



**HAL**  
open science

# The convergence problem in mean field games with a local coupling

Pierre Cardaliaguet

► **To cite this version:**

Pierre Cardaliaguet. The convergence problem in mean field games with a local coupling. Applied Mathematics and Optimization, In press, 216 (1), pp.1 - 62. 10.1007/s00245-017-9434-0. hal-01384333v2

**HAL Id: hal-01384333**

**<https://hal.science/hal-01384333v2>**

Submitted on 10 Oct 2017

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

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

# The convergence problem in mean field games with local coupling

P. Cardaliaguet\*

October 10, 2017

## Abstract

The paper studies the convergence, as  $N$  tends to infinity, of a system of  $N$  coupled Hamilton-Jacobi equations, the Nash system, when the coupling between the players becomes increasingly singular. The limit equation turns out to be a Mean Field Game system with a local coupling.

## Contents

<b>1</b>	<b>Notation and Assumptions</b>	<b>5</b>
1.1	Notation . . . . .	5
1.2	Derivatives with respect to the measure . . . . .	6
1.3	Assumption . . . . .	7
1.4	Main example . . . . .	8
<b>2</b>	<b>Regularity estimates</b>	<b>10</b>
2.1	Estimates on the MFG systems for smooth initial conditions . . . . .	11
2.2	Estimates on the MFG systems for general initial conditions . . . . .	14
2.3	Estimates for a linearized system . . . . .	15
2.4	Estimates for $\frac{\delta U^N}{\delta m}$ . . . . .	18
2.5	Estimate for $\frac{\delta^2 U^N}{\delta m^2}$ . . . . .	19
<b>3</b>	<b>Convergence</b>	<b>20</b>
3.1	Finite dimensional projections of $U^N$ . . . . .	20
3.2	Estimates between $v^{N,i}$ and $U^N$ . . . . .	22
3.3	Putting the estimates together . . . . .	26
3.4	Convergence of the optimal solutions . . . . .	27
<b>4</b>	<b>Appendix</b>	<b>30</b>

---

\*Université Paris-Dauphine, PSL Research University, CNRS, Ceremade, 75016 Paris, France.  
cardaliaguet@ceremade.dauphine.fr



for any sufficiently smooth density  $m$  of a measure  $m(x)dx \in \mathcal{P}(\mathbb{R}^d)$ . This assumption, which is the main difference with [2], is very natural in the context of mean field games. One expects (and we will actually prove) that the limit system is a MFG system with local interactions:

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m(t, x)) & \text{in } [t_0, T] \times \mathbb{R}^d, \\ \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } [t_0, T] \times \mathbb{R}^d, \\ u(T, x) = G(x), m(t_0, \cdot) = m_0 & \text{in } \mathbb{R}^d. \end{cases} \quad (4)$$

This system—which enjoys very nice properties—has been very much studied in the literature: see for instance [10, 21, 22]. Note that in these papers the terminal condition  $G$  may also depend on the measure. For technical reasons we cannot allow this dependence in our analysis.

To explain in what extend the local framework differs from the nonlocal one, let us recall the ideas of proof in this later setting. The main ingredient in [2] for the proof of the convergence is the existence of a classical solution to the so-called *master equation*. When  $F^{N,i}(\mathbf{x}) = F(x_i, m_{\mathbf{x}}^{N,i})$ , where  $F : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$  is sufficiently smooth (at least continuous), the master equation takes the form of a transport equation stated on the space of probability measures:

$$\begin{cases} -\partial_t U - \Delta_x U + H(x, D_x U) \\ \quad - \int_{\mathbb{R}^d} \operatorname{div}_y [D_m U] dm(y) + \int_{\mathbb{R}^d} D_m U \cdot D_p H(y, D_x U) dm(y) = F(x, m) \\ \quad \text{in } [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d), \\ U(T, x, m) = G(x) \quad \text{in } \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d). \end{cases} \quad (5)$$

In the above equation,  $U$  is a scalar function depending on  $(t, x, m) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ ,  $D_m U$  denotes the derivative with respect to the measure  $m$  (see Section 1). The interest of the map  $U$  is that the  $N$ -tuple  $(u^{N,i})$ , where

$$u^{N,i}(t, \mathbf{x}) = U(t, x_i, m_{\mathbf{x}}^{N,i}), \quad \mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N, \quad (6)$$

is an approximate solution to the Nash system (1) which enjoys very good regularity properties. In [2] we use these regularity properties in a crucial way to prove the convergence of the Nash system.

When  $F$  is a local coupling, i.e.,  $F(x, m) = F(x, m(x))$  for any absolutely continuous measure  $m$ , the meaning of the master equation is not clear: obviously one cannot expect  $U$  to be a smooth solution to (5), if only because the coupling function  $F$  blows up at singular measures. So the master equation should be defined on a subset of sufficiently smooth density measures. But then the definition of the map  $u^{N,i}$  through (6) is dubious and, even if such a definition could make sense, there would be no hope that the  $u^{N,i}$  satisfy the regularity properties required in the computation of [2].

Our starting point is the following easy remark: if the coupling  $F^N$  remains sufficiently smooth for  $N$  large, then the solution  $v^{N,i}$  should eventually become close to the solution of the master equation associated with  $F^N$ —and thus to the solution of the MFG system with nonlocal coupling  $F^N$ . On the other hand, if  $F^N$  is close to  $F$  (thus becoming singular), the solution  $(u^N, m^N)$  is also close to the solution  $(u, m)$  of the MFG system with local coupling. Thus if  $F^N$  converges very slowly to  $F$  while remaining “sufficiently” smooth one can expect a convergence

of  $v^{N,i}$  to  $u$ . The whole point consists in quantifying in a careful way this convergence.

To give a flavor of our results, let us discuss a particular case. Let us assume for a while that

$$F^N(x, m) = F(\cdot, \xi^{\epsilon_N} * m(\cdot)) * \xi^{\epsilon_N},$$

where  $\xi^\epsilon(x) = \epsilon^{-d}\xi(x/\epsilon)$ ,  $\xi$  being a symmetric smooth nonnegative kernel with compact support and  $(\epsilon_N)$  is a sequence which converges to 0. Let  $v^{N,i}$  be the solution of the Nash system (1). Given an initial condition  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{R}^d)$ , let  $(u, m)$  be the solution to the MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m(t, x)) & \text{in } [t_0, T] \times \mathbb{R}^d, \\ \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } [t_0, T] \times \mathbb{R}^d, \\ u(T, x) = G(x), m(t_0, \cdot) = m_0 & \text{in } \mathbb{R}^d. \end{cases}$$

In order to describe the convergence of  $v^{N,i}$  to  $u$ , we reduce the function  $v^{N,i}$  to a function of a single variable by averaging it against the measure  $m_0$ : for  $i \in \{1, \dots, N\}$ , let

$$w^{N,i}(t_0, x_i, m_0) := \int_{\mathbb{T}^d} \dots \int_{\mathbb{T}^d} v^{N,i}(t_0, \mathbf{x}) \prod_{j \neq i} m_0(dx_j) \quad \text{where } \mathbf{x} = (x_1, \dots, x_N).$$

Corollary 3.5 states that, if  $\epsilon_N = \ln(N)^{-\beta}$  for some  $\beta \in (0, (6d(2d+15))^{-1})$ , then

$$\|w^{N,i}(t_0, \cdot, m_0) - u(t_0, \cdot)\|_\infty \leq A(\ln(N))^{-1/B},$$

for some constants  $A, B > 0$ . Moreover, we show that the optimal trajectories  $(Y_{i,t})$  converge to the optimal trajectory  $(\tilde{X}_{i,t})$  associated with the limit MFG system and establish a propagation of chaos property (Proposition 3.7).

For a general sequence of couplings  $(F^N)$ , converging to a local coupling  $F$  as in (3), our main result (Theorem 3.4) states that, if the regularity of  $F^N$  does not deteriorate too fast, then  $w^{N,i}$  converges to  $u$  and the optimal trajectories converge as well.

Let us finally point out that, in order to avoid issue related to boundary conditions or problems at infinity, we will assume that the data are periodic in space, thus working on the torus  $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ .

Mean field game theory started with the pioneering works by Lasry and Lions [18, 19, 20] and Caines, Huang and Malhamé [15, 12, 13, 11, 14]. These authors introduced the mean field game system and discussed its properties: in particular, Lasry and Lions introduced the fundamental monotonicity condition on the coupling functions. They also discussed the various types of MFG systems (with soft or hard coupling, with or without diffusion).

The link between the MFG system (which can be seen as a differential game with infinitely many players) and the differential games with finitely many players has been the object of several contributions. Caines, Huang and Malhamé [11], and Delarue and Carmona [4] explained how to use the solution of the MFG system to build  $\epsilon$ -Nash equilibria (in open loop form) for  $N$ -person games. The convergence of the Nash system remained a puzzling issue for some time. The first results in that direction go back to [18, 20] (see also [6]), in the “ergodic case”, where the Nash system becomes a system of  $N$  coupled equation in dimension  $d$  (and not  $Nd$  as in our setting): then one can obtain estimates which allow to pass to the limit. Another particular case is obtained when one is interested in Nash equilibria in open loop form: Fischer [7] and Lacker [16] explained in what extend one can expect to obtain the MFG system at the limit. For the genuine Nash system (1), a first breakthrough was achieved by Lasry and Lions (see

the presentation in [21]) who formally explained the mechanism towards convergence *assuming* suitable *a priori* estimates on the solution. For that purpose they also introduced the master equation (equation (5)) and described (mostly formally) its main properties.

The rigorous derivation of Lasry and Lions ideas took some time. The existence of a classical solution to the master equation has been obtained by several authors in different frameworks (Buckdahn, Li, Peng and Rainer [1] for the linear master equation without coupling, Gangbo and Swiech [9] for the master equation without diffusion and in short time horizon, Chassagneux, Crisan and Delarue [5] for the first order master equation, Lions [21] for an approach by monotone operators). The most general result so far is obtained in [2], where the master equation is proved to be well-posed even for problems with *common noise*. The main contribution of [2] is, however, the convergence of the Nash system: it is obtained as a consequence of the well-posedness of the master equation. The present paper is the first attempt to show the convergence for a coupling which becomes singular.

The paper is organized in the following way: we first state our main notation (in particular for the derivatives with respect to a measure) and main assumptions. In section 2, we prove our key estimates on the solution of the master equation and on the MFG systems. The whole point is to display the dependence of the estimates with respect to the regularity of the coupling. The last part collects our convergence results.

**Acknowledgement:** The author was partially supported by the ANR (Agence Nationale de la Recherche) project ANR-16-CE40-0015-01. The author wishes to thank the anonymous referee for the very careful reading and for finding a serious gap in the previous version of the paper.

## 1 Notation and Assumptions

For the sake of simplicity, the paper is written under the assumption that all maps are periodic in space. So the underlying state space is the torus  $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ . This simplifying assumption allows to discard possible problems at infinity (or at the boundary of a domain).

### 1.1 Notation

We will need the following notations for the derivatives in space or in time of a map.

For  $u = u(x)$  and  $l = (l_1, \dots, l_d) \in \mathbb{N}^d$ , we denote by  $D^l u(x)$  the derivative  $D^l u(x) = \frac{\partial^{|l|} u(x)}{\partial x_1^{l_1} \dots \partial x_d^{l_d}}$ , where  $|l| = l_1 + \dots + l_d$ . If  $k \in \mathbb{N}$ ,  $D^k u(x)$  denotes the collection of derivatives  $(D^l u(x))_{|l|=k}$ . For  $k = 1$  and  $k = 2$ ,  $Du(x)$  and  $D^2 u(x)$  denote the gradient and the Hessian of  $u$  at  $x$ .

