



HAL
open science

On Lipschitz solutions of mean field games master equations

Charles Bertucci, Jean-Michel Lasry, Pierre-Louis Lions

► **To cite this version:**

Charles Bertucci, Jean-Michel Lasry, Pierre-Louis Lions. On Lipschitz solutions of mean field games master equations. *Journal of Functional Analysis*, 2024, 287 (5), pp.110486. 10.1016/j.jfa.2024.110486 . hal-03980892

HAL Id: hal-03980892

<https://hal.science/hal-03980892v1>

Submitted on 9 Feb 2023

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

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

ON LIPSCHITZ SOLUTIONS OF MEAN FIELD GAMES MASTER EQUATIONS

CHARLES BERTUCCI¹, JEAN-MICHEL LASRY², PIERRE-LOUIS LIONS^{2,3}

ABSTRACT. We develop a theory of existence and uniqueness of solutions of MFG master equations when the initial condition is Lipschitz continuous. Namely, we show that as long as the solution of the master equation is Lipschitz continuous in space, it is uniquely defined. Because we do not impose any structural assumptions, such as monotonicity for instance, there is a maximal time of existence for the notion of solution we provide. We analyze three cases: the case of a finite state space, the case of master equation set on a Hilbert space, and finally on the set of probability measures, all in cases involving common noises. In the last case, the Lipschitz continuity we refer to is on the gradient of the value function with respect to the state variable of the player.

CONTENTS

Introduction	2
1. The case of a finite state space	3
1.1. The notion of Lipschitz solution for master equations	4
1.2. The main result on Lipschitz solutions	5
1.3. Comparison with the usual notion of characteristics for MFG	8
1.4. Time regularity of Lipschitz solutions	8
1.5. A strong-weak like uniqueness result	9
2. The Hilbertian case	10
2.1. The deterministic case	11
2.2. The case of common noise	12
3. Master equations on the set of probability measures	15
3.1. Setting and notation	15
3.2. Main definition and result	17
3.3. Return to the solutions of the initial master equation	20
3.4. Master equations associated to common noise	21
3.5. Master equations involving the image measure	24
4. Comments and future developments on Lipschitz solutions of MFG master equations	24
4.1. Application to numerical computations	24
4.2. Another representation formula	25

¹ : CMAP, Ecole Polytechnique, UMR 7641, 91120 Palaiseau, France

²: CEREMADE, Université Paris Dauphine, UMR 7534, 75016 Paris, France

³: Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France .

4.3. Another notion of monotone solutions for MFG master equations	26
4.4. Master equations with control on the volatility	27
Acknowledgments	27
References	27
Appendix A. Proof of the estimate used in the proof of Theorem 3.7	28

INTRODUCTION

In this paper, we explain why there can be at most one "sufficiently smooth" solution of mean field games (MFG in short) master equations, in three different settings. We first consider the case of a MFG master equation associated to a finite state space model. In such a case, the master equation is a finite dimensional partial differential equation PDE (in short). We then present analogue results in the so-called Hilbertian approach, in which the master equation is a PDE set on a general Hilbert space; as well as a master equation associated to a MFG with a continuous state space. Of course the concept of a "sufficiently" smooth solution shall be detailed later on in the paper, and depends on the setting of the equation, but it turns out that Lipschitz continuity in space/measure is an important notion for the well posedness.

The mathematical study of MFG master equations have been initiated by the second and third authors. These equations are PDE, which can be written in either finite or infinite dimensional sets. They are an essential tool in the theory of MFG which are dynamic games involving non-atomic agents [18]. The main advantage of master equations is that they allow to model a wide variety of games, namely ones involving so-called common noises. We refer the reader to [19, 11, 13, 14, 3, 4] for more details on the study of MFG master equations.

The main mathematical difficulty arising in the study of MFG master equations is that, in general, such equations create shocks, or discontinuities. For instance, one of the simplest and most famous PDE which can be a non trivial master equation is the Burger's equation. The so-called monotone regime has been identified in [18]. It prevents shocks and allows to propagate some regularity for MFG master equations [19, 11, 4, 5, 7]. Hence in this monotone regime, one can study classical solutions of the problem on time intervals of arbitrary length. Moreover, a weaker notion of monotone solution has been introduced in [3, 4] and studied in [10]. It allows to deal with non-classical solutions in the monotone regime. Let us also mention the work [22] which also deals with the monotone regime. Other regimes of well-posedness exist, such as the so called displacement monotone one, see for instance [15, 16]. However no proper notion of weak solution exists in general.

Several results of well-posedness on short time horizons have been established, like for instance [19, 14, 12, 2]. In these results, the authors prove that if the time horizon is sufficiently small, then there always exists a unique classical solution of the master equation. This approach is close to the the one we adopt here. In some sense our results

state that this idea can be pushed further, namely by using a weaker notion of solution. Moreover, we can characterize this maximal time of well-posedness in terms of the Lipschitz regularity of the solution.

In this paper, we establish the fact that, starting from a sufficiently regular initial condition, there exists exactly one solution of the master equation in a certain regularity class up to a certain time, which depends on the initial condition. We believe this fact to be quite general and to rely mostly on the form of the master equations. Even though it seems unfeasible to treat all possible master equations at once, we give three examples of settings in which this phenomenon happens to convince the reader of the generality of this approach. We end the paper by discussing the implications that such results may have as well as by giving some directions of extensions of this work.

Finally, let us insist on the fact that the approach we shall propose is clearly in the spirit of the Cauchy-Lipschitz (or Picard-Lindelöf) theory of ordinary differential equations.

The rest of the paper is organized as follows. Section 1 presents our approach on the more simple case of a finite state space. Section 2 generalizes this approach to master equations set on Hilbert spaces and Section 3 treats the case of a continuous state space. Finally we have gathered additional remarks and perspectives in Section 4.

1. THE CASE OF A FINITE STATE SPACE

For MFG with a finite ($d \geq 1$) number of states, the master equation generally takes the following form

$$(1.1) \quad \partial_t U(t, x) + \langle F(x, U), \nabla_q \rangle U(t, x) + \lambda(U(t, x) - (DS)^* U(t, Sx)) = G(x, U) \text{ in } (0, \infty) \times \mathcal{O},$$

$$(1.2) \quad U(0, x) = U_0(x) \text{ in } \mathcal{O},$$

where $U : [0, \infty) \times \mathcal{O} \rightarrow \mathbb{R}^d$ and \mathcal{O} is bounded domain of \mathbb{R}^d . The terms $F : \mathcal{O} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $G : \mathcal{O} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ model strategic interactions between the players, U_0 is an initial condition and the terms involving $\lambda > 0$ and $S : \mathcal{O} \rightarrow \mathcal{O}$ model common noise in the game. We refer to [19, 5] for more details on such models. For the rest of this section, $\langle \cdot, \cdot \rangle$ denotes the euclidean scalar product of \mathbb{R}^d .

Remark 1.1. *In this setting, for most of the model studied in the literature, the variable x stands for the repartition of players. This choice of variable may seem strange since x denotes the state of a single player in lots of models. However, because in several key example of master equation this variable x does not stand for a repartition of players [6, 1], we made this choice. A more general statement is that quite often master equations can be of some importance in themselves, even though there is no proper underlying MFG.*

Concerning the boundary conditions, we assume that \mathcal{O} is smooth and that

$$(1.3) \quad \forall x \in \partial\mathcal{O}, \forall p \in \mathbb{R}^d, \langle \eta(x), F(x, p) \rangle \geq 0,$$

which ensures that no additional boundary condition is needed. In the previous assumption, $\eta(x)$ stands for the normal vector to $\partial\mathcal{O}$, at the point x , pointing outward.

Remark 1.2. *The smoothness of \mathcal{O} is by no means important here. All the following can be extended to non smooth \mathcal{O} (which is often the case in applications). However, we make this assumption to simplify the following discussion.*

Accordingly, we also assume that S is such that $S(\mathcal{O}) \subset \mathcal{O}$.

1.1. The notion of Lipschitz solution for master equations. The main remark at the origin of the notion of solution we propose here is the following: If we know the map $(t, x) \rightarrow (F(x, U(t, x)), G(x, U(t, x)))$, then computing the solution of (1.1) reduces to solving linear transport equation. Hence, a solution of (1.1) can be expressed a fixed point of the composition of solving the linear transport equation and evaluating F and G .

We now give a notion of solutions of the linear transport equation at interest, which is by now standard, and refer to [20] for a presentation of recent developments on transport equations. Consider the following transport equation of solution $V : [0, \infty) \times \mathcal{O} \rightarrow \mathbb{R}^d$

$$(1.4) \quad \partial_t V(t, x) - \langle B(t, x), \nabla_x \rangle V(t, x) = A(t, x) \text{ in } (0, \infty) \times \mathcal{O},$$

with initial condition

$$(1.5) \quad V|_{t=0}(x) = U_0(x) = \text{ in } \mathcal{O}.$$

In the previous, $A, B : (0, \infty) \times \mathcal{O} \rightarrow \mathbb{R}^d$ and $U_0 : \mathcal{O} \rightarrow \mathbb{R}^d$. The linear transport equation (1.4) can be solved by using the method of characteristics that we now recall briefly. Let us assume that B is, uniformly in t , Lipschitz continuous in x . This implies that the ordinary differential equation (ODE)

$$(1.6) \quad \frac{d}{ds} x(s) = B(t - s, x(s)),$$

has a unique solution given any initial condition $x \in \mathcal{O}$ and time interval $[-t, t]$. Furthermore, assume that B satisfies (1.3) so that \mathcal{O} is invariant by (1.6), i.e. $x(s) \in \mathcal{O}$ for all $s \in [0, t]$. The method of characteristics states that a classical solution V of (1.4) satisfies

$$(1.7) \quad V(t, x) = \int_0^t A(t - s, x(s)) ds + U_0(x(t)).$$

Here, we shall use this relation as the notion of solution of (1.4).

Definition 1.3. *Take a final time T , a time dependent vector field $B : [0, T] \times \mathcal{O} \rightarrow \mathbb{R}^d$ satisfying (1.3) as well as being, uniformly in t , Lipschitz continuous in $x \in \mathcal{O}$. For A and U_0 continuous, the solution of (1.4) is the function defined for every $(t, x) \in [0, T] \times \mathcal{O}$ by (1.7). In this case, we note*

$$(1.8) \quad V = \Psi(T, A, B, U_0).$$

Remark 1.4. *The previous definition indeed defines a mapping Ψ because we can consider a solution of (1.6) from any initial condition and arbitrary time length.*

We then provide the following definition of a Lipschitz solution of the master equation (1.1).

Definition 1.5. *A Lipschitz solution U of (1.1), on the time interval $[0, T]$, is a function such that*

- U is Lipschitz in $x \in \mathcal{O}$, uniformly in $t \in [0, \alpha]$ for any $\alpha < T$.
- For any $t < T$

$$(1.9) \quad U = \Psi\left(t, G(\cdot, U) - \lambda(U - (DS)^*U \circ S), -F(\cdot, U), U_0\right).$$

This, quite classical, definition of a solution of a non-linear PDE has the advantage that it allows us to define a concept of solution for merely Lipschitz (in the space variable) functions. The following result is immediate.

