J. Eur. Math. Soc. (Galley proof) DOI 10.4171/JEMS/1227

Published by EMS Press

Pierre Cardaliaguet · Marco Cirant · Alessio Porretta

Splitting methods and short time existence for the master equations in mean field games

Received January 24, 2020

Abstract. We develop a splitting method to prove the well-posedness, in short time, of solutions for two master equations in mean field game (MFG) theory: the second order master equation, describing MFGs with a common noise, and the system of master equations associated with MFGs with a major player. Both problems are infinite-dimensional equations stated in the space of probability measures. Our new approach simplifies, shortens and generalizes previous existence results for second order master equations and provides the first existence result for systems associated with MFG problems with a major player.

Keywords. Master equation, Mean Field Games, MFG with major player

Contents

Pierre Cardaliaguet: Ceremade, CNRS, Université Paris-Dauphine, PSL Research University, 75016 Paris, France; cardaliaguet@ceremade.dauphine.fr

Marco Cirant: Dipartimento di Matematica, Università di Padova, Via Trieste 63, 35121 Padova, Italy; cirant@math.unipd.it

Alessio Porretta: Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italy; porretta@mat.uniroma2.it

Mathematics Subject Classification (2020): 35K55, 49N70

1. Introduction

The paper is dedicated to the construction of a solution of the so-called "master equations" in mean field game theory (MFG). These equations have been introduced by Lasry and Lions and discussed by Lions in [27]. Let us recall that mean field games describe the behavior of infinitely many agents in interaction. We consider here two problems: the master equation with common noise and the master equation with a major player. We present a general approach valid for both problems.

Let us first discuss the master equation with common noise. In this problem, the agents are subject to a common source of randomness. The master equation is then a second order equation in the space of measures and reads as follows:

$$\begin{cases} -\partial_t U(t,x,m) - \mathrm{Tr}((a(t,x)+a^0(t,x))D^2_{xx}U(t,x,m)) + H(x, D_xU(t,x,m), m) \\ \qquad -\displaystyle\int_{\mathbb{R}^d} \mathrm{Tr}((a(t,y)+a^0(t,y))D^2_{ym}U(t,x,m,y))\, m(dy) \\ \qquad +\displaystyle\int_{\mathbb{R}^d} D_mU(t,x,m,y)\cdot H_p(y, D_xU(t,y,m),m)\, m(dy) \\ \qquad -2\displaystyle\int_{\mathbb{R}^d} \mathrm{Tr}[\sigma^0(t,y)(\sigma^0(t,x))^T D^2_{xm}U(t,x,m,y)]\, m(dy) \\ \qquad -\displaystyle\int_{\mathbb{R}^{2d}} \mathrm{Tr}[\sigma^0(t,y)(\sigma^0(t,y'))^T D^2_{mm}U(t,x,m,y,y')]\, m(dy)\, m(dy') = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2, \\ U(T,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2, \end{cases} \tag{1}$$

In the above equation, the unknown $U = U(t,x,m)$ is scalar valued and depends on the time variable $t \in [0,T]$, the space variable $x \in \mathbb{R}^d$ and the distribution of the agents m in $\mathcal{P}_2$, the space of Borel probability measures with finite second order moment; the derivatives D_mU and $D^2_{mm}U$ refer to the derivative with respect to the probability measure (see Section 2.2); the maps $H = H(x,p,m)$ and $G = G(x,m)$ reflect the running and terminal costs of the agents. The matrix valued function $a = a(t,x)$ is the volatility term corresponding to idiosyncratic noise of the small players, while $a^0 = a^0(t,x) = \sigma^0(\sigma^0)^T(t,x)$ is the volatility corresponding to the common noise.

As explained by Lions [27], the master equation can be understood as a non-linear transport equation in the space of probability measures. When $a^0 = 0$ (i.e., in the so-called first order master equation), the characteristics of this transport equation are given by the MFG system: if we fix an initial time t_0 and an initial probability measure m_0 on $\mathbb{R}^d$, and if the pair (u, m) is a solution of the MFG system

$$\begin{cases} \text{(i) } -\partial_t u - \mathrm{Tr}(a(t,x)D^2u) + H(x, Du, m(t)) = 0 & \text{in } (t_0, T) \times \mathbb{R}^d, \\ \text{(ii) } \partial_t m - \sum_{i,j} D_{ij}(a_{i,j}m) - \mathrm{div}(mH_p(x, Du, m(t)) = 0 & \text{in } (t_0, T) \times \mathbb{R}^d, \\ \text{(iii) } m(t_0) = m_0, \quad u(T, x) = G(x, m(T)) & \text{in } \mathbb{R}^d, \end{cases} \tag{2}$$

then we expect the following equality to hold:

$$U(t, x, m(t)) = u(t, x) \quad \forall t \in [t_0, T]. \tag{3}$$

The interpretation of the MFG system (2) is the following: the map u is the value function of a typical small agent (anticipating the evolution of the population density $m(t)$) and accordingly solves the Hamilton–Jacobi equation (2) (i). When this agent plays in an optimal way, the drift in the dynamic of its state is given by the term $-H_p(x, Du, m(t))$. By a mean field argument (assuming that the noises of the agents are independent), the resulting evolution of the population density $\tilde{m}$ satisfies the Kolmogorov equation

$$\begin{cases} \partial_t \tilde{m} - \sum_{i,j} D_{ij}(a_{i,j}\tilde{m}) - \mathrm{div}(\tilde{m}H_p(x, Du, m(t)) = 0 \quad \text{in } (t_0, T) \times \mathbb{R}^d, \\ m(t_0) = m_0 \quad \text{in } \mathbb{R}^d. \end{cases}$$

In an equilibrium configuration, i.e., when agents correctly anticipate the evolving measure, one has $\tilde{m} = m$ and therefore the population density m solves (2) (ii).

The existence/uniqueness of the solution for the MFG system is rather well understood: it relies on Schauder estimates, fixed point methods and monotonicity arguments (see, in particular, [24, 25]). From the well-posedness of the MFG system, one can derive the existence of a solution to the first order master equation “quite easily”: one just needs to define the map U by (3) with $t = t_0$ and check that the map U thus defined is a classical solution to the first order master equation. This is the path followed in [17,29] (when there is no diffusion at all: $a = a^0 \equiv 0$) and in [11] (when $a > 0$ is constant and $a^0 = 0$). See also [10] for a similar result (for the torus) using PDE linearization techniques.

When $a^0 \not\equiv 0$ (i.e., for the second order master equation, or master equation with a common noise), the characteristics are now given by the system of SPDEs (called “stochastic MFG system”):

$$\begin{cases} du(t,x) = [-\mathrm{Tr}((a + a^0)(t,x)D^2u(t,x)) + H(x, Du(t,x), m(t)) \\ \qquad\qquad -\sqrt{2}\,\mathrm{Tr}(\sigma^0(t,x)Dv(t,x))]\,dt + v(t,x)\cdot dW_t \quad \text{in } (0,T)\times\mathbb{R}^d, \\ dm(t,x) = \Big[\sum_{i,j} D_{ij}(((a_{ij}) + a^0_{ij})(t,x)m(t,x)) \\ \qquad\qquad + \mathrm{div}(m(,x)D_pH(x, Du(t,x), m(t)))\Big]\,dt \\ \qquad\qquad - \mathrm{div}(m(t,x)\sqrt{2}\,\sigma^0(t,x)dW_t) \quad \text{in } (0,T)\times\mathbb{R}^d, \\ u(T,x) = G(x, m(T)), \quad m(0) = m_0 \quad \text{in } \mathbb{R}^d, \end{cases} \tag{4}$$

In the above system, (W_t) is the common noise (here a Brownian motion) and the unknown is the triplet (u, m, v), where the new variable v (a random vector field in $\mathbb{R}^d$)

ensures that the solution u of the backward Hamilton–Jacobi (HJ) equation is adapted to the filtration generated by the common noise (W_t). The analysis of this system is much more involved than that of the deterministic one: Schauder estimates are no longer available and the usual fixed point methods based on compactness arguments can no longer be applied. One has to replace them by continuation methods, which are much heavier to handle [10]. Besides the PDE approach we just mentioned, MFG with common noise can also be handled through a probabilistic formulation: see the pioneering result [13], as well as [2, 21] and the monograph [12]. Once the analysis of the stochastic MFG system has been performed, one can proceed with the construction of the second order master equation as in the first order case, defining the map U by (3) for $t = t_0$, where u is now the u-component of the solution of the stochastic MFG system ($u(t_0, \cdot)$ turns out to be deterministic). However, here again, the verification that the map U defined so far is smooth enough to satisfy (1) requires a lot of work: see [10] and [12].

Let us finally recall another approach, suggested by P.-L. Lions in the seminar [28]: it consists in writing the equation for the quantity $D_m U$ as a hyperbolic equation in a Hilbert space of random variables. The construction requires, however, convexity conditions on the system with respect to the space variable (but no uniform ellipticity for the matrix a).

We now discuss the second equation considered in this paper: the master equation corresponding to MFG models with a major player. It reads as follows:

$$\begin{cases} -\partial_t U^0 - \Delta_{x_0} U^0 + H^0(x_0, D_{x_0} U^0, m) - \displaystyle\int_{\mathbb{R}^d} \operatorname{div}_y D_m U^0(t, x_0, m, y)\, m(dy) \\ \qquad\qquad + \displaystyle\int_{\mathbb{R}^d} D_m U^0(t, x_0, m, y) \cdot H_p(x_0, y, D_x U(t, x_0, y, m), m)\, m(dy) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0, T) \times \mathbb{R}^{d_0} \times \mathcal{P}_2, \\ -\partial_t U - \Delta_x U - \Delta_{x_0} U + H(x_0, x, D_x U, m) - \displaystyle\int_{\mathbb{R}^d} \operatorname{div}_y D_m U(t, x_0, x, m, y)\, m(dy) \\ \qquad\qquad + D_{x_0} U \cdot H^0_p(x_0, D_{x_0} U^0(t, x_0, m), m) \\ \qquad\qquad + \displaystyle\int_{\mathbb{R}^d} D_m U(t, x_0, x, m, y) \cdot H_p(x_0, y, D_x U(t, x_0, y, m), m)\, m(dy) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0, T) \times \mathbb{R}^{d_0} \times \mathbb{R}^d \times \mathcal{P}_2, \\ U^0(T, x_0, m) = G^0(x_0, m) \qquad \text{in } \mathbb{R}^{d_0} \times \mathcal{P}_2, \\ U(T, x_0, x, m) = G(x_0, x, m) \quad \text{in } \mathbb{R}^{d_0} \times \mathbb{R}^d \times \mathcal{P}_2. \end{cases} \tag{5}$$

In the above system, $U^0 = U^0(t, x_0, m)$ corresponds to the payoff at equilibrium for a major player interacting with a crowd in which each agent has at equilibrium a payoff given by $U = U(t, x_0, x, m)$. Here m is the distribution law of the agents. Notice that each agent is influenced by the major player whereas the latter is only influenced by the distribution of the whole population. Mean field games with a major player have been first discussed by Huang [19], and several notions of equilibria, in different contexts, have been proposed in the literature since then: see [5–8, 12, 14–16, 26]. The above system has been introduced by Lasry and Lions [26]. In the companion paper [9], we explain how the above master equation is related to the approach by Carmona and al. [14–16]. Concerning the existence of a solution, [16] shows the existence of an equilibrium in short time for the case of a finite state space, [26] proves the existence of a solution to the master equation still in the finite state space framework and notes that the Hilbertian techniques described in [28] could be adapted to the master equation with a major player (5).

The purpose of this paper is to introduce a different path towards the construction of a solution to the second order master equation and to the master equation with a major player, using as a building block the construction of a solution to the first order master equation. For the second order master equation, we justify this point of view by the fact that the deterministic MFG system and the first order master equation are much easier to manipulate than the stochastic MFG system. Our approach allows one for instance to build solutions of the second order master equation (in short time) under more general assumptions than in [10, 12]. For the MFG problem with a major player, we prove for the first time the (short time) well-posedness of the associated system of master equations in continuous space.

Let us first explain our ideas for the master equation with common noise (1). In contrast to previous works, we do not use directly the representation formula (3) (for $t = t_0$) for the solution of the second order master equation. Instead, we somehow decompose the second order master equation as the superposition of the first order master equation:

$$\begin{cases} -\partial_t U - \mathrm{Tr}(a(t,x)D^2_{xx}U) + H(x, D_xU, m) - \int_{\mathbb{R}^d} \mathrm{Tr}(a(t,y)D^2_{ym}U)\, m(dy) \\ \qquad\qquad + \int_{\mathbb{R}^d} D_mU \cdot H_p(y, D_xU, m)\, m(dy) = 0 \quad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2, \\ U(T,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2 \end{cases} \tag{6}$$

and of a linear second order master equation:

$$\begin{cases} -\partial_t U - \mathrm{Tr}[\sigma^0(\sigma^0)^T(t,x)D^2_{xx}U] - \int_{\mathbb{R}^d} \mathrm{Tr}[\sigma^0(\sigma^0)^T(t,y)D^2_{ym}U]\, m(dy) \\ \qquad\qquad - 2\int_{\mathbb{R}^d} \mathrm{Tr}[\sigma^0(t,y)(\sigma^0(t,x))^T D^2_{xm}U]\, m(dy) \\ \qquad\qquad - \int_{\mathbb{R}^{2d}} \mathrm{Tr}[\sigma^0(t,y)(\sigma^0(t,y'))^T D^2_{mm}U]\, m(dy)\, m(dy') = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2, \\ U(T,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2. \end{cases} \tag{7}$$

The solution to this linear second order master equation is just given by a Feynman–Kac formula, and thus it is very easy to handle. Then we use the Trotter–Kato formula, alternating the two equations in short time intervals to build in the limit a solution of the full equation (1). Even if the technique is quite transparent, its actual implementation requires some care. Indeed, one has to check that, at each step of the process, the regularity of the solution does not deteriorate too much, meaning at least in a linear way in time. The aim of Section 5.2 is precisely to quantify this deterioration for the solution U of the first order master equation (6), as well as for its derivatives in the measure variable. As the solution of (6) is built by using the representation formula (3) (where $t = t_0$) presented above, one has first to do the analysis on the MFG system (2), and this is the aim of Section 5.1. Note that we are able to control the regularity of the linear second order equation (7) only when the matrix a^0 is constant. Hence we only prove the short time existence of a solution to (1) in that case.

For the problem with a major player, we argue in a similar way: we view equation (5) as the superposition of two systems: the first one is a first order system of master equations (for a fixed x_0):

$$\begin{cases} -\partial_t U^0 - \int_{\mathbb{R}^d} \mathrm{div}_y\, D_m U^0(t, x_0, m, y)\, m(dy) \\ \qquad + \int_{\mathbb{R}^d} D_m U^0(t, x_0, m, y) \cdot H_p(x_0, y, D_x U(t, x_0, y, m), m)\, m(dy) = 0, \\ -\partial_t U - \Delta_x U + H(x_0, x, D_x U, m) - \int_{\mathbb{R}^d} \mathrm{div}_y\, D_m U(t, x_0, x, m, y)\, m(dy) \\ \qquad + \int_{\mathbb{R}^d} D_m U(t, x_0, x, m, y) \cdot H_p(x_0, y, D_x U(t, x_0, y, m), m)\, m(dy) = 0. \end{cases} \tag{8}$$

It turns out that this system can be handled by the method of characteristics. As for the second one, it is a simple system of HJ equations (for fixed x, m):

$$\begin{cases} -\partial_t U^0 - \Delta_{x_0} U^0 + H^0(x_0, D_{x_0} U^0, m) = 0, \\ -\partial_t U - \Delta_{x_0} U + D_{x_0} U \cdot H^0_p(x_0, D_{x_0} U^0(t, x_0, m), m) = 0. \end{cases} \tag{9}$$

The idea of splitting time is not completely new in the framework of mean field games. Let us quote for instance the paper in preparation [1] in which the authors use a splitting technique similar to the one described above to compute numerically the solution of MFGs with a major player. The construction, given in [13], of (weak) equilibria for MFG problems with common noise also relies on a time splitting. The main difference is that it is done at the level of the MFG equilibrium, while we do the construction at the (stronger) level of the master equation. One consequence is that, with our approach, the construction of a solution to the stochastic MFG system (in short time, though) is straightforward once the solution of the master equation is built, while deriving a solution of the master equation from the stochastic MFG system is much trickier. Our method is particularly relevant for the problem with a major player: indeed, for this problem, the associated MFG system involves two *backward* stochastic HJ equations, a stochastic Kolmogorov equation and a McKean–Vlasov equation; the construction of a solution to the system of master equations (5), based directly on this MFG system, would therefore be extremely technical. Instead, our method relies on the one hand on the analysis of system (8) (which derives directly from the analysis of the standard first order master equation) and on the other hand on estimates for system (9) (which is just an ordinary system of HJ equations).

Let us finally point out that, in this paper, we do not address at all the problem of the existence of a solution on a large time interval. For the first and second order master equation, this question is related to the Lasry–Lions monotonicity condition [24, 25]. The existence of a solution on a large time interval can be obtained under this condition either by the Hilbertian approach, as explained in [28], or by a continuation method, as in [11] and [12], or even directly by using the long time existence of a solution for the MFG system, as in [10]. Let us recall that when the monotonicity condition is not fulfilled, the solution to the second order master equation is expected to develop shocks (i.e., discontinuities) in finite time. Note also that a structure condition similar to the monotonicity condition is not known for MFGs with a major player.

The paper is organized in the following way. In Section 2 we fix the notation and we recall the definition of derivatives in the space of measures; then we introduce our assumptions and we state the main results of this article. We also present, at the end of Section 2, the general idea of splitting method that we adopt for both systems (1) and (5). In order to prove the existence results, our strategy is put in practice in Section 3 for the second order master equation (equation (1)) and in Section 4 for the system of master equations for MFG with a major player (system (5)), respectively. Both sections require several estimates on first order master equations, which are collected in Section 5. As first order master equations are built by the method of characteristics involving the solutions of classical MFG systems (2), Section 5.1 first provides estimates for these systems. Then Section 5.2 is devoted to the analysis of first order master equations. We complete the paper by appendices in which we prove short-time estimates for the standard Hamilton–Jacobi equations (Appendix A) and we discuss several facts on maps defined on the space of measures (differentiability, interpolation and the Ascoli theorem, Appendix B).

2. Notation, assumptions and main results

2.1. Notation

Throughout the paper, we work in the euclidean space $\mathbb{R}^d$ (with $d \in \mathbb{N}$, $d \geq 1$), endowed with the scalar product $(x, y) \mapsto x \cdot y$ and the distance $|\cdot|$. Given $T > 0$ and a map $\phi : (0, T) \times \mathbb{R}^d \to \mathbb{R}$, we denote by $\partial_t \phi$ the derivative of ϕ with respect to the time variable, by $\partial_{x_i} \phi$ its partial derivative with respect to the i-th space variable ($i = 1, \ldots, d$) and by $D\phi$ the gradient with respect to the space variable.

For $n \in \mathbb{N}$, we denote by C_b^n the set of maps $\phi : \mathbb{R}^d \to \mathbb{R}$ which are n-times differentiable with continuous and bounded derivatives; in particular, C_b^0 is the set of continuous and bounded maps. Given $\phi \in C_b^n$ and a multi-index $k = (k_1, \ldots, k_d) \in \mathbb{N}^d$, with length $|k| := \sum_{i=1}^d k_i \leq n$, we denote by $\partial^k \phi = \frac{\partial^{k_1}}{\partial x_1^{k_1}} \ldots \frac{\partial^{k_d}}{\partial x_d^{k_d}} \phi$ (or briefly ϕ_k) the k-th derivative of ϕ. We also denote by $D^n \phi$ ($n \in \mathbb{N}$, $n \geq 1$) the vector $(\partial^k \phi)_{|k|=n}$. The norm of ϕ in C_b^n is

$$\|\phi\|_n := \sum_{r=0}^n \sup_x \Big(\sum_{|\alpha|=r} |\partial^\alpha \phi(x)|^2 \Big)^{1/2} = \sum_{r=0}^n \|D^r \phi\|_\infty.$$

For $n = 0$, we use interchangeably the notation $\|\phi\|_0$ or $\|\phi\|_\infty$.

For $(n_1, \ldots, n_k) \in \mathbb{N}^k$ ($k \in \mathbb{N}$, $k \geq 2$), we denote by $C_b^{n_1,\ldots,n_k}$ the space of functions $\phi : \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_k} \to \mathbb{R}$ ($d_i \geq 1$) having continuous and bounded derivatives

$D^{l_1}_{x_1}\cdots D^{l_k}_{x_k}\phi$ for all $l_1\le n_1,\ldots,l_k\le n_k$, endowed with the norm

$$\|\phi\|_{n_1,\ldots,n_k}=\|\phi(\cdot_{x_1},\ldots,\cdot_{x_k})\|_{n_1,\ldots,n_k}:=\sum_{l_1\le n_1,\ldots,l_k\le n_k}\|D^{l_1}_{x_1}\cdots D^{l_k}_{x_k}\phi\|_\infty,$$

where now $(x_1,\ldots,x_k)$ stands for a generic element of $\mathbb{R}^{d_1}\times\mathbb{R}^{d_k}$.

We denote by C^{-n} the dual space of C^n_b, endowed with the usual norm

$$\|\rho\|_{-n}:=\sup_{\|\phi\|_n\le 1}\rho(\phi)\quad\forall\rho\in C^{-n}.$$

Finally, when a map $\phi=\phi(t,x)$ depends also on time t belonging to an interval I, we often write $\sup_{t\in I}\|\phi(t)\|_n$ for $\sup_{t\in I}\|\phi(t,\cdot)\|_n$. We use a corresponding notation for a map $\rho\in C^0([0,T],C^{-k})$.

Throughout the paper, $\mathcal{P}$ stands for the set of Borel probability measures on $\mathbb{R}^d$ and for $k\ge 1$, $\mathcal{P}_k$ stands for the set of measures in $\mathcal{P}$ with finite moment of order k: namely,

$$M_k(m):=\left(\int_{\mathbb{R}^d}|x|^k\,m(dx)\right)^{1/k}<+\infty\quad\text{if } m\in\mathcal{P}_k.$$

The set $\mathcal{P}_k$ is endowed with the distance (see for instance [4,30,31])

$$\mathbf{d}_k(m,m')=\inf_\pi\left(\int_{\mathbb{R}^d}|x-y|^k\,\pi(dx,dy)\right)^{1/k}\qquad\forall m,m'\in\mathcal{P}_k,$$

where the infimum is taken over the couplings π between m and m', i.e., over the Borel probability measures π on $\mathbb{R}^d\times\mathbb{R}^d$ with first marginal m and second marginal m'. Note that $\mathcal{P}_2\subset\mathcal{P}_1$ and $\mathbf{d}_1\le\mathbf{d}_2$ by the Cauchy–Schwarz inequality. We will often use the fact that if $\phi:\mathbb{R}^d\to\mathbb{R}$ is Lipschitz continuous with a Lipschitz constant $L\ge 0$, then

$$\left|\int_{\mathbb{R}^d}\phi(x)\,(m-m')(dx)\right|\le L\mathbf{d}_1(m,m')\quad\forall m,m'\in\mathcal{P}_1.$$

Moreover, $\mathbf{d}_1(m,m')$ is the smallest constant for which the above inequality holds for any L-Lipschitz continuous map ϕ (see for instance [30,31]). Given $m\in\mathcal{P}$ and $\phi\in C^0_b$, the image $\phi_\sharp m$ of m by ϕ is the element of $\mathcal{P}$ defined by

$$\int_{\mathbb{R}^d}f(x)\,\phi_\sharp m(dx)=\int_{\mathbb{R}^d}f(\phi(x))\,m(dx)\quad\forall f\in C^0_b.$$

2.2. *Derivatives in the space of measures*

We now define the derivative in the space $\mathcal{P}_2$. For this, we mostly follow the definition and notations introduced in [10] (in a slightly different context) and which are reminiscent of earlier approaches: see [3,4] and the references in [12]. We say that a map $U:\mathcal{P}_2\to\mathbb{R}$ is C^1 if there exists a *continuous and bounded* map $\frac{\delta U}{\delta m}:\mathcal{P}_2\times\mathbb{R}^d\to\mathbb{R}$ such that

$$U(m')-U(m)=\int_0^1\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}((1-s)m+sm',y)\,(m'-m)(dy)\,ds\quad\forall m,m'\in\mathcal{P}_2. \tag{10}$$

Note that the restriction on $\frac{\delta U}{\delta m}$ to be continuous on the entire space $\mathbb{R}^d$ and globally bounded is restrictive; it will however simplify our forthcoming construction. The map $\frac{\delta U}{\delta m}$ is defined only up to an additive constant that we fix with the convention

$$\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m, y)\, m(dy) = 0 \quad \forall m \in \mathcal{P}_2. \tag{11}$$

We say that the map U is *continuously L-differentiable* (for short, L-C^1) if U is C^1 and $y \mapsto \frac{\delta U}{\delta m}(m, y)$ is everywhere differentiable with a continuous and globally bounded derivative on $\mathcal{P}_2 \times \mathbb{R}^d$. We denote by

$$D_m U(m, y) := D_y \frac{\delta U}{\delta m}(m, y)$$

this L-derivative. In view of the discussion in [10], $D_m U$ coincides with the Lions derivative as introduced in [27] and discussed in [12]. In particular, it estimates the Lipschitz regularity of U in $\mathcal{P}_2$ [12, Remark 5.27]:

$$|U(m) - U(m')| \le \mathbf{d}_2(m, m') \sup_{\mu \in \mathcal{P}_2} \left(\int_{\mathbb{R}^d} |D_m U(\mu, y)|^2 \, \mu(dy) \right)^{1/2} \quad \forall m, m' \in \mathcal{P}_2. \tag{12}$$

Of course one can also estimate the Lipschitz regularity of U through the $\mathbf{d}_1$ norm, as

$$\begin{aligned} |U(m) - U(m')| &\le \mathbf{d}_1(m, m') \sup_{\mu \in \mathcal{P}_2} \|D_m U(\mu, \cdot)\|_\infty \\ &\le \mathbf{d}_2(m, m') \sup_{\mu \in \mathcal{P}_2} \|D_m U(\mu, \cdot)\|_\infty. \end{aligned} \tag{13}$$

Note that, with our boundedness convention, if U is continuously L-differentiable, then it is automatically globally Lipschitz continuous.

When U is smooth enough, we often see the map $\frac{\delta U}{\delta m}$ as a linear map on C^{-k} by

$$\frac{\delta U}{\delta m}(m)(\rho) = \left\langle \rho, \frac{\delta U}{\delta m}(m, \cdot) \right\rangle_{C^{-k}, C^k} \qquad \forall \rho \in C^{-k}.$$

We say that U is C^2 if $\frac{\delta U}{\delta m}$ is C^1 in m with a continuous and bounded derivative, that is, $\frac{\delta^2 U}{\delta m^2} = \frac{\delta}{\delta m}(\frac{\delta U}{\delta m}) : \mathcal{P}_2 \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is continuous in all variables and bounded. We say that U is twice L-differentiable if the map $D_m U$ is L-differentiable with respect to m with a second order derivative $D^2_{mm} U = D^2_{mm} U(m, y, y')$ which is continuous and bounded on $\mathcal{P}_2 \times \mathbb{R}^d \times \mathbb{R}^d$ with values in $\mathbb{R}^{d \times d}$.

When a map $U : \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$ is of class C^n_b with respect to the space variable, uniformly with respect to the measure variable, we often set

$$|||U|||_n := \sup_{m \in \mathcal{P}_2} \|U(\cdot, m)\|_n. \tag{14}$$

Note here the use of the different symbol $||| \cdot |||$. We use similar notation for a map U depending on several space variables and on a measure:

$$|||U|||_{n_1, \dots, n_k} := \sup_{m \in \mathcal{P}_2} \|U(\cdot, m)\|_{n_1, \dots, n_k}.$$

When a map $U : \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$ is Lipschitz continuous with respect to m, uniformly with respect to the space variable in some C^n norm, we define $\mathrm{Lip}_n(U)$ as the smallest constant C such that

$$\|U(\cdot, m_1) - U(\cdot, m_2)\|_n \le C\mathbf{d}_2(m_1, m_2) \quad \forall m, m' \in \mathcal{P}_2.$$

That is,

$$\mathrm{Lip}_n(U) := \sup_{m_1 \neq m_2} \frac{\|U(\cdot, m_1) - U(\cdot, m_2)\|_n}{\mathbf{d}_2(m_1, m_2)}.$$

More generally, if $U : (\mathbb{R}^d)^k \times \mathcal{P}_2 \to \mathbb{R}$ (for $k \in \mathbb{N}$, $k \ge 1$) is Lipschitz continuous in the measure variable in some $C_b^{n_1,\dots,n_k}$ norm (where $n_i \in \mathbb{N}$ for $i = 1, \dots, k$), then we set

$$\mathrm{Lip}_{n_1,\dots,n_k}(U) := \sup_{m_1 \neq m_2} \frac{\|U(\cdot_{x_1}, \dots, \cdot_{x_k}, m_1) - U(\cdot_{x_1}, \dots, \cdot_{x_k}, m_2)\|_{n_1,\dots,n_k}}{\mathbf{d}_2(m_1, m_2)}.$$

We will typically use this notation for the derivatives of a map $U : \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$; indeed, we will often have to estimate quantities of the form

$$\mathrm{Lip}_{n_1,n_2}(D_m U) := \sup_{m_1 \neq m_2} \frac{\|D_m U(\cdot_x, m_1, \cdot_y) - D_m U(\cdot_x, m_2, \cdot_y)\|_{n_1,n_2}}{\mathbf{d}_2(m_1, m_2)}$$

and

$$\mathrm{Lip}_{n_1,n_2,n_3}(D^2_{mm} U) := \sup_{m_1 \neq m_2} \frac{\|D^2_{mm} U(\cdot_x, m_1, \cdot_y, \cdot_{y'}) - D^2_{mm} U(\cdot_x, m_2, \cdot_y, \cdot_{y'})\|_{n_1,n_2,n_3}}{\mathbf{d}_2(m_1, m_2)}.$$

Concerning the Lipschitz continuity with respect to one of the entries x_i, we will use the following notation:

$$\begin{aligned}
&\mathrm{Lip}^{x_i}_{n_1,\dots,n_{i-1},n_{i+1},\dots,n_k}(U) \\
&:= \sup_{m, x_i^1 \neq x_i^2} \frac{\begin{array}{l}\|U(\cdot_{x_1}, \dots, \cdot_{x_{i-1}}, x_i^1, \cdot_{x_{i+1}}, \dots, \cdot_{x_k}, m) \\ \quad - U(\cdot_{x_1}, \dots, \cdot_{x_{i-1}}, x_i^2, \cdot_{x_{i+1}}, \dots, \cdot_{x_k}, m)\|_{n_1,\dots,n_{i-1},n_{i+1},\dots,n_k}\end{array}}{|x_i^1 - x_i^2|}.
\end{aligned}$$

Further norms. In order to estimate the y-dependence of a derivative with respect to the measure of a map $U = U(x, m)$, we systematically proceed by duality method, testing this derivative against distributions. This leads to the following norms, for $n, k \in \mathbb{N}$:

$$\begin{aligned}
\left\| \frac{\delta U}{\delta m} \right\|_{n;k} &:= \sup_{m \in \mathcal{P}_2} \sum_{r=0}^{n} \sup_{\substack{x \in \mathbb{R}^d,\, \rho \in C_c^0 \\ \|\rho\|_{-k} = 1}} \left(\sum_{|\alpha| = r} \left| \partial_x^\alpha \frac{\delta U}{\delta m}(x, m)(\rho) \right|^2 \right)^{1/2} \\
&= \sup_{m \in \mathcal{P}_2} \sum_{r=0}^{n} \sup_{\substack{x \in \mathbb{R}^d,\, \rho \in C_c^0, \\ \|\rho\|_{-k} = 1}} \left| D_x^r \frac{\delta U}{\delta m}(x, m)(\rho) \right|, \\
\left\| \frac{\delta^2 U}{\delta m^2} \right\|_{n;k,k'} &:= \sup_{m \in \mathcal{P}_2} \sum_{r=1}^{n} \sup_{\substack{x \in \mathbb{R}^d,\, \rho, \rho' \in C_c^0, \\ \|\rho\|_{-k} = \|\rho'\|_{-k'} = 1}} \left| D_x^r \frac{\delta^2 U}{\delta m^2}(x, m)(\rho, \rho') \right|.
\end{aligned}$$

We point out the subtle difference in notation between, say, $|||\frac{\delta U}{\delta m}|||_{n,k}$ (which involves a supremum over y) and $\|\frac{\delta U}{\delta m}\|_{n;k}$ (in which the dependence is estimated by duality). The same difference holds between $|||\frac{\delta^2 U}{\delta m^2}|||_{n,k,k}$ and $\|\frac{\delta^2 U}{\delta m^2}\|_{n;k,k}$.

For maps $U = U(x_1, x_2, m)$ depending on two (or more) space variables, we use the transparent notation $\|\cdot\|_{n_1,n_2;k}$ (and, if $n_1 = 0$ (say), we simply set $\|\cdot\|_{n_2,k} = \|\cdot\|_{0,n_2;k}$). Finally, we use similar notation for the Lipschitz norms, setting, for instance for a map $U = U(x, m)$,

$$\begin{aligned}\mathrm{Lip}_{n;k,k'}\left(\frac{\delta^2 U}{\delta m^2}\right) := \sup_{m_1\neq m_2} \mathbf{d}_2(m_1, m_2)^{-1} \\ \times \sum_{r=0}^{n} \sup_{\substack{x\in\mathbb{R}^d,\, \rho,\rho'\in C_c^0 \\ \|\rho\|_{-k}=\|\rho'\|_{-k'}=1}} \left| D_x^r \frac{\delta^2 U}{\delta m^2}(x, m_2)(\rho, \rho') - D_x^r \frac{\delta^2 U}{\delta m^2}(x, m_1)(\rho, \rho')\right|.\end{aligned}$$

Some comments about the norms we have just introduced are now in order. We discuss the norm $\|\cdot\|_{n;k}$ to fix ideas. With these notations, if $U = U(x, m)$ is smooth enough, we have

$$\left\|\frac{\delta U}{\delta m}(\cdot, m)(\rho)\right\|_n \le \left\|\frac{\delta U}{\delta m}\right\|_{n;k} \|\rho\|_{-k}$$

for every fixed $m \in \mathcal{P}_2$. Inequalities of this type are used throughout the text. Next we note that the norms $|||\cdot|||_{n,k}$ and $\|\cdot\|_{n;k}$ are equivalent if we know a priori that $\frac{\delta U}{\delta m} = \frac{\delta U}{\delta m}(x, m, y)$ is in $C_b^{n,k}$. In general we do not have this information, but only know that $\frac{\delta U}{\delta m}$ is (at least) continuous. In this case, we use the following result:

Lemma 2.1. *Let $k \in \mathbb{N}$ with $k \ge 1$ and $u \in C^0$ be such that*

$$\theta := \sup_{\rho\in C_c^0,\ \|\rho\|_{-k}=1} \int_{\mathbb{R}^d} u(y)\rho(y)\, dy < +\infty. \tag{15}$$

Then $u \in C_b^{k-1}$ with $\|u\|_{k-1} \le C_k\theta$ (where C_k depends on d and k) and, for any $\beta \in \mathbb{N}^d$ with $|\beta| = k-1$, $\partial^\beta u$ is θ-Lipschitz continuous.

Remark 2.2. In particular, if $\frac{\delta U}{\delta m} \in C_b^{n,0}$ and $\|\frac{\delta U}{\delta m}\|_{n;k}$ is finite for some $n, k \in \mathbb{N}$ with $k \ge 1$, then $\frac{\delta U}{\delta m} \in C_b^{n,k-1}$ and

$$\left|\left|\left|\frac{\delta U}{\delta m}\right|\right|\right|_{n,k-1} \le C_{n,k}\left\|\frac{\delta U}{\delta m}\right\|_{n;k}$$

for some constant $C_{n,k}$ depending in addition on dimension only. Moreover, the derivatives of the form $\partial_x^\alpha \partial_y^\beta \frac{\delta U}{\delta m}$ for $|\alpha| \le n$ and $\beta \le k-1$ are Lipschitz continuous with respect to y and thus—by (13)—also with respect to m, with a Lipschitz constant bounded by $\|\frac{\delta U}{\delta m}\|_{n;k}$.

Proof of Lemma 2.1. For $k = 1$, approximating Dirac masses by continuous maps with compact support, for any $x, y \in \mathbb{R}^d$ we have

$$|u(x)| \le \theta\|\delta_x\|_{-1} = \theta \quad \text{and} \quad |u(x) - u(y)| \le \theta\|\delta_x - \delta_y\|_{-1} = \theta|x - y|.$$

This proves the claim for $k = 1$. Let now assume that (15) holds for $k = 2$. Then u can be extended to an element T in $(C^{-2})'$ with norm $\|T\| \leq \theta$ such that $T(\rho) = \int u \, d\rho$ for any Radon measure ρ. As, for any $v \in \mathbb{R}^d$,

$$\lim_{h\to 0,\, v'\to v} h^{-1}(\delta_{x+hv'} - \delta_x) = -\partial_v \delta_x \quad \text{in } C^{-2},$$

we infer that

$$\lim_{h\to 0,\, v'\to v} h^{-1}(u(x+hv') - u(x)) = -T(\partial_v \delta_x).$$

The map $(x, v) \mapsto \partial_v \delta_x$ being continuous in C^{-2} with $\|\partial_v \delta_x\|_{-2} \leq |v|$, u is in C^1 with $\|Du\| \leq \theta$. Then, arguing as for $k = 1$, one can easily check that Du is θ-Lipschitz continuous. So the result also holds for $k = 2$. The proof can be completed in the same way for any k by induction. ■

Finally, note that the norm $|||\cdot|||_{n,k}$ will only be used to state the assumptions on the data $H, H^0, \ldots$, as it is more standard. On the other hand, the "equivalent" norm $\|\cdot\|_{n;k}$, being the natural one for the methods used to get the estimates, will be extensively used throughout the paper.

2.3. Assumptions on the data

We state here the assumptions needed on a, H and G for the existence of a classical solution to the second order master equation (1) and to the master equation (5) for the MFG problem with a major player. These assumptions are in force throughout the paper. Note that they are common to both problems (1) and (5) since both require the same kind of estimates on the first order master equation (see Section 5.2).

We assume that the map $a : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ can be written as $a = \sigma\sigma^T$ where $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ ($M \in \mathbb{N}$, $M \geq 1$) is bounded in C^n_b with respect to the space variable, uniformly with respect to the time variable, for some $n \geq 4$. We also assume that the following uniform ellipticity condition holds:

$$a(t, x) \geq C_0^{-1} I_d, \quad \|Da\|_\infty \leq C_0, \tag{16}$$

for some $C_0 > 0$, where I_d is the $d \times d$ identity matrix.

We assume that the map $H : \mathbb{R}^{d_0} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$ satisfies the growth condition

$$\sup_{x_0\in\mathbb{R}^{d_0},\, x\in\mathbb{R}^d,\, m\in\mathcal{P}_2} |D_x H(x_0, x, p, m)| \leq C_0(1 + |p|^\gamma) \quad \forall p \in \mathbb{R}^d, \tag{17}$$

for some $\gamma > 1$. We also suppose that, for any $R > 0$, the quantities

$$|||H(\cdot_{x_0}, \cdot_x, \cdot_p, m)|||_{3,n,n+1}, \quad \left|\left|\left| \frac{\delta H}{\delta m}(\cdot_{x_0}, \cdot_x, \cdot_p, m, \cdot_y) \right|\right|\right|_{2,n-1,n,k},$$

$$\left|\left|\left| \frac{\delta^2 H}{\delta m^2}(\cdot_{x_0}, \cdot_x, \cdot_p, m, \cdot_y, \cdot_{y'}) \right|\right|\right|_{1,n-2,n-1,k-1,k-1},$$

and $\mathrm{Lip}_{1,n-3,n-2,k-1,k-1}(\frac{\delta^2 H}{\delta m^2})$ are bounded for $|p| \leq R$, $m \in \mathcal{P}_2$ and $x_0 \in \mathbb{R}^{d_0}$, for any

$k \in \{2, \dots, n-1\}$. Note that we could also allow for a time dependence for H without changing at all the arguments; we will not do so to simplify the notation a little. For the second order master equation, the Hamiltonian H actually does not depend on x_0, but this dependence is important to handle the MFG problem with a major player.

As for the initial condition $G : \mathbb{R}^{d_0} \times \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$, we assume that G is of class C^2 with respect to all variables and that the quantities

$$\|\!|G(\cdot_{x_0}, \cdot_x, m)|\!\|_{3,n}, \quad \left|\!\left|\!\left| \frac{\delta G}{\delta m}(\cdot_{x_0}, \cdot_x, m, \cdot_y) \right|\!\right|\!\right|_{2,n-1,k},$$
$$\left|\!\left|\!\left| \frac{\delta^2 G}{\delta m^2}(\cdot_{x_0}, \cdot_x, m, \cdot_y, \cdot_{y'}) \right|\!\right|\!\right|_{1,n-2,k-1,k-1}, \quad \mathrm{Lip}_{1,n-3,k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2}\right)(\cdot_{x_0}, \cdot_x, m, \cdot_y, \cdot_{y'}),$$

are bounded uniformly with respect to $m \in \mathcal{P}_2$. Here again, for the second order master equation, the terminal condition G does not depend on x_0, but this dependence is needed in the MFG problem with a major player.

Additional assumptions for the MFG problem with a major player. This problem involves in addition a Hamiltonian $H^0 : \mathbb{R}^{d_0} \times \mathbb{R}^{d_0} \times \mathcal{P}_2 \to \mathbb{R}$ and a terminal condition $G^0 : \mathbb{R}^{d_0} \times \mathcal{P}_2 \to \mathbb{R}$. We assume that the map H^0 satisfies the growth property

$$\sup_{x_0 \in \mathbb{R}^{d_0}, m \in \mathcal{P}_2} \left(|D_{x_0,p} H^0(x_0, p, m)| + |D^2_{x_0,p} H^0(x_0, p, m)|\right) \le C_0(|p|^\gamma + 1) \tag{18}$$

for some $\gamma > 1$. We also suppose that, for any $R > 0$, the quantities

$$\|\!|H^0(\cdot_{x_0}, \cdot_p, m)|\!\|_{3,4}, \quad \left|\!\left|\!\left| \frac{\delta H^0}{\delta m}(\cdot_{x_0}, \cdot_p, m, \cdot_y) \right|\!\right|\!\right|_{2,3,k},$$
$$\left|\!\left|\!\left| \frac{\delta^2 H^0}{\delta m^2}(\cdot_{x_0}, \cdot_p, m, \cdot_y, \cdot_{y'}) \right|\!\right|\!\right|_{1,2,k-1,k-1},$$

and $\mathrm{Lip}_{0,1,k-2,k-2}(\frac{\delta^2 H}{\delta m^2})$ are bounded for $|p| \le R$, $m \in \mathcal{P}_2$ and $x_0 \in \mathbb{R}^{d_0}$, for any $k \in \{2, \dots, n-1\}$.