For  $k \in \mathbb{N}$  and  $\alpha \in (0, 1)$ , we denote by  $C^{k+\alpha}$  the set of maps  $u = u(x)$  which are of class  $C^k$  and such that  $D^l u$  is  $\alpha$ -Hölder continuous for any  $l \in \mathbb{N}^d$  with  $|l| = k$ . We set

$$\|u\|_{k+\alpha} := \sum_{|l| \leq k} \|D^l u\|_\infty + \sum_{|l| \leq k} \sup_{x \neq y} \frac{|D^l u(x) - D^l u(y)|}{|x - y|^\alpha}.$$

When a map  $u$  depends on several space variables, say 2 for instance, we set in the same way

$$\|u\|_{k+\alpha, k'+\alpha} := \sum_{|l| \leq k, |l'| \leq k'} \|D_{x,x'}^{l,l'} u\|_\infty + \sum_{|l| \leq k, |l'| \leq k'} \sup_{(x,x') \neq (y,y')} \frac{|D_{x,x'}^{l,l'} u(x,x') - D_{x,x'}^{l,l'} u(y,y')|}{|(x,x') - (y,y')|^\alpha}.$$

For  $p \geq 1$ , the  $L^p$  norm of  $u$  is denoted by  $\|u\|_{L^p}$ . However, by abuse of notation, we denote by  $\|u\|_\infty$  the  $L^\infty$  norm of  $u$  (instead of  $\|u\|_{L^\infty}$ ). When  $\phi$  is a distribution, we set

$$\|\phi\|_{-(k+\alpha)} := \sup_{\|u\|_{k+\alpha} \leq 1} |\phi(u)|.$$

Finally, when  $u = u(t, x)$  is also time dependent, we denote by  $\partial_t u$  the time derivative of  $u$  and, as previously, by  $D^l u$  its space derivative of order  $l \in \mathbb{N}^d$ . If  $\alpha \in (0, 1)$ , we say that  $u$  is in  $C^{\alpha/2, \alpha}$  if

$$\|u\|_{C^{\alpha/2, \alpha}} := \|u\|_\infty + \sup_{(t,x), (t',x')} \frac{|u(t, x) - u(t', x')|}{|x - x'|^\alpha + |t - t'|^{\alpha/2}} < +\infty.$$

We say that  $u$  is in  $C^{1+\alpha/2, 2+\alpha}$  if  $\partial_t u$  and  $D^2 u$  belong to  $C^{\alpha/2, \alpha}$ .

## 1.2 Derivatives with respect to the measure

We follow here [2]. We denote by  $\mathcal{P}(\mathbb{T}^d)$  the set of Borel probability measures on the torus  $\mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d$ . It is endowed with the Monge-Kantorovitch distance:

$$\mathbf{d}_1(m, m') = \sup_{\phi} \int_{\mathbb{T}^d} \phi(y) d(m - m')(y),$$

where the supremum is taken over all 1-Lipschitz continuous maps  $\phi : \mathbb{T}^d \rightarrow \mathbb{R}$ .

**Definition 1.1.** Let  $U : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  be a map.

- We say that  $U$  is  $C^1$  if there exists a continuous map  $\frac{\delta U}{\delta m} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}$  such that, for any  $m, m' \in \mathcal{P}(\mathbb{T}^d)$ ,

$$U(m') - U(m) = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}((1-s)m + sm', y) d(m' - m)(y) ds.$$

The map  $\frac{\delta U}{\delta m}$  being defined up to an additive constant, we adopt the normalization convention

$$\int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(m, y) dm(y) = 0. \quad (7)$$

- If  $\frac{\delta U}{\delta m}$  is of class  $C^1$  with respect to the second variable, the intrinsic derivative  $D_m U : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}^d$  is defined by

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

It is known [2] that the map  $D_m U$  measures the Lipschitz regularity of  $U$ :

$$|U(m') - U(m)| \leq \sup_{m''} \|D_m U(m'', \cdot)\|_\infty \mathbf{d}_1(m, m') \quad \forall m, m' \in \mathcal{P}(\mathbb{T}^d).$$

We will also need second order derivatives with respect to the measure. If  $U$  and  $\frac{\delta U}{\delta m}$  are of class  $C^1$  with respect to the measure  $m$ , we denote by  $\frac{\delta^2 U}{\delta m^2} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \times \mathbb{T}^d \rightarrow \mathbb{R}$  the derivative of  $\frac{\delta U}{\delta m}$  with respect to  $m$ . If  $\frac{\delta^2 U}{\delta m^2}$  is sufficiently smooth, we also set

$$D_{mm}^2 U(m, y, y') = D_{y, y'}^2 \frac{\delta^2 U}{\delta m^2}(m, y, y').$$

### 1.3 Assumption

Throughout the paper, we suppose that the following conditions are in force.

- The Hamiltonian  $H : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is smooth, globally Lipschitz continuous in both variables and locally uniformly convex with respect to the second variable:

$$D_{pp}^2 H(x, p) > 0 \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d. \quad (8)$$

Moreover,  $D_p H$  and its derivatives are globally Lipschitz continuous.

- $F : \mathbb{T}^d \times [0, +\infty) \rightarrow \mathbb{R}$  is smooth, with bounded derivatives in both variables. Moreover,  $F$  is increasing with respect to the second variable, with  $\frac{\partial F}{\partial m} \geq \delta > 0$  for some  $\delta > 0$  (note that  $\frac{\partial F}{\partial m}$  stands for the usual derivative of the map  $F = F(x, m)$  with respect to the last variable).
- The terminal cost  $G : \mathbb{T}^d \rightarrow \mathbb{R}$  is a smooth map.
- For any  $N \in \mathbb{N}$ ,  $F^N : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$  is monotone:

$$\int_{\mathbb{T}^d} (F^N(x, m) - F^N(x, m')) d(m - m')(x) \geq 0 \quad \forall m, m' \in \mathcal{P}(\mathbb{T}^d).$$

- (difference between  $F^N$  and  $F$ ) For any  $R > 0$  and  $\alpha \in (0, 1)$ , there exists  $k_N^{R, \alpha}$  such that  $k_N^{R, \alpha} \rightarrow 0$  as  $N \rightarrow +\infty$  and

$$\|F^N(\cdot, m dx) - F(\cdot, m(\cdot))\|_\infty \leq k_N^{R, \alpha}, \quad (9)$$

for any density  $m$  such that  $\|m\|_{C^\alpha} \leq R$ .

- (uniform regularity of  $F^N$ ) For any  $R > 0$  and  $\alpha \in (0, 1)$ , there exists  $\kappa_{R, \alpha} > 0$  such that, for any  $N \in \mathbb{N}$ ,

$$|F^N(x, m dx) - F^N(y, m' dx)| \leq \kappa_{R, \alpha} (|x - y|^\alpha + \|m - m'\|_\infty) \quad (10)$$

for any density  $m, m'$  with  $\|m\|_{C^\alpha}, \|m'\|_{C^\alpha} \leq R$ .

- (regularity assumptions on  $F^N$ ) For any  $N \in \mathbb{N}$ ,  $F^N$  is of class  $C^2$  with respect to the  $m$  variable and, for any  $\alpha \in (0, 1)$ , there exists a constant  $K_{N, \alpha}$  such that

$$\|F^N(\cdot, m)\|_{4+\alpha} + \left\| \frac{\delta F^N}{\delta m}(\cdot, m, \cdot) \right\|_{(4+\alpha, 4+\alpha)} + \left\| \frac{\delta^2 F^N}{\delta m^2}(\cdot, m, \cdot, \cdot) \right\|_{(4+\alpha, 4+\alpha, 4+\alpha)} \leq K_{N, \alpha}, \quad (11)$$

for any  $m \in \mathcal{P}(\mathbb{T}^d)$ .

Some comments are in order. The Lipschitz regularity condition of the Hamiltonian  $H$  is not very natural in the context of MFG, but we do not know how to avoid it: if it is probably not necessary in the estimates of Section 2, it is required for the convergence of the Nash system. Let us just note that it simplifies at lot the existence of solutions for the Nash system (see, for instance, [17]) as well as for the limit MFG system (see [20]): indeed, without the assumption that  $D_p H$  is bounded, existence of classical solution to (4) is related on a subtle interplay between



the growth of  $H$  and of  $F$ . Note however that [3] overcomes this issue for nonlocal couplings functions.

The monotonicity of  $F^N$  and  $F$  are natural to ensure the uniqueness of the solution to the respective MF systems. However, the strict monotonicity of  $F$  is unusual: it is used here to give a good control between the MFG limit system (4) and the (nonlocal) MFG system with coupling term  $F^N$  (given in (15) below): see Proposition 2.3.

In MFG problems, one often assumes that the terminal cost also depends on the measure. It seems difficult to allow this dependence in our context, since in this case the uniform regularity of the solution of the MFG system (and hence of the master equation) near time  $T$  could be lost (see Proposition 2.3).

Note that  $K_{N,\alpha} \rightarrow +\infty$  as  $N \rightarrow +\infty$  because  $F^N(x, m)$  blows up if  $m$  is a singular measure. So  $F^N$  becomes closer and closer to  $F$  for smooth densities while its regularity at general probability measures deteriorates. However, assumption (10) states that the  $F^N$  are uniformly Höder continuous when evaluated at probability densities which are Höder continuous. Assumption (11) explains how fast the regularity of  $F^N$  degrades as  $N \rightarrow +\infty$ . One could have had different constants for  $F^N$ ,  $\frac{\delta F^N}{\delta m}$  and  $\frac{\delta^2 F^N}{\delta m^2}$ : the choice to have a unique constant is only made for simplicity of presentation.

## 1.4 Main example

Here is a typical example for  $F^N$  when  $F$  satisfies our standing conditions. We assume that  $F^N$  is of the form  $F^N = F^{\epsilon_N}$  where  $(\epsilon_N)$  is a positive sequence which tends to 0 and

$$F^\epsilon(x, m) = (\xi^\epsilon * F(\cdot, \xi^\epsilon * m(\cdot)))(x), \quad (12)$$

with  $\xi^\epsilon(x) = \epsilon^{-d}\xi(x/\epsilon)$ ,  $\xi$  being a symmetric smooth nonnegative kernel with compact support. This example was introduced in [21].

**Proposition 1.2.** *If  $F^N$  is defined by (12), then  $F^N$  is monotone and satisfies (10). Moreover the constants  $k_N^{R,\alpha}$  and  $K_{N,\alpha}$  associated with  $F^N$  as in (9) and in (11) can be estimated by*

$$k_N^{R,\alpha} \leq C(1 + R)\epsilon_N^\alpha, \quad K_{N,\alpha} \leq C\epsilon_N^{-2d-12-3\alpha}, \quad (13)$$

where  $C$  depends on  $F$  and  $\xi$ .