Proposition 1.6. *A classical solution of (1.1) is a Lipschitz solution of (1.1) in the sense of Definition 1.5 on any time interval. Moreover a Lipschitz solution is a classical solution if it is smooth.*

Proof. First part of the claim. The proof of this result is a simple computation. Considering a classical solution U . Take the time derivative of both U and $\Psi(t, G(\cdot, U) - \lambda(U - (DT)^*U \circ T), F(\cdot, U), U_0)$. It is immediate to verify that they are actually the same, hence that U is also a Lipschitz solution of (1.1).

Second part of the claim. Consider a smooth Lipschitz solution of (1.1) on $[0, T)$. Take $t > 0, x \in \mathcal{O}$ and denote by $(x(s))_{s \in [0, t]}$ the solution of

$$(1.10) \quad \frac{dx(s)}{ds} = -F(x(s), U(t-s, x(s)))$$

with initial condition x . Observe now that, for $dt > 0$ sufficiently small

$$(1.11) \quad U(t, x) - U(t-dt, x(dt)) = \int_0^{dt} G(x(s), U(t-s, x(s))) - \lambda(U(t-s, x(s)) - (DS(x(s)))^*U(t-s, Sx(s))) ds.$$

Hence dividing the previous relation by dt and letting $dt \rightarrow 0$, we recover that U indeed solves (1.1). \square

1.2. The main result on Lipschitz solutions. In the spirit of the usual theory of ODE, we are able to prove the following result for such Lipschitz solutions.

Theorem 1.7. *Assume that F and G are locally Lipschitz functions, that S is smooth and consider a Lipschitz initial condition $U_0 : \mathcal{O} \rightarrow \mathbb{R}^d$.*

- *There always exists a time $T > 0$ such that there exists a unique solution of (1.1) in the sense of Definition 1.5 on $[0, T)$.*
- *Moreover, there exists a maximal time $T^c \in [0, \infty]$ and a solution U associated to T^c such that, for any solution V of the problem on an interval $[0, T]$: we have that $T \leq T^c$ and the restriction of U to $[0, T)$ is equal to V .*
- *If $T^c < \infty$, then $\|D_x U(t, \cdot)\|_\infty \rightarrow \infty$ as $t \rightarrow T^c$.*

Remark 1.8. *Let us insist on the fact that \mathcal{O} is assumed to be bounded here, and hence that U_0 is bounded and that, for $C > 0$, F and G are Lipschitz continuous in $x \in \mathcal{O}$, uniformly in $|p| \leq C$. Cases in which the domain of the equation is unbounded are treated in the next section.*

Proof. Consider $T > 0$ and introduce the mapping Φ defined by

$$(1.12) \quad \begin{aligned} \Phi : E_\infty &\rightarrow E_\infty, \\ U &\rightarrow \Psi \left(T, G(\cdot, U) - \lambda(U - (DS)^*U \circ S), -F(\cdot, U), U_0 \right), \end{aligned}$$

where

$$(1.13) \quad E_C := \{U : [0, T] \times \mathcal{O} \rightarrow \mathbb{R}^d, \|U\|_\infty < C, \|D_x U\|_\infty < C\}.$$

Step 1: Φ is well defined. First we want to show that Φ is well defined, i.e. that it takes its values in E_∞ . Take $U \in E_\infty$. The fact that $\Phi(U)$ is bounded is a direct consequence of (1.7). We now show that $\|D_x \Phi(U)\|_\infty < \infty$. In fact, we are going to show that as soon as A, B and U_0 are Lipschitz in the space variable x , uniformly in $[0, T)$, then $\|D_x \Psi(T, A, B, U_0)\|_\infty < \infty$.

Let us denote by $\xi(t, s, x)$ the flow of the ODE (1.6). That is, given $t > 0$, an initial condition $x \in \mathcal{O}$, $\xi(t, s, x) = x(s)$ where $(x(s))_{s \geq 0}$ is the unique solution of (1.6).

Differentiating (1.7) with respect to x yields for $t < T, x \in \mathcal{O}$

$$(1.14) \quad D_x \Psi(T, A, B, U_0)(t, x) = \int_0^t D_x A(s, \xi(t, s, x)) D_x \xi(t, s, x) ds + D_x U_0(x(t)) D_x \xi(t, t, x).$$

Let us recall that, since B is uniformly Lipschitz, the flow ξ is Lipschitz in x , uniformly in t, s . Then the bound on $\|D_x \Psi(T, A, B, U_0)\|_\infty$ follows, from which we deduce that $\Phi(U) \in E_\infty$. Moreover, if T is chosen sufficiently small, then there exists C such that

$$(1.15) \quad \sup_{t \leq t_1} \|D_x U\|_\infty \leq C \Rightarrow \sup_{t \leq t_1} \|D_x \Phi(U)\|_\infty \leq C.$$

Hence, if T is sufficiently small and C sufficiently large, Φ maps E_C into itself.

Step 2: Φ is a contraction. We now show that Φ is a contraction on E_C for the $\|\cdot\|_\infty$ norm, if T is chosen small enough and C is chosen large enough. From (1.7), we deduce that, for any $U, V \in E$, $t < T, x \in \mathcal{O}$

$$(1.16) \quad \begin{aligned} (\Phi(U) - \Phi(V))(t, x) &= \int_0^t G(\xi_1(t, s, x), U(s, \xi_1(t, s, x))) - G(\xi_2(t, s, x), V(s, \xi_2(t, s, x))) ds \\ &\quad - \lambda \int_0^t ((DS(\xi_1(t, s, x)))^* U(t-s, \xi_1(s-t, x)) - (DS(\xi_2(t, s, x)))^* V(t-s, \xi_2(s-t, x))) ds \\ &\quad + \lambda \int_0^t V(t-s, \xi_2(t, s, x)) - U(t-s, \xi_1(t, s, x)) ds, \\ &\quad + U_0(\xi_1(t, s, x)) - U_0(\xi_2(t, s, x)), \end{aligned}$$

where ξ_1 and ξ_2 are the flows associated to respectively U and V , that is $\xi_1(t, \cdot, \cdot)$ is the flow of the ODE

$$(1.17) \quad \frac{d}{ds} x(s) = -F(x(s), U(t-s, x(s))),$$

while $\xi_2(t, \cdot, \cdot)$ is the flow of the same ODE when U is replaced by V . To deduce the contraction property from (1.16), it suffices to show that

$$(1.18) \quad |\xi_1(t, s, x) - \xi_2(t, s, x)| \leq tC\|U - V\|_\infty.$$

Indeed if (1.18) holds, then we can bound the right hand side of (1.16) with

$$(1.19) \quad \begin{aligned} & t(C_G C\|U - V\|_\infty + C_G(\|U - V\|_\infty + \|D_x U\|_\infty C\|U - V\|_\infty)) + \\ & + \lambda t(\|D^2 S\|_\infty \|U\|_\infty C\|U - V\|_\infty + \|DS\|_\infty^2(\|U - V\|_\infty + \|D_x U\|_\infty C\|U - V\|_\infty)) \\ & + \lambda t(\|U - V\|_\infty + \|D_x U\|_\infty C\|U - V\|_\infty) + \|D_x U_0\|_\infty tC\|U - V\|_\infty, \end{aligned}$$

where C_G is the Lipschitz constant of G on $\{(x, p) | x \in \mathcal{O}, |p| \leq C\}$. Hence, we deduce that for any $C > 0$ such that $\Phi(E_C) \subset \Phi(E_C)$, there exists $T > 0$ such that Φ is a contraction on E_C for the $\|\cdot\|_\infty$ norm. It then remains to show (1.18). Let us compute, for $U, V \in E_C$

$$(1.20) \quad \begin{aligned} \left| \frac{d}{ds}(\xi_1(t, s, x) - \xi_2(t, s, x)) \right| &= |F(\xi_1(t, s, x), U(t - s, \xi_1(t, s, x))) - F(\xi_2(t, s, x), V(t - s, \xi_2(t, s, x)))| \\ &\leq |F(\xi_1(t, s, x), U(t - s, \xi_1(t, s, x))) - F(\xi_1(t, s, x), V(t - s, \xi_1(t, s, x)))| \\ &\quad + |F(\xi_1(t, s, x), V(t - s, \xi_1(t, s, x))) - F(\xi_1(t, s, x), V(t - s, \xi_2(t, s, x)))| \\ &\quad + |F(\xi_1(t, s, x), V(t - s, \xi_2(t, s, x))) - F(\xi_2(t, s, x), V(t - s, \xi_2(t, s, x)))| \\ &\leq C_F\|U - V\|_\infty + C_F\|D_x V\|_\infty |\xi_1(t, s, x) - \xi_2(t, s, x)| \\ &\quad + C_F |\xi_1(t, s, x) - \xi_2(t, s, x)|. \end{aligned}$$

Defining $I(s) = |\xi_1(t, s, x) - \xi_2(t, s, x)|$, we recognize that it satisfies an inequality of the form

$$(1.21) \quad \frac{d}{ds}I(s) \leq C(1 + I(s)).$$

Hence, since $I(0) = 0$, we obtain from Grönwall's Lemma that $I(s) \leq e^{Cs} - 1$, from which (1.18) follows.

Step 3: Existence of a fixed point. Let us now remark that for $C > 0$, there exists $L > 0$ such that for any $U \in E_C$, $\|\partial_t U\|_\infty \leq L$. Hence, $\Phi(E_C)$ is contained in subset of uniformly Lipschitz and bounded functions on $[0, T] \times \mathcal{O}$.

We now take $C > 0$ large enough so that $\Phi(E_C) \subset E_C$ and $T > 0$ small enough so that Φ is a contraction on $(E_C, \|\cdot\|_\infty)$. Following the classical proof of Picard's fixed point theorem, we deduce, using **Step 2**, that for any $U \in E_C$, the sequence $(\Phi^n(U))_{n \geq 0}$ is a Cauchy sequence in $(E_C, \|\cdot\|_\infty)$. Hence it converges, thanks to Ascoli-Arzelà Theorem, to a limit $U^* \in E_C$. From the continuity of Φ , we deduce that U^* is indeed a fixed point of Φ .

Step 4: Uniqueness of solutions and critical time of existence. To prove the rest of the claim, let us consider

$$(1.22) \quad T^c = \sup\{T > 0 | \exists \text{ a Lipschitz solution of (1.1) on } [0, T]\},$$

and sequences $(T_n)_{n \geq 0}, (U_n)_{n \geq 0}$ such that for all $n \geq 0$, U_n is a Lipschitz solution of (1.1) on $[0, T_n]$ and $T_n \uparrow_{n \rightarrow \infty} T^c$. For $n \leq m$, consider $T_* = \inf\{t \in [0, T_n), \exists x \in \mathcal{O}, U_n(t, x) \neq U_m(t, x)\}$. If $T_* > -\infty$, then by using **Step 2** on the initial condition $U_0(x) = U_n(t - \epsilon, x)$ for $\epsilon > 0$ sufficiently small, we arrive at a contradiction. Hence we deduce that all the Lipschitz solutions coincides up to time T^c .