The initial condition $G^0 : \mathbb{R}^{d_0} \times \mathcal{P}_2 \to \mathbb{R}$ is assumed to be of class C^2 with respect to the measure variable, and the quantities

$$\|\!|G^0(\cdot, m)|\!\|_3, \quad \left|\!\left|\!\left| \frac{\delta G^0}{\delta m}(\cdot, m, \cdot) \right|\!\right|\!\right|_{2,k}, \quad \left|\!\left|\!\left| \frac{\delta^2 G^0}{\delta m^2}(\cdot_{x_0}, m, \cdot_y, \cdot_{y'}) \right|\!\right|\!\right|_{1,k-1,k-1},$$
$$\mathrm{Lip}_{0,k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}\right)(\cdot_{x_0}, m, \cdot_y, \cdot_{y'}),$$

are supposed to be bounded uniformly with respect to $m \in \mathcal{P}_2$.

Throughout the proofs, we assume that the time horizon T is small, say $T \le 1$. We denote by C and C_M constants which might change from line to line and which depend only on the data of the problem, i.e., on a, H and H^0—the dependence on G and G^0 being always explicitly written—and, for C_M, on the additional real number M. In some proofs, when there is no ambiguity, we drop the M dependence of C_M to simplify the expressions.

2.4. Main results

In this section we state the two main results on the short-time existence and uniqueness of the second order master equation and the master equation with a major player. We also state, as a corollary, the existence of solutions to the stochastic MFG system.

Let us start with the second order master equation, which reads as follows:

$$\begin{cases} -\partial_t U(t,x,m) - \mathrm{Tr}((a(t,x)+a^0)D^2_{xx}U(t,x,m)) + H(x, D_xU(t,x,m), m) \\ \quad - \displaystyle\int_{\mathbb{R}^d} \mathrm{Tr}((a(t,y)+a^0)D^2_{ym}U(t,x,m,y))\, m(dy) \\ \quad + \displaystyle\int_{\mathbb{R}^d} D_mU(t,x,m,y)\cdot H_p(y, D_xU(t,y,m), m)\, m(dy) \\ \quad - 2\displaystyle\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{xm}U(t,x,m,y)]\, m(dy) \\ \quad - \displaystyle\int_{\mathbb{R}^{2d}} \mathrm{Tr}[a^0 D^2_{mm}U(t,x,m,y,y')]\, m(dy)\, m(dy') = 0 \quad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2, \\ U(T,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2, \end{cases} \tag{19}$$

where a^0 is a symmetric positive definite $d\times d$ matrix (independent of time and space). We say that $U : [0,T]\times\mathbb{R}^d\times\mathcal{P}_2 \to \mathbb{R}$ is a *classical solution* of (19) if U and its derivatives involved in (19) exist, are continuous in all variables and are bounded, and if (19) holds.

Our first main result is the following short time existence theorem:

Theorem 2.3. *Under the assumptions of Section 2.3, there exists a time $T>0$ such that the second order master equation (19) has a unique classical solution U on $[0,T]$.*

The proof of Theorem 2.3 is given at the end of Section 3.2, after some preliminary steps. We shall not prove here the uniqueness of the solution to (19), which holds under our assumptions; this point has often been discussed in the literature (see [10, 12] for instance). The reader may notice that we cannot handle a second order master equation with a space dependent matrix $a^0 = a^0(t,x)$. The reason is that we do not know how to extend the estimate in Proposition 3.1 to the space dependent case.

An easy consequence of the existence of a solution to the master equation is the well-posedness of the stochastic MFG system

$$\begin{cases} du(t,x) = [-\mathrm{Tr}((a+a^0)(t,x)D^2u(t,x)) + H(x, Du(t,x), m(t)) \\ \qquad\qquad -\sqrt{2}\,\mathrm{Tr}(\sigma^0 Dv(t,x))]\,dt + v(t,x)\cdot dW_t \quad \text{in } (0,T)\times\mathbb{R}^d, \\ dm(t,x) = \Big[\displaystyle\sum_{i,j} D_{ij}(((a_{ij}) + a^0_{ij})(t,x)m(t,x)) \\ \qquad\qquad + \mathrm{div}(m(t,x)H_p(x, Du(t,x), m(t)))\Big]\,dt \\ \qquad\qquad - \mathrm{div}(m(t,x)\sqrt{2}\,\sigma^0 dW_t) \quad \text{in } (0,T)\times\mathbb{R}^d, \\ u(T,x) = G(x, m(T)), \quad m(0) = m_0 \quad \text{in } \mathbb{R}^d. \end{cases} \tag{20}$$

We say that (u,m,v) is a *classical solution* to (20) if u, m and v are random with values in $C^0([0,T], C^2_b)$, $C^0([0,T], \mathcal{P}_2)$ and $C^0([0,T], C^1_b(\mathbb{R}^d, \mathbb{R}^d))$ respectively and adapted to the filtration generated by W and if the backward HJ equation is satisfied in the classical sense:

$$u(t,x) = G(x,m(T)) - \int_t^T \big(-\mathrm{Tr}((a+a^0)(s,x)D^2u(s,x)) + H(x,Du(s,x),m(s)) - \sqrt{2}\,\mathrm{Tr}(\sigma^0 Dv(s,x))\big)\,ds - \int_t^T v(s,x)\cdot dW_s$$

while the Fokker–Planck equation is satisfied in the sense of distributions: for any $\phi \in C_c^\infty([0,T)\times\mathbb{R}^d)$,

$$\begin{aligned}0 &= \int_{\mathbb{R}^d} \phi(0,x)\,m_0(dx)\\ &+\int_0^T\int_{\mathbb{R}^d} \big(\mathrm{Tr}((a+a^0)(s,x)D^2\phi(s,x)) - D\phi(s,x)\cdot H_p(x,Du(s,x),m(s))\big)\,m(s,dx)\,ds\\ &+\sqrt{2}\int_0^T\int_{\mathbb{R}^d}(\sigma^0)^T D\phi(s,x)\,m(s,dx)\cdot dW_s.\end{aligned}$$

Theorem 2.4. *Under the assumptions of Theorem 2.3, there exists a time $T>0$ for which the stochastic MFG system (20) has a classical solution (u,m,v) in $[0,T]$. Moreover,*

$$v(t,x) = \sqrt{2}\int_{\mathbb{R}^d} (\sigma^0)^T D_m U(t,x,m(t),y)\,m(t,dy), \tag{21}$$

where U is the solution to the second order master equation (19).

The proof of Theorem 2.4 is given in Section 3.3.

Then, we investigate the well-posedness of the master equation associated with the MFG problem with a major player. Here the unknown (U^0,U) solves the system of master equations

$$\begin{cases} -\partial_t U^0(t,x_0,m) - \Delta_{x_0}U^0(t,x_0,m) + H^0(x_0,D_{x_0}U^0(t,x_0,m),m) \\ \qquad - \int_{\mathbb{R}^d} \mathrm{div}_y\, D_mU^0(t,x_0,m,y)\,m(dy) \\ \qquad + \int_{\mathbb{R}^d} D_mU^0(t,x_0,m,y)\cdot H_p(x_0,y,D_xU(t,x_0,y,m),m)\,m(dy) = 0 \\ \qquad\qquad \text{in } (0,T)\times\mathbb{R}^{d_0}\times\mathcal{P}_2, \\ -\partial_t U(t,x_0,x,m) - \Delta_x U(t,x_0,x,m) - \Delta_{x_0}U(t,x_0,x,m) \\ \quad + H(x_0,x,D_xU(t,x_0,x,m),m) \\ \quad - \int_{\mathbb{R}^d} \mathrm{div}_y\, D_mU(t,x_0,x,m,y)\,m(dy) + D_{x_0}U\cdot H^0_p(x_0,D_{x_0}U^0(t,x_0,m),m) \\ \quad + \int_{\mathbb{R}^d} D_mU(t,x_0,x,m,y)\cdot H_p(x_0,y,D_xU(t,x_0,y,m),m)\,m(dy) = 0 \\ \qquad\qquad \text{in } (0,T)\times\mathbb{R}^{d_0}\times\mathbb{R}^d\times\mathcal{P}_2, \\ U^0(T,x_0,m) = G^0(x_0,m) \quad \text{in } \mathbb{R}^{d_0}\times\mathcal{P}_2, \\ U(T,x_0,x,m) = G(x_0,x,m) \quad \text{in } \mathbb{R}^{d_0}\times\mathbb{R}^d\times\mathcal{P}_2. \end{cases} \tag{22}$$

Let $U^0 : [0,T] \times \mathbb{R}^{d_0} \times \mathcal{P}_2 \to \mathbb{R}$ and $U : [0,T] \times \mathbb{R}^{d_0} \times \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$. We say that (U^0, U) is a *classical solution* of (22) if U^0 and U and their derivatives involved in (22) exist, are continuous in all variables and are bounded, and if (22) holds. Our main result is the following:

Theorem 2.5. *Under the assumptions of Section 2.3, there exists a time $T > 0$ and a classical solution (U^0, U) to (22) on the time interval $[0,T]$ such that $D_{x_0}U^0$ and $D_{x_0,x}U$ are uniformly Lipschitz continuous in the space and measure variables.*

The proof of Theorem 2.5 is given in Section 4. The result can be easily extended to non-constant diffusions. We work here with a constant diffusion to simplify the notation.

The constructions of solutions to the two master equations share a common strategy. The key idea is to use a Trotter–Kato scheme alternating two simpler evolutive problems on vanishing time intervals. This is commonly referred to as a splitting method; according to this approach, the solution u of the evolution equation

$$u_t = Au + Bu$$

can be built by alternating, in smaller and smaller time-steps, the evolution driven by A and the evolution driven by B, respectively. Indeed, if A and B were generating semigroups e^{tA}, e^{tB} acting on a common Banach space X, then the Trotter–Kato product formula implies

$$e^{t(A+B)}U = \lim_{n\to\infty} (e^{\frac{t}{n}A} e^{\frac{t}{n}B})^n U.$$

Notice that, for this formula to hold (i.e. for this scheme to be convergent), it is crucial to have estimates of the form

$$\|e^{\tau A}U\|_X \le (1 + c\tau)\|U\|_X, \quad \|e^{\tau B}U\|_X \le (1 + c\tau)\|U\|_X,$$

which yield in the limit $\|e^{t(A+B)}U\|_X \le e^{2ct}\|U\|_X$. One may even allow c in the above estimate to depend on $\|U\|_X$ itself; if so, one has convergence of the scheme for short time t only, which will be the case in our settings.

The idea of using a splitting method needs to be carefully rephrased in our context. The main point is to choose suitable pairs (A, B) in order to decompose our master equations into simpler and efficient problems. In our settings, the second order master equation will be obtained as the superposition of the first order master equation (6) and a linear second order master equation (7). The system of master equations with a major player will be seen as the superposition of a system of first order master equations (8) (where the major player is “frozen”), and a system of HJ equations (9) (where the population of minor players is “frozen”).

We do not need to prove that the two separate problems driven by A, B actually generate semigroups. However, we need to identify a suitable norm, or some other meaningful quantity, which is not deteriorated more than linearly in time by *both* the alternating problems. This is the main technical issue in our proofs. Indeed, the quantities that we estimate, and the corresponding norms that we use, turn out to be quite involved in the two settings that we address, and especially for the major player problem. We postpone the details on the technicalities to the next two sections.

3. The second order master equation

This section is devoted to the proof of Theorem 2.3. The assumptions of Section 2.3 will be in force. Following the discussion at the end of the previous section, we are going to apply a Trotter–Kato alternating scheme using the first order master equation (problem (6)) and the linear second order master equation (problem (7)). The key step will be to show that both problems provide suitable estimates for the solution U in the following norm:

$$\sup_{t\in[0,T]}\left(\|U(t)\|_n+\left\|\frac{\delta U}{\delta m}(t)\right\|_{n-1;k}+\left\|\frac{\delta^2 U}{\delta m^2}(t)\right\|_{n-2;k-1,k-1}+\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U}{\delta m^2}(t)\right)\right)$$

where, we recall, the above quantities are defined in Section 2. In particular, for a function $U(t,x,m)$, the first term means an estimate on n-derivatives with respect to x, while the second and third terms yield an estimate on first and second derivatives with respect to m, represented by $\frac{\delta U}{\delta m}(t,x,y)$ and $\frac{\delta^2 U}{\delta m^2}(t,x,y,y')$ respectively: in this case, n refers to regularity in x, while k refers to regularity in y, y'.

Note that we will not only need bounds on derivatives, but also to establish some compactness for the scheme to converge. This motivates the presence of the Lipschitz norm in the above quantity.

Thus, the main technical issue for the proof of Theorem 2.3 will be to establish the following estimate (for some $T, C > 0$ depending on the upper bound M of norms of G)

$$\begin{aligned}&\sup_{t\in[0,T]}\left(\|U(t)\|_n+\left\|\frac{\delta U}{\delta m}(t)\|_{n-1;k}+\right\|\frac{\delta^2 U}{\delta m^2}(t)\right\|_{n-2;k-1,k-1}+\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U}{\delta m^2}(t)\right)\Bigg)\\&\le \|G\|_n+\left\|\frac{\delta G}{\delta m}\right\|_{n-1;k}+\left\|\frac{\delta^2 G}{\delta m^2}\right\|_{n-2;k-1,k-1}+\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2}\right)+CT, \qquad (23)\end{aligned}$$

for both the solution of problem (6) and the solution of problem (7). The analysis of the former, being quite technical, is postponed to Section 5.2 below. The latter is considered in the next subsection. The bounds on the four terms appearing in (23) will be obtained in different propositions, which will in turn require several steps (especially those for the first order master equation). For a quick reference on each estimate on the individual terms above, one may have a look at the table below.

	$\|U\|_n$ and $\|\frac{\delta U}{\delta m}\|_{n-1;k}$	$\|\frac{\delta^2 U}{\delta m^2}\|_{n-2;k-1,k-1}$	$\mathrm{Lip}_{n-3;k-2,k-2}(\frac{\delta^2 U}{\delta m^2})$
1st order master equation	Prop. 5.11	Prop. 5.16	Prop. 5.18
linear 2nd order master eq.	Proposition 3.1		

Finally, Lemma 3.5 shows how these estimates are chained together in the Trotter–Kato scheme, with some additional control of Hölder/Lipschitz seminorms which is needed for the convergence of the scheme and is obtained by interpolation.

3.1. The linear second order master equation

In this section we consider the (forward) second order linear master equation

$$\begin{cases} \partial_t U(t,x,m) - \mathrm{Tr}[a^0 D^2_{xx} U(t,x,m)] - \displaystyle\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{ym} U(t,x,m,y)]\, m(dy) \\ \qquad\qquad - 2 \displaystyle\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{xm} U(t,x,m,y)]\, m(dy) \\ \qquad\qquad - \displaystyle\int_{\mathbb{R}^{2d}} \mathrm{Tr}[a^0 D^2_{mm} U(t,x,m,y,y')]\, m(dy)\, m(dy') = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2, \\ U(0,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2. \end{cases} \tag{24}$$

Let Γ be the fundamental solution of the equation associated with a^0:

$$\begin{cases} \partial_t \Gamma(t,x) - \mathrm{Tr}[a^0 D^2_{xx}\Gamma(t,x)] = 0 \quad \text{in } (0,+\infty)\times\mathbb{R}^d, \\ \Gamma(0,x) = \delta_0(x) \quad \text{in } \mathbb{R}^d, \end{cases}$$

and, given a map $G : \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$ of class C^2 in (x,m), let us set

$$U(t,x,m) = \int_{\mathbb{R}^d} G(\xi, (\mathrm{id} - x + \xi)_\sharp m)\Gamma(t, x-\xi)\, d\xi \quad \forall (t,x,m)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2.$$

Proposition 3.1. *The map U is a classical solution to the second order equation* (24). *Moreover, there exists a constant $C > 0$ (depending only on n, k and a^0) such that*

$$\sup_{t\in[0,T]} \|U(t)\|_n \le (1+CT) \sup_{m\in\mathcal{P}_2} \|G\|_n$$

and, for $k \in \{2,\dots,n-1\}$,

$$\sup_{t\in[0,T]} \left\| \frac{\delta U}{\delta m}(t) \right\|_{n-1;k} \le (1+CT) \left\| \frac{\delta G}{\delta m} \right\|_{n-1;k},$$

$$\sup_{t\in[0,T]} \left\| \frac{\delta^2 U}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} \le (1+CT) \left\| \frac{\delta^2 G}{\delta m^2} \right\|_{n-2;k-1,k-1}$$

$$\sup_{t\in[0,T]} \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U}{\delta m^2}(t) \right) \le (1+CT)\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2} \right).$$

Remark 3.2. If, for some constant M,

$$\|G\|_n + \left\| \frac{\delta G}{\delta m} \right\|_{n-1;k} + \left\| \frac{\delta^2 G}{\delta m^2} \right\|_{n-2;k-1,k-1} + \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2} \right) \le M,$$

then the above estimates can be rewritten in the form

$$\sup_{t\in[0,T]}\left(\|U(t)\|_n+\left\|\frac{\delta U}{\delta m}(t)\right\|_{n-1;k}+\left\|\frac{\delta^2 U}{\delta m^2}(t)\right\|_{n-2;k-1,k-1}+\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U}{\delta m^2}(t)\right)\right)$$
$$\leq \|G\|_n+\left\|\frac{\delta G}{\delta m}\right\|_{n-1;k}+\left\|\frac{\delta^2 G}{\delta m^2}\right\|_{n-2;k-1,k-1}+\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2}\right)+C_M T,$$

for some constant C_M depending on n, k, a^0 and M.

In order to prove this proposition, we need two lemmas, the proofs of which are easy and left to the reader.

Lemma 3.3. *Let $U:\mathcal{P}_2\to\mathbb{R}$ be L-C^1 and let $\phi:\mathbb{R}^d\to\mathbb{R}^d$ be of class C^1 with bounded derivative. Set $V(m)=U(\phi_\sharp m)$. Then V is L-C^1 with*

$$D_m V(m,y)=(D\phi(y))^T D_m U(\phi_\sharp m,\phi(y)).$$

Lemma 3.4. *Let $U:\mathcal{P}_2\to\mathbb{R}$ be L-C^1 and let $V(x,m)=U((\mathrm{id}+x)_\sharp m)$. Then V is of class C^1 with*

$$D_x V(x,m)=\int_{\mathbb{R}^d} D_m U((\mathrm{id}+x)_\sharp m, x+y)\,m(dy).$$

Proof of Proposition 3.1. Let us first note that

$$\begin{aligned} U(t,x,m)&=\int_{\mathbb{R}^d} G(\xi,(\mathrm{id}-x+\xi)_\sharp m)\Gamma(t,x-\xi)\,d\xi\\ &=\int_{\mathbb{R}^d} G(x-z,(\mathrm{id}-z)_\sharp m)\Gamma(t,z)\,dz.\end{aligned}$$

In particular, U is C^1 in t, C^2 in x and has second order derivatives which are C^2 in the space variables with, in view of Lemmas 3.3 and 3.4,

$$\begin{aligned} D_x U(t,x,m)&=\int_{\mathbb{R}^d} D_x G(x-y,(\mathrm{id}-y)_\sharp m)\Gamma(t,y)\,dy,\\ D_x^2 U(t,x,m)&=\int_{\mathbb{R}^d} D_x^2 G(x-y,(\mathrm{id}-y)_\sharp m)\Gamma(t,y)\,dy,\\ D_m U(t,x,m,y)&=\int_{\mathbb{R}^d} D_m G(x-z,(\mathrm{id}-z)_\sharp m,y-z)\Gamma(t,z)\,dz,\\ D_m^2 U(t,x,m,y,y')&=\int_{\mathbb{R}^d} D_m^2 G(x-z,(\mathrm{id}-z)_\sharp m,y-z,y'-z)\Gamma(t,z)\,dz.\end{aligned}$$

This easily implies the estimates on U and its derivatives.

On the other hand, since $(\mathrm{id}-w)_\sharp[(\mathrm{id}-z)_\sharp m]=(\mathrm{id}-z-w)_\sharp m$, we have, for any $t\in(0,T)$ and $h\in(0,T-t)$,

$$\begin{aligned}
&\int_{\mathbb{R}^d} U(t, x - z, (\mathrm{id} - z)_\sharp m)\Gamma(h, z)\, dz \\
&\qquad = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} G(x - z - w, (\mathrm{id} - z - w)_\sharp m)\Gamma(h, z)\Gamma(t, w)\, dw\, dz \\
&\qquad = \int_{\mathbb{R}^d} G(x - u, (\mathrm{id} - u)_\sharp m)\Big(\int_{\mathbb{R}^d} \Gamma(h, u - w)\Gamma(t, w)\, dw\Big)\, du \\
&\qquad = \int_{\mathbb{R}^d} G(x - u, (\mathrm{id} - u)_\sharp m)\Gamma(t + h, u)\, du = U(t + h, x, m).
\end{aligned}$$

So, taking the derivative with respect to $h > 0$ in the above expression we get

$$\partial_t U(t + h, x, m) = \int_{\mathbb{R}^d} U(t, x - z, (\mathrm{id} - z)_\sharp m)\partial_t\Gamma(h, z)\, dz.$$

Integrating by parts and using Lemmas 3.3 and 3.4 yields

$$\begin{aligned}
&\partial_t U(t + h, x, m) = \int_{\mathbb{R}^d} U(t, x - z, (\mathrm{id} - z)_\sharp m)\big(\mathrm{Tr}[a^0 D^2_{zz}\Gamma(h, z)]\big)\, dz \\
&= \int_{\mathbb{R}^d}\Big(\mathrm{Tr}[a^0 D^2_{xx} U(t, x - z, (\mathrm{id} - z)_\sharp m)] \\
&\quad + 2\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{xm} U(t, x-z, (\mathrm{id}-z)_\sharp m, y-z)]\, m(dy) \\
&\quad + \int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{ym} U(t, x-z, (\mathrm{id}-z)_\sharp m, y-z)]\, m(dy) \\
&\quad + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{mm} U(t, x-z, (\mathrm{id}-z)_\sharp m, y-z, y'-z)]\, m(dy)\, m(dy')\Big)\Gamma(h, z)\, dz.
\end{aligned}$$

Letting $h \to 0$ we obtain

$$\begin{aligned}
\partial_t U(t, x, m) = \mathrm{Tr}[a^0 D^2_{xx} U(t, x, m)] &+ 2\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{xm} U(t, x, m, y)]\, m(dy) \\
&+ \int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{ym} U(t, x, m, y)]\, m(dy) \\
&+ \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{mm} U(t, x, m, y, y')]\, m(dy)\, m(dy').
\end{aligned}$$

So U is a solution to (24). ■

3.2. Existence of a solution

3.2.1. Definition of the semi-discrete scheme. Let us fix some horizon $T > 0$ (small) and a step size $\tau := T/(2N)$ (where $N \in \mathbb{N}$, $N \geq 1$). We set $t_k = kT/(2N)$, $k \in \{0, 2N\}$. We define by backward induction a continuous map $U^N = U^N(t, x, m)$, with $U^N : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2 \to \mathbb{R}$, as follows: we require that

(i) U^N satisfies the terminal condition

$$U^N(T, x, m) = G(x, m) \quad \forall (x, m) \in \mathbb{R}^d \times \mathcal{P}_2,$$

(ii) U^N solves the backward linear second order master equation

$$-\partial_t U^N - 2\,\mathrm{Tr}[a^0 D^2_{xx} U^N] - 2\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{ym} U^N]\, m(dy) \\ - 4\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{xm} U^N]\, m(dy) - 2\int_{\mathbb{R}^{2d}} \mathrm{Tr}[a^0 D^2_{mm} U^N]\, m(dy)\, m(dy') = 0 \tag{25}$$

on time intervals of the form (t_{2j+1}, t_{2j+2}) for $j = 0, \ldots, N-1$,

(iii) U^N solves the first order master equation

$$-\partial_t U^N - 2\,\mathrm{Tr}(a D^2_{xx} U^N) + 2H(x, D_x U^N, m) - 2\int_{\mathbb{R}^d} \mathrm{Tr}(a D^2_{ym} U^N)\, m(dy) \\ + 2\int_{\mathbb{R}^d} D_m U^N \cdot H_p(y, D_x U^N, m)\, m(dy) = 0 \tag{26}$$

on time intervals of the form (t_{2j}, t_{2j+1}), for $j = 0, \ldots, N-1$.

Our aim is to show that if the time horizon is short enough, U^N converges to a solution of the second order master equation as $N \to +\infty$.

3.2.2. Estimates of U^N. For $n \geq 4$ and $k \in \{3, \ldots, n-1\}$, let

$$M := \|G\|_n + \left\| \frac{\delta G}{\delta m} \right\|_{n-1;k} + \left\| \frac{\delta^2 G}{\delta m^2} \right\|_{n-2;k-1,k-1} + \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2} \right) + 1. \tag{27}$$

Lemma 3.5. *There exists $T_M > 0$ such that, for any $T \in (0, T_M]$ and $N \geq 1$, we have*

$$\sup_{t\in[0,T]} \Bigg(\|U^N(t)\|_n + \left\| \frac{\delta U^N}{\delta m}(t) \right\|_{n-1;k} + \left\| \frac{\delta^2 U^N}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} \\ + \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U^N}{\delta m^2}(t) \right) \Bigg) \leq M.$$

Moreover:

- *The maps U^N, $D_x U^N$, $D^2_{xx} U^N$ are globally Lipschitz continuous in (t, x, m), uniformly with respect to N.*
- *The maps $D_m U$, $D_m D_x U^N$, $D_y D_m U^N$ are Hölder continuous in (t, x, m, y), uniformly with respect to N, in any set of the form*

$$\{(t, x, m, y) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_2 \times \mathbb{R}^d : M_2(m) \leq R,\ |y| \leq R\}, \tag{28}$$

  where $M_2(m) = (\int_{\mathbb{R}^d} |y|^2\, m(dy))^{1/2}$.
- *The map $D^2_m U^N$ is Hölder continuous in (t, x, m, y, y'), uniformly with respect to N, in any set of the form*

$$\{(t, x, m, y, y') \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_2 \times \mathbb{R}^d \times \mathbb{R}^d : M_2(m) \leq R,\ |y|, |y'| \leq R\}. \tag{29}$$

Proof. In order to prove the estimate, we use Proposition 3.1 as well as Propositions 5.11, 5.16, 5.18 (in Section 5.2 below). Let T_M be the smallest positive constant associated with these propositions. Let also C_M be the largest constant in Propositions 3.1, 5.11, 5.16 and 5.18. We assume without loss of generality that $T_M < 1/(2C_M)$ and we fix $T \in (0, T_M]$.

We define the sequence $(\theta_k)_{k=0}^{2N}$ by

$$\theta_{2j} = M - 1 + C_M \frac{T}{N}(N-j), \quad j = 0, \dots, N.$$

As $T_M \le 1/(2C_M)$, we have $\theta_{2j} \le M$ for any $T \in (0, T_M]$ and $N \ge 1$.

Now, using Propositions 5.11, 5.16, 5.18 and 3.1 one checks by backward induction that

$$\sup_{t\in[t_{2j},t_{2j+2}]} \left\{ \|U^N(t)\|_n + \left\| \frac{\delta U^N}{\delta m}(t) \right\|_{n-1;k} + \left\| \frac{\delta^2 U^N}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} + \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U^N}{\delta m^2}(t) \right) \right\} \le \theta_{2j} \le M \quad \forall j = 0, \dots, N-1. \tag{30}$$

Indeed, assume that this is true for $j+1$; Proposition 3.1 (see also Remark 3.2), applied in the interval $[t_{2j+1}, t_{2j+2}]$ and with the terminal condition $U^N(t_{2j+2}, \cdot, \cdot)$ which satisfies (30) by assumption, implies that

$$\sup_{t\in[t_{2j+1},t_{2j+2}]} \left\{ \|U^N(t)\|_n + \left\| \frac{\delta U^N}{\delta m}(t) \right\|_{n-1;k} + \left\| \frac{\delta^2 U^N}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} + \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U^N}{\delta m^2}(t) \right) \right\} \le \theta_{2j+2} + \frac{C_M T}{2N}.$$

Then using Propositions 5.11, 5.16, 5.18 for the interval $[t_{2j}, t_{2j+1}]$ and the terminal condition $U^N(t_{2j+1}, \cdot, \cdot)$ for which (30) now holds, one gets

$$\sup_{t\in[t_{2j},t_{2j+1}]} \left\{ \|U^N(t)\|_n + \left\| \frac{\delta U^N}{\delta m}(t) \right\|_{n-1;k} + \left\| \frac{\delta^2 U^N}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} + \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U^N}{\delta m^2}(t) \right) \right\} \le \theta_{2j+2} + \frac{C_M T}{N} = \theta_{2j},$$

so (30) holds for j. Since the first step ($j = N-1$) can be proved similarly using the very definition of M in (27), we can conclude that (30) holds for every $j = 0, \dots, N-1$.

We now prove the second part of the lemma. As U^N solves (25) on the time intervals (t_{2j+1}, t_{2j+2}) and (26) on (t_{2j}, t_{2j+1}), we obtain directly, by the space estimates proved above,

$$\sup_{t,m} \|\partial_t U(t, \cdot, m)\|_{n-2} \le C_M, \tag{31}$$

where C_M does not depend on N.

Let now $l \in \mathbb{N}^d$ with $|l| \le 2$. By (31) and the fact that $\|U^N\|_n$ is bounded for $n > |l|$, $D^l U^N$ is uniformly Lipschitz continuous in t and x. Moreover, since $\|\frac{\delta U^N}{\delta m}\|_{n-1;k}$ is bounded (for $k \ge 1$), $D^l U^N$ is uniformly Lipschitz continuous in m as well by Remark 2.2 since $|l| \le n-1$.

Next we prove the uniform continuity of $D_x^l D_y^r D_m U^N$ for $|l|, |r| \le 1$. First we recall that $\|\frac{\delta U^N}{\delta m}\|_{n-1;k}$ is bounded, so that $\|D_m U^N\|_{n-1;k-1}$ is bounded, with $n-1 \ge 2$ and $k-1 \ge 2$. Therefore $D_x^l D_y^r D_m U^N$ is uniformly Lipschitz continuous in (x, y) (for y, this is Remark 2.2). Second, recall that $\|\frac{\delta^2 U^N}{\delta m^2}\|_{n-2;k-1,k-1}$ is bounded, so that $\|\frac{\delta}{\delta m} D_m U^N\|_{n-2;k-2,k-1}$ is bounded as well, with $n \ge 3$ and $k \ge 3$; therefore $D_x^l D_y^r D_m U^N$ is uniformly Lipschitz continuous in m. As we have already proved that U^N is uniformly Lipschitz continuous in t, we can deduce from Lemma B.4 below applied to U^N that $D_m U^N$ is also Hölder continuous in time in any set of the form (28).

Finally, we consider $D^2_{mm} U^N = D^2_{mm} U^N(t, x, m, y, y')$. Since $\|\frac{\delta^2 U^N}{\delta m^2}\|_{n-2;k-1,k-1}$ and $\mathrm{Lip}_{n-3;k-2,k-2}(\frac{\delta^2 U^N}{\delta m^2})$ are bounded, with $n \ge 4$ and $k \ge 3$, $D^2_{mm} U^N$ is uniformly Lipschitz continuous in (x, m, y, y'). Applying Lemma B.4 to the map $D_m U^N$, which is Hölder continuous in time in sets of the form (28) (as we have seen above) and such that $D^2_{mm} U$ is uniformly Lipschitz in (m, y, y'), we deduce that $D^2_{mm} U^N$ is also Hölder continuous in time, uniformly in N, in sets of the form (29). So we conclude that $D^2_{mm} U^N$ is uniformly Hölder continuous in all variables. ∎

3.2.3. Proof of Theorem 2.3. In view of Lemma 3.5, the maps U^N, $D_x U^N$, $D^2_{xx} U^N$, $D_m U^N$, $D_m D_x U^N$, $D_y D_m U^N$ and $D^2_m U^N$ are locally Hölder continuous in all variables, uniformly with respect to N. So, by a version of the Arzelà–Ascoli theorem (see Lemma B.5 below), there is a subsequence denoted in the same way such that U^N, $D_x U^N$, $D^2_{xx} U^N$, $D_m U^N$, $D_m D_x U^N$, $D_y D_m U^N$ and $D^2_m U^N$ converge pointwise in m and locally uniformly in time-space to some maps U, $D_x U$, $D^2_{xx} U$, V, $D_x V$, $D_y V$ and W. Moreover, using the integral formula (10), it is easy to check that $V = D_m U$ and $W = D^2_m U$.

By the equation satisfied by U^N we have, for any $0 \le s < t \le T$,

$$
\begin{aligned}
&U^N(t, x, m) - U^N(s, x, m) \\
&= -\sum_{k=0}^{N-1} \int_{t_{2k+1}}^{t_{2k+2}} 2\Big\{ \mathrm{Tr}[a^0 D^2_{xx} U^N] + \int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{ym} U^N]\, m(dy) \\
&\quad + 2\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{xm} U^N]\, m(dy) + \int_{\mathbb{R}^{2d}} \mathrm{Tr}[a^0 D^2_{mm} U^N]\, m(dy)\, m(dy') \Big\} \mathbf{1}_{[s,t]}(\tau)\, d\tau \\
&\quad - \sum_{k=0}^{N-1} \int_{t_{2k}}^{t_{2k+1}} 2\Big\{ \mathrm{Tr}(a D^2_{xx} U^N) - H(x, D_x U^N, m) \\
&\quad + \int_{\mathbb{R}^d} \mathrm{Tr}(a D^2_{ym} U^N)\; m(dy) - \int_{\mathbb{R}^d} D_m U^N \cdot H_p(y, D_x U^N, m)\, m(dy) \Big\} \mathbf{1}_{[s,t]}(\tau)\, d\tau.
\end{aligned}
$$

Since, as N tends to infinity, the maps

$$t \mapsto \sum_{k=0}^{N-1} \mathbf{1}_{[t_{2k+1}, t_{2k+2}]}(t) \quad \text{and} \quad t \mapsto \sum_{k=0}^{N-1} \mathbf{1}_{[t_{2k}, t_{2k+1}]}(t)$$

weakly converge to the constant $1/2$ and since the space integrals in the above equation converge pointwise to the corresponding quantities for the limit U, by the dominated convergence theorem we obtain

$$\begin{aligned}
&U(t,x,m) - U(s,x,m) \\
&\quad = -\int_s^t \Big(\mathrm{Tr}[a^0 D^2_{xx} U] + \int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{ym} U]\, dm \\
&\qquad + 2\int_{\mathbb{R}^d} \mathrm{Tr}[a^0 D^2_{xm} U]\, dm + \int_{\mathbb{R}^{2d}} \mathrm{Tr}[a^0 D^2_{mm} U]\, dm \otimes dm \\
&\qquad + \mathrm{Tr}(a D^2_{xx} U) - H(x, D_x U, m) \\
&\qquad + \int_{\mathbb{R}^d} \mathrm{Tr}(a D^2_{ym} U)\, dm - \int_{\mathbb{R}^d} D_m U \cdot H_p(y, D_x U, m)\, dm \Big)\, d\tau,
\end{aligned}$$

so that U is a classical solution to (19). ∎

3.3. *Existence of the solution to the stochastic MFG system*

This section is devoted to the (short) proof of Theorem 2.4.

Proof of Theorem 2.4. Let m be the solution to the stochastic McKean–Vlasov equation

$$\begin{cases} dm(t,x) = \Big[\sum_{i,j} D_{ij}((a_{i,j} + a^0_{i,j})(t,x) m(t,x)) \\ \qquad\qquad + \mathrm{div}(m(t,x) H_p(x, DU(t,x,m(t)), m(t))) \Big] dt \\ \qquad\qquad - \mathrm{div}(m(t,x)\sqrt{2}\,\sigma^0 dW_t) \quad \text{in } (0,T)\times\mathbb{R}^d, \\ m(0,dx) = m_0 \quad \text{in } \mathbb{R}^d. \end{cases} \tag{32}$$

Existence of a solution for this system can be obtained, for instance, as the mean field limit of the SDE

$$\begin{cases} dX^{N,i}_s = -H_p(X^{N,i}_s, D_x U(t, X^{N,i}_s, m^N_{\mathbf{X}^N_s}), m^N_{\mathbf{X}^N_s})\, ds \\ \qquad\qquad + \sqrt{2}\,\sigma(s, X^{N,i}_s)\, , dB^i_s + \sqrt{2}\,\sigma^0(s, X^{N,i}_s)\, dW_s \\ X^{N,i}_0 = \bar{X}^{N,i}_0, \end{cases}$$

where $\bar{X}^{N,i}_0$ is a family of i.i.d. r.v. of law m_0 and where $m^N_{\mathbf{X}^N_s} = \frac{1}{N}\sum_{i=1}^N \delta_{X^{N,i}_s}$. Indeed, one can show that the family of laws of $(m^N_{\mathbf{X}^N_s})$ is tight in $C^0([0,T], \mathcal{P}_2)$ and that its limit is a solution to (32). Uniqueness for (32) comes from the regularity of U and Gronwall's lemma.

Then one can use Itô's formula [10, Theorem A.1] (see also [12, Theorem 11.13]) to derive that $u(t,x) := U(t,x,m(t))$ solves the backward stochastic HJ equation

$$\begin{cases} du(t,x) = [-\operatorname{Tr}((a+a^0)(t,x)D^2u(t,x)) + H(x,Du(t,x),m(t)) \\ \qquad\qquad - \sqrt{2}\operatorname{Tr}(\sigma^0 Dv(t,x))]\,dt + v(t,x)\cdot dW_t \quad \text{in } (0,T)\times\mathbb{R}^d, \\ u(T,x) = G(x,m(T)) \quad \text{in } \mathbb{R}^d, \end{cases}$$

where v is given by (21). Note that, by the regularity of U, u and v have the required regularity. ■

4. The master equation for MFGs with a major player

We now discuss the proof of Theorem 2.5. We recall that, throughout the whole section, the assumptions in Section 2.3 are in force.

The idea of the proof follows a similar splitting method as we did in Section 3, by dividing the time interval $[0,T]$ into $[t_{2k},t_{2k+1})$ and $[t_{2k+1},t_{2k+2})$, where $t_k = kT/(2N)$, $k\in\{0,2N\}$. This time we alternate the following two problems: in $[t_{2k+1},t_{2k+2})$ we solve, for a fixed $x_0\in\mathbb{R}^{d_0}$, the first order system of master equations in $\mathbb{R}^d\times\mathcal{P}_2$:

$$\begin{cases} -\partial_t U^0 - 2\displaystyle\int_{\mathbb{R}^d} \operatorname{div}_y D_m U^0(t,x_0,m,y)\,m(dy) \\ \quad +2\displaystyle\int_{\mathbb{R}^d} D_m U^0(t,x_0,m,y)\cdot H_p(x_0,y,D_xU(t,x_0,y,m),m)\,m(dy) = 0, \\ -\partial_t U - 2\Delta_x U + 2H(x_0,x,D_xU,m) - 2\displaystyle\int_{\mathbb{R}^d} \operatorname{div}_y D_m U(t,x_0,x,m,y)\,m(dy) \\ \quad +2\displaystyle\int_{\mathbb{R}^d} D_m U(t,x_0,x,m,y)\cdot H_p(x_0,y,D_xU(t,x_0,y,m),m)\,m(dy) = 0, \end{cases} \tag{33}$$

while on $[t_{2k},t_{2k+1})$ we solve for a fixed $(x,m)\in\mathbb{R}^d\times\mathcal{P}_2$ the system of HJ equations in $\mathbb{R}^{d_0}$:

$$\begin{cases} \text{(i) } -\partial_t U^0 - 2\Delta_{x_0}U^0 + 2H^0(x_0,D_{x_0}U^0,m) = 0, \\ \text{(ii) } -\partial_t U - 2\Delta_{x_0}U + 2D_{x_0}U\cdot H^0_p(x_0,D_{x_0}U^0(t,x_0,m),m) = 0. \end{cases} \tag{34}$$

As explained at the end of Section 2, we need to introduce a suitable norm which is preserved in the estimates of *both* problems. To this end, we need to treat the pair of maps (U^0,U) simultaneously; this requires specific notation that we discuss first.

4.1. Notation for the norms

We will be dealing with pairs of maps $(V^0,V) = (V^0(x_0,m),V(x_0,x,m))$ which might also depend on time t, not indicated here. The way we compute the norms is crucial in order to match all the estimates. We use the following norms:

$$\|(V^0,V)\|_n := \sup_{m\in\mathcal{P}_2}\sum_{r=0}^n \sup_{x_0\in\mathbb{R}^{d_0},\, x\in\mathbb{R}^d} \left(|V^0(x_0,m)|^2+|D_x^r V(x_0,x,m)|^2\right)^{1/2},$$

$$\left\|\frac{\delta(V^0,V)}{\delta m}\right\|_{n;k} := \sup_{m\in\mathcal{P}_2}\sum_{r=0}^n \sup_{\substack{x_0\in\mathbb{R}^{d_0},\, x\in\mathbb{R}^d,\\ \rho\in C_b^0,\, \|\rho\|_{-k}=1}} \left(\left|\frac{\delta V^0}{\delta m}(x_0,m)(\rho)\right|^2+\left|D_x^r\frac{\delta V}{\delta m}(x_0,x,m)(\rho)\right|^2\right)^{1/2},$$

$$\left\|\frac{\delta^2(V^0,V)}{\delta m^2}\right\|_{n;k,k} := \sup_{m\in\mathcal{P}_2}\sum_{r=0}^n \sup_{\substack{x_0\in\mathbb{R}^{d_0},\, x\in\mathbb{R}^d,\\ \rho,\rho'\in C_b^0,\, \|\rho\|_{-k}=\|\rho'\|_{-k}=1}} \left(\left|\frac{\delta^2 V^0}{\delta m^2}(x_0,m)(\rho,\rho')\right|^2+\left|D_x^r\frac{\delta^2 V}{\delta m^2}(x_0,x,m)(\rho,\rho')\right|^2\right)^{1/2}$$

and

$$\begin{aligned}
&\mathrm{Lip}_{n;k,k}\left(\frac{\delta^2(V^0,V)}{\delta m^2}\right)\\
&:= \sup_{m_1\neq m_2} \mathbf{d}_2(m_1,m_2)^{-1}\left\|\frac{\delta^2}{\delta m^2}\left(V^0(m_2)-V^0(m_1), V(m_2)-V(m_1)\right)\right\|_{n;k,k}\\
&= \sup_{m_1\neq m_2} \mathbf{d}_2(m_1,m_2)^{-1}\sum_{r=0}^n \sup_{\substack{x_0\in\mathbb{R}^{d_0},\, x\in\mathbb{R}^d,\\ \rho,\rho'\in C_b^0,\, \|\rho\|_{-k}=\|\rho'\|_{-k}=1}} \left(\left|\frac{\delta^2 V^0}{\delta m^2}(x_0,m_2)(\rho,\rho')-\frac{\delta^2 V^0}{\delta m^2}(x_0,m_1)(\rho,\rho')\right|^2\right.\\
&\qquad\left.+\left|D_x^r\frac{\delta^2 V}{\delta m^2}(x_0,x,m_2)(\rho,\rho')-D_x^r\frac{\delta^2 V}{\delta m^2}(x_0,x,m_1)(\rho,\rho')\right|^2\right)^{1/2}.
\end{aligned}$$

We define in a similar way the quantities

$$\mathrm{Lip}_n^{x_0}(D_{x_0}^2V^0, D_{x_0}^2V),\quad \mathrm{Lip}_{n;k}\left(\frac{\delta V_{x_0}^0}{\delta m}, \frac{\delta V_{x_0}}{\delta m}\right),\quad \mathrm{Lip}_n(D_{x_0}^2V^0, D_{x_0}^2U).$$

Note that arguing as in Remark 2.2, control of $\|\frac{\delta(V^0,V)}{\delta m}\|_{n;k}$ yields control of $\|\frac{\delta V^0}{\delta m}\|_{n,k-1}$ and $\|\frac{\delta V}{\delta m}\|_{n,k-1}$, and similarly for $\|\frac{\delta^2(V^0,V)}{\delta m^2}\|_{n;k,k}$, $\mathrm{Lip}_{n;k,k}(\frac{\delta^2(V^0,V)}{\delta m^2})$, ...

