through a finite dimensional parameter. We here treat the case of a master equation set on a subset of $\mathbb{R}^d$. The case of a master equation in infinite dimensions shall be tackled in a forthcoming work.

In a deterministic setting, the typical form of a master equation associated to a mean field game in a finite state space is

$$\begin{cases} -\partial_t V + (F(x,V)\cdot\nabla_x)V = G(x,V) \text{ for } (t,x)\in(0,T)\times\Omega, \\ V(T,x) = U_0(x) \text{ for } x\in\Omega. \end{cases} \tag{1.1}$$

In this framework, $T>0$ is the horizon of the game, and Ω a bounded subset of $\mathbb{R}^d$, $d\geq 1$. We are solving the previous equation for $V:[0,T]\times\Omega\to\mathbb{R}^d$, given $G,F:\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ which describe

respectively the evolution of the value and of the mean field quantities, and $U_0 : \Omega \to \mathbb{R}^d$ is the final value. The solution V of the master equation shall be called the value of the game.

The equation (1.1) is naturally associated to the system of forward backward ordinary differential equations set on $[t_0, T]$.

$$\begin{cases} \frac{dX_t}{dt} = -F(X_t, U_t), \\ \frac{dU_t}{dt} = -G(X_t, U_t), \\ X_{t_0} = x, \\ U_T = U_0(X_T). \end{cases} \tag{1.2}$$

Given a solution of V of (1.1) and $(X_t, U_t)_{t\in[t_0,T]}$ of (1.2), we get that for all $t \in [t_0, T], V(t, X_t) = U_t$. Hence, the existence of a solution (resp. regularity) to (1.1) is linked to uniqueness (resp. stability) properties of (1.2). At this level, it might seem strange that a forward-backward structure is imposed on (1.2), instead of a forward-forward or backward-backward one for instance. This is mainly due to the fact that the interpretation of (1.2) is natural in terms of the underlying game.

In this paper, we aim at investigating situations in which the characteristics are in fact stochastic, i.e. when (1.2) is a forward-backward stochastic differential equation (SDE). Namely, we are going to consider that the couplings G and F are going to depend on the value of a stochastic process $(p_t)_{t\geq 0}$ defined by

$$dp_t = -b(p_t)dt + \sqrt{2\sigma}dW_t,$$

where $(W_t)_{t\geq 0}$ is a m-dimensional Brownian motion on a standard (fixed) probability space, $b : \mathbb{R}^m \to \mathbb{R}^m$ is a data of the problem and $\sigma \geq 0$ is also given. In such a case, we look for a value which also depends on the value of this stochastic process (p_t). Thus the problem now becomes the one of finding a solution $V : [0, T] \times \Omega \times \mathbb{R}^m \to \mathbb{R}^d$ of

$$\begin{cases} -\partial_t V + (F(x, p, V) \cdot \nabla_x)V + b(p) \cdot \nabla_p V - \sigma\Delta_p V = G(x, p, V) \text{ in } (0, T) \times \Omega \times \mathbb{R}^m, \\ V(T, x, p) = U_0(x, p) \text{ in } \Omega \times \mathbb{R}^m. \end{cases} \tag{1.3}$$

This is what we mean by "noise through an additional variable": the presence of this new variable p in the master equation is a consequence of the dependency of the coefficients on the noise process $(p_t)_{t\geq 0}$. This is a different approach to noise (or randomness) in mean field game from the usual additive common noise, studied for instance in [11]. We believe this modeling of common noise to be natural in a wide variety of applications, see [10, 12, 22] for examples.

In what follows we are going to analyze (1.3) in two different contexts. The first one is mainly the one we just presented, in which the dynamics of b are purely exogenous to the underlying game. Taking proper care of the non-linearity $(F(x, U) \cdot \nabla_x)U$, we are able to extend the result existing on (1.1) to the case of (1.3).

We then proceed to study cases in which we allow the vector field b, which drives the evolution of the stochastic process $(p_t)_{t\geq 0}$ to depend on the state of the mean fields and on the value of the game, i.e. we allow b to depend on x and V. To the extent of our knowledge, this is the first attempt at studying MFG in which the dynamics of the noise depend on the MFG. This kind of dependence is extremely natural. Indeed, consider for instance the case of $(p_t)_{t\geq 0}$ modeling environmental variables, we are then modeling the fact that the players can impact on these environmental variables. Consider also the case in which $(p_t)_{t\geq 0}$ stands for certain prices affecting a financial system, or indexes (such as GDP) influencing an economy. We are then also modeling the facts that the players can have an impact on those quantities, which seems particularly natural.

Bibliographical comments Mean field games were first introduced by Lasry and Lions in [20, 23]. Since then, many works have contributed to the field, both extending it with new modelling concerns and contributing to pushing the mathematical analysis of these models. The master equation approach was introduced in [23] and lead to new mathematical questions. In general, the study of the master equation is done under structural assumptions, often on the monotonicity of the couplings G and F or the fact that we can integrate the master equation into an Hamilton-Jacobi-Bellman (HJB) equation (the so-called potential regime). A first well-posedness theory was obtained in [11] for classical solutions. Since then, several notions of weak solutions have been introduced, we can cite for instance the notion of monotone solution [4, 3, 13] which allows to define a concept of solution of master equation for values which are merely continuous. Note that the type of noise we are interested in was also discussed in [3]. Lipschitz solutions have been introduced in [8] and allow to work with well-defined notion of solution up to an eventual explosion time. Other regimes of well-posedness were studied in the works [28, 18].

Concerning master equations on a finite state space, since the initial findings presented in [23], a lot of results have been obtained. Without being exhaustive, we can mention that two different types of common noise were introduced and studied in [9, 2]. A regularizing effect was established in [7] under strong monotonicity assumptions. The paper [15] made the link between the master equation and the associated HJB equation in the potential case and [5] dealt with the convergence of the master equation in a finite state toward the master equation associated with a continuous state space when the number of states goes to infinity. Finally [25] explored existence and non uniqueness of solutions with irregular coefficients. None of these papers addresses the type of noise we propose to model here. Nonetheless, we shall follow an approach in the spirit of [8, 23]. Similar problems were encountered in the study of forward-backward stochastic differential equations (FBDSE) [30], we come back on the link between their approach and ours at the end of the paper.

Modelling comments Throughout this paper, we are not going to enter too much in details on the MFG interpretation of the master equation. While the main reason for this choice is that we focus here on the mathematical analysis of (1.3), we believe important to mention another reason for this. In general, the class of MFG master equation naturally models situations which are not proper MFG, but general economic equilibria of large populations. This is for instance the case in [10, 1, 6]. We can also mention [26] which explains why the class of so-called extended MFG can naturally arise and lead to similar master equations. Hence, we shall not require here

$$\Omega = \{x \in \left(\mathbb{R}^+\right)^d \quad \sum_i x_i = 1\},$$

even if this set is extremely natural for finite state MFG. We refer to the appendix of [4] for an explanation of how we can pass from our framework to this one.

Organization of the paper. In Section 2, we study the master equation when the noise process is solution of an autonomous stochastic differential equation. We start by establishing an a priori estimate and then prove results of existence and uniqueness of a global in time Lipschitz solution of the master equation.

In Section 3, we assume the dynamics of the noise may depend on the distribution of players and the value function of the game. We start by explaining the main difficulties arising from this dependence. We then proceed to justify some structural estimates on the drift and prove results of existence and uniqueness of a global Lipschitz solution in this case as well.

2 The case of an autonomous noise process

In this section, we assume b is a function of p only. In other words, we are interested in the well-posedness of

$$\begin{cases} \partial_t U + F(x,p,U)\cdot\nabla_x U + b(p)\cdot\nabla_p U - \sigma\Delta_p U = G(x,p,U) \text{ for } t\in(0,T), x\in\Omega, p\in\mathbb{R}^m, \\ U(0,x,p) = U_0(x,p) \text{ for } (x,p)\in\Omega\times\mathbb{R}^m. \end{cases} \tag{2.1}$$

Note that we have reversed time compared to (1.1), mainly to lighten notations. Recall that the unknown of the previous equation is a function $U : [0,T]\times\Omega\times\mathbb{R}^m \to \mathbb{R}^d$. The functions (F,G,b) are always assumed to be Lipschitz continuous (at least locally) in all the variables. We further assume that Ω is a smooth bounded domain of $\mathbb{R}^d$ and that F satisfies the following condition on its boundary:

$$\forall x\in\partial\Omega, (u,p)\in\mathbb{R}^d\times\mathbb{R}^m \quad \eta(x)\cdot F(x,p,u)\geq 0, \tag{2.2}$$

where $\eta(x)$ denotes the normal vector to $\partial\Omega$ in x directed toward the outside of the domain. This condition ensures we do not need to concern ourselves with boundary conditions for U on $\partial\Omega$. Furthermore, this kind of assumptions is natural in the context of MFG.

2.1 A propagation of monotonicity argument

It is well known that nonlinear PDE of the form of (2.1) can create shocks or singularity in finite time, such as Burger's equation. In general, structural conditions are required to prevent such explosive phenomena. We refer to P.L. Lions lectures at Collège de France [23] for extensive results concerning the PDE

$$\partial_t U + F(x,U)\cdot\nabla_x U = G(x,U) \text{ for } (t,x)\in(0,\infty)\times\mathbb{R}^d.$$

In particular, in a monotone regime, he showed, by a propagation of monotonicity argument, that a priori estimates can be obtained for such an equation. We shall say that we are in the monotone regime when the following Hypothesis holds.

Hypothesis 2.1. There exists $\alpha>0$ such that U_0 is $\alpha-$monotone and (G,F) is strongly monotone in x, i.e. for almost every (x,u) in $\Omega\times\mathbb{R}^d$, $\forall(\xi,\nu)\in\mathbb{R}^d\times\mathbb{R}^d$

$$\xi^T(D_x U_0)\xi \geq \alpha|D_x U_0\cdot\xi|^2, \tag{2.3}$$

$$(\xi,\nu)^T\begin{pmatrix} D_x G & D_u G \\ D_x F & D_u F \end{pmatrix}(\xi,\nu)\geq\alpha|\xi|^2. \tag{2.4}$$

Remark 2.2. It would be sufficient to require (G,F) to be $\alpha-$monotone in x instead of strong monotonicity. To avoid technical difficulties we here make a slightly stronger assumption as we delay such extension to a later section. In [23], other variants of the previous regime were also studied. The same approach can be carried on here. Indeed, we could also require (G,F) to be α-monotone with respect to u instead of x, in which case the assumption on U_0 can be weakened. As this extension does not raise any particular difficulty, we leave it to the interested reader.

A similar approach is also valid in the case of (2.1) as the next result presents

Proposition 2.3. *Assume that there exists $\alpha > 0$ such that Hypothesis 2.1 holds and let U be a smooth solution of (2.1), uniformly Lipschitz in (x,p). Then, for all $t \in [0,T], p \in \mathbb{R}^m$, $x \to U(t,x,p)$ is monotone. Furthermore, there exists a constant $C > 0$ depending only on F, G, b, U_0 and T such that*

$$\forall t \leq T, \quad \|D_x U(t,\cdot)\|_\infty \leq C.$$

Proof.
An equation on the gradient

First of all, by writing $W = U \cdot \xi$ let us remark that W is solution of

$$\partial_t W + F \cdot \nabla_x W + b \cdot \nabla_p W - \sigma \Delta_p W = G \cdot \xi.$$

Consider:

$$\begin{aligned} Z(t,x,p,\xi) &= \xi \cdot \nabla_x (U \cdot \xi) - \beta(t) |D_x U \cdot \xi|^2 \\ &= \xi \cdot \nabla_x W - \beta(t) |\nabla_x W|^2, \end{aligned}$$

for a smooth function of time β which has yet to be chosen. We can express the derivatives of Z in terms of (W, β):

$$\begin{cases} \nabla_x Z = (\xi - 2\beta(t) \nabla_x W) D_x^2 W, \\ \nabla_p Z = (\xi - 2\beta \nabla_x W) \cdot D_{px}^2 W, \\ \Delta_p Z = (\xi - 2\beta(t) \nabla_x W) \nabla_x \Delta_p W - 2\beta(t) |D_{xp}^2 W|^2, \\ \nabla_\xi Z = \nabla_x W + D_x U \cdot (\xi - 2\beta \nabla_x W). \end{cases}$$

Since U is smooth, we can derive directly the equation satisfied by W to get the one satisfied by its gradient

$$\begin{aligned} \partial_t (\nabla_x W) &= \nabla_x \left(G \cdot \xi - F \cdot \nabla_x W - b \cdot \nabla_p W + \sigma \Delta_p W \right) \\ &= \nabla_x (G \cdot \xi) - \nabla_x (F) \cdot \nabla_x W - F \cdot D_x^2 W - b \cdot D_{xp}^2 W + \sigma \Delta_p \nabla_x W. \end{aligned}$$

From which we get

$$\begin{aligned} \partial_t Z &= -\beta' |\nabla_x W|^2 + (\xi - 2\beta \nabla_x W) \cdot \partial_t (\nabla_x W) \\ &= -\beta' |\nabla_x W|^2 \\ &\quad + (\xi - 2\beta(t) \nabla_x W) \cdot \left(D_x G^T \xi - D_x F^T \nabla_x W \right) \\ &\quad + (\xi - 2\beta(t) \nabla_x W) \cdot \left(D_x U^T D_u G^T \xi - D_x U^T D_u F^T \nabla_x W \right) \\ &\quad + (\xi - 2\beta(t) \nabla_x W) \cdot \left(-F D_x^2 W - b D_{xp}^2 W + \sigma \Delta_p \nabla_x W \right). \end{aligned}$$

Remark that:

$$\begin{aligned} (\xi - 2\beta(t) \nabla_x W) \cdot \left(D_x U^T D_u G^T \xi \right) &= (D_x U \cdot (\xi - 2\beta(t) \nabla_x W)) \cdot \left(D_u G^T \xi \right) \\ &= (\nabla_\xi Z - \nabla_x W) \cdot \left(D_u G^T \xi \right). \end{aligned}$$

and that a similar calculation may be performed for the term in $D_u F$.
We may also notice that

$$(\xi - 2\beta(t) \nabla_x W) \cdot \left(-F D_x^2 W - b D_{xp}^2 W + \sigma \Delta_p \nabla_x W \right) = -F \cdot \nabla_x Z - b \cdot \nabla_p Z + \sigma \Delta_p Z + 2\beta(t) \sigma |D_{xp}^2 W|^2.$$

Looking at the equation satisfied by Z we get

$$\begin{aligned}
&\partial_t Z - \sigma\Delta_p Z + b\cdot\nabla_p Z + F\cdot\nabla_x Z + (\nabla_x W D_u F - \xi D_u G)\cdot\nabla_\xi Z \\
&= 2\beta\sigma|D^2_{xp}W|^2 - \frac{d\beta}{dt}|\nabla_x W|^2 \\
&+ \xi D_x G\xi - \xi D_u G\nabla_x W - \nabla_x W D_x F\xi + \nabla_x W D_u F\nabla_x W \\
&+ 2\beta\left(\nabla_x W D_x F\nabla_x W - \xi D_x G\nabla_x W\right).
\end{aligned}$$

We can now make use of the monotonicity of the couple (G,F) by noticing that

$$\begin{aligned}
\xi D_x G\xi - \xi D_u G\nabla_x W - \nabla_x W D_x F\xi + \nabla_x W D_u F\nabla_x W &= \begin{pmatrix} -\xi \\ \nabla_x W\end{pmatrix}^T \begin{pmatrix} D_x G & D_u G \\ D_x F & D_u F\end{pmatrix}\begin{pmatrix} -\xi \\ \nabla_x W\end{pmatrix} \\
&\geq \alpha|D_x G\xi|^2
\end{aligned}$$

by assumption. Now,

$$\begin{aligned}
&\partial_t Z - \sigma\Delta_p Z + b\cdot\nabla_p Z + F\cdot\nabla_x Z + (\nabla_x W D_u F - \xi D_u G)\cdot\nabla_\xi Z \\
&\geq \alpha|D_x G\cdot\xi|^2 - \frac{d\beta}{dt}|\nabla_x W|^2 + 2\beta\left(\nabla_x W D_x F\nabla_x W - \xi D_x G\nabla_x W\right) \\
&\geq (\alpha-\beta(t))|D_x G\cdot\xi|^2 + \left(\beta(t)(\|D_x G\|^2 + 2\|D_x F\|) - \frac{d\beta}{dt}(t)\right)|\nabla_x W|^2.
\end{aligned}$$