Let us now assume that $T^c < \infty$ and that there exists $C > 0$ such that $\|D_x U_n\|_\infty \leq C$ for all $n \geq 0$. Consider $\delta > 0$. By applying **Step 3** to the initial condition $U_n(T^c - \delta, x)$, for n such that $T_n \geq T^c - \delta$, we deduce that there exists a time of existence of a solution $\epsilon > 0$, which is bounded from below by a constant which depends only on C , thanks to the computation of the previous part of the proof. Hence, choosing $\delta > 0$ sufficiently small, we arrive at a contradiction since this allows to extend the solution U_n on $[0, T_n - \delta + \epsilon)$, while being a solution of (1.1) on this time interval. Hence, we necessary have that $\lim_{t \rightarrow T^c} \|D_x U(t, \cdot)\|_\infty = +\infty$ (in the case $T^c < \infty$). \square

Remark 1.9. *Obviously the previous proof is very much in the spirit of the standard theory of ODE. Maybe the main difference here is that we show that Φ is a contraction for the $\|\cdot\|_\infty$ norm while we need to verify that it is defined from a space of Lipschitz functions into itself.*

1.3. Comparison with the usual notion of characteristics for MFG. Let us recall that, in the MFG community, the notion of characteristics associated to (1.1) is in general a forward backward system. In this setting, this system takes the form

$$(1.23) \quad \begin{cases} \frac{d}{dt} Y(t) = G(y(t), Y(t)) \text{ for } t \in (0, T), \\ -\frac{d}{dt} y(t) = -F(y(t), Y(t)) \text{ for } t \in (0, T), \end{cases}$$

which is often associated with boundary conditions of the form $Y(0) = U_0(y(0)), y(T) = x_0$, where $x_0 \in \mathcal{O}$, and $T > 0$.

In our approach, we do not need such complex forward-backward system of characteristics, since, in particular, we can deal with only the forward equation (1.6). We believe that this makes our approach easier to work with than most of the existing literature on the construction of solution of the master equation in short time intervals.

1.4. Time regularity of Lipschitz solutions. We now explain why Lipschitz solutions of (1.1) are necessary locally Lipschitz in t . Even though this was present in the proof, we present the following argument as it is more intrinsic. To simplify the notation, we consider the case $\lambda = 0$, although this does not bear any importance. This fact is quite simple and maybe the best understanding is through the fact, once a bound exists on $D_x U$, we can read a bound on $\partial_t U$ on (1.1). A more rigorous approach consists in making the computation, for U a Lipschitz solution of (1.1) and $t \geq s > 0, x \in \mathcal{O}$, and

$x(\cdot)$ the solution of (1.7) with initial condition x

$$\begin{aligned}
(1.24) \quad |U(t, x) - U(s, x)| &\leq |U(t, x) - U(s, x(t-s))| + |U(s, x(t-s)) - U(s, x)| \\
&\leq \left| \int_s^t G(x(s'), U(t-s', x(s'))) ds' \right| + C|x(t-s) - x| \\
&\leq \left| \int_s^t G(x(s'), U(t-s', x(s'))) ds' \right| + C \left| \int_0^{t-s} F(x(s'), U(t-s', x(s'))) ds' \right| \\
&\leq C|t-s|.
\end{aligned}$$

From this computation, we obtain the

Proposition 1.10. *Under the assumptions of Theorem 1.7, consider the Lipschitz solution U of (1.1) on the maximal time of existence T^c . Then, for any $T < T^c$, U is Lipschitz on $[0, T] \times \mathcal{O}$.*

Remark 1.11. *To avoid repeating the same argument multiple times, we shall not present this result for the next cases, although they actually hold. Maybe the only point of difference is that in general, the Lipschitz regularity in time is only local in space, which does not appear since \mathcal{O} is bounded.*

1.5. A strong-weak like uniqueness result. Even though there does not exist a general notion of weak solution of MFG master equations, we explain why the existence of a Lipschitz solution may yield uniqueness of solutions in a wider class of notion of solutions. To present this idea, we show that any limit V of a sequence of smooth functions $(V_\epsilon)_{\epsilon>0}$ which are almost solution of (1.1), is actually equal to the Lipschitz solution of (1.1) U , when U exists of course. Once again, we are in this section in the case $\lambda = 0$, to simplify notation. We can prove the following.

Proposition 1.12. *Consider a Lipschitz solution U of (1.1) on the time interval $[0, T^c]$. Assume that, for any $T \in (0, T^c)$, there exists $C > 0$ such that for any $\epsilon > 0$, any classical solution V_ϵ of*

$$\begin{aligned}
(1.25) \quad |\partial_t V_\epsilon + \langle F(x, V_\epsilon), \nabla_x \rangle V_\epsilon - G(x, V_\epsilon)| &\leq \epsilon \text{ in } [0, T] \times \mathcal{O}, \\
|V_\epsilon(0, x) - U_0(x)| &\leq \epsilon \text{ in } \mathcal{O}.
\end{aligned}$$

Then the following holds for some constant C depending only on ϵ, T, U_0, F and G

$$(1.26) \quad \sup_{t \leq T, x \in \mathcal{O}} |U(t, x) - V_\epsilon(t, x)| \leq C\epsilon.$$

Proof. To simplify notation we omit the index ϵ on V . Let us consider $t > 0$ and $i \in \{1, \dots, d\}, x^* \in \mathcal{O}$ such that

$$(1.27) \quad U^i(t, x^*) - V^i(t, x^*) = \max_{j, x} \{U^j(t, x) - V^j(t, x)\}$$

Consider $\kappa > 0$ and the solution $x(\cdot)$ of the ODE

$$(1.28) \quad \frac{d}{ds} x(s) = -F(x(s), U(t + \kappa - s, x(s))),$$

with initial condition x^* and $y(\cdot)$ the solution of the ODE

$$(1.29) \quad \frac{d}{ds} y(s) = -F(y(s), V(t + \kappa - s, y(s))),$$

with also initial condition x^* . Let us now compute

$$\begin{aligned}
& U^i(t + \kappa, x^*) - V^i(t + \kappa, x^*) - U^i(t, x^*) + V^i(t, x^*) = \\
& = U^i(t + \kappa, x^*) - U^i(t, x(\kappa)) - V^i(t + \kappa, x^*) + V^i(t, y(\kappa)) \\
(1.30) \quad & + U^i(t, x(\kappa)) - U^i(t, x^*) + V^i(t, x^*) - V^i(t, y(\kappa)) \\
& \leq \int_0^\kappa G^i(s, U(t + \kappa - s, x(s))) - G^i(s, V(t + \kappa - s, y(s))) + \epsilon ds \\
& + U^i(t, x(\kappa)) - U^i(t, y(\kappa)).
\end{aligned}$$

Let us now remark that, since U is uniformly Lipschitz, we deduce that

$$(1.31) \quad |U^i(t, x(\kappa)) - U^i(t, y(\kappa))| \leq C|x(\kappa) - y(\kappa)|.$$

From (1.28) and (1.29) we finally deduce that

$$(1.32) \quad |U^i(t, x(\kappa)) - U^i(t, y(\kappa))| \leq C\kappa\|U - V\|_\infty.$$

Hence, dividing by κ and taking the limit $\kappa \rightarrow 0$ in (1.30), we obtain

$$(1.33) \quad \frac{d}{dt}(U^i(t, x^*) - V^i(t, x^*)) \leq C\|U(t, \cdot) - V(t, \cdot)\|_\infty + \epsilon.$$

From this, we obtain that

$$(1.34) \quad \frac{d}{dt}\|U(t, \cdot) - V(t, \cdot)\|_\infty \leq C\|U(t, \cdot) - V(t, \cdot)\|_\infty + \epsilon.$$

Hence, we deduce from Grönwall's Lemma that

$$(1.35) \quad \|U(t, \cdot) - V(t, \cdot)\|_\infty + \epsilon \leq 2\epsilon e^{Ct},$$

from which the Proposition immediately follows. \square

Remark 1.13. *From the previous result we can indeed deduce that the Lipschitz solutions attracts the limits of approximations of (1.1), since the constant C in (1.35) does not depend on V_ϵ .*

2. THE HILBERTIAN CASE

In the so-called Hilbertian approach, introduced by the third author, the typical form of a MFG master equation with common noise, is

$$(2.1) \quad \partial_t U(t, X) - \sum_{i=1}^{\infty} \lambda_i \partial_{ii} U(t, X) + \langle F(X, U), \nabla \rangle U(t, X) = G(X, U) \text{ in } (0, \infty) \times H,$$

where $(H, \langle \cdot, \cdot \rangle)$ is a real separable Hilbert space, with an orthonormal family $(e_i)_{i \geq 1}$ and $F, G : H \times H \rightarrow H$. In the previous, the term $\langle F(x, U), \nabla \rangle U$ is understood in the sense that we are taking the Gateaux derivative of U in the direction $F(x, U)$ and ∂_{ii} refers to the second order derivative with respect to e_i .

The master equation (2.13) is related to more "classical" master equations set on the space of probability measures. A typical example of the latter is written below in

equation (3.34). The two approaches are concerned with the same model when:

$$(2.2) \quad \begin{aligned} H &= L^2(\Omega, \mathbb{R}^d), B = D_p H \\ \sigma &= \sigma' = 0, \lambda_i = \sigma_0 \mathbb{1}_{i \leq d}, F(X, U) = D_p H(X, U), \\ G(X, U) &= -D_x H(X, U, \mathcal{L}(X)), \end{aligned}$$

where $(\Omega, \mathcal{A}, \mathbb{P})$ is a standard probability space and $\mathcal{L}(X)$ denotes the law of the random variable X , and when the first d components of $(e_i)_{i \geq 1}$ are the component of the canonical basis of \mathbb{R}^d . In this case, we expect that given classical solutions U and \mathcal{U} of respectively (2.13) and (3.34), the following holds for all $X \in L^2(\Omega, \mathbb{R}^d)$,

$$(2.3) \quad \nabla_x \mathcal{U}(t, X, \mathcal{L}(X)) = U(t, X).$$

2.1. The deterministic case. As the theory of ODE in Hilbert spaces does not raise any particular difficulty, we can easily adapt the results of the previous part to the so-called deterministic equation (i.e. the case $\forall i, \lambda_i = 0$)

$$(2.4) \quad \partial_t U + \langle F(x, U), \nabla \rangle U = G(x, U) \text{ in } (0, \infty) \times H,$$

$$(2.5) \quad U(0, x) = U_0(x) \text{ in } H.$$

In this context, there is also an associated linear transport equation. Given two vectors fields $A, B : [0, T] \times H \rightarrow H$, we say that $V : [0, T] \times H \rightarrow H$ is a solution of the linear transport equation

$$(2.6) \quad \partial_t V - \langle B(t, x), \nabla \rangle V = A(t, x) \text{ in } (0, T) \times H$$

if V satisfies for all $t \in [0, T], x \in H$

$$(2.7) \quad V(t, x) = \int_0^t A(t-s, \xi(t, s, x)) ds + U_0(\xi(t, t, x)),$$

where $\xi(t, \cdot, \cdot)$ is the flow of the ODE

$$(2.8) \quad \frac{d}{ds} x(s) = B(t-s, x(s)).$$

Note that, as in the finite state space case, the flow is well defined as soon as B is Lipschitz in x , uniformly in t for instance. Hence in this case, the formula (2.7) makes sense as soon as A and U_0 are continuous for instance. We also define the operator Ψ with $\Psi(T, A, B, U_0)$ is the function given by (2.7). As in the previous section we can define a notion of Lipschitz solution and give a result of existence and uniqueness.

