Monotone solutions for mean field games master equations : continuous state space and common noise

Charles Bertucci

MONOTONE SOLUTIONS FOR MEAN FIELD GAMES MASTER EQUATIONS : CONTINUOUS STATE SPACE AND COMMON NOISE

CHARLES BERTUCCI [1]

ABSTRACT. We present the notion of monotone solution of mean field games master equations in the case of a continuous state space. We establish the existence, uniqueness and stability of such solutions under standard assumptions. This notion allows us to work with solutions which are merely continuous in the measure argument, in the case of first order master equations. We study several structures of common noises, in particular ones in which common jumps (or aggregate shocks) can happen randomly, and ones in which the correlation of randomness is carried by an additional parameter.

CONTENTS

[1] : CMAP, Ecole Polytechnique, UMR 7641, 91120 Palaiseau, France .

Introduction

This paper introduces the notion of monotone solution for mean field games (MFG in short) master equations in the case of a continuous state space. Using this notion, we establish results of existence and uniqueness for merely continuous solutions of master equations, which are non-linear first order infinite dimensional partial differential equations (PDE in short) at the core of the MFG theory. Even though this paper is self-contained, it is the follow-up of [2] in which we presented a similar notion in the simpler case of a finite state space.

General introduction. MFG are dynamic games involving a crowd of non-atomic agents. If such games have a tremendous number of applications in several fields, they naturally arise in Economics, and they actually did so in the eighties and nineties. A general mathematical framework to study such games (as well as the terminology MFG) has been introduced by J.-M. Lasry and P.-L. Lions in [19, 20]. We present here some general aspects of this theory, focusing on Nash Equilibria of such games. The study of the Nash equilibria of a MFG reduces to the analysis of a PDE called the master equation [20, 6]. This fact may seem surprising since Nash equilibria are defined as fixed points of certain abstract mappings and thus possess in general few properties. The link between those equilibria and the master equation is a consequence of strong uniqueness properties of this monotone regime, in which, formally, players are adversarial to one another. The master equation is generally an infinite dimensional non linear PDE as soon as the state space of the players is continuous. Another striking property of MFG is that when the randomness (or noise) affecting the players is distributed in an i.i.d. fashion among them, Nash equilibria can be characterized with a system of forward-backward PDE in finite dimension [19, 20, 15, 24]. Let us note that in this situation, several Nash equilibria can coexist. An important aspect of the MFG theory is the so-called probabilistic approach [8, 16], which we shall not particularly use in this paper. Finally, we end this general introduction by mentioning the question, that we do not treat here, of the convergence of N-players games toward MFG, which is a possible way to justify the PDE arising in MFG theory and which has been partially solved at this point [6, 17].

Regularity of the solution of the master equation. In the aforementioned monotone regime, the uniqueness result established by Lasry and Lions [19] makes it meaningful to define a value function associated to a MFG. This value function associates tthe value of the game $U(t, x, m)$ to a player in a state x, when the remaining time in the game is t and the measure m describes the repartition of the other players. In this monotone regime, if it is smooth, the value function can be characterized as the unique solution of the master equation [20, 6]. A natural and fundamental question in the MFG theory is the following : If the value function is not smooth, can it still be characterized as the unique weak solution of the master equation ? The difficulty here lies in the definition of weak solutions one has to choose. In this paper, we answer positively to this question in the monotone regime. Namely, to define our monotone solutions, we only need continuity of the value function in the measure argument for first order master equations and first order regularity with respect to the measure argument for second order master equations. This paper is the extension of [2] in which we treated this problem in the case of a finite state space.

In the last years, the question of the regularity of solutions of master equations has raised quite a lot of interest. In [6], the authors shew that, under strong assumptions, the value function turns out to be smooth and, thus, the unique solution of the master equation. Alternatively, the monotone regime proved to be regularizing in the finite state space case [20, 3, 4]. More recently, several teams have addressed the issue of defining weak solutions of master equations in several context (which are not the monotone regime) : [10, 9] propose ways of selecting a weak solution in finite state space, particularly in the potential case ; [22, 13, 14] introduce notions of weak solutions of the master equation which do not rely on monotonicity assumptions. Up to this point, no general framework has been proposed.

Modeling of common noise. An objective of this paper is to introduce a notion of weak solution for the MFG master equations (in the monotone regime). As already mentioned, in the absence of a common noise, the study of the master equation is not necessary. Hence, it is natural to present our notion of solution in cases involving a common noise. Up to now, most of the mathematical literature on MFG [6, 8, 7, 5] is concerned with the following common noise: the state of all players is affected by a common Brownian motion. This noise has several specificities: it yields second order terms (with respect to the measure variable) in the master equation; it is not particularly regularizing (at least this has not been established up to now and it is quite unlikely since the arising terms are only "degenerate elliptic"); it induces a singular behavior for the underlying measure which describes the repartition of players in the state space. Indeed in this context, this measure is randomly pushed with a force which is a Brownian motion. If this type of noise is certainly helpful to model numerous situations, we here argue that it is undeniably not the most general situation and that several other models can be of interest in many applications, and that those models do not raise mathematical questions as difficult as the one we just mentioned. Let us precise that the second to last Section of this paper is concerned with this often-studied common noise.

A first type of common noise we insist on later on in the paper, is a setting in which the correlation in the randomness affecting the players is carried by an additional parameter. Typical examples for such kinds of models are MFG involving players which interact on a market through stochastic environmental variables, such as price for instance. In such a context, it is natural to expect that the value of the game depends on this additional parameter, and therefore that the associated master equation depends on derivatives (possibly second order ones) of the value function with respect to the price. Clearly, if this additional parameter is finite dimensional, then the master equation stays a first order PDE (with respect to the measure variable) despite modeling a MFG with common noise.

A second type of noise we want to model is one similar to the common noise introduced in [3] in the case of a finite state space. This type of noise consists in assuming that at random times, which are common to all players, a transformation is going to affect all players in the game. Such a type of noise is adequate to model aggregate shocks (to use a terminology from Economics) which may occur at random times. In addition to its intrinsic (mathematical) interest, this type of noise is helpful to approximate other types of common noises, such as the one in which all the players are pushed by the same Brownian motion.

Structure of the paper. In Section 1, we introduce some notation as well as recall some results concerning derivatives in the space of probability measure. In Section 2, we present the main MFG model underlying the master equations we study, as well as some known results concerning MFG master equations. We proceed by introducing our notion of monotone solution in Section 3. We discuss the question of the existence of monotone solutions in Section 4. Sections 5 and 6 deal with respectively monotone solutions of master equations of first and second order, in the presence of a common noise. We conclude this paper and present perspectives and extensions of this work in Section 7.

1. Notation and derivatives in the space of probability measure

In this somehow introductive section, we present some notations, especially concerning derivatives in a set of probability measures, as well as some basic results on those derivatives that we shall need later on.

- We denote by d an integer greater than 1 which refers to the dimension of the players' state space.
- We denote by $\mathbb{T}^d$ the d dimensional torus whose inner scalar product is denoted by $\cdot$, i.e. $x \cdot y$ denotes the scalar product between $x, y \in \mathbb{T}^d$.
- The set of measures on $\mathbb{T}^d$ is denoted by $\mathcal{M}(\mathbb{T}^d)$. For $m \in \mathcal{M}(\mathbb{T}^d)$, we denote its support by $Supp(m)$.
- We denote by $\mathcal{P}(\mathbb{T}^d)$ the set of probability measures on $\mathbb{T}^d$. This set is equipped with the Monge-Kantorovich (or 1-Wasserstein) distance $\mathbf{d}_1$ defined with

$$\forall \mu, \nu \in \mathcal{P}(\mathbb{T}^d), \mathbf{d}_1(\mu, \nu) = \sup_{\phi} \int_{\mathbb{T}^d} \phi(x)(\mu - \nu)(dx), \tag{1.1}$$

where the supremum is taken over all Lipschitz functions $\phi : \mathbb{T}^d \to \mathbb{R}$ with a Lipschitz constant at most one. We recall that $\big(\mathcal{P}(\mathbb{T}^d), \mathbf{d}_1\big)$ is a compact metric set.
- We denote by $\langle \cdot, \cdot \rangle$ the scalar product of $L^2(\mathbb{T}^d)$ and by a slight abuse of notation, all its extensions on functional spaces in duality. That is, if f and μ are in $L^2(\mathbb{T}^d)$, then $\langle f, \mu \rangle$ is their $L^2(\mathbb{T}^d)$ scalar product, but if $f \in \mathcal{C}^0(\mathbb{T}^d)$ and $\mu \in \mathcal{P}(\mathbb{T}^d)$, then $\langle f, \mu \rangle$ is the integral of f against the measure μ ; and if for instance $f \in \mathcal{C}^2(\mathbb{T}^d)$ and $\mu \in \mathcal{P}(\mathbb{T}^d)$, then $\langle f, \Delta\mu \rangle$ is the evaluation of the distribution of second order $\Delta\mu$ on f.
- For a function of two variables $f : (\mathbb{T}^d)^2 \to \mathbb{R}$, we define, whenever it makes sense

$$\langle \mu | f(\cdot, \cdot) | \nu \rangle = \langle f_1, \nu \rangle, \tag{1.2}$$

where $f_1 : y \to \langle f(\cdot, y), \mu \rangle$. The important remark is that μ is tested against the first argument of f whereas ν is tested against its second argument.
- An application $f : \mathcal{P}(\mathbb{T}^d) \to \mathcal{C}^0(\mathbb{T}^d)$ is said to be monotone if

$$\forall \mu, \nu \in \mathcal{P}(\mathbb{T}^d), \langle f(\mu)(\cdot) - f(\nu)(\cdot), \mu - \nu \rangle \geq 0. \tag{1.3}$$

- For $n \in \mathbb{N}, \alpha \in [0, 1)$ and a function $\phi : \mathbb{T}^d \to \mathbb{R}$ we denote by $\|\phi\|_{n+\alpha}$ its $C^{n,\alpha}$ norm.
- For $n \in \mathbb{N}, \alpha \in [0, 1)$ and a function $\phi : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{R}$ we denote by $\|\phi\|_{(n+\alpha, n+\alpha)}$ its $C^{n,\alpha}$ norm.
- We introduce the space $\mathcal{B}$ of functions $U : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that U is globally continuous and

$$\sup_{m \in \mathcal{P}(\mathbb{T}^d)} \|U(\cdot, m)\|_2 < \infty. \tag{1.4}$$

- We also introduce the space $\mathcal{B}_t$ of functions $U : [0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that U is globally continuous and for all $T > 0$:

$$\sup_{m \in \mathcal{P}(\mathbb{T}^d), t \in [0,T]} \|U(t, \cdot, m)\|_2 < \infty. \tag{1.5}$$

- We also introduce the space $\mathcal{B}'_t$ of functions $U : [0,\infty)\times\mathbb{T}^d\times\mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that U is globally continuous and for all $T > 0$:

$$\sup_{m\in\mathcal{P}(\mathbb{T}^d),t\in[0,T]} \|U(t,\cdot,\cdot,m)\|_2 < \infty. \tag{1.6}$$

- The usual convolution product in $\mathbb{T}^d$ is denoted by $\star$.
- The image measure of a measure m by a map T is denoted $T_\# m$.

1.1. **Derivatives in the space of probability measures.** We say that a function $U : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is derivable at m if there exists a continuous map $\phi : \mathbb{T}^d \to \mathbb{R}$ such that for all $\mu \in \mathcal{P}(\mathbb{T}^d)$

$$\lim_{\theta\to 0,\theta>0} \frac{U((1-\theta)m+\theta\mu)-U(m)}{\theta} = \langle \phi, \mu - m\rangle. \tag{1.7}$$

Clearly, there is no uniqueness of such a function ϕ as it is defined up to a constant. We denote $\phi = \frac{\delta U}{\delta m}(m)$ when it is such that $\langle \phi, m\rangle = 0$.

We say that U is $\mathcal{C}^1$ if the map $m \to \frac{\delta U}{\delta m}(m)$ is continuous.

When $\frac{\delta U}{\delta m}(m)$ is a $\mathcal{C}^1$ function of $\mathbb{T}^d$, we denote its gradient by $D_m U(m,x) := \nabla_x \frac{\delta U}{\delta m}(m)(x)$. The function $D_m U$ is the intrinsic derivative of U at m. It satisfies

$$\lim_{h\to 0} \frac{U((Id+h\phi)_\# m)-U(m)}{h} = \langle D_m U(m)\cdot\phi, m\rangle, \tag{1.8}$$

where $\phi : \mathbb{T}^d \to \mathbb{T}^d$ is continuous function and $T_\# m$ denotes the image measure of m by the map T.

If it exists, the second order derivatives of $U : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ at m is a function $\psi : (\mathbb{T}^d)^2 \to \mathbb{R}$ such that for any $x \in \mathbb{T}^d$, $\psi(x,\cdot) = \frac{\delta}{\delta m}\left(\frac{\delta U}{\delta m}(\cdot,x)\right)(\cdot)$. We denote by $\frac{\delta^2 U}{\delta m^2}$ the map ϕ such that for any $x \in \mathbb{T}^d$, $\left\langle \frac{\delta^2 U}{\delta m^2}(m,x,\cdot), m\right\rangle = 0$.

Let us finally introduce the following norm on functions $U : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ for $n \in \mathbb{N}$ and $\alpha \in [0,1)$

$$\mathrm{Lip}_{n+\alpha}(U) = \sup_{\mu,\nu\in\mathcal{P}(\mathbb{T}^d)} \frac{\|U(\mu)-U(\nu)\|_{n+\alpha}}{\mathbf{d}_1(\mu,\nu)}. \tag{1.9}$$

We define in the same way for $U : (\mathbb{T}^d)^2 \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$

$$\mathrm{Lip}_{n+\alpha}(U) = \sup_{\mu,\nu\in\mathcal{P}(\mathbb{T}^d)} \frac{\|U(\mu)-U(\nu)\|_{(n+\alpha,n+\alpha)}}{\mathbf{d}_1(\mu,\nu)}. \tag{1.10}$$

1.2. **First order conditions in $\mathcal{P}(\mathbb{T}^d)$.** Consider a C^1 function $U : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ and $m_0 \in \mathcal{P}(\mathbb{T}^d)$ such that

$$U(m_0) = \inf_{\mathcal{P}(\mathbb{T}^d)} U(m). \tag{1.11}$$

One would like to have that $\frac{\delta U}{\delta m}(m_0) = 0$, however this is not true in general. This is mainly due to the fact that , formally, $\mathcal{P}(\mathbb{T}^d)$ has many boundaries and that the optimality conditions associated to (1.11) only yield an inequality in general. We can establish the following.

Proposition 1.1. *Let U be a $\mathcal{C}^1$ function on $\mathcal{P}(\mathbb{T}^d)$ which attains its minimum at m_0. Then $0 \leq \frac{\delta U}{\delta m}(m_0)$ attains its minimum 0 on Supp(m_0). Moreover, if $\frac{\delta U}{\delta m}(m_0)$ is C^1, then*

$$\langle D_m U(m_0), m_0 \rangle = -\left\langle \frac{\delta U}{\delta m}(m_0), \nabla m_0 \right\rangle = 0, \tag{1.12}$$

if $\frac{\delta U}{\delta m}(m_0)$ is C^2, then

$$\left\langle \frac{\delta U}{\delta m}(m_0), \Delta m_0 \right\rangle \geq 0, \tag{1.13}$$

Proof. Let $x \in \mathbb{T}^d$ and, recalling the definition of $\frac{\delta U}{\delta m}$ (1.7) for $\mu = \delta_x$, we obtain that $\frac{\delta U}{\delta m}(m_0, x) \geq 0$. Because we have the normalization condition $\langle \frac{\delta U}{\delta m}(m_0), m_0 \rangle = 0$, we deduce that U reaches its minimum 0 on *Supp*(m_0). The rest of the claim follows quite easily from the optimality conditions in $\mathbb{T}^d$. □

Remark 1.1. *The relation* (1.12) *could have been directly established using* (1.8) *for a sufficiently large choice of functions ϕ.*

One could also provide general results for second order conditions in the spirit of what we just did. Such results are not presented because they are of no need in the following.

2. Main model and preliminaries

In this section we present the typical master equations we are going to study as well as the underlying MFG model. We also give the main assumptions for the rest of the paper and recall an existing result of uniqueness and a variation of a Lemma form Stegall.