Let us take $\beta_0 < \alpha$ and $\beta(t) = \beta_0 e^{-\lambda_\beta t}$ with λ_β chosen in such fashion that

$$\lambda_\beta \geq \lambda^*_\beta = \|D_x G\|^2 + 2\|D_x F\|.$$

By assumption on U_0 and for this well chosen β we have:

$$\begin{cases} Z|_{t=0} \geq 0, \\ \partial_t Z - \sigma\Delta_p Z + b\cdot\nabla_p Z + F\cdot\nabla_x Z + (\nabla_x W D_u F - \xi D_u G)\cdot\nabla_\xi Z \geq 0 \text{ for } (0,T)\times\Omega\times\mathbb{R}^m. \end{cases} \tag{2.5}$$

Z is non-negative The non-negativity of Z is standard and follows classical comparison results. We recall the main argument for the sake of completeness. Recall that U is Lipschitz and $D_x U$ is hence bounded. As a consequence the SDE

$$\begin{cases} p_t = p_0 - \displaystyle\int_0^t b(p_s)ds + \sqrt{2\sigma}W_t, \\ X_t = x - \displaystyle\int_0^t F(X_s,p_s,U(\theta-s,X_s,p_s))ds, \\ \xi_t = \xi_0 - \displaystyle\int_0^t \left(D_x U(\theta-s,X_s,p_s)\cdot\xi_s D_u F(X_s,p_s,U(\theta-s,X_s,p_s)) - \xi_s D_u G(X_s,p_s,U(\theta-s,X_s,p_s))\right)ds. \end{cases}$$

are well defined on $[0,\theta]$ for any $\theta < T$. Since Z has at most quadratic growth in ξ uniformly in (x,p), applying Ito's Lemma to

$$Z(\theta,x,p_0,\xi_0) - Z(0,X_\theta,p_\theta,\xi_\theta)$$

and using (2.5) yields

$$\forall t < T \quad \forall (x, p_0, \xi_0) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^d \quad Z(t, x, p_0, \xi_0) \geq \mathbb{E}[Z_0(X_t, p_t, \xi_t)] \geq 0$$

Estimate on the gradient Of course, the non-negativity of Z on $[0, T]$ gives an estimate on the Lipschitz norm of U in x:

$$\begin{aligned} Z(t, x, p, \xi) \geq 0 &\iff \xi \cdot D_x U(t, x, p) \cdot \xi \geq \beta(t) \|D_x U(t, x, p) \cdot \xi\|^2 \\ &\implies \frac{\|D_x U(t, x, p) \cdot \xi\|}{\|\xi\|} \leq \frac{1}{\beta(t)}. \end{aligned}$$

Taking the supremum over all (ξ, x, p) yields the required estimate:

$$\forall t \leq T, \quad \|D_x U(t, \cdot)\|_\infty \leq \frac{1}{\beta(t)} \leq \frac{2}{\alpha} e^{\lambda_\beta T}.$$

The monotonicity of U in x is a trivial consequence of $Z \geq 0$.

□

We shall see later on that this estimate is crucial to obtain the well-posedness of (2.1).

2.2 On Lipschitz solutions

Lipschitz solutions are weak solutions of the master equation. The main idea here is that should the vector fields $F \circ U$ and $G \circ U$ be known and Lipschitz continuous, finding the solution of the system reduce to a simple application of the Feynman-Kac formula. As a consequence we can define the solution of the master equation as a fixed point of a Feynman-Kac inspired functional. The main advantage of this approach is that it allows to define a concept of solution of (2.1) for solutions which are merely Lipschitz continuous. We refer the reader to the paper [8] for a global presentation of this notion. This notion was only stated in the absence of the noise process p. Here, we only indicate how to adapt the results of [8] to the case of (2.1) as the extension is quite immediate.

Consider two vector fields $A : \begin{cases} \mathbb{R}^{+*} \times \mathbb{R}^d \times \mathbb{R}^m \longrightarrow \mathbb{R}^d \\ (t, x, p) \longrightarrow A(t, x, p) \end{cases}$ and $B : \begin{cases} \mathbb{R}^{+*} \times \mathbb{R}^d \times \mathbb{R}^m \longrightarrow \mathbb{R}^d \\ (t, x, p) \longrightarrow B(t, x, p) \end{cases}$ Lipschitz in (x, p) uniformly in t, such that Ω is left invariant by A, an initial condition U_0 Lipschitz in (x, p) and a Lipschitz continuous $b : \mathbb{R}^m \to \mathbb{R}^m$. Let $(\Omega_{\mathcal{P}}, \mathcal{F}, \mathbb{P})$ be a sufficiently rich probability space, we define the functional ψ by $\psi(T, A, B, U_0) = V$ if

$$\begin{cases} \forall t \in [0, T), (x, p) \in \Omega \times \mathbb{R}^m \quad V(t, x, p) = \mathbb{E}\left[\int_0^t A(t - s, X_s, p_s) ds + U_0(X_t, p_t) \middle| X_0 = x, p_0 = p \right], \\ dX_s = -B(t - s, X_s, p_s) dt, \\ dp_s = -b(p_s) dt + \sqrt{2\sigma} dW_s \text{ for } W \text{ a Brownian motion on } \mathbb{R}^m \text{ under } \mathbb{P}. \end{cases}$$

When the function V defined in such a way is smooth we expect it to be solution of:

$$\begin{cases} \partial_t V + B(t, x, p) \cdot \nabla_x V + b(p) \cdot \nabla_p V - \sigma \Delta_p V = A(t, x, p) \text{ in } (0, T) \times \Omega \times \mathbb{R}^m, \\ V|_{t=0} = U_0. \end{cases}$$

Which leads to the following definition of a Lipschitz solution of the master equation (2.1):

Definition 2.4. *Let $T > 0$, $U : [0,T) \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ is said to be a Lipschitz solution of* (2.1) *if :*

- *U is Lipschitz in (x,p) uniformly in $t \in [0,\alpha]$ for all α in $[0,T)$*
- *for all $t < T$:*
$$U = \psi(t, G(\cdot, U), F(\cdot, U), U_0).$$

Let us now state the following result on Lipschitz solutions in this context, which is a local existence theorem analogous to the Cauchy-Lipschitz theorem one could get for ordinary differential equations.

Theorem 2.5. *Let (F, G, b, U_0) be Lipschitz in all the variables, then:*

- *There always is an existence time $T > 0$ such that we have a unique solution in the sense of the aforementioned definition on $[0,T)$.*
- *There exist $T_c > 0$ and a maximal solution U defined on $[0,T_c[$ such that for all Lipschitz solutions V defined on $[0,T)$ we have: $T \leq T_c$ and $U|_{[0,T)} \equiv V$.*
- *If $T_c < \infty$ then $\lim_{t \to T_c} \|U(t,\cdot)\|_{Lip} = +\infty$.*

We do not prove this Theorem as it is a straightforward extension of previously existing results on Lipschitz solutions (see [8] Theorem 1.7).

Remark 2.6. For a function $f : \mathbb{R}^k \to \mathbb{R}^k$ we make use of the notation $\|f\|_{Lip}$ for its Lipschitz semi-norm:
$$\|f\|_{Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$
For a function of several variables $f(x,p)$, we make the following misuse of notation for its Lipschitz constant in x only:
$$\|D_x f\|_\infty = \sup_{x \neq y, p} \frac{|f(x,p) - f(y,p)|}{|x - y|},$$
as they are equal for C^1 Lipschitz functions.

Note that the previous result holds even if Hypothesis 2.1 is not satisfied. The next Section is devoted to generalizing the a priori estimate of Proposition 2.3 to functions which are only Lipschitz solutions of the equation. Namely we want to show that since we have an a priori estimate on $\|D_x U\|_\infty$, the maximal time T_c given in Theorem 2.5 is equal to $+\infty$. We end the present Section by explaining how, using the equation, we can deduce a bound on $D_p U$ from a bound on $D_x U$, thus making the estimation of $D_x U$ the crucial step of the argument.

Proposition 2.7. *Assume (F,G,b,U_0) are Lipschitz in all the variables. For any $T > 0$, there exists $C > 0$ such that for any U Lipschitz solution of* (2.1) *on $[0,T)$:*
$$\forall t < T, \quad \sup_{s \in [0,t]} \|D_p U(s,\cdot)\|_\infty \leq C \left(1 + \exp\left(t \sup_{s \in [0,t]} \|D_x U(s,\cdot)\|_\infty\right)\right).$$

Proof. Let us fix $T > 0$ and consider a Lipschitz solution U on $[0, T)$.

First, we can remark that, since b depends only on p, $(p_s)_{s\geq 0}$ satisfies an autonomous stochastic differential equation independent of U. As b is Lipschitz, the classical results for the continuity of SDE with respect to initial conditions are available.
let $t < T$, we define:

$$\forall u \in [0,t] \quad X_u^{p_0} = x - \int_0^u F(X_s^{p_0}, p_s, U(t-s, X_s^{p_0}, p_s^{p_0}))ds,$$

where

$$p_u^{p_0} = p_0 - \int_0^u b(p_s)ds + \sqrt{2\sigma}W_u$$

and

$$M_t = \sup_{s\in[0,t]} \|D_x U(s,\cdot)\|_\infty,$$

which is finite as $t < T$.

$$\mathbb{E}[|X_u^{p_0} - X_u^{q_0}|] \leq \|F\|_{Lip} \int_0^u ((1+M_t)\mathbb{E}[|X_s^{p_0} - X_s^{q_0}|] + (1 + \|D_p U(t-s,\cdot)\|_\infty)\mathbb{E}[|p_s^{p_0} - p_s^{q_0}|])\, ds.$$

From Gronwall's Lemma applied to standard Lipschitz SDE we know there exists a constant $C_b > 0$ depending only on $\|b\|_{Lip}$ and T such that

$$\forall s \in [0,T], \quad \mathbb{E}[|p_s^{p_0} - p_s^{q_0}|] \leq C_b |p_0 - q_0|.$$

We can conclude by yet another Gronwall's Lemma that

$$\forall u \in [0,t], \quad \mathbb{E}[|X_u^{p_0} - X_u^{q_0}|] \leq C_b e^{(1+M_t)u}|p_0 - q_0|(u + \int_0^u \|D_p U(t-s,\cdot)\|_\infty ds).$$

Using the fact that U is a Lipschitz solution of the master equation on $[0,t]$ gives

$$\begin{aligned}|U(t,x,p_0) - U(t,x,q_0)| \leq \mathbb{E}\Big[\|G\|_{Lip} \int_0^t ((1+M_t)|X_s^{p_0} - X_s^{q_0}| + (1 + \|D_p U(t-s)\|_\infty)|p_s^{p_0} - p_s^{q_0}|)ds\Big] \\ + \|U_0\|_{Lip}\mathbb{E}[|X_t^{p_0} - X_t^{q_0}| + |p_t^{p_0} - p_t^{q_0}|]\,.\end{aligned}$$

Hence

$$\frac{|U(t,x,p_0) - U(t,x,q_0)|}{|p_0 - q_0|} \leq C(1 + e^{M_t t} + \int_0^t \|D_p U(t-s,\cdot)\|_\infty ds),$$

where C is a positive constant depending only on $\|(F,G,b,U_0)\|_{Lip}$ and T. Since

$$\sup_{p,q} \frac{|U(t,x,p) - U(t,x,q)|}{|p-q|} = \|D_p U(t,\cdot)\|_\infty,$$

we conclude by Gronwall's lemma that

$$\sup_{s\in[0,t]} \|D_p U(s,\cdot)\|_\infty \leq C\left(1 + e^{tM_t}\right).$$

As this holds true for any $t < T$, the proof is complete. □

Remark 2.8. A similar estimate on $\|D_pU\|_\infty$ can still be obtained when b is assumed to a Lipschitz function of (x,p). The proof is slightly more technical but follows from the exact same argument. Moreover, under the assumption $\sigma > 0$, we may be able to get a better bound on the gradient of U in p. Indeed, we could get estimates on D_pU once D_xU is bounded by freezing x and considering the semi-linear parabolic system in p only:

$$\partial_t U + b(p)\cdot\nabla_p U - \sigma\Delta_p U = \underbrace{G(x,p,U) - F(x,p,U)\cdot\nabla_x U.}_{\text{source term Lipschitz in } U\text{, uniformly in } x}$$

However because our estimate on $\|D_xU\|_\infty$ is only valid when b depends on p only, we would rather not rely on parabolic regularity. This allows us to consider the fully degenerate case $\sigma = 0$.

2.3 Existence of global solutions

In this section we give sufficient conditions for the existence of a Lipschitz solution of the master equation (2.1) independently of the time horizon T. Namely, we show how we can make rigorous the a priori estimate of Proposition 2.3 for Lipschitz solutions of (2.1). From Proposition 2.7, we know that this would implies a bound on D_pU. Hence, since there is always uniqueness of a maximal Lipschitz solution, existence and uniqueness of a global Lipschitz solution shall follow. The proof of Proposition 2.3 is based on a derivation of the system (2.1). Obviously this does not generalise well to Lipschitz solutions, a first step to adapt the argument of Proposition 2.3 is to state Hypothesis 2.1 directly at the level of the functions (U_0, F, G) rather than on their gradient. We start with a Lemma,

Lemma 2.9. *Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a function of class C^1, the following two proposition are equivalent:*

- $\exists\alpha, \forall x \quad \xi\cdot Df(x)\cdot\xi \geq \alpha|Df\cdot\xi|^2.$ (i)
- $\exists\alpha, \forall(x,y) \quad \langle f(x)-f(y), x-y\rangle \geq \alpha|f(x)-f(y)|^2.$ (ii)

Remark 2.10. This elementary result is already well-known, we included a proof in Appendix for the sake of completeness.