*Proof.* Under the monotonicity assumption on  $F$ , it is known that the  $F^\epsilon$  are monotone (see [21]). Next we prove that the  $F^\epsilon$  satisfy (10). Let  $m dx, m' dx \in \mathcal{P}(\mathbb{T}^d)$  with  $\|m\|_{C^\alpha}, \|m'\|_{C^\alpha} \leq R$ . Then

$$\begin{aligned} & |F^\epsilon(x, m dx) - F^\epsilon(x', m' dx)| \\ & \leq \sup_y |F(x - y, m * \xi^\epsilon(x - y)) - F(x' - y, m' * \xi^\epsilon(x' - y))| \\ & \leq C \sup_y [|x - x'| + |m * \xi^\epsilon(x - y) - m' * \xi^\epsilon(x' - y)|] \\ & \leq C \sup_y [|x - x'| + R|x - x'|^\alpha + |m * \xi^\epsilon(x' - y) - m' * \xi^\epsilon(x' - y)|] \\ & \leq C [R|x - x'|^\alpha + \|m - m'\|_\infty]. \end{aligned}$$

We now estimate the constants  $k_N^{R,\alpha}$  and  $K_{N,\alpha}$ . It is enough to estimate  $k_\epsilon^{R,\alpha}$  and  $K_{\epsilon,\alpha}$ , where

$$k_\epsilon^{R,\alpha} := \sup_m \|F^\epsilon(\cdot, m dx) - F(\cdot, m(\cdot))\|_\infty,$$

the supremum being taken over the densities  $m$  such that  $\|m\|_{C^\alpha} \leq R$ , and

$$K_{\epsilon,\alpha} = \max\{K_{\epsilon,\alpha}^{(0)}, K_{\epsilon,\alpha}^{(1)}, K_{\epsilon,\alpha}^{(2)}\},$$

with

$$K_{\epsilon, \alpha}^{(0)} := \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \|F^\epsilon(\cdot, m)\|_{4+\alpha}, \quad K_{\epsilon, \alpha}^{(1)} := \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \left\| \frac{\delta F^\epsilon}{\delta m}(\cdot, m, \cdot) \right\|_{(4+\alpha, 4+\alpha)}$$

and

$$K_{\epsilon, \alpha}^{(2)} := \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \left\| \frac{\delta^2 F^\epsilon}{\delta m^2}(\cdot, m, \cdot, \cdot) \right\|_{(4+\alpha, 4+\alpha, 4+\alpha)}.$$

Let  $m \in \mathcal{P}(\mathbb{T}^d)$  be such that  $\|m\|_\alpha \leq R$  and  $x \in \mathbb{T}^d$ . As  $\|m\|_\alpha \leq R$ , we have:

$$\|\xi^\epsilon * m - m\|_\infty \leq R\epsilon^\alpha \int_{\mathbb{R}^d} \xi(y)|y|^\alpha \leq CR\epsilon^\alpha,$$

so that

$$\|F(\cdot, \xi^\epsilon * m(\cdot)) - F(\cdot, m(\cdot))\| \leq CR\epsilon^\alpha,$$

because  $F$  is Lipschitz continuous. Thus

$$\begin{aligned} & |F^\epsilon(x, m(x)) - F(x, m(x))| \\ & \leq |\xi^\epsilon * F(\cdot, \xi^\epsilon * m(\cdot))(x) - F(x, \xi^\epsilon * m(x))| + |F(x, \xi^\epsilon * m(x)) - F(x, m(x))| \leq CR\epsilon^\alpha. \end{aligned}$$

Therefore

$$k_N^{R, \alpha} \leq C(1 + R)\epsilon^\alpha,$$

where  $C$  depends on the Lipschitz constant of  $F$  and on  $\xi$ .

For any  $l \in \mathbb{N}^d$ , we have

$$D_x^l F^\epsilon(x, m) = \int_{\mathbb{R}^d} F(y, \xi^\epsilon * m(y)) D^l \xi^\epsilon(x - y) dy.$$

Hence

$$\left| D_x^l F^\epsilon(x, m) - D_x^l F^\epsilon(x', m) \right| \leq \int_{\mathbb{R}^d} |F(y, \xi^\epsilon * m(y))| \left| D^l \xi^\epsilon(x - y) - D^l \xi^\epsilon(x' - y) \right| dy.$$

As  $F$  is Lipschitz continuous, it has a linear growth:

$$|F(y, \xi^\epsilon * m(y))| \leq C(1 + \xi^\epsilon * m(y)).$$

Thus, as  $\xi^\epsilon$  has a support with a uniformly bounded diameter,

$$\|D_x^l F^\epsilon(\cdot, m)\|_\alpha \leq C \|D^l \xi^\epsilon(\cdot)\|_\alpha (1 + \int_{\mathbb{R}^d} |\xi^\epsilon * m(y)| dy) \leq C\epsilon^{-(d+|l|+\alpha)}.$$

So

$$K_{\epsilon, \alpha}^{(0)} \leq C\epsilon^{-(d+4+\alpha)}.$$

On the other hand,

$$\frac{\delta F^\epsilon}{\delta m}(x, m, z) = \int_{\mathbb{R}^d} \sum_{k \in \mathbb{Z}^d} \xi^\epsilon(y - z - k) \frac{\partial F}{\partial m}(y, \xi^\epsilon * m(y)) \xi^\epsilon(x - y) dy$$

and

$$\frac{\delta^2 F^\epsilon}{\delta m^2}(x, m, z, z') = \int_{\mathbb{R}^d} \sum_{k, k' \in \mathbb{Z}^d} \xi^\epsilon(y - z - k) \xi^\epsilon(y - z' - k') \partial_{mm}^2 F(y, \xi^\epsilon * m(y)) \xi^\epsilon(x - y) dy.$$

Hence, for any  $l, l' \in \mathbb{N}^d$ ,

$$\left| D_{x,z}^{l,l'} \frac{\delta F^\epsilon}{\delta m}(x, m, z) \right| \leq C \int_{\mathbb{R}^d} \sum_{k \in \mathbb{Z}^d} |D^{l'} \xi^\epsilon(y - z - k)| |D^l \xi^\epsilon(x - y)| dy,$$

where  $C$  is the Lipschitz constant of  $F$ . Thus, if  $|l|, |l'| \leq 4$  and if the support of  $\xi$  is contained in the ball  $B_R$ , we have

$$\left| D_{x,z}^{l,l'} \frac{\delta F^\epsilon}{\delta m}(x, m, z) \right| \leq C \sum_{k \in \mathbb{Z}^d} \left\| D^{l'} \xi^\epsilon(\cdot - z - k) \right\|_{L^\infty(B_{R\epsilon}(x))} \left\| D^l \xi^\epsilon(x - \cdot) \right\|_{L^1} \leq C \epsilon^{-d-8},$$

where  $B_{R\epsilon}(x)$  is the ball centered at  $x$  and of radius  $R\epsilon$ . In the same way,

$$\left\| \frac{\delta^2 F^\epsilon}{\delta m^2}(\cdot, m, \cdot, \cdot) \right\|_{4,4,4} \leq C \epsilon^{-2d-12}.$$

We can estimate in the same way the Höder norms of  $\delta F^\epsilon / \delta m$  and  $\delta^2 F^\epsilon / \delta m^2$ .  $\square$

## 2 Regularity estimates

In this section, we prove estimates on the solutions of the MFG systems and on the solutions of the master equation with the smoothen coupling  $F^N$ : the whole point is to keep track of the dependence with respect to  $N$  in these estimates.

Let  $U^N$  be the solution to the master equation

$$\begin{cases} -\partial_t U^N(t, x, m) - \Delta_x U^N + H(x, D_x U^N(t, x, m)) \\ \quad - \int_{\mathbb{R}^d} \operatorname{div}_y [D_m U^N(t, x, m, y)] dm(y) \\ \quad + \int_{\mathbb{R}^d} D_m U^N(t, x, m, y) \cdot D_p H(y, D_x U^N) dm(y) = F^N(x, m) \\ \quad \text{in } [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{T}^d), \\ U^N(T, x, m) = G(x) \quad \text{in } \mathbb{R}^d \times \mathcal{P}(\mathbb{T}^d). \end{cases} \quad (14)$$

The existence and the uniqueness of the classical solution  $U^N$  to (14) are established in [2]. Namely:

**Theorem 2.1** ([2]). *Assume that  $F^N$ ,  $G$  and  $H$  satisfy our standing assumptions. Then the first order master equation (14) has a unique classical solution which is of class  $C^2$  with respect to the  $m$  variable.*

As the coupling  $F^N$  becomes increasingly singular as  $N \rightarrow +\infty$ , so does  $U^N$ . The next result collects the upper bounds on the derivatives of  $U^N$ .

**Theorem 2.2.** *Under our standing assumptions, we have, for any  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$  and for any  $\alpha \in (0, 1)$ :*

$$\|U(t_0, \cdot, m_0)\|_{4+\alpha} \leq CK_{N,\alpha}^3,$$

$$\left\| \frac{\delta U^N}{\delta m}(t_0, \cdot, m_0, \cdot) \right\|_{(k+\alpha, k+\alpha)} + \|D_m U^N(t_0, \cdot, m_0, \cdot)\|_{(k+\alpha, k-1+\alpha)} \leq CK_{N,\alpha}^{3k-2},$$

if  $k \in \{1, \dots, 4\}$ , and, if  $k = 2, 3$ ,

$$\left\| \frac{\delta^2 U^N}{\delta m^2}(t_0, \cdot, m_0, \cdot, \cdot) \right\|_{k+\alpha, k-1+\alpha, k-1+\alpha} + \|D_{mm}^2 U^N(t_0, \cdot, m_0, \cdot, \cdot)\|_{k+\alpha, k-2+\alpha, k-2+\alpha} \leq CK_{N,\alpha}^{12k},$$

where  $C$  depends on  $\alpha$  and on the data but not on  $N$ ,  $t_0$  or  $m_0$ .

As the bound on the derivative  $D_m U^N$  (respectively  $D_{mm}^2 U^N$ ) provides a bound on the Lipschitz continuity of  $U^N$  (respectively  $D_m U^N$ ), we have:

$$\|U^N(t, \cdot, m_1) - U^N(t, \cdot, m_2)\|_{k+\alpha} \leq CK_{N,\alpha}^{3k-2} \mathbf{d}_1(m_1, m_2)$$

and

$$\|D_m U^N(t, \cdot, m_1, \cdot) - D_m U^N(t, \cdot, m_2, \cdot)\|_{k+\alpha, k-2+\alpha} \leq CK_{N,\alpha}^{12k} \mathbf{d}_1(m_1, m_2),$$

for any  $m_1, m_2 \in \mathcal{P}(\mathbb{T}^d)$  and  $k = 2, 3$ .

The proof of Theorem 2.1 consists in estimating carefully the various steps in the construction of  $U^N$  in [2]. It is given through a series of statements: Proposition 2.5 for the space regularity of  $U^N$ , Corollary 2.8 for the bound on  $\frac{\delta U^N}{\delta m}$  and Corollary 2.10 for the estimate on  $\frac{\delta^2 U^N}{\delta m^2}$ .

Let us recall [2] that the map  $U^N$  is given by the representation formula:

$$U^N(t_0, x, m_0) = u^N(t_0, x)$$

for any  $(t_0, x, m_0) \in [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)$ , where  $(u^N, m^N)$  is the unique solution to the MFG system

$$\begin{cases} -\partial_t u^N - \Delta u^N + H(x, Du^N) = F^N(x, m^N(t)) & \text{in } [t_0, T] \times \mathbb{T}^d, \\ \partial_t m^N - \Delta m^N - \operatorname{div}(m^N D_p H(x, Du^N)) = 0 & \text{in } [t_0, T] \times \mathbb{T}^d, \\ u^N(T, x) = G(x), m^N(t_0, \cdot) = m_0 & \text{in } \mathbb{T}^d. \end{cases} \quad (15)$$

## 2.1 Estimates on the MFG systems for smooth initial conditions

Fix an initial condition  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$ . We consider the MFG system:

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m(t, x)) & \text{in } [t_0, T] \times \mathbb{T}^d, \\ \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } [t_0, T] \times \mathbb{T}^d, \\ u(T, x) = G(x), m(t_0, \cdot) = m_0 & \text{in } \mathbb{T}^d, \end{cases} \quad (16)$$

and compare its solution with the solution to the MFG system (15).

**Proposition 2.3.** *Under our standing assumptions, let  $t_0 \in [0, T]$ ,  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  be a positive density of class  $C^\alpha$  (where  $\alpha \in (0, 1)$ ) and  $(u^N, m^N)$  and  $(u, m)$  be the solution to the MFG systems (15) and (16) respectively. Then there exists  $\beta \in (0, \alpha]$  such that the  $(u^N, m^N)$  are bounded in  $C^{1+\beta/2, 2+\beta} \times C^{\beta/2, \beta}$  independently of  $N$ . Moreover,*

$$\forall t, t' \in [t_0, T], \quad \mathbf{d}_1(m^N(t), m^N(t')) + \mathbf{d}_1(m(t), m(t')) \leq C|t - t'|^{1/2} \quad (17)$$

and

$$\sup_{t \in [0, T]} \|u^N(t, \cdot) - u(t, \cdot)\|_{H^1(\mathbb{T}^d)} + \|m^N - m\|_{L^2} \leq Ck_N^{R,\alpha}, \quad (18)$$

where the constants  $C$  and  $R$  depend on the data and on  $m_0$ , but not on  $N$ . In particular,

$$\sup_{t \in [0, T]} \|u^N(t, \cdot) - u(t, \cdot)\|_{W^{1,\infty}} \leq C \left( k_N^{R,\alpha} \right)^{\frac{2}{(d+2)}}.$$

*Proof.* Existence of a solution to (15) and to (16) is well-known: see [20]. Estimates (17) is a known consequence of the  $L^\infty$  bound on  $D_p H$ .