2.1. **Mean Field Games and master equations.** We recall, on a well known example, the links between MFG and master equations. We assume that a crowd of non-atomic agents evolves in $\mathbb{T}^d$ during the time interval $[0, t_f]$. The state $(X_t)_{t\geq 0}$ of a player follows

$$dX_t = \alpha_t dt + \sqrt{2\sigma} dW_t, \tag{2.1}$$

where $(\alpha_t)_{t\geq 0}$ is the control of the player and $(W_t)_{t\geq 0}$ is a standard d dimensional Brownian motion on $\mathbb{T}^d$ which models individual noise. By individual noise, we mean that two players' states are going to evolve according to the previous stochastic differential equation for two independent realizations of $(W_t)_{t\geq 0}$. We assume that the cost of a player whose state and control are given by $(X_t)_{t\geq 0}$ and $(\alpha_t)_{t\geq 0}$ is given by

$$\int_0^{t_f} L(X_t, \alpha_t) + f(X_t, m_t)dt + U_0(X_{t_f}, m_{t_f}), \tag{2.2}$$

where L, f and U_0 are cost functions and $(m_t)_{t\geq 0}$ is the process which describes the evolution of the measure describing the repartition of the players in the state space. Hence L represents the part of the cost the player pays which depends on its control, whereas f is the part which depends on the other players. The function U_0 represents a final cost.

Remark 2.1. *In this paper we work only on the so-called decoupled case, in which the dependence on α and m are separated in the previous equation. All the following is adaptable to the general case in which those dependences can be more intricate, under some appropriated additional assumptions.*

Denoting by $U(t, x, m)$ the value of the game (which is not clearly defined at this moment) for a player in the state x when it remains t time in the game and the distribution of the other players in the state space is currently m, we obtain that U solves (if it smooth enough) the so-called master equation

$$\begin{aligned} &\partial_t U - \sigma\Delta U + H(x, \nabla_x U) - \left\langle \frac{\delta U}{\delta m}(x, m, \cdot), \operatorname{div}(D_p H(\cdot, \nabla U(\cdot, m))m) \right\rangle \\ &\quad - \sigma \left\langle \frac{\delta U}{\delta m}(x, m, \cdot), \Delta m \right\rangle = f(x, m), \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \\ &U(0, x, m) = U_0(x, m) \text{ in } \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned} \tag{2.3}$$

where H is the Hamiltonian of the players given by $H(x, p) := \sup_\alpha\{-\alpha \cdot p - L(x, \alpha)\}$.

In the present case, because the noise is only distributed in an i.i.d. fashion among the players, we can characterize Nash equilibria of the game which lasts a time t_f and starts with an initial distribution of player given by m_0, using the following system of finite dimensional PDE

$$\begin{cases} -\partial_t u - \sigma\Delta u + H(x, \nabla_x u) = f(x, m) \text{ in } (0, t_f) \times \mathbb{T}^d, \\ \partial_t m - \sigma\Delta m - \operatorname{div}(D_p H(x, \nabla_x u)m) = 0 \text{ in } (0, t_f) \times \mathbb{T}^d, \\ u(t_f, x) = U_0(x, m(t_f)), m(0, x) = m_0(x) \text{ in } \mathbb{T}^d. \end{cases} \tag{2.4}$$

In the previous system, a solution (u, m) is associated to a Nash equilibria of the game in the following way. The distribution of players m evolves according to the second equation of (2.4) and under the anticipation that the distribution of players is indeed going to be m, the value of the game for the players is given by u. A particular set of MFG are the one called monotone, i.e. for which the following assumption is satisfied.

Hypothesis 1. *The Hamiltonian H is convex in its second argument. The couplings f and U_0 are monotone, i.e. they verify for all $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$*

$$\langle f(\cdot,\mu) - f(\cdot,\nu), \mu - \nu \rangle \geq 0, \tag{2.5}$$

$$\langle U_0(\cdot,\mu) - U_0(\cdot,\nu), \mu - \nu \rangle \geq 0. \tag{2.6}$$

If the previous assumption is satisfied, and that for either (2.5) or (2.6) the inequality is strict as soon as $\mu \neq \nu$, then there exists at most one solution of (2.4) for any initial condition m_0 and any duration of the game t_f. Hence we deduce from this strong uniqueness result for Nash equilibria of the MFG, that a concept of value of a game can be defined. By this we mean that we can indeed talk about the value $U(t,x,m)$ of the MGF for a player in the state x, when the time remaining in the game is t, and the repartition of players in the state space is described by m. In this context, the value U obviously satisfies for all $t_f \geq 0, x \in \mathbb{T}^d, m_0 \in \mathcal{P}(\mathbb{T}^d), U(t_f,x,m_0) = u(0,x)$, where u is such that (u,m) is the unique solution of (2.4). Clearly if U, defined in this way, is smooth, then it is a solution of (2.3).

One of the main objectives of this paper is to generalize the previous approach to a situation in which the use of a system of characteristics such as (2.4) is not clear, for instance in the presence of common noise (i.e. a noise which is not distributed in an i.i.d. fashion among the players). Mainly, we are going to establish that we can characterize, under Hypothesis 1, a value function U for the MFG as the sole weak solution of the master equation, without needing derivability of U with respect to the measure argument.

Even though we do not detail the model underlying the following stationary counterpart of (2.3), it could have been presented in the same manner.

$$\begin{aligned} rU - \sigma\Delta U + H(x,\nabla_x U) - \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \operatorname{div}(D_p H(\cdot,\nabla U(\cdot,m))m) \right\rangle \\ - \sigma \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \Delta m \right\rangle = f(x,m) \text{ in } \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d). \end{aligned} \tag{2.7}$$

This stationary master equation is also a subject of study for this paper.

Remark 2.2. *In the rest of the paper, the presence of the i.i.d. noise between the players plays a crucial role in our study. In the case in which such a noise is not present, let us mention what seems to be the most natural way to formulate the master equation. It is the so-called Hilbertian approach introduced by P.-L. Lions* *in [20]. In this context, the master equation is posed on an Hilbert space and the problem is closer to the finite dimensional setting introduced in [2].*

2.2. **Preliminary results.** In this section we recall the two main results of existence and uniqueness on master equations which we can find in [6], as well as a variant of a Lemma of Stegall on approximated optimization.

The following Theorem of existence of classical solutions is borrowed from [6]. We do not reproduce its rather long proof, but let us mention that it relies on a precise study of the system (2.4) and its dependence on the initial conditions. In some sense, a contribution of this paper is to provide another existence result for master equations, which relies on weaker assumptions.

Theorem 2.1. *Assume that there exists* $C > 0, \alpha \in (0,1)$ *such that :*

- *The Hamiltonian* H *satisfies*

$$\forall x \in \mathbb{T}^d, p \in \mathbb{R}^d, 0 < D^2_{pp}H(x,p) \leq CId, \tag{2.8}$$

  in the sense of symmetric matrices.

-

$$\sup_{m\in\mathcal{P}(\mathbb{T}^d)} \left(\|f(\cdot,m)\|_{2+\alpha} + \left\| \frac{\delta f(\cdot,m,\cdot)}{\delta m} \right\|_{(2+\alpha,2+\alpha)} \right) + Lip_{2+\alpha}\left(\frac{\delta f}{\delta m}\right) \leq C. \tag{2.9}$$

-

$$\sup_{m\in\mathcal{P}(\mathbb{T}^d)} \left(\|U_0(\cdot,m)\|_{3+\alpha} + \left\| \frac{\delta U_0(\cdot,m,\cdot)}{\delta m} \right\|_{(3+\alpha,3+\alpha)} \right) + Lip_{3+\alpha}\left(\frac{\delta U_0}{\delta m}\right) \leq C. \tag{2.10}$$

Then there exists a classical solution $U, \mathcal{C}^1$ *in all the variables,* $\mathcal{C}^2$ *in the space variable* x*, of the master equation* (2.3)*.*

The next result is concerned with uniqueness of solutions of master equations. We present its proof mainly for two reasons, the first one is that it is rarely given in this form, the second one is that this proof is at the core of the definition of monotone solutions that we give in the next section.

Proposition 2.1. *Under Hypothesis 1 : i) there exists at most one smooth solution* U *of* (2.3)*, moreover if it exists,* $U(t)$ *is monotone for all* $t \geq 0$ *; ii) there exists at most one classical solution of* (2.7) *and if it exists, it is monotone.*

Proof. We only detail the proof of the uniqueness property for the stationary equation, the time dependent being treated using similar arguments. Moreover, we prove more general results in the next section.

Let U and V be two smooth solutions of the master equation (2.7). Let us define the function W on $\mathcal{P}(\mathbb{T}^d)^2$ by

$$W(\mu,\nu) = \langle U(\cdot,\mu) - V(\cdot,\nu), \mu - \nu\rangle := \int_{\mathbb{T}^d} U(x,\mu) - V(x,\nu)(\mu-\nu)(dx). \tag{2.11}$$

Using the equations satisfied by both U and V, we deduce that W satisfies on $\mathcal{P}(\mathbb{T}^d)^2$

(2.12)
$$rW + \langle H(x,\nabla_x U) - H(x,\nabla_x V), \mu-\nu\rangle - \left\langle \frac{\delta W}{\delta \mu}, \mathrm{div}\,(D_pH(\cdot,\nabla U(\cdot,\mu))\mu)\right\rangle$$
$$-\sigma\left\langle \frac{\delta W}{\delta\mu},\Delta\mu\right\rangle - \sigma\left\langle\frac{\delta W}{\delta \nu},\Delta\nu\right\rangle + \langle U-V, \mathrm{div}\,(D_pH(\cdot,\nabla U(\cdot,\mu))\mu - D_pH(\cdot,\nabla V(\cdot,\nu))\nu)\rangle$$
$$-\left\langle\frac{\delta W}{\delta\nu}, \mathrm{div}\,(D_pH(\cdot,\nabla V(\cdot,\nu))\nu)\right\rangle = \langle f(\cdot,\mu) - f(\cdot,\nu),\mu-\nu\rangle.$$

To establish the previous equation, we have used the relations (which are true up to a function $c : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$)

(2.13)
$$U(x,\mu) - V(x,\nu) + \left\langle \frac{\delta U}{\delta m}(\cdot,\mu,x), \mu-\nu\right\rangle = \frac{\delta W}{\delta \mu}(\mu,\nu,x) \text{ for all } x\in\mathbb{T}^d, \mu,\nu\in\mathcal{P}(\mathbb{T}^d),$$

(2.14)
$$V(x,\nu) - U(x,\mu) + \left\langle \frac{\delta V}{\delta m}(\cdot,\nu,x), \nu-\mu\right\rangle = \frac{\delta W}{\delta \nu}(\mu,\nu,x) \text{ for all } x\in\mathbb{T}^d, \mu,\nu\in\mathcal{P}(\mathbb{T}^d).$$

The continuous function W reaches its minimum at some point (μ^*,ν^*) at which the following holds

(2.15)
$$\begin{aligned} rW(\mu^*,\nu^*) &+ \langle H(\cdot,\nabla_x U) - H(x,\nabla_x V), \mu^*-\nu^*\rangle - \langle \nabla_x(U-V)\cdot\nabla_x D_pH(\nabla_x U),\mu^*\rangle \\ &- \langle\nabla_x(V-U)\cdot\nabla_x D_pH(\nabla_x V),\nu^*\rangle \geq \langle f(\cdot,\mu^*) - f(\cdot,\nu^*),\mu^*-\nu^*\rangle, \end{aligned}$$

where we have used the optimality conditions given by Proposition 1.1. Hence we deduce from Hypothesis 1 that $rW(\mu^*,\nu^*) \geq 0$ and thus that W is a non-negative function. From the non-negativity of W, we obtain that $U = V + c(m)$ for a function $c : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$, otherwise W should change sign around the diagonal of $\mathcal{P}(\mathbb{T})^2$. Indeed we have here more or less established that $U - V$ is orthogonal to the set of measure of mass 0. Evaluating (2.7) for U and V immediately implies that $c(m) = 0$ on $\mathcal{P}(\mathbb{T}^d)$. Since $U = V$, we finally obtain that U is monotone.

□

We end these preliminary results with the following variation of Stegall variational principle. Although this extension seems to be new, it is a rather immediate adaptation of existing results the interested could find in the monologue [23] for instance.

Lemma 2.1. *Let $f : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ be a continuous function. Take $m \in \mathbb{N}, m > 6d$. For any $\epsilon > 0$, there exists ϕ in the Sobolev space $H^m(\mathbb{T}^d), \|\phi\|_{H^m} \leq \epsilon$ such that $\mu \to f(\mu) + \langle \phi, \mu \rangle$ has a strict minimum on $\mathcal{P}(\mathbb{T}^d)$.*

Proof. Let us consider the multivalued operator $A : H^m(\mathbb{T}^d) \rightrightarrows \mathcal{P}(\mathbb{T}^d)$ which is defined by $A(\phi) = \text{argmin}\{f(\mu) + \langle \phi, \mu \rangle | \mu \in \mathcal{P}(\mathbb{T}^d)\}$. By construction $-A$ is cyclically monotone in the sense that, for finite sequences $\phi_0, \phi_1, ..., \phi_n = \phi_0$, $\mu_i \in A(\phi_i)$,

$$\sum_{i=1}^{n} \langle \phi_i - \phi_{i-1}, \mu_i \rangle \leq 0. \tag{2.16}$$

Indeed for such a sequence,

$$\begin{aligned} \sum_{i=1}^{n} \langle \phi_i - \phi_{i-1}, \mu_i \rangle &= \sum_{i=1}^{n} \langle \phi_i, \mu_i - \mu_{i+1} \rangle \\ &\leq \sum_{i=1}^{n} f(\mu_{i+1}) - f(\mu_i) \\ &= 0. \end{aligned} \tag{2.17}$$

We can take $\mu_0 \in A(0)$ and construct a function $\psi : H^m(\mathbb{T}^d) \to \mathbb{R}$ by setting

$$\psi(\phi) = \sup\{\langle \phi - \phi_n, \mu_n \rangle + \langle \phi_n - \phi_{n-1}, \mu_{n-1} \rangle + ... + \langle \phi_1, \mu_0 \rangle\}, \tag{2.18}$$

where the supremum is taken over all finite sequences satisfying $-\mu_n \in A(\phi_n)$. The function ψ is proper, convex and continuous over the separable Hilbert space $H^m(\mathbb{T}^d)$. Moreover, defining by $\partial\psi$ the sub-differential of ψ, $-A \subset \partial\psi$ by construction. Hence the result is proved since ψ is Fréchet differentiable on a dense subset of the Hilbert space H^m (since it is convex and continuous). □

Remark 2.3. *The result is stated for $m > 6d$ so that $\mathcal{C}^2(\mathbb{T}^d) \subset H^m(\mathbb{T}^d)$. This point will be of use later on in the paper.*

2.3. **On the choice of writing the master equation in $\mathcal{P}(\mathbb{T}^d)$.** Before passing to the core Section of this paper, we take some time to comment the modeling choice we make to write the master equation on $\mathcal{P}(\mathbb{T}^d)$ instead of on $\{m \in \mathcal{M}(\mathbb{T}^d) | m \geq 0\} =: \mathcal{M}_+(\mathbb{T}^d)$. Because, in the problem we are interested in, the number or mass of players stays constant, it is natural to consider the master equation only on $\mathcal{P}(\mathbb{T}^d)$, even if this situation is not the most general one. For instance, one can think about optimal stopping problem such as in [1, 2]. On the other hand, it is natural to define a value for the MFG whatever the total mass of players is. Of course in the situation of interest here, we can write the master equation on $\mathcal{M}_+(\mathbb{T}^d)$ and only the derivatives in the space of measure in directions which preserve the mass of the measure are needed. This previous remark makes the extension from $\mathcal{P}(\mathbb{T}^d)$ to $\mathcal{M}_+(\mathbb{T}^d)$ relatively easy. Moreover, working on the whole $\mathcal{M}_+(\mathbb{T}^d)$ is easier to treat the question of uniqueness of solutions.

For instance, recalling the proof of Proposition 2.1 and its notations, the non-negativity of the function W on $\mathcal{M}_+(\mathbb{T}^d)$ would have been sufficient to conclude. (This remark has higher implications later on in the paper.)

However, even though it seems more profitable to work on $\mathcal{M}_+(\mathbb{T}^d)$ than on $\mathcal{P}(\mathbb{T}^d)$, we prefer the second option as it allows us to use some existing results of the literature. We apologize for this inconvenience and hope that the interested reader shall be able to extend quite easily the results of this paper to the case of $\mathcal{M}_+(\mathbb{T}^d)$.

3. Monotone solutions

In this section, we extend the notion of monotone solution introduced in [2] to the equations (2.7) and (2.3). We shall not be concerned with the existence of such solutions here, as we delay this question to the next section. We start this section with the case of (2.7) before treating (2.3).

3.1. The stationary case.

Even though we refer to [2] for more details on why the notion of monotone solution is natural for MFG master equations, let us briefly recall the main idea behind this notion.

The proof of Proposition 2.1 clearly suggests that uniqueness of solutions can be obtained by looking at points of minimum of a function W defined by $W = \langle U(\mu) - V(\nu), \mu - \nu\rangle$ for U and V two solutions. We then use the information that one has from the fact that U and V solve a master equation to proceed with the proof.

An important remark is that, at points of minimum of W, if W is smooth, we have a relation to express some terms involving the derivatives of U and V, uniquely through U and V (without derivatives). This is observed by combining (2.13) and Proposition 1.1 or formally by taking $\frac{\delta W}{\delta \mu} = 0$ in (2.13).