We are going to show that the two systems (33) and (34) preserve with a linear rate (as in (23) for the second order master equation) the following norms:

$$
\begin{aligned}
&\|(U^0,U)(t)\|_n + \|D_{x_0}(U^0,U)(t)\|_{n-1} + \|D^2_{x_0}(U^0,U)(t)\|_{n-2} \\
&+\mathrm{Lip}^{x_0}_{n-3}((D^2_{x_0}U^0, D^2_{x_0}U)(t)) \\
&+\left\|\frac{\delta(U^0,U)}{\delta m}(t)\right\|_{n-1;k} + \left\|\frac{\delta(U^0_{x_0},U_{x_0})}{\delta m}(t)\right\|_{n-2;k-1} + \mathrm{Lip}^{x_0}_{n-3;k-2}\left(\left(\frac{\delta U^0_{x_0}}{\delta m},\frac{\delta U_{x_0}}{\delta m}\right)(t)\right) \\
&+\left\|\frac{\delta^2(U^0,U)}{\delta m^2}(t)\right\|_{n-2;k-1,k-1} + \mathrm{Lip}^{x_0}_{n-3;k-2,k-2}\left(\left(\frac{\delta^2 U^0}{\delta m^2},\frac{\delta^2 U}{\delta m^2}\right)(t)\right) \\
&+\mathrm{Lip}_{n-3;k-2,k-2}\left(\left(\frac{\delta^2 U^0}{\delta m^2},\frac{\delta^2 U}{\delta m^2}\right)(t)\right) + \mathrm{Lip}_{n-3;k-2}\left(\left(\frac{\delta U^0_{x_0}}{\delta m},\frac{\delta U_{x_0}}{\delta m}\right)(t)\right).
\end{aligned}
$$

The (technical) analysis of the system of master equations (33) is postponed to Section 5.2. We rather concentrate on system (34) in the next subsection. We stress again that the only difference from the second order master equation problem is the derivation of suitable bounds. Once these are given, the proof of the convergence of the Trotter–Kato scheme is identical. Since these bounds are collected in several propositions, we give a short guidance for the reader in the following table, where for each term in the above defined norms, we refer to the proposition in which this term is estimated.

	First order system of master eqns.	System of HJ equations
$\|(U^0,U)\|_n$	Proposition 5.15	Proposition 4.1
$\|D_{x_0}(U^0,U)\|_{n-1}$		
$\|D^2_{x_0}(U^0,U)\|_{n-2}$	Proposition 5.17	
$\mathrm{Lip}^{x_0}_{n-3}(D^2_{x_0}U^0, D^2_{x_0}U)$	Proposition 5.19	
$\|\frac{\delta(U^0,U)}{\delta m}\|_{n-1;k}$	Proposition 5.15	Proposition 4.2
$\|\frac{\delta(U^0_{x_0},U_{x_0})}{\delta m}\|_{n-2;k-1}$	Proposition 5.17	
$\mathrm{Lip}^{x_0}_{n-3;k-2}(\frac{\delta U^0_{x_0}}{\delta m},\frac{\delta U_{x_0}}{\delta m})$	Proposition 5.19	
$\|\frac{\delta^2(U^0,U)}{\delta m^2}\|_{n-2;k-1,k-1}$	Proposition 5.17	Proposition 4.3
$\mathrm{Lip}^{x_0}_{n-3;k-2,k-2}(\frac{\delta^2 U^0}{\delta m^2},\frac{\delta^2 U}{\delta m^2})$	Proposition 5.19	
$\mathrm{Lip}_{n-3;k-2,k-2}(\frac{\delta^2 U^0}{\delta m^2},\frac{\delta^2 U}{\delta m^2})$		Proposition 4.4
$\mathrm{Lip}_{n-3;k-2}(\frac{\delta U^0_{x_0}}{\delta m},\frac{\delta U_{x_0}}{\delta m})$		Proposition 4.5

4.2. Analysis of the simple system of HJ equations

In this section we consider the system

$$
\begin{cases}
-\partial_t U^0(t,x_0;m) - \Delta_{x_0} U^0(t,x_0;m) + H^0(x_0, D_{x_0}U^0(t,x_0;m), m) = 0 \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T)\times\mathbb{R}^{d_0}, \\
-\partial_t U(t,x_0;x,m) - \Delta_{x_0} U(t,x_0;x,m) \\
\quad + D_{x_0}U(t,x_0;x,m)\cdot H^0_p(x_0, D_{x_0}U^0(t,x_0,m),m) = 0 \quad \text{in } (0,T)\times\mathbb{R}^{d_0}, \\
U^0(T,x_0;m) = G^0(x_0,m) \text{ in } \mathbb{R}^{d_0},\ U(T,x_0;x,m) = G(x_0,x,m) \quad \text{in } \mathbb{R}^{d_0},
\end{cases}
\tag{35}
$$

where $(x,m)\in\mathbb{R}^d\times\mathcal{P}_2$ are fixed. The main part of this subsection consists in proving estimates on the solution (U^0,U) to (35).

4.2.1. Basic regularity of (U^0,U). We recall that H^0 satisfies the assumptions of Section 2.3, in particular condition (18) is in force.

Proposition 4.1. *Fix $M>0$ and $n\geq 3$. There are constants $K_M, T_M>0$, depending on M, C_0 and γ, and a constant $C_M>0$ depending on*

$$\sup_{|p|\leq K_M}\sup_{m\in\mathcal{P}_2}\sum_{k=0}^{3}\|D^k_{(x_0,p)}H^0(\cdot,p,m)\|_\infty + \sum_{k=0}^{3}\|D^k_{(x_0,p)}H^0_p(\cdot,p,m)\|_\infty,$$

such that if

$$\|(G^0,G)\|_n + \|D_{x_0}(G^0,G)\|_{n-1} + \|D^2_{x_0}(G^0,G))\|_{n-2} + \mathrm{Lip}^{x_0}_{n-3}(D^2_{x_0}G^0, D^2_{x_0}G)\leq M,$$

then, for any $T\in(0,T_M)$, we have

$$
\begin{aligned}
\sup_t\big(&\|(U^0,U)(t)\|_n + \|D_{x_0}(U^0,U)(t)\|_{n-1} + \|D^2_{x_0}(U^0,U)(t)\|_{n-2} \\
&\qquad\qquad\qquad\qquad + \mathrm{Lip}^{x_0}_{n-3}(D^2_{x_0}(U^0,U)(t))\big) \\
\leq\ & \|(G^0,G)\|_n + \|D_{x_0}(G^0,G)\|_{n-1} + \|D^2_{x_0}(G^0,G))\|_{n-2} + \mathrm{Lip}^{x_0}_{n-3}(D^2_{x_0}(G^0,G)) + C_M T.
\end{aligned}
$$

Proof. To estimate $\|(U^0,U)\|_n$ it suffices to apply successively Proposition A.8 with $r=0$ and $l\leq n$, and to sum over l. The argument to estimate first and higher order derivatives with respect to x_0 is identical: apply successively Proposition A.8 with $r=1$ and $l\leq n-1$ (for $\|(D_{x_0}U^0, D_{x_0}U)\|_{n-1}$), with $r=2$ and $l\leq n-2$ (for $\|(D^2_{x_0}U^0, D^2_{x_0}U)\|_{n-2}$) and finally with $r=3$ and $l\leq n-3$ (for the Lipschitz bound in x_0 of $D^2_{x_0}(U^0,U)$). ■

4.2.2. First order differentiability in m

Proposition 4.2. *Under the assumptions of Proposition 4.1, the pair (U^0,U) is of class C^1 with respect to m, as also are its derivatives with respect to x appearing below, and, for any fixed $(x,m,\rho)\in\mathbb{R}^d\times\mathcal{P}_2\times C^{-k}$ the derivative*

$$(v^0,v) = \left(\frac{\delta U^0}{\delta m}(t,x_0;m)(\rho),\ \frac{\delta U}{\delta m}(t,x_0;x,m)(\rho)\right)$$

solves

$$
\begin{cases}
-\partial_t v^0 - \Delta_{x_0} v^0 + \dfrac{\delta H^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho) \\
\qquad\qquad\qquad + H_p^0(x_0, D_{x_0}U^0, m) \cdot D_{x_0} v^0 = 0 \quad \text{in } (0,T) \times \mathbb{R}^{d_0}, \\
-\partial_t v - \Delta_{x_0} v + D_{x_0} v \cdot H_p^0(x_0, D_{x_0}U^0, m) \\
\quad + D_{x_0} U \cdot \left(\dfrac{\delta H_p^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho) + H_{pp}^0(x_0, D_{x_0}U^0, m) D_{x_0} v^0 \right) = 0 \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T) \times \mathbb{R}^{d_0}, \\
v^0(T, x_0; m) = \dfrac{\delta G^0}{\delta m}(x_0, m)(\rho), \;\; v(T, x_0, x; m) = \dfrac{\delta G}{\delta m}(x_0, x, m)(\rho) \quad \text{in } \mathbb{R}^{d_0}.
\end{cases}
\tag{36}
$$

Suppose in addition that, for $k \geq 2$,

$$
\left\| \frac{\delta(G^0, G)}{\delta m} \right\|_{n-1;k} + \left\| \frac{\delta(G_{x_0}^0, G_{x_0})}{\delta m} \right\|_{n-2;k-1} + \mathrm{Lip}_{n-3;k-2}^{x_0}\left(\frac{\delta G_{x_0}^0}{\delta m}, \frac{\delta G_{x_0}}{\delta m} \right) \leq M.
$$

Then there exist $T_M, C_M > 0$ such that, for any $T \in (0, T_M)$, we have

$$
\begin{aligned}
&\sup_t \left(\left\| \frac{\delta(U^0, U)}{\delta m}(t) \right\|_{n-1;k} + \left\| \frac{\delta(U_{x_0}^0, U_{x_0})}{\delta m}(t) \right\|_{n-2;k-1} + \mathrm{Lip}_{n-3;k-2}^{x_0}\left(\frac{\delta U_{x_0}^0}{\delta m}, \frac{\delta U_{x_0}}{\delta m} \right)(t) \right) \\
&\leq \left(\left\| \frac{\delta(G^0, G)}{\delta m} \right\|_{n-1;k} + \left\| \frac{\delta(G_{x_0}^0, G_{x_0})}{\delta m} \right\|_{n-2;k-1} + \mathrm{Lip}_{n-3;k-2}^{x_0}\left(\frac{\delta G_{x_0}^0}{\delta m}, \frac{\delta G_{x_0}}{\delta m} \right) + C_M T,
\end{aligned}
$$

where C_M depends on M, r, n, k and on the regularity of H^0.

Proof. In order to show that U^0 is C^1 with respect to m, let us define

$$
\hat{U}^0(t, x_0; s, m, y) := U^0(t, x_0, (1-s)m + s\delta_y).
$$

Then, as $\hat{H}^0 := H^0(x_0, p, (1-s)m + s\delta_y)$ and $\hat{g}^0 := G^0(x_0, (1-s)m + s\delta_y)$ are of class C^1 with respect to the parameter $s \in [0,1]$, the map $\hat{U}^0$ is C^1 in s and its derivative $\hat{v}^0(t, x_0; m, y) := (d\hat{U}^0/ds)(t, x_0; 0, m, y)$ solves the linearized equation

$$
\begin{cases}
-\partial_t \hat{v}^0 - \Delta_{x_0} \hat{v}^0 + \dfrac{\delta H^0}{\delta m}(x_0, D_{x_0}U^0, m, y) + H_p^0(x_0, D_{x_0}U^0, m) \cdot D_{x_0} \hat{v}^0 = 0 \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T) \times \mathbb{R}^{d_0}, \\
\hat{v}^0(T, x_0) = \dfrac{\delta G^0}{\delta m}(x_0, m, y) \quad \text{in } \mathbb{R}^{d_0}.
\end{cases}
$$

By uniqueness and parabolic regularity, the solution to this equation depends continuously on the parameters (m, y). Hence Lemma B.1 states that U^0 is C^1 in m with $\frac{\delta U^0}{\delta m}(t, x_0, m, y) = \hat{v}^0(t, x_0; m, y)$.

Next we consider the linear equation satisfied by U. By our previous discussion on U^0, the vector field

$$
(t, x_0) \mapsto H_p^0(x_0, D_{x_0}U^0(t, x_0, m), m)
$$

is C^1 with respect to m. For $(s, m, y) \in [0,1] \times \mathcal{P}_2 \times \mathbb{R}^{d_0}$, the map $\hat{U}(t, x_0; s, x, m, y) := U(t, x_0, x, (1-s)m + s\delta_y)$ solves a linear equation in which the vector field

$$\hat{V}(t, x_0; s, m, y) := H_p^0(x_0, D_{x_0} U^0(t, x_0; (1-s)m + s\delta_y), (1-s)m + s\delta_y)$$

and the terminal condition $\hat{g}(x_0; x, s, m, y) := G(x_0, x, (1-s)m + s\delta_y)$ are C^1 in s. Then $\hat{U}$ is C^1 in s and its derivative $\hat{v}(t, x_0; x, m, y) := (d/ds)\hat{U}(t, x_0; 0, x, m, y)$ solves the linear equation

$$\begin{cases} -\partial_t \hat{v} - \Delta_{x_0} \hat{v} + D_{x_0} \hat{v} \cdot H_p^0(x_0, D_{x_0} U^0, m) \\ \quad + D_{x_0} U \cdot \left(\dfrac{\delta H_p}{\delta m}(x_0, D_{x_0} U^0, m, y) + H_{pp}(x_0, D_{x_0} U^0, m) D_{x_0} \hat{v}^0 \right) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T) \times \mathbb{R}^{d_0}, \\ \hat{v}(T, x_0; x, m, y) = \dfrac{\delta G}{\delta m}(x_0, x, m, y) \quad \text{in } \mathbb{R}^{d_0}. \end{cases}$$

As the solution to this equation depends continuously on the parameters (m, y), Lemma B.1 states that U is C^1 in m with $\frac{\delta U}{\delta m}(t, x_0, x, m, y) = \hat{v}(t, x_0; x, m, y)$. This proves that the derivative $(\hat{v}^0, \hat{v}) = (\frac{\delta U^0}{\delta m}, \frac{\delta U}{\delta m})(t, x_0, x, m, y)$ solves (36) with $\rho = \delta_y$.

Hence, for any $\rho \in C_b^0$, the pair $(v^0, v) = (\frac{\delta U^0}{\delta m}(t, x_0; m)(\rho), \frac{\delta U}{\delta m}(t, x_0; x, m)(\rho))$ solves a linear system of the form (116) in which the drifts

$$\begin{aligned} V^0(t, x^0; m) &:= H_p^0(x_0, D_{x_0} U^0(t, x_0, m), m), \\ V(t, x^0; x, m) &:= H_{pp}^0(x_0, D_{x_0} U^0(t, x_0, m), m) D_{x_0} U(t, x_0; x) \end{aligned}$$

are bounded of class C_b^1 and $C_b^{0,n-1} \cap C_b^{1,n-2}$ respectively, while the source terms

$$\begin{aligned} f^0(t, x^0; m) &:= \frac{\delta H^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho), \\ f(t, x_0; x, m) &:= D_{x_0} U(t, x_0; x) \cdot \frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho) \end{aligned}$$

are in C_b^1 and $C_b^{0,n-1} \cap C_b^{1,n-2}$ respectively, thanks to Proposition 4.1. We then use Proposition A.9 successively to obtain the estimates: first with $r = 0$ and $l \le n-1$, we get

$$\begin{aligned} &\left(\left| \frac{\delta U^0}{\delta m}(t, x_0; m)(\rho) \right|^2 + \left| D_x^l \frac{\delta U}{\delta m}(t, x_0; x, m)(\rho) \right|^2 \right)^{1/2} \\ &\quad \le (1 + CT) \sup_{x_0, x} \left(\left| \frac{\delta G^0}{\delta m}(x_0; m)(\rho) \right|^2 + \left| D_x^l \frac{\delta G}{\delta m}(x_0; x, m)(\rho) \right|^2 \right)^{1/2} + CT. \end{aligned}$$

Then by taking the supremum over $\|\rho\|_{-k} = 1, x_0, x$ and summing over $l \le n-1$ we find the estimate for $\| \frac{\delta (U^0, U)}{\delta m} \|_{n-1;k}$. An analogous application of Proposition A.9 with $r = 1$ and $l \le n-2$ provides the bound for $\| \frac{\delta (U^0_{x_0}, U_{x_0})}{\delta m} \|_{n-2;k-1}$, while the Lipschitz estimate in x_0 for $(\frac{\delta U^0_{x_0}}{\delta m}, \frac{\delta U_{x_0}}{\delta m})$ is obtained similarly with $r = 2$ and $l \le n-3$. ∎

4.2.3. Second order differentiability with respect to m

Proposition 4.3. *Under the assumptions of Proposition 4.2, $k \geq 3$, the pair (U^0, U) (together with its derivatives with respect to x) is of class C^2 with respect to m and, for any fixed $(x, m, \rho, \rho') \in \mathbb{R}^d \times \mathcal{P}_2 \times C^{-(k-1)} \times C^{-(k-1)}$ the derivative*

$$(w^0, w) = \left(\frac{\delta^2 U^0}{\delta m^2}(t, x_0; m)(\rho, \rho'), \frac{\delta^2 U}{\delta m^2}(t, x_0; x, m)(\rho, \rho') \right)$$

solves

$$\begin{cases} -\partial_t w^0 - \Delta_{x_0} w^0 + H_p^0(x_0, D_{x_0} U^0, m) \cdot D_{x_0} w^0 + \dfrac{\delta^2 H^0}{\delta m^2}(x_0, D_{x_0} U^0, m)(\rho, \rho') \\ \qquad + H_{pp}^0(x_0, D_{x_0} U^0, m) D_{x_0} v^0 \cdot D_{x_0} (v')^0 + \dfrac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho) \cdot D_{x_0} (v')^0 \\ \qquad + \dfrac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho') \cdot D_{x_0} v^0 = 0 \quad \text{in } (0, T) \times \mathbb{R}^{d_0}, \\ -\partial_t w - \Delta_{x_0} w + H_p^0(x_0, D_{x_0} U^0, m) \cdot D_{x_0} w \\ \qquad + D_{x_0} v \cdot \left(\dfrac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho') + H_{pp}^0(x_0, D_{x_0} U^0, m) D_{x_0} (v')^0 \right) \\ \qquad + D_{x_0} v' \cdot \left(\dfrac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho) + H_{pp}^0(x_0, D_{x_0} U^0, m) D_{x_0} v^0 \right) \\ \qquad + D_{x_0} U \cdot \Bigg(\dfrac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho) D_{x_0} (v')^0 + \dfrac{\delta^2 H_p^0}{\delta m^2}(x_0, D_{x_0} U^0, m)(\rho, \rho') \\ \qquad + H_{ppp}^0(x_0, D_{x_0} U^0, m) D_{x_0} v^0 D_{x_0} (v')^0 + \dfrac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0} U^0, m)(\rho') D_{x_0} v^0 \\ \qquad\qquad + H_{pp}^0(x_0, D_{x_0} U^0, m) D_{x_0} w^0 \Bigg) = 0 \quad \text{in } (0, T) \times \mathbb{R}^{d_0}, \\ w^0(T, x_0; m) = \dfrac{\delta^2 G^0}{\delta m^2}(x_0, m)(\rho, \rho'), \ w(T, x_0; x, m) = \dfrac{\delta^2 G}{\delta m^2}(x_0, x, m)(\rho, \rho') \text{ in } \mathbb{R}^{d_0}, \end{cases} \tag{37}$$

where (v^0, v), $((v')^0, v')$ are the solutions to (36) associated with ρ and ρ' respectively. Moreover, if

$$\left\| \frac{\delta^2 (G^0, G)}{\delta m^2} \right\|_{n-2;k-1,k-1} + \mathrm{Lip}^{x_0}_{n-3;k-2,k-2} \left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2} \right) \leq M,$$

then there exist $T_M, C_M > 0$ such that, for any $T \in (0, T_M)$,

$$\sup_t \left(\left\| \frac{\delta^2 (U^0, U)}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} + \mathrm{Lip}^{x_0}_{n-3;k-2,k-2} \left(\frac{\delta^2 U^0}{\delta m^2}, \frac{\delta^2 U}{\delta m^2} \right)(t) \right) \leq \left\| \frac{\delta^2 (G^0, G)}{\delta m^2} \right\|_{n-2;k-1,k-1} + \mathrm{Lip}^{x_0}_{n-3;k-2,k-2} \left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2} \right) + C_M T.$$

Proof. The differentiability of $\frac{\delta U^0}{\delta m}$ and of $\frac{\delta U}{\delta m}$ and the representation formula (37) can be established as for U^0 and U in Proposition 4.2. To prove the estimate, we use Proposition

A.9 with

$$V^0(t,x^0;m) := H_p^0(x_0, D_{x_0}U^0(t,x_0,m),m),$$
$$V(t,x^0x;m) := H_{pp}^0(x_0, D_{x_0}U^0(t,x_0,m),m)D_{x_0}U(t,x_0,x),$$

which are bounded of class C_b^1 and $C_b^{0,n-1} \cap C_b^{1,n-2}$ respectively, while the source terms

$$f^0(t,x^0;m) := \frac{\delta^2 H^0}{\delta m^2}(x_0, D_{x_0}U^0, m)(\rho,\rho') + H_{pp}^0(x_0, D_{x_0}U^0, m)D_{x_0}v^0 \cdot D_{x_0}(v')^0 + \frac{\delta H_p^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho) \cdot D_{x_0}(v')^0 + \frac{\delta H_p^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho') \cdot D_{x_0}v^0$$

and

$$\begin{aligned} f(t,x_0,x;m) :=\ & D_{x_0}v \cdot \left(\frac{\delta H_p^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho') + H_{pp}^0(x_0, D_{x_0}U^0, m)D_{x_0}(v')^0\right) \\ & + D_{x_0}v' \cdot \left(\frac{\delta H_p^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho) + H_{pp}^0(x_0, D_{x_0}U^0, m)D_{x_0}v^0\right) \\ & + D_{x_0}U \cdot \left(\frac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho)D_{x_0}(v')^0 + \frac{\delta^2 H_p^0}{\delta m^2}(x_0, D_{x_0}U^0, m)(\rho,\rho')\right. \\ & \left. + H_{ppp}^0(x_0, D_{x_0}U^0, m)D_{x_0}v^0 D_{x_0}(v')^0 + \frac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0}U^0, m)(\rho')D_{x_0}v^0\right) \end{aligned}$$

are in C_b^0 and $C_b^{0,n-2}$ respectively, thanks to Propositions 4.1 and 4.2. By Proposition A.9, with $r = 0$ and $n-2$ we obtain the estimates for $\|\frac{\delta^2(U^0,U)}{\delta m^2}\|_{n-2;k-1,k-1}$. The Lipschitz bound in x_0 of $(\frac{\delta^2 U^0}{\delta m^2}, \frac{\delta^2 U}{\delta m^2})$ follows analogously. ■

4.2.4. Lipschitz regularity of second order derivatives. We finally address the Lipschitz regularity of second order derivatives of U^0 and U with respect to m and x_0.

Proposition 4.4. *Under the assumptions of Proposition 4.3 and if, in addition,*

$$\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2}\right) \le M,$$

then

$$\sup_t \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U^0}{\delta m^2}, \frac{\delta^2 U}{\delta m^2}\right)(t) \le \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2}\right) + C_M T,$$

where the constant C_M depends on the regularity of H and H^0 and on M.

Proof. Let $(x,\rho,\rho') \in \mathbb{R}^d \times C^{-(k-2)} \times C^{-(k-2)}$, $m^1, m^2 \in \mathcal{P}_2$, $(U^{0,1}, U^1)$ be the solution to (35) associated with (x, m^1), and $(U^{0,2}, U^2)$ be the solution associated with (x, m^2). We denote by $(v^{0,1}, v^1)$, $((v')^{0,1}, (v')^1)$ (resp. $(v^{0,2}, v^2)$, $((v')^{0,2}, (v')^2)$) the corresponding solutions to the first order linearized system (36) associated with ρ and ρ', and by $(w^{0,1}, w^1)$ (resp. $(w^{0,2}, w^2)$) the corresponding solution of the second order linearized system (37). We want to estimate the difference $(z^0, z) := (w^{0,2} - w^{0,1}, w^2 - w^1)$. We have

$$\begin{cases} -\partial_t z^0 - \Delta_{x_0} z^0 + H_p^0(x_0, D_{x_0} U^{0,1}(t, x_0, m^1), m^1) \cdot D_{x_0} z^0 + f^0 = 0, \\ -\partial_t z - \Delta_{x_0} z + D_{x_0} z \cdot H_p^0(x_0, D_{x_0} U^{0,1}, m^1) \\ \qquad\qquad - H_{pp}^0(x_0, D_{x_0} U^{0,1}, m) D_{x_0} U^1 \cdot D_{x_0} z^0 + f = 0, \\ z^0(T) = \dfrac{\delta^2 G^0}{\delta m^2}(x_0, m^2)(\rho, \rho') - \dfrac{\delta^2 G^0}{\delta m^2}(x_0, m^1)(\rho, \rho'), \\ z(T) = \dfrac{\delta^2 G}{\delta m^2}(x_0, x, m^2)(\rho, \rho') - \dfrac{\delta^2 G}{\delta m^2}(x_0, x, m^1)(\rho, \rho'), \end{cases}$$

where

$$\begin{aligned} f^0 := {} & (H_p^0(x_0, D_{x_0} U^{0,2}, m^2) - H_p^0(x_0, D_{x_0} U^{0,1}, m^1)) \cdot D_{x_0} w^{0,2} \\ & + \frac{\delta^2 H^0}{\delta m^2}(x_0, D_{x_0} U^{0,2}, m^2)(\rho, \rho') - \frac{\delta^2 H^0}{\delta m^2}(x_0, D_{x_0} U^{0,1}, m^1)(\rho, \rho') \\ & + H_{pp}^0(x_0, D_{x_0} U^{0,2}, m^2) D_{x_0} v^{0,2} \cdot D_{x_0} (v')^{0,2} \\ & - H_{pp}^0(x_0, D_{x_0} U^{0,1}, m^1) D_{x_0} v^{0,1} \cdot D_{x_0} (v')^{0,1} \\ & + \frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,2}, m^2)(\rho) \cdot D_{x_0} (v')^{0,2} - \frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,1}, m^1)(\rho) \cdot D_{x_0} (v')^{0,1} \\ & + \frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,2}, m^2)(\rho') \cdot D_{x_0} v^{0,2} - \frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,1}, m^1)(\rho') \cdot D_{x_0} v^{0,1} \end{aligned}$$

and

$$\begin{aligned} f := {} & D_{x_0} w^2 \cdot (H_p^0(x_0, D_{x_0} U^{0,2}, m^2) - H_p^0(x_0, D_{x_0} U^{0,1}, m^1)) \\ & + D_{x_0} v^2 \cdot \left(\frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,2}, m^2)(\rho') + H_{pp}^0(x_0, D_{x_0} U^{0,2}, m^2) D_{x_0} (v')^{0,2} \right) \\ & - D_{x_0} v^1 \cdot \left(\frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,1}, m^1)(\rho') + H_{pp}^0(x_0, D_{x_0} U^{0,1}, m^1) D_{x_0} (v')^{0,1} \right) \\ & + D_{x_0} (v')^2 \cdot \left(\frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,2}, m^2)(\rho) + H_{pp}^0(x_0, D_{x_0} U^{0,2}, m^2) D_{x_0} v^{0,2} \right) \\ & - D_{x_0} (v')^1 \cdot \left(\frac{\delta H_p^0}{\delta m}(x_0, D_{x_0} U^{0,1}, m^1)(\rho) + H_{pp}^0(x_0, D_{x_0} U^{0,1}, m^1) D_{x_0} v^{0,1} \right) \\ & + D_{x_0} U^2 \cdot \left(\frac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0} U^{0,2}, m^2)(\rho) D_{x_0} (v')^{0,2} + \frac{\delta^2 H_p^0}{\delta m^2}(x_0, D_{x_0} U^{0,2}, m^2)(\rho, \rho') \right. \\ & \left. + H_{ppp}^0(x_0, D_{x_0} U^{0,2}, m^2) D_{x_0} v^{0,2} D_{x_0} (v')^{0,2} + \frac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0} U^{0,2}, m^2)(\rho') D_{x_0} v^{0,2} \right) \\ & + \left(H_{pp}^0(x_0, D_{x_0} U^{0,2}, m) D_{x_0} U^2 - H_{pp}^0(x_0, D_{x_0} U^{0,1}, m) D_{x_0} U^1 \right) \cdot D_{x_0} w^{0,2} \\ & - D_{x_0} U^1 \cdot \left(\frac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0} U^{0,1}, m^1)(\rho) D_{x_0} (v')^{0,1} + \frac{\delta^2 H_p^0}{\delta m^2}(x_0, D_{x_0} U^{0,1}, m^1)(\rho, \rho') \right. \\ & \left. + H_{ppp}^0(x_0, D_{x_0} U^{0,1}, m^1) D_{x_0} v^{0,1} D_{x_0} (v')^{0,1} + \frac{\delta H_{pp}^0}{\delta m}(x_0, D_{x_0} U^{0,1}, m^1)(\rho') D_{x_0} v^{0,1} \right). \end{aligned}$$

Proposition 4.2 (for the representation of $(v^{0,i}, v^i)$) and Proposition 4.3 (for their Lipschitz regularity in m and in x_0) imply in particular that

$$\sup_t \big(\| D_{x_0}(v^{0,2} - v^{0,1}) \|_\infty + \| D_{x_0}(v^2 - v^1) \|_{0,n-3} \big) \le C \mathbf{d}_2(m_1, m_2)$$

and hence we have, using also Proposition 4.3,

$$\sup_t (\| f^0 \|_\infty + \| f \|_{0,n-3}) \le C \mathbf{d}_2(m^1, m^2).$$

Using Proposition A.9 (with $r = 0$), we obtain, for any $l \le n-3$,

$$\begin{aligned} &\sup_{t,x_0,x} (|z^0(t,x_0)|^2 + |D_x^l z(t,x_0,x)|^2)^{1/2} \\ &\le (1+CT) \sup_{x_0,x} \Bigg(\left| \frac{\delta^2 G^0}{\delta m^2}(x_0, m^2)(\rho,\rho') - \frac{\delta^2 G^0}{\delta m^2}(x_0, m^1)(\rho,\rho') \right|^2 \\ &\quad + \left| D_x^l \frac{\delta^2 G}{\delta m^2}(x_0, x, m^2)(\rho,\rho') - D_x^l \frac{\delta^2 G}{\delta m^2}(x_0, x, m^1)(\rho,\rho') \right|^2 \Bigg)^{1/2} + CT \mathbf{d}_2(m^1, m^2), \end{aligned}$$

which gives the claim. ■

We complete this section by stating similar estimates on the Lipschitz regularity of the other second order derivatives:

Proposition 4.5. *Under the assumptions of Proposition 4.3 and if, in addition,*

$$\mathrm{Lip}_{n-3;k-2}\left(\frac{\delta G^0_{x_0}}{\delta m}, \frac{\delta G_{x_0}}{\delta m} \right) + \mathrm{Lip}_{n-3}(D^2_{x_0} G^0, D^2_{x_0} G) \le M,$$

then

$$\begin{aligned} \sup_t \mathrm{Lip}_{n-3;k-2}\left(\frac{\delta U^0_{x_0}}{\delta m}, \frac{\delta U_{x_0}}{\delta m} \right)(t) &\le \mathrm{Lip}_{n-3;k-2}\left(\frac{\delta G^0_{x_0}}{\delta m}, \frac{\delta G_{x_0}}{\delta m} \right) + C_M T, \\ \sup_t \mathrm{Lip}_{n-3}(D^2_{x_0} U^0, D^2_{x_0} U)(t) &\le \mathrm{Lip}_{n-3}(D^2_{x_0} G^0, D^2_{x_0} G) + C_M T, \end{aligned}$$

where the constant C_M depends on the regularity of H and H^0 and on M.

As the proof is completely similar to the proof of Proposition 4.4, we omit it.

4.3. Existence of a solution

4.3.1. Definition of the semi-discrete scheme. Let us fix some horizon $T > 0$ (small) and a step size $\tau := T/(2N)$ (where $N \in \mathbb{N}$, $N \ge 1$). We set $t_k = kT/(2N)$, $k \in \{0, 2N\}$. We define by backward induction the continuous maps $U^{0,N} = U^{0,N}(t, x_0, m)$ and $U^N = U^N(t, x_0, x, m)$ as follows: we require that

(i) $(U^{0,N}, U^N)$ satisfies the terminal condition:

$$U^{0,N}(T, x_0, m) = G^0(x_0, m), \quad U^N(T, x_0, x, m) = G(x_0, x, m)$$

for all $(x_0, x, m) \in \mathbb{R}^d \times \mathbb{R}^{d_0} \times \mathcal{P}_2$,

(ii) for $x_0 \in \mathbb{R}^{d_0}$ fixed, $(U^{0,N}, U^N)$ solves the backward system of first order master equations:

$$\begin{cases} -\partial_t U^0 - 2\int_{\mathbb{R}^d} \operatorname{div}_y D_m U^0(t, x_0, m, y)\, m(dy) \\ \quad + 2\int_{\mathbb{R}^d} D_m U^0(t, x_0, m, y) \cdot H_p(x_0, y, D_x U(t, x_0, y, m), m)\, m(dy) = 0, \\ -\partial_t U - 2\Delta_x U + 2H(x_0, x, D_x U, m) - 2\int_{\mathbb{R}^d} \operatorname{div}_y D_m U(t, x_0, x, m, y)\, m(dy) \\ \quad + 2\int_{\mathbb{R}^d} D_m U(t, x_0, x, m, y) \cdot H_p(x_0, y, D_x U(t, x_0, y, m), m)\, m(dy) = 0 \end{cases} \tag{38}$$

on time intervals of the form (t_{2j+1}, t_{2j+2}) for $j = 0, \ldots, N-1$,

(iii) for $(x, m) \in \mathbb{R}^d \times \mathcal{P}_2$ fixed, $(U^{0,N}, U^N)$ solves the backward system of HJ equations

$$\begin{cases} -\partial_t U^0 - 2\Delta_{x_0} U^0 + 2H^0(x_0, D_{x_0} U^0, m) = 0, \\ -\partial_t U - 2\Delta_{x_0} U + 2D_{x_0} U \cdot H_p^0(x_0, D_{x_0} U^0(t, x_0, m), m) = 0 \end{cases} \tag{39}$$

on time intervals of the form (t_{2j}, t_{2j+1}) for $j = 0, \ldots, N-1$.

Our aim is to show that if the time horizon is short enough, $(U^{0,N}, U^N)$ converges to a solution of the master equation for MFGs with a major player as $N \to +\infty$.

4.3.2. Proof of the existence of a solution. For $n \geq 4$ and $k \in \{3, \ldots, n-1\}$, let

$$\begin{aligned} M := {} & 1 + \|(G^0, G)\|_n + \|D_{x_0}(G^0, G)\|_{n-1} + \|D_{x_0}^2(G^0, G)\|_{n-2} \\ & + \mathrm{Lip}_{n-3}^{x_0}(D_{x_0}^2 G^0, D_{x_0}^2 G) + \left\| \frac{\delta(G^0, G)}{\delta m} \right\|_{n-1;k} + \left\| \frac{\delta(G_{x_0}^0, G_{x_0})}{\delta m} \right\|_{n-2;k-1} \\ & + \mathrm{Lip}_{n-3;k-2}^{x_0}\left(\frac{\delta G_{x_0}^0}{\delta m}, \frac{\delta G_{x_0}}{\delta m} \right) + \left\| \frac{\delta^2(G^0, G)}{\delta m^2} \right\|_{n-2;k-1,k-1} \\ & + \mathrm{Lip}_{n-3;k-2,k-2}^{x_0}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2} \right) + \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2} \right) \\ & + \mathrm{Lip}_{n-3;k-2}\left(\frac{\delta G_{x_0}^0}{\delta m}, \frac{\delta G_{x_0}}{\delta m} \right) + \mathrm{Lip}_{n-3}(D_{x_0}^2 G^0, D_{x_0}^2 G). \end{aligned}$$

Lemma 4.6. *There exists $T_M > 0$, depending on the regularity of H^0, H and on M, such that, for any $T \in (0, T_M]$ and $N \geq 1$, we have, for any $t \in [0, T]$,*

$$\begin{aligned}
&\|(U^0, U)(t)\|_n + \|D_{x_0}(U^0, U)(t)\|_{n-1} + \|D^2_{x_0}(U^0, U)(t)\|_{n-2} \\
&\quad + \mathrm{Lip}^{x_0}_{n-3}((D^2_{x_0}U^0, D^2_{x_0}U)(t)) + \left\| \frac{\delta(U^0, U)}{\delta m}(t) \right\|_{n-1;k} + \left\| \frac{\delta(U^0_{x_0}, U_{x_0})}{\delta m}(t) \right\|_{n-2;k-1} \\
&\quad + \mathrm{Lip}^{x_0}_{n-3;k-2}\left(\left(\frac{\delta U^0_{x_0}}{\delta m}, \frac{\delta U_{x_0}}{\delta m} \right)(t) \right) + \left\| \frac{\delta^2(U^0, U)}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} \\
&\quad + \mathrm{Lip}^{x_0}_{n-3;k-2,k-2}\left(\left(\frac{\delta^2 U^0}{\delta m^2}, \frac{\delta^2 U}{\delta m^2} \right)(t) \right) + \mathrm{Lip}_{n-3;k-2,k-2}\left(\left(\frac{\delta^2 U^0}{\delta m^2}, \frac{\delta^2 U}{\delta m^2} \right)(t) \right) \\
&\quad + \mathrm{Lip}_{n-3;k-2}\left(\left(\frac{\delta U^0_{x_0}}{\delta m}, \frac{\delta U_{x_0}}{\delta m} \right)(t) \right) + \mathrm{Lip}_{n-3}((D^2_{x_0}U^0, D^2_{x_0}U)(t)) \leq M. \qquad (40)
\end{aligned}$$

Moreover:

- *The maps $U^{0,N}$ and U^N are globally Lipschitz continuous in all variables and their first and second space derivatives are globally Hölder continuous in all variables, uniformly with respect to N.*
- *The maps $D_m U^{0,N}$ and $D_m U^N$ are Hölder continuous in (t, x_0, m, y) and (t, x_0, x, m, y) respectively, uniformly with respect to N, in any set of the form*

$$\begin{aligned}
&\{(t, x_0, m, y) \in [0, T] \times \mathbb{R}^{d_0} \times \mathcal{P}_2 \times \mathbb{R}^d : M_2(m) \leq R,\ |y| \leq R\}, \\
&\{(t, x_0, x, m, y) \in [0, T] \times \mathbb{R}^{d_0} \times \mathbb{R}^d \times \mathcal{P}_2 \times \mathbb{R}^d : M_2(m) \leq R,\ |y| \leq R\}
\end{aligned} \qquad (41)$$

respectively, where $M_2(m) = (\int_{\mathbb{R}^d} |y|^2 m(dy))^{1/2}$.

Proof. We only sketch the proof, since it is exactly the same as for the second order master equation (see Lemma 3.5). The proof of (40) can be established by collecting the estimates in Propositions 5.15, 5.17 and 5.19 in Section 5.2 below, which provide the bounds on intervals of the form (t_{2j+1}, t_{2j+2}), and, for the intervals of the form (t_{2j}, t_{2j+1}), by Propositions 4.1–4.4.

The Lipschitz regularity in space of $U^{0,N}$ and U^N and of their first and second order space derivatives follows immediately from (40). As $D_m U^{0,N}$ and $D_m U^N$ are bounded according to (40), $U^{0,N}$ and U^N and their first and second order space derivatives are also Lipschitz continuous in m. Finally, since $U^{0,N}$ and U^N satisfy (38) and (39), the bounds in (40) show that $\partial_t U^{0,N}$ and $\partial_t U^N$ are bounded and therefore $U^{0,N}$ and U^N are also Lipschitz continuous in time. The global Hölder regularity of the first and second space derivatives of $U^{0,N}$ and U^N then follows by interpolation (Lemma B.2).

The Lipschitz regularity in space and in measure of $D_m U^{0,N}$ and $D_m U^N$ is a consequence of (40), while the Hölder regularity in time in sets of the form (41) comes from interpolation (Lemma B.4). ∎

Proof of Theorem 2.5. The argument is exactly the same as in the proof of Theorem 2.3 and we omit it. ∎

4.4. Uniqueness of the solution

We finally address the uniqueness of the solution of the master equation for MFGs with a major player:

Theorem 4.7. *Let* $(U^{0,1}, U^1)$ *and* $(U^{0,2}, U^2)$ *be two classical solutions to* (22) *defined on the time interval* $[0, T]$ *and such that* $D_{x_0} U^{0,1}$ *and* $D_{x_0,x} U^1$ *are uniformly Lipschitz continuous in the space and measure variables. Then* $(U^{0,1}, U^1) = (U^{0,2}, U^2)$.

Proof. Let $(t_0, \bar{x}_0, \bar{m}_0) \in [0, T) \times \mathbb{R}^{d_0} \times \mathcal{P}_2$ be an initial condition, Z a random variable with law $\bar{m}_0$, and (X_t^0, m_t, X_t) the solution to

$$\begin{cases} dX_t^0 = -H_p^0(X_t^0, D_{x_0} U^{0,1}(t, X_t^0, m_t), m_t)\, dt + \sqrt{2}\, dW_t^0 \quad \text{in } (0, T), \\ dm_t = (\Delta m_t + \operatorname{div}(m_t H_p(X_t^0, x, D_x U^1(t, X_t^0, x, m_t), m_t)))\, dt \quad \text{in } (0, T) \times \mathbb{R}^d, \\ dX_t = -H_p(X_t^0, X_t, D_x U^1(t, X_t^0, X_t, m_t)\, dt + \sqrt{2}\, dW_t \quad \text{in } (0, T), \\ X_{t_0}^0 = \bar{x}_0, \quad m_{t_0} = \bar{m}_0, \quad X_{t_0} = Z, \end{cases}$$

where (W_t^0) and (W_t) are Brownian motions, (W_t^0), (W_t) and Z being independent. As $D_x U^{0,1}$ and $D_x U^1$ are globally Lipschitz continuous, the above system has a unique solution. Note that m_t is the conditional law of X_t given $(W_s^0)_{s \leq t}$.