Definition 2.1. *A Lipschitz solution U of (2.4) on the time interval $[0, T]$ is a function such that*

- U is Lipschitz in $x \in H$, uniformly for $t \in [0, \alpha]$ for any $\alpha < T$.
- For any $t < T$

$$(2.9) \quad U = \Psi \left(t, G(\cdot, U), -F(\cdot, U), U_0 \right).$$

Theorem 2.2. *Assume that F and G are Lipschitz functions. For any Lipschitz initial condition $U_0 : H \rightarrow H$:*

- There always exists a time $T > 0$ such that there exists a unique solution of (1.1) in the sense of Definition 2.1.
- There exists a maximal time $T^c \in [0, \infty]$ and a solution U associated to T^c such that, for any solution V of the problem on an interval $[0, T]$: we have that $T \leq T^c$ and the restriction of U to $[0, T)$ is equal to V .
- If the maximal time T^c is such that $T^c < \infty$, then $\|D_x U(t)\|_\infty \rightarrow \infty$ as $t \rightarrow T^c$.

Proof. This proof is very similar to the one in the finite state space case, hence we only detail the main difference which is the fact that we consider an unbounded initial condition here. Since we assumed that F and G are globally Lipschitz, the only point where we need some uniform estimates on Lipschitz solutions is when we consider the set E_C on which we want to use some fixed points results. In this setting, it is natural to consider the set defined by

$$(2.10) \quad E_C = \left\{ U : [0, T) \times H \rightarrow U, \sup_{R>0} \sup_{t \in [0, T), |x| \leq R} R^{-1} |U(t, x)| < C, \|D_X U\|_\infty < C \right\}.$$

The argument of the previous proof can be carried on on those sets simply by remarking that now, all the functional convergence shall be locally uniformly in $x \in H$. \square

As an application of the previous result, let us consider the case of the master equation

$$(2.11) \quad \partial_t U + \langle U, \nabla_x \rangle U = 0 \text{ in } (0, \infty) \times H,$$

with initial condition $U_0(x) = A(x)$ for some linear operator A such that $A^* = A$. Remark that the solution of this master equation is simply given by $U(t, x) = A(t)x$, where $(A(t))_{t \geq 0}$ is the solution of

$$(2.12) \quad \frac{d}{dt} A(t) + A(t)^2 = 0.$$

Hence we are here in the situation $T^c < \infty$ except in the cases in which $A \geq 0$, which corresponds to the monotone regime which is known to propagate Lipschitz regularity for MFG master equation.

2.2. The case of common noise. We now turn to master equations involving a common noise, which is slightly more involved than the previous one. Moreover it will help us to understand how to extend this mathematical analysis to the case of master equations set on the space of probability measures. Recall that we are interested in the master equation

$$(2.13) \quad \partial_t U(t, x) - \sum_{i=1}^{\infty} \lambda_i \partial_{ii} U(t, x) + \langle F(x, U), \nabla \rangle U(t, x) = G(x, U) \text{ in } (0, \infty) \times H.$$

The presence of second order terms in this master equation imposes to use stochastic characteristics of the associated linear transport equation. Indeed, consider the linear (possibly degenerate) parabolic equation

$$(2.14) \quad \partial_t V - \sum_{i=1}^{\infty} \lambda_i \partial_{ii} V - \langle B(t, x), \nabla \rangle V = A(t, x) \text{ in } (0, \infty) \times H,$$

with initial condition

$$(2.15) \quad V|_{t=0}(x) = U_0(x) \text{ in } H,$$

where $A, B : [0, \infty) \times H \rightarrow H$ and $U_0 : H \rightarrow H$. Let us consider, for $t > 0$ the stochastic differential equation (SDE in short) in H that we write component wise on the family $(e_i)_{i \geq 1}$

$$(2.16) \quad dX_s^i = B^i(t-s, X_s)ds + \sqrt{2\lambda_i}dW_s^i,$$

where $(W^i)_{1 \leq i}$ is a collection of independent real Brownian motions on $(\Omega, \mathcal{A}, \mathbb{P})$. In the case in which B is Lipschitz in $x \in H$, uniformly in t , the SDE is well defined, it even admits strong solutions. In this context, it is thus meaningful to introduce the following Feynman-Kac representation formula for $t < T, x \in H$

$$(2.17) \quad V(t, x) = \mathbb{E} \left[\int_0^t A(t-s, X_s)ds + U_0(X_t) | X_0 = x \right],$$

where $(X_s)_{s \geq 0}$ is of course the solution of (2.16), conditioned here to take initial value x . The function V defined by this formula is denoted once again by

$$(2.18) \quad V = \Psi(T, A, B, U_0).$$

This remarks naturally leads us to the

Definition 2.3. *Given a time $T > 0$, a function $U : [0, T] \times H \rightarrow H$ is called a Lipschitz solution of (2.13) if*

- *U is Lipschitz in $x \in H$, uniformly in $t \in [0, \alpha]$ for any $\alpha \in [0, T)$,*
- *For any $t < T$,*

$$(2.19) \quad U = \Psi(t, G(\cdot, U), -F(\cdot, U), U_0).$$

As in the previous cases, we can establish the following Cauchy-Lipschitz like result.

Theorem 2.4. *Assume that F and G are Lipschitz functions and that $\sum_{i=1}^{\infty} \lambda_i < \infty$. For any Lipschitz initial condition U_0 :*

- *There always exists a time $T > 0$ such that there exists a unique solution of (2.13) in the sense of Definition 2.3.*
- *There exists a maximal time $T^c \in [0, \infty]$ and a solution U associated to T^c such that, for any solution V of the problem on an interval $[0, T)$: we have that $T \leq T^c$ and the restriction of U to $[0, T)$ is equal to V .*
- *If $T^c < \infty$, then $\|D_x U(t)\|_{\infty} \rightarrow \infty$ as $t \rightarrow T^c$.*

Proof. Consider $T > 0$, a bounded Lipschitz function U_0 and the function Φ defined by

$$(2.20) \quad \Phi(U) = \Psi(T, G(\cdot, U), -F(\cdot, U), U_0).$$

Let us consider the set E_C for $C > 0$, defined in (2.10).

Step 1: Φ is well defined. Consider a function $U : [0, T) \times H \rightarrow H$ which is, uniformly in t , Lipschitz in $x \in H$. Then, uniformly in t , it is also the case for $x \rightarrow F(x, U(t, x))$ and $x \rightarrow G(x, U(t, x))$. This implies that, given $t > 0$ and an initial condition, the SDE

$$(2.21) \quad dX_s^i = -F^i(X_s, U(t-s, X_s))ds + \sqrt{2\lambda_i}dW_s^i,$$

is well defined. Denote by $(X_s^x)_{t \in [0, T]}$ the strong solution of the previous SDE with initial conditions $x \in H$. Remark that, for any $x \in H$, $(X_s^x)_{s \geq 0}$ satisfies

$$(2.22) \quad \begin{aligned} \mathbb{E}[|X_s^x|] &\leq C \int_0^s \mathbb{E}[|X_{s'}|] ds' + \mathbb{E} \left[\left(\sum_i 2\lambda_i (W_s^i)^2 \right)^{\frac{1}{2}} \right] \\ &\leq C \int_0^s \mathbb{E}[|X_{s'}|] ds' + C(1 + \sum_i \lambda_i). \end{aligned}$$

Hence $\sup_{s \geq 0} \mathbb{E}[|X_s^x|] < \infty$ and thus the function Φ is indeed well defined on the set of functions U such that $\|D_x U\|_\infty < \infty$. It is valued in E .

Step 2: $\Phi(E_C) \subset E_C$ for C large enough and T small enough. From (2.22), we immediately deduce that if T is small enough and C is large enough, then $U \in E_C \Rightarrow \|\Phi(U)\|_\infty \leq C$. Let us now compute for $x, y \in H$ and $t > 0$

$$(2.23) \quad \begin{aligned} |\Phi(U)(t, x) - \Phi(U)(t, y)| &= \left| \mathbb{E} \left[\int_0^t G(X_s^x, U(t-s, X_s^x)) - G(X_s^y, U(t-s, X_s^y)) ds \right] \right| \\ &\quad + |\mathbb{E}[U_0(X_t^x) - U_0(X_t^y)]| \\ &\leq C \mathbb{E}[|X_t^x - X_t^y|] + \mathbb{E} \left[\int_0^t C |X_s^x - X_s^y| ds \right]. \end{aligned}$$

The following estimate holds almost surely

$$(2.24) \quad d(X_t^x - X_t^y) \leq \|D_X F\|_\infty |X_t^x - X_t^y| + \|D_p F\|_\infty \|D_x U\|_\infty |X_t^x - X_t^y|$$

Hence we deduce from Grönwall's Lemma and from (2.23) that $D_x \Phi(U)$ is uniformly bounded and that, choosing T sufficiently small, we can guarantee that there exists $C > 0$ such that

$$(2.25) \quad \sup_{t \leq T} \|D_X U(t)\|_\infty \leq C \Rightarrow \sup_{t \leq T} \|D_X \Phi(U)(t)\|_\infty \leq C.$$

Step 3: Φ is a contraction if T is small enough. Consider $U, V \in E_C$, and let us compute for $t \leq T, x \in H$

$$(2.26) \quad \begin{aligned} &|\Phi(U)(t, x) - \Phi(V)(t, x)| = \\ &= \left| \mathbb{E} \left[U_0(X_t) - U_0(\tilde{X}_t) + \int_0^t G(X_s, U(t-s, X_s)) - G(\tilde{X}_s, V(t-s, \tilde{X}_s)) ds \right] \right|, \end{aligned}$$

where $(X_s)_{s \in [0, T]}$ is the strong solution of (2.16) with initial condition x and $(\tilde{X}_s)_{s \in [0, T]}$ is the strong solution of (2.16) with initial condition x when U has been replaced by V . Hence, almost surely, we have the estimate

$$(2.27) \quad d|X_s - \tilde{X}_s| \leq \|D_x F\|_\infty |X_s - \tilde{X}_s| + \|D_p F\|_\infty (\|U - V\|_\infty + \|D_x U\|_\infty |X_s - \tilde{X}_s|).$$

From this estimate, we immediately deduce from (2.26) that Φ is indeed a contraction for the $\|\cdot\|_\infty$ norm if T is small enough. The rest of the proof follows exactly the same argument as the proof of Theorem 1.7. □

3. MASTER EQUATIONS ON THE SET OF PROBABILITY MEASURES

3.1. Setting and notation. In this section, we address master equations set on the space of probability measures in cases with and without common noise. We start by presenting the case of the master equation without common noise and we turn to the general case later on. Hence we study first the PDE

$$(3.1) \quad \begin{aligned} & \partial_t U(t, x, m) + H(x, \nabla_x U(t, x, m), m) - \sigma \Delta_x U(t, x, m) \\ & + \int_{\mathbb{R}^d} B(y, \nabla_x U(t, y, m)) D_m U(t, x, m, y) m(dy) \\ & - \sigma' \int_{\mathbb{R}^d} \operatorname{div}_y (D_m U(t, x, m, y)) dm(y) = 0 \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned}$$

$$(3.2) \quad U(0, x, m) = U_0(x, m) \text{ in } \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d).$$

Here, \mathbb{T}^d is the d dimensional torus, $\mathcal{P}(\mathbb{T}^d)$ is the set of probability measures on \mathbb{T}^d , $H : (x, p, m) \rightarrow \mathbb{R}$ and $B : (x, p, m) \rightarrow \mathbb{R}^d$ are given functions, $\sigma, \sigma' > 0$ are constants.