Hence, from the point of view of U, we only need information at points of minimum of $\mu \to \langle U(\mu) - V, \mu - \nu\rangle$, for some function $V \in \mathcal{C}^2(\mathbb{T}^d)$ and measure $\nu \in \mathcal{P}(\mathbb{T}^d)$. But at this points of minimum, the terms involving the derivatives of U with respect to μ in (2.7) can be expressed without using derivatives in the space of measures. This leads us to the following notion of solution of (2.7).

Definition 3.1. *A function $U \in \mathcal{B}$ is a monotone solution of* (2.7) *if for any $\mathcal{C}^2$ function $\phi : \mathbb{T}^d \to \mathbb{R}$, for any measure $\nu \in \mathcal{M}(\mathbb{T}^d)$, any point m_0 of strict minimum of $m \to \langle U(\cdot, m) - \phi, m - \nu\rangle$, the following holds*

$$\begin{aligned} r\langle U(\cdot, m_0), m_0 - \nu\rangle - \langle \sigma\Delta U + H(\cdot, \nabla_x U), m_0 - \nu\rangle \geq \langle f(\cdot, m_0), m_0 - \nu\rangle \\ - \langle U - \phi, div(D_p H(\nabla_x U) m_0)\rangle - \sigma\langle \Delta(U - \phi), m_0\rangle. \end{aligned} \tag{3.1}$$

Remark 3.1. *Let us insist on the fact that we take $\nu \in \mathcal{M}(\mathbb{T}^d)$ and not in $\mathcal{P}(\mathbb{T}^d)$. Would we not have done so, for any monotone solution U, any function $c : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$, $U + c$ would also be a monotone solution. Hence, to prevent* *such situations, we enlarge the choice of ν for which the definition holds. We refer to the proof of Theorem 3.1 for more on this fact.*

Remark 3.2. *We only ask for information at points of strict minimum for stability reasons.*

This notion of monotone solution is reminiscent of the notion of viscosity solution introduced by Crandall and Lions in [11], although the equation (2.7) does not have a proper comparison principle.

Let us also remark that Definition 3.1 demands regularity in the space variable x. We shall not comment a lot on this fact except for the fact that, if it may be possible to consider less regular functions in the space variable, it does not seem obvious since the notion of solution we propose is not particularly local in the space variable x.

The two following results justify in some sense the notion of solution we propose. The first one states that classical solutions are also monotone solutions, and the second one that there is uniqueness of a monotone solution in the monotone regime.

Proposition 3.1. *Assume that U is a classical solution of* (2.7)*, then it is also a monotone solution of* (2.7)*.*

We do not detail the short proof of this result. Let us mention that it is a direct application of Proposition 1.1.

As the interested reader may have observed at this point, the uniqueness of a solution of (2.7) may be obtained naturally modulo a function $c : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. Indeed, recalling the proof of Proposition 2.1, proving the non-negativity of $\mu, \nu \to \langle U(\mu) - V(\nu), \mu - \nu \rangle$, for U and V two solutions, yields $U = V + c$. In some sense, this weak uniqueness result is sufficient to obtain (formally) the uniqueness of a Nash equilibria. Indeed the optimal strategy, given at the equilibrium by the solution U of the master equation, only depends on the gradient in the spatial variable $x \in \mathbb{T}^d$ of U. Hence the addition of a function $c : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ does not alter the induced strategies.

In any case, we give an additional mild assumption on the monotonicity of f under which the uniqueness of monotone solutions can be established.

Hypothesis 2. *The coupling f is monotone and satisfies for all $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$*

$$\langle f(\cdot, \mu) - f(\cdot, \nu), \mu - \nu \rangle = 0 \Rightarrow f(\cdot, \mu) = f(\cdot, \nu). \tag{3.2}$$

We can now prove the

Theorem 3.1. *Under Hypothesis 1, two monotone solutions of* (2.7) *in the sense of Definition 3.1 only differ by a function $c : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. If a monotone solution exists it is a monotone application. Moreover, if, in addition, Hypothesis 2 is* *satisfied, then there is uniqueness of a monotone solution of* (2.7) *in the sense of Definition 3.1.*

Proof. Let us assume that there exists two such solutions U and V. Let us define $W : \mathcal{P}(\mathbb{T}^d)^2 \to \mathbb{R}$ with

$$W(\mu,\nu) = \langle U(\cdot,\mu) - V(\cdot,\nu), \mu - \nu\rangle := \int_{\mathbb{T}^d} U(x,\mu) - V(x,\nu)(\mu-\nu)(dx). \tag{3.3}$$

We want to show that W is a non-negative function. Assume that this is not the case and that there exists (μ_1,ν_1) such that $W(\mu_1,\nu_1) < 0$. From this we deduce that there exists $\epsilon > 0$ such that for all $\phi,\psi \in \mathcal{C}^0(\mathbb{T}^d)$, $\|\phi\|_0 + \|\psi\|_0 \le \epsilon$

$$\inf_{(\mu,\nu)\in\mathcal{P}(\mathbb{T}^d)^2} \{W(\mu,\nu) + \langle\phi,\mu\rangle + \langle\psi,\nu\rangle\} < \frac{W(\mu_1,\nu_1)}{2} < 0. \tag{3.4}$$

On the other hand, from Lemma 2.1, we deduce that there exist $\phi,\psi \in \mathcal{C}^2(\mathbb{T}^d)$, $\|\phi\|_2 + \|\psi\|_2 \le \epsilon$ such that $(\mu,\nu) \to W(\mu,\nu) + \langle\phi,\mu\rangle + \langle\psi,\nu\rangle$ has a strict minimum at (μ_0,ν_0) on $\mathcal{P}(\mathbb{T}^d)^2$. Using the definition of monotone solutions for U, we deduce that

(3.5)

$$\begin{aligned} r\langle U(\cdot,\mu_0),\mu_0-\nu_0\rangle - \langle\sigma\Delta U + H(\cdot,\nabla_x U),\mu_0-\nu_0\rangle &\ge \langle f(\cdot,\mu_0),\mu_0-\nu_0\rangle \\ &- \langle U(\mu_0) - V(\nu_0) + \phi, \operatorname{div}(D_pH(\nabla_x U)\mu_0)\rangle - \sigma\langle\Delta(U-V+\phi),\mu_0\rangle, \end{aligned}$$

and the corresponding relation for V :

(3.6)

$$\begin{aligned} r\langle V(\cdot,\nu_0),\nu_0-\mu_0\rangle - \langle\sigma\Delta V + H(\cdot,\nabla_x V),\nu_0-\mu_0\rangle &\ge \langle f(\cdot,\nu_0),\nu_0-\mu_0\rangle \\ &- \langle V(\nu_0) - U(\mu_0) + \psi, \operatorname{div}(D_pH(\nabla_x V)\nu_0)\rangle - \sigma\langle\Delta(V-U+\psi),\nu_0\rangle. \end{aligned}$$

Combining the two previous relations, using the convexity of H and the monotonicity of f we deduce that

(3.7)

$$rW(\mu_0,\nu_0) \ge -\langle\phi,\operatorname{div}(D_pH(\nabla_x U)\mu_0)\rangle - \sigma\langle\Delta\psi,\nu_0\rangle - \langle\psi,\operatorname{div}(D_pH(\nabla_x V)\nu_0)\rangle - \sigma\langle\Delta\phi,\mu_0\rangle,$$

which is a contradiction because ϕ and ψ can be chosen arbitrary small. Hence, we obtain that $W \ge 0$. This established the first part of the Theorem (the equality of the spatial gradients of U and V and the monotonicity of U).

Let us now assume that Hypothesis 2 holds and take two monotone solutions U and V of (2.7). From the first part of the proof, we know that there exists $c : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that $V = U - c$. We want to show that $c = 0$. Consider $\rho \in \mathcal{M}(\mathbb{T}^d)$, $\epsilon > 0$ and define $W : \mathcal{P}(\mathbb{T}^d)^2 \to \mathbb{R}$ by

$$W(\mu,\nu) = \langle U(\mu) - U(\nu) + c(\nu), \mu - \nu + \epsilon\rho\rangle. \tag{3.8}$$

Assume first that there exists $\bar\nu$ such that $c(\bar\nu)\langle 1,\rho\rangle = -\delta_0 < 0$. In this situation, for ϕ and ψ sufficiently small, the minimum of $\tilde W : \mu,\nu \to W(\mu,\nu) + \langle\phi,\mu\rangle + \langle\psi,\nu\rangle$ is less that $-\epsilon\delta_0/2$. From Lemma 2.1, for any $\epsilon > 0$, there exists ϕ_ϵ and ψ_ϵ sufficiently small such that the previous function has a strict minimum at $(\mu_\epsilon, \nu_\epsilon)$ on $\mathcal{P}(\mathbb{T}^d)^2$. Using the fact that both U and V are monotone solutions, we obtain, using the convexity of H, that

(3.9)
$$\begin{aligned} rW(\mu_\epsilon, \nu_\epsilon) \geq &\langle f(\mu_\epsilon) - f(\nu_\epsilon), \mu_\epsilon - \nu_\epsilon + \epsilon\rho\rangle - \langle \phi_\epsilon, \mathrm{div}(D_pH(\nabla_x U)\mu_\epsilon)\rangle - \sigma\langle \Delta\psi_\epsilon, \nu_\epsilon\rangle \\ &- \langle \psi_\epsilon, \mathrm{div}(D_pH(\nabla_x V)\nu_\epsilon)\rangle - \sigma\langle \Delta\phi_\epsilon, \mu_\epsilon\rangle. \end{aligned}$$

Consider an accumulating point (μ_0, ν_0) of $((\mu_\epsilon, \nu_\epsilon))_{\epsilon>0}$. Taking the limit $\epsilon \to 0$ in the previous equation (up to the correct subsequence) yields that $f(\mu_0) = f(\nu_0)$, because of Hypothesis 2. From (3.9), we deduce that

$$-\frac{\epsilon\delta_0}{2} \geq rW(\mu_\epsilon, \nu_\epsilon) \geq \epsilon\langle f(\mu_\epsilon) - f(\nu_\epsilon), \rho\rangle + o(\epsilon), \tag{3.10}$$

where we assumed that ϕ_ϵ and ψ_ϵ were indeed chosen sufficiently small. Dividing the previous relation by ϵ and letting $\epsilon \to 0$, we arrive at a contradiction. Hence $\langle c(\nu), \rho\rangle \geq 0$ for any $\nu \in \mathcal{P}(\mathbb{T}^d)$. Since ρ was chosen arbitrary in $\mathcal{M}(\mathbb{T}^d)$, we deduce that $c = 0$ and thus that $U = V$.

□

We now give a result of stability of monotone solutions.

Proposition 3.2. *Assume that there exist sequences $(H_n)_{n\geq 0}$ and $(f_n)_{n\geq 0}$ in respectively $\mathcal{C}(\mathbb{T}^d \times \mathbb{R}^d, \mathbb{R})$ and $\mathcal{C}(\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \mathbb{R})$ which converge locally uniformly toward respectively H and f. Assume that there is a sequence $(U_n)_{n\geq 0}$ of monotone solutions of* (2.7) *(where U_n is the solution associated with H_n and f_n). Assume that $(U_n)_{\geq 0}$ converges locally uniformly toward some function U (for the topology of $\mathcal{B}$), then U is a monotone solution of* (2.7) *associated with H and f.*

Proof. Let us consider $\phi \in \mathcal{C}^2$, $\nu \in \mathcal{M}(\mathbb{T}^d)$ and μ_* a point of strict maximum of $m \to \langle U(\cdot, m) - \phi, m - \nu\rangle$ on $\mathcal{P}(\mathbb{T}^d)$. From Lemma 2.1, we can consider a sequence of functions $(\phi_n)_{n\geq 0}$ such that $\|\phi_n\|_2 \to 0$ as $n \to \infty$ and $m \to \langle U_n(\cdot, m) - \phi + \phi_n, m - \nu\rangle$ admits a strict minimum at μ_n on $\mathcal{P}(\mathbb{T}^d)$. Because U_n is a monotone solution of (2.7), we obtain that

(3.11)
$$\begin{aligned} r\langle U_n(\cdot, \mu_n), \mu_n - \nu\rangle - \langle \sigma\Delta U_n + H_n(\cdot, \nabla_x U_n), \mu_n - \nu\rangle \geq \langle f_n(\cdot, \mu_n), \mu_n - \nu\rangle \\ - \langle U_n - \phi + \phi_n, \mathrm{div}(D_pH_n(\nabla_x U_n)\mu_n)\rangle - \sigma\langle \Delta(U_n - \phi + \phi_n), \mu_n\rangle. \end{aligned}$$

Since $(\mu_n)_{n\geq 0}$ is a compact sequence, extracting a subsequence if necessary, it converges toward a measure $\tilde{\mu}$. By construction of $(\mu_n)_{n\geq 0}$ and convergence of $(U_n)_{n\geq 0}$ toward U, we deduce that for any $m \in \mathcal{P}(\mathbb{T}^d)$,

$$\langle U(\cdot, \tilde{\mu}) - \phi, \tilde{\mu} - \nu\rangle \leq \langle U(\cdot, m) - \phi, m - \nu\rangle. \tag{3.12}$$

From which we obtain $\tilde{\mu} = \mu_*$ (and the convergence of the whole sequence $(\mu_n)_{n\geq 0}$). Let us now remark that since $(U_n)_{n\geq 0}$ converges toward U in $\mathcal{B}$, then $\|\Delta(U_n -$

$U)\|_0 \to 0$ as $n \to \infty$, thus, we can actually pass to the limit in all the terms in (3.11) to obtain

$$
\begin{aligned}
(3.13)\qquad r\langle U(\cdot,\mu_*),\mu_*-\nu\rangle - \langle \sigma\Delta U + H(\cdot,\nabla_x U),\mu_*-\nu\rangle \geq \langle f(\cdot,\mu_*),\mu_*-\nu\rangle \\
-\langle U_*-\phi, \mathrm{div}(D_pH(\nabla_x U)\mu_*)\rangle - \sigma\langle \Delta(U-\phi),\mu_*\rangle.
\end{aligned}
$$

Hence U is a monotone solution of (2.7) (associated with H and f). □

We conclude this section on stationary monotone solutions of master equations by giving a result concerning smooth monotone solutions.

Proposition 3.3. *Assume that U is a $\mathcal{C}^1$ monotone solution such that $\frac{\delta U}{\delta m}$ is Lipschitz continuous and which satisfies for some $c > 0$*

$$
(3.14)\qquad \langle U(m) - U(m'), m - m'\rangle \geq c(\boldsymbol{d}_1(m,m'))^2.
$$

Then U satisfies (2.7) *for all $m \in \mathcal{P}(\mathbb{T}^d), x \in \overset{\circ}{Supp}(m)$ and more generally*

$$
\begin{aligned}
(3.15)\qquad rU - \sigma\Delta U + H(x,\nabla_x U) - \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), div(D_pH(\cdot,\nabla U(\cdot,m))m)\right\rangle \\
-\sigma\left\langle \frac{\delta U}{\delta m}(x,m,\cdot),\Delta m\right\rangle \geq f(x,m) \text{ in } \mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d).
\end{aligned}
$$

Remark 3.3. *It is a general feature of MFG that the value function may not satisfy a PDE outside the support of the measure describing the repartition of players [24].*

Remark 3.4. *The stronger requirements on the monotonicity of U can probably be weakened while preserving the same result.*

Proof. Let us fix $\bar{m} \in \mathcal{P}(\mathbb{T}^d)$. If we choose ν sufficiently close to $\bar{m}$ and define V by

$$
(3.16)\qquad V(x) = U(x,\bar{m}) + \left\langle \frac{\delta U}{\delta m}(x,\bar{m},\cdot),\nu-\bar{m}\right\rangle,
$$

then $W : \mu \to \langle U(\mu) - V, \mu - \nu\rangle$ has a strict minimum at $\mu = \bar{m}$. Writing the relation satisfied by monotone solutions for U at $\bar{m}$, we deduce that it satisfies the announced properties. □

3.2. **The time dependent case.** Let us now introduce the definition of monotone solution in the time dependent setting. The approach is extremely similar except for the fact that, because we do not want to ask for time regularity outside of continuity for a solution U, we use technique of viscosity solutions to treat the time derivative.