We now check the regularity of  $m^N$ . As  $D_p H$  is bounded and  $m_0$  is in  $C^\alpha$ , standard estimates for parabolic equations in divergence form (Theorem III.10.1 of [17]) state that the  $m^N$  are

bounded in  $C^{\beta/2, \beta}$  for some  $\beta \in (0, \alpha]$ . Note that the bound and  $\beta$  depend on  $\alpha$ ,  $\|D_p H\|_\infty$  and  $\|m_0\|_{C^\alpha}$  only.

We now plug this estimate into the parabolic equation for  $u^N$ . For this we note that the map  $(t, x) \rightarrow F^N(x, m^N(t))$  is uniformly Hölder continuous. Indeed, in view of assumption (10) and the uniform regularity of  $m^N$ ,

$$\begin{aligned} |F^N(x, m^N(t)) - F^N(x', m^N(t'))| &\leq \kappa_{R, \beta} \left( |x - x'|^\beta + \|m^N(t, \cdot) - m^N(t', \cdot)\|_\infty \right) \\ &\leq \kappa_{R, \beta} \left( |x - x'|^\beta + |t - t'|^{\beta/2} \right) \end{aligned}$$

where  $R := \sup_N \|m^N\|_{C^{\beta/2, \beta}} + \|m\|_{C^{\beta/2, \beta}}$ . Since the terminal condition  $G$  is  $C^{2+\beta}$  and is independent of  $m^N$  and since  $H$  is Lipschitz continuous, standard estimates on Hamilton-Jacobi equations imply that the  $u^N$  are bounded in  $C^{1+\beta/2, 2+\beta}$ .

We now establish (18): following [19, 20], we have

$$\begin{aligned} &\left[ \int_{\mathbb{T}^d} (u^N - u)(m^N - m) \right]_0^T \\ &= - \int_0^T \int_{\mathbb{T}^d} m (H(x, Du^N) - H(x, Du) - D_p H(x, Du) \cdot D(u^N - u)) \\ &\quad - \int_0^T \int_{\mathbb{T}^d} m^N (H(x, Du) - H(x, Du^N) - D_p H(x, Du^N) \cdot D(u - u^N)) \\ &\quad - \int_0^T \int_{\mathbb{T}^d} (F^N(x, m^N(t)) - F(x, m(t, x)))(m^N(t, x) - m(t, x)). \end{aligned}$$

Note that, on the one hand,  $m^N(0) = m(0) = m_0$  and  $u^N(T) = u(T) = G$ . So the left-hand side vanishes. On the other hand, by strong maximum principle,  $m$  is bounded below by a positive constant since  $m_0$  is positive. As the  $u^N$  and  $u$  are uniformly Lipschitz continuous and assumption (8) holds, we obtain:

$$C^{-1} \int_0^T \int_{\mathbb{T}^d} |Du^N - Du|^2 \leq - \int_0^T \int_{\mathbb{T}^d} (F^N(x, m^N(t)) - F(x, m(t, x)))(m^N(t, x) - m(t, x)).$$

As  $F = F(x, m)$  is increasing in the second variable with  $\frac{\partial F}{\partial m} \geq \delta$  and as assumption (9) holds, we have:

$$\begin{aligned} &\int_0^T \int_{\mathbb{T}^d} (F^N(x, m^N(t)) - F(x, m(t, x)))(m^N(t, x) - m(t, x)) \\ &= \int_0^T \int_{\mathbb{T}^d} (F^N(x, m^N(t)) - F(x, m^N(t, x)))(m^N(t, x) - m(t, x)) \\ &\quad + \int_0^T \int_{\mathbb{T}^d} (F(x, m^N(t, x)) - F(x, m(t, x)))(m^N(t, x) - m(t, x)) \\ &\geq -Ck_N^{R, \beta} \|m^N - m\|_{L^1} + \delta \int_0^T \int_{\mathbb{T}^d} (m^N(t, x) - m(t, x))^2. \end{aligned}$$

We obtain therefore

$$C^{-1} \int_0^T \int_{\mathbb{T}^d} |Du^N - Du|^2 + \delta \int_0^T \int_{\mathbb{T}^d} (m^N(t, x) - m(t, x))^2 \leq Ck_N^{R, \beta} \|m^N - m\|_{L^1} \leq Ck_N^{R, \beta} \|m^N - m\|_{L^2}.$$

Hence

$$\|Du^N - Du\|_{L^2} + \|m^N - m\|_{L^2} \leq Ck_N^{R, \beta}.$$

In particular

$$\begin{aligned} & \|F^N(\cdot, m^N(t)) - F(\cdot, m(t, \cdot))\|_{L^2} \\ & \leq \|F^N(\cdot, m^N(t)) - F(\cdot, m^N(t, \cdot))\|_{\infty} + \|F(\cdot, m^N(t, \cdot)) - F(\cdot, m(t, \cdot))\|_{L^2} \\ & \leq Ck_N^{R,\beta} + C\|m^N - m\|_{L^2} \leq Ck_N^{R,\beta}. \end{aligned}$$

Therefore the difference  $w := u^N - u$  satisfies

$$-\partial_t w - \Delta w = h(t, x)$$

with  $h(t, x) = F^N(x, m^N(t)) - F(x, m(t, x)) - H(x, Du^N(t, x)) + H(x, Du(t, x))$ . By our previous bounds, we have  $\|h\|_{L^2} \leq Ck_N^{R,\beta}$ , so that standard estimates on the heat equation imply that

$$\sup_{t \in [t_0, T]} \|u^N(t, \cdot) - u(t, \cdot)\|_{H^1(\mathbb{T}^d)} \leq Ck_N^{R,\beta}.$$

As, for any smooth map  $\phi : \mathbb{T}^d \rightarrow \mathbb{R}$ , one has:  $\|\phi\|_{\infty} \leq C\|\phi\|_{L^2}^{\frac{2}{d+2}} \|D\phi\|_{\infty}^{\frac{d}{d+2}}$ , and since  $u^N$  and  $u$  are bounded in  $C^{1+\beta/2, 2+\beta}$ , we get

$$\|u^N - u\|_{\infty} + \|Du^N - Du\|_{\infty} \leq C \left(k_N^{R,\beta}\right)^{\frac{2}{d+2}}.$$

We conclude by recalling that  $k_N^{R,\beta} \leq k_N^{R,\alpha}$ . □

A straightforward consequence of Proposition 2.3 is the following estimate on optimal trajectories associated with the MFG systems (15) and (16).

**Corollary 2.4.** *Let  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,  $(u^N, m^N)$  and  $(u, m)$  be the solution to the MFG system (15) and (16) respectively. Let  $t_0 \in [0, T)$  and  $Z$  be a random variable independent of a Brownian motion  $(B_t)$ . If  $(X_t)$  and  $(X_t^N)$  are the solution to*

$$\begin{cases} dX_t = -D_p H(X_t, Du(t, X_t))dt + \sqrt{2}dB_t & \text{in } [t_0, T], \\ X_{t_0} = Z, \end{cases}$$

and

$$\begin{cases} dX_t^N = -D_p H(X_t, Du^N(t, X_t^N))dt + \sqrt{2}dB_t & \text{in } [t_0, T], \\ X_{t_0}^N = Z, \end{cases}$$

then

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |X_t - X_t^N| \right] \leq C \left(k_N^{R,\alpha}\right)^{\frac{2}{d+2}},$$

where  $C$  and  $R$  are as in Proposition 2.3.

*Proof.* By Proposition 2.3, we have

$$\|Du^N - Du\|_{\infty} \leq C \left(k_N^{R,\alpha}\right)^{\frac{2}{d+2}}.$$

The conclusion follows easily since  $D_p H$  is Lipschitz continuous. □

## 2.2 Estimates on the MFG systems for general initial conditions

We now establish regularity estimates for the MFG system which are valid for any initial conditions.

**Proposition 2.5.** *Let  $U^N$  be the solution to the master equation (14),  $(u^N, m^N)$  be the solution to (15) for an arbitrary initial condition  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$ . Then, for any  $\alpha \in (0, 1)$ , we have*

$$\sup_{t \in [t_0, T]} \|u^N(t)\|_{4+\alpha} \leq CK_{N,\alpha}^3,$$

where  $C$  depends on  $\alpha$ . In particular,

$$\sup_{t \in [0, T], m \in \mathcal{P}(\mathbb{T}^d)} \|U^N(t, \cdot, m)\|_{4+\alpha} \leq CK_{N,\alpha}^3.$$

*Proof.* Assumption (11) implies:

$$\sup_{t \in [0, T]} \|D^l F^N(\cdot, m^N(t))\|_\infty \leq K_{N,\alpha} \quad \text{for any } l \in \mathbb{N}^d, \text{ with } |l| \leq 4.$$

By maximum principle we have

$$\|u^N\|_\infty \leq C (\|H(\cdot, 0)\|_\infty + \|G\|_\infty + \|F^N\|_\infty) \leq CK_{N,\alpha}.$$

Standard Lipschitz estimates for Hamilton-Jacobi equations (with a globally Lipschitz continuous Hamiltonian  $H$ ) lead to

$$\|Du^N\|_\infty \leq C (1 + \|DG\|_\infty + \|D_x F^N\|_\infty) \leq CK_{N,\alpha}.$$

For any  $l \in \mathbb{N}^d$  with  $|l| = 1$ , the map  $w_l := D^l u^N$  solves the linear equation with bounded coefficient:

$$\begin{cases} -\partial_t w_l - \Delta w_l + D_p H(x, Du^N) \cdot Dw_l = D_x^l F^N(x, m^N) - D_x^l H(x, Du^N) & \text{in } (t_0, T) \times \mathbb{T}^d, \\ w_l(T, x) = D^l G(x) & \text{in } \mathbb{T}^d. \end{cases}$$

So, for any  $\alpha \in (0, 1)$ , we have (Proposition 4.1 in appendix)

$$\sup_{t \in [t_0, T]} \|w_l(t)\|_{1+\alpha} \leq C \left[ \|D_x^l F^N(\cdot, m^N)\|_\infty + \|D^l G\|_{1+\alpha} + \|D_x^l H(\cdot, Du^N)\|_\infty \right] \leq CK_{N,\alpha}.$$

This implies that

$$\sup_{t \in [t_0, T]} \|u^N(t)\|_{2+\alpha} \leq CK_{N,\alpha}.$$

We now estimate the second order derivative. Let  $w_l := D^l u^N$  with  $l = l_1 + l_2$ ,  $l_1, l_2 \in \mathbb{N}^d$  and  $|l_1| = |l_2| = 1$ . Then  $w_l$  solves the linear equation with bounded coefficients:

$$\begin{cases} -\partial_t w_l - \Delta w_l + D_p H \cdot Dw_l = D_{xx}^l F^N(x, m^N) - D_{xp}^2 H e_{l_1} \cdot Dw_{l_2} \\ \quad - D_{xp}^2 H e_{l_2} \cdot Dw_{l_1} - D_{pp}^2 H Dw_{l_1} \cdot Dw_{l_2} - D_{xx}^2 H e_{l_1} \cdot e_{l_2} & \text{in } (t_0, T) \times \mathbb{T}^d, \\ w_l(T, x) = D^l G(x) & \text{in } \mathbb{T}^d, \end{cases}$$