Remark 3.1. *In a lot of cases studied in the literature, B is equal to $D_p H$ and $\sigma = \sigma'$. Because these assumptions play no role here, we remove them, just as it was the case in the previous sections.*

Remark 3.2. *The choice of the d dimensional torus \mathbb{T}^d as the state space does not play any particular role except the one of simplifying the formulation of some statements in the following. Moreover, the setting at hand is sometimes refers to as the one of "extended MFG" as in [21].*

Remark 3.3. *The same study could be carried on if the coefficients σ and σ' depend on x and m , provided that this dependence is sufficiently smooth, but we do not consider this case to keep the following more understandable.*

The derivatives with respect to the measure argument are defined in the following way. For a function $F : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$, when it is defined, we denote for $m \in \mathcal{P}(\mathbb{T}^d)$, $x \in \mathbb{T}^d$

$$(3.3) \quad \nabla_m F(m, x) = \lim_{h \rightarrow 0} \frac{F((1-h)m + h\delta_x) - F(m)}{h},$$

where δ_x is the Dirac mass at x . Furthermore, when it is defined, we note for $m \in \mathcal{P}(\mathbb{T}^d)$, $x \in \mathbb{T}^d$,

$$(3.4) \quad D_m F(m, x) = \nabla_x \nabla_m F(m, x).$$

Let us remark that if $\nabla_m F(m, \cdot)$ and $D_m F(m, \cdot)$ are well defined, then for $m, m' \in \mathcal{P}(\mathbb{T}^d)$, $\phi : \mathbb{T}^d \rightarrow \mathbb{R}^d$

$$(3.5) \quad \int_{\mathbb{T}^d} \nabla_m F(m, x) (m' - m)(dx) = \lim_{h \rightarrow 0} \frac{F((1-h)m + hm') - F(m)}{h},$$

$$(3.6) \quad \int_{\mathbb{T}^d} D_m F(m, x) \cdot \phi(x) m(dx) = \lim_{h \rightarrow 0} \frac{F((Id + h\phi)_\# m) - F(m)}{h},$$

where $T_\# m$ denotes the image measure of the measure m by the map T .

In all the following, we are going to equip $\mathcal{P}(\mathbb{T}^d)$ with the Monge-Kantorovich distance \mathbf{d}_1 defined by

$$(3.7) \quad \mathbf{d}_1(m, m') = \sup_{f, \|\nabla_x f\|_\infty \leq 1} \int_{\mathbb{T}^d} f d(m - m').$$

This distance is a metric for the weak convergence of measures. Moreover, it is a norm when extended to the set of measures on \mathbb{T}^d . Considering a function $F : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d$, we also introduce the notation

$$(3.8) \quad \text{Lip}(F) = \sup_{x, y, m} |x - y|^{-1} |F(x, m) - F(y, m)| + \sup_{x, m, m'} \mathbf{d}_1(m, m')^{-1} |F(x, m) - F(x, m')|.$$

Remark 3.4. *Having chosen this distance, adapting the following results to master equations which are set on sets of measures with different masses is quite immediate. Even though we are not going to discuss anymore this fact, we believe it is worth insisting on the fact that the choice of \mathbf{d}_1 is particularly natural because of all the MFG in which the mass of players does not remain constant.*

Several approaches can be taken here. We present one in details and sketch a second one later on. This first approach is hinted by the previous section which suggests to consider the PDE satisfies by the function $W(t, x, m) := \nabla_x U(t, x, m)$ where U is the value function of the MFG. If the value function U is indeed a (say classical) solution of (3.1), then its spatial gradient W is a solution of

$$(3.9) \quad \begin{aligned} & \partial_t W + D_p H(x, W, m) \cdot \nabla_x W - \sigma \Delta_x W(t, x, m) \\ & + \int_{\mathbb{R}^d} B(y, W(t, y, m), m) D_m W(t, x, m, y) m(dy) \\ & - \sigma' \int_{\mathbb{R}^d} \text{div}_y (D_m W(t, x, m, y)) dm(y) = -\nabla_x H(x, W, m) \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned}$$

The transport equation naturally associated to this previous nonlinear PDE is simply

$$(3.10) \quad \begin{aligned} & \partial_t V(t, x, m) - b(t, x, m) \cdot \nabla_x V(t, x, m) - \sigma \Delta_x V(t, x, m) \\ & + \int_{\mathbb{R}^d} F(t, x, m) \cdot D_m V(t, x, m, y) m(dy) \\ & - \sigma' \int_{\mathbb{R}^d} \text{div}_y (D_m V(t, x, m, y)) dm(y) = A(t, x, m) \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned}$$

where $b, B, A : [0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d$ are given vector fields.

As in the previous settings, the solutions of the transport equation (3.10) are naturally given by a representation formula. Moreover, from the presence of a second order term in x , the evolution equation associated to this variable is stochastic. This does not raise any particular mathematical difficulty. The associated Feynman-Kac representation formula

is then

$$(3.11) \quad \forall t \in [0, T], x \in \mathbb{T}^d, m_* \in \mathcal{P}(\mathbb{T}^d), V(t, x, m_*) = \mathbb{E} \left[\int_0^t A(t-s, X_s, m(s)) ds + U_0(X_t, m(t)) \right],$$

where $(X_s, m(s))_{s \geq 0}$ is the solution of the following SDE-PDE system

$$(3.12) \quad \begin{aligned} dX_s &= b(t-s, X_s, m(s)) ds + \sqrt{2\sigma} dW_s \text{ for } s \leq t, \\ \partial_s m &= -\operatorname{div}(F(t-s, x, m(s))m(s)) + \sigma' \Delta_x m \text{ in } (0, t) \times \mathbb{T}^d, \end{aligned}$$

with initial conditions x and m_* , where $(W_s)_{s \geq 0}$ is a d dimensional Brownian motion on a standard probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let us recall that the link with equation (3.10) can be observed by computing

$$(3.13) \quad (dt)^{-1} V(t, x, m_*) - \mathbb{E}[V(t-dt, X_{dt}, m(dt))]$$

and letting $dt \rightarrow 0$.

We denote by Ψ the operator which is defined by the representation formula (3.11). That is, the function V given by the right hand side of (3.11) is denoted by $V = \Psi(T, b, F, A, U_0)$.

3.2. Main definition and result. We can now easily state the following definition of solution of (3.9).

Definition 3.5. *Given an initial condition U_0 , a Lipschitz solution of (3.9) on $[0, T]$ is a function $W : [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d$ such that*

- *W is Lipschitz in x, m , uniformly in $[0, t]$ for any $t < T$.*
- *For any $t < T$,*

$$(3.14) \quad W = \Psi(t, D_p H(x, W, m), B(x, W, m), -D_x H(x, W, m), U_0).$$

Remark 3.6. *In this context, W is not asked to be a gradient in x .*

We can produce the same type of result for this notion of solution.

Theorem 3.7. *Assume that:*

- *The function H is such that $D_x H$ and $D_p H$ are globally Lipschitz functions.*
- *The function B is a globally Lipschitz function.*

Then, for any initial condition U_0 such that $\nabla_x U_0$ is Lipschitz:

- *There always exists a time $T > 0$ such that there exists a unique solution W of (3.1) in the sense of Definition 3.5.*
- *There exists a maximal time $T^c \in [0, \infty]$ and a solution W associated to T^c such that, for any solution V of the problem on an interval $[0, T]$: we have that $T \leq T^c$ and the restriction of W to $[0, T)$ is equal to V .*
- *If $T^c < \infty$, then $\operatorname{Lip}(W(t, \cdot, \cdot)) \rightarrow \infty$ as $t \rightarrow T^c$.*

Proof. The proof of this result follows the same line of argument as the previous ones. Consider a time $T > 0$, the map

$$(3.15) \quad \Phi(W) = \Psi(T, D_p H(x, W), B(x, W, m), -D_x H(x, W, m), U_0).$$

and the set

$$(3.16) \quad E_C := \{W : [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d, \|W\|_\infty \leq C, \sup_{t \leq T} \text{Lip}(W(t, \cdot, \cdot)) \leq C\}.$$

Step 1: Φ is well defined. Take $W \in E_\infty$. Setting $b(t, x, m) = -D_p H(x, W(t, x, m))$ and $F(t, x, m) = B(x, U(t, x, m), m)$, we remark that the system (3.12) always has strong solutions for $t < T$. Hence from the continuity of $\nabla_x U_0$ and $\nabla_x H$, we deduce that Φ is indeed well defined.

Let us now recall a standard estimate on Fokker-Planck equations. Consider m_1 and m_2 solutions of

$$(3.17) \quad \begin{aligned} \partial_s m_i - \sigma' \Delta_x m_i + \text{div}(b_i(s, x) m_i) &= 0 \text{ in } (0, T) \times \mathbb{T}^d \text{ for } i = 1, 2, \\ m_i|_{s=0} &= \mu, \end{aligned}$$

where $b_1, b_2 : [0, T] \times \mathbb{T}^d \rightarrow \mathbb{R}^d$ are both bounded in s, x , and Lipschitz continuous in x , uniformly on $[0, t]$ for $t < T$. Then, for any $s \leq t < T$, there exists C_t depending only on the bounds on b such that

$$(3.18) \quad \mathbf{d}_1(m_1(s), m_2(s)) \leq C \int_0^t \|b_1(s) - b_2(s)\|_\infty ds.$$

This estimate is classical and a proof is provided in appendix for the sake of completeness.