According to Lemma 2.9, Hypothesis 2.1 implies that there exists $\alpha > 0$ such that for all x, y in Ω, u, v in $\mathbb{R}^d$ and p in $\mathbb{R}^m$

$$\langle U_0(t,x,p) - U_0(t,y,p), x-y\rangle \geq \alpha|U_0(t,x,p) - U_0(t,y,p)|^2, \tag{2.6}$$

$$\langle G(x,p,u) - G(y,p,v), x-y\rangle + \langle F(x,p,u) - F(x,p,v), u-v\rangle \geq \alpha|x-y|^2. \tag{2.7}$$

The main idea of this Section is that, to overcome the lack of regularity of a Lipschitz solution, we need to adapt the proof of Proposition 2.3 for another auxiliary function. Following the equivalence of Lemma 2.9, we define the following:

$$Z(t,x,y,p) = \langle U(t,x,p) - U(t,y,p), x-y\rangle - \beta(t)|U(t,x,p) - U(t,y,p)|^2, \tag{2.8}$$

for a Lipschitz solution U of (2.1) and a smooth function of time β to be chosen later on. This function shall play the role of Z in the proof of Proposition 2.3. If U was smooth we would expect this new function Z defined in (2.8) to be a solution of

$$\begin{aligned}&\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b(p) \cdot \nabla_p Z - \sigma \Delta_p Z \\ &= \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle \\ &+ 2\beta(t)\sigma|\nabla_p U^x - \nabla_p U^y|^2 - 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2,\end{aligned} \tag{2.9}$$

where we have used the notation $U^x = U(t,x,p)$,$F^x = F(x,p,U^x)$ and $G^x = G(x,p,U^x)$. Recall that in Proposition 2.3, we wanted to apply a comparison principle on Z. We keep the same strategy here. We are going to show that Z is a super-solution of

$$\begin{aligned}&\partial_t W + F^x \cdot \nabla_x W + F^y \cdot \nabla_y W + b(p) \cdot \nabla_p W - \sigma \Delta_p W \\ &\geq \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle - 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2.\end{aligned} \tag{2.10}$$

We now show that Z satisfies equation (2.10) in the viscosity sense. Let us remind the definition of a viscosity solution. Let $\mathcal{O}$ be an open subset of $\mathbb{R}^k$ for some $k \in \mathbb{N}$ and B be a continuous mapping from $\mathbb{R}^+ \times \mathcal{O} \times \mathbb{R} \times \mathbb{R}^m \times \mathcal{M}_k(\mathbb{R}) \to \mathbb{R}$. Consider the following non linear partial differential equation

$$\partial_t u(t,x) + B(t,x,u(t,x),\nabla_x u(t,x), D_x^2 u(t,x)) = 0 \text{ for } (t,x) \text{ in } (0,T)\times\mathcal{O} \tag{E}$$

Definition 2.11. *Let $u : (0,T)\times\mathcal{O} \to \mathbb{R}$ be a continuous function.*
(i) We say that u is a viscosity supersolution of (E) *if for any $\varphi \in C^{1,2}((0,T)\times\mathcal{O})$ such that a point of minimum (t^*,x^*) of $u-\varphi$ is achieved in $(0,T)\times\mathcal{O}$ the following holds*

$$\partial_t\varphi(t^*,x^*) + B(t^*,x^*,u(t^*,x^*),\nabla_x\varphi(t^*,x*), D_x^2\varphi(t^*,x^*)) \geq 0.$$

(ii) We say that u is a viscosity subsolution of (E) *if for any $\varphi \in C^{1,2}((0,T)\times\mathcal{O})$ such that a point of maximum (t^*,x^*) of $u-\varphi$ is achieved in $(0,T)\times\mathcal{O}$ the following holds*

$$\partial_t\varphi(t^*,x^*) + B(t^*,x^*,u(t^*,x^*),\nabla_x\varphi(t^*,x*), D_x^2\varphi(t^*,x^*)) \leq 0.$$

(iii) We say that u is a viscosity solution of (E) *if u is both a viscosity supersolution and a viscosity subsolution of* (E).

Note that in all generality, the definition of a viscosity solution can be extended beyond the continuous setting to function u that are only locally bounded. One of the key features of viscosity solutions is that under some structural assumptions on the operator B they satisfy a comparison principle. For more details on viscosity solutions we refer to [16] for a general introduction and to [31] for an introduction to viscosity solutions in the context of optimal control.

Lemma 2.12. *Assume that (F,G,b,U_0) are Lipschitz. Let U be a Lipschitz solution of the master equation* (2.1) *on $[0,T)$ for $T>0$ and Z be defined as in* (2.8).
Then Z is a viscosity supersolution of (2.10) *on $(0,T)$*

Proof.
A relation satisfied by Z
To prove this lemma, we make extensive use of the representation formula for Lipschitz solutions. For $t < T$ U is given by

$$\begin{cases} U(t,x,p) = \mathbb{E}\left[\int_0^t G(X_s,p_s,U(t-s,X_s,p_s))ds + U_0(X_t,p_t)\middle| X_0 = x, p_0 = p\right], \\ dX_s = -F(X_s,p_s,U(t-s,X_s,p_s))dt, \\ dp_s = -b(p_s)dt + \sqrt{2\sigma}dW_s. \end{cases}$$

To simplify computations we make use of the notation

$$G_s^{x,p} = G(X_s,p_s,U(t-s,X_s,p_s)),$$

for the process $(X_t,p_t)_{t\geq 0}$ starting in (x,p). Let us first remark that U satisfies a dynamic programming principle

$$U(t,x,p_0) = \mathbb{E}\Big[U(s,X_{t-s}^x,p_{t-s}^{p_0}) + \int_0^{t-s} G(X_u^x,p_u^{p_0},U(t-u,X_u^x,p_u^{p_0}))du\Big].$$

This follows from the definition of a Lipschitz solution. Let us introduce

$$W(t,x,y,p) = \langle U(t,x,p) - U(t,y,p), x-y\rangle.$$

From the previous dynamic programming principle, the following holds

$$W(t,x,y,p_0) =$$
$$\mathbb{E}\Big[W(s,X_{t-s}^x Y_{t-s}^y,p_{t-s}^{p_0}) + \int_0^{t-s} \left(\langle G_u^{x,p_0} - G_u^{y,p_0}, x-y\rangle + \langle F_u^{x,p_0} - F_u^{y,p_0}, U_s^{x,p_0} - U_s^{y,p_0}\rangle\right) du\Big]. \quad (2.11)$$

Remark that

$$M_{t-s}^{x,p_0} = U(s,X_{t-s}^x,p_{t-s}^{p_0}) + \int_0^{t-s} G(X_u^x,p_u^{p_0},U(t-u,X_u^x,p_u^{p_0})du$$

is a martingale on $[0,t]$. This is a classical result for Feynman-Kac representation formulae. To establish it, notice that for any $w > s$

$$\begin{aligned} &U(s,X_{t-s}^x,p_{t_s}^{p_0}) + \int_{t-w}^{t-s} G(X_u^x,p_u^{p_0},U(t-u,X_u^x,p_u^{p_0})du \\ &= U(s,\tilde{X}_{s-w}^{X_{t-w}^x},\tilde{p}_{s-w}^{p_{t-w}^{p_0}}) + \int_0^{s-w} G(\tilde{X}_u^{X_{t-w}^x},\tilde{p}_u^{p_{t-w}^{p_0}},U(w-u,\tilde{X}_u^{X_{t-w}^x},\tilde{p}_u^{p_{t-w}^{p_0}})du, \end{aligned}$$

where the process $(\tilde{X},\tilde{p})$ have been constructed with a Brownian motion independent of $\mathcal{F}_{t-w} = \sigma\left((W_s)_{s\leq t-w}\right)$. Taking the conditional expectation with respect to $\mathcal{F}_{t-w}$ and applying the dynamic programming principle satisfied by U we then easily get that for any $0 \leq s \leq w \leq t$

$$\mathbb{E}[M_{t-s}^{x,p_0} - M_{t-w}^{x,p_0}|\mathcal{F}_{t-w}] = 0.$$

Let $M_s = M_s^{x,p_0} - M_s^{y,p_0}$,

$$\begin{aligned} |U(t,x,p_0) - U(t,y,p_0)|^2 = |M_0|^2 &= \mathbb{E}\Big[|M_s|^2 - |M_s - M_0|^2\Big] \\ &\leq \mathbb{E}\Big[|M_s|^2\Big]. \end{aligned}$$

By definition of $(M_s)_{s\geq 0}$

$$|U(t,x,p_0)-U(t,y,p_0)|^2 \leq \mathbb{E}\left[\left|U_s^{x,p_0}-U_s^{y,p_0}+\int_0^{t-s}(G_u^{x,p_0}-G_u^{y,p_0})du\right|^2\right]. \tag{2.12}$$

We are now ready to prove the statement.

$$\begin{aligned} Z(t,x,y,p) &= W(t,x,y,p)-\beta(t)|U(t,x,p)-U(t,y,p)|^2 \\ &= \underbrace{W(t,x,y,p)}_{I} -\beta(s)\underbrace{|U(t,x,p)-U(t,y,p)|^2}_{II} -(\beta(t)-\beta(s))|U(t,x,p)-U(t,y,p)|^2. \end{aligned}$$

Applying (2.11) to I and (2.12) to II, and developing the square in inequality (2.12) implies that

$$\begin{aligned} Z(t,x,y,p_0) \geq\ & \mathbb{E}\Big[Z(s,X_{t-s}^x,Y_{t-s}^y,p_{t-s}^{p_0})-\beta(s)\|\int_0^{t-s}(G_u^{x,p_0}-G_u^{y,p_0})du\|^2\Big] \\ &+\mathbb{E}\Big[\int_0^{t-s}\left(\langle G_u^{x,p_0}-G_u^{y,p_0},x-y\rangle+\langle F_u^{x,p_0}-F_u^{y,p_0},U_s^{x,p_0}-U_s^{y,p_0}\rangle\right)du\Big] \\ &-(\beta(t)-\beta(s))\|U(t,x,p_0)-U(t,y,p_0)\|^2-2\beta(s)\mathbb{E}\Big[\int_0^{t-s}\langle U_s^{x,p_0}-U_s^{y,p_0},G_u^{x,p_0}-G_u^{y,p_0}\rangle|\Big]. \end{aligned} \tag{2.13}$$

Moreover, remark that such an inequality would still holds if instead of a deterministic time s we were to consider an almost surely bounded stopping time τ in $[0,t]$.

Z is a viscosity supersolution
Let us now assume that for some $(t,x,y,p_0)\in(0,T)\times\Omega^2\times\mathbb{R}^m$ there exists a smooth function φ such that

$$Z(t,x,y,p_0)-\varphi(t,x,y,p_0)=\min(Z-\varphi)=0,$$

and define the stopping time

$$\tau_h=\inf\{s>0 \quad (s,p_s)\notin[0,h]\times B(p_0,1)\}.$$

Where $B(p_0,1)=\{q\in\mathbb{R}^m \quad \|q-p_0\|\leq 1\}$. Because $Z(t-\tau_h,X_{\tau_h}^x,Y_{\tau_h}^y,p_{\tau_h}^{p_0})\geq\varphi(t-\tau_h,X_{\tau_h}^x,Y_{\tau_h}^y,p_{\tau_h}^{p_0})$, holds almost surely and it becomes an equality at (t,x,y,p_0) it is easy to see that φ satisfies inequality (2.13) with the stopping time for any $h\leq 1$. We may then apply Ito's lemma to φ and divide by h. Since $\varphi, D\varphi, D^2\varphi, F, G, b, U$ are all continuous function on $[t-1,t]\times\Omega\times B(p_0,1)$ which is compact, they are bounded on this set. By the dominated convergence theorem, we finally arrive at

$$\begin{aligned} &\partial_t\varphi+F^x\cdot\nabla_x\varphi+F^y\cdot\nabla_y\varphi+b(p)\cdot\nabla_p\varphi-\sigma\Delta_p\varphi \\ &\geq\langle G^x-G^y,x-y\rangle+\langle F^x-F^y,U^x-U^y\rangle-2\beta\langle G^x-G^y,U^x-U^y\rangle-\frac{d\beta}{dt}|U^x-U^y|^2. \end{aligned}$$

Because this holds at all points of minimum $(t,x,y,p_0)\in(0,T)\times\Omega^2\times\mathbb{R}^m$ of $Z-\varphi$ for any smooth function φ, Z is indeed a viscosity supersolution of (2.10). □

This slightly technical Lemma is the cornerstone that will allows us to show that monotonicity estimates are still valid for Lipschitz solution. We can now state the main result of this Section:

Theorem 2.13. *Assume that (U_0,G,F,b) are Lipschitz in all variables, under Hypothesis 2.1 there exists a unique global Lipschitz solution of* (2.1).

Proof. Because we already got an estimate on $\|D_pU\|_\infty$ in terms of $\|D_xU\|_\infty$ we need only to show that $\|D_xU\|_\infty$ does not blow up in finite time. All functions (U_0, G, F, b) are Lipschitz, hence there exists a time $T > 0$ and an associated Lipschitz solution U to (2.1) on $[0, T)$. Let Z be defined as in (2.8), we know thanks to Lemma 2.12 that on any time interval $[0, \tau], \quad \tau < T$, Z is a viscosity supersolution of (2.10). Now, using assumption (2.7) on the monotonicity of (G, F), we get that for any (t, x, y, p)

$$\begin{aligned}&\langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle - 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2\\ &\geq (\alpha - \|D_xG\|_\infty^2\beta)|x - y|^2 - (\beta\|D_uG\|_\infty + \frac{d\beta}{dt})|U^x - U^y|^2.\end{aligned}$$

Fixing $\beta_0 \leq \alpha$ and $\beta(t) = \beta_0 e^{-\lambda t}$ with $\lambda \geq \|D_uG\|_\infty$, we finally get that Z is a viscosity supersolution of

$$\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b(p) \cdot \nabla_p Z - \sigma\Delta_p Z \geq 0. \tag{2.14}$$

Since β_0 was choosen smaller than α, our assumption on U_0 also ensure that

$$Z_{t=0} \geq 0.$$

We may now conclude by a comparison principle that Z stay positive over the course of time for any time interval $[0, \tau]$ with $\tau < T$. We can apply Theorem 2.5 of [19] on such time interval which implies:

$$\inf_{[0,\tau]} Z \geq 0.$$

It only remains to show that the non-negativity of Z implies an estimate on its Lipschitz bound in x. Take $t < T, (x, y, p) \in \Omega^2 \times \mathbb{R}^m$, from the non-negativity of Z

$$|U(t, x, p) - U(t, y, p)\|x - y| \geq \beta(t)|U(t, x, p) - U(t, y, p)|^2.$$

We then deduce

$$\frac{1}{\beta(t)} \geq \frac{|U(t, x, p) - U(t, y, p)|}{|x - y|}.$$

Taking the supremum over (x, y, p), we finally get

$$\forall t < T, \quad \|D_xU(t, \cdot)\|_\infty \leq \frac{1}{\beta(t)}.$$

Since we have an uniform bound on the Lipschitz norm of U in x, it is then easy to see that the blow up time T of our Lipschitz solution U must satisfies $T = +\infty$ to avoid any contradiction, hence the existence of a global Lipschitz solution of the problem. □

3 Noise depending on the state of the game

In this Section we investigate the more intricate case of the master equation

$$\begin{cases} \partial_t U + F(x, p, U) \cdot \nabla_x U + b(x, p, U) \cdot \nabla_p U - \sigma\Delta_p U = G(x, p, U) \text{ for } t \in (0, T), x \in \Omega, p \in \mathbb{R}^m, \\ U(0, x, p) = U_0(x, p) \text{ for } (x, p) \in \Omega \times \mathbb{R}^m. \end{cases} \tag{3.1}$$

The only difference with (2.1) is that now we allow b, the drift of the stochastic process p, to depend both on x and the value U itself. We keep up with the assumption we made on $\Omega \subset \mathbb{R}^d$.

3.1 Preliminary remarks

3.1.1 Modeling

In our opinion, the fact that b can depend on (x, U) is natural in a wide variety of applications. Indeed, consider for instance the case in which p models the price of a stock. In such a situation, it seems fair for its evolution to depend on the distribution of players x and their decisions (captured by U itself in this finite states setting). In general, common noise models exogenous or environmental factors. We believe realistic to assume that in certain contexts, the players have an effect on their environment!

Let us insist that at the moment, there seems to have been remarkably few interest in the systematic study of PDE arising from this kind of model. In fact to the extent of our knowledge the only preexisting results were given on Forward backward stochastic differential equations (FBSDE) rather than on the associated partial differential equation [30].

3.1.2 Necessity of new estimates

In this setting, the definition of what is a Lipschitz solution has to be slightly changed to account for non linearity introduced by the dependency of b in U. Indeed we now define $\tilde{\psi}$ for three vector fields (A, B^1, B^2):

$$\begin{cases} & \tilde{\psi}(T, A, B^1, B^2, U_0) = V, \\ \text{where } \forall t \in [0,T), (x,p) \in \Omega \times \mathbb{R}^m & V(t,x,p) = \mathbb{E}\left[\int_0^t A(t-s, X_s, p_s) ds + U_0(X_t, p_t) \middle| X_0 = x, p_0 = p \right], \\ & dX_s = -B^1(t-s, X_s, p_s) dt, \\ & dp_s = -B^2(t-s, X_s, p_s) dt + \sqrt{2\sigma} dW_s. \end{cases}$$

and we define a Lipschitz solution as a fixed point of this new functional.