(where  $H = H(x, Du^N)$ ) so that

$$\begin{aligned} \sup_{t \in [t_0, T]} \|w_l(t)\|_{1+\alpha} \leq & C \left[ \|D_{xx}^2 F^N\|_\infty + \|D_{xp}^2 H\|_\infty \|D^2 u^N\|_\infty + \|D_{pp}^2 H\|_\infty \|D^2 u^N\|_\infty^2 \right. \\ & \left. + \|D_{xx}^2 H(\cdot, Du^N)\|_\infty + \|D^l G(\cdot)\|_{1+\alpha} \right] \leq CK_{N,\alpha}^2, \end{aligned}$$

where we used our previous estimate on  $\sup_t \|u^N(t)\|_{2+\alpha}$ . We infer that

$$\sup_{t \in [0, T]} \|u^N(t)\|_{3+\alpha} \leq CK_{N, \alpha}^2.$$

Finally, for  $l \in \mathbb{N}^d$  with  $|l| = 3$ , one can prove in the same way that  $w_l := D^l u^N$  satisfies

$$\begin{aligned} \sup_{t \in [t_0, T]} \|w_l(t)\|_{1+\alpha} &\leq C \left[ \|D_{xxx}^3 F^N\|_\infty + \|D_{xp}^2 H\|_\infty \|D^3 u^N\|_\infty + \|D_{pp}^2 H\|_\infty \|D^2 u^N\|_\infty \|D^3 u^N\|_\infty \right. \\ &\quad + \|D_{xpp}^3 H\|_\infty \|D^2 u^N\|_\infty + \|D_{xpp}^3 H\|_\infty \|D^2 u^N\|_\infty^2 + \|D_{xp}^2 H\|_\infty \|D^3 u^N\|_\infty \\ &\quad + \|D_{xpp}^3 H\|_\infty \|D^2 u^N\|_\infty^2 + \|D_{ppp}^3 H\|_\infty \|D^2 u^N\|_\infty^3 + \|D_{pp}^2 H\|_\infty \|D^2 u^N\|_\infty \|D^3 u^N\|_\infty \\ &\quad \left. + \|D_{xxx}^3 H\|_\infty + \|D_{xpp}^3 H\|_\infty \|D^2 u^N\|_\infty \right] \leq CK_{N, \alpha}^3. \end{aligned}$$

Therefore

$$\sup_{t \in [t_0, T]} \|u^N\|_{4+\alpha} \leq CK_{N, \alpha}^3.$$

□

## 2.3 Estimates for a linearized system

We consider systems of the form

$$\begin{cases} (i) & -\partial_t z - \Delta z + V(t, x) \cdot Dz = \frac{\delta F^N}{\delta m}(x, m^N(t))(\rho(t)) + b(t, x) & \text{in } [t_0, T] \times \mathbb{T}^d, \\ (ii) & \partial_t \rho - \Delta \rho - \operatorname{div}(\rho V(t, x)) - \operatorname{div}(m^N \Gamma Dz + c) = 0 & \text{in } [t_0, T] \times \mathbb{T}^d, \\ (iii) & z(T, x) = 0, \rho(t_0) = \rho_0 & \text{in } \mathbb{T}^d, \end{cases} \quad (19)$$

where  $V : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a given vector field,  $m^N \in C^0([0, T], \mathcal{P}(\mathbb{T}^d))$ ,  $\Gamma : [0, T] \times \mathbb{T}^d \rightarrow \mathbb{R}^{d \times d}$  is a continuous map with values into the family of symmetric matrices and where the maps  $b : [t_0, T] \times \mathbb{T}^d \rightarrow \mathbb{R}$  and  $c : [t_0, T] \times \mathbb{T}^d \rightarrow \mathbb{R}^d$  are given. We assume that, for any  $\alpha \in (0, 1)$ , there is a constant  $\bar{C} > 0$  (depending on  $\alpha$ ) such that

$$\begin{aligned} \|D^k V\|_\infty &\leq \bar{C} K_{N, \alpha}^{3k} & \forall k \in \{0, \dots, 3\}, \\ \forall (t, x) \in [t_0, T] \times \mathbb{T}^d, & \quad 0 \leq \Gamma(t, x) \leq \bar{C} I_d. \end{aligned} \quad (20)$$

Typically,  $V(t, x) = D_p H(x, Du^N(t, x))$ ,  $\Gamma(t, x) = D_{pp}^2 H(x, Du^N(t, x))$  for some solution  $(u^N, m^N)$  of the MFG system (15) starting from some initial data  $m(t_0) = m_0$ . Proposition 2.5 then implies that (20) holds.

Following [2], given  $\rho_0 \in C^{-(k+\alpha)}(\mathbb{T}^d)$ ,  $(b, c) \in C^0([t_0, T], C^{k+\alpha}(\mathbb{T}^d) \times C^{-(k-1+\alpha)}(\mathbb{T}^d))$  (where  $k \in \{1, \dots, 4\}$ ), there exists a unique solution  $(z, \rho)$  to system (19) in the sense of distribution in  $C^0([0, T], C^{k+\alpha}(\mathbb{T}^d) \times C^{-(k+\alpha)}(\mathbb{T}^d))$ .

**Proposition 2.6.** *For any  $k \in \{1, \dots, 4\}$  and any  $\alpha \in (0, 1)$ , we have,*

$$\sup_{t \in [0, T]} \|z(t)\|_{k+\alpha} \leq CK_{N, \alpha}^{3k-2} M_k \quad \text{and} \quad \sup_{t \in [t_0, T]} \|\rho(t)\|_{-(k+\alpha)} \leq CK_{N, \alpha}^{3(k-1)} M_k,$$

where  $C$  depends on  $\bar{C}$  and  $\alpha$ , but not on  $N$ ,  $V$ ,  $\Gamma$ ,  $m^N$ , and where

$$M_k := \|\rho_0\|_{-(k+\alpha)} + \sup_{t \in [t_0, T]} \|c(t)\|_{-(k-1+\alpha)} + \sup_{t \in [t_0, T]} \|b(t)\|_{k+\alpha}.$$



*Proof.* To simplify the notation, we argue as if the solution  $(z, \rho)$  is smooth. The general estimate is obtained by approximation (see [2]).

*Step 1: Structure estimate.* Computing  $\frac{d}{dt} \int_{\mathbb{T}^d} z \rho$ , we find:

$$\frac{d}{dt} \int_{\mathbb{T}^d} z \rho = - \int_{\mathbb{T}^d} \left\{ \left( \frac{\delta F^N}{\delta m}(x, m^N)(\rho) + b \right) \rho + (m^N \Gamma D z + c) \cdot D z \right\}.$$

So, using the initial and terminal conditions for  $z$  and  $\rho$  and the fact that  $F^N$  is monotone, we get

$$\int_{t_0}^T \int_{\mathbb{T}^d} m^N \Gamma D z \cdot D z \leq \int_{\mathbb{T}^d} z(t_0) \rho(t_0) - \int_{t_0}^T \int_{\mathbb{T}^d} \{b \rho + c \cdot D z\}. \quad (21)$$

*Step 2: Estimate for a linear backward equation.* In order to estimate  $\rho$ , we use a duality method requiring estimates on a backward system: given  $t_1 \in (t_0, T]$  and  $w^1 \in C^\infty$ , let  $w$  solve

$$\begin{cases} -\partial_t w - \Delta w + V(t, x) \cdot D w = 0 & \text{in } [t_0, t_1] \times \mathbb{T}^d, \\ w(t_1, x) = w^1(x) & \text{in } \mathbb{T}^d. \end{cases} \quad (22)$$

where  $w^1 \in C^\infty(\mathbb{T}^d)$ . Proposition 4.1 in the Appendix states that there exists a constant  $C$  depending on  $\|V\|_\infty$ ,  $d$ ,  $\alpha$  only such that

$$\sup_{t \in [0, T]} \|w(t)\|_{1+\alpha} \leq C \|w^1\|_{1+\alpha}.$$

Note that, for any  $l \in \mathbb{N}^d$  with  $k := |l| \in \{1, \dots, 3\}$ , the map  $\hat{w} := D^l w$  solves an equation of the form:

$$\begin{cases} -\partial_t \hat{w} - \Delta \hat{w} + V(t, x) \cdot D \hat{w} = g_l & \text{in } [t_0, t_1] \times \mathbb{T}^d, \\ \hat{w}(t_1, x) = D^l w^1(x) & \text{in } \mathbb{T}^d, \end{cases}$$

where  $g_l$  is a linear combination of the  $D^{l'} w$  with  $1 \leq |l'| \leq k$  and where the coefficient in front of  $D^{l'} w$  is proportional to a derivative of order  $k - |l'| + 1$  of  $V$  in the space variable. By Proposition 4.1 and (20) we get therefore

$$\begin{aligned} \sup_{t \in [0, T]} \|D^l w(t)\|_{1+\alpha} &\leq C \left[ \|D^l w^1\|_{1+\alpha} + \|g_l\|_\infty \right] \\ &\leq C \left[ \|D^l w^1\|_{1+\alpha} + \sum_{1 \leq |l'| \leq k} K_{N, \alpha}^{3(k-|l'|+1)} \|D^{l'} w\|_\infty \right]. \end{aligned}$$

By induction, this implies that, for  $k \in \{1, \dots, 4\}$ ,

$$\sup_{t \in [0, T]} \|w(t)\|_{k+\alpha} \leq C K_{N, \alpha}^{3(k-1)} \|w^1\|_{k+\alpha}. \quad (23)$$

*Step 3: Estimate of  $\rho$  by duality.* Let us fix  $t_1 \in (t_0, T]$ ,  $w^1 \in C^\infty$  and let  $w$  be the solution to (22). As  $\rho$  solves (19), we have, for  $k \in \{1, \dots, 4\}$ ,

$$\begin{aligned} \int_{\mathbb{T}^d} w^1 \rho(t_1) &= \int_{\mathbb{T}^d} w(t_0) \rho_0 - \int_{t_0}^{t_1} \int_{\mathbb{T}^d} (m^N \Gamma D z + c) \cdot D w \\ &\leq \|\rho_0\|_{-(k+\alpha)} \|w(t_0)\|_{k+\alpha} + \left( \int_{t_0}^T \int_{\mathbb{T}^d} m^N \Gamma D z \cdot D z \right)^{\frac{1}{2}} \left( \int_{t_0}^{t_1} \int_{\mathbb{T}^d} m^N \Gamma D w \cdot D w \right)^{\frac{1}{2}} \\ &\quad + C \sup_t \|c(t)\|_{-(k-1+\alpha)} \sup_t \|D w\|_{k-1+\alpha} \\ &\leq \|\rho_0\|_{-(k+\alpha)} \|w(t_0)\|_{k+\alpha} + C \|D w\|_\infty \left( \int_{t_0}^T \int_{\mathbb{T}^d} m^N \Gamma D z \cdot D z \right)^{\frac{1}{2}} \\ &\quad + C \sup_t \|c(t)\|_{-(k-1+\alpha)} \sup_t \|D w\|_{k-1+\alpha}, \end{aligned}$$

where we used the fact that  $\Gamma$  is bounded and  $\int_{\mathbb{T}^d} m^N(t) = 1$  in the last inequality. Recalling (23), we get:

$$\int_{\mathbb{T}^d} w^1 \rho(t_1) \leq C \|w^1\|_{k+\alpha} \left\{ K_{N,\alpha}^{3(k-1)} (\|\rho_0\|_{-(k+\alpha)} + \sup_t \|c(t)\|_{-(k-1+\alpha)}) + \left( \int_{t_0}^T \int_{\mathbb{T}^d} m^N \Gamma D z \cdot D z \right)^{\frac{1}{2}} \right\}.$$

Taking the supremum with respect to  $t_1$  and to  $w^1$  with  $\|w^1\|_{k+\alpha} \leq 1$ , we obtain therefore

$$\sup_t \|\rho(t)\|_{-(k+\alpha)} \leq C K_{N,\alpha}^{3(k-1)} \left( \|\rho_0\|_{-(k+\alpha)} + \sup_t \|c(t)\|_{-(k-1+\alpha)} \right) + C \left( \int_{t_0}^T \int_{\mathbb{T}^d} m^N \Gamma D z \cdot D z \right)^{\frac{1}{2}}.$$

For  $r \geq 1$ , we plug (21) into the above estimate:

$$\begin{aligned} \sup_t \|\rho(t)\|_{-(k+\alpha)} &\leq C K_{N,\alpha}^{3(k-1)} \left( \|\rho_0\|_{-(k+\alpha)} + \sup_t \|c(t)\|_{-(k-1+\alpha)} \right) + C \|z(t_0)\|_r^{\frac{1}{2}} \|\rho_0\|_{-r}^{\frac{1}{2}} \\ &\quad + C \left( \sup_t \|b(t)\|_{k+\alpha}^{\frac{1}{2}} \sup_t \|\rho(t)\|_{-(k+\alpha)}^{\frac{1}{2}} + \sup_t \|c(t)\|_{-(r-1)}^{\frac{1}{2}} \sup_t \|z(t)\|_r^{\frac{1}{2}} \right). \end{aligned}$$

Rearranging we find:

$$\sup_{t \in [t_0, T]} \|\rho(t)\|_{-(k+\alpha)} \leq C K_{N,\alpha}^{3(k-1)} M_k + C \sup_{t \in [t_0, T]} \|z(t)\|_r^{\frac{1}{2}} \left( \|\rho_0\|_{-r}^{\frac{1}{2}} + \sup_{t \in [t_0, T]} \|c(t)\|_{-(r-1)}^{\frac{1}{2}} \right),$$

where  $M_k$  is defined in the Proposition.