We compute the variation of $U^{0,1}$ along (t, X_t^0, m_t):

$$\begin{aligned} dU^{0,1}(t, X_t^0, m_t) = \Big(& \partial_t U^{0,1} + \Delta_{x_0} U^{0,1} - H_p^0(X_t^0, D_{x_0} U^{0,1}, m_t) \cdot D_{x_0} U^{0,1} \\ & - \int_{\mathbb{R}^d} D_m U^{0,1} \cdot H_p(X_t^0, y, D_x U^1(t, X_t^0, y, m_t), m_t)\, m_t(dy) \\ & + \int_{\mathbb{R}^d} \operatorname{div}_y D_m U^{0,1}\, m_t(dy) \Big) dt + \sqrt{2}\, D_{x_0} U^{0,1} \cdot dW_t^0, \end{aligned}$$

where, unless specified otherwise, $U^{0,1}$ and its space derivatives are computed at (t, X_t^0, m_t) while $D_m U^{0,1}$ and its space derivatives are computed at (t, X_t^0, m_t, y). In view of the equation satisfied by $U^{0,1}$, we find

$$\begin{aligned} dU^{0,1}(t, X_t^0, m_t) = & \big(H^0(X_t^0, D_{x_0} U^{0,1}, m_t) - H_p^0(X_t^0, D_{x_0} U^{0,1}, m_t) \cdot D_{x_0} U^{0,1}\big)\, dt \\ & + \sqrt{2}\, D_{x_0} U^{0,1} \cdot dW_t^0. \end{aligned}$$

We proceed in the same way for $U^{0,2}$ and obtain, in view of the equation satisfied by $U^{0,2}$,

$$\begin{aligned} dU^{0,2}(t, X_t^0, m_t) = \Big(& H^0(X_t^0, D_{x_0} U^{0,2}, m_t) - H_p^0(X_t^0, D_{x_0} U^{0,1}, m_t) \cdot D_{x_0} U^{0,2} \\ & + \int_{\mathbb{R}^d} D_m U^{0,2} \cdot \big(H_p(X_t^0, y, D_x U^2(t, X_t^0, y, m_t), m_t) \\ & - H_p(X_t^0, y, D_x U^1(t, X_t^0, y, m_t)), m_t\big)\, m_t(dy) \Big) dt + \sqrt{2}\, D_{x_0} U^{0,2} \cdot dW_t^0, \end{aligned}$$

where, unless specified otherwise, $U^{0,2}$ and its space derivatives are computed at (t, X_t^0, m_t) while $D_m U^{0,2}$ and its space derivatives are computed at (t, X_t^0, m_t, y). Therefore

$$
\begin{aligned}
d(U^{0,2} - U^{0,1})^2 = 2(U^{0,2} - U^{0,1})\Big(& H^0(X_t^0, D_{x_0}U^{0,2}, m_t) - H^0(X_t^0, D_{x_0}U^{0,1}, m_t) \\
& - H_p^0(X_t^0, D_{x_0}U^{0,1}, m_t) \cdot (D_{x_0}U^{0,2} - D_{x_0}U^{0,1}) \\
& + \int_{\mathbb{R}^d} D_m U^{0,2} \cdot (H_p(X_t^0, y, D_x U^2(t, X_t^0, y, m_t), m_t) \\
& - H_p(X_t^0, y, D_x U^1(t, X_t^0, y, m_t), m_t))\, m_t(dy)\Big)\, dt \\
+ 2(D_{x_0}U^{0,2} - D_{x_0}U^{0,1})^2\, dt & + 2\sqrt{2}\,(U^{0,2} - U^{0,1})(D_{x_0}U^{0,2} - D_{x_0}U^{0,1}) \cdot dW_t^0.
\end{aligned}
$$

Let us set $U_t^{0,i} = U^{0,i}(t, X_t^0, m_t)$ (for $i = 1, 2$). We integrate in time between $s \in [t_0, T]$ and T, take expectation and use the fact that $U_T^{0,1} = U_T^{0,2} = G^0(X_T^0, m_T)$:

$$
\begin{aligned}
0 = \mathbb{E}\Big[& (U_s^{0,2} - U_s^{0,1})^2 \\
& + \int_s^T 2(U_t^{0,2} - U_t^{0,1})\Big(H^0(X_t^0, D_{x_0}U^{0,2}, m_t) - H^0(X_t^0, D_{x_0}U^{0,1}, m_t) \\
& - H_p^0(X_t^0, D_{x_0}U^{0,1}, m_t) \cdot (D_{x_0}U^{0,2} - D_{x_0}U^{0,1}) \\
& + \int_{\mathbb{R}^d} D_m U^{0,2} \cdot (H_p(X_t^0, y, D_x U^2(t, X_t^0, y, m_t), m_t) \\
& - H_p(X_t^0, y, D_x U^1(t, X_t^0, y, m_t), m_t))\, m_t(dy)\Big)\, dt \\
& + 2\int_s^T |D_{x_0}U^{0,2} - D_{x_0}U^{0,1}|^2\, dt\Big].
\end{aligned}
$$

Thanks to the regularity of the solutions, by the Cauchy–Schwarz inequality and for any $\epsilon > 0$ we have

$$
\begin{aligned}
0 \geq \mathbb{E}\Big[& (U_s^{0,2} - U_s^{0,1})^2 - \int_s^T \Big(C_\epsilon (U_t^{0,2} - U_t^{0,1})^2 + \epsilon |D_{x_0}(U^{0,2} - U^{0,1})|^2 \\
& + \epsilon \int_{\mathbb{R}^d} |D_x(U^2 - U^1)(t, X_t^0, y, m_t)|^2\, m_t(dy)\Big)\, dt + 2\int_s^T |D_{x_0}(U^{0,2} - U^{0,1})|^2\, dt\Big].
\end{aligned}
$$

So, for ϵ small enough, we obtain

$$
\begin{aligned}
0 \geq \mathbb{E}\Big[& (U_s^{0,2} - U_s^{0,1})^2 - \int_s^T \Big(C_\epsilon (U_t^{0,2} - U_t^{0,1})^2 \\
& + \epsilon \int_{\mathbb{R}^d} |D_x(U^2 - U^1)(t, X_t^0, y, m_t)|^2\, m_t(dy)\Big)\, dt + \int_s^T |D_{x_0}(U^{0,2} - U^{0,1})|^2\, dt\Big].
\end{aligned}
$$

We argue in the same way for $U_t^i := U^i(t, X_t^0, X_t, m_t)$ $(i = 1, 2)$ and find that

$$0 \geq \mathbb{E}\Big[(U_s^2 - U_s^1)^2 - \int_s^T \Big(C_\epsilon (U_t^2 - U_t^1)^2 + \epsilon |D_{x_0}(U^{0,2} - U^{0,1})|^2 + \epsilon \int_{\mathbb{R}^d} |D_x(U^2 - U^1)(t, X_t^0, y, m_t)|^2 \, m_t(dy)\Big) dt + \int_s^T (|D_{x_0}(U^2 - U^1)|^2 + |D_x(U^2 - U^1)|^2) \, dt\Big].$$

We add the last two inequalities to obtain

$$0 \geq \mathbb{E}\Big[(U_s^{0,2} - U_s^{0,1})^2 + (U_s^2 - U_s^1)^2 - \int_s^T \Big(C_\epsilon ((U_t^{0,2} - U_t^{0,1})^2 + (U_t^2 - U_t^1)^2) + \epsilon |D_{x_0}(U^{0,2} - U^{0,1})|^2 + 2\epsilon \int_{\mathbb{R}^d} |D_x(U^2 - U^1)(t, X_t^0, y, m_t)|^2 \, m_t(dy)\Big) dt + \int_s^T \big(|D_{x_0}(U^{0,2} - U^{0,1})|^2 + |D_{x_0}(U^2 - U^1)|^2 + |D_x(U^2 - U^1)|^2\big) \, dt\Big]. \tag{42}$$

Note that, as m_t is the conditional law of X_t given $(W_u^0)_{u \leq t}$, we have

$$\mathbb{E}[|D_x(U^2 - U^1)(t, X_t^0, X_t, m_t)|^2] = \mathbb{E}\big[\mathbb{E}[|D_x(U^2 - U^1)(t, X_t^0, X_t, m_t)|^2 \,|\, (W_u^0)_{u \leq t}]\big] = \mathbb{E}\Big[\int_{\mathbb{R}^d} |D_x(U^2 - U^1)(t, X_t^0, y, m_t)|^2 \, m_t(dy)\Big]$$

since X_t^0 and X_t are adapted to $(W_u^0)_{u \leq t}$. Plugging this relation into (42) we find therefore, for $\epsilon > 0$ small enough,

$$0 \geq \mathbb{E}\Big[(U_s^{0,2} - U_s^{0,1})^2 + (U_s^2 - U_s^1)^2 - \int_s^T C_\epsilon ((U_t^{0,2} - U_t^{0,1})^2 + (U_t^2 - U_t^1)^2) \, dt + \frac{1}{2} \int_s^T \big(|D_{x_0}(U^{0,2} - U^{0,1})|^2 + |D_{x_0}(U^2 - U^1)|^2 + |D_x(U^2 - U^1)|^2\big) \, dt\Big].$$

We conclude by Gronwall's inequality that, for any $t \in [t_0, T]$,

$$\mathbb{E}[(U^{0,2}(t, X_t^0, m_t) - U^{0,1}(t, X_t^0, m_t))^2 + (U^2(t, X_t^0, X_t, m_t) - U^1(t, X_t^0, X_t m_t))^2] = 0.$$

For $t = t_0$, we therefore have $U^{0,2}(t_0, \bar{x}_0, \bar{m}_0) = U^{0,1}(t_0, \bar{x}_0, \bar{m}_0)$ and

$$U^1(t_0, \bar{x}_0, Z, \bar{m}_0) = U^2(t_0, \bar{x}_0, Z, \bar{m}_0) \quad \text{a.s.}$$

If $\bar{m}_0$ has a positive density, the fact that the law of Z is $\bar{m}_0$ easily implies the equality of U^1 and U^2 at any point $(t_0, \bar{x}_0, x, \bar{m}_0)$ for $x \in \mathbb{R}^d$. We conclude by density of such laws and by continuity of the U^i's. ■

5. Analysis of the first order master equations

In this section, we complete our program by proving regularity results for the solutions of the various first order master equations encountered in the previous sections. We mainly consider the first order master equation

$$\begin{cases} -\partial_t U(t,x_0,x,m) - \mathrm{Tr}(a(t,x)D^2_{xx}U(t,x_0,x,m)) + H(x_0,x,D_xU(t,x_0,x,m),m) \\ \quad - \int_{\mathbb{R}^d} \mathrm{Tr}(a(t,y)D^2_{ym}U(t,x_0,x,m,y))\, m(dy) \\ \quad + \int_{\mathbb{R}^d} D_mU(t,x_0,x,m,y)\cdot H_p(x_0,y,D_xU(t,x_0,y,m),m)\, m(dy) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2, \\ U(T,x_0,x,m) = G(x_0,x,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2. \end{cases} \tag{43}$$

In the above equation, $x_0 \in \mathbb{R}^{d_0}$ is considered as a parameter. Our aim is to build a solution to this equation and study its regularity. The method for finding a solution to (43) is well-known, and is based on looking at its characteristics: if we set

$$U(t_0,x_0,x,m_0) := u(t_0,x) \tag{44}$$

where (u,m) is the solution to the MFG system

$$\begin{cases} -\partial_t u(t,x) - \mathrm{Tr}(a(t,x)D^2u(t,x)) + H(x_0,x,Du(t,x),m(t)) = 0 \quad \text{in } (t_0,T)\times\mathbb{R}^d, \\ \partial_t m(t,x) - \sum_{i,j} D_{ij}(a_{i,j}(t,x)m(t,x)) - \mathrm{div}(m(t,x)H_p(x_0,x,Du(t,x),m(t))) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (t_0,T)\times\mathbb{R}^d, \\ m(t_0) = m_0, \quad u(T,x) = G(x_0,x,m(T)) \quad \text{in } \mathbb{R}^d \end{cases} \tag{45}$$

(here $x_0 \in \mathbb{R}^{d_0}$ is again treated as a fixed parameter), then U is a solution to (43).

In order to study the Major–Minor agents' problem, we also have to consider a linear master equation

$$\begin{cases} -\partial_t U^0 - \int_{\mathbb{R}^d} \mathrm{Tr}(a(t,y)D^2_{ym}U^0(t,x_0,m,y))\, m(dy) \\ \quad + \int_{\mathbb{R}^d} D_mU^0(t,x_0,m,y)\cdot H_p(x_0,y,D_xU(t,x_0,y,m),m)\, m(dy) = 0, \\ U^0(T,x_0,m) = G^0(x_0,m) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2, \end{cases} \tag{46}$$

where U is the solution to (43). In this case, we build the solution U^0 by the simple formula

$$U^0(t_0,x_0,m_0) = G^0(x_0,m(T)), \tag{47}$$

where (u,m) is also the solution to (45).

Our aim is to show that if G and G^0 are regular enough, then (43) and (46) have classical solutions, given by the above representation formulas. Moreover, we show that the regularity of these solutions only deteriorates linearly in time. This last point is the key result in order to build later solutions to the second order master equation and to the master equation for the Major–Minor agents' problem.

To guide the reader, the plan of the section is as follows: Section 5.1 is devoted to the study of the regularity of the MFG system (45), together with its linearizations, in particular:

- Basic estimates on (u, m) solving the *MFG system* are given in Section 5.1.1. Note that these heavily rely on technical bounds for Hamilton–Jacobi equations, that will be proven in Appendix A.
- Estimates on the *first order linearized system* are given in Section 5.1.2, and these are again based on results in Appendix A.
- Estimates on the *second order linearized system* are given in Section 5.1.3, and their proofs basically follow the scheme of the first order linearized system.

Then, we will use these regularity results to obtain bounds on solutions to the master equations in Section 5.2:

- Basic estimates and first order differentiability of U and U^0 are shown in Section 5.2.1. These are consequences of estimates for the MFG system (in Section 5.1.1) and its first order linearization (in Section 5.1.2). Note that here we also need a general criterion for differentiability of functions depending on measures (Lemma B.1).
- Second order differentiability of U and U^0 is established in Section 5.2.2. Here we will need several bounds for the MFG system, its first and second order linearizations (Sections 5.1.1–5.1.3).
- Finally, Section 5.2.3 is devoted to uniform continuity estimates on second order derivatives, that are crucial to obtain compactness for the Trotter–Kato scheme. Again, we rely on estimates on the linearized system (Sections 5.1.2–5.1.3).

A complete roadmap of Section 5 is given in Figure 1.

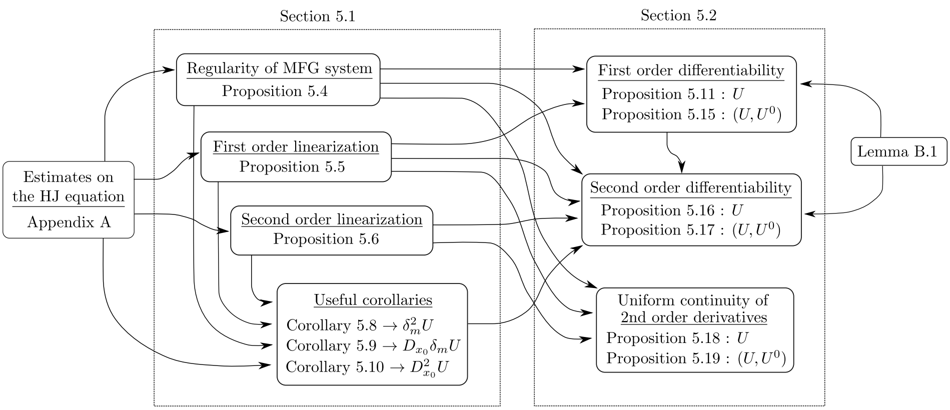

Fig. 1. A roadmap of Section 5.

5.1. Estimates on the MFG system

Let us first explain the notion of solution to (45). Fix $(t_0, m_0) \in [0,T] \times \mathcal{P}_2$ and $x_0 \in \mathbb{R}^{d_0}$. We say that (u, m) is a *solution* to (45) if $u \in C^0([t_0,T], C_b^2)$ satisfies

$$u(t,x) = G(x_0, x, m(T)) + \int_t^T \big(\mathrm{Tr}(a(s,x) D^2 u(s,x)) - H(x_0, x, Du(s,x), m(s))\big)\, ds$$

for all $t \in [t_0, T]$ and if $m \in C^0([t_0,T], \mathcal{P}_2)$ solves the Fokker–Planck equation in the sense of distributions: for any $\phi \in C_c^\infty([0,T) \times \mathbb{R}^d)$,

$$0 = \int_{\mathbb{R}^d} \phi(0,x)\, m_0(dx) + \int_0^T \int_{\mathbb{R}^d} (\mathrm{Tr}(a(s,x) D^2 \phi(s,x)) - D\phi(s,x) \cdot H_p(x_0, x, Du(s,x), m(s)))\, m(s,dx)\, ds.$$

The assumptions on a, H and G given in Section 2.3 are in force throughout the section.

5.1.1. Well-posedness and regularity of the MFG system. We discuss here the well-posedness of the MFG system (45) and provide several estimates. Let us start with the Hamilton–Jacobi (HJ) equation (general estimates on this equation are given in Appendix A).

Proposition 5.1. *For any $M > 0$, there exist $T_M, L_M > 0$, depending on C_0 and γ given in assumptions (16) and (17), such that if $\sup_{x_0,m} \|G(x_0,\cdot,m)\|_1 \le M$, then, for any $T \in (0, T_M)$ and any $m \in C^0([0,T], \mathcal{P}_2)$, the solution u to the HJ equation*

$$\begin{cases} -\partial_t u(t,x) - \mathrm{Tr}(a(t,x) D^2 u(t,x)) + H(x_0, x, Du(t,x), m(t)) = 0 & \text{in } (t_0,T) \times \mathbb{R}^d, \\ u(T,x) = G(x_0, x, m(T)) & \text{in } \mathbb{R}^d \end{cases} \tag{48}$$

satisfies

$$\sup_{t \in [t_0,T]} \|u\|_1 \le \sup_{x_0,m} \|G(x_0,\cdot,m)\|_1 + L_M T.$$

Henceforth, we set $K_M := \sup_{x_0,m} \|G(x_0,\cdot,m)\|_1 + L_M T_M$.

If, in addition, $\sup_{x_0,m} \|G(x_0,\cdot,m)\|_n \le M$, then there exists $C_M > 0$, depending on n, C_0, γ and

$$\sup_{t \in [0,T_M]} \|a(t)\|_n + \sup_{|p| \le K_M,\, x_0 \in \mathbb{R}^{d_0},\, m \in \mathcal{P}_2} \sum_{k=0}^n \|D^k_{(x,p)} H(x_0,\cdot,p,m)\|_\infty,$$

such that, for any $T \in (0, T_M)$, $x_0 \in \mathbb{R}^{d_0}$ and $r \le n$,

$$\sup_{t \in [t_0,T],\, x \in \mathbb{R}^d} |D_x^r u(t,x)| \le \sup_{x \in \mathbb{R}^d} |D_x^r G(x_0, x, m(T))| + C_M T.$$

Therefore, for any $x_0 \in \mathbb{R}^{d_0}$,

$$\sup_{t\in[t_0,T]} \|u(t)\|_n \le \sup_m \|G(x_0,\cdot,m)\|_n + C_M T. \tag{49}$$

Proof. Use Propositions A.1 and A.6. ∎

Next we discuss the dependence of the solution u of (48) on $(m(t))_{t\in[t_0,T]}$ and $x_0 \in \mathbb{R}^{d_0}$. We stress that, hereafter, we use the preliminary gradient estimate $\sup_{t\in[t_0,T_M]} \|u(t)\|_1 \le K_M$ which is obtained as a first step in Proposition 5.1. In particular, the Hamiltonian $H(x_0,x,p,m)$ will be systematically estimated for $|p| \le K_M$.

Proposition 5.2. *If the assumptions of Proposition 5.1 are satisfied so that (49) holds true, then there exists $T_M > 0$ such that, for $T \in (0,T_M)$ and any $t_0 \in [0,T]$, for any $m^1, m^2 \in C^0([0,T],\mathcal{P}_2)$ and any $x_0^1, x_0^2 \in \mathbb{R}^{d_0}$, if u^1 and u^2 are the corresponding solutions to the HJ equation (48), then we have, for $n \ge 2$,*

$$\begin{aligned}\sup_{t\in[t_0,T]} \|u^1(t)-u^2(t)\|_{n-1} &\le C_M T\Big(\sup_{t\in[t_0,T]} \mathbf{d}_2(m^1(t),m^2(t)) + |x_0^1-x_0^2|\Big)\\ &\quad + (1+C_M T)\big\{[\mathrm{Lip}_{0,n-1}(G)]\mathbf{d}_2(m^1(T),m^2(T)) + [\mathrm{Lip}^{x_0}_{n-1}(G)]|x_0^1-x_0^2|\big\}\end{aligned}$$

where C_M depends on the same quantities as in Proposition 5.1 as well as on $\mathrm{Lip}_{n-1,n}(H(x_0,\cdot,\cdot,m))$, $\mathrm{Lip}^{x_0}_{n-1,n}(H(x_0,\cdot,\cdot,m))$ *(for $x \in \mathbb{R}^d$ and $|p| \le K_M$) and* $\sup_{x_0,m} \|G(x_0,\cdot,m)\|_n$.

Proof. The map $v := u^1 - u^2$ satisfies

$$\begin{cases}-\partial_t v - \mathrm{Tr}(a(t,x)D^2 v) + V(t,x)\cdot Dv + f(t,x) = 0,\\ v(T,x) = G(x_0^1,x,m^1(T)) - G(x_0^2,x,m^2(T)),\end{cases}$$

where

$$\begin{aligned}V(t,x) &:= \int_0^1 H_p\big(x,x_0^2,sDu^1(t,x)+(1-s)Du^2(t,x),m^2(t)\big)\,ds,\\ f(t,x) &:= H(x_0^1,x,Du^1(t,x),m^1(t)) - H(x_0^2,x,Du^1(t,x),m^2(t)).\end{aligned}$$

By Proposition A.7 (applied with $k=1$ and $n-1$), we have

$$\begin{aligned}&\sup_{t\in[0,T]} \|u^1(t)-u^2(t)\|_{n-1}\\ &\quad\le (1+CT)\|G(x_0^1,\cdot,m^1(T)) - G(x_0^2,\cdot,m^2(T))\|_{n-1} + CT\sup_{t\in[t_0,T]}\|f(t)\|_{n-1}\\ &\quad\le (1+CT)\{[\mathrm{Lip}_{0,n-1}(G)]\mathbf{d}_2(m^1(T),m^2(T)) + [\mathrm{Lip}^{x_0}_{n-1}(G)]|x_0^1-x_0^2|\}\\ &\qquad + CT\Big(\sup_{t\in[t_0,T]} \mathbf{d}_2(m^1(t),m^2(t)) + |x_0^1-x_0^2|\Big),\end{aligned}$$

where the constant C depends on H and on $\sup_{t\in[0,T]}\|V(t)\|_{n-1}$, hence on $\sup_{t\in[0,T]}\|u^1(t)\|_n$, $\sup_{t\in[0,T]}\|u^2(t)\|_n$, which are estimated thanks to Proposition 5.1. ∎

In our next step, we consider the solution to the Fokker–Planck equation

$$\begin{cases} \partial_t \tilde m(t,x) - \sum_{i,j} D_{ij}(a_{i,j}(t,x)\tilde m(t,x)) \\ \quad - \operatorname{div}(\tilde m(t,x) H_p(x_0, x, Du(t,x), m(t)) = 0 \quad \text{in } (t_0,T)\times\mathbb{R}^d, \\ \tilde m(t_0) = m_0 \quad \text{in } \mathbb{R}^d, \end{cases} \tag{50}$$

where $(m(t))_{t\in[t_0,T]}$ is given and u satisfies (48). Let us recall that, under the assumptions of Proposition 5.1, there exists a unique weak solution $\tilde m \in C^0([t_0,T],\mathcal{P}_2)$ to (50).

Proposition 5.3. *Assume that*

$$\|D_x G\|_\infty \le M, \quad \|D^2_{xx} G\|_\infty \le M, \quad \mathrm{Lip}_{0,1}(G) + \mathrm{Lip}_1^{x_0}(G) \le M. \tag{51}$$

Then there exists a constant $C_M > 0$, only depending on M, $\|a\|_2$ and the regularity of H, such that, for any $m^1, m^2 \in C^0([0,T],\mathcal{P}_2)$, $x_0^1, x_0^2 \in \mathbb{R}^{d_0}$ and $m_0^1, m_0^2 \in \mathcal{P}_2$, if u^1 and u^2 are the corresponding solutions to the HJ equation (48) with $x_0 = x_0^i$ and if $\tilde m_1, \tilde m_2$ are the corresponding solutions to (50) starting from m_0^1 and m_0^2 respectively, then

$$\sup_{t\in[t_0,T]} \mathbf{d}_2^2(\tilde m^1(t), \tilde m^2(t)) \\ \le (1 + C_M T)\mathbf{d}_2^2(m_0^1, m_0^2) + C_M T\Big(\sup_{t\in[t_0,T]} \mathbf{d}_2^2(m^1(t), m^2(t)) + |x_0^1 - x_0^2|^2\Big).$$

Proof. We can represent $\tilde m^i(t)$ as the law of X_t^i where $\mathbb{E}[|X_0^1 - X_0^2|^2] = \mathbf{d}_2^2(m_0^1, m_0^2)$ and X^i solves

$$X_t^i = X_0^i - \int_0^t H_p(x_0^i, X_s^i, Du^i(s, X_s^i), m^i(s))\,ds + \sqrt{2}\int_0^t \sigma(s, X_s^i)\,dB_s,$$

so that

$$\mathbb{E}[|X_t^1 - X_t^2|^2] \le \mathbb{E}[|X_0^1 - X_0^2|^2] \\ +2\mathbb{E}\Big[\int_0^t (X_s^1 - X_s^2)\cdot\big(H_p(x_0^1, X_s^1, Du^1, m^1(t)) - H_p(x_0^2, X_s^2, Du^2, m^2(t))\big)\,ds\Big] \\ +\mathbb{E}\Big[\int_0^t \mathrm{Tr}\big((\sigma(s,X_s^1)-\sigma(s,X_s^2))(\sigma(s,X_s^1)-\sigma(s,X_s^2))^*\big)\,ds|\Big] \le \mathbb{E}[|X_0^1 - X_0^2|^2] \\ + C_M\mathbb{E}\Big[\int_0^t (|X_s^1 - X_s^2|^2 + |D(u^1-u^2)(s,X_s^1)|^2 + \mathbf{d}_2^2(m^1(s), m^2(s)) + |x_0^1 - x_0^2|^2)\,ds\Big],$$

where C_M depends on the Lipschitz regularity of H_p in $\mathbb{R}^{d_0}\times\mathbb{R}^d\times B(K_M)\times\mathcal{P}_2$ (where K_M is defined in Proposition 5.1), on $\sup_t \|u^1(t)\|_2$, and on the Lipschitz regularity of σ. We infer from Gronwall's lemma that

$$\mathbb{E}[|X_t^1 - X_t^2|^2] \le (1 + C_M T)\mathbb{E}[|X_0^1 - X_0^2|^2] \\ + C_M T\Big(\sup_t \|D(u^1 - u^2)(t)\|_\infty^2 + \sup_{t\in[t_0,T]} \mathbf{d}_2^2(m^1(t), m^2(t)) + |x_0^1 - x_0^2|^2\Big).$$

As $\mathbb{E}[|X_0^1 - X_0^2|^2] = \mathbf{d}_2^2(m_0^1, m_0^2)$ and $\mathbf{d}_2^2(\tilde m^1(t), \tilde m^2(t)) \le \mathbb{E}[|X_t^1 - X_t^2|^2]$, we obtain

$$\sup_{t\in[t_0,T]} \mathbf{d}_2^2(\tilde m^1(t), \tilde m^2(t)) \le (1 + C_M T)\mathbf{d}_2^2(m_0^1, m_0^2) \\ + C_M T(\sup_t \|D(u^1 - u^2)(t)\|_\infty^2 + \sup_{t\in[t_0,T]} \mathbf{d}_2^2(m^1(t), m^2(t)) + |x_0^1 - x_0^2|^2).$$

We estimate the term $\sup_t \|D(u^1 - u^2)(t)\|_\infty^2$ by Proposition 5.2 (with $n = 2$): since $\mathrm{Lip}_{0,1}(G)$ and $\mathrm{Lip}_1^{x_0}(G)$ are estimated by (51), we deduce, for some (possibly different) constant C_M:

$$\sup_{t\in[t_0,T]} \mathbf{d}_2^2(\tilde m^1(t), \tilde m^2(t)) \\ \le (1 + C_M T)\mathbf{d}_2^2(m_0^1, m_0^2) + C_M T\Big(\sup_{t\in[t_0,T]} \mathbf{d}_2^2(m^1(t), m^2(t)) + |x_0^1 - x_0^2|^2\Big). \quad \blacksquare$$

Collecting the estimates in Propositions 5.1–5.3 yields the well-posedness of the MFG system and estimates on the solution:

Proposition 5.4. *Fix $M > 0$ and assume that (51) holds true and that $\|G\|_n \le M$ holds. Then there exist $T_M, C_M > 0$, depending on M, n, C_0, γ and*

$$\sup_{t\in[0,T_M]} \|a(t)\|_n + \sup_{|p|\le K_M,\, x_0\in\mathbb{R}^{d_0},\, m\in\mathcal{P}_2} \sum_{k=0}^{n} \|D^k_{(x,p)} H(x_0, \cdot, p, m)\|_\infty$$

(where K_M is given in Proposition 5.1) such that, for any $T \in (0, T_M)$ and any $(t_0, m_0) \in [0, T] \times \mathcal{P}_2$, there exists a unique solution to the MFG system (45). This solution satisfies

$$\sup_{t\in[t_0,T]} \|u(t)\|_n \le \|G(x_0, \cdot, m(T))\|_n + C_M T.$$

Moreover, if (t_0, m_0^1) and (t_0, m_0^2) are two initial conditions in $[0, T] \times \mathcal{P}_2$ and $x_0^1, x_0^2 \in \mathbb{R}^{d_0}$, and if (u^1, m^1) and (u^2, m^2) are the corresponding solutions to the MFG system (45) with $x_0 = x_0^1$ and $x_0 = x_0^2$ respectively, then

$$\sup_{t\in[t_0,T]} \mathbf{d}_2(m^1(t), m^2(t)) \le (1 + C_M T)\mathbf{d}_2(m_0^1, m_0^2) + C_M T|x_0^1 - x_0^2|,$$

and

$$\sup_{t\in[t_0,T]} \|u^1(t) - u^2(t)\|_{n-1} \le C_M T(\mathbf{d}_2(m_0^1, m_0^2) + |x_0^1 - x_0^2|) \\ + (1 + C_M T)\{[\mathrm{Lip}_{0,n-1}(G)](\mathbf{d}_2(m_0^1, m_0^2) + |x_0^1 - x_0^2|) + [\mathrm{Lip}_{n-1}^{x_0}(G)]|x_0^1 - x_0^2|\}.$$

Proof. The existence and uniqueness come from a standard fixed point argument on $C^0([t_0, T], \mathcal{P}_2)$ for T small enough (say $T \le T_M$ where $C_M T_M \le 1/2$, C_M being given by the previous propositions). For the stability with respect to the initial condition, one first uses the estimate in Proposition 5.3 with $\tilde{m}^i = m^i$:

$$\sup_{t\in[t_0,T]} \mathbf{d}_2^2(m^1(t), m^2(t))$$
$$\leq (1 + C_M T)\mathbf{d}_2^2(m_0^1, m_0^2) + C_M T\Big(\sup_{t\in[t_0,T]} \mathbf{d}_2^2(m^1(t), m^2(t)) + |x_0^1 - x_0^2|^2\Big).$$

Thus, as $C_M T \leq 1/2$, one obtains

$$\sup_{t\in[t_0,T]} \mathbf{d}_2(m^1(t), m^2(t)) \leq (1 + C_M T)\mathbf{d}_2(m_0^1, m_0^2) + C_M T|x_0^1 - x_0^2|,$$

modifying C_M if necessary. Plugging this estimate into the estimate for u^i in Proposition 5.2 gives the result. ■

5.1.2. The first order linearized system. Next we consider the linearized system

$$\begin{cases} \text{(i)}\ -\partial_t v - \mathrm{Tr}(a(t,x)D^2 v) + H_p(x_0, x, Du, m(t)) \cdot Dv \\ \qquad\qquad + \dfrac{\delta H}{\delta m}(x_0, x, Du, m(t))(\rho(t)) = R_1(t,x) \quad \text{in } (t_0, T) \times \mathbb{R}^d, \\ \text{(ii)}\ \partial_t \rho - \displaystyle\sum_{i,j} D_{ij}(a_{i,j}\rho) - \mathrm{div}(\rho H_p(x_0, x, Du, m(t))) - \mathrm{div}(m H_{pp} Dv) \\ \qquad\qquad - \mathrm{div}\Big(m \dfrac{\delta H_p}{\delta m}(\rho)\Big) = \mathrm{div}(R_2(t,x)) \quad \text{in } (t_0, T) \times \mathbb{R}^d, \\ \text{(iii)}\ \rho(t_0) = \rho_0, \quad v(T,x) = \dfrac{\delta G}{\delta m}(x_0, x, m(T))(\rho(T)) + R_3(x) \quad \text{in } \mathbb{R}^d, \end{cases} \tag{52}$$

where (u, m) solves (45) and H and its derivatives are evaluated at $(x_0, x, Du(t,x), m(t))$. In this section, we work under the conditions given in Proposition 5.4 so that (45) admits a unique solution, in particular we always assume that $T \leq T_M$, where T_M is given by Proposition 5.4. Our goal now is to establish estimates for (v, ρ) in dependence on G and u; we implicitly assume that $G(x_0, \cdot, m)$ is sufficiently regular (say, C_b^n) so that u inherits the same regularity (from (49)).

The data of equation (52) are $x_0 \in \mathbb{R}^{d_0}$, $\rho_0 \in C^{-k}$, $R_1 \in C^0([0,T], C_b^{n-1})$, $R_2 \in C^0([0,T], C^{-(k-1)})$ and $R_3 \in C_b^{n-1}$. Here $n \geq 2$ and $k \geq 1$. By a *solution* to (52), we mean a pair (v, ρ) such that $v \in C^0([0,T], C_b^{n-1})$ satisfies (52) (i) (integrated in time) with terminal condition $v(T, \cdot) = \frac{\delta G}{\delta m}(x_0, \cdot, m(T))(\rho(T)) + R_3(\cdot)$ and $\rho \in C^0([0,T], C_b^{-(k-1)})$ is a solution in the sense of distributions to (52) (ii) with initial condition $\rho(t_0) = \rho_0$.

Proposition 5.5. *Fix $M > 0$, $n \geq 2$ and $k \geq 1$. Under the assumptions of Proposition 5.4, and if*

$$\left\| \frac{\delta G}{\delta m} \right\|_{1;k} \leq M, \tag{53}$$

then there exist constants $T_M, C_M > 0$, depending on M, n, k, $\sup_{t\in[0,T]} \|u(t)\|_n$, $\sup_{t\in[0,T]} \|u(t)\|_{k+1}$, such that for $T \leq T_M$ there exists a unique solution (v, ρ) to (52), *and this solution satisfies*

$$\sup_{t\in[t_0,T]} \|v(t)\|_{n-1} \leq$$
$$(1+C_M T)\left\|\frac{\delta G}{\delta m}(x_0,\cdot_x,m(T),\cdot_y)\right\|_{n-1;k}\Big(\|\rho_0\|_{-k}+T\sup_t\|R_2(t)\|_{-(k-1)}+T\sup_t\|R_1(t)\|_1\Big)$$
$$+ (1+C_M T)\|R_3\|_{n-1} + C_M T\Big(1+\sup_t\|R_1(t)\|_{n-1}+\|R_2\|_{-(k-1)}\Big), \tag{54}$$

as well as

$$\sup_{t\in[t_0,T]} \|\rho(t)\|_{-k} \leq (1+C_M T)\|\rho_0\|_{-k} + C_M T\Big(\sup_t\|R_1(t)\|_1 + \sup_t\|R_2(t)\|_{-(k-1)} + \|R_3\|_{n-1}\Big). \tag{55}$$

Moreover, for any $r \leq n-1$,

$$\sup_{t\in[0,T]} \|D_x^r v(t)\|_\infty \leq (1+C_M T)\left(\left\|D_x^r\frac{\delta G}{\delta m}(x_0,\cdot,m(T))(\rho(T))\right\|_\infty + \|D_x^r R_3\|_\infty\right) + C_M T\Big(\|\rho_0\|_{-k} + \sup_t\|R(t)\|_{n-1} + \sup_t\|R_2(t)\|_{-(k-1)} + \|R_3\|_{n-1}\Big). \tag{56}$$

Proof. After proving the a priori estimates, the existence of a solution can be obtained using a continuation argument (see [10] for details). The uniqueness is an obvious consequence of the estimates. So it remains to prove the estimates. To simplify the expression, we omit the dependence of the constant C on M. Fix $t_1 \in [t_0,T]$ and $z_1 \in C_b^k$ with $k \in \{1,\dots,n-1\}$. Let z be the solution to

$$\begin{cases} -\partial_t z - \mathrm{Tr}(a(t,x)D^2 z) + H_p(x_0,x,Du,m(t))\cdot Dz = 0 & \text{in } (t_0,t_1)\times\mathbb{R}^d,\\ z(t_1,\cdot) = z_1(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{57}$$

According to Proposition A.7 (with $k=1$), we have

$$\sup_{t\in[t_0,t_1]} \|z(t)\|_k \leq (1+CT)\|z_1\|_k,$$

where C depends on the regularity of a and H and on $\sup_t \|u(t)\|_{k+1}$. Then, by duality,

$$\begin{aligned}
&\int_{\mathbb{R}^d} z_1\rho(t_1)\\
&= \int_{\mathbb{R}^d} z(t_0)\rho_0 - \int_{t_0}^{t_1}\int_{\mathbb{R}^d}\Big(H_{pp}Dv\cdot Dz + \frac{\delta H_p}{\delta m}(\rho)\cdot Dz\Big)m - \int_{t_0}^{t_1}\int_{\mathbb{R}^d} Dz\cdot R_2\\
&\leq \|z(t_0)\|_k\|\rho_0\|_{-k} + C\|Dz\|_\infty\Big(T\|Dv\|_\infty + \int_{t_0}^{t_1}\|\rho(t)\|_{-k}\Big) + T\sup_t\|z(t)\|_k\|R_2\|_{-(k-1)}\\
&\leq (1+CT)\|z_1\|_k\Big(\|\rho_0\|_{-k} + C(T\|Dv\|_\infty + \int_{t_0}^{t_1}\|\rho(t)\|_{-k}\Big) + T\|R_2\|_{-(k-1)}),
\end{aligned}$$

where $\|R_2\|_{-(k-1)} := \sup_t \|R_2(t)\|_{-(k-1)}$. Thus, taking the supremum over $\|z_1\|_k \le 1$, we obtain

$$\|\rho(t_1)\|_{-k} \le (1+CT)\|\rho_0\|_{-k} + CT(\|Dv\|_\infty + \|R_2\|_{-(k-1)}) + C\int_{t_0}^{t_1} \|\rho(t)\|_{-k}.$$

Since this holds for all $t_1 \in (t_0, T]$, by Gronwall's inequality we obtain

$$\sup_{t\in[t_0,T]} \|\rho(t)\|_{-k} \le (1+CT)\|\rho_0\|_{-k} + CT(\|Dv\|_\infty + \|R_2\|_{-(k-1)}). \tag{58}$$

Next we apply Proposition A.7 (with $k = 1$) to the HJ equation satisfied by v: we have, for any $r \le n-1$,

$$\sup_t \|v(t)\|_r \le (1+CT)\|v(T)\|_r + CTC_1, \tag{59}$$

where C depends on $\sup_t \|a(t)\|_{n-1}$, on the regularity of H, on $\sup_t \|u(t)\|_n$, and where C_1 is estimated by

$$\begin{aligned} C_1 &= \sup_t \left\| \frac{\delta H}{\delta m}(x_0, \cdot, Du(t,\cdot), m(t))(\rho(t)) \right\|_{n-1} + \|R_1\|_{n-1} \\ &\le C \sup_t \|\rho(t)\|_{-k} + \|R_1\|_{n-1}, \end{aligned} \tag{60}$$

where we have used the inequality

$$\left\| \frac{\delta H}{\delta m}(x_0, \cdot, Du(t,\cdot), m(t))(\rho(t)) \right\|_{n-1} \le \left\| \frac{\delta H}{\delta m}(x_0, \cdot, Du(t,\cdot_x), m(t), \cdot)\|_{n-1,k} \|\rho(t) \right\|_{-k}.$$

Again we notice here that the right-hand side is estimated through the regularity of H and $\sup_t \|u(t)\|_n$. Similarly we estimate, for $r \le n-1$,

$$\|v(T)\|_r \le \left\| \frac{\delta G}{\delta m}(x_0, \cdot, m(T)) \right\|_{r;k} \sup_t \|\rho(t)\|_{-k} + \|R_3\|_r. \tag{61}$$

Collecting the estimates in (58)–(61), we find, for $r \le n-1$,

$$\begin{aligned} &\sup_t \|v(t)\|_r \\ &\le (1+CT)\left\| \frac{\delta G}{\delta m}(x_0, \cdot, m(T), \cdot) \right\|_{r;k} \{(1+CT)\|\rho_0\|_{-k} + CT(\|Dv\|_\infty + \|R_2\|_{-(k-1)})\} \\ &\quad + \|R_3\|_r(1+CT) + CT\big(\|\rho_0\|_{-k} + T(\|Dv\|_\infty + \|R_2\|_{-(k-1)}) + \|R_1\|_{n-1}\big). \end{aligned} \tag{62}$$

We first consider this inequality for $r = 1$. Recall that $\|\frac{\delta G}{\delta m}\|_{1;k} \le M$. So, if we choose $T_M > 0$ such that

$$(1+CT_M)MCT_M + CT_M^2 < 1,$$

we obtain (54) for $T \le T_M$ and $n = 2$. Then from (58) we infer (55) (with a constant only depending on $\sup_t \|u(t)\|_{k+1}$). Having now estimated $\|Dv\|_\infty$, we deduce from (62) that (54) holds.