Definition 3.1. *Let $T > 0$, $U : [0,T) \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ is said to be a Lipschitz solution of* (3.1) *if :*

- *U is Lipschitz in (x,p) uniformly in $t \in [0,\alpha]$ for all α in $[0,T)$*
- *for all $t < T$:*

$$U = \tilde{\psi}(t, G(\cdot, U), F(\cdot, U), b(\cdot, U), U_0)$$

Previous results of local existence for Lipschitz solutions still hold in this more involved case (such as in [8], Theorem 2.4). We want to stress that global gradient estimates obtained in the autonomous case do not hold in general in this new model. While we expect that we could still get estimates on the gradient in p in function of the gradient in x using parabolic regularity, it is not so clear that $D_x U$ does not blow up in finite time.

Difficulties arising from the dependency of b on U

Before giving an example of an explosion of $D_x U$, we state a Lemma on solutions of the master equation which are invertible in x. This is a natural extension of a result presented by P.L. Lions during his lectures [23] to the master equation we are studying.

Lemma 3.2. *Let U be a regular solution of*

$$\begin{cases} \partial_t U + F(x,p,U)\cdot\nabla_x U + b(x,p,U)\cdot\nabla_p U - \sigma\Delta_p U = G(x,p,U) \quad (t,x,p)\in(0,T)\times\mathbb{R}^d\times\mathbb{R}^m, \\ U(0,x,p) = U_0(x,p). \end{cases}$$

Assume that for all (t,p) U is invertible in x and that V, its inverse in x, is smooth in all variables. Then V is a solution of

$$\begin{cases} \partial_t V + G(V,p,y)\cdot\nabla_y V + b(V,p,y)\cdot\nabla_p V - \sigma\Delta_p V = F(V,p,y) \quad (t,y,p)\in(0,T)\times\mathbb{R}^d\times\mathbb{R}^m, \\ V(0,y,p) = U_0^{-1}(y,p). \end{cases}$$

with the misuse of the notation $U_0^{-1}(y,p) = (U_0(\cdot,p))^{-1}(y)$.

Proof. Let $1\le i\le d$ and $t<T$, we know that:

$$V^i(t,U(t,x,p),p) = x^i.$$

As a consequence:

$$\begin{aligned} 0 &= \frac{d}{dp_k}V^i(t,U(t,x,p),p), \\ 0 &= \partial_{p_k}V^i + \partial_{y_j}V^i\partial_{p_k}U^j, \end{aligned} \tag{3.2}$$

where we used Einstein notation for summation. Similarly,

$$\begin{aligned} 0 &= \frac{d}{dt}V^i(t,U(t,x,p),p), \\ &= \partial_t V^i + \partial_{y_j}V^i\partial_t U^j. \end{aligned}$$

Plugging the equation satisfied by U in this last equality leads to

$$0 = \partial_t V^i + \partial_{y_j}V^i\left(G^j - F^k\partial_{x_k}U^j - b^k\partial_{p_k}U^j + \sigma\sum_k \partial^2_{p_k}U^j.\right).$$

Defining $y = U(t,x,p)$ and $x = V(t,y,p)$ gives

$$\partial_{y_j}V^i(t,U(t,x,p),p)G^j(x,p,U(t,x,p)) = G(V,p,y)\cdot\nabla_y V^i.$$

Because $\partial_{y_j}V^i\partial_{x_k}U^j = \delta_{ik}$ for the Kronecker delta symbol,

$$F^k\partial_{y_j}V^i\partial_{x_k}U^j = F^i(x,p,U(t,x,p)) = F^i(V(t,y,p),p,y).$$

For the last term involving σ,

$$\begin{aligned} \sigma\sum_k \partial^2_{p_k}U^j\partial_{y_j}V^i &= \sigma\sum_k \partial_{p_k}\left(\partial_{p_k}U^j\partial_{y_j}V^i\right) - \partial_{p_k}U^j\partial^2_{p_k,y_j}V^i - \partial_{p_k}U^j\partial^2_{y_j,y_l}V^i\partial_{p_k}U^l \\ &= \sigma\sum_k \partial_{p_k}\left(\partial_{p_k}U^j\partial_{y_j}V^i\right) - \partial_{p_k}U^j\partial_{y_j}\left(\partial_{p_k}V^i + \partial_{y_l}V^i\partial_{p_k}U^l\right). \end{aligned}$$

Making use of (3.2) we can write

$$\sigma\sum_k \partial^2_{p_k}U^j\partial_{y_j}V^i = -\sigma\sum_k \partial^2_{p_k}V^i.$$

Finally, we get

$$\partial_t V^i + G(V,p,y)\cdot\nabla_y V^i - F^i + b(V,p,y)\cdot\nabla_p V^i - \sigma\Delta_p V^i = 0.$$

□

Remark 3.3. While we decided to state this result on $\mathbb{R}^d$, in the proof we only use the fact that U is a one to one function. Taking an equation on the whole space is a choice of convenience as otherwise the domain of $y \to V(t, y, p)$ depends on (t, p).

The previous Lemma yields an analogy between the dependence of b on x and on U. Before giving an example of explosion in the case in which b depends on x, we present an example in which we highlight this formal change of variable.

Example 3.4. Consider the equation

$$\begin{gathered}\partial_t U + F(U) \cdot \nabla_x U + b(U) \cdot \nabla_p U - \sigma \Delta_p U = 0 \quad (t, x, p) \in (0, T) \times \mathbb{R}^d \times \mathbb{R}^m, \\ U(0, x, p) = U_0(x, p) \quad x \in \mathbb{R}^d, p \in \mathbb{R}^m.\end{gathered}$$

For U_0 invertible in x, we note V the inverse of U in x. Following Lemma 3.2 we expect V to be a solution of:

$$\begin{gathered}\partial_t V + b(y) \cdot \nabla_p U - \sigma \Delta_p U = F(y) \quad (t, y, p) \in (0, T) \times \mathbb{R}^d \times \mathbb{R}^m \\ V(0, y, p) = U_0^{-1}(y, p) \quad y \in \mathbb{R}^d, p \in \mathbb{R}^m,\end{gathered}$$

when it is well defined. From which we can deduce the expression of V by integrating along the characteristics

$$V(t, y, p) = \mathbb{E}\left[tF(y) + V_0(y, p - tb(y) + \sqrt{2\sigma} W_t)\right].$$

It then follows that

$$\begin{aligned}D_y V(t, y, p) &= \mathbb{E}\left[tD_u F(y) + D_y V_0(y, p - tb(y) + \sqrt{2\sigma} W_t) - tD_p V_0(y, p - tb(y) + \sqrt{2\sigma} W_t) D_u b\right] \\ &= \mathbb{E}\left[(D_x U_0)^{-1} \left(I_d + t(D_x U_0 D_u F(y) + D_p U_0 D_u b)\right)(y, p - tb(y) + \sqrt{2\sigma} W_t)\right],\end{aligned}$$

where the last line was obtained using the fact that U and V are inverse of each others,

$$D_p(U \circ V) = D_p U(V) + D_x U(V) D_p V = 0.$$

If there exists a point (t^*, y^*, p^*) such that $D_y V(t^*, y^*, p^*)$ is not invertible, then we expect that

$$\lim_{(t,y,p) \to (t^*, y^*; p^*)} \|D_x U(t, V(t, y, p), p)\| \to +\infty.$$

In the case $\sigma = 0$, a sufficient condition to avoid such an explosion could be

$$\forall (x, p) \in \mathbb{R}^d \times \mathbb{R}^m \quad D_x U_0(x, p) D_u F(U_0(x, p)) + D_p U_0(x, p) D_u b(U_0(x, p)) \geq 0.$$

This clearly imposes structural conditions on b and U_0.

Difficulties arising from the dependency of b on x It is quite natural that further assumptions and restrictions be required when b depends on U. However, the dependence of b in x, even if it does not add any additional non-linearity in the problem, can lead to the apparition of singularities in finite time. We now give such an example.

Example 3.5. Let us a consider a toy model in dimension one with

$$\begin{cases} F(x, p, U) = U, \\ G(x, p, U) = x, \\ b(x, p) = \lambda x, \\ U_0(x, p) = \alpha_0 x + \beta_0 p. \end{cases}$$

Looking for a linear solution $U(t,x,p) = \alpha(t)x + \beta(t)p$ of

$$\partial_t U + U\partial_x U + \lambda x \partial_p U - \sigma \partial_p^2 U = x \text{ in } (0,\infty) \times \mathbb{R}^2,$$

we get

$$\begin{cases} \frac{d\alpha}{dt}(t) + \alpha(t)^2 + \lambda\beta(t) = 1, \\ \frac{d\beta}{dt}(t) + \alpha(t)\beta(t) = 0, \\ \alpha(0) = \alpha_0, \\ \beta(0) = \beta_0. \end{cases}$$

We may easily solve for β assuming α is well behaved

$$\beta(t) = \beta_0 e^{-\int_0^t \alpha(s)ds}.$$

Hence α is solution to

$$\begin{cases} \frac{d\alpha}{dt}(t) + \alpha(t)^2 + \lambda\beta_0 e^{-\int_0^t \alpha(s)ds} = 1, \\ \alpha(0) = \alpha_0. \end{cases}$$

Let us now take $\lambda\beta_0 \geq 1$. Assume that $\alpha_0 < 0$ at time 0, then

$$\frac{d\alpha}{dt}(0) \leq -\alpha^2(0) < 0.$$

As a consequence α must be decreasing in a neighbourhood of zero and we deduce that on this set

$$\frac{d\alpha}{dt}(t) \leq -\alpha^2(t) < 0.$$

Because this leads to α further decreasing, we deduce that the inequality must be true for all times and hence that there exist $T < +\infty$ such that $\lim_{t\to T} \alpha(t) = -\infty$. This was to be expected as in this very simple case the monotonicity condition on U_0 reduces to $\alpha_0 \geq 0$.

Let us now assume that $\alpha_0 > 0$ so that U_0 is strongly monotone in x, it only remains to show that for a well chosen pair (λ, β_0) there exists a time t such that $\alpha(t) < 0$, using the previous result this will then be enough to ensure α blows up in finite time. Take a couple (λ, β_0) satisfying

$$1 - \lambda\beta_0 e^{-\alpha_0(\alpha_0+1)} \leq -1.$$

For this choice, α must be decreasing around 0 and by a similar argument as the one used earlier we deduce that

$$\forall t \in [0, \alpha_0 + 1], \frac{d\alpha}{dt}(t) \leq -1.$$

Hence α becomes negative in finite time.

Since $\partial_x U(t,x,p) = \alpha(t)$ this is indeed a blow up of the gradient of U in finite time. This shows that despite the regularizing effects from the (strong in this case) monotonicity[1] of both F and G, the dependency of b in x might be enough to perturb the well-posedness of the equation, even though it did not add further non linearity. Let us also remark the regularizing effect the monotonicity of b in p would have on the equation. In this linear setting this translates as

$$b(x,p) = \lambda x + \mu p \quad \mu > 0.$$

Hence β would become:

$$\beta(t) = \beta_0 e^{-\mu t - \int_0^t \alpha(s)ds}.$$

Adding the term in $e^{-\mu t}$ tends to make $\beta(t)$ smaller in absolute value (though it might not be enough to compete with the term in α).

[1] Indeed, it was shown in [7] that in the absence of b, such an equation regularizes the initial condition.

3.2 A heuristic on the mean field game master equation

As we just saw, the estimates of Section 2 are in general not true anymore when b depends on (x, U). However under additional assumptions, we may recover an estimate stemming from a similar idea to the one that led to Proposition 2.3. As the proof is somewhat calculation intensive (although very similar to the one in the case where b depends only on p), we present first an heuristic that shows from where both the result, and the hypothesis which might not seems very intuitive at first glance, come from.

Let us first take a look at first order Hamilton Jacobi Bellman (HJB) equations. Let φ be a solution of:

$$\begin{cases} \partial_t \varphi + H(x, \nabla_x \varphi) = 0 \text{ in } (0,\infty) \times \mathbb{R}^d, \\ \varphi|_{t=0} = \varphi_0. \end{cases}$$

If we fix $U = \nabla_x \varphi$, then we know that U is solution of

$$\partial_t U + F(x, U) \cdot \nabla_x U = G(x, U) \text{ in } (0,\infty) \times \mathbb{R}^d, \tag{3.3}$$

for $F(x, u) = D_u H(x, u)$ and $G(x, u) = -D_x H(x, u)$. In particular this shows that the propagation of monotonicity (and as such of regularity) on U is equivalent to the propagation of convexity on φ. In general, we don't have to consider a potential game (ie the gradient of a HJB equation) so long as the equations keeps the same structure. Let us also mention that equation (2.1) falls into this category of equation as we can easily see it by taking the gradient in x of the equation

$$\begin{cases} \partial_t \varphi + H(x, p, \nabla_x \varphi) + b(p) \cdot \nabla_p \varphi - \sigma \Delta_p \varphi = 0, \\ \varphi|_{t=0} = \varphi_0. \end{cases}$$

This explains why estimates previously obtained without the variable p still held true in this setting: the fundamental structure of the equations was not changed. However it is easy to see that such considerations do not hold anymore when b depends on more than just p. In this case, we would see a term in $\nabla_p \varphi$ come out in the equation satisfied by $\nabla_x \varphi$ which cannot be expressed in term of this last function alone. We remark, nonetheless, that we would obtain an equation with the same structure if we were to look at the one satisfied by $(\nabla_x \varphi, \nabla_p \varphi)$. With this insight, we may try to "complete" equation (3.1) by another. Let V be a solution of

$$\begin{cases} \partial_t V + F(x, p, U) \cdot \nabla_x V + b(x, p, U) \cdot \nabla_p V - \sigma \Delta_p V = Q(x, p, U, V) \text{ in } (0,\infty) \times \mathbb{R}^d \times \mathbb{R}^m \\ V|_{t=0} = V_0(x, p), \end{cases} \tag{3.4}$$

where Q and V_0 are $\mathbb{R}^m$ valued Lipschitz functions. Considering the equation satisfied by $\bar{U} = (U, V)$ we get

$$\partial_t \bar{U} + \bar{F}(X, \bar{U}) \cdot \nabla_X \bar{U} - \sigma \Delta_p \bar{U} = \bar{G}(X, \bar{U}),$$

with $X = (x, p)$, $\bar{F}(X, \bar{U}) = (F, b)(x, p, U)$ and $\bar{G}(X, \bar{U}) = (G, Q)(x, p, U, V)$.
Following the proof of Proposition 2.3 we expect that under the assumptions

$$\langle \bar{U}_0(X) - \bar{U}_0(Y), X - Y \rangle \geq \alpha |\bar{U}_0(X) - \bar{U}_0(X)|^2 \tag{3.5}$$

$$\begin{aligned} \langle \bar{G}(X, \bar{U}(t, X)) - \bar{G}(Y, \bar{U}(t, Y)), X - Y \rangle + \langle \bar{F}(X, \bar{U}(t, X)) - \bar{G}(Y, \bar{U}(t, Y)), \bar{U}(t, X) - \bar{U}(t, Y) \rangle \\ \geq \alpha |X - Y|^2, \end{aligned} \tag{3.6}$$

we will be able to recover global Lipschitz estimates from monotonicity on the couple (U, V) and hence on U. Let us now discuss a bit on the choice of the function V to "complete" U. So far we have stayed rather ambiguous on whether (V_0, Q) are fixed and V is to be considered as an additional unknown or V is a chosen smooth function and (V_0, Q) are defined through (3.4). We claim that both approach are possible. However because condition (3.6) is also a condition on V, it is hard to keep track of it without some a priori knowledge on V. This is why the second approach is much more practical. Condition (3.5) implies that V_0 must be monotonous in p, the simplest choice of such a function V is to take:

$$V(t, x, p) = V_0(x, p) = Ap,$$

with $A \in \mathcal{M}_m(\mathbb{R})$ a positive definite matrix. Defining Q through (3.4) we deduce that

$$Q(x, p, U) = Ab(x, p, U).$$

In light of this fact and the linearity of V, Conditions (3.5) and (3.6) reduce to

Hypothesis 3.6. There exist a matrix $A \in \mathcal{M}_m(\mathbb{R})$ and $\alpha > 0$ such that for all x, y in Ω, p, q in $\mathbb{R}^m$ and u, v in $\mathbb{R}^d$:

$$\langle U_0(x, p) - U_0(y, q), x - y\rangle + \langle p - q, A(p - q)\rangle \geq \alpha |U_0(x, p) - U_0(y, q)|^2, \tag{3.7}$$

$$\langle G(x, p, u) - G(y, q, v), x - y\rangle + \langle F(x,p, u) - F(x, q, v), u - v\rangle + \langle b(x, p, u) - b(y, q, v), (A + A^T)(p - q)\rangle \geq \alpha(|x - y|^2 + |p - q|^2). \tag{3.8}$$

Because both conditions can be expressed only in terms of $A + A^T$, we may assume without loss of generality in the rest of the paper that A is a symmetric matrix. The condition (3.8) is then simply a monotonicity condition on (G, F, Ab). If there exists a matrix A such that both conditions hold, we expect that we are able to get estimates on U. Let us insist on the fact that both conditions on U_0 and A are not trivial. For example (3.7) implies that the gradient of U in p must vanish when the gradient in x is itself 0. Of course, if we have a strong non-degeneracy condition on $D_x U_0$, say for instance, uniform coercivity in (x, p) then we may always find a suitable A. Note that for potential games, this means U_0 is the gradient of a strongly convex function in x uniformly in p. We can also give the following example that do not require coercivity of $D_x U_0$.