*Step 4: Estimate of  $z$ .* Fix  $l \in \mathbb{N}^d$  with  $k := |l| \in \{0, \dots, 4\}$ . In view of the equation satisfied by  $z$ , the map  $\hat{z} := D^l z$  solves an equation of the form

$$\begin{cases} -\partial_t \hat{z} - \Delta \hat{z} + V \cdot D \hat{z} = D^l \frac{\delta F^N}{\delta m}(x, m^N(t))(\rho(t)) + D^l b + g_l & \text{in } [t_0, T] \times \mathbb{T}^d, \\ \hat{z}(T, x) = 0 & \text{in } \mathbb{T}^d, \end{cases}$$

where  $g_l$  is as in step 2 with  $z$  replacing  $w$ . Proposition 4.1 thus implies that

$$\sup_{t \in [0, T]} \|D^l z(t)\|_{1+\alpha} \leq C \left[ \|D^l \frac{\delta F^N}{\delta m}(\cdot, m^N(\cdot))(\rho(\cdot))\|_\infty + \|D^l b\|_\infty + \sum_{1 \leq |l'| \leq k} K_{N,\alpha}^{3(k-|l'|+1)} \|D^{l'} z\|_\infty \right],$$

where, from assumption (11) and Step 3 for  $r = k + 1 + \alpha$ :

$$\begin{aligned} \left\| D^l \frac{\delta F^N}{\delta m}(\cdot, m^N(\cdot))(\rho(\cdot)) \right\|_\infty &\leq \sup_m \left\| \frac{\delta F^N}{\delta m}(\cdot, m, \cdot) \right\|_{k+\alpha, k+1+\alpha} \sup_t \|\rho(t)\|_{-(k+1+\alpha)} \\ &\leq C K_{N,\alpha}^{3k+1} M_{k+1} + C K_{N,\alpha} \sup_t \|z(t)\|_{k+1+\alpha}^{\frac{1}{2}} \left( \|\rho_0\|_{-(k+1+\alpha)}^{\frac{1}{2}} + \sup_t \|c(t)\|_{-(k+\alpha)}^{\frac{1}{2}} \right). \end{aligned}$$

So we find

$$\begin{aligned} \sup_{t \in [0, T]} \|z(t)\|_{k+1+\alpha} &\leq C \left[ K_{N,\alpha}^{3k+1} M_{k+1} + K_{N,\alpha} \sup_t \|z(t)\|_{k+1+\alpha}^{\frac{1}{2}} \left( \|\rho_0\|_{-(k+1+\alpha)}^{\frac{1}{2}} + \sup_t \|c(t)\|_{-(k+\alpha)}^{\frac{1}{2}} \right) \right. \\ &\quad \left. + \|b\|_k + \sum_{1 \leq n \leq k} K_{N,\alpha}^{3(k-n+1)} \sup_{t \in [0, T]} \|z(t)\|_{n+\alpha} \right]. \end{aligned}$$

We rearrange the expression to obtain

$$\begin{aligned} \sup_{t \in [0, T]} \|z(t)\|_{k+1+\alpha} &\leq C \left[ K_{N, \alpha}^{3k+1} M_{k+1} + K_{N, \alpha}^2 \left( \|\rho_0\|_{-(k+1+\alpha)} + \sup_t \|c(t)\|_{-(k+\alpha)} \right) \right. \\ &\quad \left. + \|b\|_k + \sum_{1 \leq n \leq k} K_{N, \alpha}^{3(k-n+1)} \sup_{t \in [0, T]} \|z(t)\|_{n+\alpha} \right] \\ &\leq C \left[ K_{N, \alpha}^{3k+1} M_{k+1} + \sum_{1 \leq n \leq k} K_{N, \alpha}^{3(k-n+1)} \sup_{t \in [0, T]} \|z(t)\|_{n+\alpha} \right]. \end{aligned}$$

By induction we infer that, for  $k \in \{1, \dots, 4\}$ ,

$$\sup_{t \in [0, T]} \|z(t)\|_{k+\alpha} \leq C K_{N, \alpha}^{3k-2} M_k.$$

Plugging this inequality into our estimate for  $\rho$  (in step 3) gives:

$$\sup_{t \in [t_0, T]} \|\rho(t)\|_{-(k+\alpha)} \leq C K_{N, \alpha}^{3(k-1)} M_k.$$

□

## 2.4 Estimates for $\frac{\delta U^N}{\delta m}$

In this section we provide estimates for  $\frac{\delta U^N}{\delta m}$  where  $U^N$  is the solution of the master equation (14). Following the construction of [2], we can express this derivative in terms of a linearized system. Let us fix  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$  and let  $(m^N, u^N)$  be the solution to the MFG system (15) with initial condition  $m(t_0) = m_0$ . Recall that, by definition,  $U^N(t_0, x, m_0) = u^N(t_0, x)$ .

For any  $\mu_0 \in C^\infty(\mathbb{T}^d)$ , we consider the solution  $(z, \rho)$  to the linearized system

$$\begin{cases} -\partial_t z - \Delta z + D_p H(x, Du^N) \cdot Dz = \frac{\delta F^N}{\delta m}(x, m^N(t))(\rho(t)) \text{ in } (0, T) \times \mathbb{T}^d, \\ \partial_t \rho - \Delta \rho - \operatorname{div}(\rho D_p H(x, Du^N)) - \operatorname{div}(m^N D_{pp}^2 H(x, Du^N) Dz) = 0 \text{ in } (0, T) \times \mathbb{T}^d, \\ z(T, \cdot) = 0, \rho(t_0, \cdot) = \rho_0 \text{ in } \mathbb{T}^d. \end{cases} \quad (24)$$

We proved in [2] the identity

$$z(t_0, x) = \int_{\mathbb{T}^d} \frac{\delta U^N}{\delta m}(t_0, x, m_0, y) \rho_0(y) dy.$$

In order to estimate  $\frac{\delta U^N}{\delta m}$ , which just need to estimate  $z$ : this is the aim of the next statement.

**Proposition 2.7.** *The unique solution  $(z, \rho)$  of (24) satisfies, for  $k \in \{1, \dots, 4\}$  and any  $\alpha \in (0, 1)$ ,*

$$\sup_{t \in [t_0, T]} \|z(t, \cdot)\|_{k+\alpha} \leq C K_{N, \alpha}^{3k-2} \|\rho_0\|_{-(k+\alpha)}, \quad (25)$$

$$\sup_{t \in [t_0, T]} \|\rho(t)\|_{-(k+\alpha)} \leq C K_{N, \alpha}^{3(k-1)} \|\rho_0\|_{-(k+\alpha)}, \quad (26)$$

where the constant  $C$  does not depend on  $(t_0, m_0)$  nor on  $N$ .

*Proof.* It is a straightforward application of Proposition 2.6, with  $V(t, x) = D_p H(x, Du^N(t, x))$ ,  $\Gamma(t, x) = D_{pp}^2 H(x, Du^N(t, x))$  and  $b = c = 0$ . □

As in [2] we can derive from the Proposition an estimate on the first order derivative of  $U$  with respect to the measure:

**Corollary 2.8.** *For any  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$  and  $k \in \{1, \dots, 4\}$ , we have:*

$$\left\| \frac{\delta U^N}{\delta m}(t_0, \cdot, m_0, \cdot) \right\|_{(k+\alpha, k+\alpha)} + \|D_m U^N(t_0, \cdot, m_0, \cdot)\|_{(k+\alpha, k-1+\alpha)} \leq CK_{N, \alpha}^{3k-2}.$$

## 2.5 Estimate for $\frac{\delta^2 U^N}{\delta m^2}$

We now estimate the second order derivative with respect to  $m$  of the solution  $U^N$  to the master equation (14). Let us fix  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$  and let  $(m^N, u^N)$  be the solution to the MFG system (15) with initial condition  $m(t_0) = m_0$ . Let  $(z, \rho)$  be a solution of the linearized system (24) with initial condition  $\rho_0$ . The second order linearized system reads

$$\left\{ \begin{array}{l} -\partial_t w - \Delta w + D_p H(x, Du^N) \cdot Dw = \frac{\delta F^N}{\delta m}(x, m^N(t))(\mu(t)) \\ \quad + \frac{\delta^2 F^N}{\delta m^2}(x, m^N(t))(\rho(t), \rho(t)) - D_{pp}^2 H(x, Du^N) Dz \cdot Dz \text{ in } (0, T) \times \mathbb{T}^d, \\ \partial_t \mu - \Delta \mu - \operatorname{div}(\mu D_p H(x, Du^N)) - \operatorname{div}(m^N D_{pp}^2 H(x, Du^N) Dw) \\ \quad = \operatorname{div}(m^N D_{ppp}^3 H(x, Du^N) Dz Dz) + 2 \operatorname{div}(\rho D_{pp}^2 H(x, Du^N) Dz) \text{ in } (0, T) \times \mathbb{T}^d, \\ w(T, \cdot) = 0, \mu(t_0, \cdot) = 0 \text{ in } \mathbb{T}^d. \end{array} \right. \quad (27)$$

Following [2], we have

$$w(t_0, x) = \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\delta^2 U^N}{\delta m^2}(t_0, x, m_0, y, y') \rho_0(y) \rho_0(y') dy dy'.$$

**Proposition 2.9.** *We have, for  $k = 2, 3$ ,*

$$\sup_{t \in [t_0, T]} \|w(t)\|_{k+\alpha} \leq CK_{N, \alpha}^{12k} \|\rho_0\|_{-(k-1+\alpha)}^2.$$

As a consequence, we have:

**Corollary 2.10.** *For any  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$  and  $k = 2, 3$ , we have:*

$$\left\| \frac{\delta^2 U^N}{\delta m^2}(t_0, \cdot, m_0, \cdot, \cdot) \right\|_{k+\alpha, k-1+\alpha, k-1+\alpha} + \|D_{mm}^2 U^N(t_0, \cdot, m_0, \cdot, \cdot)\|_{k+\alpha, k-2+\alpha, k-2+\alpha} \leq CK_{N, \alpha}^{12k}.$$

*Proof of Proposition 2.9.* We apply Proposition 2.6 to  $(w, \mu)$  with initial condition  $\mu(t_0) = 0$  and

$$\begin{aligned} b(t) &= \frac{\delta^2 F^N}{\delta m^2}(x, m^N(t))(\rho(t), \rho(t)) - D_{pp}^2 H(x, Du^N) Dz \cdot Dz, \\ c(t) &= (m^N D_{ppp}^3 H(x, Du^N) Dz Dz) + 2\rho D_{pp}^2 H(x, Du^N) Dz. \end{aligned}$$