To obtain (56), we apply again Proposition A.7 to the HJ equation satisfied by v, together with estimates (60) and (55). ∎

5.1.3. The second order linearized system. Next we study the second order linearization of the MFG system. Given a solution (u,m) to (45) and (v,ρ) and two solutions (v',ρ') to (52) with arbitrary R_1, R_2, R_3 and R'_1, R'_2, R'_3, we consider a solution (w,μ) to

$$\begin{cases} \text{(i) } -\partial_t w - \mathrm{Tr}(aD^2 w) + H_p\cdot Dw + \dfrac{\delta H}{\delta m}(\mu(t)) + \dfrac{\delta^2 H}{\delta m^2}(\rho(t),\rho'(t)) \\ \qquad +H_{pp}Dv\cdot Dv' + \dfrac{\delta H_p}{\delta m}(\rho)\cdot Dv' + \dfrac{\delta H_p}{\delta m}(\rho')\cdot Dv = \tilde R_1(t,x) \\ \qquad\qquad \text{in } (t_0,T)\times\mathbb{R}^d, \\ \text{(ii) } \partial_t\mu - \sum_{i,j} D_{ij}(a_{i,j}\mu) - \mathrm{div}(\mu H_p) - \mathrm{div}(mH_{pp}Dw) - \mathrm{div}\left(m\dfrac{\delta H_p}{\delta m}(\mu)\right) \\ \qquad -\mathrm{div}(\rho H_{pp}Dv') - \mathrm{div}\left(\rho\dfrac{\delta H_p}{\delta m}(\rho')\right) - \mathrm{div}(\rho' H_{pp}Dv) - \mathrm{div}(mH_{ppp}DvDv') \\ \qquad -\mathrm{div}\left(m\dfrac{\delta H_{pp}}{\delta m}(\rho')Dv\right) - \mathrm{div}\left(\rho'\dfrac{\delta H_p}{\delta m}(\rho)\right) - \mathrm{div}\left(m\dfrac{\delta H_{pp}}{\delta m}(\rho)Dv'\right) \\ \qquad -\mathrm{div}\left(m\dfrac{\delta^2 H_p}{\delta m^2}(\rho,\rho')\right) = \mathrm{div}(\tilde R_2(t,x)) \quad \text{in } (t_0,T)\times\mathbb{R}^d, \\ \text{(iii) } \mu(t_0)=0, \quad w(T,x) = \dfrac{\delta^2 G}{\delta m^2}(x_0,x,m(T))(\rho(T),\rho'(T)) \\ \qquad\qquad + \dfrac{\delta G}{\delta m}(x_0,x,m(T))(\mu(T)) + \tilde R_3(x) \quad \text{in } \mathbb{R}^d, \end{cases} \tag{63}$$

where H and its derivatives are evaluated at $(x_0, x, Du(t,x), m(t))$. Here again we work under the conditions assumed in the previous sections which guarantee the existence, uniqueness and enough regularity for (u,m) as well as for the solutions of the linearized system. In particular, we always assume that $T \le T_M$, where T_M is now given by Proposition 5.5. The goal now is to establish estimates for (w,μ) in terms of G as well as of (u,m) and (v,ρ), (v',ρ').

The data of the problem are $\tilde R_1 \in C^0([0,T], C_b^{n-2})$, $\tilde R_2 \in C^0([0,T], C^{-(k-1)})$ and $\tilde R_3 \in C_b^{n-2}$. By a *solution* to (63), we mean a pair (w,μ) such that $w \in C^0([0,T], C_b^{n-2})$ satisfies (63) (i) (integrated in time) with the terminal condition in (63) (iii) and $\mu \in C^0([0,T], C^{-k})$ solves (63) (ii), in the sense of distributions with vanishing initial condition. Here we assume $n \ge 3$ and $k \ge 2$; the reason for this condition is just that we wish to keep the regularity threshold of (w,μ) consistent with what stated previously for (u,m) and for (v,ρ). In general, the estimates below apply to any degree of k, n but this is obviously a cascade regularity: an estimate of w in C_b^{n-2} requires an estimate of v in C_b^{n-1} and of u in C_b^n, while an estimate of μ in C^{-k} requires an estimate of ρ in $C^{-(k-1)}$.

Proposition 5.6. *Fix $M > 0$, $n \ge 3$ and $k \ge 1$. Under the assumptions of Proposition 5.4, and if (53) holds, there exist $T_M > 0$, depending on M and the regularity of H, such* *that for any* $T \in (0, T_M]$*, system* (63) *has a unique solution which satisfies*

$$\sup_t \|w(t)\|_{n-2} \le (1 + C_M T) \times \left(\left\| \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T), \cdot, \cdot) \right\|_{n-2;k-1,k-1} \|\rho(T)\|_{-(k-1)} \|\rho'(T)\|_{-(k-1)} + \|\tilde R_3\|_{n-2} \right) + C_M T \left(1 + \left\|\frac{\delta G}{\delta m}\right\|_{n-2,k}\right)\Big(\sup_t \|\tilde R_1(t)\|_{n-2} + \sup_t \|\tilde R_2(t)\|_{-(k-1)} + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k} + \mathcal{R}_{k-1,n-1}\mathcal{R}'_{k-1,n-1}\Big) \tag{64}$$

for some C_M *depending on* M*, on the regularity of* H *as well as on* n, k*,* $\sup_{t\in[0,T]} \|u\|_{n-1}$*,* $\sup_{t\in[0,T]} \|u\|_{k+1}$*, and*

$$\sup_t \|\mu(t)\|_{-k} \le \tilde C_M T \Bigg(\left(1 + \left\| \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T), \cdot, \cdot)\right\|_{1;k-1,k-1}\right) \|\rho(T)\|_{-(k-1)}\|\rho'(T)\|_{-(k-1)} + \sup_{t\in[0,T]} \|\tilde R_1(t)\|_1 + \sup_{t\in[0,T]} \|\tilde R_2(t)\|_{-(k-1)} + \|\tilde R_3\|_1 + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k} + \mathcal{R}_{k-1,2}\mathcal{R}'_{k-1,2}\Bigg), \tag{65}$$

where $\tilde C_M$ *depends on* M*, the regularity of* H*,* $n, k, \sup_{t\in[0,T]} \|u\|_{k+1}$*, and where we have set, for* $k, j \ge 1$*,*

$$\mathcal{R}_{k-1,j} := \sup_t (\|\rho(t)\|_{-(k-1)} + \|v(t)\|_j), \quad \mathcal{R}'_{k-1,j} := \sup_t (\|\rho'(t)\|_{-(k-1)} + \|v'(t)\|_j).$$

In addition, if

$$\left\|\frac{\delta G}{\delta m}\right\|_{n-2;k} \le M,$$

then for any $r \le n-2$ *and* $(t, x_0) \in [0, T] \times \mathbb{R}^{d_0}$*,*

$$\|D^r w(t,\cdot)\|_\infty \le \left(\left\| D_x^r \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T))(\rho(T), \rho'(T)) \right\|_\infty + \|D_x^r \tilde R_3(\cdot)\|_\infty \right) + C_M T \Bigg(\left\| \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T), \cdot, \cdot)\right\|_{n-2;k-1,k-1} \|\rho(T)\|_{-(k-1)} \|\rho'(T)\|_{-(k-1)} + \sup_t \|\tilde R_1(t)\|_{n-2} + \sup_t \|\tilde R_2(t)\|_{-(k-1)} + \|\tilde R_3\|_{n-2} + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k} + \mathcal{R}_{k-1,n-1}\mathcal{R}'_{k-1,n-1}\Bigg). \tag{66}$$

Remark 5.7. We recall that the quantities $\|\rho(T)\|_{-(k-1)}$ and $\mathcal{R}_{k-1,j}$ are estimated from (54) and (55). In particular, we have

$$\mathcal{R}_{k-1,k} \leq (1+C_M T)C\Big(\|\rho_0\|_{-(k-1)} + \sup_t \|R_2(t)\|_{-(k-2)} + \sup_t \|R_1(t)\|_k + \|R_3\|_k\Big)$$

for some constant C depending on $\|\frac{\delta G}{\delta m}\|_{k;k-1}$ and $\sup_t \|u(t)\|_{k+1}$, and similarly

$$\begin{aligned}&\mathcal{R}_{k-1,n-1}\\ &\quad\leq (1+C_M T)C\Big(\|\rho_0\|_{-(k-1)} + \sup_t \|R_2(t)\|_{-(k-2)} + \|R_3\|_{n-1} + \sup_t \|R_1(t)\|_{n-1}\Big)\end{aligned}$$

for a constant C depending on $\|\frac{\delta G}{\delta m}\|_{n;k-1}$ and $\sup_t \|u(t)\|_n$. Of course the same holds for ρ', v' accordingly.

Proof. We omit the proof of the well-posedness of the system, which is a consequence of the estimates (as for Proposition 5.5). To simplify the expression, we also omit the dependence of the constant C on M. We first estimate μ by duality. Fix $t_1 \in [t_0, T]$ and $z_1 \in C_b^k$ for $k \in \{1, \dots, n-1\}$. Let z be the solution to (57). Recall that Proposition A.7 (with $k=1$) implies that there is a constant $C>0$, depending on $\sup_t \|u(t)\|_{k+1}$, such that

$$\sup_{t\in[t_0,t_1]} \|z(t)\|_k \leq (1+CT)\|z_1\|_k.$$

Then

$$\begin{aligned}\int_{\mathbb{R}^d} \mu(t_1)z_1 = -\Big\{\int_{t_0}^{t_1}\int_{\mathbb{R}^d} Dz\cdot\Big(& mH_{pp}Dw + m\frac{\delta H_p}{\delta m}(\mu) + \rho H_{pp}Dv' + \rho' H_{pp}Dv\\ &+\rho\frac{\delta H_p}{\delta m}(\rho') + \rho'\frac{\delta H_p}{\delta m}(\rho) + mH_{ppp}DvDv'\\ &+m\frac{\delta H_{pp}}{\delta m}(\rho')Dv + m\frac{\delta H_{pp}}{\delta m}(\rho)Dv' + m\frac{\delta^2 H_p}{\delta m^2}(\rho,\rho') + \tilde{R}_2(t,x)\Big)\Big\}.\end{aligned}$$

Hence

$$\begin{aligned}&\int_{\mathbb{R}^d} \mu(t_1)z_1 \leq CT\|Dw\|_\infty\|Dz\|_\infty + C\|Dz\|_\infty\int_{t_0}^{t_1}\|\mu(s)\|_{-k}\,ds\\ &+CT\Big(\sup_t\|\rho(t)\|_{-(k-1)}\sup_t\|v'(t)\|_k + \sup_t\|\rho'(t)\|_{-(k-1)}\sup_t\|v(t)\|_k\Big)\sup_t\|z(t)\|_k\\ &+CT\Big(\sup_t\|\rho(t)\|_{-(k-1)}\sup_t\|\rho'(t)\|_{-(k-1)}\Big)\sup_t\|z(t)\|_k + CT\|Dv\|_\infty\|Dv'\|_\infty\|Dz\|_\infty\\ &+CT\Big(\sup_t\|\rho(t)\|_{-(k-1)}\|Dv'\|_\infty + \sup_t\|\rho'(t)\|_{-(k-1)}\|Dv\|_\infty\\ &\qquad\qquad + \sup_t\|\rho(t)\|_{-(k-1)}\sup_t\|\rho'(t)\|_{-(k-1)}\Big)\|Dz\|_\infty\\ &+CT\|\tilde{R}_2\|_{-(k-1)}\sup_t\|z(t)\|_k,\end{aligned}$$

where the constant C depends on the regularity of the function H and on $\sup_t \|u(t)\|_k$. Taking the supremum over $\|z_1\|_k \leq 1$, we infer that

$$\|\mu(t_1)\|_{-k} \leq C \int_{t_0}^{t_1} \|\mu(s)\|_{-k}\, ds + CT\Big\{\|Dw\|_\infty + \|\tilde{R}_2\|_{-(k-1)} + \Big(\sup_t \|\rho(t)\|_{-(k-1)} + \sup_t \|v(t)\|_k\Big)\Big(\sup_t \|\rho'(t)\|_{-(k-1)} + \sup_t \|v'(t)\|_k\Big)\Big\}.$$

By Gronwall's inequality, we obtain

$$\sup_t \|\mu(t)\|_{-k} \leq CT\Big\{\|Dw\|_\infty + \sup_t \|\tilde{R}_2(t)\|_{-(k-1)} + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k}\Big\}, \tag{67}$$

where C depends on the regularity of the function H and on $\sup_t \|u(t)\|_{k+1}$. From Proposition A.7 (with $k = 1$), we have

$$\sup_t \|w(t)\|_{n-2} \leq (1+CT)\left(\left\|\frac{\delta^2 G}{\delta m^2}(\rho(T), \rho'(T))\right\|_{n-2} + \left\|\frac{\delta G}{\delta m}(\mu(T))\right\|_{n-2} + \|\tilde{R}_3\|_{n-2}\right) + CT \sup_t \|f(t)\|_{n-2}, \tag{68}$$

where

$$\begin{aligned} f(t,x) \;&= \frac{\delta H}{\delta m}(\mu(t)) + \frac{\delta^2 H}{\delta m^2}(\rho(t), \rho'(t)) + H_{pp} Dv \cdot Dv' \\ &\quad + \frac{\delta H_p}{\delta m}(\rho) \cdot Dv' + \frac{\delta H_p}{\delta m}(\rho') \cdot Dv - \tilde{R}_1(t,x). \end{aligned}$$

We estimate

$$\begin{aligned} \sup_t \|f(t)\|_{n-2} \leq \Bigg(&\left\|\frac{\delta H}{\delta m}(x_0, \cdot_x, Du(t,\cdot_x), m(t), \cdot_y)\right\|_{n-2;k} \sup_t \|\mu(t)\|_{-k} + \|\tilde{R}_1\|_{n-2} \\ &+ C \sup_t (\|\rho(t)\|_{-(k-1)} + \|v(t)\|_{n-1})(\|\rho'(t)\|_{-(k-1)} + \|v'(t)\|_{n-1})\Bigg) \end{aligned}$$

for a constant C depending on the regularity of H and on $\sup_t \|u(t)\|_{n-1}$. So we conclude, using also (67), that

$$\begin{aligned} \sup_t \|f(t)\|_{n-2} \leq\; & CT(\|Dw\|_\infty + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k} + \|\tilde{R}_2\|_{-(k-1)}) \\ & + \sup_t \|\tilde{R}_1\|_{n-2} + C\, \mathcal{R}_{k-1,n-1}\mathcal{R}'_{k-1,n-1}. \end{aligned}$$

Similarly, again from (67) we get

$$\begin{aligned} &\left\|\frac{\delta G}{\delta m}(\mu(T))\right\|_{n-2} \\ &\quad \leq CT \left\|\frac{\delta G}{\delta m}(x_0, \cdot, m(T), \cdot)\right\|_{n-2;k} (\|Dw\|_\infty + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k} + \|\tilde{R}_2\|_{-(k-1)}) \end{aligned}$$

and

$$\left\| \frac{\delta^2 G}{\delta m^2}(\rho(T), \rho'(T)) \right\|_{n-2} \leq \left\| \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T), \cdot, \cdot) \right\|_{n-2;k-1,k-1} \|\rho(T)\|_{-(k-1)} \|\rho'(T)\|_{-(k-1)}.$$

Then, we find

$$\begin{aligned}
&\sup_t \|w(t)\|_{n-2} \leq \\
&(1+CT)\left(\left\| \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T), \cdot, \cdot) \right\|_{n-2;k-1,k-1} \|\rho(T)\|_{-(k-1)} \|\rho'(T)\|_{-(k-1)} + \|\tilde{R}_3\|_{n-2} \right) \\
&+ CT\left(1 + \left\| \frac{\delta G}{\delta m} \right\|_{n-2;k} \right) (\|Dw\|_\infty + \|\tilde{R}_2\|_{-(k-1)} + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k}) \\
&+ CT(\|\tilde{R}_1\|_{n-2} + \mathcal{R}_{k-1,n-1}\mathcal{R}'_{k-1,n-1}),
\end{aligned}$$

where now the constant C depends on both $\sup_t \|u(t)\|_{k+1}$ and $\sup_t \|u(t)\|_{n-1}$.

For $n = 3$, if we choose T small enough (depending on $\|\frac{\delta G}{\delta m}\|_{1,k}$ and $\sup_t \|u(t)\|_2$) we estimate $\|Dw\|_\infty$. Then, plugging this estimate into (67) gives (65) (with a constant only depending on $\sup_t \|u(t)\|_{k+1}$). Finally, we deduce (64) for $n > 3$.

For any $r \leq n-2$, $x_0 \in \mathbb{R}^{d_0}$ and $t \in [0, T]$, the estimate (66) on $D_x^r w$ follows again from Proposition A.7 (with $k = 1$), which gives, arguing as before,

$$\begin{aligned}
&\|D_x^r w(t, \cdot)\|_\infty \\
&\leq (1+CT)\left(\left\| D_x^r \frac{\delta^2 G}{\delta m^2}(\rho(T), \rho'(T)) \right\|_\infty + \left\| D_x^r \frac{\delta G}{\delta m}(\mu(T)) \right\|_\infty + \|D_x^r \tilde{R}_3\|_\infty \right) \\
&\quad + CT \sup_t \|f(t)\|_{n-2} \\
&\leq \left(\left\| D_x^r \frac{\delta^2 G}{\delta m^2}(\rho(T), \rho'(T)) \right\|_\infty + \|D_x^r \tilde{R}_3\|_\infty \right) \\
&\quad + (1+CT) \left\| \frac{\delta G}{\delta m}(x_0, \cdot, m(t), \cdot) \right\|_{n-2;k} \sup_t \|\mu(t)\|_{-k} \\
&\quad + CT\Bigg(\|\tilde{R}_1\|_{n-2} + \|\tilde{R}_2\|_{-(k-1)} + \|\tilde{R}_3\|_{n-2} \\
&\quad + \left\| \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T), \cdot, \cdot) \right\|_{n-2;k-1,k-1} \|\rho(T)\|_{-(k-1)} \|\rho'(T)\|_{-(k-1)} \\
&\quad + \mathcal{R}_{k-1,k}\mathcal{R}'_{k-1,k} + \mathcal{R}_{k-1,n-1}\mathcal{R}'_{k-1,n-1} \Bigg),
\end{aligned}$$

which yields the desired claim using (65). ∎

By gathering together Propositions 5.5 and 5.6, we deduce the following three corollaries, which will be useful in the derivation of second order estimates for the solution of the master equation.

Corollary 5.8. *Let $M > 0$, $n \geq 3$ and $k \in \{2, \dots, n-1\}$, and assume that*

$$\|G\|_n + \left\| \frac{\delta G}{\delta m} \right\|_{n-1;k} + \left\| \frac{\delta^2 G}{\delta m^2} \right\|_{n-2;k-1,k-1} \leq M.$$

Let (u, m) be the unique solution to (45) in some interval $[0, T_M]$ given by Proposition 5.4, and let (v, ρ) and (v', ρ') be two solutions to (52) with $R_1 = R_2 = R_3 = 0$ and initial conditions ρ_0, ρ'_0 respectively.

Then there exists a constant C_M such that the solution (w, μ) to (63) corresponding to (u, m), (v, ρ) and (v', ρ') and with $\tilde{R}_1 = \tilde{R}_2 = \tilde{R}_3 = 0$ satisfies, for any $T \in (0, T_M)$, $r \leq n-2$,

$$\sup_{t,x} |D_x^r w(t,x)| \leq \sup_x \left| D_x^r \frac{\delta^2 G}{\delta m^2}(x_0, x, m(T))(\rho(T), \rho'(T)) \right| + C_M T \|\rho_0\|_{-(k-1)} \|\rho'_0\|_{-(k-1)},$$

where C_M depends on M, as well as on $\|a\|_n$ and the regularity of H.

Proof. We first notice that

$$\mathrm{Lip}_{0,1}(G) \leq \sup_{x_0, m} \left\| \frac{\delta G}{\delta m}(x_0, \cdot, m, \cdot) \right\|_{1,1} \leq M,$$

hence we are in a position to apply Proposition 5.4, and there exists a time $T_M > 0$ such that the unique solution (u, m) to (45) satisfies $u \in C_b^n$ with an estimate depending on M and $\sup_{x_0, m} \|G(x_0, \cdot, m)\|_n$.

From Proposition 5.6, we have

$$\|D_x^r w(t, \cdot)\|_\infty \leq \left\| D_x^r \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T))(\rho(T), \rho'(T)) \right\|_\infty + C_M T (\|\rho(T)\|_{-(k-1)} \|\rho'(T)\|_{-(k-1)} + \mathcal{R}_{k-1,k} \mathcal{R}'_{k-1,k} + \mathcal{R}_{k-1,n-1} \mathcal{R}'_{k-1,n-1}).$$

On the other hand, we know from Proposition 5.5 that

$$\sup_t \|v(t)\|_{n-1} \leq (1 + C_M T) \left\| \frac{\delta G}{\delta m} \right\|_{n-1;k-1} \|\rho_0\|_{-(k-1)},$$
$$\sup_t \|\rho(t)\|_{-(k-1)} \leq (1 + C_M T) \|\rho_0\|_{-(k-1)},$$

which allows us to estimate $\mathcal{R}_{k-1,k}$ and $\mathcal{R}_{k-1,n-1}$. Here the constant depends on $\sup_t \|u(t)\|_n$. A similar estimate holds for (v', ρ'). Therefore, we conclude the desired estimate. ∎

Corollary 5.9. *Under the assumptions of Corollary 5.8, suppose in addition that*

$$\left\| D_{x_0} \frac{\delta G}{\delta m} \right\|_{n-2;k-1} \leq M.$$

Let (u, m) be the unique solution to (45) in $[0, T_M]$, let (v, ρ) be a solution to (52) with $R_1 = R_2 = R_3 = 0$ and initial condition ρ_0, and, for any $l \in \mathbb{R}^{d_0}$ with $|l| = 1$, let (v^l, ρ^l) be a solution to (52) with zero initial condition and with

$$\begin{aligned} R_1(t, x) &= -\partial^l_{x_0} H(y_0, x, Du(t, x), m(t)), \\ R_2(t, x) &= m(t, x)\partial^l_{x_0} H_p(y_0, x, Du(t, x), m(t)), \\ R_3(t, x) &= \partial^l_{x_0} G(y_0, x, m(T)). \end{aligned} \tag{69}$$

Then there exists a constant C_M such that the solution (w^l, μ^l) to (63) corresponding to (u, m), (v, ρ) and (v^l, ρ^l) and with

$$\begin{aligned} \tilde{R}_1(t, x) &= -\partial^l_{x_0} H_p(x_0, x, Du, m(t))Dv - \partial^l_{x_0} \frac{\delta H}{\delta m}(x_0, x, Du, m(t))(\rho(t)), \\ \tilde{R}_2(t, x) &= \rho\partial^l_{x_0} H_p(x_0, x, Du, m(t)) + m\partial^l_{x_0} H_{pp}(x_0, x, Du, m(t))Dv + m\partial^l_{x_0} \frac{\delta H_p}{\delta m}(\rho), \\ \tilde{R}_3(x) &= \partial^l_{x_0} \frac{\delta G}{\delta m}(x_0, x, m(T))(\rho(T)), \end{aligned} \tag{70}$$

satisfies, for any $T \in (0, T_M)$ and $r \leq n - 2$,

$$\sup_{t,x} \Big(\sum_{|l|=1} |D^r_x w^l(t, x)|^2 \Big)^{1/2} \leq \sup_x \left| D^r_x D_{x_0} \frac{\delta G}{\delta m}(x_0, x, m(T))(\rho(T)) \right| + C_M T \|\rho_0\|_{-(k-1)},$$

where C_M depends on M, as well as on $\|a\|_n$ and the regularity of H.

Proof. We first notice that

$$\sup_t \|\tilde{R}_1(t)\|_{n-2} + \sup_t \|\tilde{R}_2(t)\|_{-(k-1)} \leq C \sup_t (\|v(t)\|_{n-1} + \|\rho(t)\|_{-(k-1)})$$

for a constant depending on the regularity of H, on $\sup_t \|u(t)\|_{n-1}$ and on $\sup_t \|u(t)\|_k$. However, the last term is bounded by $\sup_t \|u(t)\|_{n-1}$ since $k \leq n - 1$. Next we estimate the terms (v, ρ), (v^l, ρ^l) and μ^l: we have, from Propositions 5.5 and 5.6,

$$\begin{aligned} \sup_t \|v(t)\|_{n-1} &\leq (1 + C_M T) \left\| \frac{\delta G}{\delta m} \right\|_{n-1;k-1} \|\rho_0\|_{-(k-1)} \leq C_M \|\rho_0\|_{-(k-1)}, \\ \sup_t \|\rho(t)\|_{-(k-1)} &\leq (1 + C_M T)\|\rho_0\|_{-(k-1)}, \\ \sup_t \|v^l(t)\|_{n-1} &\leq (1 + C_M T) \left\| \frac{\delta G}{\delta m} \right\|_{n-1;k-1} + C_M T \leq C_M, \\ \sup_t \|\rho^l(t)\|_{-(k-1)} &\leq C_M T, \quad \|\mu^l\|_{-k} \leq C_M T \|\rho_0\|_{-(k-1)}. \end{aligned}$$

We note that the functions w^l solve linear equations with the same diffusion and the same drift. So, combining Proposition A.7 with the inequalities above and arguing as in the proof of Proposition 5.6 gives, for any $r \le n-2$,

$$\begin{aligned}
&\sup_x \Big(\sum_{|l|=1} |D^r_x w^l(t,x)|^2 \Big)^{1/2} \le (1+CT) \\
&\times \sup_x \left(\sum_{|l|=1} \left(\left| D^r_x \frac{\delta^2 G}{\delta m^2}(\rho(T), \rho^l(T)) \right| + \left| D^r_x \frac{\delta G}{\delta m}(\mu^l(T)) \right| + \left| D^r_x \partial^l_{x_0} \frac{\delta G}{\delta m}(\rho^l(T)) \right| \right)^2 \right)^{1/2} \\
&\quad + C_M T \|\rho_0\|_{-(k-1)} \\
&\le \sup_x \left(\sum_{|l|=1} \left(\left| D^r_x \partial^l_{x_0} \frac{\delta G}{\delta m}(\rho^l(T)) \right| + C_M T \|\rho_0\|_{-(k-1)} \right)^2 \right)^{1/2} + C_M T \|\rho_0\|_{-(k-1)},
\end{aligned}$$

where we have omitted the dependence of G on $(x_0, x, m(T))$. This gives the result. ∎

Corollary 5.10. *Under the assumptions of Corollary 5.9, suppose in addition that* $\|D^2_{x_0} G(x_0, \cdot, m)\|_{n-2} \le M$. *Fix* $l, l' \in \mathbb{N}^{d_0}$ *with* $|l| = |l'| = 1$. *Let* (u, m) *be the unique solution to (45) in* $[0, T_M]$ *and let* (v^l, ρ^l), $(v^{l'}, \rho^{l'})$ *be the solution to (52) with zero initial condition and with* R_1, R_2, R_3 *and* R'_1, R'_2, R'_3 *given by (69) for* l *and* l' *respectively.*

Let $(w^{l,l'}, \mu^{l,l'})$ *be the solution to (63) corresponding to* (u, m), (v^l, ρ^l) *and* $(v^{l'}, \rho^{l'})$ *and with*

$$\begin{aligned}
\tilde R^{l,l'}_1(t,x) &= -\Big(\partial^{l+l'}_{x_0} H + \partial^l_{x_0} H_p D v^{l'} + \partial^{l'}_{x_0} H_p D v^l + \partial^l_{x_0} \frac{\delta H}{\delta m}(\rho^{l'}(t)) + \partial^{l'}_{x_0} \frac{\delta H}{\delta m}(\rho^l(t)) \Big), \\
\tilde R^{l,l'}_2(t,x) &= \rho^{l'} \partial^l_{x_0} H_p + \rho^l \partial^{l'}_{x_0} H_p + m(\partial^l_{x_0} H_{pp} D v^{l'} + \partial^{l'}_{x_0} H_{pp} D v^l) \\
&\qquad + m\Big(\partial^l_{x_0} \frac{\delta H_p}{\delta m}(\rho^{l'}) + \partial^{l'}_{x_0} \frac{\delta H_p}{\delta m}(\rho^l) \Big) + m \partial^{l+l'}_{x_0} H_p \qquad (71) \\
\tilde R^{l,l'}_3(t,x) &= \partial^{l+l'}_{x_0} G(x_0, x, m(T)) + D^l_{x_0} \frac{\delta G}{\delta m}(x_0, x, m(T))(\rho^{l'}(T)) \\
&\qquad + \partial^{l'}_{x_0} \frac{\delta G}{\delta m}(x_0, x, m(T))(\rho^l(T)),
\end{aligned}$$

where H *and its derivatives are computed at* $(x_0, x, Du(t,x), m(t))$. *Then there exists a constant* C_M *such that, for any* $T \in (0, T_M)$ *and* $r \le n-2$,

$$\sup_{t,x} \Big(\sum_{l,l'} |D^r_x w^{l,l'}(t,x)|^2 \Big)^{1/2} \le \sup_x |D^r_x D^2_{x_0} G(x_0, \cdot, m(T))| + C_M T,$$

where C_M *depends on* M, *as well as on* $\|a\|_n$ *and the regularity of* H.

Proof. We can estimate (v^l, ρ^l) and $(v^{l'}, \rho^{l'})$ and $\mu^{l,l'}$—and therefore $\tilde R^{l,l'}_1$ and $\tilde R^{l,l'}_2$—exactly as in the previous corollary. Moreover, as the functions $w^{l,l'}$ solve a HJ with the same diffusion and the same drift term, we can use Proposition A.7 to bound $(\sum_{l,l'} |D_x^r w^{l,l'}(t,x)|^2)^{1/2}$:

$$\sup_{t,x} \Big(\sum_{l,l'} |D_x^r w^{l,l'}(t,x)|^2\Big)^{1/2} \le \sup_x \Big(\sum_{l,l'} \big(|D_x^r \partial_{x_0}^{l+l'} G(x_0,\cdot,m(T))| + C_M T\big)^2\Big)^{1/2} + C_M T,$$

which gives the required estimate after rearranging. ■

5.2. *Estimates on the first order master equation*

Armed with the regularity results for the MFG system, we finally prove the estimates on the solutions to the master equations that appear in our splitting schemes. As before, throughout the section, the assumptions of Section 2.3 on a, H, G and G^0 are in force.

5.2.1. *First order differentiability of U and U^0*

Proposition 5.11. *For any $M > 0$, there exist $T_M, K_M > 0$, depending on C_0, γ and $\|Da\|_\infty$, and there exists $C_M > 0$, depending also on n, $k \in \{2,\dots,n-1\}$, $\sup_t \|a(t)\|_n$ and the regularity of H, such that, if*

$$\|G\|_n + \left\|\frac{\delta G}{\delta m}\right\|_{n-1;k} \le M, \tag{72}$$

and if $T \in (0,T_M]$, then the map U defined by (44) is a classical solution to (43), and satisfies

$$\sup_{t\in[0,T]} \|U(t)\|_n \le \|G\|_n + C_M T.$$

Moreover, for any $|\alpha| \le n-1$, $\partial_x^\alpha \frac{\delta U}{\delta m}$ is of class C^1 in m, and for $k \in \{2,\dots,n-1\}$,

$$\sup_{t\in[0,T]} \left\|\frac{\delta U}{\delta m}(t)\right\|_{n-1;k} \le \left\|\frac{\delta G}{\delta m}\right\|_{n-1;k} + C_M T.$$

Remark 5.12. In the proof we show the following representation:

$$\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t_0,x_0,x,m_0,y)\,\rho_0(dy) = v(t_0,x) \tag{73}$$

where (u,m) is the solution of the MFG system (45) and (v,ρ) is the solution of the linearized system (52) with right-hand side $R_1 = R_2 = R_3 = 0$ and with initial condition (t_0,ρ_0). Note that the normalization condition (11) is satisfied, because if one chooses $\rho_0 = m_0$, then $(v,\rho) = (0,m)$.

The proof relies on the following lemma, in which we also provide estimates to obtain the differentiability of U with respect to x_0 later on.

Lemma 5.13. *Under the assumptions of Proposition 5.11, fix $(t_0, m_0), (t_0, m_1) \in [0, T) \times \mathcal{P}_2$ and $y_0, \xi \in \mathbb{R}^d$ with $|\xi| \le 1$. Let (u, m) be the solution to (45) with $x_0 = y_0$ and with initial condition (t_0, m_0), and, for $h \in (0, 1)$, let (u_h, m_h) be the solution to (45) with $x_0 = y_0 + \xi h$ and with initial condition $(t_0, (1-h)m_0 + hm_1)$. Let also (v, ρ) be the solution to (52) associated with (u, m), $x_0 = y_0$ and with*

$$\begin{aligned} R_1(t,x) &= -H_{x_0}(y_0, x, Du(t,x), m(t)) \cdot \xi, \\ R_2(t,x) &= m(t,x) H_{x_0 p}(y_0, x, Du(t,x), m(t)) \cdot \xi, \\ R_3(t,x) &= G_{x_0}(y_0, x, m(T)) \cdot \xi, \end{aligned} \tag{74}$$

and initial condition $(t_0, m_1 - m_0)$. Then there exists a constant C (independent of h) such that

$$\sup_{t \in [t_0, T]} \|u_h(t) - u(t) - hv(t)\|_{n-1} \le Ch^2, \tag{75}$$

$$\sup_{t \in [t_0, T]} \|m_h(t) - m(t) - h\rho(t)\|_{-k} \le Ch^2. \tag{76}$$

Remark 5.14. The goal of this lemma is to identify the first order derivatives $\frac{\delta U}{\delta m}$ and $D_{x_0} U$. The constant C above will depend on the regularity of H and G, as well as on $\sup_{t \in [t_0, T]} \|u(t)\|_n$; however, this is not detailed later since it will not be relevant; indeed, (75) and (76) are only used when letting $h \to 0$.

Proof of Lemma 5.13. We set

$$v_h(t,x) = u_h(t,x) - u(t,x) - hv(t,x), \quad \rho_h(t,x) = m_h(t,x) - m(t,x) - h\rho(t,x).$$

Then the pair (v_h, ρ_h) solves

$$\begin{cases} -\partial_t v_h - \mathrm{Tr}(a(t,x) D^2 v_h) + H_p(y_0, x, Du, m(t)) \cdot Dv_h \\ \qquad\qquad + \dfrac{\delta H}{\delta m}(y_0, x, Du, m(t))(\rho_h(t)) = R_{h,1}(t,x) \quad \text{in } (t_0, T) \times \mathbb{R}^d, \\ \partial_t \rho_h - \sum_{i,j} D_{ij}(a_{i,j} \rho_h) - \mathrm{div}(\rho_h H_p(x, Du, m(t)) - \mathrm{div}(m H_{pp}(x, Du, m(t)) Dv_h) \\ \qquad\qquad - \mathrm{div}\left(m \dfrac{\delta H_p}{\delta m}(x, Du, m(t))(\rho_h) \right) = \mathrm{div}(R_{h,2}(t,x)) \quad \text{in } (t_0, T) \times \mathbb{R}^d, \\ \rho_h(t_0) = 0, \quad v_h(T, x) = \dfrac{\delta G}{\delta m}(x, m(T))(\rho_h(T)) + R_{h,3}(x) \quad \text{in } \mathbb{R}^d, \end{cases}$$

where

$$\begin{aligned} R_{h,1}(t,x) = -\Big(& H(y_0 + \xi h, x, Du_h(t,x), m_h(t)) - H(y_0, x, Du(t,x), m(t)) \\ & - H_p(y_0, x, Du(t,x), m(t)) \cdot D(u_h(t,x) - u(t,x)) \\ & - \frac{\delta H}{\delta m}(y_0, x, Du(t,x), m(t))(m_h(t) - m(t)) \\ & - hH_{x_0}(y_0, x, Du(t,x), m(t)) \cdot \xi \Big), \end{aligned}$$

$$
\begin{aligned}
R_{h,2}(t,x) &= m_h(t,x)H_p(y_0+\xi h, x, Du_h(t,x), m_h(t)) \\
&\quad - m(t,x)H_p(y_0, x, Du(t,x), m(t)) \\
&\quad - (m_h(t,x)-m(t,x))H_p(y_0, x, Du(t,x), m(t)) \\
&\quad - m(t,x)H_{pp}(y_0, x, Du(t,x), m(t))D(u_h-u)(t,x) \\
&\quad - hm(t,x)H_{x_0 p}(y_0, x, Du(t,x), m(t))\cdot\xi \\
&\quad - m(t,x)\frac{\delta H_p}{\delta m}(y_0, x, Du(t,x), m(t))(m_h(t)-m(t)), \\
R_{h,3}(x) &= G(y_0+\xi h, x, m_h(T)) - G(y_0, x, m(T)) \\
&\quad - \frac{\delta G}{\delta m}(y_0, x, m(T))(m_h(T)-m(T)) - hG_{x_0}(y_0, x, m(T))\cdot\xi.
\end{aligned}
$$

Next we estimate $R_{h,1}$, $R_{h,2}$ and $R_{h,3}$. As

$$
\begin{aligned}
&R_{h,1} = \\
&-\int_0^1 \Big\{ (H_p(x_\tau, x, p_\tau(t,x), m_\tau(t)) - H_p(y_0, x, Du(t,x), m(t)))\cdot D(u_h(t,x)-u(t,x)) \\
&+ (H_{x_0}(x_\tau, x, p_\tau(t,x), m_\tau(t)) - H_{x_0}(y_0, x, Du(t,x), m(t)))\cdot h\xi \\
&+ \int_{\mathbb{R}^d}\Big(\frac{\delta H}{\delta m}(x_\tau, x, p_\tau(t,x), m_\tau(t), y) - \frac{\delta H}{\delta m}(y_0, x, Du(t,x), m(t), y)\Big) \\
&\qquad\qquad \cdot (m_h(t)-m(t))(dy)\Big\}\, d\tau,
\end{aligned}
$$

where $x_\tau := (1-\tau)y_0 + \tau(y_0+\xi h)$, $p_\tau := (1-\tau)Du(t,x) + \tau Du_h(t,x)$ and $m_\tau(t,x) := (1-\tau)m(t,x) + \tau m_h(t,x)$, we have

$$
\|R_{h,1}(t)\|_{n-1} \le C\big(\|u_h(t)-u(t)\|_n^2 + h^2 + \mathbf{d}_2^2(m_h(t), m(t))\big).
$$

In the same way,

$$
\begin{aligned}
\|R_{h,3}\|_{n-1} &\le C(\mathbf{d}_2^2(m_h(T), m(T)) + h^2) \\
&\le C\big(\|u_h(T)-u(T)\|_n^2 + \mathbf{d}_2^2(m_h(T), m(T)) + h^2\big).
\end{aligned}
$$

Finally, for $k \ge 2$, we have

$$
\begin{aligned}
&\|R_{h,2}(t)\|_{-(k-1)} \\
&\quad = \sup_{\|\phi\|_{k-1}\le 1}\int_{\mathbb{R}^d}\phi(t,x)\big(H_p(x_0, x, Du_h(t,x), m_h(t)) - H_p(y_0, x, Du(t,x), m(t))\big) \\
&\qquad\qquad \times (m_h(t,dx) - m(t,dx)) \\
&\qquad + \int_{\mathbb{R}^d}\phi(t,x)\Big(H_p(x_0, x, Du_h(t,x), m_h(t)) - H_p(y_0, x, Du(t,x), m(t)) \\
&\qquad - H_{x_0 p}(y_0, x, Du(t,x), m(t))\cdot h\xi - H_{pp}(y_0, x, Du(t,x), m(t))D(u_h-u)(t,x) \\
&\qquad - \frac{\delta H_p}{\delta m}(y_0, x, Du(t,x), m(t))(m_h(t)-m(t))\Big)\, m(t,dx) \\
&\quad \le C(\|u_h-u\|_2^2 + \mathbf{d}_2^2(m_h(t), m(t)) + h^2).
\end{aligned}
$$

By Proposition 5.5, there exist constants $T_M, C_M > 0$, depending on M, n, k, $\sup_{t\in[0,T]} \|u\|_n$, such that if $T \le T_M$ and if (72) holds, then

$$\begin{aligned}\sup_{t\in[0,T]} &\|v_h(t)\|_{n-1} \\ &\le (1+C_M T)\|R_{h,3}\|_{n-1} + C_M T\Big(\sup_t \|R_{h,1}(t)\|_{n-1} + \sup_t \|R_{h,2}(t)\|_{-(k-1)}\Big) \\ &\le C\Big(\sup_t \|u_h(t)-u(t)\|_n^2 + \sup_t \mathbf{d}_2^2(m_h(t),m(t)) + h^2\Big).\end{aligned}$$

We then infer by Proposition 5.4 and the definition of v_h that

$$\sup_{t\in[t_0,T]} \|u_h(t)-u(t)-hv(t)\|_{n-1} \le C\left(\mathbf{d}_2^2((1-h)m_0+hm_1,m_0)+h^2\right) \le Ch^2.$$

The estimate of ρ_h comes from Proposition 5.5 in the same way. ∎

Proof of Proposition 5.11. Proposition 5.4 and the representation formula (44) imply the estimate on $\|U(t,\cdot,m)\|_n$. Let us now show that the map U given by (44) is differentiable with respect to m. Fix $x_0 \in \mathbb{R}^{d_0}$, $(t_0,m_0), (t_0,m_1) \in [0,T)\times\mathcal{P}_2$, let (u,m), (u_h,m_h) and (v,ρ) be as in Lemma 5.13 with $\xi = 0$, so $R_1 = R_2 = R_3 = 0$. Then

$$\sup_{t\in[t_0,T]} \|u_h(t)-u(t)-hv(t)\|_{n-1} \le o(h).$$

Taking $t = t_0$, this implies that

$$\|U(t_0,x_0,\cdot,(1-h)m_0+hm_1) - U(t_0,x_0,\cdot,m_0) - hv(t_0,\cdot)\|_{n-1} \le o(h).$$

So, if we choose $m_1 = \delta_y$ for a fixed $y \in \mathbb{R}^d$, we have just proved that the map $\hat{U}(h;m_0,y) = U(t_0,x_0,x,(1-h)m_0+h\delta_y)$ has a derivative at $h=0$ and that this derivative is given by $v(t_0,x)$. Note that the map $(m_0,y)\mapsto v(t_0,x;m_0,y)$ is continuous and bounded thanks to the estimates in Proposition 5.5 and the uniqueness of the solution. So we can apply Lemma B.1 which says that U is C^1 in m with

$$v(t_0,x) = \frac{\delta U}{\delta m}(t_0,x_0,x,m_0,y).$$

Then by linearity and continuity one easily checks that (73) and the normalization condition (11) hold. A similar argument applies to derivatives of $\frac{\delta U}{\delta m}$ with respect to x.