Example 3.7. Let $\Psi : \mathbb{R}^d \to \mathbb{R}^d$ be such that

$$\exists \alpha > 0 \quad \forall (x, y) \in \mathbb{R}^{2d} \quad \langle \Psi(x) - \Psi(y), x - y\rangle \geq \alpha |\Psi(x) - \Psi(y)|^2.$$

and $f : \mathbb{R}^m \to \mathbb{R}^d$ be Lipschitz. Then there exists $A \in \mathcal{M}_m(\mathbb{R})$ such that

$$U_0(x, p) = \Psi(x + f(p))$$

satisfies the condition (3.7).

It does becomes much simpler though, if we assume that U_0 is not a function of p, as the condition on (U_0, A) then reduces to the $\alpha-$monotonicity of U_0 in x, and because we can choose any matrix A positive semi-definite, the joint monotony of (G, F, b) become a sufficient condition to propagate estimates on U. While this was a fairly obvious observation, we still present it as we believe that it is natural in some applications to have $U_0(x, p) \equiv U_0(x)$ (remember that p models noise). Nevertheless, we don't believe this monotonicity condition on b to be unreasonable. Consider the following example in dimension 1

Example 3.8. Let

$$b(x,p) = (r_0 + \frac{r(x)}{1+p})p \text{ for } (x,p) \in \Omega \times \mathbb{R}^+,$$

with $r_0 > 0$ and $\inf_x(r_0 + r(x)) = \alpha > 0$ and assume (F, G) satisfy Hypothesis 2.1. Then if we assume either that $\|D_pF\|_\infty$ and $\|D_pG\|_\infty$ or $\|r\|_{Lip}$ are sufficiently small, there exist an $A \in \mathbb{R}^+$ such that the condition (3.8) is fulfilled.

This is an interesting example as it allows us to consider SDEs of the form

$$dp_t = p_t(r_t dt + \sigma dW_t),$$

which may be used to model market prices or other positive quantities of interest. In this example r_0 could model a flat rate for inflation while $r(x)$ correspond to the dependence of the asset price on the distribution of the players x and the term $\frac{1}{1+p}$ accounts for inertia of the price. The condition on $r+r_0$ can then be understood as the fact that the market tends to drive up the price of the asset which is arguably a strong assumption. In this particular case a value function can be defined for p in $\mathbb{R}^+$ only without additional boundary condition as 0 is an absorbing point for the process (p_t). Of course since the Brownian term depends on p this model does not quite correspond to equation (3.1). Let us state that for σ sufficiently small, this is of no concern as we delay the analysis of this more convoluted case to section 3.5.4

3.3 Gradient estimate

Following the heuristic we just described, we state and prove the following gradient estimate. As in the previous Section, we are not going to use directly this estimate as we shall work with Lipschitz solutions. We still present it so that the link with the autonomous case is made more transparent.

Proposition 3.9. *Let U be a smooth solution of the master equation (3.1) on $(0,T)$, uniformly Lipschitz continuous in (x,p). Assume that (G,F,b) is smooth, Lipschitz in all the variables and satisfies (3.8). Further assume that U_0 is such that:*

$$\exists A \in \mathcal{M}_m(\mathbb{R}), \alpha > 0, \quad \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^T \begin{pmatrix} D_xU_0 & D_pU_0 \\ 0 & A \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \geq \alpha |D_xU_0 \cdot \xi_1|^2, \tag{3.9}$$

Then U is monotone in x on $[0,T)$ and there exists constants C_x, C_p depending on G,F,b,U_0 and T only such that:

$$\forall t < T \quad \|D_xU(t,\cdot)\|_\infty \leq C_x, \|D_pU(t,\cdot)\|_\infty \leq C_p.$$

Proof. Set $\xi = (\xi_1, \xi_2) \in \mathbb{R}^{d+m}$, $W(t,x,p,\xi) = U(t,x,p) \cdot \xi_1$ and

$$\begin{aligned} Z(t,x,p,\xi) &= \xi_1 \cdot \nabla_x(U \cdot \xi_1) + \xi_2 \cdot \nabla_p(U \cdot \xi_1) - \beta(t)|\nabla_x U \cdot \xi_1|^2 + \xi_2 \cdot A\xi_2 \\ &= \xi_1 \cdot \nabla_x W + \xi_2 \cdot \nabla_p W - \beta(t)|\nabla_x W|^2 + \xi_2 \cdot A\xi_2 \end{aligned}$$

Remark that for $\beta_0 \leq \alpha$ we have $Z|_{t=0} \geq 0$ by assumption on U_0.
Let us now show that Z stays non-negative over time for a well chosen function β. Looking at the

equation satisfied by Z we get

$$\begin{aligned}
&\partial_t Z - \sigma\Delta_p Z + b\cdot\nabla_p Z + F\cdot\nabla_x Z\\
&+(-D_p b\xi_2 + D_x b\xi_1 - D_u b\nabla_x W + 2\beta D_x b\nabla_x W)\cdot\nabla_{\xi_2} Z + (\nabla_x W D_u F + \nabla_p W D_u b - \xi_1 D_u G)\cdot\nabla_{\xi_1} Z\\
&= 2\beta\sigma|D^2_{xp}W|^2 - \frac{d\beta}{dt}|\nabla_x W|^2\\
&+(\xi_1 D_x G\xi_1 - \xi_1 D_u G\nabla_x W - \nabla_x W D_x F\xi_1 + \nabla_x W D_u F\nabla_x W)\\
&+(\nabla_x W D_p F - \xi_1 D_p G)\cdot\xi_2 + 2\xi_2\left(AD_p b\xi_2 + AD_x b\xi_1 - AD_u b\nabla_x W\right)\\
&+2\beta\left(-2\xi_2 AD_x b\nabla_x W + \nabla_x W D_x F\nabla_x W - \xi_1 D_x G\nabla_x W\right)
\end{aligned}$$

Remark that

$$\begin{aligned}
&(\xi_1 D_x G\xi_1 - \xi_1 D_u G\nabla_x W - \nabla_x W D_x F\xi_1 + \nabla_x W D_u F\nabla_x W)\\
&+(\nabla_x W D_p F - \xi_1 D_p G)\cdot\xi_2 + 2\xi_2\left(AD_p b\xi_2 + AD_x b\xi_1 - AD_u b\nabla_x W\right)\\
&=\begin{pmatrix} -\xi_1\\ \nabla_x W\\ -\xi_2\end{pmatrix}^T\begin{pmatrix} D_xG & D_uG & D_pG\\ D_xF & D_uF & D_pF\\ 2AD_xb & 2AD_ub & 2AD_pb\end{pmatrix}\begin{pmatrix} -\xi_1\\ \nabla_x W\\ -\xi_2\end{pmatrix} \geq \alpha(|\xi_1|^2+|\xi_2|^2)
\end{aligned}$$

by (3.8). Hence, we end up with:

$$\begin{aligned}
&\partial_t Z - \sigma\Delta_p Z + b\cdot\nabla_p Z + F\cdot\nabla_x Z\\
&+(-D_p b\xi_2 + D_x b\xi_1 - D_u b\nabla_x W + 2\beta D_x b\nabla_x W)\cdot\nabla_{\xi_2} Z + (\nabla_x W D_u F + \nabla_p W D_u b - \xi_1 D_u G)\cdot\nabla_{\xi_1} Z\\
&\geq (\alpha-\beta(t))\left(|\xi_1|^2+|\xi_2|^2\right) + \left(4\beta\|2AD_xb\|^2_\infty + 4\beta\|D_xG\|^2_\infty + 2\beta\|D_xF\|_\infty - \frac{d\beta}{dt}\right)|\nabla_x W|^2.
\end{aligned}$$

Let now β_0 be chosen so that

$$\beta_0 \leq \alpha$$

and $\beta(t) = \beta_0 e^{-\lambda_\beta t}$ with

$$\lambda_\beta \geq 4\|2AD_xb\|^2_\infty + 4\|D_xG\|^2_\infty + 2\|D_xF\|_\infty.$$

For such a β we have

$$\begin{aligned}
&\partial_t Z - \sigma\Delta_p Z + b\cdot\nabla_p Z + F\cdot\nabla_x Z\\
&+(-D_p b\xi_2 + D_x b\xi_1 - D_u b\nabla_x W + 2\beta D_x b\nabla_x W)\cdot\nabla_{\xi_2} Z + (\nabla_x W D_u F + \nabla_p W D_u b - \xi_1 D_u G)\cdot\nabla_{\xi_1} Z \geq 0.
\end{aligned}$$

The property $Z \geq 0$ follows from the same argument we used in Proposition 2.3. It only remains to show that this gives estimates on the gradient of the solution U. Remark first that by taking $\xi_2 = 0$, we get back

$$\forall t\in[0,T), \langle\xi_1, D_xU, \xi_1\rangle - \beta(t)\|\langle D_xU,\xi_1\rangle\|^2 \geq 0.$$

As in Proposition 2.3, this implies:

$$\|D_xU(t,\cdot)\|_\infty \leq \frac{1}{\beta_t} = \frac{1}{\beta_0}e^{\lambda_\beta t} \leq \frac{1}{\beta_0}e^{\lambda_\beta T}.$$

For the gradient of U in p, we know that

$$\forall t\in[0,T), \xi_1,\xi_2, \quad \langle\xi_1, D_xU,\xi_1\rangle + \langle\xi_2, D_pU,\xi_1\rangle + \xi_2\cdot A\xi_2 \geq 0.$$

Let $t < T$ be fixed, $\xi_1 = xr_i^x$ and $\xi_2 = yr_j^p$ where (r_i) indicate basis elements. Hence, for all $x, y \in \mathbb{R}$,

$$x^2 \left(\frac{\partial U_i}{\partial x_i}\right) + xy\frac{\partial U_i}{\partial p_j} + A_{jj}y^2 \geq 0.$$

Thus

$$\left(\frac{\partial U_i}{\partial p_j}\right)^2 \leq 4A_{jj}\left(\frac{\partial U_i}{\partial x_i}\right) \leq \frac{4A_{jj}}{\beta_0}e^{\lambda_\beta T}.$$

□

3.4 Existence of solutions

3.4.1 Main result

We are now going to proceed as we did in Section 2 to justify the estimate of Proposition 3.9 for Lipschitz solutions of the master equation (3.1). The heuristic presented earlier hints at the idea of studying

$$Z(t,x,y,p,q) = \langle U(t,x,p) - U(t,y,q), x-y\rangle + \langle p-q, A, p-q\rangle - \beta(t)|U(t,x,p) - U(t,y,q)|^2. \quad (3.10)$$

Due to the doubling of variable we did in (p,q), the equation satisfied for smooth U (and hence smooth Z) is slightly different. Indeed, in this case Z would be solution of:

$$\begin{aligned} &\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - \sigma\Delta_p Z - \sigma\Delta_q Z \\ &= \langle G^x - G^y, x-y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p-q\rangle \\ &- 2\beta\langle G^x - G^y, U^x - U^y\rangle + 2\sigma\beta(|D_p U^x|^2 + |D_q U^y|^2) - 4\sigma TrA - \frac{d\beta}{dt}|U^x - U^y|^2, \end{aligned} \quad (3.11)$$

where we used the notation $U^x = U(t,x,p)$, $U^y = U(t,y,q)$, $F^x = F(x,p,U^x)$ and so on. However as is, equation (3.11) does not seem suited for the use of a comparison principle. Even under the assumption that (3.8) holds for some $A \in \mathcal{M}_m(\mathbb{R})$, it seems hard to get rid of the term $-4\sigma Tr(A)$ which is always negative as A must be a positive semi-definite matrix. However, remark that

$$D_{pq}Z = -2A + 2\beta D_q U^y D_p U^x.$$

Hence, for $B = \begin{pmatrix} I_m & I_m \\ I_m & I_m \end{pmatrix}$ we obtain

$$\sigma Tr(BD^2_{(p,q)}Z) = \sigma\Delta_p Z + \sigma\Delta_q Z - 4\sigma TrA + 4\beta\sigma Tr(D_p U^x D_q U^y)$$

where $D^2_{(p,q)}Z$ is the hessian matrix of Z in (p,q) only. It follows that the equation satisfied by Z might be rewritten as:

$$\begin{aligned} &\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - \sigma Tr(BD^2_{(p,q)}Z) \\ &= \langle G^x - G^y, x-y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p-q\rangle \\ &- 2\beta\langle G^x - G^y, U^x - U^y\rangle + 2\sigma\beta(|D_p U^x - D_q U^y|^2) - \frac{d\beta}{dt}|U^x - U^y|^2. \end{aligned} \quad (3.12)$$

As $B \geq 0$ it is now more clear that we can use a comparison principle for Equation (3.12). As in Section 2 we may now state the following key Lemma.

Lemma 3.10. *Assume that (F, G, b, U_0) is Lipschitz continuous. Let U be a Lipschitz solution of the master equation (3.1) on $[0, T)$ for $T > 0$ and Z be defined as in (3.10). Then Z is a viscosity supersolution of:*

$$\begin{aligned}
&\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - \sigma Tr(BD^2_{(p,q)} Z) \\
&\geq \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle \\
&- 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2,
\end{aligned} \tag{3.13}$$

on $(0, T) \times \Omega^2 \times \mathbb{R}^{2m}$.