Definition 3.2. *A function $U \in \mathcal{B}_t$ is a monotone solution of* (2.3) *if*

- *for any* $\mathcal{C}^2$ *function* $\phi : \mathbb{T}^d \to \mathbb{R}$, *for any measure* $\nu \in \mathcal{M}(\mathbb{T}^d)$, *for any smooth function* $\vartheta : [0, \infty) \to \mathbb{R}$ *and any point* $(t_0, m_0) \in (0, \infty) \times \mathcal{P}(\mathbb{T}^d)$ *of strict minimum of* $(t, m) \to \langle U(t, \cdot, m) - \phi, m - \nu\rangle - \vartheta(t)$, *the following holds*

$$\begin{aligned}\frac{d\vartheta}{dt}(t_0) - \langle \sigma\Delta U + H(\cdot, \nabla_x U), m_0 - \nu\rangle &\geq \langle f(\cdot, m_0), m_0 - \nu\rangle \\ &- \langle U - \phi, div(D_p H(\nabla_x U)m_0)\rangle - \sigma\langle \Delta(U - \phi), m_0\rangle.\end{aligned} \tag{3.17}$$

- *the initial condition holds*

$$U(0, \cdot, \cdot) = U_0(\cdot, \cdot). \tag{3.18}$$

As we did in the stationary case, we now present results of consistency and uniqueness of such solutions. The consistency result being straightforward, we do not prove it. Moreover, let us recall that we postpone the question of existence to the next section.

Proposition 3.4. *Assume that* U *is a smooth solution of* (2.3)*, then it is also a monotone solution of* (2.3)*.*

Theorem 3.2. *Under Hypothesis 1, two monotone solutions of* (2.3) *in the sense of Definition 3.2 only differ by a function* $c : [0, \infty) \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$*. If a monotone solution* U *exists,* $U(t)$ *is a monotone application for all* $t \geq 0$*. Moreover, if, in addition, Hypothesis 2 is satisfied, then there is uniqueness of the monotone solutions of* (2.3) *in the sense of Definition 3.2.*

Proof. Let us consider U and V two such solutions. We define W by

$$W(t, s, \mu, \nu) = \langle U(t, \cdot, \mu) - V(s, \cdot, \nu), \mu - \nu\rangle := \int_{\mathbb{T}^d} U(t, x, \mu) - V(s, x, \nu)(\mu - \nu)(dx). \tag{3.19}$$

We want to prove that $W(t, t, \mu, \nu) \geq 0$ for all $t \geq 0$, $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$. Assume it is not the case, hence there exists $t_*, \delta, \bar{\epsilon} > 0$, such that for all $\epsilon \in (0, \bar{\epsilon})$, $\alpha > 0$, $\phi, \psi \in C^2$ such that $\|\phi\|_2 + \|\psi\|_2 \leq \epsilon$ and $\gamma_1, \gamma_2 \in (\frac{\bar{\epsilon}}{2}, \bar{\epsilon})$,

$$\inf_{t,s\in[0,t_*],\mu,\nu\in\mathcal{P}(\mathbb{T}^d)} \left\{ W(t, s, \mu, \nu) + \langle \phi, \mu\rangle + \langle \psi, \nu\rangle + \frac{1}{2\alpha}(t - s)^2 + \gamma_1 t + \gamma_2 s \right\} \leq -\delta. \tag{3.20}$$

From Lemma 2.1, we know that there exists (for any value of α) ϕ, ψ, γ_1 and γ_2 such that $(t, s, \mu, \nu) \to W(t, s, \mu, \nu) + \langle \phi, \mu\rangle + \langle \psi, \nu\rangle + \frac{1}{2\alpha}(t - s)^2 + \gamma_1 t + \gamma_2 s$ has a strict minimum on $[0, t_*]^2 \times \mathcal{P}(\mathbb{T}^d)^2$ at (t_0, s_0, μ_0, s_0).

We assume first that $t_0 > 0$ and $s_0 > 0$. Using the fact that U is a monotone solution of (2.3) we obtain that

$$(3.21)\quad \begin{aligned} -\gamma_1 - \frac{t_0 - s_0}{\alpha} - \langle \sigma\Delta U(\mu_0) + H(\cdot, \nabla_x U), \mu_0 - \nu_0\rangle &\geq \langle f(\cdot,\mu_0), \mu_0 - \nu_0\rangle \\ &- \langle U(t_0,\mu_0) - V(s_0,\nu_0) + \phi, \operatorname{div}(D_pH(\nabla_x U)\mu_0)\rangle \\ &- \sigma\langle \Delta(U(t_0,\mu_0) - V(s_0,\nu_0) + \phi), \mu_0\rangle, \end{aligned}$$

and similarly for V :

$$(3.22)\quad \begin{aligned} -\gamma_2 - \frac{s_0 - t_0}{\alpha} - \langle \sigma\Delta V(s_0,\nu_0) + H(\cdot, \nabla_x V), \nu_0 - \mu_0\rangle &\geq \langle f(\cdot,\nu_0), \nu_0 - \mu_0\rangle \\ &- \langle V(s_0,\nu_0) - U(t_0,\mu_0) + \phi, \operatorname{div}(D_pH(\nabla_x V)\nu_0)\rangle \\ &- \sigma\langle \Delta(V(s_0,\nu_0) - U(t_0,\mu_0) + \phi), \nu_0\rangle. \end{aligned}$$

Summing the two previous relations, using the monotonicity of f and the convexity of H, we deduce that
(3.23)

$$-\gamma_1-\gamma_2 \geq -\langle \phi, \operatorname{div}(D_pH(\nabla_x U)\mu_0)\rangle - \sigma\langle\Delta\psi,\nu_0\rangle - \langle \psi, \operatorname{div}(D_pH(\nabla_x V)\nu_0)\rangle - \sigma\langle \Delta\phi, \mu_0\rangle.$$

The previous relation is a contradiction (provided that ϵ had been chosen sufficiently small compared to $\bar\epsilon$).

Let us now turn to the case $t_0 = 0$ (the case $s_0 = 0$ being treated in exactly the same fashion). By construction s_0 satisfies $|s_0 - t_0| \leq C\sqrt{\alpha}$ for some $C > 0$ independent of ϵ. Thus choosing $\alpha > 0$ sufficiently small, we easily manage to contradict (3.20).

Hence we have proven that $W(t,t,\mu,\nu) \geq 0$ for $t \geq 0, \mu,\nu \in \mathcal{P}(\mathbb{T}^d)$. This proves the first part of the claim. Let us now consider $c : [0,\infty) \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that $V = U - c$. We place ourselves in the case in which Hypothesis 2 is satisfied. Assume that there exists t_*, ν_* such that $c(t_*,\nu_*) = -\delta_0 < 0$ and consider a non-negative non-zero measure $\rho \in \mathcal{M}(\mathbb{T}^d)$. Because the initial condition is satisfied for both U and V, we know that $t_* > 0$. Furthermore, from Lemma 2.1, we know that for any $\epsilon > 0$ there exists $t_0, \delta, \bar\epsilon > 0$, such that for all $\epsilon' \in (0,\bar\epsilon), \alpha > 0, \phi, \psi \in C^2$ such that $\|\phi\|_2 + \|\psi\|_2 \leq \epsilon'$ and $\gamma_1,\gamma_2 \in (\frac{\bar\epsilon}{2}, \bar\epsilon)$,
(3.24)

$$\inf_{t,s\in[0,t_0],\mu,\nu\in\mathcal{P}(\mathbb{T}^d)} \left\{ \langle U(t,\mu) - V(s,\nu), \mu - \nu + \epsilon\rho\rangle + \langle\phi,\mu\rangle + \langle\psi,\nu\rangle + \frac{1}{2\alpha}(t-s)^2 + \gamma_1 t + \gamma_2 s \right\} \leq -\epsilon\delta_0/2.$$

and the infimum is attained at a unique point $(t_\epsilon, s_\epsilon, \mu_\epsilon, \nu_\epsilon)$. Moreover, we can choose $\bar\epsilon$ such that $\bar\epsilon/\epsilon \to \kappa > 0$ as $\epsilon \to 0$. Proceeding as we did in the first part of the proof in the case $s_\epsilon, t_\epsilon > 0$ for almost all $\epsilon > 0$, we arrive at the relation

$$(3.25)\qquad -\gamma_1 - \gamma_2 \geq \langle f(\mu_\epsilon) - f(\nu_\epsilon), \mu_\epsilon - \nu_\epsilon + \epsilon\rho\rangle + o(\epsilon).$$

Concluding as in the proof of Theorem 3.1, we arrive at a contradiction in this case. The case where either s_ϵ or t_ϵ is equal to 0 for sufficiently many $\epsilon > 0$ can be treated in a similar way. Thus we have proved that $c \geq 0$ and thus by symmetry that $U = V$. □

We now give a result of stability of monotone solutions.

Proposition 3.5. *Assume that there exist sequences $(H_n)_{n\geq 0}$ and $(f_n)_{n\geq 0}$ in respectively $\mathcal{C}(\mathbb{T}^d \times \mathbb{R}^d, \mathbb{R})$ and $\mathcal{C}(\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \mathbb{R})$ which converge locally uniformly toward respectively H and f. Assume that there is a sequence $(U_n)_{n\geq 0}$ of monotone solutions of* (2.3) *(where U_n is the solution associated with H_n and f_n). Assume that $(U_n)_{\geq 0}$ converges locally uniformly toward some function U (for the topology of $\mathcal{B}_t$), then U is a monotone solution of* (2.3) *associated with H and f.*

Proof. Let us consider $T > 0, \phi \in \mathcal{C}^2, \nu \in \mathcal{M}(\mathbb{T}^d)$ and a smooth function $\vartheta : \mathbb{R} \to \mathbb{R}$. Consider also (t_*, μ_*) a point of strict maximum of $m \to \langle U(\cdot, m) - \phi, m - \nu\rangle - \vartheta(t)$ on $[0, T] \times \mathcal{P}(\mathbb{T}^d)$. From Lemma 2.1, we can consider a sequence of functions $(\phi_n)_{n\geq 0}$ and of real numbers $(\delta_n)_{n\geq 0}$ such that $\|\phi_n\|_2 + \delta_n \to 0$ as $n \to \infty$ and $(t, m) \to \langle U_n(\cdot, m) - \phi + \phi_n, m - \nu\rangle - \vartheta(t) - \delta_n t$ admits a strict minimum at (t_n, μ_n) on $[0, T] \times \mathcal{P}(\mathbb{T}^d)$. Because U_n is a monotone solution of (2.3), we obtain that

$$\begin{aligned}\frac{d\vartheta}{dt}(t_n) + \delta_n - \langle \sigma\Delta U_n + H_n(\cdot, \nabla_x U_n), \mu_n - \nu\rangle &\geq \langle f_n(\cdot, \mu_n), \mu_n - \nu\rangle \\ &- \langle U_n - \phi + \phi_n, \operatorname{div}(D_p H_n(\nabla_x U_n)\mu_n)\rangle - \sigma\langle \Delta(U_n - \phi + \phi_n), \mu_n\rangle.\end{aligned} \tag{3.26}$$

Following the same arguments as in the proof of Proposition 3.2 we obtain first that $(t_n, \mu_n) \to (t_*, \mu_*)$ and then that

$$\begin{aligned}\frac{d\vartheta}{dt}(t_*) - \langle \sigma\Delta U + H(\cdot, \nabla_x U), \mu_* - \nu\rangle &\geq \langle f(\cdot, \mu_*), \mu_* - \nu\rangle \\ &- \langle U_* - \phi, \operatorname{div}(D_p H(\nabla_x U)\mu_*)\rangle - \sigma\langle \Delta(U - \phi), \mu_*\rangle.\end{aligned} \tag{3.27}$$

Hence U is a monotone solution of (2.3). □

4. Existence of monotone solutions

In this section, we establish the existence of a monotone solution of (2.3), in cases for which the assumptions of Theorem 2.1 are not satisfied. We first prove an estimate for classical solutions of (2.3) and then use a stability result to prove our existence result. Let us insist on the fact that we believe this section to be more pedagogical than anything else. We want to give an example of how to establish the existence of monotone solutions ; more than weakening optimally the assumptions of the existing literature on a priori regularity of solutions of MFG master equations. We can prove the following

Proposition 4.1. *Assume that U is a classical solution of* (2.3) *and that there exists $C > 0, \alpha, \beta \in (0,1)$ such that*

- $$\sup_{\mu,\nu\in\mathcal{P}(\mathbb{T}^d)} \frac{\|f(\mu)-f(\nu)\|_\alpha}{\boldsymbol{d}_1(\mu,\nu)^\beta} + \frac{\|U_0(\mu)-U_0(\nu)\|_{1+\alpha}}{\boldsymbol{d}_1(\mu,\nu)^\beta} \leq C. \tag{4.1}$$
- $$\sup_{m\in\mathcal{P}(\mathbb{T}^d)} \|f(\cdot,m)\|_\alpha + \|U_0(\cdot,m)\|_{2+\alpha} \leq C. \tag{4.2}$$
- *H satisfies* (2.8) *with the same C.*

Then there exists $C' > 0$ depending only on C, α and β such that

$$|U(t,x,m) - U(t',x',m')| \leq C'(|t-t'|^{\frac{\gamma}{2}} + |x-x'| + \boldsymbol{d}_1(m,m')^\gamma), \tag{4.3}$$

where $\gamma = (2(\beta^{-1} - \frac{1}{2}))^{-1} \in (0,1)$.

The following proof is largely inspired from a similar result in [6] which establishes the global Lipschitz regularity of U.

Proof. Lets us take $t \geq 0$, $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{T}^d)$. We define for $i \in \{1;2\}$, $(u_i, m_i) \in \mathcal{C}^{1,2,\alpha} \times \mathcal{C}([0,T], \mathcal{P}(\mathbb{T}^d))$ the unique solution ([19]) of

$$\begin{cases} -\partial_t u_i - \sigma\Delta u_i + H(x,\nabla u_i) = f(x,m_i) \text{ in } [0,t]\times\mathbb{T}^d, \\ \partial_t m_i - \sigma\Delta m_i - \operatorname{div}(D_pH(x,\nabla u_i)m_i) = 0 \text{ in } [0,t]\times\mathbb{T}^d, \\ m_i(0) = \mu_i, u_i(t) = U_0(m_i(t)), \text{ in } \mathbb{T}^d. \end{cases} \tag{4.4}$$

Since U is a classical solution of (2.3), it follows that U satisfies

$$\forall x \in \mathbb{T}^d, U(t,x,\mu_1) = u_1(0,x). \tag{4.5}$$

From this and the regularity assumptions on f, H and U_0, we deduce that there exists $C > 0$, such that for all $s \geq 0, \mu \in \mathcal{P}(\mathbb{T}^d)$, $\|U(s,\cdot,\mu)\|_{2+\alpha} \leq C$, from which we deduce the estimate in the space variable in (4.3).

We now come back to the Hölder estimate in the measure argument. Let us compute

(4.6)

$$\begin{aligned} 0 \leq &\int_0^t \int_{\mathbb{T}^d} [-\partial_t(u_1-u_2) - \sigma\Delta(u_1-u_2) + H(x,\nabla u_1) - H(x,\nabla u_2)]\, d(m_1(s)-m_2(s))ds, \\ = &\int_0^t \int_{\mathbb{T}^d} (H(x,\nabla u_1) - H(x,\nabla u_2) + D_pH(x,\nabla u_1)\cdot\nabla(u_1-u_2))dm_1ds \\ &+ \int_0^t \int_{\mathbb{T}^d} (H(x,\nabla u_2) - H(x,\nabla u_1) + D_pH(x,\nabla u_2)\cdot\nabla(u_2-u_1))dm_2ds \\ &- \int_{\mathbb{T}^d} (U_0(m_1(t)) - U_0(m_2(t)))d(m_1(t)-m_2(t)) + \int_{\mathbb{T}^d} (u_1(0)-u_2(0))d(\mu_1-\mu_2). \end{aligned}$$

Here, we have used the monotonicity of f for the inequality. Using the convexity of H and the monotonicity of U_0, we obtain

(4.7)
$$\int_0^t \int_{\mathbb{T}^d} |\nabla(u_1 - u_2)|^2 d(m_1(s) + m_2(s))ds \leq C \int_{\mathbb{T}^d} (u_1(0) - u_2(0)) d(\mu_1 - \mu_2) \leq C d_1(\mu_1, \mu_2).$$

Let us remark that since $\nabla_x u_i$ is indeed uniformly bounded, we can use a strict-like convexity of H to obtain the previous inequality. The estimate (4.7) is extremely helpful to establish the next estimate on the trajectories m_1 and m_2 that we now provide using a coupling argument. Let X_1 and X_2 be two random variables of law μ_1 and μ_2 such that $\mathbb{E}[|X_1 - X_2|] = d_1(\mu_1, \mu_2)$. Let us define $(X_{i,s})_{s \geq 0}$ for $i \in \{1; 2\}$ the strong solutions of

$$\begin{cases} dX_{i,s} = -D_p H(X_{i,s}, \nabla u_i(X_{i,s}))ds + \sqrt{2\sigma} dB_s, \\ X_{i,0} = X_i, \end{cases} \tag{4.8}$$

for $(B_s)_{s \geq 0}$ a standard Brownian motion. We now compute using Itô's Lemma