We have, for any  $k \in \{1, \dots, 3\}$ ,

$$\begin{aligned} \sup_t \|b(t)\|_{k+\alpha} &\leq CK_{N, \alpha} \sup_t \|\rho(t)\|_{-(k-1+\alpha)}^2 + C \sup_t \|D_{pp}^2 H(\cdot, Du^N(t))\|_{k+\alpha} \sup_t \|z(t)\|_{k+1+\alpha}^2 \\ &\leq CK_{N, \alpha}^{1+6(k-2)} \|\rho_0\|_{-(k-1+\alpha)}^2 + CK_{N, \alpha}^{3k+2(3(k+1)-2)} \|\rho_0\|_{-(k+1+\alpha)}^2 \\ &\leq CK_{N, \alpha}^{9k+2} \|\rho_0\|_{-(k-1+\alpha)}^2, \end{aligned}$$

where we used Proposition 2.5, (25) and (26). Next we estimate  $c$  for  $k \in \{2, 3\}$ :

$$\begin{aligned}
\|c(t)\|_{-(k-1+\alpha)} &\leq C \sup_{\|\phi\|_{k-1+\alpha} \leq 1} \left( \int_{\mathbb{T}^d} m^N \phi D_{ppp}^3 H(x, Du^N) Dz Dz + \int_{\mathbb{T}^d} \phi \rho D_{pp}^2 H(x, Du^N) Dz \right) \\
&\leq C \|Dz\|_\infty^2 + C \sup_{\|\phi\|_{k-1+\alpha} \leq 1} \|\phi D_{pp}^2 H(x, Du^N) Dz\|_{k-1+\alpha} \sup_t \|\rho(t)\|_{-(k-1+\alpha)} \\
&\leq CK_{N,\alpha}^{3(k-1)-2} \|\rho_0\|_{-(k-1+\alpha)}^2 + CK_{N,\alpha}^{3(k-1)+3k-2+3(k-2)} \|\rho_0\|_{-(k+\alpha)} \|\rho_0\|_{-(k-1+\alpha)} \\
&\leq CK_{N,\alpha}^{9(k-1)-2} \|\rho_0\|_{-(k-1+\alpha)}^2.
\end{aligned}$$

Therefore, by Proposition 2.6, we obtain, for  $k \in \{2, 3\}$ :

$$\sup_t \|w(t)\|_{k+\alpha} \leq CK_{N,\alpha}^{3k-2} \left[ \sup_t \|b(t)\|_{k+\alpha} + \sup_t \|c(t)\|_{-(k-1+\alpha)} \right] \leq CK_{N,\alpha}^{12k} \|\rho_0\|_{-(k-1+\alpha)}^2.$$

□

### 3 Convergence

In this section, we consider, for an integer  $N \geq 2$ , a classical solution  $(v^{N,i})_{i \in \{1, \dots, N\}}$  of the Nash system:

$$\left\{ \begin{array}{l} -\partial_t v^{N,i}(t, \mathbf{x}) - \sum_j \Delta_{x_j} v^{N,i}(t, \mathbf{x}) + H(x_i, D_{x_i} v^{N,i}(t, \mathbf{x})) \\ \quad + \sum_{j \neq i} D_p H(x_j, D_{x_j} v^{N,j}(t, \mathbf{x})) \cdot D_{x_j} v^{N,i}(t, \mathbf{x}) = F^N(x_i, m_{\mathbf{x}}^{N,i}) \text{ in } [0, T] \times (\mathbb{T}^d)^N, \\ v^{N,i}(T, \mathbf{x}) = G(x_i) \quad \text{in } (\mathbb{T}^d)^N, \end{array} \right. \quad (28)$$

where we have set, for  $\mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{T}^d)^N$ ,  $m_{\mathbf{x}}^{N,i} = \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}$ .

As  $H$  is Lipschitz continuous, system (28) has a unique classical solution [17]. By uniqueness, the  $v^{N,i}$  enjoy strong symmetry properties. On the one hand,  $v^{N,i}(t, x_1, \dots, x_N)$  is symmetric with respect to the variables  $(x_j)_{j \neq i}$ . On the other hand, for  $j \neq i$ ,  $v^{N,i}(t, \mathbf{x}) = v^{N,j}(t, \mathbf{y})$ , where  $\mathbf{x} = (x_1, \dots, x_N)$  and  $\mathbf{y}$  is obtained from  $\mathbf{x}$  by permuting the  $x_i$  and  $x_j$  variables.

Our aim is to quantify the convergence rate of  $v^{N,i}$  to the solution  $U^N$  of the master equation (14) as  $N$  tends to  $\infty$ .

#### 3.1 Finite dimensional projections of $U^N$

Let  $U^N = U^N(t, x, m)$  be the solution of the second order master equation (14). For  $N \geq 2$  and  $i \in \{1, \dots, N\}$  we set

$$u^{N,i}(t, \mathbf{x}) = U^N(t, x_i, m_{\mathbf{x}}^{N,i}) \quad \text{where } \mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{T}^d)^N, \quad m_{\mathbf{x}}^{N,i} = \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}.$$

Following [2], we know that the  $u^{N,i}$  are of class  $C^2$  with respect to the space variables and  $C^1$  with respect to the time variable, with, for  $i, j, k$  distinct:

$$\begin{aligned}
\partial_t u^{N,i}(t, \mathbf{x}) &= \partial_t U^N(t, x_i, m_{\mathbf{x}}^{N,i}), & D_{x_i} u^{N,i}(t, \mathbf{x}) &= D_x U^N(t, x_i, m_{\mathbf{x}}^{N,i}), \\
D_{x_i x_i}^2 u^{N,i}(t, \mathbf{x}) &= D_{xx}^2 U^N(t, x_i, m_{\mathbf{x}}^{N,i}), & D_{x_j} u^{N,i}(t, \mathbf{x}) &= \frac{1}{N-1} D_m U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j), \\
D_{x_j x_k}^2 u^{N,i}(t, \mathbf{x}) &= \frac{1}{(N-1)^2} D_{mm}^2 U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j, x_k), \\
D_{x_j x_j}^2 u^{N,i}(t, \mathbf{x}) &= \frac{1}{(N-1)^2} D_{mm}^2 U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j, x_j) + \frac{1}{N-1} D_{ym}^2 U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j).
\end{aligned} \tag{29}$$

We estimate how far  $(u^{N,i})_{i \in \{1, \dots, N\}}$  is to be a solution to the Nash system (28):

**Proposition 3.1.** *The map  $(v^{N,i})$  satisfies*

$$\left\{ \begin{array}{l} -\partial_t u^{N,i} - \sum_j \Delta_{x_j} u^{N,i} + H(x_i, D_{x_i} u^{N,i}) \\ \quad + \sum_{j \neq i} D_{x_j} u^{N,i}(t, \mathbf{x}) \cdot D_p H(x_j, D_{x_j} u^{N,j}(t, \mathbf{x})) = F^N(x_i, m_{\mathbf{x}}^{N,i}) + r^{N,i}(t, \mathbf{x}) \\ \quad \text{a.e. in } (0, T) \times \mathbb{T}^{Nd}, \\ u^{N,i}(T, \mathbf{x}) = G(x_i) \quad \text{in } \mathbb{T}^{Nd}, \end{array} \right. \tag{30}$$

where  $r^{N,i} \in C^0([0, T] \times \mathbb{T}^d)$  with

$$\|r^{N,i}\|_{\infty} \leq \frac{C}{N} (\|D_m U^N\|_{\infty} \|D_{m,x} U^N\|_{\infty} + \|D_{mm}^2 U^N\|_{\infty}).$$

*Proof.* As  $U^N$  solves (14), one has at a point  $(t, x_i, m_{\mathbf{x}}^{N,i})$ :

$$\begin{aligned}
& -\partial_t U^N - \Delta_x U^N + H(x_i, D_x U^N) - \int_{\mathbb{T}^d} \operatorname{div}_y [D_m U^N](t, x_i, m_{\mathbf{x}}^{N,i}, y) dm_{\mathbf{x}}^{N,i}(y) \\
& + \int_{\mathbb{T}^d} D_m U^N(t, x_i, m_{\mathbf{x}}^{N,i}, y) \cdot D_p H(y, D_x U^N(t, y, m_{\mathbf{x}}^{N,i})) dm_{\mathbf{x}}^{N,i}(y) = F^N(x_i, m_{\mathbf{x}}^{N,i}).
\end{aligned}$$

So  $u^{N,i}$  satisfies:

$$\begin{aligned}
& -\partial_t u^{N,i} - \Delta_{x_i} u^{N,i} + H(x_i, D_{x_i} u^{N,i}) - \int_{\mathbb{T}^d} \operatorname{div}_y [D_m U^N](t, x_i, m_{\mathbf{x}}^{N,i}, y) dm_{\mathbf{x}}^{N,i}(y) \\
& + \frac{1}{N-1} \sum_{j \neq i} D_m U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) \cdot D_p H(x_j, D_x U^N(t, x_j, m_{\mathbf{x}}^{N,i})) = F^N(x_i, m_{\mathbf{x}}^{N,i}).
\end{aligned}$$

Note that, by (29), for any  $j \neq i$ , we have:

$$\frac{1}{N-1} D_m U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) = D_{x_j} u^{N,i}(t, \mathbf{x}).$$

In particular,

$$\|D_{x_j} u^{N,i}\|_{\infty} \leq \frac{1}{N} \|D_m U^N\|_{\infty}. \tag{31}$$

By the Lipschitz continuity of  $D_x U^N$  with respect to  $m$ , we have

$$|D_x U^N(t, x_j, m_{\mathbf{x}}^{N,i}) - D_x U^N(t, x_j, m_{\mathbf{x}}^{N,j})| \leq \|D_{m,x} U^N\|_{\infty} \mathbf{d}_1(m_{\mathbf{x}}^{N,i}, m_{\mathbf{x}}^{N,j}) \leq \frac{C}{N} \|D_{m,x} U^N\|_{\infty},$$

so that, by Lipschitz continuity of  $D_p H$ ,

$$|D_p H(x_j, D_x U^N(t, x_j, m_{\mathbf{x}}^{N,i})) - D_p H(x_j, D_x u^{N,j}(t, \mathbf{x}))| \leq \frac{C}{N} \|D_{m,x} U^N\|_{\infty}. \quad (32)$$

Therefore

$$\begin{aligned} & \frac{1}{N-1} \sum_{j \neq i} D_m U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) \cdot D_p H(x_j, D_x U^N(t, x_j, m_{\mathbf{x}}^{N,i})) \\ &= \sum_{j \neq i} D_{x_j} u^{N,i}(t, \mathbf{x}) \cdot D_p H(x_j, D_{x_j} u^{N,j}(t, \mathbf{x})) + r_1^{N,i}(t, \mathbf{x}) = 0, \end{aligned}$$

where, by (31) and (32),

$$\|r_1^{N,i}\|_{\infty} \leq \frac{C}{N} \|D_m U^N\|_{\infty} \|D_{m,x} U^N\|_{\infty}.$$