Proof. We first show that Z satisfies the following inequality

$$\begin{aligned}
Z(t, x, y, p_0, q_0) \geq\ & \mathbb{E}\Big[Z(s, X^x_{t-s}, Y^y_{t-s}, p^{p_0}_{t-s}, q^{q_0}_{t-s}) - \beta(s)\left\|\int_0^{t-s}(G^{x,p_0}_u - G^{y,q_0}_u)du\right\|^2\Big] \\
&+ \mathbb{E}\Big[\int_0^{t-s}\left(\langle G^{x,p_0}_u - G^{y,q_0}_u, x - y\rangle + \langle F^{x,p_0}_u - F^{y,q_0}_u, U^{x,p_0}_s - U^{y,q_0}_s\rangle\right)du\Big] \\
&+ \mathbb{E}\Big[\int_0^{t-s}\langle p - q, 2A(b^{x,p_0}_u - b^{y,q_0}_u)\rangle du\Big] \\
&- (\beta(t) - \beta(s))\|U(t, x, p_0) - U(t, y, q_0)\|^2 - 2\beta(s)\mathbb{E}\Big[\int_0^{t-s}\langle U^{x,p_0}_s - U^{y,q_0}_s, G^{x,p_0}_u - G^{y,q_0}_u\rangle\Big],
\end{aligned} \tag{3.14}$$

with

$$\begin{cases}
dX_s = -F(X_s, p_s, U(t-s, X_s, p_s)ds, \\
dY_s = -F(Y_s, q_s, U(t-s, Y_s, q_s)ds, \\
dp_s = -b(X_s, p_s, U(t-s, X_s, p_s)ds + \sqrt{2\sigma}dW_s, \\
dq_s = -b(Y_s, q_s, U(t-s, Y_s, q_s)ds + \sqrt{2\sigma}dW_s, \\
U^{x,p_0}_s = U(t-s, X^x_s, p^{p_0}_s), \\
G^{x,p_0}_s = G(X^x_s, p^{p_0}_s, U^{x,p_0}_s),
\end{cases}$$

and the notation F^{x,p_0}_s and b^{x,p_0}_s being analogous to the ones for G. As usual in this kind of coupling techniques, or in comparison questions in the theory of viscosity solutions, it is fundamental for the two processes $(p_s)_{s\geq 0}, (q_s)_{s\geq 0}$ to be generated with the same Brownian motion. Indeed if the computations were done for two independent Brownian motions, we would end up with Z being a supersolution of (3.11) (i.e there will be a term in $Tr(A)$). The link between (3.11) and (3.14) is obvious for smooth Z but it does not seems so easy to switch from one to the other in the viscosity sense.

Let

$$W(t, x, y, p, q) = \langle U(t, x, p) - U(t, y, q), x - y\rangle.$$

It follows that

$$\begin{aligned}
Z(t, x, y, p, q) = \underbrace{W(t, x, y, p, q)}_{I} + \underbrace{\langle p - q, A(p - q)\rangle}_{II} - \beta(s)\underbrace{|U(t, x, p) - U(t, y, p)|^2}_{III} \\
- (\beta(t) - \beta(s))|U(t, x, p) - U(t, y, p)|^2.
\end{aligned}$$

For II, we can write

$$\begin{aligned}\langle p_0 - q_0, A, p_0 - q_0\rangle &= \mathbb{E}\Big[\langle p^{p_0}_{t-s} - q^{q_0}_{t-s}, A, p^{p_0}_{t-s} - q^{q_0}_{t-s}\rangle + \int_0^{t-s}\langle p - q, A + A^T, b^{x,p_0}_u - b^{y,q_0}_u\rangle du\Big] \\ &\quad + \mathbb{E}\Big[\langle \int_0^{t-s}(p^{p_0}_u - q^{q_0}_u)du, A, \int_0^{t-s}(p^{p_0}_u - q^{q_0}_u)du\rangle\Big] \\ &\geq \mathbb{E}\Big[\langle p^{p_0}_{t-s} - q^{q_0}_{t-s}, A, p^{p_0}_{t-s} - q^{q_0}_{t-s}\rangle + \int_0^{t-s}\langle p - q, A + A^T, b^{x,p_0}_u - b^{y,q_0}_u\rangle du\Big].\end{aligned}$$

The terms I and III are treated as in Lemma 2.12. Hence, because we assumed that A is symmetric, using the inequalities satisfied by those 3 terms, Z satisfies (3.14).

As in the autonomous case, considering a test function φ such that

$$(Z - \varphi)(t, x, y, p, q) = \min(Z - \varphi) = 0,$$

we get that at (t, x, y, p, q), φ satisfies

$$\begin{aligned}&\partial_t\varphi + F^x \cdot \nabla_x\varphi + F^y \cdot \nabla_y\varphi + b^x \cdot \nabla_p\varphi + b^y \cdot \nabla_q\varphi - \sigma Tr(BD^2_{(p,q)}\varphi) \\ &\geq \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle \\ &- 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2.\end{aligned}$$

Observe that the term $\sigma Tr(BD^2_{(p,q)}\varphi)$ appears as a direct consequence of Ito's lemma because we used the same Brownian motion for p and q.

One can also check that for smooth U, the quadratic variation of the martingale term we discarded in the inequality satisfied by III is exactly

$$\langle M\rangle_t = 2\sigma\int_0^t |D_pU^{x,p_0}_s - D_qU^{y,q_0}_s|^2 ds,$$

which obviously corresponds to the missing term in $|D_pU(t, x, p) - D_qU(t, y, q)|^2$ in Equation (3.13) compared to (3.12). This causes Z to only be a supersolution. □

Armed with this Lemma, we may now state a global existence result for Lipschitz solutions of the master equation (3.1). We recall that there is always uniqueness of a maximal Lipschitz solution so it will imply a well-posedness result.

Theorem 3.11. *Assume that (U_0, G, F, b) is Lipschitz continuous in all the variables and further assume that Hypothesis 3.6 holds for some $\alpha > 0$ and $A \in \mathcal{M}_m(\mathbb{R})$. Then there exists a unique global Lipschitz solution of the master equation* (3.1).

Proof. Since (U_0, G, F, b) is Lipschitz continuous, we have existence of a Lipschitz solution at least on $[0, T_c)$ for some $T_c > 0$. Let $T < T_c$, thanks to Lemma 3.10, we already know that

$$Z(t, x, y, p, q) = \langle U(t, x, p) - U(t, y, q), x - y\rangle + \langle p - q, A, p - q\rangle - \beta(t)|U(t, x, p) - U(t, y, q)|^2$$

is a viscosity supersolution of Equation (3.13) on $(0, T)$. It only remains to choose a smooth function β such that both $Z_{t=0} \geq 0$ and

$$\begin{aligned}&\langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle \\ &-2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2 \geq 0\end{aligned}$$

holds. The fact that (F, G, b) satisfies (3.8) yields

$$\begin{aligned}
&\langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle - 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2 \\
&\geq \alpha\left(|x - y|^2 + |p - q|^2\right) - \frac{d\beta}{dt}|U^x - U^y|^2 - 2\beta\langle G^x - G^y, U^x - U^y\rangle \\
&\geq (\alpha - \beta)\left(|x - y|^2 + |p - q|^2\right) - \left(\frac{d\beta}{dt} + \beta(2\|D_u G\|_\infty + \|D_x G\|_\infty^2 + \|D_p G\|_\infty^2)\right)|U^x - U^y|^2.
\end{aligned}$$

We also know that taking $\beta_0 \leq \alpha$ implies that $Z_{t=0} \geq 0$ as U_0 satisfies (3.7). For $\beta(t) = \beta_0 e^{-\lambda_\beta t}$ with $\lambda_\beta \geq 2\|D_u G\|_\infty + \|D_x G\|_\infty^2 + \|D_p G\|_\infty^2$ and $\beta_0 \leq \alpha$ both inequalities are satisfied.

Finally we get that Z is a viscosity supersolution of

$$\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - \sigma Tr(BD^2_{(p,q)} Z) \geq 0,$$

on $(0, T)$ with $Z_{t=0} \geq 0$. We may then conclude that Z stays non-negative on $[0, T)$ with Theorem 2.5 of [19].

Evaluating Z for $p = q$ we recover an estimate on the Lipschitz norm of U in x on $[0, T)$ as in the autonomous case. For $\|D_p U\|_\infty$ we know that

$$\langle U(t, x, p) - U(t, y, q), x - y\rangle + \langle p - q, A, p - q\rangle \geq 0.$$

Taking $x = y + \lambda e_i$ and using the fact that U is Lipschitz in x we get

$$\lambda^2\|D_x U\|_\infty + \lambda(U^i(t, x, p) - U^i(t, x, q)) + \langle p - q, A, p - q\rangle \geq 0,$$

which implies that

$$\forall (t, x, p, q) \in [0, T) \times \Omega \times \mathbb{R}^{2m}, \quad (U^i(t, x, p) - U^i(t, x, q))^2 - 4\|D_x U\|_\infty \langle p - q, A, p - q\rangle \leq 0.$$

Denoting by $\lambda_a > 0$ be the biggest eigenvalue of A, we obtain

$$\forall t < T, \quad \|D_p U(t, \cdot)\|_\infty \leq 2\sqrt{\|D_x U(t, \cdot)\|_\infty \lambda_a}.$$

Suppose now that the maximum time of existence satisfies $T_c < +\infty$. By the local existence Theorem 2.4 of [8] we must have

$$\lim_{t \to T_c} \|D_{(x,p)} U(t, \cdot)\|_\infty = +\infty.$$

However we have an uniform bound in time on the Lipschitz norm of U in (x, p) for $t < T_c$. As a consequence this is absurd, hence $T_c = +\infty$ and the result is proved. □

Remark 3.12. Take note that when $D_x b, D_u b = 0$ we recover the hypothesis for the autonomous case. Let Z be defined as in Section 2 by

$$Z(t, x, y, p) = \langle U(t, x, p) - U(t, y, p), x - y\rangle - \beta(t)|U(t, x, p) - U(t, y, p)|^2.$$

Define

$$M = \min_{[0,T] \times \Omega^2 \times \mathbb{R}^{2m}} Z,$$

and consider

$$Z^\varepsilon(t,x,y,p,q) = \langle U(t,x,p)-U(t,y,q), x-y\rangle - \beta(t)|U(t,x,p)-U(t,y,q)|^2 + \frac{1}{2\varepsilon}|p-q|^2 + \alpha|p|^2+\alpha|q|^2.$$

Looking at point of minimum of Z^ε we can show that as ε tends to 0 there exists a subsequence $(t_{\varepsilon_n}, x_{\varepsilon_n}, y_{\varepsilon_n}, p_{\varepsilon_n}, q_{\varepsilon_n})_{n\geq 0}$ such that

$$\begin{cases} (t_{\varepsilon_n}, x_{\varepsilon_n}, y_{\varepsilon_n}, p_{\varepsilon_n}, q_{\varepsilon_n}) \to (t^*, x^*, y^*, p^*, q^*) \in [0,T]\times\Omega^2\times\mathbb{R}^{2m} \\ \lim\limits_{n\to+\infty} Z^{\varepsilon_n}(t_{\varepsilon_n}, x_{\varepsilon_n}, y_{\varepsilon_n}, p_{\varepsilon_n}, q_{\varepsilon_n}) = M \\ \frac{1}{\varepsilon_n}|p_{\varepsilon_n} - q_{\varepsilon_n}|^2 \to 0 \\ |p_{\varepsilon_n} - q_{\varepsilon_n}| \to 0 \end{cases}$$

Because $\frac{1}{\varepsilon}\langle b(p)-b(q), p-q\rangle \leq \frac{C}{\varepsilon}|p-q|^2$, we may conclude by taking first $\varepsilon_n \to 0$ and then $\alpha \to 0$ that we only need conditions (3.7) and (3.8) to hold for $p=q$ to deduce that $M \geq 0$. This way, we indeed recover the hypothesis we had in the autonomous case.

3.5 Extensions

Now that we have established the main setting, let us discuss on some extensions

3.5.1 On the strong monotonicity of (G,F,b)

As we stated it at the beginning the strong monotonicity assumptions on (G,F,b), condition (3.8) can be weakened quite a bit. Following the proof of Theorem 3.11, the term we control with strong monotonicity depends on G only. This means that strong monotonicity is only needed whenever G is not degenerate. In particular if G does not depend on p then strong monotonicity in this variable is not needed. We may summarize this observation in the following corollary:

Corollary 3.13. *Suppose there exist a matrix $A \in \mathcal{M}_m(\mathbb{R})$ and $\alpha > 0$ such that for all x, y in Ω, p, q in $\mathbb{R}^m$ and u, v in $\mathbb{R}^d$ there is $t = t(x,y,p,q,u,v) \in [0,1]$ such that*

$$\langle U_0(x,p) - U_0(y,q), x-y\rangle + \langle p-q, A(p-q)\rangle \geq \alpha|U_0(x,p)-U_0(y,q)|^2,$$

$$\begin{aligned}\langle G(x,p,u)-G(y,q,v), x-y\rangle + \langle F(x,p,u)-F(x,q,v), u-v\rangle + \langle b(x,p,u)-b(y,q,v), (A+A^T)(p-q)\rangle \\ \geq \alpha|G(x,p,tu+(1-t)v) - G(y,q,tu+(1-t)v)|^2.\end{aligned}$$

Then there exist a Lipschitz solution to (3.1) *on any time interval.*

Proof. The only difference lies in how we deal with the term

$$\langle G(x,p,u) - G(y,q,v), u-v\rangle.$$

Observe that there is no difficulty in adapting the proof of Theorem 3.11 as for any $t \in [0,1]$,

$$\langle G(x,p,u)-G(y,q,v), u-v\rangle \leq \frac{1}{2}|G(x,p,tu+(1-t)v)-G(y,q,tu+(1-t)v)|^2 + \left(\frac{1}{2} + \|D_u G\|_\infty\right)|u-v|^2.$$

□

Remark 3.14. The addition of this $t \in [0,1]$ in the monotonicity condition might seems cryptic at first. Essentially it is a way of translating the fact that we do not need strong monotonicity along the variable associated to U.

Similarly, the gradient estimate of Proposition 3.9 may be stated under weaker monotonicity assumptions. Following the proof it is evident that we only really need (G, F, b) to satisfy

$$\forall \nu \in \mathbb{R}^{2d+m} \quad \nu^T \begin{pmatrix} D_x G & D_u G & D_p G \\ D_x F & D_u F & D_p F \\ 2AD_x b & 2AD_u b & 2AD_p b \end{pmatrix} \nu \geq \alpha \left| \begin{pmatrix} D_x G & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D_x b \end{pmatrix} \nu \right|^2,$$

for some $\alpha > 0$ and a symmetric matrix A. Obviously, this condition is not equivalent to the one we stated in Corollary 3.13. We believe the reason for that to be sligtly technical, indeed in general condition (3.7)

$$\langle U_0(t,x,p) - U_0(t,y,q), x - y \rangle + \langle p - q, A, p - q \rangle \geq \alpha |U_0(t,x,p) - U_0(t,y,q)|^2$$

and (3.9)

$$\forall (\xi_1, \xi_2) \quad \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^T \begin{pmatrix} D_x U_0 & D_p U_0 \\ 0 & A \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \geq \alpha |D_x U_0 \cdot \xi_1|^2,$$

are not equivalent even though we believe them both to be natural assumptions. They are however both sufficient condition to recover existence and uniqueness of solution to the master equation under the strong monotonicity condition (3.8). Indeed notice that that whenever (3.9) is fulfilled for some (α, A), for any $\varepsilon > 0$ there exist an $\alpha_\varepsilon > 0$ such that (3.7) holds for the couple $(\alpha_\varepsilon, \varepsilon I_d + A)$. The idea being that whenever strong monotonicity of (G, F, Ab) holds, the addition of this small perturbation ε is non consequential and we may recover bounds on the gradient as in the proof of Theorem 3.11 without much additional difficulty.

3.5.2 A uniqueness result

Let us first present a uniqueness property. It may seem surprising that we are now concerned with uniqueness while we insisted that it always holds in the class of Lipschitz solutions. Even though in this context uniqueness follows from the Lipschitz regularity of solutions we provide a proof of uniqueness to indicate why developments similar to monotone solutions [4] are also immediate for equation (3.1). Monotone solutions only require the continuity of the value function U and yields uniqueness and stability properties under monotonicity assumptions on the coefficients.

Lemma 3.15. *Assume that U_0 and (G, F, b) satisfy:*

$$\exists A \in \mathcal{M}(\mathbb{R}^m), \quad \forall (x, y, u, v, p, q) \in \Omega^2 \times \mathbb{R}^{2d} \times \mathbb{R}^{2m},$$

$$\langle U_0(x,p) - U_0(y,q), x - y \rangle + \langle p - q, A(p - q) \rangle \geq 0,$$

$$\langle G(x,p,u) - G(y,q,v), x - y \rangle + \langle F(x,p,u) - F(x,q,v), u - v \rangle + \langle b(x,p,u) - b(y,q,v), A(p - q) \rangle \geq 0.$$

Then there exists at most one smooth solution of the master equation (3.1).