(4.9)
$$\begin{aligned} \mathbb{E}[|X_{1,s} - X_{2,s}|] \leq \mathbb{E}[|X_1 - X_2|] + \mathbb{E}\left[\int_0^s |D_p H(X_{1,s}, \nabla u_1(X_{1,s})) - D_p H(X_{2,s}, \nabla u_1(X_{2,s}))| ds\right] \\ + \mathbb{E}\left[\int_0^s |D_p H(X_{2,s}, \nabla u_1(X_{2,s})) - D_p H(X_{2,s}, \nabla u_2(X_{2,s}))| ds\right]. \end{aligned}$$

We now deduce, using the Lipschitz continuity of $D_p H$ and the Lipschitz continuity of ∇u_1 for the second term and (4.7) for the third term, that

(4.10)
$$\mathbb{E}[|X_{1,s} - X_{2,s}|] \leq \mathbb{E}[|X_1 - X_2|] + C \int_0^s \mathbb{E}[|X_{1,s} - X_{2,s}|] ds + C \left(\int_{\mathbb{T}^d} (u_1(0) - u_2(0)) d(\mu_1 - \mu_2)\right)^{\frac{1}{2}}.$$

From which we deduce using Gronwall's Lemma that

(4.11)
$$\sup_{s \in [t,T]} d_1(m_1(s), m_2(s)) \leq C \left(d_1(\mu_1, \mu_2) + \left(\int_{\mathbb{T}^d} (u_1(0) - u_2(0)) d(\mu_1 - \mu_2)\right)^{\frac{1}{2}}\right).$$

Let us now remark that, using Lemma 3.2.2 in [6], we deduce that

(4.12)
$$\sup_{s \in [0,t]} \|(u_1 - u_2)(s)\|_{1+\alpha} \leq C \left(\sup_{s \in [0,t]} \|f(m_1(s)) - f(m_2(s))\|_\alpha + \|U_0(m_1(t)) - U_0(m_2(t))\|_{1+\alpha}\right).$$

Hence we deduce, using (4.11) and the assumptions on f and U_0 that, that there exists $C > 0$ such that

$$\sup_{s \in [0,t]} \|(u_1 - u_2)(s)\|_{1+\alpha} \leq C \left(\mathbf{d}_1(\mu_1, \mu_2) + \|(u_1 - u_2)(0)\|_{1+\alpha}^{\frac{1}{2}} \mathbf{d}_1(\mu_1, \mu_2)^{\frac{1}{2}}\right)^\beta. \tag{4.13}$$

We now easily obtain that

$$|U(t,x,\mu_1) - U(t,x,\mu_2)| = |u_1(0,x) - u_2(0,x)| \leq C d_1(\mu_1,\mu_2)^\gamma, \tag{4.14}$$

for $\gamma = (2(\beta^{-1} - \frac{1}{2}))^{-1} \in (0,1)$. Let us now recall that in view of the Lipschitz continuity of D_pH, we have the classical estimate for the solution of the Fokker-Planck equation :

$$\forall s,s' \in [t,T], d_1(m_1(s), m_1(s')) \leq C\sqrt{|s-s'|}, \tag{4.15}$$

where C is a constant independent of μ_1 and $t \geq 0$. Moreover, the following relation holds

$$\forall s \in [t,T], U(s,x,m_1(s)) = u_1(s,x). \tag{4.16}$$

Recalling (4.11), we finally obtain that there exists $C > 0$ such that for any $t,s \in [0,T], x \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d)$

$$|U(t,x,m) - U(s,x,m)| \leq C|t-s|^{\frac{\gamma}{2}}, \tag{4.17}$$

which concludes the proof. □

Remark 4.1. *The extension of this result to value function being defined on* $\mathcal{M}_+(\mathbb{T}^d) := \{m \in \mathcal{M}(\mathbb{T}^d) | m \geq 0\}$ *is straightforward when equipping the previous convex set with the metric* $\tilde{\boldsymbol{d}}_1(\mu,\nu) := \sup\langle\phi, \mu - \nu\rangle$ *where the supremum is taken over Lipschitz functions on* $\mathbb{T}^d$ *whose Lipschitz constant is at most* 1 *and which verify* $\phi(0) = 0$.

Remark 4.2. *Let us remark, following exactly the proof of Proposition 3.2 in [6], that, assuming*

$$Lip_\alpha(f) + Lip_{1+\alpha}(U_0) \leq C, \tag{4.18}$$

instead of the Hölder continuity estimates in the previous proposition, one arrives at the conclusion that for some $C' > 0$ *and any* $(t,x,m),(t',x',m') \in [0,\infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)$

$$|U(t,x,m) - U(t',x',m')| \leq C'(|t-t'|^{\frac{1}{2}} + |x-x'| + \boldsymbol{d}_1(m,m')). \tag{4.19}$$

Having established this a priori estimate, we are now in position to prove the existence of monotone solutions which are not necessary classical solutions.

Theorem 4.1. *Assume that Hypothesis 1 holds and that*

- *The Hamiltonian* H *satisfies* (2.8).
- f *and* U_0 *satisfy the assumption of Proposition 4.1 for some* $\beta > 0$ *and* f *and* U_0 *are in the closure (with respect to the uniform convergence) of the set of couplings* f *and* U_0 *satisfying the assumptions of Theorem 2.1.*

Then, there exists a (unique) monotone solution of the master equation (2.3) *in the sense of Definition 3.2.*

Remark 4.3. *As we do not want to enter into the problem of regularization of functions on $\mathcal{P}(\mathbb{T}^d)$, we add the assumption that f and U_0 have to be in the closure of more regular functions. This assumption seems necessary. Furthermore, such a set of couplings is not trivial. Indeed, consider for instance a coupling f defined on $\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)$ by*

$$f(x,m) := \int_{\mathbb{T}^d} \phi(z,(m \star \rho)(z))\rho(x-z)dz, \tag{4.20}$$

where $\phi : \mathbb{R}^2 \to \mathbb{R}$ is a continuous function, smooth in its first argument and Hölder continuous in its second one, with ρ a smooth non-negative even function. By regularizing ϕ, one obtain a regularization of f. Hence such a coupling f is the closure of couplings satisfying the assumptions of Theorem 2.1. Moreover it satisfies the required assumptions.

Proof. Let us consider sequences $(f_n)_{n\geq 0}$ and $(U_{0,n})_{n\geq 0}$ which approximate f and U_0. For any $n \geq 0$, thanks to Theorem 2.1, there exists a (unique) solution U_n of (2.3) associated to f_n and $U_{0,n}$. Using Proposition 4.1, we obtain that the sequence $(U_n)_{n\geq 0}$ is a uniformly continuous sequence of functions. Since it is bounded uniformly for $t = 0$, we deduce from Ascoli-Arzela Theorem that $(U_n)_{n\geq 0}$ is a compact sequence for the uniform convergence. Hence, extracting a subsequence if necessary, it converges toward a limit U_*. Since for all $n \geq 0$, U_n is a classical solution of (2.3) (associated to f_n and $U_{0,n}$) we deduce that it is also a monotone solution of the same equation. We finally conclude using Proposition 3.5 that U_* is a monotone solution of (2.3). □

5. First order master equations with common noise

The interest of this section is twosome. First we want to present some structures of common noise in MFG, which, despite being entirely new, have attracted little attention in the literature even though they cover a wide range of applications. Secondly, we explain, without entering into the same amount of details as we did for (2.7) or (2.3), how the notion of monotone solution is helpful for such cases. We mainly look at two situations. The first one in which an additional parameter is added which is stochastic and which affects all the players in the same way (like a price on a market for instance). The second one is a situation in which players jump in a coordinated manner at random times, as the one introduced in [3]. For pedagogical reasons, we introduce first a case in which the additional parameter has only two states.

As most of the following analysis does not rely on the fact that we are in a time dependent or a stationary situation, we choose here to focus on the time dependent setting.

Finally, let us state that this section is not particularly concerned with the existence of monotone or classical solutions of the master equations we shall introduce. We present in subsection 5.4 an a priori estimate which is valid in all the cases we shall introduce and only detail a proof of existence of monotone solutions in the case of common jumps.

5.1. **Additional two-state stochastic parameter.** Consider a MFG which is similar to the one presented in section 2 except for the fact that there is an additional parameter p, totally exogenous from the rest of the game, which can take two values p_1 and p_2. We assume that this parameter affects the player in the following way : when $p = p_i$ the running cost of the player is a function $f_i : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. We assume that p is a random process which jump from p_1 to p_2 with a transition rate $\lambda_1 > 0$, and from p_2 to p_1 with a transition rate λ_2. In such a situation, the associated master equation is in fact the system of two master equations

(5.1)

$$\begin{aligned} &\partial_t U_i - \sigma\Delta U_i + H(x, \nabla_x U_i) - \left\langle \frac{\delta U_i}{\delta m}(x, m, \cdot), \operatorname{div}(D_p H(\cdot, \nabla U_i(\cdot, m))m) \right\rangle \\ &\quad - \sigma \left\langle \frac{\delta U_i}{\delta m}(x, m, \cdot), \Delta m \right\rangle + \lambda_i(U_i - U_j) = f_i(x, m) \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \\ &U_i(0,x, m) = U_0(x, m)\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \end{aligned}$$

where $i, j\{1; 2\}, i \neq j$.

Remark 5.1. *Not only f, but H, σ and U_0 could have depended on i, without changing any of the following. We only restricted ourselves to the case of f to lighten the notation.*

Following section 3, we naturally propose the following definition of monotone solution for (5.1).

Definition 5.1. *A pair of functions $U_1, U_2 \in \mathcal{B}_t$ is a monotone solution of* (5.1) *if*

- *for any $\mathcal{C}^2$ functions $\phi_1, \phi_2 : \mathbb{T}^d \to \mathbb{R}$, for any measure $\nu \in \mathcal{M}(\mathbb{T}^d)$, for any smooth function $\vartheta : [0, \infty) \to \mathbb{R}$ and any point $(i_0, t_0, m_0) \in \{1; 2\}\times(0, \infty)\times \mathcal{P}(\mathbb{T}^d)$ of strict minimum of $(i, t, m) \to \langle U_i(t, \cdot, m) - \phi_i, m - \nu\rangle - \vartheta(t)$, the following holds*

$$\begin{aligned} &\frac{d\vartheta}{dt}(t_0) - \langle \sigma\Delta U_{i_0} + H(\cdot, \nabla_x U_{i_0}), m_0 - \nu\rangle + \lambda_{i_0}\langle U_{i_0} - U_{j_0}, m_0 - \nu\rangle \\ &\quad \geq \langle f_{i_0}(\cdot, m_0), m_0 - \nu\rangle - \langle U_{i_0} - \phi_{i_0}, div(D_p H(\nabla_x U_{i_0})m_0)\rangle \\ &\qquad - \sigma\langle \Delta(U_{i_0} - \phi_{i_0}), m_0\rangle, \end{aligned} \tag{5.2}$$

  where $j_0 \neq i_0$.
- *the initial condition holds*

$$U_1(0, \cdot, \cdot) = U_2(0, \cdot, \cdot) = U_0(\cdot, \cdot). \tag{5.3}$$

We now present a result of uniqueness of monotone solutions of (5.1), which does not rely on any particular assumption on the evolution of the process p. This is not surprising since the uniqueness of solutions arises from monotonicity properties which models a sort of competition between the players, and this additional parameter does not perturb the competition between the players.

Theorem 5.1. *Under Hypothesis 1, two pairs of functions U_1, U_2 monotone solution of* (5.1) *in the sense of Definition 5.1 only differ by a constant pair of functions $c_1, c_2 : [0, \infty) \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. Under Hypothesis 2, there is uniqueness of such a pair of functions.*

Remark 5.2. *We here understand Hypotheses 1 and 2 in the sense that $f(\cdot, p_1)$ and $f(\cdot, p_2)$ both satisfy the monotonicity assumptions.*

Proof. The proof of this result is very similar to the one of Theorem 3.2. Let us take two solutions (U_1, U_2) and (V_1, V_2) and define $W : \{1; 2\} \times [0, T]^2 \times \mathcal{P}(\mathbb{T}^d)^2$ by

$$W_i(t, s, \mu, \nu) = \langle U_i(t, \mu) - V_i(s, \nu), \mu - \nu \rangle. \tag{5.4}$$

Let us now remark, following the same argument as in the proof of Theorem 3.2, that at a point $(i_0, t_0, s_0, \mu_0, \nu_0)$ of strict minimum of W (up to the addition small perturbations using Lemma 2.1), the term arising in the relation of monotone solutions from the additional term in λ_{i_0} is of the form

$$\lambda_{i_0}(W_{i_0}(t_0, s_0, \mu_0, \nu_0) - W_{j_0}(t_0, s_0, \mu_0, \nu_0)) \tag{5.5}$$

for $j_0 \neq i_0$. Because we are at a point of minimum of W, this term has a sign and the rest of the proof follows quite easily. □

Remark 5.3. *Even though they are true, we do not write once again precise results of stability or consistency for this particular master equation because they are merely trivial adaptations of the ones we already gave.*

5.2. **Additional stochastic parameter following a stochastic differential equation.** We now place ourselves in the same framework as in the previous section, except for the fact that now the parameter p is supposed to evolve according to

$$dp_t = b(p_t)dt + \sqrt{2\sigma'}dB_t \tag{5.6}$$

where $(B_t)_{t\geq 0}$ is a standard k dimensional Brownian motion and $b : \mathbb{R}^k \to \mathbb{R}^k$ a smooth function. For simplicity, we assume that p is valued in $\mathbb{T}^k$ (using the classical quotient $\mathbb{T}^k = \mathbb{R}^k/\mathbb{Z}^k$). This simplification does not play a role if for the fact that it simplifies the notation and the formulation of some results. Following the previous subsection on the two states case, we want to define the value function of the MFG as a function of p (in addition to the other variables). Assuming now that f is also a function of p, we naturally arrive at the master equation

(5.7)

$$\partial_t U - \sigma\Delta_x U + H(x,\nabla_x U) - \sigma'\Delta_p U - b\cdot\nabla_p U - \sigma\left\langle \frac{\delta U}{\delta m}(x,m,p,\cdot), \Delta m\right\rangle$$
$$- \left\langle \frac{\delta U}{\delta m}(x,m,p,\cdot), \mathrm{div}\,(D_pH(\cdot,\nabla_x U(\cdot,m,p))m)\right\rangle = f(x,m,p) \text{ in } (0,\infty)\times\mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d),$$
$$U(0,x,m,p) = U_0(x,m,p) \text{ in } \mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d).$$

In the following, we are not going to enter into much details about the regularity of U with respect to p. Let us only remark that if f and U_0 are smooth functions of p, satisfying, uniformly in p, the assumptions of Theorem 2.1, then we expect that there exists a classical solution of the master equation (5.7). The following definition should by now seems natural to the reader.

Definition 5.2. *A function U in $\mathcal{B}'_t$ is a monotone solution of* (5.7) *if*

- *for any $\mathcal{C}^2$ function $\phi:\mathbb{T}^d\times\mathbb{T}^k\to\mathbb{R}^d$, for any measure $\nu\in\mathcal{M}(\mathbb{T}^d)$, for any smooth function $\vartheta:[0,\infty)\to\mathbb{R}$ and any point $(t_0,m_0,p_0)\in(0,\infty)\times\mathcal{P}(\mathbb{T}^d)\times\mathbb{T}^k$ of strict minimum of*

$$(t,m,p)\to\langle U(t,\cdot,m,p)-\phi(\cdot,p), m-\nu\rangle - \vartheta(t), \tag{5.8}$$

  the following holds

(5.9)

$$\frac{d\vartheta}{dt}(t_0) + \langle -\sigma\Delta_x U - b\cdot\nabla_p U - \sigma'\Delta_p U + H(\cdot,\nabla_x U), m_0-\nu\rangle$$
$$\geq \langle f(\cdot,m_0,p_0), m_0-\nu\rangle - \langle U-\phi, div(D_pH(\nabla_x U)m_0)\rangle - \sigma\langle\Delta_x(U-\phi), m_0\rangle.$$

- *The initial condition holds*

$$U(0,\cdot,\cdot,\cdot) = U_0(\cdot,\cdot,\cdot). \tag{5.10}$$

As it was the case in the two states model, a result of uniqueness can be established without much assumptions on the evolution of the stochastic process $(p_t)_{t\geq 0}$ or on the dependence of f on it.

Theorem 5.2. *Under Hypothesis 1, two monotone solutions of* (5.7) *in the sense of Definition 5.2 differ only by a function $c:[0,\infty)\times\mathcal{P}(\mathbb{T}^d)\times\mathbb{T}^k\to\mathbb{R}$. If such a solution U exists, for any $t\geq 0, p\in\mathbb{T}^k$, $U(t,\cdot,\cdot,p)$ is monotone. Furthermore, if, in addition, Hypothesis 2 is satisfied, then there exists at most one monotone solution of* (5.7).