Next we check that U solves (43). Let us start with $m(t_0) = m_0$ with a smooth density. Then (u,m) is a classical solution and, as

$$U(t,x_0,x,m(t)) = u(t,x) \quad \forall (t,x)\in[t_0,T]\times\mathbb{R}^d,$$

we have, for any $h>0$ and in view of the equation for m,

$$
\begin{aligned}
&u(t_0+h,x)-u(t_0,x) = U(t_0+h,x_0,x,m(t_0+h)) - U(t_0,x_0,x,m(t_0))\\
&= \int_{t_0}^{t_0+h}\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t_0+h,x_0,x,m(t),y)\partial_t m(t,y)\,dy\,dt\\
&\quad + U(t_0+h,x_0,x,m(t_0)) - U(t_0,x_0,x,m(t_0))\\
&= -\int_{t_0}^{t_0+h}\int_{\mathbb{R}^d} D_m U(t_0+h,x_0,x,m(t),y)\cdot H_p(x_0,y,D_x u(t,y),m(t))m(t,y)\,dy\,dt\\
&\quad + \int_{t_0}^{t_0+h}\int_{\mathbb{R}^d} \mathrm{Tr}(a(t,y)D^2_{ym}U(t,x_0,x,m,y))\,m(dy)\,dt\\
&\quad + U(t_0+h,x_0,x,m(t_0)) - U(t_0,x_0,x,m(t_0)).
\end{aligned}
$$

On the other hand, by the equation for u,

$$
\begin{aligned}
&u(t_0+h,x)-u(t_0,x)\\
&= \int_{t_0}^{t_0+h}\big(-\mathrm{Tr}(a(t,x)D^2u(t,x)) + H(x_0,x,Du(t,x),m(t))\big)\,dt\\
&= \int_{t_0}^{t_0+h}\big(-\mathrm{Tr}(a(t,x)D^2_{xx}U(t,x_0,x,m(t))) + H(x_0,x,D_xU(t,x_0,x,m(t)),m(t))\big)\,dt.
\end{aligned}
$$

So

$$
\begin{aligned}
&U(t_0+h,x_0,x,m_0) - U(t_0,x_0,x,m_0)\\
&= \int_{t_0}^{t_0+h}\int_{\mathbb{R}^d} D_m U(t_0+h,x_0,x,m(t),y)\\
&\qquad\qquad \cdot H_p(x_0,y,D_xU(t,x_0,y,m(t)),m(t))m(t,y)\,dy\,dt\\
&- \int_{t_0}^{t_0+h}\int_{\mathbb{R}^d} \mathrm{Tr}(a(t,y)D^2_{ym}U(t,x_0,x,m,y))\,m(dy)\,dt\\
&+ \int_{t_0}^{t_0+h}\big(-\mathrm{Tr}(a(t,x)D^2_{xx}U(t,x_0,x,m(t))) + H(x_0,x,D_xU(t,x_0,x,m(t)),m(t))\big)\,dt.
\end{aligned}
$$

Therefore U has a time derivative at (t_0,x_0,x,m_0) and

$$
\begin{aligned}
&\partial_t U(t_0,x_0,x,m_0)\\
&\quad = \int_{\mathbb{R}^d} D_m U(t_0,x_0,x,m_0,y)\cdot H_p(x_0,y,D_xU(t_0,x_0,y,m_0),m_0))m_0(y)\,dy\\
&\quad\quad - \int_{\mathbb{R}^d} \mathrm{Tr}(a(t_0,y)D^2_{ym}U(t_0,x_0,x,m,y))\,m(dy)\\
&\quad\quad - \mathrm{Tr}(a(t_0,x)D^2_{xx}U(t_0,x_0,x,m_0)) + H(x_0,x,D_xU(t,x_0,x,m_0),m_0).
\end{aligned}
$$

This shows that U satisfies (43) at any point (t_0,x_0,x,m_0) where m_0 has a smooth density. The general case can be treated by a density argument, since the right-hand side of the above equation is continuous in (t_0,x_0,x,m_0).

Let us now explain the estimates on $\frac{\delta U}{\delta m}$. In view of (72), (73) and Proposition 5.5, we have, for any $r \le n-1$,

$$\left\| D_x^r \frac{\delta U}{\delta m}(t_0, x_0, \cdot, m_0)(\rho_0) \right\|_\infty = \| D_x^r v(t_0, x_0, \cdot) \|_\infty$$
$$\le (1 + C_M T) \left\| D_x^r \frac{\delta G}{\delta m}(x_0, \cdot, m(T), \cdot) \right\|_{0;k} \|\rho_0\|_{-k} + C_M T \|\rho_0\|_{-k}.$$

Taking the sup over ρ_0 with $\|\rho_0\|_{-k} \le 1$, $x_0 \in \mathbb{R}^{d_0}$, summing over $r \le n-1$ and then taking the sup over t, m gives the estimate on $\frac{\delta U}{\delta m}$. Notice that the estimate given by Proposition 5.5 depends on $\sup_t \|u(t)\|_n$ (we use here $k \le n-1$); but this last term is estimated in terms of M only, because of Proposition 5.4 and since $\|G(x_0, \cdot, m)\|_n \le M$. ■

Proposition 5.15. *Under the assumptions of Proposition 5.11, let $M, T_M, C_M > 0$ be given accordingly. Assume, in addition, that $T \in (0, T_M]$ and*

$$\sup_{x_0, m} |G^0(x_0, m)| + |D_{x_0} G^0(x_0, m)| + \left\| \frac{\delta G^0}{\delta m}(x_0, m, \cdot) \right\|_{n-1;k} + \|D_{x_0} G(x_0, \cdot, m)\|_{n-1} \le M. \quad (77)$$

Then the map U^0 defined by (47) is a classical solution to (46). In addition, U^0 and U are differentiable with respect to x_0 and satisfy

$$\sup_t \|(U^0, U)(t)\|_n \le \|(G^0, G)\|_n + C_M T, \quad (78)$$

$$\sup_t \|D_{x_0}(U^0, U)(t)\|_{n-1} \le \|(D_{x_0} G^0, D_{x_0} G)\|_{n-1} + C_M T, \quad (79)$$

$$\sup_t \left\| \frac{\delta(U^0, U)}{\delta m}(t) \right\|_{n-1;k} \le \left\| \frac{\delta(G^0, G)}{\delta m} \right\|_{n-1;k} + C_M T. \quad (80)$$

As we will see in the proof, it is possible to estimate U^0 and U separately. However, we will need the specific form of the estimate in the analysis of the MFG problem with a major player.

Proof of Proposition 5.15. Differentiability of U with respect to x_0 can be checked just as its differentiability with respect to m. Let ξ be any unit vector of $\mathbb{R}^{d_0}$, and (u, m), (u_h, m_h) and (v, ρ) be as in Lemma 5.13 with $m_1 = m_0$. Then, by Proposition 5.5 and the fact that

$$\sup_t \|R_1(t)\|_{n-1} + \sup_t \|R_2(t)\|_{-(k-1)} \le C, \quad \|R_3\|_{n-1} \le \sup_{x_0, m} \|G_{x_0}(x_0, \cdot, m(T))\|_{n-1}, \quad (81)$$

one has

$$\|U(t_0, x_0 + h\xi, \cdot, m_0) - U(t_0, x_0, \cdot, m_0) - h v(t_0, \cdot)\|_{n-1} \le o(h),$$

and so

$$U_{x_0}(t_0, x_0, x, m_0) \cdot \xi = v(t_0, x). \tag{82}$$

To show the differentiability of U^0 with respect to m we proceed as in the proof of Proposition 5.11. Fix $x_0 \in \mathbb{R}^d$, $(t_0, m_0), (t_0, m_1) \in [0, T) \times \mathscr{P}_2$, and let (u, m), (u_h, m_h) and (v, ρ) be as in Lemma 5.13 with $\xi = 0$, so $R_1 = R_2 = R_3 = 0$. Then

$$\sup_{t \in [t_0, T]} \|\rho_h(t)\|_{-k} \leq o(h),$$

where $\rho_h(t,x) = m_h(t,x) - m(t,x) - h\rho(t,x)$. This inequality and Proposition 5.4 imply

$$\begin{aligned}
\left| G^0(x_0, m_h(T)) - G^0(x_0, m(T)) - h\frac{\delta G^0}{\delta m}(x_0, m(T))(\rho(T)) \right| &\leq \left| \frac{\delta G^0}{\delta m}(x_0, m(T))(\rho_h(T)) \right| \\
+ \Bigg| \int_0^1 \int_{\mathbb{R}^d} \left(\frac{\delta G^0}{\delta m}(x_0, (1-\tau)m(T) + \tau m_h(T), y) - \frac{\delta G^0}{\delta m}(x_0, m(T), y) \right) & \\
\times (m_h(t) - m(t))(dy)\, d\tau \Bigg| & \\
\leq o(\mathbf{d}_2(m_h(T), m(T)) + h) \leq o(h). &
\end{aligned} \tag{83}$$

For $y \in \mathbb{R}^d$ choose now $m_1 = \delta_y$; then

$$\left| U^0(t_0, x_0, (1-h)m_0 + h\delta_y) - U^0(t_0, x_0, m_0) - h\frac{\delta G^0}{\delta m}(x_0, m(T))(\rho(T)) \right| \leq o(h).$$

Note that $\rho_0 \mapsto \rho(T)$ is linear and continuous as a map from C^{-k} onto itself. Apply then Lemma B.1 to deduce that U^0 is C^1 in m with

$$\frac{\delta U^0}{\delta m}(t_0, x_0, m_0, y) = \frac{\delta G^0}{\delta m}(x_0, m(T))(\rho(T)). \tag{84}$$

Moreover, one can check as in the proof of Proposition 5.11 that U^0 solves (46) (here it is even simpler, and based on the fact that by definition of U^0, $U^0(t_0 + h, x_0, m(t_0 + h)) - U^0(t_0, x_0, m(t_0)) = 0$).

Concerning the differentiability of U^0 with respect to x_0, let ξ be any unit vector of $\mathbb{R}^d$, and let (u, m), (u_h, m_h) and (v, ρ) be as in Lemma 5.13 with $m_1 = m_0$. Then

$$\begin{aligned}
&\left| G^0(y_0 + \xi h, m_h(T)) - G^0(y_0, m(T)) - hG^0_{x_0}(y_0, m(T)) \cdot \xi - h\frac{\delta G^0}{\delta m}(y_0, m(T))(\rho(T)) \right| \\
&\quad \leq |G^0(y_0 + \xi h, m_h(T)) - G^0(y_0, m_h(T)) - hG^0_{x_0}(y_0, m_h(T)) \cdot \xi| \\
&\qquad + h|G^0_{x_0}(y_0, m_h(T)) - G^0_{x_0}(y_0, m(T))| \\
&\qquad + \left| G^0(y_0, m_h(T)) - G^0(y_0, m(T)) - h\frac{\delta G^0}{\delta m}(y_0, m(T))(\rho(T)) \right|.
\end{aligned}$$

The third term of this inequality can be treated as in (83). Therefore,

$$\begin{aligned}
\left| U^0(t_0, y_0 + \xi h, m_0) - U^0(t_0, y_0, m_0) - hG^0_{x_0}(y_0, m(T)) \cdot \xi - h\frac{\delta G^0}{\delta m}(y_0, m(T))(\rho(T)) \right| & \\
\leq o(h), &
\end{aligned}$$

and hence it follows that

$$D_{x_0}U^0(t_0,x_0,m_0)\cdot\xi = G^0_{x_0}(x_0,m(T))\cdot\xi + \frac{\delta G^0}{\delta m}(x_0,m(T))(\rho(T)). \tag{85}$$

We now prove the estimates. By Proposition 5.1 and the representation formulas (44) and (47), we have, for any $x_0\in\mathbb{R}^{d_0}$, $m\in\mathcal{P}_2$ and $r\le n$,

$$\begin{aligned}|U^0(t,x_0,m)|^2+|D^r_xU(t,x_0,x,m)|^2 &= |G^0(x_0,m(T))|^2+|D^r_xu(t,x)|^2\\ &\le |G^0(x_0,m(T))|^2+\Big(\sup_x|D^r_xG(x_0,x,m(T))|+C_MT\Big)^2\\ &\le \Big(\big(|G^0(x_0,m(T))|^2+\sup_x|D^r_xG(x_0,x,m(T))|^2\big)^{1/2}+C_MT\Big)^2\end{aligned}$$

(where we have used $x^2+(y+z)^2\le((x^2+y^2)^{1/2}+z)^2$ for nonnegative reals x,y,z), which gives (78). Next we prove (79). For $|l|=1$, $l\in\mathbb{N}^{d_0}$, we represent $\partial^l_{x_0}U^0$ and $\partial^l_{x_0}U$ by (85) and (82) respectively, where (v^l,ρ^l) is as in Lemma 5.13 with $\xi=e_l$, $m_1=m_0$ (so that $\rho^l_0=0$). Then, for $r\le n-1$,

$$\begin{aligned}&\sum_{|l|=1}|\partial^l_{x_0}U^0(t,x_0,m)|^2+|D^r_x\partial^l_{x_0}U(t,x_0,x,m)|^2\\ &\qquad=\sum_{|l|=1}\left|\partial^l_{x_0}G^0(x_0,m(T))+\frac{\delta G^0}{\delta m}(x_0,m(T))(\rho^l(T))\right|^2+|D^r_xv^l(t,x)|^2.\end{aligned}$$

Note that $\sup_t\|\rho(t)\|_{-k}\le C_MT$ by Proposition 5.5. As the v^l solve HJ equations with the same diffusion and the same drift, Proposition A.7, (77) and (81) imply that

$$\begin{aligned}&\sup_x\Big(\sum_{|l|=1}|D^rv^l|^2\Big)^{1/2}\le(1+CT)\sup_x\Big(\sum_{|l|=1}|D^rv^l(T)|^2\Big)^{1/2}+C_MT\\ &\le(1+CT)\\ &\times\sup_x\left(\sum_{|l|=1}\left(\left\|D^r_x\frac{\delta G}{\delta m}(x_0,\cdot,m(T),\cdot)\right\|_{0;k}\|\rho(T)\|_{-k}+|D^r_x\partial^l_{x_0}G(x_0,x,m(T))|\right)^2\right)^{1/2}\\ &+C_MT\\ &\le\sup_x\Big(\sum_{|l|=1}(|D^r_x\partial^l_{x_0}G(x_0,x,m(T))|+C_MT)^2\Big)^{1/2}+C_MT\\ &\le\sup_x\Big(\sum_{|l|=1}|D^r_x\partial^l_{x_0}G(x_0,x,m(T))|^2\Big)^{1/2}+C_MT,\end{aligned}$$

while

$$\begin{aligned}&\sum_{|l|=1}|\partial^l_{x_0}U^0(t,x_0,m)|^2\le\sum_{|l|=1}\left(|\partial^l_{x_0}G^0(x_0,m(T))|+\left|\frac{\delta G^0}{\delta m}(x_0,m(T))(\rho^l(T))\right|\right)^2\\ &\le\sum_{|l|=1}\big(|\partial^l_{x_0}G^0(x_0,m(T))|+C_MT\big)^2\le\Big(\Big(\sum_{|l|=1}|\partial^l_{x_0}G^0(x_0,m(T))|^2\Big)^{1/2}+C_MT\Big)^2.\end{aligned}$$

Using $((x+z)^2+(y+z)^2)^{1/2} \le (x^2+y^2)^{1/2} + \sqrt{2}\,z$, we obtain

$$\sup_x \Big(\sum_{|l|=1} \big(|\partial^l_{x_0} U^0(t,x_0,m)|^2 + |D^r_x \partial^l_{x_0} U(t,x_0,x,m)|^2 \big) \Big)^{1/2} \\ \le \sup_x \Big(\sum_{|l|=1} \big(|\partial^l_{x_0} G^0(x_0,m(T))|^2 + |D^r_x \partial^l_{x_0} G(x_0,x,m(T))|^2 \big) \Big)^{1/2} + C_M T,$$

from which we derive (79), by taking the sup over x_0, summing over r and finally taking the sup over m.

For (80), let (v,ρ) be as in Lemma 5.13 with $m_1 - m_0 = \rho_0 \in C^{-k}$ and $\xi = 0$, as in (84) and (73). We have, for any $r \le n-1$,

$$\Big| \frac{\delta U^0}{\delta m}(t,x_0,m)(\rho_0) \Big|^2 + \Big| D^r_x \frac{\delta U}{\delta m}(t,x_0,x,m)(\rho_0) \Big|^2 \\ = \Big| \frac{\delta G^0}{\delta m}(x_0,m(T))(\rho(T)) \Big|^2 + |D^r_x v(t,x)|^2.$$

So again by Proposition 5.5,

$$\Big| \frac{\delta U^0}{\delta m}(t,x_0,m)(\rho_0) \Big|^2 + \Big| D^r_x \frac{\delta U}{\delta m}(t,x_0,x,m)(\rho_0) \Big|^2 \\ \le \Big| \frac{\delta G^0}{\delta m}(x_0,m(T))(\rho(T)) \Big|^2 + \Big(\sup_x \Big| D^r_x \frac{\delta G}{\delta m}(x_0,\cdot,m(T))(\rho(T)) \Big| + C_M T \|\rho(T)\|_{-k} \Big)^2 \\ \le \Bigg[\sup_x \frac{1}{\|\rho(T)\|_{-k}} \Big(\Big| \frac{\delta G^0}{\delta m}(x_0,m(T))(\rho(T)) \Big|^2 + \Big| D^r_x \frac{\delta G}{\delta m}(x_0,x,m(T))(\rho(T)) \Big|^2 \Big)^{1/2} \\ + C_M T \Bigg]^2 \|\rho(T)\|^2_{-k} \\ \le (1+C_M T)^2 \Bigg[\sup_{x,\|\rho\|_{-k}=1} \Big(\Big| \frac{\delta G^0}{\delta m}(x_0,m(T))(\rho) \Big|^2 + \Big| D^r_x \frac{\delta G}{\delta m}(x_0,x,m(T))(\rho) \Big|^2 \Big)^{1/2} \\ + C_M T \Bigg]^2 \|\rho_0\|^2_{-k},$$

This gives (80). ∎

5.2.2. *Second order differentiability of U and U^0*

Proposition 5.16. *Let U be the solution of (43) given by (44). Let $n \ge 3$ and $k \in \{2,\dots,n-1\}$. Suppose, in addition to the assumptions of Proposition 5.11, that G is of class C^2 and*

$$\Big\| \frac{\delta^2 G}{\delta m^2}(x_0,\cdot,m,\cdot,\cdot) \Big\|_{n-2;k-1,k-1} \le M.$$

Then there exists $T_M > 0$ (depending on M and on the data but not on G) such that if $T \in (0,T_M]$, then the map U is C^2 with respect to the measure variable and the parameter x_0, and satisfies

$$\sup_{t\in[0,T]} \Big\| \frac{\delta^2 U}{\delta m^2}(t) \Big\|_{n-2;k-1,k-1} \le \Big\| \frac{\delta^2 G}{\delta m^2} \Big\|_{n-2;k-1,k-1} + C_M T.$$

Proof. Our first goal is to show that $\frac{\delta U}{\delta m}$ is differentiable with respect to m. Let $(t_0, m_0) \in [0, T) \times \mathcal{P}_2$, $y, y' \in \mathbb{R}^d$ and let

- (u, m) (respectively (u_h, m_h)) be the solution of the MFG system (45) with initial condition (t_0, m_0) (respectively $(t_0, (1-h)m_0 + h\delta_{y'})$),
- (v, ρ) (respectively (v', ρ')) be the solution of the first order linearized system (52) with zero right-hand side, initial condition (t_0, δ_y) (respectively $(t_0, \delta_{y'})$) and where the Hamiltonian and its derivatives are evaluated at $(x_0, x, Du(t, x), m(t))$,
- $(\tilde{v}_h, \tilde{\rho}_h)$ be the solution to the first order linearized system (52) with zero right-hand side, with initial condition (t_0, δ_y) and where the Hamiltonian and its derivatives are evaluated at $(x_0, x, Du_h(t, x), m_h(t))$,
- (w, μ) be the solution to the second order linearized system (63) associated with (u, m), (v, ρ), (v', ρ') and with right-hand side 0.

Recall (see (73)) that

$$\begin{aligned} \tilde{v}_h(t_0, x) &= \frac{\delta U}{\delta m}(t_0, x_0, x, (1-h)m_0 + h\delta_{y'}, y), \\ v(t_0, x) &= \frac{\delta U}{\delta m}(t_0, x_0, x, m_0, y), \quad v'(t_0, x) = \frac{\delta U}{\delta m}(t_0, x_0, x, m_0, y') \end{aligned} \tag{86}$$

so we expect $w(t_0, \cdot)$ to represent the derivative in m of $\frac{\delta U}{\delta m}$, that is, $\frac{\delta^2 U}{\delta m^2}(t_0, x_0, x, m_0, y, y')$.

We consider

$$(\hat{v}_h, \hat{\rho}_h) := (\tilde{v}_h, \tilde{\rho}_h) - (v, \rho) - h(w, \mu).$$

Let us first note that, by Proposition 5.4, we have

$$\begin{aligned} \sup_{t \in [t_0, T]} \left(\|\tilde{u}_h(t, x) - u(t, x)\|_{n-1} + \mathbf{d}_2(m_h(t), m(t)) \right) &\le C\mathbf{d}_2((1-h)m_0 + h\delta_{y'}, m_0) \\ &\le Ch. \end{aligned} \tag{87}$$

Next we claim that

$$\sup_{t \in [t_0, T]} \left(\|\tilde{v}_h(t, x) - v(t, x)\|_{n-2} + \|\tilde{\rho}_h(t) - \rho(t)\|_{-(k-1)} \right) \le Ch. \tag{88}$$

Indeed, the pair $(\tilde{v}_h, \tilde{\rho}_h) - (v, \rho)$ solves the first order linearized system (52), associated with (u, m), initial condition $(t_0, 0)$ and with a right-hand side given by

$$\begin{aligned} R_{h,1}(t, x) = &-\Bigg(\big(H_p(x_0, x, Du_h, m_h(t)) - H_p(x_0, x, Du, m(t)) \big) \cdot D\tilde{v}_h \\ &+ \left(\frac{\delta H}{\delta m}(x_0, x, Du_h, m_h(t)) - \frac{\delta H}{\delta m}(x_0, x, Du, m(t)) \right)(\tilde{\rho}_h(t)) \Bigg) \\ R_{h,2}(t, x) = &\ \tilde{\rho}_h \big(H_p(x_0, x, Du_h, m_h(t)) - H_p(x_0, x, Du, m(t)) \big) \\ &+ \big(m_h H_{pp}(x_0, x, Du_h, m_h) - m H_{pp}(x_0, x, Du, m) \big) \cdot D\tilde{v}_h \\ &+ \left(m_h \frac{\delta H_p}{\delta m}(x_0, x, Du_h, m_h) - m \frac{\delta H_p}{\delta m}(x_0, x, Du, m) \right)(\tilde{\rho}_h) \end{aligned}$$

$$R_{h,3}(t,x) = \left(\frac{\delta G}{\delta m}(x_0, x, m_h(T)) - \frac{\delta G}{\delta m}(x_0, x, m(T))\right)(\tilde{\rho}_h(T)).$$

Applying Proposition 5.5 and using (87) we infer that (88) holds.

In view of the equations satisfied by $(\tilde{v}_h, \tilde{\rho}_h)$, (v, ρ) and (w, μ), the pair $(\hat{v}_h, \hat{\rho}_h)$ solves the first order linearized system (52), associated with (u, m), initial condition $(t_0, 0)$ and with

$$\begin{aligned}
R_{h,1}(t,x) = &-\Big[\big(H_p(x_0, x, Du_h, m_h(t)) - H_p(x_0, x, Du, m(t))\big) \cdot D\tilde{v}_h \\
&- hH_{pp}(x_0, x, Du, m(t))Dv \cdot Dv' - h\frac{\delta H_p}{\delta m}(x_0, x, Du, m(t))(\rho'(t)) \cdot Dv \\
&+ \left(\frac{\delta H}{\delta m}(x_0, x, Du_h, m_h(t)) - \frac{\delta H}{\delta m}(x_0, x, Du, m(t))\right)(\tilde{\rho}_h(t)) \\
&- h\frac{\delta^2 H}{\delta m^2}(x_0, x, Du, m(t))(\rho(t), \rho'(t)) \\
&- h\frac{\delta H_p}{\delta m}(x_0, x, Du, m(t))(\rho(t)) \cdot Dv'\Big],
\end{aligned}$$

$$\begin{aligned}
R_{h,2}(t,x) = &\ \tilde{\rho}_h(H_p(x_0, x, Du_h, m_h(t)) - H_p) - h\rho\left(H_{pp}Dv' + \frac{\delta H_p}{\delta m}(\rho')\right) \\
&+ D\tilde{v}_h \cdot (m_h H_{pp}(x_0, x, Du_h, m_h(t)) - mH_{pp}) \\
&- hDv \cdot \left(\rho' H_{pp} + m\frac{\delta H_{pp}}{\delta m}(\rho') + mH_{ppp}Dv'\right) \\
&+ \left(m_h\frac{\delta H_p}{\delta m}(x_0, x, Du_h, m_h(t)) - m\frac{\delta H_p}{\delta m}\right)(\tilde{\rho}_h) \\
&- h\left(\rho'\frac{\delta H_p}{\delta m} + mDv' \cdot \frac{\delta H_{pp}}{\delta m}\right)(\rho) - hm\frac{\delta^2 H_p}{\delta m^2}(\rho, \rho'),
\end{aligned}$$

$$\begin{aligned}
R_{h,3}(x) = &\ \frac{\delta G}{\delta m}(x_0, x, m_h(T))(\tilde{\rho}_h(T)) - \frac{\delta G}{\delta m}(x_0, x, m(T))(\tilde{\rho}_h(T)) \\
&- h\frac{\delta^2 G}{\delta m^2}(x_0, x, m(T))(\rho(T), \rho'(T))
\end{aligned}$$

(H_p and its derivatives in $R_{h,2}$ are evaluated at $(x_0, x, Du, m(t))$, unless otherwise specified). Using

$$\begin{aligned}
&\sup_{t\in[t_0,T]} \|u_h(t) - u(t) - hv'(t)\|_{n-2} \leq Ch^2, \\
&\sup_{t\in[t_0,T]} \|m_h(t) - m(t) - h\rho'(t)\|_{-(k-1)} \leq Ch^2
\end{aligned} \tag{89}$$

(see (75) and (76) in Lemma 5.13) as well as the above estimate (87), we have

$$\sup_t \big(\|R_{h,1}(t,\cdot)\|_{n-2} + \|R_{h,2}(t,\cdot)\|_{-(k-1)} + \|R_{h,3}(t,\cdot)\|_{n-2}\big) \leq Ch^2.$$

Then Proposition 5.5 and the representation formula (86) imply that

$$\left\| \frac{\delta U}{\delta m}(t_0, x_0, \cdot, (1-h)m_0 + h\delta_{y'}, y) - \frac{\delta U}{\delta m}(t_0, x_0, \cdot, m_0, y) - hw(t_0, \cdot) \right\|_{n-2}$$
$$= \|\tilde{v}_h(t_0, \cdot) - v(t_0, \cdot) - hw(t_0, \cdot)\|_{n-2} \le \sup_t \|\hat{v}_h(t)\|_{n-2} \le Ch^2.$$

Note that we also have

$$\sup_{t\in[t_0,T]} \|\tilde{\rho}_h(t) - \rho(t) - h\mu(t)\|_{-k} \le Ch^2. \tag{90}$$

Hence, we can apply Lemma B.1 as in the proof of Proposition 5.11 and infer that $\frac{\delta U}{\delta m}$ has a derivative in m given by w:

$$\frac{\delta^2 U}{\delta m^2}(t_0, x_0, x, m_0, y, y') = w(t_0, x).$$

If, in general, w is the solution to the second order linearized system (63) associated with (v, ρ), (v', ρ') (having initial data (t_0, ρ_0) and (t_0, ρ_0') respectively) and with $R_i = 0$, $\tilde{R}_i = 0$, $i = 1, 2, 3$, then by a linearity argument one may also conclude that

$$\int_{\mathbb{R}^d} \frac{\delta^2 U}{\delta m^2}(t_0, x_0, x, m_0, y, y')\, \rho_0(dy)\, \rho_0'(dy') = w(t_0, x). \tag{91}$$

Thus, the estimate of $\frac{\delta^2 U}{\delta m^2}$ follows from Corollary 5.8:

$$\left\| \frac{\delta^2 U}{\delta m^2}(t_0, x_0, \cdot, m_0, \cdot, \cdot) \right\|_{n-2;k-1,k-1} \le \left\| \frac{\delta^2 G}{\delta m^2}(x_0, \cdot, m(T), \cdot, \cdot) \right\|_{n-2;k-1,k-1} + C_M T, \tag{92}$$

using the fact that $\sup_t \|\rho(t)\|_{-(k-1)} \le (1 + C_M T)\|\rho_0\|_{-(k-1)}$ and that the same holds for ρ'. ■

Next we discuss the second order regularity of U and U^0 with respect to m and x_0.

Proposition 5.17. *Let U^0 and U be the solutions of (46) and (43) respectively. Suppose, in addition to the assumptions of Propositions 5.15 and 5.16, that*

$$\|D^2_{x_0}(G^0, G)\|_{n-2} + \left\| D_{x_0} \frac{\delta(G^0, G)}{\delta m} \right\|_{n-2;k-1} + \left\| \frac{\delta^2(G^0, G)}{\delta m^2} \right\|_{n-2;k-1,k-1} \le M.$$

Then there exists $T_M > 0$ (depending on M and on the data but not on G) such that if $T \in (0, T_M]$, the maps U^0 and U are C^2 with respect to the measure variable and x_0, and

$$\sup_t \|D^2_{x_0}(U^0, U)(t)\|_{n-2} \le \|(D^2_{x_0} G^0, D^2_{x_0} G)\|_{n-2} + C_M T,$$

$$\sup_t \left\| D_{x_0} \frac{\delta(U^0, U)}{\delta m}(t) \right\|_{n-2;k-1} \le \left\| D_{x_0} \frac{\delta(G^0, G)}{\delta m} \right\|_{n-2;k-1} + C_M T.$$

Moreover,

$$\sup_t \left\| \frac{\delta^2(U^0, U)}{\delta m^2}(t) \right\|_{n-2;k-1,k-1} \leq \left\| \frac{\delta^2(G^0, G)}{\delta m^2} \right\|_{n-2;k-1,k-1} + C_M T.$$

Proof. Step 1. The *differentiability of* $\frac{\delta U}{\delta m}$ *with respect to* x_0 can be achieved exactly as its differentiability with respect to m in Proposition 5.16. For any direction $\xi \in \mathbb{R}^{d_0}$, let

- (u, m) (respectively (u_h, m_h)) be the solution of the MFG system (45) with initial condition (t_0, m_0) and parameters x_0 and $x_0 + h\xi$ respectively,
- (v, ρ) (respectively (v', ρ')) be the solution of the first order linearized system (52) with zero right-hand side (respectively right-hand side as in (74)), initial condition (t_0, δ_y) (respectively $(t_0, 0)$) and where the Hamiltonian and its derivatives are evaluated at $(x_0, x, Du(t, x), m(t))$,
- $(\tilde{v}_h, \tilde{\rho}_h)$ be the solution to the first order linearized system (52) with zero right-hand side, with initial condition (t_0, δ_y) and where the Hamiltonian and its derivatives are evaluated at $(x_0 + h\xi, x, Du_h(t, x), m_h(t))$,
- (w, μ) be the solution to the second order linearized system (63) associated with (v, ρ), (v', ρ') (and (u, m)), and with right-hand side

$$\begin{aligned}
\tilde{R}_1(t, x) &= -H_{x_0 p}(x_0, x, Du, m(t))\xi \cdot Dv - \frac{\delta H_{x_0}}{\delta m}(x_0, x, Du, m(t))(\rho(t)) \cdot \xi, \\
\tilde{R}_2(t, x) &= \rho H_{x_0 p}(x_0, x, Du, m(t))\xi + m H_{x_0 pp}(x_0, x, Du, m(t))\xi Dv \\
&\quad + m \frac{\delta H_{x_0 p}}{\delta m}(\rho)\xi, \\
\tilde{R}_3(x) &= \frac{\delta G_{x_0}}{\delta m}(x_0, x, m(T))(\rho(T)) \cdot \xi,
\end{aligned}$$

so that

$$\begin{aligned}
\tilde{v}_h(t_0, x) &= \frac{\delta U}{\delta m}(t_0, x_0 + h\xi, x, m_0, y), \\
v(t_0, x) &= \frac{\delta U}{\delta m}(t_0, x_0, x, m_0, y), \quad v'(t_0, x) = U_{x_0}(t_0, x_0, x, m_0) \cdot \xi.
\end{aligned}$$

Then we find $\frac{\delta U_{x_0}}{\delta m}(t_0, x_0, x, m_0, y) \cdot \xi = w(t_0, x)$, and if one replaces δ_y by an arbitrary $\rho_0 \in C^{-(k-1)}$ as the initial datum for ρ, the following representation holds:

$$\frac{\delta U_{x_0}}{\delta m}(t_0, x_0, x, m_0)(\rho_0) \cdot \xi = w(t_0, x). \tag{93}$$

Step 2. The *second order differentiability of* U *with respect to* x_0 can be checked in a similar way. Let (u, m) and (u_h, m_h) be as before, and let

- (v, ρ), $(\tilde{v}_h, \tilde{\rho}_h)$ be the solutions of (52) with right-hand side as in (74), initial condition $(t_0, 0)$, and Hamiltonian and its derivatives evaluated at $(x_0, x, Du(t, x), m(t))$ and $(x_0 + h\xi, x, Du_h(t, x), m_h(t))$ respectively,

- (w,μ) be the solution to (63) associated with (v,ρ), $(v',\rho')=(v,\rho)$ (and (u,m)), and with right-hand side $\tilde{R}_1, \tilde{R}_2, \tilde{R}_3$ given by (71).

Then we find

$$D^2_{x_0}U(t_0,x_0,x,m_0)\xi\cdot\xi = w(t_0,x). \tag{94}$$

Step 3. We now prove the regularity of U^0. To show that $\frac{\delta U^0}{\delta m}$ is differentiable with respect to m, let $(t_0,m_0)\in[0,T)\times\mathcal{P}_2$, $y,y'\in\mathbb{R}^d$ and

- (u,m) (respectively (u_h,m_h)) be the solution of (45) with initial condition (t_0,m_0) (respectively $(t_0,(1-h)m_0+h\delta_{y'})$),
- (v,ρ) (respectively (v',ρ')) be the solution of (52) with zero right-hand side, initial condition (t_0,δ_y) (respectively $(t_0,\delta_{y'})$) and where the Hamiltonian and its derivatives are evaluated at $(x_0,x,Du(t,x),m(t))$,
- $(\tilde{v}_h,\tilde{\rho}_h)$ be the solution to (52) with zero right-hand side, with initial condition (t_0,δ_y) and where the Hamiltonian and its derivatives are evaluated at $(x_0,x,Du_h(t,x),m_h(t))$,
- (w,μ) be the solution to (63) associated with (v,ρ), (v',ρ') (and (u,m)), and with right-hand side 0,

as in the proof of differentiability of $\frac{\delta U}{\delta m}$ with respect to m in Proposition 5.16. Note that

$$\frac{\delta U^0}{\delta m}(t_0,x_0,(1-h)m_0+h\delta_{y'},y) = \frac{\delta G^0}{\delta m}(x_0,m_h(T))(\tilde{\rho}_h(T)),$$
$$\frac{\delta U^0}{\delta m}(t_0,x_0,m_0,y) = \frac{\delta G^0}{\delta m}(x_0,m(T))(\rho(T)).$$

Therefore, using (89) and (90),

$$\Bigg|\frac{\delta G^0}{\delta m}(x_0,m_h(T))(\tilde{\rho}_h(T)) - \frac{\delta G^0}{\delta m}(x_0,m(T))(\rho(T)) - h\left(\frac{\delta^2 G^0}{\delta m^2}(x_0,m(T))(\rho(T),\rho'(T)) + \frac{\delta G^0}{\delta m}(x_0,m(T))(\mu(T))\right)\Bigg| \le Ch^2.$$

Lemma B.1 then implies that $\frac{\delta U^0}{\delta m}(t_0,x_0,\cdot,y)$ has a derivative, and by linearity, if μ is the solution to (63) associated with (v,ρ), (v',ρ') (which in turn have initial data (t_0,ρ_0) and (t_0,ρ_0') respectively and with zero right-hand side), then

$$\int_{\mathbb{R}^d}\frac{\delta^2 U^0}{\delta m^2}(t_0,x_0,m_0,y,y')\,\rho_0(dy)\,\rho_0'(dy') = \frac{\delta^2 G^0}{\delta m^2}(x_0,m(T))(\rho(T),\rho'(T)) + \frac{\delta G^0}{\delta m}(x_0,m(T))(\mu(T)). \tag{95}$$

Hence, by the representation formula (91) for $\frac{\delta^2 U}{\delta^2 m}$, Propositions 5.5 and 5.6 and Corollary 5.8, we have, for $r\le n-2$,

$$\begin{aligned}
&\left|\frac{\delta^2 U^0}{\delta m^2}(t,x_0,m_0)(\rho_0,\rho_0')\right|^2+\left|D_x^r\frac{\delta^2 U}{\delta m^2}(t,x_0,x,m_0)(\rho_0,\rho_0')\right|^2\\
&\quad=\left(\left|\frac{\delta^2 G^0}{\delta m^2}(x_0,m(T))(\rho(T),\rho'(T))\right|+\left|\frac{\delta G^0}{\delta m}(x_0,m(T))(\mu(T))\right|\right)^2+|D_x^r w(t,x)|^2\\
&\quad\le\left(\left|\frac{\delta^2 G^0}{\delta m^2}(x_0,m(T))(\rho(T),\rho'(T))\right|+C_M T\|\rho_0\|_{-(k-1)}\|\rho_0'\|_{-(k-1)}\right)^2\\
&\qquad+\left(\sup_x\left|D_x\frac{\delta^2 G}{\delta m^2}(x_0,x,m(T))(\rho(T),\rho'(T))\right|+C_M T\|\rho_0\|_{-(k-1)}\|\rho_0'\|_{-(k-1)}\right)^2\\
&\quad\le\left\{\sup_x\frac{1}{\|\rho(T)\|_{-(k-1)}\|\rho'(T)\|_{-(k-1)}}\right.\\
&\qquad\times\left(\left|\frac{\delta^2 G^0}{\delta m^2}(x_0,m(T))(\rho(T),\rho'(T))\right|^2+\left|D_x^r\frac{\delta^2 G}{\delta m^2}(x_0,x,m(T))(\rho(T),\rho'(T))\right|^2\right)^{1/2}\\
&\qquad\left.\times(1+C_M T)+C_M T\right\}^2\|\rho_0\|_{-(k-1)}^2\|\rho_0'\|_{-(k-1)}^2,
\end{aligned}$$

(where we use $(x+z)^2+(y+z)^2\le((x^2+y^2)^{1/2}+2z)^2$ for $x,y,z\ge 0$). Taking the square root, then sup over x_0, ρ_0 and ρ_0' and summing over $r\le n-2$ gives the estimate on $\|\frac{\delta^2(U^0,U)}{\delta m^2}\|_{n-2;k-1,k-1}$.

Differentiability of $\frac{\delta U^0}{\delta m}$ with respect to x_0 follows analogous lines: (v,ρ), (v',ρ'), $(\tilde v_h,\tilde\rho_h)$ and (w,μ) have to be changed according to Step 1. By (93), we have, using the notations of Corollary 5.9 and for any $r\le n-2$,

$$\begin{aligned}
&\sum_{|l|=1}\left(\left|\partial_{x_0}^l\frac{\delta U^0}{\delta m}(t_0,x_0,m_0,y)(\rho_0)\right|^2+\left|D_x^r\partial_{x_0}^l\frac{\delta U}{\delta m}(t,x_0,x,m_0)(\rho_0)\right|^2\right)\\
&\quad=\sum_{|l|=1}\left(\left|\partial_{x_0}^l\frac{\delta G^0}{\delta m}(\rho(T))+\frac{\delta^2 G^0}{\delta m^2}(\rho(T),\rho^l(T))+\frac{\delta G^0}{\delta m}(\mu^l(T))\right|^2+|D_x^r w^l(t,x)|^2\right),
\end{aligned}$$

where G^0 and its derivatives are all evaluated at $(x_0,m(T))$. We obtain the bounds on $(\frac{\delta U^0_{x_0}}{\delta m},\frac{\delta U_{x_0}}{\delta m})$ by using Propositions 5.5 and 5.6 and Corollary 5.9.

Finally, second order differentiability of U^0 with respect to x_0, and the corresponding bound, can be obtained similarly. Let $l,l'\in\mathbb{R}^{d_0}$ with $|l|=|l'|=1$, (v^l,ρ^l), $(v^{l'},\rho^{l'})$ and $(w^{l,l'},\mu^{l,l'})$ be as in Corollary 5.10. Note that

$$\begin{aligned}
\partial_{x_0}^{l+l'}U^0(t_0,x_0,m_0)&=\partial_{x_0}^{l+l'}G^0+\partial_{x_0}^l\frac{\delta G^0_{x_0}}{\delta m}(\rho^{l'}(T))+\partial_{x_0}^{l'}\frac{\delta G^0_{x_0}}{\delta m}(\rho^l(T))\\
&\quad+\frac{\delta^2 G^0}{\delta m^2}(\rho^l(T),\rho^{l'}(T))+\frac{\delta G^0}{\delta m}(\mu^{l,l'}(T)),
\end{aligned}$$

while $\partial_{x_0}^{l+l'}U^0(t_0,x_0,m_0)$ is given by polarizing the representation formula (94). We can then conclude the proof by using Propositions 5.5 and 5.6 and Corollary 5.10. ■

5.2.3. Uniform continuity estimates on second order derivatives

Proposition 5.18. *Let U be the solution of (43) given by (44) and $n \geq 4$, $k \in \{3, \dots, n-1\}$. Suppose, in addition to the assumptions of Proposition 5.16, that*

$$\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2}\right) \leq M. \tag{96}$$

Then there exists $T_M > 0$ (depending on M and on the data but not on G) such that

$$\sup_t \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U}{\delta m^2}(t)\right) \leq \sup_{x_0} \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G}{\delta m^2}\right) + C_M T.$$

Proof. We establish for later use a slightly stronger estimate involving the dependence on x_0. This is used in Proposition 5.19 below. Fix $(t_0, m_1, m_2) \in [0, T] \times \mathcal{P}_2^2$ and $x_0^1, x_0^2 \in \mathbb{R}^{d_0}$. We use the representation formula (91) for $\frac{\delta^2 U}{\delta m^2}(t_0, x_0^1, m_1)$ and $\frac{\delta^2 U}{\delta m^2}(t_0, x_0^2, m_2)$. In particular, for $i = 1, 2$, we let

- (u^i, m^i) be the solution to (45) starting from m_i at time t_0 with H (and G) evaluated at $(x_0^i, x, Du^i(t,x), m^i(t))$ (and $(x_0^i, x, m^i(T))$),
- (v_i, ρ_i) (respectively (v_i', ρ_i')) be the solution of (52) with zero right-hand side, initial condition (t_0, ρ_0) (respectively (t_0, ρ_0')) and where the Hamiltonian and its derivatives are evaluated at $(x_0^i, x, Du^i(t,x), m^i(t))$,
- (w^i, μ^i) be the solution to (63) associated with (v_i, ρ_i), (v_i', ρ_i') (and x_0^i, u^i, m^i), and with zero right-hand side.