Let us show that $\Phi(E_\infty) \subset E_\infty$. Consider $W \in E_\infty$, m_1 and m_2 and let us compute for $t < T, x \in \mathbb{T}^d$,

$$(3.19) \quad \begin{aligned} |\Phi(W)(t, x, m_1) - \Phi(W)(t, x, m_2)| &= \mathbb{E} \left[\int_0^t -\nabla_x H(X_{1,s}, W(t-s, X_{1,s}, m_1(s)), m(s)) \right. \\ &\quad \left. + \nabla_x H(X_{2,s}, W(s, X_{2,s}, m_2(s)), m_2(s)) ds \right. \\ &\quad \left. + \nabla_x U_0(X_{1,t}, m_1(t)) - \nabla_x U_0(X_{2,t}, m_2(t)) \right], \end{aligned}$$

where the $(X_{i,s}, m_i(s))_{s \geq 0}$ are the solutions of

$$(3.20) \quad \begin{aligned} dX_{i,s} &= D_p H(X_{i,s}, W(t-s, X_{i,s}, m_i(s)), m_i(s)) ds + \sqrt{2\sigma} dW_s \text{ for } s \leq t, \\ \partial_s m_i &= -\text{div}(B(x, W(t-s, x, m_i)) m_i) + \sigma' \Delta_x m_i \text{ in } (0, t) \times \mathbb{T}^d, \end{aligned}$$

with initial conditions m_i and $X_{i,0} = x$. Furthermore, since $W \in E_\infty$, we can use the estimate (3.18) to deduce that, for $t \leq T$, there exists $C > 0$, depending only on the data of the problem, such that

$$(3.21) \quad \mathbf{d}_1(m_1(t), m_2(t)) \leq e^{C \text{Lip}(W)t} \mathbf{d}_1(m_1, m_2).$$

Moreover, the following holds almost surely

$$(3.22) \quad \begin{aligned} d|X_{1,s} - X_{2,s}| &\leq \|D_{pp} H\|_\infty \text{Lip}(W) (\mathbf{d}_1(m_1(s), m_2(s)) + |X_{1,s} - X_{2,s}|) \\ &\quad + \|D_{px} H\|_\infty |X_{1,s} - X_{2,s}| \\ &\leq C \text{Lip}(W) |X_{1,s} - X_{2,s}| + e^{C \text{Lip}(W)T} \mathbf{d}_1(m_1, m_2). \end{aligned}$$

Hence, we obtain using Grönwall's Lemma that, almost surely, for $s \leq T$

$$(3.23) \quad |X_{1,s} - X_{2,s}| \leq (e^{C\text{Lip}(W)s} - 1)e^{C\text{Lip}(W)T} \mathbf{d}_1(m_1, m_2).$$

Using (3.21) and (3.23) in (3.19) and using the regularity assumptions on H and f , we finally deduce that

$$(3.24) \quad |\Phi(W)(t, x, m_1) - \Phi(W)(t, x, m_2)| \leq Ce^{C\text{Lip}(W)T} \mathbf{d}_1(m_1, m_2),$$

where $C > 0$ is a constant which depends only on U_0, H and B . The same type of result is also true for estimating the Lipschitz constant of $\Phi(W)$ in x but we do not present it here. Thus, it follows that $\Phi(E_\infty) \subset E_\infty$.

Moreover, if C is sufficiently large, then for T sufficiently small, $\Phi(E_C) \subset E_C$.

Step 2: Φ is a contraction. We now show that, if T is small enough, then Φ is a contraction. Take $W_1, W_2 \in E$ and compute for $t \leq T, x \in \mathbb{T}^d, m \in \mathcal{P}(\mathbb{T}^d)$

$$(3.25) \quad |\Phi(W_1)(t, x, m) - \Phi(W_2)(t, x, m)| = \left| \mathbb{E} \left[\int_0^t -\nabla_x H(X_{1,s}, W_1(t-s, X_{1,s}, m_1(s)), m_1(s)) \right. \right. \\ \left. \left. + \nabla_x H(X_{2,s}, W_2(t-s, X_{2,s}, m_2(s)), m_2(s)) ds \right. \right. \\ \left. \left. + \nabla_x U_0(X_{1,t}, m_1(t)) - \nabla_x U_0(X_{2,t}, m_2(t)) \right] \right|,$$

where the $(X_{i,s}, m_i(s))_{s \geq 0}$ are the solutions of

$$(3.26) \quad \begin{aligned} dX_{i,s} &= D_p H(X_{i,s}, W_i(t-s, X_{i,s}, m_i(s)), m_i(s)) ds + \sqrt{2\sigma} dW_s \text{ for } s \leq t, \\ \partial_s m_i &= -\text{div}(B(x, W_i(t-s, x, m_i), m_i)m_i) + \sigma' \Delta_x m_i \text{ in } (0, t) \times \mathbb{T}^d, \end{aligned}$$

with initial conditions $m_i = m$ and $X_{i,0} = x$. From estimate (3.18) and the regularity of B , we obtain that for $t \leq T$, there exists $C > 0$ such that

$$(3.27) \quad \begin{aligned} \mathbf{d}_1(m_1(t), m_2(t)) &\leq C \int_0^t \|W_1(t-s, \cdot, m_1(s)) - W_2(t-s, \cdot, m_2(s))\|_\infty ds \\ &\leq C \int_0^t \text{Lip}(W_1) \mathbf{d}_1(m_1(s), m_2(s)) + \|W_1(t-s, \cdot, \cdot) - W_2(t-s, \cdot, \cdot)\|_\infty ds. \end{aligned}$$

Using once again Grönwall's Lemma, we obtain that

$$(3.28) \quad \mathbf{d}_1(m_1(t), m_2(t)) \leq C(e^{C\text{Lip}(W_1)t} - 1) \|W_1 - W_2\|_\infty.$$

Hence it follows that, for any $\alpha \in (0, 1)$, if T is chosen small enough (where C was already chosen sufficiently large), for all $W_1, W_2 \in E_C$

$$(3.29) \quad \mathbf{d}_1(m_1(t), m_2(t)) \leq \alpha \|W_1 - W_2\|_\infty.$$

Using the same type of argument as in **Step 1**, we can obtain the same estimate (almost surely) on $|X_{1,t} - X_{2,t}|$. From this, recalling (3.25), we finally obtain that

$$(3.30) \quad \|\Phi(W_1) - \Phi(W_2)\|_\infty \leq \alpha \|W_1 - W_2\|_\infty.$$

Hence Φ is a contraction from $(E_C, \|\cdot\|_\infty)$ into itself.

Step 3: Existence of a fixed point. From the previous step, we deduce that for $W \in E_C$, $(\Phi^n(W))_{n \geq 0}$ is a Cauchy sequence. Moreover it is valued in E_C and $(\|\partial_t \Phi^n(W)\|_\infty)_{n \geq 1}$ is a bounded sequence. From Ascoli-Arzelà Theorem, we deduce that $(\Phi^n(W))_{n \geq 0}$ converges uniformly to some $W_* \in E_C$. From the continuity of Φ , $\Phi(W_*) = W_*$.

The rest of the result follows quite easily from standard arguments. \square

Moreover, we also have the

Corollary 3.8. *The result of the Theorem remains true if the dependence on W in H and B is non local, as long as the Lipschitz regularity holds for the $\|\cdot\|_\infty$ norm.*

Proof. It suffices to follow the previous proof and remark that we never used explicitly that the dependence was local, and that we always used the $\|\cdot\|_\infty$ norm anyway. \square

3.3. Return to the solutions of the initial master equation. To complete the study of (3.1), we now explain how we can use the knowledge of the solutions of (3.9) to define solutions of (3.1).

On the time interval $[0, T]$ for $T > 0$, if W is the unique Lipschitz solution of (3.9) and U is the unique solution of (3.1), then we expect that $\nabla_x U = W$. Hence we expect that U is a solution of

$$(3.31) \quad \begin{aligned} & \partial_t U + H(x, W(t, x, m), m) - \sigma \Delta_x U(t, x, m) \\ & + \int_{\mathbb{R}^d} B(y, W(t, y, m), m) D_m U(t, x, m, y) m(dy) \\ & - \sigma' \int_{\mathbb{R}^d} \operatorname{div}_y (D_m U(t, x, m, y)) dm(y) = 0 \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d). \end{aligned}$$

The previous PDE is a linear transport equation in U . Hence it can be dealt with by means of the following representation formula, for $t < T$, $x \in \mathcal{P}(\mathbb{T}^d)$, $\mu \in \mathcal{P}(\mathbb{T}^d)$

$$(3.32) \quad U(t, x, \mu) = \mathbb{E} \left[\int_0^t -H(X_s, W(t-s, X_s, m(s))) ds + U_0(X_t, m(t)) \right],$$

where $(X_s, m(s))_{s \in [0, t]}$ is the solution of

$$(3.33) \quad \begin{aligned} & dX_s = \sqrt{2\sigma} dW_s \text{ for } s \in (0, t), \\ & \partial_t m = -\operatorname{div}(B(x, W(t-s, x, m))m) + \sigma' \Delta_x m \text{ in } (0, t) \times \mathbb{T}^d, \end{aligned}$$

with initial conditions x and μ . This leads us to the definition

Definition 3.9. *A bounded function $U : [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ is a Lipschitz solution of (3.1) on the time interval $[0, T]$ if there exists W , Lipschitz solution of (3.9) on $[0, T]$ such that U satisfies (3.32).*

Remark 3.10. *Let us insist that we do not need to impose the facts that W satisfies $W = \nabla_x U$ nor that it is a gradient. Somehow, we translate here the fact that, given the controls of the players, i.e. the function W , we can simply compute the value by following the characteristics.*

Remark 3.11. *This Definition makes clear that, in this setting, the appropriate Lipschitz regularity is on $W = \nabla_x U$ and not on U . This type of fact is often interpreted in the literature as the fact that (3.1) is an equation on the controls of the players as well as on the value.*

As a consequence of Theorem 3.7, we obtain the

Theorem 3.12. *Under the assumptions of Theorem 3.7:*

- *There exists $T > 0$ such that there is a Lipschitz solution U of (3.1) on $[0, T)$.*
- *There is a maximal time of existence $T^c \in (0, \infty]$ and it is such that any Lipschitz solution V of (3.1) on a time interval $[0, T)$ is such that $T \leq T^c$ and the restriction of U to $[0, T)$ is equal to V .*

Proof. This result is a direct application of Theorem 3.7. □

Remark 3.13. *Another notion of solution of (3.1) could have been introduced, based on the existence of a solution of (3.9). Indeed we could have replaced all the terms involving U in (3.1), except $\partial_t U$, by terms involving W . It then suffices to check that all the terms in W are bounded and thus that this new equation characterizes $\partial_t U$. However, we did not use this route since the one we choose allows us to avoid technical problems such as giving a precise sense to the term $\nabla_m W$, which should appear with this new method.*

3.4. Master equations associated to common noise. As in the Hilbertian case, the notion of Lipschitz solutions can easily be adapted to cases involving so-called common noises. Let us recall that the presence of a common noise (which should probably be called common shocks) in a MFG usually translates into the addition of terms in the master equation, which can be either non-local or of higher order. We refer to [4] for several examples of possible common noises. Even though the previous approach seems quite general and should work in all the cases, we focus here on the following master equation.