Proof. Suppose there exists two smooth solutions U and V on $[0,T]$ for some $T>0$ and take

$$Z(t,x,y,p,q) = \langle U(t,x,p) - V(t,y,q), x-y\rangle + \langle p-q, A, p-q\rangle.$$

First, the assumption formulated on U_0 ensure that $Z_0 \geq 0$. Since U and V are smooth, so is Z. Using the equations satisfied by U and V, we get that Z is solution of

$$\begin{aligned}&\partial_t Z + F^x\cdot\nabla_x Z + F^y\cdot\nabla_y Z + b^x\cdot\nabla_p Z + b^y\cdot\nabla_q Z - \sigma Tr(BD^2_{(p,q)}Z)\\ &= \langle G^x - G^y, x-y\rangle + \langle F^x - F^y, U^x - V^y\rangle + \langle 2A(b^x-b^y), p-q\rangle,\end{aligned}$$

where we used the notation $G^x = G(x,p,U(t,x,p))$, $G^y = G(y,q,V(t,y,q))$, similarly for F^x, F^y, b^x, b^y and $B = \begin{pmatrix} I_m & I_m \\ I_m & I_m\end{pmatrix}$.
Finally, the monotonicity condition on (G,F,Ab) implies (recall that Z is smooth) that Z satisfies

$$\begin{cases} \partial_t Z + F^x\cdot\nabla_x Z + F^y\nabla_y Z + b^x\nabla_p Z + b^y\nabla_q Z - \sigma Tr(BD^2_{(p,q)}Z) \geq 0 \quad \forall t\in(0,T),\\ Z_{t=0}\geq 0.\end{cases}$$

By a comparison principle with 0 we deduce that $Z\geq 0$. Fixing $p=q\in\mathbb{R}^m$ yields that for all $t\geq 0, x,y\in\Omega$

$$\langle U(t,x,p) - V(t,y,p), x-y\rangle \geq 0.$$

Take $x\in\overset{\circ}{\Omega}$ and set $y = x+\varepsilon\xi$ with $\|\xi\|=1$ such that for ε sufficiently small $y\in\Omega$. Dividing by ε and taking the limit as ε tends to 0 in the previous inequality gives

$$\forall\|\xi\|\leq 1 \quad \langle U(t,x,p) - V(t,x,p), \xi\rangle \geq 0,$$

which implies $U(t,x,p) = V(t,x,p)$. By continuity the equality holds for any $x\in\Omega$. □

3.5.3 Recovering monotonicity estimates under different assumptions

It was first observed in [23] that we may trade the $\alpha-$monotonicity of U_0 against stronger assumptions on the coefficients of the equation. Following this idea we state the following Theorem.

Theorem 3.16. *Suppose (U_0,G,F,b) is Lipschitz continuous. Further assume that*

$$\begin{aligned}&\forall(x,y,p,q,u,v)\in\Omega^2\times\mathbb{R}^{2m}\times\mathbb{R}^{2d},\\ &\langle U_0(x,p) - U_0(y,q), x-y\rangle + \langle p-q, A(p-q)\rangle \geq 0,\end{aligned} \tag{3.15}$$

$$\begin{aligned}&\langle G(x,p,u) - G(y,q,v), x-y\rangle + \langle F(x,p,u) - F(x,q,v), u-v\rangle + \langle b(x,p,u) - b(y,q,v), 2A(p-q)\rangle\\ &\geq \alpha|u-v|^2.\end{aligned} \tag{3.16}$$

Then there exists a global Lipschitz solution.

Proof. We only provide a sketch of the proof as we believe the extension to be natural. Define

$$\begin{aligned}Z(t,x,y,p,q) = \langle U(t,x,p) - U(t,y,q), x-y\rangle + \langle p-q, A, p-q\rangle - \beta(t)|U(t,x,p) - U(t,y,q)|^2\\ + \gamma(t)(|x-y|^2 + |p-q|^2).\end{aligned}$$

By a proof similar to the one of Lemma 3.10, Z is a viscosity supersolution of

$$\begin{aligned}
&\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - \sigma Tr(BD^2_{(p,q)}Z) \\
&\geq \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle \\
&- 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2 \\
&+ 2\gamma(t)(\langle F^x - F^y, x - y\rangle + \langle (b^x - b^y), p - q\rangle) + \frac{d\gamma}{dt}(|x - y|^2 + |p - q|^2).
\end{aligned}$$

For $\max(\|D_x U_0\|^2_\infty, \|D_p U_0\|^2_\infty)\beta_0 \leq \gamma_0$, (3.15) implies that $Z_{t=0} \geq 0$. Using (3.16) we know that if we can find a couple (β, γ) such that

$$\begin{cases} \frac{d\gamma}{dt} \geq 2\gamma(\|F\|_{Lip} + \|b\|_{Lip}) + \beta\|G\|_{Lip}, \\ \alpha \geq \frac{d\beta}{dt} + 2\beta\|G\|_{Lip} + \gamma(\|F\|_{Lip} + \|b\|_{Lip}), \end{cases}$$

on $(0, T)$ for the Lipschitz semi-norm $\|F\|_{Lip} = \|D_x F\|_\infty + \|D_p F\|_\infty + \|D_u F\|_\infty$, then Z will be a viscosity supersolution of

$$\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - \sigma Tr(BD^2_{(p,q)}Z) \geq 0$$

on $(0, T) \times \Omega^2 \times \mathbb{R}^{2m}$. To find such a couple, it suffices to take $\beta(t) = \frac{1}{\max(\|D_x U_0\|^2_\infty, \|D_p U_0\|^2_\infty)}\gamma(t)$ with

$$\gamma(t) = \gamma_0 e^{\lambda(t-T)},$$

for λ sufficiently big and γ_0 sufficiently small chosen independently of T. We may then recover $Z \geq 0$ on $[0, T)$ by a comparison principle, which of course implies an estimate on the Lipschitz norm of U in (x, p). □

3.5.4 The case of a non-constant volatility

Let us also comment on monotonocity estimates for the more general equation

$$\begin{gathered}
\partial_t U + F(x, p, U) \cdot \nabla_x U + b(x, p, U) \cdot \nabla_p U - Tr(\Gamma(x, p, U)D^2_p U) = G(x, p, U) \text{ in } (0, T) \times \Omega \times \mathbb{R}^m, \\
U(0, x, p) = U_0(x, p) \text{ in } \Omega \times \mathbb{R}^m,
\end{gathered} \tag{3.17}$$

for $\Gamma(x, p, U) = \Sigma(x, p, U)\Sigma^T(x, p, U)$ with $\Sigma : \Omega \times \mathbb{R}^m \times \mathbb{R}^d \to \mathcal{M}_m(\mathbb{R})$. Let us first remark that when Σ is Lipschitz in all variables, we may still define a Lipschitz solution of equation (3.17). Let us now state a global existence result for Lipschitz solution of this equation

Proposition 3.17. *Suppose U_0, F, G, b, Σ are Lipschitz continuous in all variables. If they furthermore satisfy the following*

$$\begin{aligned}
&\forall (x, y, p, q, u, v) \in \Omega^2 \times \mathbb{R}^{2m} \times \mathbb{R}^{2d} \\
&\langle U_0(x, p) - U_0(y, q), x - y\rangle + \langle p - q, A(p - q)\rangle \geq \alpha|U_0(x, p) - U_0(y, q)|^2, \\
&\langle G(x, p, u) - G(y, q, v), x - y\rangle + \langle F(x, p, u) - F(x, q, v), u - v\rangle + \langle b(x, p, u) - b(y, q, v), 2A(p - q)\rangle, \\
&\geq \alpha(|x - y|^2 + |p - q|^2) + Tr\left((\Sigma(x, p, u) - \Sigma(y, q, v))^T 2A\,(\Sigma(x, p, u) - \Sigma(y, q, v))\right),
\end{aligned} \tag{3.18}$$

for a matrix $A \in \mathcal{M}_m(\mathbb{R})$, then there exists a global Lipschitz solution of equation (3.17) *on any time interval.*

Proof. We only sketch the proof as it follows from the one of Theorem 3.11. Defining Z as in this Theorem, for a smooth U, it is a solution of

$$\begin{aligned}
&\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - Tr(\Gamma^x D_p^2 Z) - Tr(\Gamma^y D_q^2 Z) \\
&= \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle \\
&- 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2 \\
&+ 2\beta\left(|\Sigma^x D_p U^x|^2 + |\Sigma^y D_q U^y|^2\right) - 2Tr((\Gamma^x + \Gamma^y)A),
\end{aligned}$$

with the notation $|M|^2 = Tr(M^T M) = \sum_{i,j} M_{ij}^2$ for a square matrix and the same notation U^x, F^x and so on we used in earlier proofs. Remark that

$$D_p(D_q Z) = -2A + 2\beta(D_p U^x)^T D_q U^y.$$

We are also going to need the formula

$$(\Gamma^x + \Gamma^y)A = (\Sigma^x - \Sigma^y)(\Sigma^x - \Sigma^y)^T A + \left(\Sigma^x(\Sigma^y)^T + \Sigma^y(\Sigma^x)^T\right) A.$$

Thanks to those two equations, we may rewrite the PDE satisfied by Z as

$$\begin{aligned}
&\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - Tr(\Lambda D^2_{(p,q)} Z) \\
&= \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle \\
&- 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2 \\
&+ 2\beta|\Sigma^x D_p U^x - \Sigma^y D_q U^y|^2 - 2Tr\left((\Sigma^x - \Sigma^y)^T A(\Sigma^x - \Sigma^y))\right),
\end{aligned}$$

for $\Lambda = \begin{pmatrix} \Gamma^x & \Sigma^x(\Sigma^y)^T \\ \Sigma^y(\Sigma^x)^T & \Gamma^y \end{pmatrix} \geq 0$.

Adapting Lemma 3.13, we may prove that when U is not smooth, Z is still a viscosity supersolution of

$$\begin{aligned}
&\partial_t Z + F^x \cdot \nabla_x Z + F^y \cdot \nabla_y Z + b^x \cdot \nabla_p Z + b^y \cdot \nabla_q Z - Tr(\Lambda D_q^2 Z) \\
&\geq \langle G^x - G^y, x - y\rangle + \langle F^x - F^y, U^x - U^y\rangle + \langle 2A(b^x - b^y), p - q\rangle \\
&- 2\beta\langle G^x - G^y, U^x - U^y\rangle - \frac{d\beta}{dt}|U^x - U^y|^2 - 2Tr\left((\Sigma^x - \Sigma^y)^T A(\Sigma^x - \Sigma^y))\right).
\end{aligned}$$

It then only remains to use the assumption (3.18) and to conclude for a well chosen β. □

Remark 3.18. This shows that when Σ depends on (x, p) only, there always exists a global Lipschitz solution under Hypothesis 3.6 for $\|\Sigma\|_{Lip}$ sufficiently small.

3.5.5 Stationary Master equation

We could add a discount term λU in equation (3.1) and the analysis would stay unchanged. In fact such term is regularizing for the equation as soon as $\lambda > 0$. However the presence of such factor seems to be a necessary condition to be able to talk about steady states of equation (3.1) and the associated hypoelliptic PDE

$$\lambda U + F(x, p, U) \cdot \nabla_x U + b(x, p, U) \cdot \nabla_p U - \sigma\Delta_p U = G(x, p, U) \text{ for } x \in \Omega, p \in \mathbb{R}^m. \tag{3.19}$$

Without the additional variable p, Equation (3.19) was studied in [4]. Let us also briefly mention that this equation has intrinsic significance independently of its interpretation as the large time limit of the solution of its parabolic counterpart. Indeed, it has been observed that the mean field limit of general economic equilibrium may fall into this category of equation [6, 10, 1]. Just as we did in the proof of Lemma 3.15, it is possible to obtain uniqueness of solutions to this equation by using the auxiliary function

$$Z(x,p) = \langle U(x,p) - U(y,q), x-y\rangle + \langle A(p-q), p-q\rangle.$$

We do not details the proof all over again as it follows naturally from previously introduced arguments. However, contrary to the parabolic case, there seems to be no other way than looking at the equation satisfied by the gradient of U to get Lipschitz estimates. We believe this may be done without much difficulties by taking inspiration from [4] and Proposition 3.9 so we do not prove it here. Let us however state this result for the associated parabolic PDE and explain how it relates to long time convergence.

Lemma 3.19. *Under hypothesis 3.6 there exist a Lipschitz solution U of the following equation*

$$\begin{cases} \partial_t U + \lambda U + F(x,p,U)\cdot\nabla_x U + b(x,p,U)\cdot\nabla_p U - \sigma\Delta_p U = G(x,p,U) \text{ for } t\in(0,T), x\in\Omega, p\in\mathbb{R}^m, \\ U(0,x,p) = U_0(x,p) \text{ for } (x,p)\in\Omega\times\mathbb{R}^m, \end{cases} \tag{3.20}$$

on any time interval. Moreover, if we assume that $\lambda \geq 2\|D_u G\|_\infty$, that U_0 is bounded and that G has linear growth in U then there exist a constant C which depends on the data of the problem such that

$$\forall t < +\infty \quad \|U(t,\cdot)\|_\infty + \|(D_x U, D_p U)(t,\cdot)\|_\infty \leq C.$$

Proof. The Lipschitz estimate follows quite easily from the proof of Theorem 3.11. It suffices to notice that because $\lambda \geq 2\|D_u G\|_\infty$ we may take a constant β which gives exactly this uniform in time Lipschitz estimate. The bound on $\|U\|_\infty$ is obtained by mean of Gronwall Lemma whenever U_0 is bounded and G has linear growth in U. □

Remark 3.20. Strong monotonicity conditions on (F,G,b) in (x,p) only as in Hypothesis 3.6 do not in general help for long time convergence: the condition $\lambda \geq 2\|D_u G\|_\infty$ appears to be necessary to control the Lipschitz norm of U uniformly in time if no strong monotonicity of the system is imposed in U. Consider the following example in dimension 1

$$\begin{cases} \lambda > 0, \\ F(x,U) = -\gamma x, \\ G(x,U) = \alpha x + \gamma U. \end{cases}$$

Obviously for this choice

$$\begin{pmatrix} D_x G & D_u G \\ D_x F & D_u F \end{pmatrix} \geq \alpha \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

in the order of positive semi-definite matrix. However, the solution of the associated PDE on $\mathbb{R}$

$$\partial_t U + \lambda U - \gamma x\cdot\nabla_x U = \alpha x + \gamma U,$$

is given by

$$U(t,x) = e^{(\gamma-\lambda)t} U_0(x e^{\gamma t}) + \frac{\alpha}{\lambda}\left(e^{(2\gamma-\lambda)t} - e^{(\gamma-\lambda)t}\right) x.$$

It is evident in this particular example that there is no hope to recover uniform in time Lispchitz estimate unless the condition

$$\lambda \geq 2\gamma,$$

is fulfilled. Under stronger assumptions on the monotonicity of (G, F, b) in U the conclusion of Lemma 3.19 may hold even for $\lambda = 0$ (though we expect that $\lambda > 0$ might be necessary to recover uniqueness of the limit). We do not mean to be exhaustive here and leave such developments to the interested reader.