We aim at estimating $(\bar{w}, \bar{\mu}) := (w^1 - w^2, \mu^1 - \mu^2)$, since

$$\bar{w}(t_0, x) = \frac{\delta^2 U}{\delta m^2}(t_0, x_0^1, x, m_1)(\rho_0, \rho_0') - \frac{\delta^2 U}{\delta m^2}(t_0, x_0^2, x, m_2)(\rho_0, \rho_0'). \tag{97}$$

We first set $(\bar{v}, \bar{\rho}) := (v_1 - v_2, \rho_1 - \rho_2)$ and $(\bar{v}', \bar{\rho}') := (v_1' - v_2', \rho_1' - \rho_2')$. The pair $(\bar{v}, \bar{\rho})$ solves (52) with zero initial datum, with H and its derivatives evaluated at $(x_0^1, x, Du^1(t,x), m^1(t))$, and with right-hand side

$$R_1(t,x) = -(H_p^1 - H_p^2) \cdot Dv_2 - \left(\frac{\delta H^1}{\delta m} - \frac{\delta H^2}{\delta m}\right)(\rho_2(t)),$$

$$R_2(t,x) = \rho_2(H_p^1 - H_p^2) + (m^1 H_{pp}^1 - m^2 H_{pp}^2) Dv_2 + \left(m^1 \frac{\delta H_p^1}{\delta m} - m^2 \frac{\delta H_p^2}{\delta m}\right)(\rho_2),$$

$$R_3(x) = \left(\frac{\delta G^1}{\delta m} - \frac{\delta G^2}{\delta m}\right)(\rho_2(T)),$$

where H^i and its derivatives correspond to H and its derivatives evaluated at $(x_0^i, x, Du^i(t,x), m^i(t))$.

By Proposition 5.5 we have

$$\sup_{t \in [t_0, T]} \|v_i(t)\|_{n-1} \leq C \|\rho_0\|_{-(k-2)}, \qquad \sup_{t \in [t_0, T]} \|\rho_i(t)\|_{-(k-2)} \leq (1 + CT) \|\rho_0\|_{-(k-2)}, \tag{98}$$

where C depends on the regularity of $\frac{\delta G}{\delta m}$, H_{x_0}, $H_{x_0 p}$, m^i and $\sup_t \|u^i\|_n$. Note that, by the above estimates and Proposition 5.4,

$$\sup_t \|R_1(t)\|_{n-2} + \sup_t \|R_2(t)\|_{-(k-2)} + \|R_3\|_{n-2} \\ \leq C(\mathbf{d}_2(m_1, m_2) + |x_0^1 - x_0^2|)\|\rho_0\|_{-(k-2)},$$

and therefore by Proposition 5.5 (applied to $n-1 \geq 2$ and $k-2 \geq 1$) we obtain

$$\sup_t \|\bar{v}(t)\|_{n-2} \leq CT(\mathbf{d}_2(m_1, m_2) + |x_0^1 - x_0^2|)\|\rho_0\|_{-(k-2)}, \tag{99}$$

$$\sup_t \|\bar{\rho}(t)\|_{-(k-1)} \leq CT(\mathbf{d}_2(m_1, m_2) + |x_0^1 - x_0^2|)\|\rho_0\|_{-(k-2)}. \tag{100}$$

Completely analogous estimates hold for v_i', ρ_i' and their differences $\bar{v}'$, $\bar{\rho}'$.

We now proceed by estimating $(\bar{w}, \bar{\mu})$, which solves the first order linearized system with zero initial datum, with H and its derivatives evaluated at $(x_0^1, x, Du^1(t,x), m^1(t))$, and with right-hand side

$$\begin{aligned}
\overline{R}_1(t,x) := -\Bigg(& (H_p^1 - H_p^2) \cdot Dw^2 + \left(\frac{\delta H^1}{\delta m} - \frac{\delta H^2}{\delta m}\right)(\mu^2(t)) \\
& + \frac{\delta^2 H^1}{\delta m^2}(\rho_1(t), \rho_1'(t)) - \frac{\delta^2 H^2}{\delta m^2}(\rho_2(t), \rho_2'(t)) + H_{pp}^1 Dv_1 \cdot Dv_1' \\
& - H_{pp}^2 Dv_2 \cdot Dv_2' + \frac{\delta H_p^1}{\delta m}(\rho_1) \cdot Dv_1' - \frac{\delta H_p^2}{\delta m}(\rho_2) \cdot Dv_2' + \frac{\delta H_p^1}{\delta m}(\rho_1') \cdot Dv_1 \\
& - \frac{\delta H_p^2}{\delta m}(\rho_2') \cdot Dv_2 \Bigg),
\end{aligned}$$

$$\begin{aligned}
\overline{R}_2(t,x) := {} & \mu^2(H_p^1 - H_p^2) + (m^1 H_{pp}^1 - m^2 H_{pp}^2) Dw^2 + \left(m^1 \frac{\delta H_p^1}{\delta m} - m^2 \frac{\delta H_p^2}{\delta m}\right)(\mu^2) \\
& + \rho_1 H_{pp}^1 Dv_1' - \rho_2 H_{pp}^2 Dv_2' + \rho_1' H_{pp}^1 Dv_1 - \rho_2' H_{pp}^2 Dv_2 \\
& + m^1 H_{ppp}^1 Dv_1 Dv_1' - m^2 H_{ppp}^2 Dv_2 Dv_2' + m^1 \frac{\delta^2 H_p^1}{\delta m^2}(\rho_1, \rho_1') \\
& - m^2 \frac{\delta^2 H_p^2}{\delta m^2}(\rho_2, \rho_2') + \rho_1 \frac{\delta H_p^1}{\delta m}(\rho_1') - \rho_2 \frac{\delta H_p^2}{\delta m}(\rho_2') + \rho_1' \frac{\delta H_p^1}{\delta m}(\rho_1) \\
& - \rho_2' \frac{\delta H_p^2}{\delta m}(\rho_2) + m^1 \frac{\delta H_{pp}^1}{\delta m}(\rho_1') Dv_1 - m^2 \frac{\delta H_{pp}^2}{\delta m}(\rho_2') Dv_2 \\
& + m^1 \frac{\delta H_{pp}^1}{\delta m}(\rho_1) Dv_1' - m^2 \frac{\delta H_{pp}^2}{\delta m}(\rho_2) Dv_2',
\end{aligned}$$

$$\overline{R}_3(x) := \frac{\delta^2 G^1}{\delta m^2}(\rho_1(T), \rho_1'(T)) - \frac{\delta^2 G^2}{\delta m^2}(\rho_2(T), \rho_2'(T)) + \left(\frac{\delta G^1}{\delta m} - \frac{\delta G^2}{\delta m}\right)(\mu^2(T)).$$

Recall also that Proposition 5.6 and Remark 5.7 (applied to $n-1$ and $k-1$) yield

$$\begin{aligned}
\sup_t \|w^i(t)\|_{n-3} &\leq (1 + CT)\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)}, \\
\sup_t \|\mu^i(t)\|_{-(k-1)} &\leq CT\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)}.
\end{aligned} \tag{101}$$

By (98)–(101) we get

$$\sup_t \|\overline{R}_1(t)\|_{n-3} + \sup_t \|\overline{R}_2(t)\|_{-(k-1)} \leq CT(\mathbf{d}_2(m_1,m_2) + |x_0^1 - x_0^2|)\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)}.$$

Similarly, using also the Lipschitz regularity of $\frac{\delta G}{\delta m}$,

$$\|\overline{R}_3\|_{n-3} \leq (1+CT)\left\|\frac{\delta^2 G}{\delta m^2}(x_0^2, m^2(T)) - \frac{\delta^2 G}{\delta m^2}(x_0^1, m^1(T))\right\|_{n-3;k-2,k-2} \times \|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)} + CT(\mathbf{d}_2(m_1,m_2) + |x_0^1 - x_0^2|)\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)},$$

Then, recalling that $\bar{w} = w^1 - w^2$ satisfies (97), we obtain in view of (56) in Proposition 5.5 and for any $r \leq n-3$,

$$\begin{aligned}\left\|D_x^r \frac{\delta^2 U}{\delta m^2}(t_0, x_0^2, m_2) - D_x^r \frac{\delta^2 U}{\delta m^2}(t_0, x_0^1, m_1)\right\|_{0;k-2,k-2} &\leq (1+C_M T)\left\|D_x^r \frac{\delta^2 G}{\delta m^2}(x_0^2, m_2(T)) - D_x^r \frac{\delta^2 G}{\delta m^2}(x_0^1, m_1(T))\right\|_{0;k-2,k-2} \\ &\quad + C_M T(\mathbf{d}_2(m_1,m_2) + |x_0^1 - x_0^2|).\end{aligned} \tag{102}$$

Choosing $x_0^1 = x_0^2$, summing over $r \leq n-3$ and recalling Proposition 5.4 and (96) then gives the claim.

Note that we also have the following inequality for $\bar{\mu} = \mu^1 - \mu^2$, which will be useful in the next proposition:

$$\sup_{t\in[t_0,T]} \|\mu^1(t) - \mu^2(t)\|_{-k} \leq CT(\mathbf{d}_2(m_1,m_2) + |x_0^1 - x_0^2|)\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)}. \tag{103}$$

■

Finally, we establish the Lipschitz regularity of the second order derivatives of G^0 and G with respect to x_0 and m.

Proposition 5.19. *Let U be the solution of (43) given by (44) and U^0 be the solution to (46) given by (47). Suppose that the assumptions of Proposition 5.18 hold and that in addition*

$$\mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2}\right) + \mathrm{Lip}_{n-3;k-2}\left(\frac{\delta G^0_{x_0}}{\delta m}, \frac{\delta G_{x_0}}{\delta m}\right) + \mathrm{Lip}_{n-3}(D^2_{x_0}G^0, D^2_{x_0}G) \leq M,$$

$$\mathrm{Lip}^{x_0}_{n-3;k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2}\right) + \mathrm{Lip}^{x_0}_{n-3;k-2}\left(\frac{\delta G^0_{x_0}}{\delta m}, \frac{\delta G_{x_0}}{\delta m}\right) + \mathrm{Lip}^{x_0}_{n-3}(D^2_{x_0}G^0, D^2_{x_0}G) \leq M,$$

for some $n \geq 4$ *and* $k \in \{3, \dots, n-1\}$. *Then*

$$\sup_t \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 U^0(t)}{\delta m^2}, \frac{\delta^2 U(t)}{\delta m^2}\right) \leq \mathrm{Lip}_{n-3;k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2}\right) + C_M T,$$
$$\sup_t \mathrm{Lip}^{x_0}_{n-3;k-2,k-2}\left(\frac{\delta^2 U^0(t)}{\delta m^2}, \frac{\delta^2 U(t)}{\delta m^2}\right) \leq \mathrm{Lip}^{x_0}_{n-3;k-2,k-2}\left(\frac{\delta^2 G^0}{\delta m^2}, \frac{\delta^2 G}{\delta m^2}\right) + C_M T,$$
$$\sup_t \mathrm{Lip}_{n-3;k-2}\left(\frac{\delta U^0_{x_0}(t)}{\delta m}, \frac{\delta U_{x_0}(t)}{\delta m}\right) \leq \mathrm{Lip}_{n-3;k-2}\left(\frac{\delta G^0_{x_0}}{\delta m}, \frac{\delta G_{x_0}}{\delta m}\right) + C_M T,$$
$$\sup_t \mathrm{Lip}^{x_0}_{n-3;k-2}\left(\frac{\delta U^0_{x_0}(t)}{\delta m}, \frac{\delta U_{x_0}(t)}{\delta m}\right) \leq \mathrm{Lip}^{x_0}_{n-3;k-2}\left(\frac{\delta G^0_{x_0}}{\delta m}, \frac{\delta G_{x_0}}{\delta m}\right) + C_M T,$$
$$\mathrm{Lip}^{x_0}_{n-3}(D^2_{x_0} U^0(t), D^2_{x_0} U(t)) \leq \mathrm{Lip}^{x_0}_{n-3}(D^2_{x_0} G^0, D^2_{x_0} G) + C_M T.$$

Proof. We will detail only the proof of Lipschitz estimates of $(\frac{\delta^2 U^0}{\delta m^2}, \frac{\delta^2 U}{\delta m^2})$. Lipschitz regularity of $\frac{\delta U^0_{x_0}}{\delta m}$ and $D^2_{x_0} U^0$, $\frac{\delta U_{x_0}}{\delta m}$ and $D^2_{x_0} U$ can be proven along the same lines using the representation formulas that appear in the proof of Proposition 5.17.

Let us start with $\frac{\delta^2 U^0}{\delta m^2}$. Fix $(t_0, m_1, m_2) \in [0, T] \times \mathcal{P}_2^2$ and $x_0^1, x_0^2 \in \mathbb{R}^d$. Also, as in the proof of Proposition 5.17, for $i = 1, 2$ let

- (u^i, m^i) be the solution to the MFG system (45) starting from m_i at time t_0 with H (and G) evaluated at $(x_0^i, x, Du^i(t, x), m^i(t))$ (and $(x_0^i, x, m^i(T))$),
- (v_i, ρ_i) (respectively (v_i', ρ_i')) be the solution of the first order linearized system (52) with zero right-hand side, initial condition (t_0, ρ_0) (respectively (t_0, ρ_0')) and where the Hamiltonian and its derivatives are evaluated at $(x_0^i, x, Du^i(t, x), m^i(t))$,
- (w^i, μ^i) be the solution to the second order linearized system (63) associated with $(v_i, \rho_i), (v_i', \rho_i')$ (and (u^i, m^i)), and with zero right-hand side.

Recall that (95) provides a representation formula for $\frac{\delta^2 U^0}{\delta m^2}$:

$$\frac{\delta^2 U^0}{\delta m^2}(t_0, x_0^i, m_i)(\rho_0, \rho_0')$$
$$= \frac{\delta^2 G^0}{\delta m^2}(x_0^i, m^i(T))(\rho_i(T), \rho_i'(T)) + \frac{\delta G^0}{\delta m}(x_0^i, m^i(T))(\mu^i(T)),$$

and $\frac{\delta^2 U}{\delta m^2}(t_0, x_0^i, x, m_i)(\rho_0, \rho_0') = w^i(t_0, x)$. Let us recall the inequalities

$$\sup_{t \in [t_0, T]} \mathbf{d}_2(m^1(t), m^2(t)) \leq (1 + CT)\mathbf{d}_2(m_0^1, m_0^2) + CT|x_0^1 - x_0^2|,$$
$$\sup_{t \in [t_0, T]} \|\rho_i(t)\|_{-(k-2)} \leq (1 + CT)\|\rho_0\|_{-(k-2)},$$
$$\sup_{t \in [t_0, T]} \|\rho_i'(t)\|_{-(k-2)} \leq (1 + CT)\|\rho_0'\|_{-(k-2)},$$
$$\sup_{t \in [t_0, T]} \|\rho_1(t) - \rho_2(t)\|_{-(k-1)} \leq CT\big(\mathbf{d}_2(m_1, m_2) + |x_0^1 - x_0^2|\big)\|\rho_0\|_{-(k-2)},$$

$$\sup_{t\in[t_0,T]} \|\rho_1'(t)-\rho_2'(t)\|_{-(k-1)} \le CT(\mathbf{d}_2(m_1,m_2)+|x_0^1-x_0^2|)\|\rho_0'\|_{-(k-2)},$$
$$\sup_{t\in[t_0,T]} \|\mu^i(t)\|_{-(k-1)} \le CT\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)},$$
$$\sup_{t\in[t_0,T]} \|\mu^1(t)-\mu^2(t)\|_{-k} \le CT(\mathbf{d}_2(m_1,m_2)+|x_0^1-x_0^2|)\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)},$$

which are consequences of Proposition 5.4, (98), (100), (101) and (103). Setting

$$\theta_T := CT(\mathbf{d}_2(m_1,m_2)+|x_0^1-x_0^2|)\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)},$$

we obtain, using also (102), for any $r \le n-3$,

$$\begin{aligned}
&\left|\left(\frac{\delta^2 U^0}{\delta m^2}(t,x_0^1,m_1)-\frac{\delta^2 U^0}{\delta m^2}(t,x_0^2,m_2)\right)(\rho_0,\rho_0')\right|^2 \\
&\qquad + \sup_x \left|D_x^r\left(\frac{\delta^2 U}{\delta m^2}(t,x_0^1,x,m_1)-\frac{\delta^2 U}{\delta m^2}(t,x_0^2,x,m_2)\right)(\rho_0,\rho_0')\right|^2 \\
&\le (1+CT) \\
&\quad \times \left\{\left|\frac{\delta^2 G^0}{\delta m^2}(x_0^1,m^1(T))(\rho_1(T),\rho_1'(T))-\frac{\delta^2 G^0}{\delta m^2}(x_0^2,m^2(T))(\rho_1(T),\rho_1'(T))\right|+\theta_T\right\}^2 \\
&\quad + (1+CT)\left\{\sup_x \left|D_x^r\frac{\delta^2 G}{\delta m^2}(x_0^1,x,m^1(T))(\rho_1(T),\rho_1'(T))\right.\right. \\
&\qquad\qquad \left.\left. - D_x^r\frac{\delta^2 G}{\delta m^2}(x_0^2,x,m^2(T))(\rho_1(T),\rho_1'(T))\right|+\theta_T\right\}^2.
\end{aligned}$$

Choosing $m_1=m_2=m$ and rearranging gives Lipschitz estimates in x_0:

$$\begin{aligned}
&\left|\left(\frac{\delta^2 U^0}{\delta m^2}(t,x_0^1,m)-\frac{\delta^2 U^0}{\delta m^2}(t,x_0^2,m)\right)(\rho_0,\rho_0')\right|^2 \\
&\qquad + \sup_x \left|D_x^r\left(\frac{\delta^2 U}{\delta m^2}(t,x_0^1,x,m)-\frac{\delta^2 U}{\delta m^2}(t,x_0^2,x,m)\right)(\rho_0,\rho_0')\right|^2 \\
&\le (1+CT)\left\{\left(\left|\frac{\delta^2 G^0}{\delta m^2}(x_0^1,m^1(T))(\rho_1(T),\rho_1'(T))-\frac{\delta^2 G^0}{\delta m^2}(x_0^2,m^1(T))(\rho_1(T),\rho_1'(T))\right|^2\right.\right. \\
&\quad + \sup_x \left|D_x^r\frac{\delta^2 G}{\delta m^2}(x_0^1,x,m^1(T))(\rho_1(T),\rho_1'(T))\right. \\
&\quad \left.\left. - D_x^r\frac{\delta^2 G}{\delta m^2}(x_0^2,x,m^1(T))(\rho_1(T),\rho_1'(T))\right|^2\right)^{1/2} \\
&\quad \left. + CT|x_0^1-x_0|\|\rho_0\|_{-(k-2)}\|\rho_0'\|_{-(k-2)}\right\}^2,
\end{aligned}$$

while the choice $x_0^1=x_0^2$ similarly gives Lipschitz estimates in m. ∎

Appendix A. Estimates for solutions to HJ equations

A.1. Main estimates

In this section, we assume that the data a, h and g are smooth and we are looking for a priori estimates on the smooth and globally bounded solution u to the HJ equation

$$\begin{cases} -\partial_t u(t,x) - \mathrm{Tr}(a(t,x)D^2u(t,x)) + h(t,x,Du(t,x)) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ u(T,x) = g(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{104}$$

We always assume that there exist $C_0 > 0$ and $\gamma \geq 1$ such that

$$a(t,x) \geq C_0^{-1} I_d, \quad \|Da\|_\infty \leq C_0$$

and

$$|D_x h(t,x,p)| \leq C_0(1+|p|^\gamma)$$

for every $(t,x,p) \in (0,T)\times\mathbb{R}^d\times\mathbb{R}^d$.

Proposition A.1 (Lipschitz estimates). *For any $M > 0$ there exists $T_M, C_M > 0$, depending on M, C_0 and γ, such that if $T \in (0, T_M)$ and $\|Dg\|_\infty \leq M$, then*

$$\sup_{t\in[0,T]} \|Du(t)\|_\infty \leq \|Dg\|_\infty + C_M T.$$

Proof. We use the standard Bernstein method. Let $v(t,x) = \sum_i u_i^2(t,x)$. Then

$$\partial_t v(t,x) = 2\sum_i u_i(t,x)u_{i,t}(t,x), \quad v_j(t,x) = 2\sum_i u_i(t,x)u_{ij}(t,x),$$

$$v_{jk}(t,x) = 2\sum_i \big(u_{ik}(t,x)u_{ij}(t,x) + u_i(t,x)u_{ijk}(t,x)\big).$$

Thus

$$\begin{aligned} &-\partial_t v - \mathrm{Tr}(a(t,x)D^2v(t,x)) \\ &\quad = -2\sum_i u_i(t,x)u_{i,t}(t,x) - 2\sum_{i,j,k} a_{jk}(t,x)(u_{ik}(t,x)u_{ij}(t,x) + u_i(t,x)u_{ijk}(t,x) \\ &\quad = -2\sum_{i,j,k} a_{jk}(t,x)u_{ik}(t,x)u_{ij}(t,x) - 2\sum_i u_i(t,x)D_i\big(\partial_t u + \mathrm{Tr}(a(t,x)D^2u(t,x))\big) \\ &\qquad + 2\sum_{i,j,k} u_i(t,x)(a_{jk})_i(t,x)u_{jk}(t,x), \end{aligned}$$

where $(a_{jk})_i$ denotes the x_i-derivative of the element a_{jk} of the matrix $a(t,x)$.

Using the equation for u we find

$$
\begin{aligned}
-\partial_t v &- \mathrm{Tr}(a(t,x)D^2 v(t,x)) \\
&= -2\sum_{i,j,k} a_{jk}(t,x)u_{ik}(t,x)u_{ij}(t,x) \\
&\quad - 2\sum_i u_i(t,x)(h_i(t,x,Du(t,x)) + h_p(t,x,Du(t,x))\cdot Du_i(t,x)) \\
&\quad + 2\sum_{i,j,k} u_i(t,x)(a_{jk})_i(t,x)u_{jk}(t,x).
\end{aligned} \tag{105}
$$

Using our assumptions on a and h, we infer that

$$
\begin{aligned}
-\partial_t v - \mathrm{Tr}(a(t,x)D^2 v(t,x)) + h_p(t,x,Du(t,x))\cdot Dv(t,x) \\
\le -2C_0^{-1}|D^2u|^2 + 2C_0|Du|(1+|Du|^\gamma) + \|Da\|_\infty |Du|\,|D^2u| \\
\le 2C_0|Du|(1+|Du|^\gamma) + c_d\|Da\|_\infty^2 C_0|Du|^2
\end{aligned}
$$

for some constant c_d only depending on the dimension d. In particular, by the maximum principle we estimate

$$
\begin{aligned}
\|v\|_{L^\infty(Q_T)} &\le \|Dg\|^2_{L^\infty(Q_T)} \\
&+ T[2C_0\|Du\|_{L^\infty(Q_T)}(1+\|Du\|^\gamma_{L^\infty(Q_T)}) + c_d\|Da\|^2_\infty C_0\|Du\|^2_{L^\infty(Q_T)}],
\end{aligned} \tag{106}
$$

which implies

$$
\begin{aligned}
\|v\|_{L^\infty(Q_T)} \le \|Dg\|^2_{L^\infty(Q_T)} + 4T^2C_0^2 + \tfrac{1}{4}\|Du\|^2_{L^\infty(Q_T)} \\
+ \hat{C}T\|Du\|^2_{L^\infty(Q_T)}[\|Du\|^{\gamma-1}_{L^\infty(Q_T)} + 1]
\end{aligned} \tag{107}
$$

for some $\hat{C}$ only depending on d and C_0. Recall that

$$
\|v\|_{L^\infty(Q_T)} = \|Du\|^2_{L^\infty(Q_T)},
$$

and define T_M as

$$
T_M = \min\left\{\frac{1}{2C_0}M, \frac{1}{4\hat{C}(1+(2M)^{\gamma-1})}\right\}.
$$

Then it is easy to see that

$$
\|Du\|_{L^\infty(Q_T)} \le 2M \quad \forall T \le T_M. \tag{108}
$$

Indeed, for $T < T_M$ and $\|Du\|_{L^\infty(Q_T)} \le 2M$, (107) implies

$$
\begin{aligned}
\|Du\|^2_{L^\infty(Q_T)} &\le \|Dg\|^2_{L^\infty(Q_T)} + 4T_M^2C_0^2 + \tfrac{1}{4}\|Du\|^2_{L^\infty(Q_T)} \\
&\quad + \hat{C}T_M\|Du\|^2_{L^\infty(Q_T)}[(2M)^{\gamma-1} + 1] \\
&< \|Dg\|^2_{L^\infty(Q_T)} + M^2 + \tfrac{1}{2}\|Du\|^2_{L^\infty(Q_T)},
\end{aligned}
$$

hence

$$\|Du\|_{L^\infty(Q_T)} < 2M$$

whenever $T < T_M$ and $\|Du\|_{L^\infty(Q_T)} \le 2M$. A continuity argument implies that

$$\sup\{T : \|Du\|_{L^\infty(Q_T)} \le 2M\} = T_M,$$

so (108) holds true. Using this information, we deduce from (106) that

$$\|Du\|^2_{L^\infty(Q_T)} \le \|Dg\|^2_{L^\infty(Q_T)} + C_M T\|Du\|_{L^\infty(Q_T)},$$

where $C_M = 2C_0(1 + (2M)^\gamma) + c_d\|Da\|^2_\infty C_0 2M$. Hence

$$\left(\|Du\|_{L^\infty(Q_T)} - \tfrac{1}{2}C_M T\right)^2 \le \|Dg\|^2_{L^\infty(Q_T)} + \tfrac{1}{4}C_M^2 T^2$$

which implies $\|Du\|_{L^\infty(Q_T)} \le C_M T + \|Dg\|_{L^\infty(Q_T)}$. ■

Proposition A.2 (Lipschitz estimates, linear case). *Assume that $T \le 1$ and*

$$|D_x h(t,x,p)| \le C_1 + C_2|p| \quad \forall (t,x,p) \in (0,T) \times \mathbb{R}^d \times \mathbb{R}^d,$$

for some constants $C_1, C_2 > 0$. Then there exists a constant C, depending on C_0, C_2 and $\|Da\|_\infty$ only, such that

$$\sup_{t\in[0,T]} \|Du(t)\|_\infty \le \|Dg\|_\infty(1 + CT) + CC_1T.$$

Proof. Our starting point is inequality (105) in the previous proof. Using our assumptions on a and h we get

$$\begin{aligned} -\partial_t v - \mathrm{Tr}(a(t,x)D^2 v(t,x)) + h_p(t,x,Du(t,x)) \cdot Dv(t,x) \\ \le -2C_0^{-1}|D^2u|^2 + 2|Du|(C_1 + C_2|Du|) + \|Da\|_\infty |Du|\,|D^2u| \\ \le 2|Du|(C_1 + C_2|Du|) + c_d\|Da\|^2_\infty C_0|Du|^2, \end{aligned}$$

which implies

$$-\partial_t v - \mathrm{Tr}(a(t,x)D^2 v(t,x)) + h_p(t,x,Du(t,x)) \cdot Dv(t,x) \le \lambda v + 2C_1 v^{1/2},$$

where $\lambda = 2C_2 + c_d\|Da\|^2_\infty C_0$. By the maximum principle we get

$$\|v\|_{L^\infty(Q_T)} \le e^{\lambda T}(2C_1 T\|v\|^{1/2}_{L^\infty(Q_T)} + \|Dg\|^2_\infty),$$

from which we derive

$$\|v\|^{1/2}_{L^\infty(Q_T)} \le 2C_1 T e^{\lambda T} + e^{\lambda T/2}\|Dg\|_\infty.$$

Since $T \le 1$ (and so $e^{\lambda T/2} \le 1 + c_\lambda T$), the conclusion follows. ■

Proposition A.3 (Second order estimate). *Assume that h and a are of class C^2_b. Then, for any $M > 0$, there are constants $T_M, C_M > 0$, depending on M and on*

$$\sup_{t\in[0,T]} \|a(t)\|_2 + \sup_{|p|\le\|Du\|_\infty} \|D^2_{xp}h(\cdot,\cdot,p)\|_\infty, \tag{109}$$

such that if $\|D^2 g\|_\infty \le M$ and $T \in (0, T_M)$, then

$$\sup_{t\in[0,T]} \|D^2u(t)\|_\infty \le \|D^2 g\|_\infty + C_M T.$$

If in addition h is affine in p, then there is a constant C, depending only on C_0, $\sup_{t\in[0,T]} \|a(t)\|_2$ and on $\|D^2_{xp}h\|_\infty$, such that, for any $T \in (0, 1]$,

$$\sup_{t\in[0,T]} \|D^2u(t)\|_\infty \le (1 + CT)\|D^2 g\|_\infty + CT \sup_{|p|\le\|Du\|_\infty} \|D^2_{xx}h(\cdot,\cdot,p)\|_\infty.$$

Proof. We use the Bernstein method again. Let $w(t,x) = \sum_{i,j} u^2_{ij}$. Then

$$\begin{aligned}
&-\partial_t w - \mathrm{Tr}(a(t,x)D^2w(t,x))\\
&\quad = -2\sum_{i,j,k,l} a_{kl}(t,x)u_{ijk}(t,x)u_{ijl}(t,x) - 2\sum_{i,j} u_{ij}(t,x)D_{i,j}\Big(\partial_t u + \sum_{k,l} a_{kl}u_{kl}\Big)\\
&\qquad + 2\sum_{i,j,k,l} u_{ij}(t,x)((a_{kl})_i(t,x)u_{jkl}(t,x) + (a_{kl})_j(t,x)u_{ikl}(t,x) + (a_{kl})_{ij}u_{kl}).
\end{aligned}$$

So

$$\begin{aligned}
&-\partial_t w - \mathrm{Tr}(a(t,x)D^2w(t,x)) =\\
&-2\sum_{i,j,k,l} a_{kl}u_{ijk}u_{ijl} - 2\sum_{i,j} u_{ij}(h_{ij} + h_{i,p}\cdot Du_j + h_{j,p}\cdot Du_i + h_{pp}Du_i\cdot Du_j + h_pDu_{ij})\\
&+2\sum_{i,j,k,l} u_{ij}(t,x)\big((a_{kl})_i(t,x)u_{jkl}(t,x) + (a_{kl})_j(t,x)u_{ikl}(t,x) + (a_{kl})_{ij}u_{kl}\big),
\end{aligned} \tag{110}$$

which yields, using the ellipticity of $a(t,x)$,

$$\begin{aligned}
&-\partial_t w - \mathrm{Tr}(a(t,x)D^2w(t,x)) + h_p(t,x,Du(t,x))\cdot Dw(t,x)\\
&\le -2C_0^{-1}|D^3u|^2 + C_h|D^2u|(1+|D^2u|+|D^2u|^2) + C|D^2u|(\|a\|_1|D^3u| + \|a\|_2|D^2u|)
\end{aligned}$$

for some constant C_h depending on $\sup_{|p|\le\|Du\|_\infty} \|D^2_{x,p}h(\cdot,\cdot,p)\|_\infty$. Young's inequality leads to

$$\begin{aligned}
-\partial_t w - \mathrm{Tr}(a(t,x)D^2w(t,x)) + h_p(t,x,Du(t,x))\cdot Dw(t,x)&\\
\le C|D^2u|(1 + |D^2u| + |D^2u|^2)&,
\end{aligned}$$

where now C depends on $\|a\|_2$ as well. We conclude the proof using the maximum principle as in the proof of Proposition A.1.

If h is affine in p, then with the same estimates we deduce from (110) that

$$
\begin{aligned}
&-\partial_t w - \mathrm{Tr}(a(t,x)D^2 w(t,x)) + h_p(t,x,Du(t,x)) \cdot Dw(t,x) \\
&\qquad\qquad \leq |D^2 u|(2\|D_{xx}h\|_\infty + C|D^2 u|) \leq Cw + 2\|D_{xx}h\|_\infty |D^2 u|,
\end{aligned}
$$

where C depends on $\|a\|_2$, C_0 and $\sup_{|p|\leq \|Du\|_\infty} \|D^2_{x,p} h(\cdot,\cdot,p)\|_\infty$. The conclusion follows as in Lemma A.2. ∎

Proposition A.4 (Third order estimate). *Assume that h and a (and the solution u) are of class C^3_b. Then there is a constant C, depending on $\|D^2 u\|_\infty$, on $\|Da\|_\infty + \|D^2 a\|_\infty + \|D^3 a\|_\infty$ and on*

$$
\sup_{|p|\leq \|Du\|_\infty} \{\|D^3_{(x,p)} h(\cdot,\cdot,p)\|_\infty + \|h_{pp}(\cdot,\cdot,p)\|_\infty\},
$$

such that, for any $T \in (0,1]$,

$$
\sup_{t\in[0,T]} \|D^3 u(t)\|_\infty \leq (1+CT)\|D^3 g\|_\infty + CT.
$$

Proof. Let $w = \sum_{ijk} u_{ijk}^2$. Then

$$
\begin{aligned}
&-\partial_t w - \mathrm{Tr}(a(t,x)D^2 w(t,x)) = \\
&-2 \sum_{i,j,k,l,m} a_{lm}(t,x)u_{ijkl}(t,x)u_{ijkm}(t,x) - 2\sum_{i,j} u_{ijk}(t,x)D_{i,j,k}\Big(\partial_t u + \sum_{l,m} a_{lm}u_{lm}\Big) \\
&+2 \sum_{i,j,k,l,m} u_{ijk}(t,x)\big((a_{lm})_{ijk}u_{lm} + (a_{lm})_{ij}u_{klm} + (a_{lm})_{ik}u_{jlm} + (a_{lm})_{jk}u_{ilm} \\
&\qquad\qquad\qquad\qquad + (a_{lm})_i u_{jklm} + (a_{lm})_j u_{iklm} + (a_{lm})_k u_{ijlm}\big)
\end{aligned}
$$

So

$$
\begin{aligned}
&-\partial_t w - \mathrm{Tr}(a(t,x)D^2 w(t,x)) \\
&\quad = -2 \sum_{i,j,k,l,m} a_{lm}(t,x)u_{ijkl}(t,x)u_{ijkm}(t,x) - 2\sum_{i,j} u_{ijk}(t,x)D_{i,j,k}\{h\} \\
&\qquad +2 \sum_{i,j,k,l,m} u_{ijk}(t,x)\big((a_{lm})_{ijk}u_{lm} + (a_{lm})_{ij}u_{klm} + (a_{lm})_{ik}u_{jlm} + (a_{lm})_{jk}u_{ilm} \\
&\qquad\qquad\qquad\qquad + (a_{lm})_i u_{jklm} + (a_{lm})_j u_{iklm} + (a_{lm})_k u_{ijlm}\big). \qquad (111)
\end{aligned}
$$

As before, the coercivity of a implies

$$
-2 \sum_{i,j,k,l,m} a_{lm}(t,x)u_{ijkl}(t,x)u_{ijkm}(t,x) \leq -2C_0^{-1}|D^4 u|^2,
$$

whereas the last term in (111) is estimated as

$$2\sum_{i,j,k,l,m} u_{ijk}(t,x)\big((a_{lm})_{ijk}u_{lm} + (a_{lm})_{ij}u_{klm} + (a_{lm})_{ik}u_{jlm} + (a_{lm})_{jk}u_{ilm} + (a_{lm})_i u_{jklm} + (a_{lm})_j u_{iklm} + (a_{lm})_k u_{ijlm}\big) \le C_0^{-1}|D^4u|^2 + |D^3u|(2\|D^3a\|_\infty |D^2u| + C|D^3u|),$$

for some C depending on C_0 and $\|D^2a\|_\infty$. Finally, a direct computation of $D_{i,j,k}\{h\}$ and a straightforward estimate of all terms involved imply

$$-2\sum_{i,j} u_{ijk}(t,x)D_{i,j,k}\{h\} \le -h_p(t,x,Du(t,x))\cdot Dw(t,x) + C|D^3u|[\|D^2h\|_\infty|D^3u|(1+|D^2u|) + \|D^3h\|_\infty(1+|D^2u|^3)].$$

Hence, putting all together we deduce from (111) that

$$\begin{aligned}-\partial_t w - \mathrm{Tr}(a(t,x)D^2w(t,x)) + h_p(t,x,Du(t,x))\cdot Dw(t,x) &\le C|D^3u|[\|D^2h\|_\infty|D^3u|(1+|D^2u|) + \|D^3h\|_\infty(1+|D^2u|^3)]\\ &\quad + |D^3u|(2\|D^3a\|_\infty|D^2u| + C|D^3u|)\\ &\le C|D^3u|^2 + C|D^3u|,\end{aligned}$$

where C depends also on $\|D^k u\|_\infty$ for $k \le 2$. We conclude the proof as in Proposition A.2. ∎

Lemma A.5 (Higher order estimate). *Let $n \in \mathbb{N}$ with $n \ge 3$ and assume that h and a (and the solution u) are of class C_b^n. There is a constant C, depending on n, d, $\sup_t \|u(t)\|_{n-1}$, $\sup_t \|a(t)\|_n$ and*

$$\sup_{|p|\le\|Du\|_\infty} \sum_{k=0}^{n} \|D^k_{(x,p)}h(\cdot,\cdot,p)\|_\infty, \tag{112}$$

such that, for any $T \in (0,1]$,

$$\sup_t \|D^n u(t)\|_\infty \le (1+CT)\|D^n g\|_\infty + CT.$$

Proof. Let $w := \sum_{|k|=n} u_k^2$, where $k = (k_1,\dots,k_d) \in \mathbb{N}^d$ and $|k| = \sum_i k_i$. Then

$$\begin{aligned}&-\partial_t w - \mathrm{Tr}(a(t,x)D^2w(t,x))\\ &\quad= -2\sum_{|k|=n}\sum_{i,j} a_{ij}(t,x)u_{k,i}(t,x)u_{k,j}(t,x) - 2\sum_{|k|=n} u_k(t,x)D_k\{\partial_t u + \mathrm{Tr}(aD^2u)\}\\ &\qquad + 2\sum_{|k|=n} u_k\big(D_k(\mathrm{Tr}(aD^2u)) - \mathrm{Tr}(aD^2u_k)\big).\end{aligned}$$

As $n \ge 3$, a simple induction argument shows that $D_k\{h\}$ is of the form

$$D_k\{h\} = f_k + g_k\cdot D^n u + h_p\cdot Du_k$$

where the map

$$f_k = f_k(t, x, Du(t, x), \dots, D^{n-1}u(t, x))$$

is a polynomial function of the derivatives of u up to order $n - 1$ with coefficients involving derivatives of h with respect to (x, p) up to order n computed at $(t, x, Du(t, x))$, while

$$g_k \cdot D^n u = \sum_{|\xi|=n-1} \sum_{z+\xi=k} D_{z,p}h(t, x, Du(t, x))Du_\xi + h_{pp}(t, x, Du(t, x))Du_z Du_\xi,$$

where ξ is any multi-index of length $n - 1$, z is a multi-index of length 1 ($z = e_j$ for some $j \in \{1, \dots, d\}$) and $\xi + z = k$. Therefore

$$\begin{aligned}
&-\partial_t w - \mathrm{Tr}(a(t, x)D^2 w(t, x)) + h_p \cdot Dw \\
&\quad = -2\sum_{i,j} \sum_{|k|=n} a_{ij}(t, x)u_{k,i}(t, x)u_{k,j}(t, x) - 2\sum_{|k|=n} u_k(t, x)(f_k + g_k \cdot D^n u) \\
&\qquad + 2\sum_{|k|=n} u_k\big(D_k(\mathrm{Tr}(aD^2u)) - \mathrm{Tr}(aD^2u_k)\big) \\
&\quad \le -2C_0^{-1} \sum_{|k|=n} |Du_k|^2 + C|u_k|(1 + |u_k|) \\
&\qquad + 2\sum_{|k|=n} u_k\big(D_k(\mathrm{Tr}(aD^2u)) - \mathrm{Tr}(aD^2u_k)\big),
\end{aligned}$$

where C depends on $\sup_t \|u(t)\|_{n-1}$ and the quantity in (112). The last term can be estimated as before: the higher order quantity involves Du_k, so by Young's inequality we have

$$2\sum_{|k|=n} u_k\big(D_k(\mathrm{Tr}(aD^2u)) - \mathrm{Tr}(aD^2u_k)\big) \le 2C_0^{-1} \sum_{|k|=n} |Du_k|^2 + C|u_k|(1 + |u_k|)$$

for some C depending on $\sup_t \|a(t)\|_n$ and $\sup_t \|u(t)\|_{n-1}$. Finally, we use the maximum principle, as in Lemma A.2. ■

Proposition A.6 (Higher order estimate, further informations). *Let $n \in \mathbb{N}$ with $n \ge 3$ and assume that h and a (and the solution u) are of class C_b^n. For any $M > 0$, there are constants $K_M, T_M > 0$, depending on M, C_0 and γ, and a constant $C_M > 0$ depending on*

$$\sup_{t\in[0,T_M]} \|a(t)\|_n + \sup_{|p|\le K_M} \sum_{k=0}^{n} \|D^k_{(x,p)}h(\cdot, \cdot, p)\|_\infty,$$

such that if $\|g\|_n \le M$, then, for any $T \in (0, T_M)$ and any $r \le n$, we have

$$\sup_{t\in[0,T]} \|D^r_x u(t)\|_\infty \le \|D^r_x g\|_\infty + C_M T$$

and therefore

$$\sup_{t\in[0,T]} \|u(t)\|_n \le \|g\|_n + C_M T. \tag{113}$$

Proof. The proof is by a straightforward combination of Propositions A.1 and A.3 and Lemma A.5. ∎

We finally address the same issue for (uncoupled) systems of linear parabolic equations. Let $(u^l)_{l=1}^k$ solve the system

$$\begin{cases} -\partial_t u^l - \mathrm{Tr}(a(t,x)D^2u^l) + V(t,x)\cdot Du^l + f^l(t,x) = 0 \quad \text{in } (0,T)\times\mathbb{R}^d, \\ u^l(T,x) = g^l(x) \quad \text{in } \mathbb{R}^d, \end{cases}$$

where a, V and the f^l are bounded in C_b^n independently of $t \in [0,1]$, for some $n \in \mathbb{N}^*$. Note that the diffusion and the drift terms are independent of l.

Proposition A.7 (Higher order estimate, systems of affine equations). *There is a constant C, depending on k, d, $\sup_t \|a(t)\|_n$ and $\sup_t \|V(t)\|_n$, such that, for any $T \in (0,1]$ and any $r \le n$,*

$$\sup_{t,x} \Big(\sum_{l=1}^k |D_x^r u^l(t,x)|^2\Big)^{1/2} \\ \le (1+CT)\sup_x \Big(\sum_{l=1}^k |D_x^r g^l(x)|^2\Big)^{1/2} + CT\sup_l (\|g^l\|_r + \sup_t \|f^l(t)\|_r).$$

In particular, if $k=1$ then for any $r \le n$,

$$\sup_{t\in[0,T]} \|D_x^r u(t)\|_\infty \le (1+CT)\|D_x^r g\|_\infty + CT\sup_t \|D_x^r f(t)\|_\infty.$$

The only small point here is that the supremum over x is outside the sum (and not inside as it would be by simply applying the previous propositions to each u^l).

Proof of Proposition A.7. The proof runs exactly as before and so we just briefly explain the idea for $r=0$. Let $v(t,x) = \sum_{l=1}^k (u^l(t,x))^2$. Then v solves

$$-\partial_t v - \mathrm{Tr}(aD^2v) + V\cdot Dv = -2\sum_{l=1}^k u^l f^l - \sum_{i,j,l} a_{ij} u_i^l u_j^l.$$

We infer the result by using the positivity of a and the maximum principle. ∎

A.2. Systems with parameters

In this section we revisit the above estimates for specific systems of Hamilton–Jacobi equations involving a parameter y. The motivation for the specific form of the system is the analysis of MFG problems with a major player. Note that here the variables-parameter couple $(x;y)$ plays the role of $(x_0;x)$ in the HJ system (35) analyzed throughout Section 4.2. As usual, we discuss linear and nonlinear systems separately.