$$\begin{aligned}
 (3.34) \quad & \partial_t U(t, x, m) + H(x, \nabla_x U(t, x, m), m) - (\sigma + \sigma_0) \Delta_x U(t, x, m) \\
 & + \int_{\mathbb{R}^d} B(y, \nabla_x U(t, y, m), m) D_m U(t, x, m, y) m(dy) \\
 & - (\sigma' + \sigma_0) \int_{\mathbb{R}^d} \operatorname{div}_y (D_m U(t, x, m, y)) dm(y) - 2\sigma_0 \int_{\mathbb{T}^d} \operatorname{div}_x (D_m U(t, x, m, y)) m(dy) \\
 & - \sigma_0 \int_{\mathbb{T}^{2d}} \operatorname{Tr}[D_{mm}^2 U(t, x, m, y, z)] m(dy) m(dz) = 0 \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d),
 \end{aligned}$$

with initial condition $U_0 : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$. A study similar to the one we just conducted can be done here. Indeed, we can consider first the equation satisfied by $W = \nabla_x U$. In

this context, this PDE is

$$\begin{aligned}
(3.35) \quad & \partial_t W(t, x, m) + D_p H(x, W(t, x, m), m) \cdot \nabla_x W(t, x, m) - (\sigma + \sigma_0) \Delta_x W(t, x, m) \\
& + \int_{\mathbb{R}^d} B(y, W(t, y, m), m) D_m W(t, x, m, y) m(dy) + \nabla_x H(x, W(t, x, m), m) \\
& - (\sigma' + \sigma_0) \int_{\mathbb{R}^d} \operatorname{div}_y (D_m W(t, x, m, y)) dm(y) - 2\sigma_0 \int_{\mathbb{T}^d} \operatorname{div}_x (D_m W(t, x, m, y)) m(dy) \\
& - \sigma_0 \int_{\mathbb{T}^{2d}} \operatorname{Tr}[D_{mm}^2 W(t, x, m, y, z)] m(dy) m(dz) = 0 \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d).
\end{aligned}$$

Furthermore, associated to this equation, we can also consider the linear transport equation

$$\begin{aligned}
(3.36) \quad & \partial_t V(t, x, m) - b(t, x, m) \cdot \nabla_x V(t, x, m) - (\sigma + \sigma_0) \Delta_x V(t, x, m) \\
& + \int_{\mathbb{R}^d} F(t, y, m) D_m V(t, x, m, y) m(dy) \\
& - (\sigma' + \sigma_0) \int_{\mathbb{R}^d} \operatorname{div}_y (D_m V(t, x, m, y)) dm(y) - 2\sigma_0 \int_{\mathbb{T}^d} \operatorname{div}_x (D_m V(t, x, m, y)) m(dy) \\
& - \sigma_0 \int_{\mathbb{T}^{2d}} \operatorname{Tr}[D_{mm}^2 V(t, x, m, y, z)] m(dy) m(dz) = A(t, x, m) \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d),
\end{aligned}$$

which is associated to the system of SDE-stochastic PDE

$$\begin{aligned}
(3.37) \quad & dX_s = b(t-s, X_s, m_s) ds + \sqrt{2\sigma} dW_s + \sqrt{2\sigma_0} dW'_s \text{ for } s \in (0, t), \\
& dm_s = [(\sigma' + \sigma_0) \Delta_x m - \operatorname{div}(F(t-s, x, m_s) m_s)] ds - \operatorname{div}(m_s \sqrt{2\sigma_0} dW'_s) \text{ in } (0, t) \times \mathbb{T}^d,
\end{aligned}$$

where $(W_t)_{t \geq 0}$ and $(W'_t)_{t \geq 0}$ are two independent Brownian motions on the standard probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Note that the fact that $(W')_{t \geq 0}$ appears in the two equations is fundamental to obtain the crossed derivatives term which is the term in $2\sigma_0$ in (3.34).

Let us insist on the fact that the previous system has a unique strong solution as soon as b and F are Lipschitz in x, m , uniformly in time. Indeed, the additional Brownian motion merely acts as a translation here and does not perturb too much the mathematical analysis. As we did several times above, we can associate a representation formula to (3.36) which reads for $t \geq 0, x \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d)$

$$(3.38) \quad V(t, x, \mu) = \mathbb{E} \left[\int_0^t A(t-s, X_s, m_s) ds + U_0(X_t, m_t) \right],$$

where $(X_s, m_s)_{s \in [0, t]}$ is the unique (strong) solution of (3.37) with initial conditions x and μ . Denoting by Ψ the operator defined by this formula, we can introduce the

Definition 3.14. *Given $T > 0$, a Lipschitz solution of (3.35) on $[0, T)$ is a function $W : [0, T) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d$ such that*

- W is Lipschitz in x, m , uniformly in $t \in [0, \alpha]$ for $\alpha < T$.
- The following holds for any $t < T$

$$(3.39) \quad W = \Psi(T, -\nabla_x H(x, W), D_p H(x, W, m), B(x, W, m), U_0).$$

We can establish the following result

Theorem 3.15. *Under the assumptions of Theorem 3.7:*

- *There always exists a time $T > 0$ such that there exists a unique solution W of (3.35) on $[0, T)$ in the sense of Definition 3.14.*
- *There exists a maximal time $T^c \in [0, \infty]$ and a solution W associated to T^c such that, for any solution V of the problem on an interval $[0, T)$: we have that $T \leq T^c$ and the restriction of W to $[0, T)$ is equal to V .*
- *If $T^c < \infty$, then $\text{Lip}(W(t, \cdot, \cdot)) \rightarrow \infty$ as $t \rightarrow T^c$.*

The previous statement is word for word the same as the one in the case $\sigma' = 0$ and the same almost holds for their proofs. That is why we only sketch the proof here, mainly by highlighting the main differences with the proof of Theorem 3.7.

Proof. We use the same notation as in the proof of Theorem 3.7.

The only key argument here consists in showing that the estimate (3.18) can also be used in this stochastic case. Consider m_1 and m_2 , two solutions of

$$(3.40) \quad dm_{i,s} = [(\sigma' + \sigma_0)\Delta_x m_{i,s} - \text{div}(b_i(s, x)m_{i,s})]ds - \text{div}(m_{i,s}\sqrt{2\sigma_0}dW'_s) \text{ in } (0, t) \times \mathbb{T}^d,$$

for $(W'_t)_{t \geq 0}$ a standard Brownian motion and $b_1, b_2 : [0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d$ two bounded vector fields. Consider \tilde{m}_1 and \tilde{m}_2 given by

$$(3.41) \quad \tilde{m}_{i,s} = (\tau_{\sqrt{2\sigma'}W'_s})_{\#} m_{i,s},$$

where $\tau_x : \mathbb{T}^d \rightarrow \mathbb{T}^d$ is the translation of x and $T_{\#}\mu$ denotes the image measure of the measure μ by the map T . Let us remark that, for any $s \geq 0$

$$(3.42) \quad \mathbf{d}_1(m_{1,s}, m_{2,s}) = \mathbf{d}_1(\tilde{m}_{1,s}, \tilde{m}_{2,s}).$$

We now observe that, by construction, for any $\omega \in \Omega$,

$$(3.43) \quad \partial_t \tilde{m}_i - \sigma' \Delta \tilde{m}_i + \text{div}(\tilde{b}_i \tilde{m}_i) = 0 \text{ in } (0, t) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d),$$

where $\tilde{b}_i(s, x) := b_i(s, x + \sqrt{2\sigma_0}W'_s)$. Remarking finally that

$$(3.44) \quad \|\tilde{b}_1 - \tilde{b}_2\|_{\infty} = \|b_1 - b_2\|_{\infty},$$

we deduce that the estimate (3.18) is satisfied almost surely (with a constant independent of $\omega \in \Omega$) in this stochastic case.

The rest of the proof follows the same argument as in the case without common noise. \square

As in the case without common noise, we can of course use a notion of solution of (3.35) to establish a definition and results on solutions of (3.34). Although we do not detail it here as it will merely be a copy of the previous case.

3.5. Master equations involving the image measure. In several MFG models, see for instance [13, 9, 5, 17], the dependence on m of the non-linearities $D_p H$ and B in (3.1) happens through the image measure of m by a certain function of the gradient in x of the value function. Hence, in such cases, the non-linearities have a form similar to $\tilde{A} : \mathbb{T}^d \times (\mathbb{T}^d \rightarrow \mathbb{R}^d) \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d$

$$(3.45) \quad \tilde{A}(x, \phi, m) := A(x, m, \phi(x), \psi(\phi)_{\#}m).$$

where $A : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d$, $\psi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $T_{\#}m$ denotes the image measure of m by the map T . The following result holds.

Proposition 3.16. *If A is Lipschitz, then so is \tilde{A} on $\mathbb{T}^d \times E_C \times \mathcal{P}(\mathbb{T}^d)$, where E_C is the set of Lipschitz functions $\mathbb{T}^d \rightarrow \mathbb{R}^d$ with Lipschitz constant at most C .*

Proof. Let us first compute, for ϕ, ϕ' two Lipschitz functions $\mathbb{T}^d \rightarrow \mathbb{R}^d$ with Lipschitz constant C and $\mu, \mu' \in \mathcal{P}(\mathbb{T}^d)$

$$(3.46) \quad \begin{aligned} \mathbf{d}_1(\psi(\phi)_{\#}\mu, \psi(\phi')_{\#}\mu') &= \sup_{\|f\|_{Lip} \leq 1} \left\{ \int_{\mathbb{T}^d} f(\psi(\phi(x)))\mu(dx) - \int_{\mathbb{T}^d} f(\psi(\phi'(x)))\mu'(dx) \right\} \\ &\leq \sup_{\|f\|_{Lip} \leq 1} \left\{ \int_{\mathbb{T}^d} f(\psi(\phi(x)))\mu(dx) - \int_{\mathbb{T}^d} f(\psi(\phi(x)))\mu'(dx) \right\} \\ &\quad + \sup_{\|f\|_{Lip} \leq 1} \left\{ \int_{\mathbb{T}^d} f(\psi(\phi(x))) - f(\psi(\phi'(x)))\mu'(dx) \right\} \\ &\leq C\|\psi\|_{Lip}\mathbf{d}_1(\mu, \mu') + \|\psi\|_{Lip}\|\phi - \phi'\|_{\infty}. \end{aligned}$$

Hence we deduce that \tilde{A} is Lipschitz continuous □

From this property, we easily deduce from Corollary 3.8 that Theorem 3.7 can be extended to situations involving the image measure, that is master equation of the form

$$(3.47) \quad \begin{aligned} &\partial_t U(t, x, m) + H(x, \nabla_x U(t, x, m), m, \psi(\nabla_x U)_{\#}m) - \sigma \Delta_x U(t, x, m) \\ &+ \int_{\mathbb{R}^d} B(y, \nabla_x U(t, y, m), m, \psi'(\nabla_x U)_{\#}m) D_m U(t, x, m, y) m(dy) \\ &- \sigma' \int_{\mathbb{R}^d} \operatorname{div}_y (D_m U(t, x, m, y)) dm(y) = 0 \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned}$$

for ψ and ψ' two Lipschitz functions.

4. COMMENTS AND FUTURE DEVELOPMENTS ON LIPSCHITZ SOLUTIONS OF MFG MASTER EQUATIONS

We present in this sections several comments and future directions of research on Lipschitz solutions of MFG master equations that we believe could be of interest.

4.1. Application to numerical computations. Let us precise what we believe to be the principal application of the previous uniqueness result : the justification of numerical computations. Indeed, if by using some abstract or black-box method, like neural networks for instance, one is able to exhibit a solution of a MFG master equation of one of

the types presented above, then if it is a Lipschitz solution, it is necessary the unique one.