Remark 5.4. *We here understand Hypotheses 1 and 2 in the sense that $f(\cdot,p)$ satisfies the monotonicity assumptions for any value of p.*

Proof. The proof of this result is once again very similar to the one of Theorem 3.2. Let us take two solutions U and V and define $W:[0,T]^2\times\mathcal{P}(\mathbb{T}^d)^2\times\mathbb{T}^k$ by

$$W(t,s,\mu,\nu,p) = \langle U(t,\cdot,\mu,p) - V(s,\cdot,\nu,p), \mu-\nu\rangle. \tag{5.11}$$

Let us now remark, following the same argument as in the proof of Theorem 3.2, that at a point $(t_0, s_0, \mu_0, \nu_0, p_0)$ of strict minimum of W (up to the addition small perturbations using Lemma 2.1), the term arising in the relation of monotone solutions from the additional terms in p is of the form

$$-b(p_0) \cdot \nabla_p W(t_0, s_0, \mu_0, \nu_0, p_0) - \sigma' \Delta_p W(t_0, s_0, \mu_0, \nu_0, p_0). \tag{5.12}$$

Because we are at a point of minimum of W (in particular it is also a minimum in p), this term has a sign and the rest of the proof follows quite easily. □

Remark 5.5. *Because few information is needed for the regularity in p of W in the previous proof, it is very likely that continuity with respect to p and viscosity solution like information are sufficient to characterize monotone solution of* (5.7), *although we do not claim that such results are trivially in the scope of this paper.*

5.3. **Common jumps.** We now introduce, in the continuous state space framework, a type of common noise similar to the one introduced in [3]. More precisely, we want to model situations in which, at random times which are given by a Poisson process of intensity $\lambda > 0$, all the players in the game are affected by a common transformation. This transformation can be deterministic, for instance all the players in the state x are transported in a state $\Lambda(x)$. It can also carry a form of randomness which is distributed in an i.i.d. fashion among the players. In such a situation, all the players in a state x are going to be transported to a new state which is drawn according to a distribution on the state space $K(\cdot, x)$, independently from one another. So that if before the jumps, the players are distributed according to $m \in \mathcal{P}(\mathbb{T}^d)$, they are distributed according to $\int_{\mathbb{T}^d} K(x, y)m(dy)$ immediately after the jump. We refer to [3] for more details on this type of noise (in the finite state space case).

In the following we assume that K is a non-negative smooth function on $(\mathbb{T}^d)^2$ such that for all $y \in \mathbb{T}^d$, $\int_{\mathbb{T}^d} K(x, y)dx = 1$. We define the operator $\mathcal{T}$ by

$$\forall m \in \mathcal{P}(\mathbb{T}^d), x \in \mathbb{T}^d, \mathcal{T}(m)(x) = \int_{\mathbb{T}^d} K(x, y)m(dy). \tag{5.13}$$

Let us recall that the adjoint $\mathcal{T}^*$ of $\mathcal{T}$ is given by

$$\forall \phi \in \mathcal{C}^0(\mathbb{T}^d), y \in \mathbb{T}^d, \mathcal{T}^*(\phi)(y) = \int_{\mathbb{T}^d} K(x, y)\phi(x)dx. \tag{5.14}$$

Because the players anticipate the noise and the fact that they are going to be transported to another state, the associated master equation is given by

(5.15)

$$\partial_t U - \sigma\Delta U + H(x, \nabla_x U) - \left\langle \frac{\delta U}{\delta m}(x, m, \cdot), \mathrm{div}\,(D_p H(\cdot, \nabla U(\cdot, m))m) \right\rangle$$
$$+ \lambda\Big(U - \mathcal{T}^*(U(t, \cdot, \mathcal{T}(m), p))\Big) - \sigma \left\langle \frac{\delta U}{\delta m}(x, m, \cdot), \Delta m \right\rangle = f(x, m) \text{ in } (0, \infty) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d),$$
$$U(0, x, m) = U_0(x, m) \text{ in } \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d).$$

This master equation is obviously reminiscent of the one studied in [2] (in the finite state space case). Once again we state an appropriate notion of solution for this equation.

Definition 5.3. *A function $U \in \mathcal{B}_t$ is a monotone solution of* (5.15) *if*

- *For any $\mathcal{C}^2$ function $\phi : \mathbb{T}^d \to \mathbb{R}$, for any measure $\nu \in \mathcal{M}(\mathbb{T}^d)$, for any smooth function $\vartheta : [0, \infty) \to \mathbb{R}$ and any point $(t_0, m_0) \in (0, \infty) \times \mathcal{P}(\mathbb{T}^d)$ of strict minimum of $(t, m) \to \langle U(t, \cdot, m) - \phi, m - \nu\rangle - \vartheta(t)$, the following holds*

(5.16)

$$\frac{d\vartheta}{dt}(t_0) + \left\langle -\sigma\Delta U + H(\cdot, \nabla_x U) + \lambda\Big(U - \mathcal{T}^*(U(t, \cdot, \mathcal{T}(m)))\Big), m_0 - \nu \right\rangle$$
$$\geq \langle f(\cdot, m_0), m_0 - \nu\rangle - \langle U - \phi, div(D_p H(\nabla_x U)m_0)\rangle - \sigma\langle \Delta(U - \phi), m_0\rangle.$$

- *The initial condition holds*

$$U(0, \cdot, \cdot) = U_0(\cdot, \cdot). \tag{5.17}$$

As we did in the previous cases, we can establish the following uniqueness result.

Theorem 5.3. *Under Hypothesis 1, two monotone solutions of* (5.15) *in the sense of Definition 5.3 only differ by a function $c : [0, \infty) \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. If such a monotone solution U exists, then $U(t)$ is actually monotone for all time $t \geq 0$. If, in addition, Hypothesis 2 is satisfied, then there is uniqueness of monotone solutions of* (5.15)*.*

Proof. The proof of this result is once again very similar to the one of Theorem 3.2. For the first part of the claim, let us take two solutions U and V and define $W : [0, T]^2 \times \mathcal{P}(\mathbb{T}^d)^2 : \mathbb{R}$ by

$$W(t, s, \mu, \nu) = \langle U(t, \cdot, \mu) - V(s, \cdot, \nu), \mu - \nu\rangle. \tag{5.18}$$

Let us now remark, following the same argument as in the proof of Theorem 3.2, that at a point (t_0, s_0, μ_0, ν_0) of strict minimum of W (up to the addition small perturbations using Lemma 2.1), the relation of monotone solutions one obtains is similar to the classical one except for the addition of a term in λ. When combining the relation from U and the one from V, one obtain the same relation except for the addition of the term

$$\lambda\Big(W(t_0, s_0, \mu_0, \nu_0) - W(t_0, s_0, \mathcal{T}(\mu_0), \mathcal{T}(\nu_0))\Big). \tag{5.19}$$

Because we are at a point of minimum of W, this term has a sign and the rest of the first part of the proof follows quite easily. Concerning the second part of the proof, i.e. when we assume that Hypothesis 2 holds, the proof is once again very similar to the proof of Theorem 3.2, although one has to be careful with the choice of the measure ρ (we reuse the same notations as in the aforementioned proof). Indeed, because an extra term may arise from the jump operator $\mathcal{T}$, one has to choose $\rho \in \mathcal{P}(\mathbb{T}^d)$ such that $\mathcal{T}\rho = \rho$. For such a measure (which always exists under the standing assumptions on $\mathcal{T}$), the argument can be made exactly as in the case without common noise. □

Remark 5.6. *In fact the previous result is still valid if $\mathcal{T}$ is any continuous linear mapping on $\mathcal{P}(\mathbb{T}^d)$ which posses a fixed point.*

5.4. **An a priori estimate for the solution of the master equation with common noise.** We now show an a priori estimate which is essentially valid for all the master equations we have written up to now. We only state (and prove) it in the case of (5.15) and we leave its generalization to the other master equations to the interested reader. Although it is not sufficient to establish general result of existence, we believe that it may be a good starting point for such results. This a priori estimate is valid only under additional monotonicity assumptions on f and U_0. Namely we assume that f and U_0 are differentiable with respect to the measure argument and that there exists $\alpha > 0$ such that for all $\mu \in \mathcal{P}(\mathbb{T}^d), \nu \in L^2(\mathbb{T}^d)$

$$\left\langle \nu \middle| \frac{\delta f}{\delta m}(\cdot, \mu, \cdot) \middle| \nu \right\rangle \geq \alpha \left\| \left\langle \frac{\delta f}{\delta m}(\cdot, \mu), \nu \right\rangle \right\|_{L^2}^2, \tag{5.20}$$

$$\left\langle \nu \middle| \frac{\delta U_0}{\delta m}(\cdot, \mu, \cdot) \middle| \nu \right\rangle \geq \alpha \left\| \left\langle \frac{\delta U_0}{\delta m}(\cdot, \mu), \nu \right\rangle \right\|_{L^2}^2. \tag{5.21}$$

This assumption is a sort of strong monotonicity assumption on f and U_0. For instance for $\alpha = 0$ this assumption reduces to usual monotonicity. Furthermore, this assumption is weaker than α monotonicity. Indeed, if f is smooth and satisfies for all $\nu \in L^2(\mathbb{T}^d)$,

$$\left\langle \nu \middle| \frac{\delta f}{\delta m}(\cdot, \mu, \cdot) \middle| \nu \right\rangle \geq \alpha \|\nu\|_{L^2}^2, \tag{5.22}$$

for some $\alpha > 0$, then it satisfies (5.20) for all $\nu \in L^2(\mathbb{T}^d)$ (possibly for another $\alpha > 0$). Let us remark that such a requirement is satisfied for functions f defined by

$$(5.23) \qquad \forall x \in \mathbb{T}^d, m \in \mathcal{P}(\mathbb{T}^d), f(x,m) = \int_{\mathbb{T}^d} \Psi(z, m \star \rho(z))\rho(x-z)dz,$$

for a smooth non negative function ρ and a smooth function $\Psi : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ whose derivative with respect to the second argument $\partial_y \Psi$ satisfies $C^{-1} \leq \partial_y \Psi \leq C$ for some constant C. We can now state the a priori estimate.

Proposition 5.1. *Under Hypothesis 1, the assumption that $\mathcal{T}$ is continuous and linear and the stronger requirements* (5.20) *and* (5.21)*, for any $t_f > 0$, there exists $C > 0$ such that a classical solution U of* (5.15) *satisfies for $t \in (0, t_f), \nu, \nu' \in L^2(\mathbb{T}^d)$:*

$$(5.24) \qquad \left| \left\langle \nu \left| \frac{\delta U}{\delta m}(t, \cdot, \mu, \cdot) \right| \nu' \right\rangle \right| \leq C \|\nu\|_{L^2} \|\nu'\|_{L^2},$$

where C depends only on α, λ and $\mathcal{T}$. If $\mathcal{T}$ is non expansive in L^2, then C only depends on α.

Proof. Let us define $W, Z_\beta : [0, t_f] \times \mathcal{P}(\mathbb{T}^d) \times L^2(\mathbb{T}^d) \to \mathbb{R}$ by

$$(5.25) \qquad W(t, \mu, \nu) = \langle U(t, \cdot, \mu), \nu \rangle,$$

$$(5.26) \qquad Z_\beta(t, \mu, \nu) = \left\langle \frac{\delta W}{\delta \mu}(t, \mu, \nu), \nu \right\rangle - \beta(t) \left\langle \frac{\delta W}{\delta \mu}(t, \mu, \nu), \frac{\delta W}{\delta \mu}(t, \mu, \nu) \right\rangle,$$

for $\beta : [0, \infty) \to \mathbb{R}$ to be defined later on. Let us remark that Z_β is a quadratic and smooth function of ν. We denote by $\frac{\delta Z_\beta}{\delta \nu}$ the gradient of Z_β in $\nu \in L^2(\mathbb{T}^d)$ (since $L^2(\mathbb{T}^d)$ is an Hilbert space, this is a usual gradient). The chain rule yields that Z_β is a solution on $(0,\infty)\times\mathcal{P}(\mathbb{T}^d)\times L^2(\mathbb{T}^d)$ of

(5.27)

$$\begin{aligned}
&\partial_t Z_\beta + \left\langle -\mathrm{div}\left(D_{pp}H(\cdot,\nabla_x U(\cdot,\mu))\nabla_x \frac{\delta Z_\beta}{\delta\nu}(t,\mu,\nu,\cdot)\mu\right), \frac{\delta W}{\delta\mu}(t,\mu,\nu)\right\rangle \\
&+ \left\langle -\mathrm{div}\left(D_pH(\cdot,\nabla U(\cdot,\mu))\mu\right) - \sigma\Delta\mu, \frac{\delta Z_\beta}{\delta\mu}\right\rangle + \lambda\left(Z_\beta - Z_\beta(t,\mathcal{T}\mu,\mathcal{T}\nu)\right) \\
&- \left\langle \Delta\frac{\delta Z_\beta}{\delta\nu}(t,\mu,\nu,\cdot) - D_pH(\cdot,\nabla U)\nabla\frac{\delta Z_\beta}{\delta\nu}(t,\mu,\nu,\cdot),\nu\right\rangle \\
=&\left\langle \nu\middle|\frac{\delta f}{\delta m}(\cdot,\mu,\cdot)\middle|\nu\right\rangle + \left\langle -\mathrm{div}\left(D_{pp}H(\cdot,\nabla_x U(\cdot,\mu))\nabla_x\frac{\delta W}{\delta\mu}(t,\mu,\nu,\cdot)\mu\right), \frac{\delta W}{\delta\mu}(t,\mu,\nu)\right\rangle \\
&- \left\langle -\sigma\Delta\frac{\delta W}{\delta\mu} - \mathrm{div}\left(D_pH(\cdot,\nabla U)\frac{\delta W}{\delta\mu}\right),\nu\right\rangle \\
&- \left\langle \sigma\Delta\frac{\delta W}{\delta\mu}(t,\mu,\nu,\cdot) - D_pH(\cdot,\nabla U)\nabla\frac{\delta W}{\delta\mu}(t,\mu,\nu,\cdot),\nu\right\rangle \\
&- 2\beta\left\langle \nu\middle|\frac{\delta f}{\delta m}(\cdot,\mu,\cdot)\middle|\frac{\delta W}{\delta\mu}\right\rangle + 2\beta\left\langle -\sigma\Delta\frac{\delta W}{\delta\mu} - \mathrm{div}\left(D_pH(\cdot,\nabla U)\frac{\delta W}{\delta\mu}\right), \frac{\delta W}{\delta\mu}\right\rangle \\
&+ \beta\lambda\left(\left\|\frac{\delta W}{\delta\mu}\right\|^2 + \left\|\frac{\delta W}{\delta\mu}(t,\mathcal{T}\mu,\mathcal{T}\nu)\right\|^2 - 2\left\langle \mathcal{T}\frac{\delta W}{\delta\mu},\frac{\delta W}{\delta\mu}(t,\mathcal{T}\mu,\mathcal{T}\nu)\right\rangle\right) - \frac{d\beta}{dt}\left\|\frac{\delta W}{\delta\mu}\right\|^2.
\end{aligned}$$

The convexity of H and calculus on the term in λ yields

$$\begin{aligned}
&\partial_t Z_\beta + \left\langle -\mathrm{div}\left(D_{pp}H(\cdot,\nabla_x U(\cdot,\mu))\nabla_x \frac{\delta Z_\beta}{\delta\nu}(t,\mu,\nu,\cdot)\mu\right), \frac{\delta W}{\delta\mu}(t,\mu,\nu)\right\rangle \\
&+ \left\langle -\mathrm{div}\left(D_pH(\cdot,\nabla U(\cdot,\mu))\mu\right) - \sigma\Delta\mu, \frac{\delta Z_\beta}{\delta\mu}\right\rangle + \lambda\left(Z_\beta - Z_\beta(t,\mathcal{T}\mu,\mathcal{T}\nu)\right) \\
&- \left\langle \Delta\frac{\delta Z_\beta}{\delta\nu}(t,\mu,\nu,\cdot) - D_pH(\cdot,\nabla U)\nabla\frac{\delta Z_\beta}{\delta\nu}(t,\mu,\nu,\cdot),\nu\right\rangle \\
\geq&\left\langle \nu\middle|\frac{\delta f}{\delta m}(\cdot,\mu,\cdot)\middle|\nu\right\rangle - 2\beta\left\langle \nu\middle|\frac{\delta f}{\delta m}(\cdot,\mu,\cdot)\middle|\frac{\delta W}{\delta\mu}\right\rangle \\
&+ \beta\lambda\left(\left\|\frac{\delta W}{\delta\mu}\right\|^2 - \left\|\mathcal{T}\frac{\delta W}{\delta\mu}\right\|^2\right) - \frac{d\beta}{dt}\left\|\frac{\delta W}{\delta\mu}\right\|^2.
\end{aligned} \tag{5.28}$$

Using the assumption on f, we deduce that

(5.29)

$$\begin{aligned}
&\left\langle \nu\middle|\frac{\delta f}{\delta m}(\cdot,\mu,\cdot)\middle|\nu\right\rangle - 2\beta\left\langle \nu\middle|\frac{\delta f}{\delta m}(\cdot,\mu,\cdot)\middle|\frac{\delta W}{\delta\mu}\right\rangle + \beta\lambda\left(\left\|\frac{\delta W}{\delta\mu}\right\|^2 - \left\|\mathcal{T}\frac{\delta W}{\delta\mu}\right\|^2\right) - \frac{d\beta}{dt}\left\|\frac{\delta W}{\delta\mu}\right\|^2 \\
&\geq \left(-\alpha^{-2}\beta^2 + \lambda\left(1 - \|\mathcal{T}\|_{\mathcal{L}(L^2)}\right)\beta - \frac{d\beta}{dt}\right)\left\|\frac{\delta W}{\delta\mu}\right\|^2.
\end{aligned}$$

From the assumption on U_0, we deduce that for any $\beta(0) \in (0,\alpha)$, Z_β is non-negative at $t=0$. Hence we deduce that there exists β such that

- β is defined on $[0,t_f]$.
- The right hand side of (5.28) is non negative for $t \in (0,t_f)$
- Z_β is non-negative at $t=0$
- There exists $C>0$ depending only on α, λ and $\mathcal{T}$ such that for $t \in (0,t_f)$, $\beta(t) \geq C^{-1}$.