We believe this Lemma to be almost sufficient just by itself to conclude to the converge of a solution U of (3.20) to a solution of (3.19) as time tends to infinity. Indeed it gives compactness of the sequence $(V_n)_{n\in\mathbb{N}}$ defined by

$$V_n(t, x, p) = U(t + n, x, p) \text{ for } (t, x, p) \in [0, 1] \times \Omega \times \mathbb{R}^m.$$

Along a subsequence, (V_n) converges locally uniformly to a Lipschitz bounded function V which does not depend on time. Now, Lipschitz solution are not adapted, at least in the form we introduced them in this paper, to equation (3.19). However we believe viscosity like information to be sufficient to characterise the limit function V as a weak solution of equation (3.19) by adapting monotone solutions [4] to equation (3.20) with the help of Lemma 3.15. The property of stability enjoyed by those solution would guarantee that the limiting function V is a monotone solution of (3.19) while uniqueness would ensure the sequence (V_n), and hence U, converges indeed to V.

3.5.6 Other notions of monotonicity

It is an observation of P.-L. Lions presented in [4] that instead of monotonicity of U_0 in x we may ask for monotonicity along a function $\phi : \mathbb{R}^d \to \mathbb{R}^d$:

$$\langle U_0(x, p) - U_0(y, p), \phi(x) - \phi(y)\rangle \geq 0.$$

In which case, conditions (3.7) and (3.8) may become for instance

$$\forall (x, y, p, q, u, v) \in \Omega^2 \times \mathbb{R}^{2m} \times \mathbb{R}^{2d},$$

$$\langle U_0(x, p) - U_0(y, q), \phi(x) - \phi(y)\rangle + \langle p - q, A(p - q)\rangle \geq \alpha |U_0(x, p) - U_0(y, q)|^2,$$

$$\begin{gathered}\langle G(x, p, u) - G(y, q, v), \phi(x) - \phi(y)\rangle + \langle F(x, p, u)D\phi(x) - F(y, q, v)D\phi(y), u - v\rangle \\ +\langle 2A(b(x, p, u) - b(y, q, b)), p - q\rangle \\ \geq \alpha \left(|x - y|^2 + |p - q|^2\right).\end{gathered}$$

We make this remark as it leads to estimates for a wider class of functions (G, F, b, U_0).

3.5.7 Return on the heuristic we presented and link with FBSDE

We previously mentioned without detailing much that we could use a wider class of function V satisfying

$$\begin{cases} \partial_t V + F(x, p, U) \cdot \nabla_x V + b(x, p, U) \cdot \nabla_p V - \sigma \Delta_p V = Q(x, p, U, V) \text{ in } (0, T) \times \mathbb{R}^d \times \mathbb{R}^m, \\ V(0, x, p) = V_0(x, p) \subset \mathbb{R}^m \text{ in } \mathbb{R}^d \times \mathbb{R}^m, \end{cases}$$

to 'complete' the master equation, instead of a linear function $V(t,x,p) = Ap$, so long as conditions (3.5) and (3.6) hold for this choice of V. If Q is Lipschitz continuous, we can construct a Lispchitz solution V to this equation which is well defined so long as U is Lipschitz continuous. Using this fact the analysis we provided in the special case of a linear function can be extended to a wider class without much difficulty. In particular we stressed out that V could be taken as an additional unknown instead of an explicit choice. We are now going to give one such example, which is based entirely on the notion of G-monotonicity [30], and give another interpretation of those wellposedness conditions

Example 3.21. Let $N \in \mathcal{M}_d(\mathbb{R})$ and $M \in \mathcal{M}_{d\times m}(\mathbb{R})$. Looking at the equation satisfied by $(\tilde{U}, V)$ for $\tilde{U} = NU$ and $V = MU$ and using the heuristic we developed, we expect that under the following two conditions:

$$\forall (x,y,p,q,u,v) \in \Omega^2 \times \mathbb{R}^{2m} \times \mathbb{R}^{2d},$$

$$\langle U_0(X) - U_0(Y), \Gamma(X-Y)\rangle \geq 0 \quad X = \begin{pmatrix} x \\ p \end{pmatrix} \quad Y = \begin{pmatrix} y \\ q \end{pmatrix} \quad \Gamma = (N^T, M^T),$$

$$\langle \Gamma(\tilde{F}(X,u) - \tilde{F}(Y,v)), u-v\rangle + \langle \Gamma^T(G(X,u) - G(Y,v)), X-Y\rangle \geq \alpha|u-v|^2 \quad \tilde{F} = \begin{pmatrix} F \\ b \end{pmatrix}.$$

we can show existence of a Lipschitz solution on any time interval. In fact, even though there seem to be few results in the PDE literature on the subject, this set of conditions is well known in the field of forward backward stochastic differential equations as G-monotonicity [30].

We are not presenting this example to show that letting $V = Ap$ was a poor choice. Rather we want to stress the following: There is no one general best choice of a function V.
Indeed we believe the conditions for G-monotonicity to hold are equally as strong as the one we previously required. However, depending on the nature of the coefficients, one or the other notion of monotonicity might hold. When (F,b) depends on U only as in Example 3.4 G-monotonicity seems better suited to the study of equation (3.1) as there is no chance of the necessary assumptions for Theorem 3.11 to hold. However when G does not depend on p but the other terms do, G-monotonicity will never be satisfied. Point being that the way to approach equation (3.1) seems to depends heavily on the nature of the different coefficients. Some monotonicity assumptions on U_0 and (F,G,b) does however seems unavoidable to propagate Lipschitz regularity.

Remark 3.22. We do not want to delve too much in the link between forward backward stochastic differential equations and semi-linear systems of PDE as this is beyond the scope of this paper (for more on the subject, we recommend [29]). Let us just state that the existence of a Lipschitz solution of (3.1) gives an existence result for the following FBSDE:

$$\begin{cases} X_t = x_0 - \int_0^t F(X_s, p_s, V_s)ds, \\ p_t = p_0 - \int_0^t b(X_s, p_s, V_s)ds + \sqrt{2\sigma}W_t, \\ V_t = U_0(X_T, p_T) + \int_t^T G(X_s, p_s, V_s)ds - \int_t^T Z_s dW_s \quad \forall t \in [0,T], \end{cases} \tag{3.21}$$

for $(W_t)_{t\geq 0}$ a Brownian motion on $\mathbb{R}^m$ and for any initial condition (x_0, p_0).
Let

$$\begin{cases} Y_t = x_0 - \int_0^t F(Y_s, q_s, U(T-s, Y_s, q_s))ds, \\ q_t = p_0 - \int_0^t b(Y_s, q_s, U(T-s, Y_s, q_s))ds + \sqrt{2\sigma}W_t, \end{cases}$$

for a Lipschitz solution U. Denoting $U_t = U(T-t, Y_t, q_t)$, we know that $(M_t)_{t\in[0,T]}$ defined by

$$M_t = U_t + \int_0^t G(Y_s, q_s, U_s)ds$$

is a $\mathcal{F}-$martingale for the natural filtration $\mathcal{F}$ associated to the Brownian motion $(W_t)_{t\geq 0}$. Since (F, G, U, b) are all Lipschitz continuous, M is also a square integrable martingale. By the martingale representation theorem, there exists a predictable $\mathcal{F}-$adapted stochastic process $(C_s)_{s\in[0,T]}$ such that

$$\begin{aligned} &\mathbb{E}\left[\int_0^T (C_s)^2 ds\right] < +\infty, \\ &\forall t \leq T \quad M_t = M_0 + \int_0^t C_s dW_s. \end{aligned}$$

As a consequence:

$$M_T - \int_t^T C_s dW_s - \int_0^t G(Y_s, q_s, U_s)ds = U_t.$$

Which ends to show that $(Y_t, q_t, U_t, C_t)_{t\in[0,T]}$ is a solution of the previous FBSDE as by definition of M:

$$M_T = U_0(Y_T, q_T) + \int_0^T G(Y_s, q_s, U_s)ds.$$

Of course this is perfectly natural if we remember what is the master equation to the associated mean field game (which in this case take the form of the FBSDE (3.21)).

References

[1] Achdou, Y., Bertucci, C., Lasry, J.-M., Lions, P.-L., Rostand, A., and Scheinkman, J.A. "A class of short-term models for the oil industry that accounts for speculative oil storage". *Finance and Stochastics* 26 (2022), pp. 631–669.

[2] Bayraktar, E., Cecchin, A., Cohen, A., and Delarue, F. "Finite state mean field games with Wright–Fisher common noise". *Journal de Mathématiques Pures et Appliquées* 147 (2021), pp. 98–162.

[3] Bertucci, C. "Monotone solutions for mean field games master equations: continuous state space and common noise". *Communications in Partial Differential Equations* 48.10-12 (2023), pp. 1245–1285.

[4] Bertucci, C. "Monotone solutions for mean field games master equations: finite state space and optimal stopping". *Journal de l'École polytechnique — Mathématiques* 8 (2021), pp. 1099–1132.

[5] Bertucci, C. and Cecchin, A. "Mean field games master equations: from discrete to continuous state space". *To appear in SIAM journal on Mathematical Analysis, 2024* (2022).

[6] Bertucci, C., Lasry, J.-M., and Lions, P.-L. "A singular infinite dimensional Hamilton-Jacobi-Bellman equation arising from a storage problem" (2023). arXiv: 2210.02780 [math.AP].

[7] Bertucci, C., Lasry, J.-M., and Lions, P.-L. "Master equation for the finite state space planning problem". *Archive for Rational Mechanics and Analysis* 242 (2021), pp. 327–342.

[8] Bertucci, C., Lasry, J.-M., and Lions, P.-L. "On Lipschitz solutions of mean field games master equations" (2023). arXiv: 2302.05218 [math.AP].

[9] Bertucci, C., Lasry, J.-M., and Lions, P.-L. "Some remarks on Mean Field Games". *Communications in Partial Differential Equations* 44.3 (2019), pp. 205–227.

[10] Bertucci., C., Bertucci, L., Lasry, J.-M., and Lions, P.-L. "Mean Field Game Approach to Bitcoin Mining" (2020). arXiv: 2004.08167 [econ.TH].

[11] Cardaliaguet, P., Delarue, F., Lasry, J.-M., and Lions, P.-L. "The master equation and the convergence problem in mean field games". *Princeton University Press* (2019).

[12] Cardaliaguet, P. and Lehalle, C.-A. "Mean Field Game of Controls and An Application To Trade Crowding". *Mathematics and Financial Economics* 12.3 (2019), pp. 335–363.

[13] Cardaliaguet, P. and Souganidis, P. "Monotone Solutions of the Master Equation for Mean Field Games with Idiosyncratic Noise". *SIAM Journal on Mathematical Analysis* 54.4 (2022), pp. 4198–4237.

[14] Carmona, R. and Delarue, F. "Mean field forward-backward stochastic differential equations". *Electronic Communications in Probability* 18.none (2013), pp. 1–15.

[15] Cecchin, A. and Delarue, F. "Weak solutions to the master equation of potential mean field games" (2022). arXiv: 2204.04315 [math.OC].

[16] Crandall, M., Ishii, H., and Lions, P.-L. "user's guide to viscosity solutions of second order partial differential equations". *Bulletin of the American Mathematical Society* 27.1 (1992), pp. 1–67.

[17] Crandall, M. and Lions, P.-L. "Quadratic growth of solutions of fully nonlinear second order equations in $\mathbb{R}^n$". *Differential and Integral Equations* 3.4 (1990), pp. 601–616.

[18] Gangbo, W., Mészáros, A., Mou, C., and Zhang, J. "Mean Field Games Master Equations with Non-separable Hamiltonians and Displacement Monotonicity". *Annals of Probability* 50 (2022).

[19] Ishii, H. "Uniqueness of Unbounded Viscosity Solution of Hamilton-Jacobi Equations". *Indiana University Mathematics Journal* 33.5 (1984), pp. 721–748.

[20] Lasry, J.-M. and Lions, P.-L. "Mean Field Games". *Japanese Journal of Mathematics* 2 (Mar. 2007), pp. 229–260.

[21] Lasry, J.-M. and Lions, P.-L. "Mean-field games with a major player". *Comptes Rendus Mathematique* 356.8 (2018), pp. 886–890.

[22] Lavigne, P., Petit, Q., and Warin, X. "A mean field game model for energy transition with common noise through pollution". *[unpublished]* (2024).

[23] Lions, P.-L. "Lessons at Collège de France" (2011). URL: http://www.college-de-france.fr.

[24] Lions, P.-L. "Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of viscosity solutions for general second-order equations". *Journal of Functional Analysis* 86.1 (1989), pp. 1–18.

[25] Lions, P.-L. and Seeger, B. "Linear and nonlinear transport equations with coordinate-wise increasing velocity fields" (2023). arXiv: 2307.05819 [math.AP].

[26] Lions, P.-L. and Souganidis, P. "Extended mean-field games". *Rendiconti Lincei* 3 (2020), pp. 611–625.

[27] Mészáros, A. and Mou, C. "Mean Field Games Systems under Displacement Monotonicity". *SIAM Journal on Mathematical Analysis* 56.1 (2024), pp. 529–553.

[28] Mou, C. and Zhang, J. "Wellposedness of Second Order Master Equations for Mean Field Games with Nonsmooth Data". *accepted in the Mem. Amer. Math. Soc., 108pp* (2022).

[29] Pardoux, E. and Tang, S. "Forward-backward stochastic differential equations and quasilinear parabolic PDEs". *Probability Theory and Related Fields* (1999).

[30] Peng, S. and Wu, Z. "Fully Coupled Forward-Backward Stochastic Differential Equations and Applications to Optimal Control". *SIAM Journal on Control and Optimization* 37.3 (1999), pp. 825–843.

[31] Touzi, N. "Optimal stochastic control, stochastic target problems, and backward SDE" (2010).

Appendix A A proof of Lemma 2.9

Lemma 2.9. *Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a function of class C^1, the following two proposition are equivalent:*

- $\exists \alpha, \forall x \quad \xi \cdot Df(x) \cdot \xi \geq \alpha |Df \cdot \xi|^2.$ (i)
- $\exists \alpha, \forall (x,y) \quad \langle f(x) - f(y), x - y \rangle \geq \alpha |f(x) - f(y)|^2.$ (ii)

Proof. $(ii) \implies (i)$ is trivial, as we only need to divide by t^2 and take the limit as $t \to 0$ for $x = y + t\xi$.

For the the other way, let us first assume that f is invertible. Its inverse g satisfies (by taking $\xi = Dg \cdot \nu$)

$$\forall \nu \quad \nu \cdot Dg \cdot \nu \geq \alpha |\nu|^2.$$

Let

$$w(t) = \langle g(th + (1-t)z) - g(z), h - z \rangle - \alpha t |h - z|^2,$$
$$\implies w'(t) = \langle h - z, Dg(th + (1-t)z), h - z \rangle - \alpha |h - z|^2 \geq 0,$$

by assumption. Because $w(0) = 0$ we can deduce that

$$w(1) = \langle g(h) - g(z), h - z \rangle - \alpha |h - z|^2 \geq 0.$$

By taking $h = f(x)$ and $z = f(y)$ we conclude that for invertible f we have

$$\langle f(x) - f(y), x - y \rangle \geq \alpha |f(x) - f(y)|^2.$$

For non invertible f, because (i) implies that f is monotonous, we know that for any ε $f^\varepsilon : x \to f(x) + \varepsilon x$ is invertible. Moreover:

$$\xi \cdot Df^\varepsilon \cdot \xi = \xi \cdot Df \cdot \xi + \varepsilon |\xi|^2$$

and

$$|Df^\varepsilon \cdot \xi|^2 = |D_x f \cdot \xi|^2 + \varepsilon^2 |\xi^2| + 2\varepsilon \xi \cdot Df \cdot \xi.$$

As a consequence, for $\varepsilon < \min(\frac{1}{2}, \frac{1}{2\alpha})$ f^ε satisfies property (i) for $\alpha^\varepsilon = \alpha - 2\alpha^2\varepsilon$, which means by what we just showed for invertible f that

$$\langle f^\varepsilon(x) - f^\varepsilon(y), x - y \rangle \geq (\alpha - 2\alpha^2\varepsilon)|f^\varepsilon(x) - f^\varepsilon(y)|^2.$$

We can then conclude by taking the limit as $\varepsilon \to 0$. □