With the growing number of works on the use of machine learning techniques to solve MFG master equation, and the lack of proof of convergence results, the results presented above justify the following heuristic : if one Lipschitz solution has been selected, then it have at least some meaning in the sense that it is the only one. Moreover, let us insist that, because several machine learning methods are parametrized (neural networks for instance). Hence establishing some regularity properties of the learned solution can be done a priori.

4.2. Another representation formula. We provide here another approach to represent solutions of (3.1), in which we linearize only the transport term in m . To be more precise, we consider the following equation

$$(4.1) \quad \begin{aligned} & \partial_t V(t, x, m) + H(x, \nabla_x V(t, x, m), m) - \sigma \Delta_x V(t, x, m) \\ & + \int_{\mathbb{R}^d} B(t, y, m) \cdot D_m V(t, x, m, y) m(dy) \\ & - \sigma' \int_{\mathbb{R}^d} \operatorname{div}_y (D_m V(t, x, m, y)) dm(y) = f(x, m) \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned}$$

where $B : [0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ is a vector field. To this equation, we naturally associates the system

$$(4.2) \quad \begin{aligned} dX_s &= \alpha_s ds + \sqrt{2\sigma} dW_s \text{ for } t \geq 0, \\ \partial_s m &= -\operatorname{div}(B(t-s, x, m)m) + \sigma' \Delta_x m \text{ in } (0, t) \times \mathbb{T}^d, \end{aligned}$$

where $(\alpha_s)_{s \geq 0}$ represents the control of a player. In this framework, the natural representation of a solution of (4.1) is

$$(4.3) \quad V(t, x, \mu) = \inf_{\alpha} \mathbb{E} \left[\int_0^t H^*(X_s^\alpha, \alpha_s, m(s)) ds + U_0(X_t^\alpha, m(t)) \right],$$

where $(X_s^\alpha, m(s))_{s \in [0, t]}$ is the solution of (4.2) with initial conditions x and μ and where H^* is the Fenchel conjugate of H with respect to its second argument. If the vector field B is Lipschitz continuous, it is a standard result of stochastic optimal control that V is indeed well defined. As in the previous case we denote by Ψ the operator defined by the previous relation. That is, if V is given by (4.3), then we denote

$$(4.4) \quad V = \Psi(t, B, U_0).$$

We can introduce the following notion of solution.

Definition 4.1. *Given a time T and an initial condition U_0 , a Lipschitz solution of (3.1) is a function $U : [0, T) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ such that*

- $\nabla_x U$ is Lipschitz in x, m , uniformly in $[0, t]$ for any $t < T$.
- For any $t < T$,

$$(4.5) \quad U = \Psi(t, D_p H(x, \nabla_x U, m), U_0).$$

We do not provide a detailed mathematical study of this notion of solution, mainly because it will be redundant with the one we provided in Section 3 and also more technical. Nonetheless, this notion of solution is easily interpretable in terms of game theory. Indeed, in this context, $\Psi(t, B, U_0)$ is simply the operator which computes the value associated to the best response strategy for a generic player, given that it anticipates the vector field B . By anticipation of the vector field B , we mean that it anticipates that the repartition of players is going to be driven by B . Hence, a fixed point of (4.5) is easily interpretable in terms of a fixed point of a best reply operator.

Finally, this operator Ψ is, in general in game theory contexts, of a practical use outside of just defining a notion of equilibria for games. It can help to study procedures such as fictitious play for instance.

4.3. Another notion of monotone solutions for MFG master equations. The notion of Lipschitz solution we just presented is based on the idea that through some Lipschitz regularity, we can define characteristics for the equation and then a value through a representation formula. Another natural regime in which the flow of an evolution equation is well defined is the monotone one, see for instance [8]. We now make a brief development on how we could use this idea to define a solution of the master equation, in the same spirit as what we did for the Lipschitz regularity. We work in the case of (2.4) to fix ideas.

In the setting of (2.13), we recall that the natural characteristics of the master equation have the form

$$(4.6) \quad \frac{dX(s)}{ds} = -F(X(s), U(t-s, X(s))),$$

with initial condition

$$(4.7) \quad X(0) = x_0 \in H.$$

The previous ODE admits a unique solution as soon as $(t, X) \rightarrow F(X, U(t, X))$ is Lipschitz in X , uniformly in t . But is also the case when

$$(4.8) \quad \forall t \geq 0, X, Y \in H, \langle F(X, U(t, X)) - F(Y, U(t, Y)), X - Y \rangle \geq 0.$$

Based on this remark, we can provide the following definition

Definition 4.2. For $T > 0$, given an initial condition U_0 , a bounded function $U : [0, T] \times H \rightarrow H$ is a solution of (2.4) on $[0, T]$ if

- For $t \leq T$, U satisfies (4.8).
- For $t \leq T, x \in H$, U satisfies

$$(4.9) \quad U(t, x) = \int_0^t G(X(s), U(t-s, X(s))) ds + U_0(X(t)),$$

where $(X(s))_{s \in [0, t]}$ is the unique solution of (4.6) with initial condition x .

Obviously, this notion of monotonicity is different from the usual one in MFG theory and one does not clearly imply the other.

4.4. Master equations with control on the volatility. In the setting of Section 3, we can also consider master equations which are associated to a MFG in which the players control the volatility of their trajectory. In such a setting, the master equations takes the form of

$$(4.10) \quad \begin{aligned} \partial_t U(t, x, m) + \int_{\mathbb{T}^d} \text{Tr}[G(y, D_x^2 U(t, y, m), m) D_y^2 \nabla_m U(t, x, m, y)] m(dy) \\ + F(x, D_x^2 U(t, x, m), m) = 0 \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned}$$

where $F : \mathbb{T}^d \times S_d(\mathbb{R}) \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ and $G : \mathbb{T}^d \times S_d(\mathbb{R}) \times \mathcal{P}(\mathbb{T}^d) \rightarrow S_d(\mathbb{R})$ are given functions. Recall that $D_x^2 \phi$ stands for the Hessian matrix of a function $\phi : \mathbb{T}^d \rightarrow \mathbb{R}$.

Previously, we developed a theory of Lipschitz solution of the master equation based on the equation satisfied by $W = \nabla_x U$. In this context, we believe that a similar approach can be developed based on properties of the equation satisfied by $W = D_x^2 U$. However, a direct use of the argument of the previous proofs is not possible and new results are here needed.

ACKNOWLEDGMENTS

The three authors acknowledge a partial support from the Lagrange Mathematics and Computing Research Center and a partial support from the chair FDD (Institut Louis Bachelier).

REFERENCES

- [1] Yves Achdou, Charles Bertucci, Jean-Michel Lasry, Pierre-Louis Lions, Antoine Rostand, and José A Scheinkman. A class of short-term models for the oil industry that accounts for speculative oil storage. *Finance and Stochastics*, pages 1–39, 2022.
- [2] David M Ambrose and Alpár R Mészáros. Well-posedness of mean field games master equations involving non-separable local hamiltonians. *arXiv preprint arXiv:2105.03926*, 2021.
- [3] Charles Bertucci. Monotone solutions for mean field games master equations: finite state space and optimal stopping. *Journal de l'École polytechnique — Mathématiques*, 8:1099–1132, 2021.
- [4] Charles Bertucci. Monotone solutions for mean field games master equations: continuous state space and common noise. *arXiv preprint arXiv:2107.09531*, 2021.
- [5] Charles Bertucci, Jean-Michel Lasry, and Pierre-Louis Lions. Some remarks on mean field games. *Communications in Partial Differential Equations*, 44(3):205–227, 2019.
- [6] Charles Bertucci, Louis Bertucci, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field game approach to bitcoin mining. *arXiv preprint arXiv:2004.08167*, 2020.
- [7] Charles Bertucci, Jean-Michel Lasry, and Pierre-Louis Lions. Master equation for the finite state space planning problem. *Archive for Rational Mechanics and Analysis*, 242(1):327–342, 2021.
- [8] Haim Brezis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. Elsevier, 1973.

- [9] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, pages 1–29, 2017.
- [10] Pierre Cardaliaguet and Panagiotis Souganidis. Monotone solutions of the master equation for mean field games with idiosyncratic noise. *SIAM Journal on Mathematical Analysis*, 54(4):4198–4237, 2022.
- [11] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The Master Equation and the Convergence Problem in Mean Field Games:(AMS-201)*, volume 201. Princeton University Press, 2019.
- [12] Pierre Cardaliaguet, Marco Cirant, and Alessio Porretta. Splitting methods and short time existence for the master equations in mean field games. *Journal of the European Mathematical Society*, 2022.
- [13] René Carmona and François Delarue. Probabilistic theory of mean field games: vol. i, mean field fbsdes, control, and games. *Springer*, 2018.
- [14] René Carmona and François Delarue. Probabilistic theory of mean field games: vol. ii, mean field games with common noise and master equations. *Springer*, 2018.
- [15] Wilfrid Gangbo and Alpár R Mészáros. Global well-posedness of master equations for deterministic displacement convex potential mean field games. *arXiv preprint arXiv:2004.01660*, 2020.
- [16] Wilfrid Gangbo, Alpár R Mészáros, Chenchen Mou, and Jianfeng Zhang. Mean field games master equations with non-separable hamiltonians and displacement monotonicity. *arXiv preprint arXiv:2101.12362*, 2021.
- [17] Ziad Kobeissi. On classical solutions to the mean field game system of controls. *Communications in Partial Differential Equations*, 47(3):453–488, 2022.
- [18] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.
- [19] Pierre-Louis Lions. Cours au college de france. *www.college-de-france.fr*, 2006-2012.
- [20] Pierre-Louis Lions. Cours au college de france. *www.college-de-france.fr*, 2021-2022.
- [21] Pierre-Louis Lions and Panagiotis E Souganidis. Extended mean-field games. *Rendiconti Lincei*, 31(3):611–625, 2020.
- [22] Chenchen Mou and Jianfeng Zhang. Wellposedness of second order master equations for mean field games with nonsmooth data. *arXiv preprint arXiv:1903.09907*, 2020.

APPENDIX A. PROOF OF THE ESTIMATE USED IN THE PROOF OF THEOREM 3.7

Remark that $\mu = m_1 - m_2$ is a solution of

$$(A.1) \quad \partial_t \mu - \sigma \Delta \mu + \operatorname{div}(b\mu) = \operatorname{div}((b_2 - b_1)m_2) \text{ in } (0, t_1) \times \mathbb{T}^d.$$

Consider now $t < t_1$ and a Lipschitz function $\phi_0 : \mathbb{T}^d \rightarrow \mathbb{R}$ and consider the solution ϕ of

$$(A.2) \quad \begin{cases} -\partial_s \phi - \sigma \Delta \phi - b \cdot \nabla_x \phi = 0 \text{ in } (0, t) \times \mathbb{T}^d, \\ \phi|_{s=t} = \phi_0 \text{ in } \mathbb{T}^d. \end{cases}$$

Using the fact that μ is in particular a weak solution of the previous PDE, we deduce that

$$(A.3) \quad \int_{\mathbb{T}^d} \phi_0 d\mu_t = \int_0^t \int_{\mathbb{T}^d} \nabla_x \phi(s, x) \cdot (b_1 - b_2)(s, x) m_2(s)(dx).$$

Then result then follows from standard parabolic estimates on (A.2).