From the first two points of the previous list, we obtain, using a maximum principle like result that we do not detail but only sketch, that Z_β is non-negative for $t \in [0,t_f)$. The maximum principle result we do not detail can be obtained by arguing by contradiction and considering t^*, the infimum of times t for which $Z_\beta \geq 0$ is not true. Because Z_β is smooth, quadratic in ν and $\mathcal{P}(\mathbb{T}^d)$ is compact, we can consider a point of minimum $(\mu^*,\nu^*) \in \mathcal{P}(\mathbb{T}^d) \times L^2(\mathbb{T}^d)$ of $Z_\beta(t^*)$. Evaluating (5.28) at this point we easily arrive at a contradiction.

Rewriting the non-negativity of Z_β, we obtain for any $t \in (0,t_f), \nu \in L^2, \mu \in \mathcal{P}(\mathbb{T}^d)$

$$\left\| \left\langle \frac{\delta U}{\delta m}(\cdot,\mu), \nu \right\rangle \right\|_{L^2}^2 \leq C \left\langle \nu \left| \frac{\delta U}{\delta m}(t,\cdot,\mu,\cdot) \right| \nu \right\rangle . \tag{5.30}$$

From Cauchy-Schwarz inequality we deduce that

$$\left\| \left\langle \frac{\delta U}{\delta m}(\cdot,\mu), \nu \right\rangle \right\|_{L^2} \leq C \left\| \nu \right\|_{L^2} . \tag{5.31}$$

The requited estimate then easily follows. The fact that when $\mathcal{T}$ is non expansive for the L^2 norm, then C does not depend on λ nor $\mathcal{T}$ can be easily observed at the level of (5.28). □

Remark 5.7. *Let us insist that the a priori estimate we just presented is also valid for the master equations* (5.1) *and* (5.7)*, the proofs of these facts follow exactly the same argument as the one we presented.*

5.5. **Existence of solutions of first order master equations with common noise.** As we already said, the question of existence of solutions of the previous master equations is not the main interest of this paper. However, in this section, we take some time to explain how the previous a priori estimate is helpful to establish the existence of a monotone solution of (5.15) under some additional assumptions on the jump operator $\mathcal{T}$. We also explain briefly afterwards how such a result can be generalized.

Theorem 5.4. *Assume that $\mathcal{T}$ is given by* (5.13) *for some smooth non negative function K and that the assumption of Theorem 2.1 and Proposition 5.1 are satisfied. Then there exists a (unique) monotone solution of* (5.15) *in the sense of definition 5.3.*

Proof. Let us define the operator $\mathcal{A}$ which associates to a function V the solution $U := \mathcal{A}(V)$ of the master equation

(5.32)

$$\partial_t U - \sigma\Delta U + H(x, \nabla_x U) - \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \mathrm{div}\,(D_p H(\cdot, \nabla U(\cdot, m))m) \right\rangle$$
$$+ \lambda U - \sigma \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \Delta m \right\rangle = f(x,m) + \lambda \mathcal{T}^*(V(t,\cdot,\mathcal{T}(m))) \text{ in } (0,\infty)\times\mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d),$$
$$U(0,x,m) = U_0(x,m) \text{ in } \mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d).$$

Considering $\mathcal{T}^*(V(t,\cdot,\mathcal{T}(m)))$ in the same way as f, the previous equation falls in the scope of the previous analysis (Theorem 2.1 and Proposition 4.1) except for the presence of the term λU which does not play any sort of role in the previous result. Thus, we admit that they extend to this situation.

From this we deduce that the operator $\mathcal{A}$ is well defined from the set of functions V which are monotone and satisfies the assumptions (for f) of Theorem 2.1. Moreover, if $V(t)$ is C Lipschitz, uniformly in t, seen as an operator from $L^2(\mathbb{T}^d)$ into itself, then from the regularity of the kernel K that we assumed (the kernel of the operator $\mathcal{T}$), then $\mathcal{T}^*(V(t,\cdot,\mathcal{T}(\cdot)))$ is in fact (uniformly in t) Lipschitz continuous from $\mathcal{P}(\mathbb{T}^d)$ to $\mathcal{C}^2(\mathbb{T}^d)$. From this we deduce that the operator $\mathcal{A}$ is compact (Proposition 4.1).

Moreover, the a priori estimate of Proposition 5.1 yields that $\{U, \exists\theta\in[0,1], \theta\mathcal{A}(U) = U\}$ is bounded in the set of Lipschitz continuous functions $[0,t_f]\times L^2(\mathbb{T}^d)\to L^2(\mathbb{T}^d)$ for any $t_f > 0$. Hence it is a uniformly continuous family of functions over $[0,t_f]\times\mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d)$ thanks to Proposition 4.1. Hence it is bounded in $\mathcal{B}_t$. Finally let us recall that from Proposition 3.5, we know that $\mathcal{A}$ is continuous for the topology of $\mathcal{B}_t$. Hence we deduce that $\mathcal{A}$ has a fixed point. □

Let us comment on extensions of this result. First, concerning master equations involving additional stochastic parameters, it is clear that the case of (5.1) can be treated in a similar fashion whereas the case of (5.7) requires an additional a priori estimate on the regularity of the solution with respect to p.

Moreover, it also seems possible to extend the previous result to more general jump operators $\mathcal{T}$. We now briefly explain how this should be possible. Let us recall that the a priori estimate of Proposition 5.1 only requires the L^2 Lipschitz continuity of $\mathcal{T}$ whereas the assumptions of Theorem 5.4 require strong regularization properties of $\mathcal{T}$. Let us now recall that obviously the presence of the i.i.d. noise in the MFG induces a regularization properties of the underlying Fokker-Planck equation. In other words, if at a time t the repartition of players is characterized by a measure $m(t)$ and if all the players play the Nash equilibrium, then instantly the measure describing the repartition of players has a density in $L^2(\mathbb{T}^d)$ (even much more regular than L^2 under the standing assumptions in fact). Then we could use the L^2 a priori estimate of Proposition 5.1 to gain additional regularity for solutions of (5.15) with bootstrapping techniques and ideas from Proposition 4.1 (or more precisely from Section 3 in [6]).

5.6. **Asymptotic differential operators from common jumps.** Another interesting features of the kind of common noise of Section 5.3 is that one can obtained differential terms in the m variable by taking limits of such noises. We briefly extend the results and ideas from [3] to this continuous state space case. Let us first mention, that the particular form of the operator $\mathcal{T}$ was of no interest in the previous section. We only used its linearity (more precisely the existence of an adjoint...). Hence, we shall only assume here that $\mathcal{T}$ is a linear operator from the space of measures on $\mathbb{T}^d$ into itself which maps $\mathcal{P}(\mathbb{T}^d)$ into itself. Let us remark that we do not establish precise result of convergence in this section but rather present the natural asymptotics one can obtain from common jumps.

5.6.1. *General case.* In a first time, let us assume that $\mathcal{T}$ is of the form

$$\mathcal{T} = Id + \lambda^{-1}\mathcal{S}, \tag{5.33}$$

for a linear operator $\mathcal{S}$ which maps $\mathcal{P}(\mathbb{T}^d)$ into the set of measure of mass zero. The previous study can be extended for such an operator $\mathcal{T}$, in particular the associated master equation is still (5.15). Observe that, at least formally,
(5.34)
$$\lambda(U(t,x,m) - \mathcal{T}^*U(t,x,\mathcal{T}(m))) \underset{\lambda\to\infty}{\longrightarrow} -\left\langle \frac{\delta U}{\delta m}(t,x,m,\cdot), \mathcal{S}(m)\right\rangle - \mathcal{S}^*(U(t,x,m)).$$

As one can guess, the limit master equation
(5.35)
$$\begin{aligned}
&\partial_t U - \sigma\Delta U + H(x,\nabla_x U) - \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \operatorname{div}(D_pH(\cdot,\nabla U(\cdot,m))m)\right\rangle - \mathcal{S}^*(U(t,x,m))) \\
&\quad - \left\langle \frac{\delta U}{\delta m}(t,x,m,\cdot), \mathcal{S}(m)\right\rangle - \sigma\left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \Delta m\right\rangle = f(x,m) \text{ in } (0,\infty)\times\mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d),\\
&U(0,x,m) = U_0(x,m) \text{ in } \mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d),
\end{aligned}$$

also propagates monotonicity and is thus adequate for a notion of monotone solution which can be stated in this situation with

Definition 5.4. *A function $U \in \mathcal{B}_t$ is a monotone solution of* (5.15) *if*

- *For any $\mathcal{C}^2$ function $\phi : \mathbb{T}^d \to \mathbb{R}$, for any measure $\nu \in \mathcal{M}(\mathbb{T}^d)$, for any smooth function $\vartheta : [0,\infty) \to \mathbb{R}$ and any point $(t_0, m_0) \in (0,\infty)\times\mathcal{P}(\mathbb{T}^d)$ of strict minimum of $(t,m) \to \langle U(t,\cdot,m) - \phi, m - \nu\rangle - \vartheta(t)$, the following* *holds*

$$\frac{d\vartheta}{dt}(t_0) + \langle -\sigma\Delta U + H(\cdot, \nabla_x U), m_0 - \nu\rangle - \langle \mathcal{S}^*\phi, m_0\rangle + \langle \mathcal{S}^* U, \nu\rangle \geq \langle f(\cdot, m_0), m_0 - \nu\rangle - \langle U - \phi, div(D_p H(\nabla_x U) m_0)\rangle - \sigma\langle \Delta(U - \phi), m_0\rangle. \tag{5.36}$$

- *The initial condition holds*

$$U(0, \cdot, \cdot) = U_0(\cdot, \cdot). \tag{5.37}$$

We do not detail another uniqueness result for such a case. Following [3], we could also obtained second order terms in a similar fashion. Because we detail such a fact on a specific example below, we do not focus on this asymptotic right now.

5.6.2. *Asymptotic terms associated to translations.* An important case of common jumps is the one in which $\mathcal{T}(m)$ is the image measure of m by some application $B : \mathbb{T}^d \to \mathbb{T}^d$. In such a case, at the random times at which the jumps occur, all the players in x are transported to $B(x)$. Let us assume that B is of the form $B = Id + \lambda^{-1}\tilde{B}$ for some smooth $\tilde{B} : \mathbb{T}^d \to \mathbb{T}^d$. In this context, formally, one obtain that

$$\lambda(U(t,x,m) - \mathcal{T}^* U(t,x,\mathcal{T}(m))) \underset{\lambda\to\infty}{\longrightarrow} -\langle D_m U(t,x,m,\cdot)\tilde{B}(\cdot), m\rangle - \tilde{B}(x)\cdot\nabla_x U(t,x,m). \tag{5.38}$$

Following [3], the addition of several terms modeling common jumps does not raise any particular issue and can be treated in a similar way. By doing so, we are able to obtain higher order asymptotic terms in the master equation as we now explain. Let us define $\mathcal{T}_+(m)$ the image measure of m by $Id + \lambda^{-1/2}\tilde{B}$ and $\mathcal{T}_-(m)$ the image measure of m by $Id - \lambda^{-1/2}\tilde{B}$. In such a situation, we can remark the following asymptotic

$$\begin{aligned}\lambda(2U(t,x,m) &- \mathcal{T}_+^* U(t,x,\mathcal{T}_+(m)) - \mathcal{T}_-^* U(t,x,\mathcal{T}_-(m)))\\ \underset{\lambda\to\infty}{\longrightarrow} &- \tilde{B}(x)\cdot D^2_{xx}U(t,x,m)\cdot\tilde{B}(x) - 2\tilde{B}(x)\cdot\nabla_x\langle\tilde{B}(\cdot)D_m U(t,x,m,\cdot), m\rangle\\ &- \int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\tilde{B}(y)D^2_{mm}U(t,x,m,y,z)\tilde{B}(z)m(dy)m(dz)\\ &- \langle\tilde{B}(\cdot)D^2_{yy}\frac{\delta U}{\delta m}(t,x,m,\cdot)\tilde{B}(\cdot), m\rangle,\end{aligned} \tag{5.39}$$

where $D^2_{mm} := D_m(D_m U)$. This gives us another way to derive the master equation with common noise studied in [6], which is written below as equation (6.1). Indeed to recover (6.1), one has to sum such terms as in (5.39) for the constant maps $\tilde{B}(y) = (0, ..., 0, \sqrt{\beta}, 0, ...0)$ where the $\sqrt{\beta}$ varies from position $i = 1$ to $i = d$. Even though the proper study of such an asymptotic is not the subject of this paper (and thus we do not it here), we believe that such an approach can be insightful in many ways for the study of (6.1). Moreover this justifies in some sense the generality of the noise of the form of Section 5.3.

6. Master equations of second order

We now turn to the case of second order master equations such as the one studied in [6]. We are here interested in the equation

(6.1)
$$\begin{aligned}
&\partial_t U - (\sigma+\beta)\Delta U + H(x,\nabla_x U) - \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \operatorname{div}\left(D_p H(\cdot,\nabla U(\cdot,m))m\right)\right\rangle \\
&\qquad - (\sigma+\beta)\left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \Delta m\right\rangle + 2\beta\nabla_x \cdot \left\langle \frac{\delta U}{\delta m}(x,m,\cdot), \nabla m\right\rangle \\
&\qquad - \beta\left\langle \nabla m \middle| \frac{\delta^2 U}{\delta m^2}(x,m,\cdot,\cdot) \middle| \nabla m\right\rangle = f(x,m), \\
&U(0,x,m) = U_0(x,m).
\end{aligned}$$

Let us precise that the last term of the left hand side of (6.1) is understood as $-\beta\sum_{i=1}^d \langle \partial_i m | \frac{\delta^2 U}{\delta m^2} | \partial_i m\rangle$. As we did for (2.3), we now state a uniqueness result for (6.1) for which we give a different proof than in [6], where it originates from.

Proposition 6.1. *Under Hypothesis 1, there exists at most one classical solution of* (6.1)*.*

Proof. Let us consider U and V two smooth solutions and define W by

$$W(t,\mu,\nu) = \langle U(t,\cdot,\mu) - V(t,\cdot,\nu), \mu-\nu\rangle. \tag{6.2}$$

A simple computation yields that W is a solution of

(6.3)
$$\begin{aligned}
&\partial_t W - (\sigma+\beta)\langle \Delta_x(U-V) + H(\cdot,\nabla U) - H(\cdot,\nabla V), \mu-\nu\rangle \\
&\quad - \left\langle \frac{\delta W}{\delta \mu}(\mu,\nu), \operatorname{div}\left(D_p H(\cdot,\nabla U(\cdot,\mu))\mu\right)\right\rangle - \left\langle \frac{\delta W}{\delta \nu}(\mu,\nu), \operatorname{div}\left(D_p H(\cdot,\nabla V(\cdot,\nu))\nu\right)\right\rangle \\
&\quad - (\sigma+\beta)\left\langle \frac{\delta W}{\delta \mu}(\mu,\nu), \Delta\mu\right\rangle - (\sigma+\beta)\left\langle \frac{\delta W}{\delta \nu}(\mu,\nu), \Delta\nu\right\rangle \\
&\quad - \beta\left(\left\langle \nabla\mu \middle| \frac{\delta^2 W}{\delta \mu^2}(\mu,\nu,\cdot,\cdot)\middle|\nabla\mu\right\rangle + \left\langle \nabla\nu \middle| \frac{\delta^2 W}{\delta \nu^2}(\mu,\nu,\cdot,\cdot)\middle|\nabla\nu\right\rangle - 2\left\langle \nabla\nu \middle| \frac{\delta^2 W}{\delta\mu\delta\nu}(\mu,\nu,\cdot,\cdot)\middle|\nabla\mu\right\rangle\right) \\
&= \langle f(\cdot,\mu) - f(\cdot,\nu), \mu-\nu\rangle - \langle U(\cdot,\mu) - V(\cdot,\nu), \operatorname{div}\left(D_p H(\cdot,\nabla U(\cdot,\mu))\mu\right)\rangle \\
&\quad - \langle V(\cdot,\nu) - U(\cdot,\mu), \operatorname{div}\left(D_p H(\cdot,\nabla V(\cdot,\nu))\nu\right)\rangle - (\sigma+\beta)\langle \Delta_x(U-V), \mu-\nu\rangle.
\end{aligned}$$

Once again, using the monotonicity of f and the convexity of H, we obtain that the right hand side of the previous equation is non-negative. Using a maximum principle like result (that we do not detail here), we deduce that W is a non-negative function, from which the required result easily follows, as in the case without common noise. Let us briefly observe the, maybe, easiest way to obtain that the particular structure of the terms in β indeed behaves as an elliptic term, which yields maximum principle like result : Those terms can be obtained from common jumps which satisfy such elliptic property. Indeed, following exactly what we did in section 5.6.2, we observe that the terms in β are the limits of terms arising from jumps which verify the property that they are non-negative at a point of minimum of W. Because we assumed that W is smooth, the passage to the limit of section 5.6.2 is here correct and thus an elliptic property is proved.