A.2.1. Nonlinear systems. Here we consider the system consisting in coupling a nonlinear HJ equation with a linear one:

$$\begin{cases} -\partial_t u^0(t,x) - \Delta u^0(t,x) + h^0(t,x,Du^0(t,x)) = 0 \quad \text{in } (0,T)\times\mathbb{R}^d, \\ -\partial_t u(t,x;y) - \Delta u(t,x;y) \\ \qquad + h^0_p(t,x,Du^0(t,x)) \cdot Du(t,x;y) + f(t,x;y) = 0 \quad \text{in } (0,T)\times\mathbb{R}^d, \\ u^0(T,x) = g^0(x),\ u(T,x) = g(x;y) \quad \text{in } \mathbb{R}^d, \end{cases} \tag{114}$$

where $h^0 : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$ and $f : [0,T]\times\mathbb{R}^d\times\mathbb{R}^{d_1} \to \mathbb{R}$ (d_1 being the space parameter of the variable y) are smooth maps satisfying in addition the bounds

$$|D_{x,p}h^0(t,x,p)| + |D^2_{x,p}h^0(t,x,p)| \le C_0(|p|^\gamma + 1) \tag{115}$$

for some $\gamma > 0$ and $C_0 > 0$.

Proposition A.8. *Let $r, n \in \mathbb{N}$ and assume (in addition to (115)) that h^0, h^0_p are of class C^r_b and that f is bounded in $C^{r,n}_b$ independently of $t \in [0,1]$ for some $n \in \mathbb{N}$. For any $M > 0$, there are constants $K_M, T_M > 0$, depending on M, C_0 and γ in (115), and a constant $C_M > 0$ depending on*

$$\sup_{|p|\le K_M} \Big(\sum_{k=0}^r \|D^k_{(x,p)}h^0(\cdot,\cdot,p)\|_\infty + \sum_{k=0}^r \|D^k_{(x,p)}h^0_p(\cdot,\cdot,p)\|_\infty\Big), \quad \sup_t \|f(t)\|_{r,n}$$

such that if $\|g^0\|_r + \|g\|_{r,n} \le M$ and $T \in (0,T_M)$, and if (u^0,u) is the solution to (114), then, for $l \le n$,

$$\sup_{t,x,y} \big(|D^r u^0(t,x)|^2 + |D^r_x D^l_y u(t,x;y)|^2\big)^{1/2} \le \sup_{x,y}\big(|D^r g^0(x)|^2 + |D^r_x D^l_y g(x;y)|^2\big)^{1/2} + C_M T.$$

Recall that $D^r_x D^l_y u = (\partial^\beta_x \partial^\alpha_y u)_{|\beta|=r,|\alpha|=l}$, hence $|D^r_x D^l_y u|^2 = \sum_{|\beta|=r,|\alpha|=l} (\partial^\beta_x \partial^\alpha_y u)^2$. Let us also point out that the main difference compared to Proposition A.6 is that we need to estimate u^0 and u at the same time.

Proof. The proof uses the same technique as for a single Hamilton–Jacobi equation without parameter. We only explain the main changes. We first prove the result for $l = 0$.

By the maximum principle we can first bound $|u^0|^2 + |u|^2$ by $\|(g^0)^2 + g^2\|_\infty + CT$. Next we address the Lipschitz estimate. We claim that, for any $M > 0$ and any $n \in \mathbb{N}$, if $\|Dg^0\|_\infty + \|D_x g\|_\infty \le M$, then there exist T_M and C_M (depending on M, C_0, n and γ in (115) only) such that

$$\sup_{t,x,y} \big(|Du^0(t,x)|^2 + |D_x u(t,x;y)|^2\big)^{1/2} \le \sup_{x,y}\big(|Dg^0(x)|^2 + |D_x g(x;y)|^2\big)^{1/2} + C_M T\Big(1 + \sup_t \|D_x f(t)\|_\infty\Big).$$

To see this, set $v(t,x) = \sum_{i=1}^d ((u_i^0)^2 + (u_i)^2)$. Then following the computation in the proof of Proposition A.1, we find

$$\begin{aligned}
&-\partial_t v - \Delta v(t,x) \\
&\quad = -2\sum_i (u_i^0 D_{x_i}(\partial_t u^0 + \Delta u^0) + u_i D_{x_i}(\partial_t u + \Delta u)) - 2(|D^2u^0|^2 + |D^2u|^2) \\
&\quad = -2\sum_i \big(u_i^0(h^0_{x_i} + h^0_p \cdot Du_i^0) + u_i(h^0_{p,x_i}\cdot Du + h^0_{pp}Du_i^0 \cdot Du + h^0_p \cdot Du_i + f_i)\big) \\
&\qquad - 2(|D^2u^0|^2 + |D^2u|^2),
\end{aligned}$$

so that

$$\begin{aligned}
&-\partial_t v - \Delta v(t,x) + h^0_p \cdot Dv \\
&\qquad = -2\sum_i \big(u_i^0 h^0_{x_i} + u_i(h^0_{p,x_i}\cdot Du + h^0_{pp} Du_i^0 \cdot Du + f_i)\big) - 2(|D^2u^0|^2 + |D^2u|^2).
\end{aligned}$$

Using our assumption on h^0 we get

$$-\partial_t v - \Delta v(t,x) + h^0_p \cdot Dv \le C v^{1/2}(|v|^\theta + 1 + \|D_x f\|_\infty)$$

for some $C > 0$ and $\theta > 0$ which depend on C_0 and γ only. We derive from this the Lipschitz estimate thanks to the maximum principle, exactly as in the proof of Proposition A.1.

The higher order estimates can be checked exactly as in Propositions A.3 and Lemma A.5, so we omit the proof. Note that the higher order estimates on $D^r u^0$ and $D^r u$ depend on $D^{r-1}u^0$ and $D^{r-1}u$, but this dependence affects the constant C_M only.

Let us finally explain how to handle the derivative with respect to y. We note that $\partial_y^\alpha u$ satisfies the same linear equation as u with f replaced by $\partial_y^\alpha f$, and the final datum g is replaced by $\partial_y^\alpha g$. So, in order to estimate $D_x D_y^l(u^0, u)$ for instance, we just set $c_l = (\sum_{|\alpha|=l} 1)^{-1}$, $v^\alpha = \sum_{i=1}^d (c_l (u_i^0)^2 + (\partial_y^\alpha u_i)^2)$ and $w = \sum_{|\alpha|=l} v^\alpha$. As above,

$$\begin{aligned}
-\partial_t v^\alpha - \Delta v^\alpha(t,x) + h^0_p \cdot Dv^\alpha &\le C(v^\alpha)^{1/2}(|v^\alpha|^\theta + 1 + \|D_x \partial_y^\alpha f\|_\infty) \\
&\le C w^{1/2}(|w|^\theta + 1 + \|D_x D_y^l f\|_\infty),
\end{aligned}$$

and summing up one gets the desired inequality, noting that $w = \sum_{|\beta|=1}(\partial_x^\beta u^0)^2 + \sum_{|\beta|=1,\,|\alpha|=l}(\partial_x^\beta \partial_y^\alpha u)^2$. ∎

A.2.2. Linear systems. We also need to quantify the regularity of linear systems of the form

$$\begin{cases}
-\partial_t u^0(t,x) - \Delta u^0(t,x) + V^0(t,x)\cdot Du^0(t,x) + f^0(t,x) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\
-\partial_t u(t,x;y) - \Delta u(t,x;y) + V^0(t,x)\cdot Du(t,x;y) + V(t,x;y)\cdot Du^0(t,x) & \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad + f(t,x;y) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\
u^0(T,x) = g^0(x), \quad u(T,x) = g(x;y) \quad \text{in } \mathbb{R}^d, &
\end{cases} \tag{116}$$

Proposition A.9. *Assume that, independently of $t \in (0,1]$, V^0, f^0 are bounded in C^r, and V, f are bounded $C_b^{r,n}$ for some $r, n \geq 0$. If (u^0, u) is a solution of (116) which is bounded in $C_b^r \times C_b^{r,n}$ and if $\|g^0\|_r + \|g\|_{r,n} \leq M$, then, for any $T \in (0,1]$, $l \leq n$,*

$$\sup_{t,x,y} \left(|D_x^r u^0(t,x)|^2 + |D_x^r D_y^l u(t,x;y)|^2\right)^{1/2} \leq \sup_{x,y} \left(|D_x^r g^0(x)|^2 + |D_x^r D_y^l g(x;y)|^2\right)^{1/2} + C_M T,$$

where C_M depends on M and the bounds on V^0, f^0 and V, f in C^r and $C_b^{r,n}$ respectively.

In addition, for $r = 0$ and $l \leq n$, we have

$$\sup_{t,x,y} \left(|u^0(t,x)|^2 + |D_y^l u(t,x;y)|^2\right)^{1/2} \leq (1 + CT) \sup_{x,y} \left(|g^0(x)|^2 + |D_y^l g(x;y)|^2\right)^{1/2} + CT(\|f^0\|_\infty + \|D_y^l f\|_\infty),$$

where C depends just on the bounds of V^0 and V.

Proof. We first note that the derivatives of u with respect to y solve a system which has the same structure as the one for u; so we just need to check the result for $n = 0$, and proceed as in the proof of Proposition A.8 for $n > 0$.

Let us start with the L^∞ bounds. We consider $\tilde{v} := (u^0)^2 + u^2$. Then v satisfies

$$\begin{aligned} -\partial_t \tilde{v} - \Delta \tilde{v} &= -2u^0(\partial_t u^0 + \Delta u^0) - 2u(\partial_t u + \Delta u) - 2(|Du^0|^2 + |Du|^2) \\ &= -2u^0(V^0(t,x) \cdot Du^0(t,x) + f^0(t,x)) \\ &\quad - 2u\big(V^0(t,x) \cdot Du(t,x) + V(t,x;y) \cdot Du^0(t,x) + f(t,x;y)\big) \\ &\quad - 2(|Du^0|^2 + |Du|^2) \\ &\leq C\tilde{v} + \tilde{v}^{1/2}(\|f^0\|_\infty + \|f\|_\infty), \end{aligned}$$

where C depends on $\|V^0\|_\infty$ and $\|V\|_\infty$ only. This implies the result for $r = n = 0$.

We now check the C^1 estimate. Set as usual $v(t,x) = \sum_{i=1}^d ((u_i^0)^2 + (u_i)^2)$. Then

$$\begin{aligned} &-\partial_t v - \Delta v(t,x) \\ &\quad = -2\sum_i \big(u_i^0 D_{x_i}(\partial_t u^0 + \Delta u^0) + u_i D_{x_i}(\partial_t u + \Delta u)\big) - 2(|D^2 u^0|^2 + |D^2 u|^2) \\ &\quad = -2\sum_i \big(u_i^0(V_{x_i}^0 \cdot Du^0 + V^0 \cdot Du_i^0 + f_i^0) + u_i(V_{x_i}^0 \cdot Du + V^0 \cdot Du_i + V_{x_i} \cdot Du^0 \\ &\qquad\qquad + V \cdot Du_i^0 + f_i)\big) - 2(|D^2 u^0|^2 + |D^2 u|^2) \\ &\quad \leq Cv + v^{1/2}(\|Df^0\|_\infty + \|D_x f\|_\infty), \end{aligned}$$

where C only depends on the C^1 bounds on V^0 and V and on d. This implies the estimate for $r = 1$ and $n = 0$.

As for the C^2 estimate, set as usual $w(t,x) = \sum_{i,j=1}^d ((u_{ij}^0)^2 + (u_{ij})^2)$. Then

$$-\partial_t w - \Delta w(t,x) \leq Cw + Cw^{1/2}(1 + \|D^2 f^0\|_\infty + \|D_x^2 f\|_\infty + \|Du^0\|_\infty + \|D_x u\|_\infty),$$

where C depends on the C^1 bound on V^0 and on V and on d only. We then get the estimate for $r = 2$ and $n = 0$ by the maximum principle and using the previous bounds for Du^0, Du.

The estimate on higher order derivatives can be checked in a similar way and we omit the proof. ■

Appendix B. Functions on $\mathscr{P}_2$

B.1. A criterion of differentiability

Here we introduce a simple criterion for a map U, depending on a measure, to be of class C^1.

Lemma B.1. *Let $U : \mathscr{P}_2 \to \mathbb{R}$ be continuous. For $(s, m, y) \in [0,1] \times \mathscr{P}_2 \times \mathbb{R}^d$ set*

$$\hat{U}(s; m, y) := U((1-s)m + s\delta_y).$$

If the map $s \mapsto \hat{U}(s; m, y)$ has a derivative at $s = 0$ and if $\frac{d}{ds}\big|_{s=0}\hat{U} : \mathscr{P}_2 \times \mathbb{R}^d \to \mathbb{R}$ is continuous and bounded, then U is of class C^1 with

$$\frac{\delta U}{\delta m}(m, y) = \frac{d}{ds}\hat{U}(0; m, y).$$

Proof. We have to show that, for any $m_0, m_1 \in \mathscr{P}_2$, we have

$$U(m_1) - U(m_0) = \int_0^1 \int_{\mathbb{R}^d} \frac{d}{ds}\hat{U}(0; (1-s)m_0 + sm_1, y)\,(m_1 - m_0)(dy).$$

Before starting the proof, let us note that the continuity of $\frac{d}{ds}\hat{U}$ at $s = 0$ implies its continuity at any $s \in [0,1]$, replacing m by $(1-s)m + s\delta_y$.

Let us start by considering the case where m_0 is fixed and m_1 is an empirical measure: $m_1 = m_{\mathbf{y}}^N := \frac{1}{N}\sum_{k=1}^N \delta_{y_k}$ for some $N \in \mathbb{N}$, $N \geq 1$, $y_k \in \mathbb{R}^d$. The general case will be treated next by approximation.

All the measures we are going to manipulate belong to the set

$$K := \Big\{\alpha_0 m_0 + \sum_{k=1}^N \alpha_k \delta_{y_k} : \alpha_k \geq 0, \sum_{k=0}^N \alpha_k = 1\Big\}$$

which is compact in $\mathscr{P}_2$. So, by continuity of $\frac{d}{ds}\hat{U}$, if we fix $\epsilon > 0$, there exists $\delta \in (0, 1/2)$ such that if $m', m'' \in K$ with $\mathbf{d}_2(m, m') < \delta$ and $s \in [0, \delta]$, then

$$\sup_k \left| \frac{d}{ds}\hat{U}(s; m, y_k) - \frac{d}{ds}\hat{U}(0; m', y_k)\right| \leq \epsilon. \tag{117}$$

Our first step consists in showing that, for $s > 0$ small enough (to be defined below) and for any $m \in K$, we have

$$\left| U((1-s)m + sm_{\mathbf{y}}^N) - U(m) - s\int_{\mathbb{R}^d} \frac{d}{ds}\hat{U}(0; m, y)\, m_{\mathbf{y}}^N(dy)\right| \leq C(\epsilon s + s^2), \tag{118}$$

where C depends on the sup norm of $\frac{d}{ds}\hat{U}$ on $[0,1]\times K\times\{y_k : k=1,\dots,N\}$. In order to prove (118), we define $\alpha_k = \frac{s}{N-(N-k)s}$ for $k=0,\dots,N$ and note that

$$\prod_{l=k}^{N}(1-\alpha_l) = 1-\frac{(N+1-k)s}{N}. \tag{119}$$

We now define by induction

$$m_0 = m, \quad m_k = (1-\alpha_k)m_{k-1} + \alpha_k\delta_{y_k}, \tag{120}$$

and using (119) we get

$$m_N = \prod_{k=1}^{N}(1-\alpha_k)m + \alpha_n\delta_{y_N} + \sum_{k=1}^{N-1}\alpha_k\delta_{y_k}\prod_{l=k+1}^{N}(1-\alpha_l)$$
$$= (1-s)m + \sum_{k=1}^{N}\delta_{y_k}\frac{s}{N-(N-k)s}\left(1-\frac{(N-k)s}{N}\right) = (1-s)m + sm_{\mathbf{y}}^N.$$

So, by the definition of m_{k+1} in terms of m_k in (120),

$$U((1-s)m+sm_{\mathbf{y}}^N) - U(m) = \sum_{k=0}^{N-1}U(m_{k+1}) - U(m_k)$$
$$= \sum_{k=0}^{N-1}(\hat{U}(\alpha_{k+1}; m_k, y_{k+1}) - \hat{U}(0; m_k, y_{k+1})) = \sum_{k=0}^{N-1}\int_0^{\alpha_{k+1}}\frac{d}{ds}\hat{U}(\tau; m_k, y_{k+1})\,d\tau.$$

Assume that $s\in(0,\delta)$. As $s<1/2$, we have $\alpha_k\le 2s/N$ for any k, and thus

$$\mathbf{d}_2(m_k, m) \le Cs$$

for a constant C which depends on m_0 and on the y_k (but not on $m\in K$ nor on $s\in(0,\delta)$). We now require that s is so small that $Cs<\delta$. Then, for any k and any $\tau\in(0,\alpha_k)$, by (117) we have

$$\left|\frac{d}{ds}\hat{U}(\tau; m_k, y_{k+1}) - \frac{d}{ds}\hat{U}(0; m, y_{k+1})\right| \le \epsilon.$$

We infer from this that

$$\left|U((1-s)m+sm_{\mathbf{y}}^N) - U(m) - \sum_{k=0}^{N-1}\alpha_{k+1}\frac{d}{ds}\hat{U}(0; m, y_{k+1})\right| \le C\epsilon\sum_{k=0}^{N-1}\alpha_{k+1}.$$

As $|\alpha_k - s/N| \le Cs^2/N$, we conclude that (118) holds.

The next step consists in showing that

$$U(e^{-1}m_0 + (1-e^{-1})m_{\mathbf{y}}^N) - U(m_0)$$
$$= \int_0^{1-e^{-1}}\int_{\mathbb{R}^d}\frac{d}{ds}\hat{U}(0; (1-\tau)m_0 + \tau m_{\mathbf{y}}^N, y)\,m_{\mathbf{y}}^N(dy)\,\frac{d\tau}{1-\tau}. \tag{121}$$

For this, let us now choose $T \in \mathbb{N}$ large and let

$$m_n = \left(1 - \frac{1}{T}\right)^n m_0 + \left(1 - \left(1 - \frac{1}{T}\right)^n\right) m_{\mathbf{y}}^N, \quad n \in \{0, \dots, T\}.$$

We have

$$m_{n+1} = \left(1 - \frac{1}{T}\right) m_n + \frac{1}{T} m_{\mathbf{y}}^N, \quad n \in \{0, \dots, T\}.$$

So, by (118),

$$\begin{aligned}
&\left| U(m_T) - U(m_0) - T^{-1} \sum_{n=0}^{T-1} \int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; m_n, y)\, m_{\mathbf{y}}^N(dy) \right| \\
&\le \sum_{n=0}^{T-1} \left| U((1-1/T)m_n + (1/T)m_{\mathbf{y}}^N) - U(m_n) - T^{-1} \int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; m_n, y)\, m_{\mathbf{y}}^N(dy) \right| \\
&\le C \sum_{n=0}^{T-1} (\epsilon/T + (1/T)^2) \le C(\epsilon + T^{-1}).
\end{aligned}$$

We let $T \to +\infty$ and then $\epsilon \to 0$ to conclude by continuity of U and of $\frac{d}{ds}\hat{U}$ that

$$\begin{aligned}
&U(e^{-1} m_0 + (1 - e^{-1}) m_{\mathbf{y}}^N) - U(m_0) \\
&\qquad = \int_0^1 \int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; e^{-s} m_0 + (1 - e^{-s}) m_{\mathbf{y}}^N, y)\, m_{\mathbf{y}}^N(dy)\, ds \\
&\qquad = \int_0^{1-e^{-1}} \int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; (1-\tau) m_0 + \tau m_{\mathbf{y}}^N, y)\, m_{\mathbf{y}}^N(dy)\, \frac{d\tau}{1-\tau}.
\end{aligned}$$

This is (121).

By continuity of U and of $\frac{d}{ds}\hat{U}$ and by density of the empirical measures, one deduces from (121) that, for any measures $m_0, m_1 \in \mathcal{P}_2$,

$$\begin{aligned}
&U(e^{-1} m_0 + (1 - e^{-1}) m_1) - U(m_0) \\
&\qquad = \int_0^{1-e^{-1}} \int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; (1-\tau) m_0 + \tau m_1, y)\, m_1(dy)\, \frac{d\tau}{1-\tau}. \qquad (122)
\end{aligned}$$

Choosing $m_1 = m_0$ then implies the normalization convention

$$\int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; m_0, y)\, m_0(dy) = 0$$

for any $m_0 \in \mathcal{P}_2$. In particular, this yields

$$\begin{aligned}
&\int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; (1-\tau) m_0 + \tau m_1, y)\, m_1(dy) \\
&\qquad = (1-\tau) \int_{\mathbb{R}^d} \frac{d}{ds} \hat{U}(0; (1-\tau) m_0 + \tau m_1, y)\, (m_1 - m_0)(dy).
\end{aligned}$$

Inserting this relation in (122) gives the more standard form

$$U(e^{-1}m_0 + (1-e^{-1})m_1) - U(m_0) \\ = \int_0^{1-e^{-1}} \int_{\mathbb{R}^d} \frac{d}{ds}\hat{U}(0;(1-\tau)m_0 + \tau m_1, y)\,(m_1 - m_0)(dy)\,d\tau.$$

Using again the continuity of U and of $\frac{d}{ds}\hat{U}$, one easily deduce from this the desired equality. ∎

B.2. Interpolation and Ascoli theorem in $\mathcal{P}_2$

In the proof of Lemma 3.5, we have used two interpolation lemmas. The first one is standard (see, for instance, [22, Lemma II.3.1]); we recall it because we need a specific setting. The second one is an adaptation to $\mathcal{P}_2$ of the same techniques.

Lemma B.2. *Let $W : [0,1] \times \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ be Hölder continuous in time locally uniformly in space: for any $R > 0$, there exist $C_{0,R}, \alpha_R > 0$ such that*

$$|W(t,y) - W(s,y)| \le C_{0,R}|t-s|^{\alpha} \\ \forall (s,t,y) \in [0,1]\times[0,1]\times\mathbb{R}^{d_1} \text{ with } |y| \le R \text{ and } |t-s| \le \alpha_R,$$

and such that D_yW is Hölder continuous in space uniformly in time: there exists $C_1 > 0$ such that

$$|D_yW(t,y_0) - D_yW(t,y_1)| \le C_1|y_0 - y_1|^{\delta} \quad \forall (t,y_1,y_2) \in [0,1]\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_1}.$$

Then D_yW is Hölder continuous in time locally uniformly in space:

$$|D_yW(t,y) - D_yW(s,y)| \le C_R|t-s|^{\frac{\alpha\delta}{1+\delta}} \\ \forall (s,t,y) \in [0,1]\times[0,1]\times\mathbb{R}^{d_1} \text{ with } |y| \le R \text{ and } |t-s| \le \alpha'_R,$$

for some constants $C_R > 0$ and α'_R only depending on $C_{0,R+1}$, α_{R+1}, C_1, α and δ.

Remark B.3. The proof below also shows that if in addition W is Hölder continuous in time uniformly in space (i.e., $C_{0,R}$ and α_R do not depend on R) and if D_yW is bounded, then D_yW is also Hölder continuous in time uniformly in space.

Proof of Lemma B.2. Fix $y_0, y_1 \in \mathbb{R}^d$ with $|y_0| \le R$ and $|y_1| \le R+1$. Let $y_\tau = (1-\tau)y_0 + \tau y_1$ for $\tau \in [0,1]$. We have

$$\left|\int_0^1 (D_yW(t,y_\tau) - D_yW(s,y_\tau))\cdot(y_1-y_0)\,d\tau\right| \\ = |W(t,y_1) - W(t,y_0) - W(s,y_1) + W(s,y_0)| \le 2C_{0,R+1}\,|t-s|^{\alpha}.$$

So

$$\begin{aligned}
&|(D_yW(t,y_0)-D_yW(s,y_0))\cdot(y_1-y_0)|\\
&\qquad\leq\left|\int_0^1(D_yW(t,y_0)-D_yW(t,y_\tau))\cdot(y_1-y_0)\,d\tau\right|\\
&\qquad\quad+\left|\int_0^1(D_yW(t,y_\tau)-D_yW(s,y_\tau))\cdot(y_1-y_0)\,d\tau\right|\\
&\qquad\quad+\left|\int_0^1(D_yW(s,y_\tau)-D_yW(s,y_0))\cdot(y_1-y_0)\,d\tau\right|\\
&\qquad\leq 2C_{0,R+1}|t-s|^\alpha+2C_1|y_1-y_0|^{1+\delta},
\end{aligned}$$

using also the Hölder continuity of D_yW. Choosing $y_1=y_0+hv$ with $|v|=1$, we get

$$\big|[D_yW(t,y)-D_yW(s,y)]\cdot v\big|\leq\frac{2C_{0,R+1}}{|h|}|t-s|^\alpha+2C_1|h|^\delta.$$

Optimizing with respect to $h\in(0,\alpha_{R+1}]$ and $|v|=1$, we find the result for $|t-s|\leq\alpha'_R$ for a suitable constant α'_R depending on $C_{0,R+1}$, α, C_1 and δ. ∎

Lemma B.4. *Let $W:[0,1]\times\mathcal{P}_2\to\mathbb{R}^{d_2}$ be Hölder continuous, locally in time and uniformly in measure: there exists $\alpha\in(0,1]$ and, for any $R>0$, there exists $C_{0,R}>0$ such that*

$$|W(t,m)-W(s,m)|\leq C_{0,R}|t-s|^\alpha\quad\forall m\in\mathcal{P}_2\text{ with }M_2(m)\leq R,\ \forall s,t\in[0,1]$$

(where $M_2(m)=(\int_{\mathbb{R}^d}|y|^2\,m(dy))^{1/2}$) and such that $\frac{\delta W}{\delta m}$ and D_mW are bounded and D_mW is Hölder continuous with respect to the measure uniformly in time: there exist $\gamma,\delta\in(0,1]$ and $C_1>0$ such that

$$|D_mW(t,m_0,y_0)-D_mW(t,m_1,y_1)|\leq C_1(\mathbf{d}_2^\gamma(m_0,m_1)+|y_0-y_1|^\delta)$$

for any $t\in[0,1]$ and any $(m_i,y_i)\in\mathcal{P}_2\times\mathbb{R}^d$. Then D_mW is Hölder continuous in time locally uniformly in $(m,y)\in\mathcal{P}_2\times\mathbb{R}^d$: for any $R>0$, there exists a constant $C_R>0$, depending on R, $\|D_mW\|_\infty$, $C_{0,R+1}$, C_1, α, γ and δ, such that

$$|D_mW(t,m,y)-D_mW(s,m,y)|\leq C_R|t-s|^{\alpha\gamma/((2+\gamma)(1+\delta))}$$

for any $s,t\in[0,1]$ and any $(m,y)\in\mathcal{P}_2\times\mathbb{R}^d$ with $|y|\leq R$ and $M_2(m)\leq R$.

Proof. Let $R\geq 1$. Fix $m_0,m_1\in\mathcal{P}_2$ with $M_2(m_i)\leq R$ and set $m_\tau=(1-\tau)m_0+\tau m_1$. Then

$$\begin{aligned}
&\left|\int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta W}{\delta m}(t,m_\tau,y)-\frac{\delta W}{\delta m}(s,m_\tau,y)\right)(m_1-m_0)(dy)\,d\tau\right|\\
&\qquad=|W(t,m_1)-W(t,m_0)-W(s,m_1)+W(s,m_0)|\leq 2C_{0,R}|t-s|^\alpha.
\end{aligned}$$

As

$$\left|\int_{\mathbb{R}^d}\left(\frac{\delta W}{\delta m}(t,m_0,y)-\frac{\delta W}{\delta m}(s,m_0,y)\right)(m_1-m_0)(dy)\right|$$
$$\leq\left|\int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta W}{\delta m}(t,m_\tau,y)-\frac{\delta W}{\delta m}(s,m_\tau,y)\right)(m_1-m_0)(dy)\,d\tau\right|$$
$$+\left|\int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta W}{\delta m}(t,m_\tau,y)-\frac{\delta W}{\delta m}(t,m_0,y)\right)(m_1-m_0)(dy)\,d\tau\right|$$
$$+\left|\int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta W}{\delta m}(s,m_\tau,y)-\frac{\delta W}{\delta m}(s,m_0,y)\right)(m_1-m_0)(dy)\,d\tau\right|,$$

we obtain, by our Hölder continuity assumption on D_mW,

$$\left|\int_{\mathbb{R}^d}\left(\frac{\delta W}{\delta m}(t,m_0,y)-\frac{\delta W}{\delta m}(s,m_0,y)\right)(m_1-m_0)(dy)\right|$$
$$\leq 2C_0|t-s|^\alpha+\sup_{\tau\in[0,1]}\|D_mW(t,m_\tau,\cdot)-D_mW(t,m_0,\cdot)\|_\infty\,\mathbf{d}_1(m_0,m_1)$$
$$+\sup_{\tau\in[0,1]}\|D_mW(s,m_\tau,\cdot)-D_mW(s,m_0,\cdot)\|_\infty\,\mathbf{d}_1(m_0,m_1)$$
$$\leq 2C_0|t-s|^\alpha+2C_1\mathbf{d}_2^\gamma(m_0,m_1)\mathbf{d}_1(m_0,m_1).$$

For any $y_0\in\mathbb{R}^d$ with $|y_0|\leq R$, let $m_1=(1-\theta)m_0+\theta\delta_{y_0}$ for some $\theta\in(0,1]$ to be chosen below. Note that

$$\mathbf{d}_1(m_1,m_0)\leq\theta\int_{\mathbb{R}^d}|y_0-x|m_0(dx)\leq\theta(|y_0|+(M_2(m_0))^{1/2})\leq 2\theta R$$

(since $R\geq 1$), while

$$\mathbf{d}_2(m_1,m_0)\leq\left(\theta\int_{\mathbb{R}^d}|y_0-x|^2\,m_0(dx)\right)^{1/2}\leq(2\theta)^{1/2}(|y_0|^2+M_2^2(m_0))^{1/2}\leq 2\theta^{1/2}R.$$

We get, by the convention on the derivative and our previous estimates,

$$\left|\frac{\delta W}{\delta m}(t,m_0,y_0)-\frac{\delta W}{\delta m}(s,m_0,y_0)\right|$$
$$=\frac{1}{\theta}\left|\int_{\mathbb{R}^d}\left(\frac{\delta W}{\delta m}(t,m_0,y)-\frac{\delta W}{\delta m}(s,m_0,y)\right)(m_1-m_0)(dy)\right|$$
$$\leq\frac{1}{\theta}[2C_{0,R}|t-s|^\alpha+cC_1R^{1+\gamma}\theta^{1+\gamma/2}],$$

where c is universal. If $|t-s|$ is small enough such that $C_{0,R}|t-s|^\alpha/(cC_1R^{1+\gamma})\leq 1$, then we choose $\theta^{1+\gamma/2}:=C_{0,R}|t-s|^\alpha/(cC_1R^{1+\gamma})$ and obtain

$$\left|\frac{\delta W}{\delta m}(t,m_0,y_0)-\frac{\delta W}{\delta m}(s,m_0,y_0)\right|$$
$$\leq cC_{0,R}^{\gamma/(2+\gamma)}C_1^{1/(1+\gamma/2)}R^{2(1+\gamma)/(2+\gamma)}|t-s|^{\alpha\gamma/(2+\gamma)},\qquad(123)$$

where c is another universal constant.

To show the regularity in time of $D_m W$, we just need to apply Lemma B.2 to $\frac{\delta W}{\delta m}$ since, by (123), $\frac{\delta W}{\delta m}$ is locally Hölder in time locally uniformly in space (the constant depending also on the measure) and $\frac{D_y \delta W}{\delta m} = D_m W$ is globally bounded and Hölder in y uniformly in time by assumption. We can remove the smallness restriction on $|t-s|$ by using the fact that $D_m W$ is globally bounded. ■

In the proof of Theorem 2.3 we also used the following version of the Arzelà–Ascoli theorem.

Lemma B.5. *Let (X, d) be a locally compact space and $W^N : X \times \mathcal{P}_2 \to \mathbb{R}$ be a family of uniformly bounded and locally uniformly continuous maps: there exists $x_0 \in X$ such that, for any $R > 0$, there exists a continuous nondecreasing modulus $\omega_R : [0, +\infty) \to [0, +\infty)$ with $\omega_R(0) = 0$ such that*

$$|W^N(x,m) - W^N(x',m')| \leq \omega_R(d(x,x') + \mathbf{d}_2(m,m')) \tag{124}$$

for any $x, x' \in X$ and $m, m' \in \mathcal{P}_2$ with $d(x, x_0) \leq R$, $d(x', x_0) \leq R$, $M_2(m) \leq R$, $M_2(m') \leq R$.

Then there exists a continuous map $W : X \times \mathcal{P}_2 \to \mathbb{R}$ and a subsequence (denoted in the same way) such that (W^N) converges to W pointwise in m and locally uniformly in x: for any $R > 0$ and any $m \in \mathcal{P}_2$,

$$\lim_{N \to +\infty} \sup_{d(x,x_0) \leq R} |W^N(x,m) - W(x,m)| = 0. \tag{125}$$

The only (very small) issue in the result is that $\mathcal{P}_2$ is not locally compact, so that the standard Arzelà–Ascoli theorem cannot be applied.

Proof of Lemma B.5. Let D be countable dense subset of $X \times \mathcal{P}_2$. By a diagonal argument we can find a subsequence (denoted in the same way) such that, for any $(x, m) \in D$, $(W^N(x, m))$ converges to some $W(x, m)$. By our regularity assumption (124) and since $X \times \mathcal{P}_2$ is complete, W can be extended to the whole space $X \times \mathcal{P}_2$ as a continuous map which satisfies

$$|W(x,m) - W(x',m')| \leq \omega_R(d(x,x') + \mathbf{d}_2(m,m')) \tag{126}$$

for any $x, x' \in X$ and $m, m' \in \mathcal{P}_2$ with $d(x, x_0) \leq R$, $d(x', x_0) \leq R$, $M_2(m) \leq R$, $M_2(m') \leq R$.

We claim that, for any $(x, m) \in X \times \mathcal{P}_2$, $(W^N(x, m))$ converges to $W(x, m)$. Indeed, fix $\epsilon > 0$ and $R = 2(1 + d(x, x_0) + M_2(m))$. Then there is $(x', m') \in D$ such that $d(x', x_0) \leq R$, $M_2(m') \leq R$ and $\omega_R((d(x, x') + \mathbf{d}_2(m, m')) \leq \epsilon/3$. Let also N_0 be so large that $|W^N(x', m') - W(x', m')| \leq \epsilon/3$ for $N \geq N_0$. Then, for $N \geq N_0$, we have

$$\begin{aligned}
&|W^N(x,m) - W(x,m)| \\
&\leq |W^N(x,m) - W^N(x',m')| + |W^N(x',m') - W(x',m')| + |W(x',m') - W(x,m)| \leq \epsilon,
\end{aligned}$$

where we have used (124) and (126) in the last inequality.

It remains to show that (125) holds. Fix $\epsilon > 0$ and let $\eta > 0$ be such that $\omega(\eta) \le \epsilon/3$. As X is locally compact, we can find $x_1, \dots, x_n$ such that any point $x \in B_X(x_0, R)$ is at a distance at most η from one of the $(x_i)_{i=1}^n$. Let N_0 be so large that $|W^N(x_i, m) - W(x_i, m)| \le \epsilon/3$ for any $i = 1, \dots, n$. Then, for any $x \in B_X(x_0, R)$ and any $N \ge N_0$, we have (for i such that $d(x, x_i) \le \eta$, so that $\omega_R(d(x, x_i)) \le \epsilon/3$)

$$\begin{aligned} &|W^N(x,m) - W(x,m)| \\ &\le |W^N(x,m) - W^N(x_i,m)| + |W^N(x_i,m) - W(x_i,m)| + |W(x_i,m) - W(x,m)| \le \epsilon, \end{aligned}$$

where we have again used (124) and (126) in the last inequality. This shows (125). ∎

Acknowledgements. The authors wish to thank the anonymous referees for their valuable comments which helped to improve the paper.

Funding. The first author was partially supported by the ANR (Agence Nationale de la Recherche) project ANR-16-CE40-0015-01 and by the AFOSR grant FA9550-18-1-0494. The second author was partially supported by the Fondazione CaRiPaRo Project "Nonlinear Partial Differential Equations: Asymptotic Problems and Mean-Field Games" and the Programme "FIL-Quota Incentivante" of University of Parma, co-sponsored by Fondazione Cariparma. The third author was partially supported by Indam (Gnampa national project 2018) and by FSMP (Fondation Sciences Mathématiques de Paris).

References

[1] Achdou, Y., Lasry, J.-M., Lions, P.-L., Rostand, A., Scheinkman, J.: Models for the economy of oil. In preparation

[2] Ahuja, S.: Wellposedness of mean field games with common noise under a weak monotonicity condition. SIAM J. Control Optim. **54**, 30–48 (2016) Zbl 1327.93403 MR 3439756

[3] Albeverio, S., Kondratiev, Y. G., Röckner, M.: Analysis and geometry on configuration spaces. J. Funct. Anal. **154**, 444–500 (1998) Zbl 0914.58028 MR 1612725

[4] Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. 2nd ed., Lectures in Math. ETH Zürich, Birkhäuser, Basel (2008) Zbl 1145.35001 MR 2401600

[5] Bensoussan, A., Chau, M. H. M., Yam, S. C. P.: Mean field Stackelberg games: aggregation of delayed instructions. SIAM J. Control Optim. **53**, 2237–2266 (2015) Zbl 1320.91028 MR 3376121

[6] Bensoussan, A., Chau, M. H. M., Yam, S. C. P.: Mean field games with a dominating player. Appl. Math. Optim. **74**, 91–128 (2016) Zbl 1348.49031 MR 3518868

[7] Buckdahn, R., Li, J., Peng, S.: Nonlinear stochastic differential games involving a major player and a large number of collectively acting minor agents. SIAM J. Control Optim. **52**, 451–492 (2014) Zbl 1292.93144 MR 3163233

[8] Caines, P. E., Kizilkale, A. C.: Mean field estimation for partially observed LQG systems with major and minor agents. In: Proc. 19th World Congres Int. Federation of Automatic Control (Cape Town, 2014), IFAC Proc. Vol. 47, 8705–8709 (2014)

[9] Cardaliaguet, P., Cirant, M., Porretta, A.: Remarks on Nash equilibria in mean field game models with a major player. Proc. Amer. Math. Soc. **148**, 4241–4255 (2020) Zbl 1446.35213 MR 4135293

[10] Cardaliaguet, P., Delarue, F., Lasry, J.-M., Lions, P.-L.: The Master Equation and the Convergence Problem in Mean Field Games. Ann. of Math. Stud. 201, Princeton Univ. Press, Princeton, NJ (2019) Zbl 1430.91002 MR 3967062

[11] Carmona, R., Delarue, F.: The master equation for large population equilibriums. In: Stochastic Analysis and Applications 2014, Springer Proc. Math. Statist. 100, Springer, Cham, 77–128 (2014) Zbl 1391.92036 MR 3332710

[12] Carmona, R., Delarue, F.: Probabilistic Theory of Mean Field Games with Applications. I. Probab. Theory Stoch. Modelling 83, Springer, Cham (2018) Zbl 1422.91014 MR 3752669

[13] Carmona, R., Delarue, F., Lacker, D.: Mean field games with common noise. Ann. Probab. **44**, 3740–3803 (2016) Zbl 1422.91083 MR 3572323

[14] Carmona, R., Wang, P.: An alternative approach to mean field game with major and minor players, and applications to herders impacts. Appl. Math. Optim. **76**, 5–27 (2017) Zbl 1378.49036 MR 3679337

[15] Carmona, R., Wang, P.: Finite state mean field games with major and minor players. arXiv:1610.05408 (2016)

[16] Carmona, R., Zhu, X.: A probabilistic approach to mean field games with major and minor players. Ann. Appl. Probab. **26**, 1535–1580 (2016) Zbl 1342.93121 MR 3513598

[17] Gangbo, W., Święch, A.: Existence of a solution to an equation arising from the theory of mean field games. J. Differential Equations **259**, 6573–6643 (2015) Zbl 1359.35221 MR 3397332

[18] Gilbarg, D., Trudinger, N. S.: Elliptic Partial Differential Equations of Second Order. Springer (2015) Zbl 0361.35003 !

[19] Huang, M.: Large-population LQG games involving a major player: the Nash certainty equivalence principle. SIAM J. Control Optim. **48**, 3318–3353 (2009/10) Zbl 1200.91020 MR 2599921

[20] Huang, M., Malhamé, R. P., Caines, P. E.: Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle. Commun. Inf. Syst. **6**, 221–251 (2006) Zbl 1136.91349 MR 2346927 !

[21] Lacker, D., Webster, K.: Translation invariant mean field games with common noise. Electron. Commun. Probab. **20**, no. 42, 13 pp. (2015) Zbl 1332.60102 MR 3358964

[22] Ladyženskaja, O. A., Solonnikov, V. A., Ural'ceva, N. N.: Linear and Quasilinear Equations of Parabolic Type. Transl. Math. Monogr. 23, Amer. Math. Soc., Providence, RI (1968) MR 0241822

[23] Lasry, J.-M., Lions, P.-L.: Jeux à champ moyen. I. Le cas stationnaire. C. R. Math. Acad. Sci. Paris **343**, 619–625 (2006) Zbl 1153.91009 MR 2269875 !

[24] Lasry, J.-M., Lions, P.-L.: Jeux à champ moyen. II. Horizon fini et contrôle optimal. C. R. Math. Acad. Sci. Paris **343**, 679–684 (2006) Zbl 1153.91010 MR 2271747

[25] Lasry, J.-M., Lions, P.-L.: Mean field games. Jpn. J. Math. **2**, 229–260 (2007) Zbl 1156.91321 MR 2295621

[26] Lasry, J.-M., Lions, P.-L.: Mean-field games with a major player. C. R. Math. Acad. Sci. Paris **356**, 886–890 (2018) Zbl 1410.91048 MR 3851543

[27] Lions, P.-L.: Cours au Collège de France. www.college-de-france.fr.

[28] Lions, P.-L.: Estimées nouvelles pour les équations quasilinéaires. Seminar in Applied Mathematics at the Collège de France. http://www.college-de-france.fr/site/pierrelouis-lions/seminar-2014-11-14-11h15.htm (2014)

[29] Mayorga, S.: Short time solution to the master equation of a first order mean field game. J. Differential Equations **268**, 6251–6318 (2020) Zbl 1447.35331 MR 4069006

[30] Rachev, S. T., Rüschendorf, L.: Mass Transportation Problems. Vol. II. Probab. Appl. (New York), Springer, New York (1998) Zbl 0990.60500 MR 1619171

[31] Villani, C.: Optimal Transport: Old and New. Grundlehren Math. Wiss. 338, Springer, Berlin (2009) Zbl 1156.53003 MR 2459454