□

Observing the similarity between this equation and the first order one (2.3), it is very tempting to formulate an adaptation of the definitions of monotone solutions we gave earlier. Two main difficulties arise at this point. The first one has to do with the fact that (6.1) is of second order in the measure argument. As monotone solutions are reminiscent of viscosity solutions, similar technical difficulties naturally arise, we refer to [12] for more details on viscosity solutions of second order equations in a finite dimensional setting. The second difficulty has to do with the precise nature of (6.1). More or less, the definition of monotone solution we give consists in reformulating that $W : (t,\mu) \to \langle U(t,\mu) - \phi, \mu - \nu\rangle$ is a sort of super solution of a certain PDE. The problem is here that in this second order case, the PDE satisfied by W cannot be expressed solely in terms of W and its derivatives, there is a term which involves directly the first order derivative of U which cannot be rewritten with W. Indeed, if U solves (6.1) then W defined above for $\phi \in \mathcal{C}^2, \nu \in \mathcal{P}(\mathbb{T}^d)$ satisfies

(6.4)

$$\begin{aligned}
&\partial_t W + \langle -(\sigma+\beta)\Delta U + H(\cdot, \nabla_x U) - f(\mu), \mu - \nu\rangle - \left\langle \frac{\delta W}{\delta \mu}, \operatorname{div}(D_p H(\cdot, \nabla U)\mu)\right\rangle \\
&\quad + \langle U - \phi, \operatorname{div}(D_p H(\cdot, \nabla U)\mu)\rangle - (\sigma+\beta)\left\langle \frac{\delta W}{\delta \mu}, \Delta \mu\right\rangle + (\sigma+\beta)\langle U - \phi, \Delta\mu\rangle \\
&\quad + 2\beta \left\langle \nabla \nu \middle| \frac{\delta U}{\delta m}(\cdot,\mu,\cdot) \middle| \nabla \mu \right\rangle - \beta \left\langle \nabla \mu \middle| \frac{\delta^2 W}{\delta \mu^2}(\cdot,\cdot) \middle| \nabla\mu\right\rangle = 0.
\end{aligned}$$

A way to define monotone solutions could be to ask W to be a super solution of this equation. If we do so, the previous proof is still valid, at least formally. The challenge here is that the second to last term cannot be written without using derivative (with respect to the measure variable) of U. This difficulty could be overcome by asking the couple (U, W) to be a super solution of this equation. This idea is probably the most general one, for instance in the finite state space case, it is immediate to verify that it yields uniqueness of solutions by proving the non-negativity of $\langle U - V, \mu - \nu\rangle$ for U and V two monotone solutions, using the techniques of [12]. However, because the technical difficulties usually arising from viscosity solutions of second order are much more difficult to treat in this infinite dimensional setting we focus on a regime which is more regular than the one we adopted up to this point. Formally we restrict ourselves to $\mathcal{C}^1$ solution in the measure argument and we use the limit (5.39) to express the second order derivatives. This discussion leads us to the following definition.

Definition 6.1. *A function U in $\mathcal{B}_t$, which is derivable with respect to the measure variable at every point, is a monotone solution of* (6.1) *if*

- *for any $\mathcal{C}^2$ functions $\phi : \mathbb{T}^d \to \mathbb{R}, \vartheta : (0,\infty) \to \mathbb{R}$, for any measure $\nu \in \mathcal{M}(\mathbb{T}^d)$, and (t_0, m_0) point of strict minimum of the function W defined by $(t,m) \to \langle U(t,\cdot,m) - \phi, m - \nu\rangle - \vartheta(t)$ the following holds*

(6.5)
$$\begin{aligned}\frac{d\vartheta}{dt}(t_0) &+ \langle -\sigma\Delta U + H(\cdot,\nabla_x U), m_0 - \nu\rangle + 2\beta\left\langle \nabla\nu \middle| \frac{\delta U}{\delta m}(\cdot,\mu,\cdot)\middle|\nabla\mu\right\rangle \\ &+ \beta \liminf_{h\to 0}\sum_{i=1}^{d} h^{-2}(2W(t_0,m_0) - W(t_0, T^h_{+i}m_0) - W(t_0, T^h_{-i}m_0)) \\ &\geq \langle f(\cdot,m_0), m_0 - \nu\rangle - \langle U - \phi, div(D_pH(\nabla_x U)m_0)\rangle - \sigma\langle\Delta(U-\phi), m_0\rangle.\end{aligned}$$

- *The initial condition holds*

(6.6)
$$U(0,\cdot,\cdot) = U_0(\cdot,\cdot).$$

In the previous definition, we used the notation

(6.7)
$$\forall 1 \leq i \leq d, m \in \mathcal{P}(\mathbb{T}^d), T^h_{\pm i}m = (Id \pm he_i)_{\#}m,$$

where for $1 \leq i \leq d$, $e_i = (0,...,1,...,0)$ and the sole 1 is in position i. For this notion of solution, we can establish the following result of uniqueness

Theorem 6.1. *Under Hypothesis 1, two monotone solutions of* (6.1) *in the sense of Definition 6.1 only differ by a function $c : [0,\infty)\times\mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$. If such a solution U exists, then $U(t)$ is monotone for all time $t \geq 0$. If, in addition, Hypothesis 2 is satisfied, then there is at most one monotone solution of* (6.1).

Proof. Let us consider U and V two solutions, and define W by

(6.8)
$$W(t,s,\mu,\nu) = \langle U(t,\cdot,\mu) - V(s,\cdot,\nu), \mu - \nu\rangle.$$

We want to show that $W(t,t,\cdot,\cdot) \geq 0$ for all $t \geq 0$. Once again we argue by contradiction. Hence, we assume that there exists $t_*, \delta, \bar\epsilon > 0$, such that for all $\epsilon \in (0,\bar\epsilon), \alpha > 0, \phi, \psi \in \mathcal{C}^2$ such that $\|\phi\|_2 + \|\psi\|_2 \leq \epsilon$ and $\gamma_1, \gamma_2 \in (\frac{\bar\epsilon}{2}, \bar\epsilon)$,

(6.9)
$$\inf_{t,s\in[0,t_*],\mu,\nu\in\mathcal{P}(\mathbb{T}^d)}\left\{W(t,s,\mu,\nu) + \langle\phi,\mu\rangle + \langle\psi,\nu\rangle + \frac{1}{2\alpha}(t-s)^2 + \gamma_1 t + \gamma_2 s\right\} \leq -\delta.$$

From Lemma 2.1, we know that there exists (for any value of α) ϕ, ψ, γ_1 and γ_2 such that $(t,s,\mu,\nu) \to W(t,s,\mu,\nu) + \langle\phi,\mu\rangle + \langle\psi,\nu\rangle + \frac{1}{2\alpha}(t-s)^2 + \gamma_1 t + \gamma_2 s$ has a strict minimum on $[0,t_*]^2 \times \mathcal{P}(\mathbb{T}^d)^2$ at (t_0, s_0, μ_0, s_0).

We now assume that $t_0, s_0 > 0$ (the case in which one of the two is zero is treated exactly as in Theorem 3.2). Let us now remark the following

$$
\begin{aligned}
0 \geq & \liminf_{h\to 0} \sum_{i=1}^{d} h^{-2}(2W(t_0, s_0, \mu_0, \nu_0) - W(t_0, s_0, T^h_{+i}\mu_0, T^h_{+i}\nu_0) - W(t_0, w_0, T^h_{-i}\mu_0, T^h_{-i}\nu_0)) \\
= & \liminf_{h\to 0} \sum_{i=1}^{d} h^{-2}(2W(t_0, s_0, \mu_0, \nu_0) - W(t_0, s_0, T^h_{+i}\mu_0, \nu_0) - W(t_0, w_0, T^h_{-i}\mu_0, \nu_0)) \\
& + \liminf_{h\to 0} \sum_{i=1}^{d} h^{-2}(2W(t_0, s_0, \mu_0, \nu_0) - W(t_0, s_0, \mu_0, T^h_{+i}\nu_0) - W(t_0, w_0, \mu_0, T^h_{-i}\nu_0)) \\
& + 2\beta\left(\left\langle \nabla\nu_0 \middle| \frac{\delta U}{\delta m}(\cdot, \mu_0, \cdot) \middle| \nabla\mu_0 \right\rangle + \left\langle \nabla\mu_0 \middle| \frac{\delta V}{\delta m}(\cdot, \nu_0, \cdot) \middle| \nabla\nu_0 \right\rangle\right)
\end{aligned} \tag{6.10}
$$

Using the definition of monotone solutions of both U and V, we then arrive at a contradiction following the same argument as in the proof of Theorem 3.2. The proof of the second part of the claim can also be established by following the same argument of the proof of Theorem 3.2, namely by choosing the measure ρ as the Lebesgue measure on the Torus $\mathbb{T}^d$ (using the same notation as the aforementioned proof). □

Remark 6.1. *Let us insist on the fact that for two smooth functions $U, V : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$, the crossed derivative of $W(\mu, \nu) = \langle U(\mu) - V(\nu), \mu - \nu\rangle$ is given, up to a constant, by*

$$
\forall x, y \in \mathbb{T}^d, \frac{\delta^2 W}{\delta\mu\delta\nu}(\mu, \nu, x, y) = \frac{\delta U}{\delta m}(\mu, y, x) + \frac{\delta V}{\delta m}(\nu, x, y). \tag{6.11}
$$

Remark 6.2. *If we were to work in an Hilbert space, then an analogous of Theorem 3.2 in [12] could probably have been established using the properties of semi convex functions in Hilbert space such as in [18], then allowing us to use a weaker notion of solutions.*

Let us remark that following Section 4, we can obtain a result of existence of monotone solutions of (6.1) by weakening the assumptions of Theorem 2.11 in [6]. As the approach is very similar to the one of Section 4, we only state the following without a demonstration.

Theorem 6.2. *Assume that the assumptions of Theorem 2.1 hold and that f and U_0 can be approximated uniformly by functions f_n and $U_{0,n}$ satisfying for some*

$C_n > 0$
(6.12)

$$\sup_{m\in\mathcal{P}(\mathbb{T}^d)} \left(\|f(\cdot,m)\|_{2+\alpha} + \left\| \frac{\delta f(\cdot,m,\cdot)}{\delta m} \right\|_{(2+\alpha,2+\alpha)} + \left\| \frac{\delta^2 f(\cdot,m,\cdot,\cdot)}{\delta m^2} \right\|_{(2+\alpha,2+\alpha,2+\alpha)} \right) + Lip_{2+\alpha}\left(\frac{\delta^2 f}{\delta m^2}\right) \leq C_n.$$

(6.13)

$$\sup_{m\in\mathcal{P}(\mathbb{T}^d)} \left(\|U_0(\cdot,m)\|_{3+\alpha} + \left\| \frac{\delta U_0(\cdot,m,\cdot)}{\delta m} \right\|_{(3+\alpha,3+\alpha)} + \left\| \frac{\delta^2 U_0(\cdot,m,\cdot,\cdot)}{\delta m^2} \right\|_{(3+\alpha,3+\alpha,3+\alpha)} \right) + Lip_{3+\alpha}\left(\frac{\delta^2 U_0}{\delta m^2}\right) \leq C_n.$$

Then there exists a monotone solution of 6.1 in the sense of Definition 6.1.

Remark 6.3. *As the second order terms in* (6.1) *can be seen as limits of common jumps for jump operators which are non expansive in $L^2(\mathbb{T}^d)$ (they are translations), let us remark that the a priori estimate established in Proposition 5.1 is also valid for* (6.1)*. This could have been observed by applying directly the same technique as in the proof of Proposition 5.1 on* (6.1)*.*

7. Conclusion and future perspectives

We have presented a notion of solution for MFG master equations which allows us to work with solutions which are merely continuous for first order equations and only one time differentiable for second order equations (each time with respect to the measure argument). This notion is built to enjoy nice uniqueness properties but it also verifies strong stability properties. Even though we do not treat exhaustively the question of existence, the stability properties is of course helpful to establish such results. Let us mention that the generalization of our results to master equation which do not model exactly MFG but which have a monotone structure (such as in [21] in the continuous setting) is of course immediate following [2] modulo some technical assumptions on the non-linearities. The extension to master equations on $\mathcal{M}_+(\mathbb{T}^d) := \{m \in \mathcal{M}(\mathbb{T}^d) | m \geq 0\}$ is also straightforward.

At this point, we believe worth mentioning several extensions or future directions to explore, that we believe to be meaningful. Perhaps the most important one concerns the extension of monotone solutions of second order to merely continuous functions of the measure argument. In this setting, such an improvement should probably pass by a better understanding of the derivatives of functions on $\mathcal{P}(\mathbb{T}^d)$, in order to establish an analogous of Theorem 3.2 in [12]. Another interesting extension is the adaptation of the previous results to MFG of optimal stopping or other singular controls, as it has been done in [2] in the finite state space case. For such problems, a uniqueness result for monotone solutions of the master equation (which is now posed on the set of non-negative measure and not on the set of probability measures) follows quite immediately from the the present study, using a similar formulation such as in [2]. However, existence questions require key estimates which have not been proven at the moment.

Other extensions of this work include the study of problems on more general domain than the Torus, thus involving boundary conditions, as well as numerical methods to approximate monotone solutions.

References

[1] Charles Bertucci. Optimal stopping in mean field games, an obstacle problem approach. *Journal de Mathématiques Pures et Appliquées*, 120:165–194, 2018.

[2] 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.

[3] 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.

[4] Charles Bertucci, Jean-Michel Lasry, and Pierre-Louis Lions. Master equation for the finite state space planning problem. *to appear in Archive for Rational Mechanics and Analysis*, 2021.

[5] Pierre Cardaliaguet and Panagiotis Souganidis. On first order mean field game systems with a common noise. *arXiv preprint arXiv:2009.12134*, 2020.

[6] 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.

[7] René Carmona, François Delarue, Daniel Lacker, et al. Mean field games with common noise. *Annals of Probability*, 44(6):3740–3803, 2016.

[8] René Carmona, François Delarue, et al. *Probabilistic Theory of Mean Field Games with Applications I-II.* Springer, 2018.

[9] Alekos Cecchin and François Delarue. Selection by vanishing common noise for potential finite state mean field games. *arXiv preprint arXiv:2005.12153*, 2020.

[10] Alekos Cecchin, Paolo Dai Pra, Markus Fischer, and Guglielmo Pelino. On the convergence problem in mean field games: a two state model without uniqueness. *SIAM Journal on Control and Optimization*, 57(4):2443–2466, 2019.

[11] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society*, 277(1): 1–42, 1983.

[12] Michael G Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American mathematical society*, 27(1):1–67, 1992.
[13] 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.
[14] 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.
[15] Minyi Huang, Roland P Malhamé, Peter E Caines, et al. Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.
[16] Daniel Lacker. A general characterization of the mean field limit for stochastic differential games. *Probability Theory and Related Fields*, 165(3-4):581–648, 2016.
[17] Daniel Lacker. On the convergence of closed-loop nash equilibria to the mean field game limit. *The Annals of Applied Probability*, 30(4):1693–1761, 2020.
[18] Jean-Michel Lasry and Pierre-Louis Lions. A remark on regularization in hilbert spaces. *Israel Journal of Mathematics*, 55(3):257–266, 1986.
[19] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.
[20] Pierre-Louis Lions. Cours au college de france. *www.college-de-france.fr*, 2011.
[21] Pierre-Louis Lions and Panagiotis E Souganidis. Extended mean-field games. *Rendiconti Lincei-Matematica e Applicazioni*, 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.
[23] Robert R Phelps. *Convex functions, monotone operators and differentiability*, volume 1364. Springer, 2009.
[24] Alessio Porretta. Weak solutions to fokker–planck equations and mean field games. *Archive for Rational Mechanics and Analysis*, 216(1):1–62, 2015.