On the other hand, by (29), we have

$$\begin{aligned} & \sum_{j=1}^N \Delta_{x_j} u^{N,i} - \Delta_x U^N(t, x_i, m_{\mathbf{x}}^{N,i}) - \frac{1}{N-1} \sum_{j \neq i} \operatorname{div}_y D_m U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) \\ &= \frac{1}{(N-1)^2} \sum_{j \neq i} \operatorname{tr}(D_{mm}^2 U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j, x_j)) =: r_2^{N,i}(t, \mathbf{x}), \end{aligned}$$

where

$$\|r_2^{N,i}\|_{\infty} \leq \frac{C}{N} \|D_{mm}^2 U^N\|_{\infty}.$$

Therefore

$$\begin{aligned} & -\partial_t u^{N,i}(t, \mathbf{x}) - \sum_j \Delta_{x_j} u^{N,i}(t, \mathbf{x}) + H(x_i, D_{x_i} u^{N,i}(t, \mathbf{x})) \\ &+ \sum_{j \neq i} D_{x_j} u^{N,i}(t, \mathbf{x}) \cdot D_p H(x_j, D_{x_j} u^{N,j}(t, \mathbf{x})) = F(x_i, m_{\mathbf{x}}^{N,i}) + r_1^{N,i} + r_2^{N,i}, \end{aligned}$$

which shows the result.  $\square$

### 3.2 Estimates between $v^{N,i}$ and $U^N$

Let us fix  $t_0 \in [0, T]$ . Let  $(Z_i)_{i \in \{1, \dots, N\}}$  be an i.i.d family of  $N$  random variables. We set  $\mathbf{Z} = (Z_i)_{i \in \{1, \dots, N\}}$ . Let also  $((B_t^i)_{t \in [0, T]})_{i \in \{1, \dots, N\}}$  be a family of  $N$  independent  $d$ -dimensional Brownian Motions which is also independent of  $(Z_i)_{i \in \{1, \dots, N\}}$ . We consider the systems of SDEs with variables  $(\mathbf{X}_t = (X_{i,t})_{i \in \{1, \dots, N\}})_{t \in [0, T]}$  and  $(\mathbf{Y}_t = (Y_{i,t})_{i \in \{1, \dots, N\}})_{t \in [0, T]}$ :

$$\begin{cases} dX_{i,t} = -D_p H(X_{i,t}, D_{x_i} u^{N,i}(t, \mathbf{X}_t)) dt + \sqrt{2} dB_t^i & t \in [t_0, T], \\ X_{i,t_0} = Z_i, \end{cases} \quad (33)$$

and

$$\begin{cases} dY_{i,t} = -D_p H(Y_{i,t}, D_{x_i} v^{N,i}(t, \mathbf{Y}_t)) dt + \sqrt{2} dB_t^i & t \in [t_0, T], \\ Y_{i,t_0} = Z_i. \end{cases} \quad (34)$$

Note that, by the symmetry properties of the  $(u^{N,i})_{i \in \{1, \dots, N\}}$  and of the  $(v^{N,i})_{i \in \{1, \dots, N\}}$ , the processes  $(X_{i,t}, Y_{i,t})_{t \in [t_0, T]}_{i \in \{1, \dots, N\}}$  are exchangeable.

Let us finally introduce notations for the error terms:

$$\alpha^N = \sup_i \|D_{x_i} u^{N,i}\|_\infty, \quad \beta^N = \sup_i \sup_{j \neq i} \|D_{x_j} u^{N,i}\|_\infty \quad \text{and} \quad r^N := \sup_i \|r^{N,i}\|_\infty, \quad (35)$$

where  $r^{N,i}$  is the error term in Proposition 3.1. In the same way, we set

$$\hat{\alpha}^N = \sup_i \|D_{x_i, x_i}^2 u^{N,i}\|_\infty, \quad \hat{\beta}^N = \sup_i \sup_{j \neq i} \|D_{x_i, x_j}^2 u^{N,i}\|_\infty.$$

Finally,

$$\theta^N := (1 + (\alpha^N)^2 + (N\beta^N)^2), \quad \hat{\theta}^N := (1 + \hat{\alpha}^N + N\hat{\beta}^N). \quad (36)$$

Note that, by symmetry, the  $\sup_i$  in the above expressions is actually superfluous. Theorem 2.2 implies that  $r^N$  and  $\theta^N$  are of the following order:

$$r^N \leq \frac{CK_{N,\alpha}^{12}}{N}, \quad \theta^N \leq CK_{N,\alpha}^6, \quad \hat{\theta}^N \leq CK_{N,\alpha}^3.$$

**Theorem 3.2.** *We have, for any  $i \in \{1, \dots, N\}$ ,*

$$\|u^{N,i} - v^{N,i}\|_\infty \leq Cr^N (\theta^N)^{1/2} \exp(C\theta^N), \quad (37)$$

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |Y_{i,t} - X_{i,t}| \right] \leq Cr^N (\theta^N \hat{\theta}^N)^{1/2} \exp(C(\theta^N + \hat{\theta}^N)) \quad (38)$$

and

$$\mathbb{E} \left[ \int_{t_0}^T |D_{x_i} v^{N,i}(t, \mathbf{Y}_t) - D_{x_i} u^{N,i}(t, \mathbf{Y}_t)|^2 dt \right] \leq C(r^N)^2 \theta^N \exp(C\theta^N), \quad (39)$$

where  $C$  is a (deterministic) constant that does not depend on  $t_0$ ,  $m_0$  and  $N$ .

*Proof.* We follow closely the proof in [2] and so indicate only the main changes. We will use the following notations: for  $t \in [t_0, T]$ ,

$$\begin{aligned} U_t^{N,i} &= u^{N,i}(t, \mathbf{Y}_t), & V_t^{N,i} &= v^{N,i}(t, \mathbf{Y}_t), \\ DU_t^{N,i,j} &= D_{x_j} u^{N,i}(t, \mathbf{Y}_t), & DV_t^{N,i,j} &= D_{x_j} v^{N,i}(t, \mathbf{Y}_t). \end{aligned}$$

As the  $(v^{N,i})_{i \in \{1, \dots, N\}}$  solve equation (28), we have by Itô's formula that, for any  $i \in \{1, \dots, N\}$ ,

$$\begin{aligned} dV_t^{N,i} &= \left[ H(Y_{i,t}, D_{x_i} v^{N,i}(t, \mathbf{Y}_t)) - D_{x_i} v^{N,i}(t, \mathbf{Y}_t) \cdot D_p H(Y_{i,t}, D_{x_i} v^{N,i}(t, \mathbf{Y}_t)) \right. \\ &\quad \left. - F^N(Y_{i,t}, m_{\mathbf{Y}_t}^{N,i}) \right] dt \\ &\quad + \sqrt{2} \sum_j D_{x_j} v^{N,i}(t, \mathbf{Y}_t) \cdot dB_t^j. \end{aligned} \quad (40)$$

Similarly, as  $(u^{N,i})_{i \in \{1, \dots, N\}}$  satisfies (30), we have:

$$\begin{aligned} dU_t^{N,i} &= \left[ H(Y_{i,t}, D_{x_i} u^{N,i}(t, \mathbf{Y}_t)) - D_{x_i} u^{N,i}(t, \mathbf{Y}_t) \cdot D_p H(Y_{i,t}, D_{x_i} u^{N,i}(t, \mathbf{Y}_t)) \right. \\ &\quad \left. - F^N(Y_{i,t}, m_{\mathbf{Y}_t}^{N,i}) - r^{N,i}(t, \mathbf{Y}_t) \right] dt \\ &\quad - \sum_j D_{x_j} u^{N,i}(t, \mathbf{Y}_t) \cdot \left( D_p H(Y_{j,t}, D_{x_j} v^{N,j}(t, \mathbf{Y}_t)) - D_p H(Y_{j,t}, D_{x_j} u^{N,j}(t, \mathbf{Y}_t)) \right) dt \\ &\quad + \sqrt{2} \sum_j D_{x_j} u^{N,i}(t, \mathbf{Y}_t) \cdot dB_t^j. \end{aligned} \quad (41)$$



Computing the difference between (40) and (41), taking the square and applying Itô's formula again, we obtain:

$$\begin{aligned}
& d[U_t^{N,i} - V_t^{N,i}]^2 \\
&= \left[ 2(U_t^{N,i} - V_t^{N,i}) \cdot \left( H(Y_{i,t}, DU_t^{N,i,i}) - H(Y_{i,t}, DV_t^{N,i,i}) \right) \right. \\
&\quad - 2(U_t^{N,i} - V_t^{N,i}) \cdot \left( DU_t^{N,i,i} \cdot [D_p H(Y_{i,t}, DU_t^{N,i,i}) - D_p H(Y_{i,t}, DV_t^{N,i,i})] \right) \\
&\quad - 2(U_t^{N,i} - V_t^{N,i}) \cdot \left( [DU_t^{N,i,i} - DV_t^{N,i,i}] \cdot D_p H(Y_{i,t}, DV_t^{N,i,i}) \right) \\
&\quad \left. - 2(U_t^{N,i} - V_t^{N,i}) r^{N,i}(t, \mathbf{Y}_t) \right] dt \\
&- 2(U_t^{N,i} - V_t^{N,i}) \sum_j DU_t^{N,i,j} \cdot \left( D_p H(Y_{j,t}, DV_t^{N,j,j}) - D_p H(Y_{j,t}, DU_t^{N,j,j}) \right) dt \\
&+ \left[ 2 \sum_j |DU_t^{N,i,j} - DV_t^{N,i,j}|^2 + \sqrt{2}(U_t^{N,i} - V_t^{N,i}) \sum_j \left[ (DU_t^{N,i,j} - DV_t^{N,i,j}) \cdot dB_t^j \right] \right].
\end{aligned} \tag{42}$$

Recall now that  $H$  and  $D_p H$  are Lipschitz continuous in the variable  $p$ . Recall also the notation  $\alpha^N$ ,  $\beta^N$  and  $r^N$  in (35). Integrating from  $t$  to  $T$  in the above formula and taking the conditional expectation given  $\mathbf{Z}$  (with the shorten notation  $\mathbb{E}^{\mathbf{Z}}[\cdot] = \mathbb{E}[\cdot|\mathbf{Z}]$ ), we deduce:

$$\begin{aligned}
& \mathbb{E}^{\mathbf{Z}}[|U_t^{N,i} - V_t^{N,i}|^2] + 2 \sum_j \mathbb{E}^{\mathbf{Z}} \left[ \int_t^T |DU_s^{N,i,j} - DV_s^{N,i,j}|^2 ds \right] \\
&\leq \mathbb{E}^{\mathbf{Z}}[|U_T^{N,i} - V_T^{N,i}|^2] + r^N \int_t^T \mathbb{E}^{\mathbf{Z}}[|U_s^{N,i} - V_s^{N,i}|] ds \\
&\quad + C(1 + \alpha^N) \int_t^T \mathbb{E}^{\mathbf{Z}}[|U_s^{N,i} - V_s^{N,i}| \cdot |DU_s^{N,i,i} - DV_s^{N,i,i}|] ds \\
&\quad + C\beta^N \sum_{j \neq i} \int_t^T \mathbb{E}^{\mathbf{Z}}[|U_s^{N,i} - V_s^{N,i}| \cdot |DU_s^{N,j,j} - DV_s^{N,j,j}|] ds.
\end{aligned} \tag{43}$$

Recall that  $U_T^{N,i} = V_T^{N,i} = G(Y_{i,T})$ . By Young's inequality, we get

$$\begin{aligned}
& \mathbb{E}^{\mathbf{Z}}[|U_t^{N,i} - V_t^{N,i}|^2] + \mathbb{E}^{\mathbf{Z}} \left[ \int_t^T |DU_s^{N,i,i} - DV_s^{N,i,i}|^2 ds \right] \\
&\leq C(r^N)^2 + C(1 + (\alpha^N)^2 + (N\beta^N)^2) \int_t^T \mathbb{E}^{\mathbf{Z}}[|U_s^{N,i} - V_s^{N,i}|^2] ds \\
&\quad + \frac{1}{2N} \sum_j \mathbb{E}^{\mathbf{Z}} \left[ \int_t^T |DU_s^{N,j,j} - DV_s^{N,j,j}|^2 ds \right].
\end{aligned} \tag{44}$$

Summing over  $i$  we obtain:

$$\begin{aligned}
& \sum_i \mathbb{E}^{\mathbf{Z}}[|U_t^{N,i} - V_t^{N,i}|^2] + \frac{1}{2} \sum_i \mathbb{E}^{\mathbf{Z}} \left[ \int_t^T |DU_s^{N,i,i} - DV_s^{N,i,i}|^2 ds \right] \\
&\leq CN(r^N)^2 + C(1 + (\alpha^N)^2 + (N\beta^N)^2) \int_t^T \sum_i \mathbb{E}^{\mathbf{Z}}[|U_s^{N,i} - V_s^{N,i}|^2] ds.
\end{aligned} \tag{45}$$

By Gronwall's Lemma, this leads to:

$$\sup_{t_0 \leq t \leq T} \left[ \sum_i \mathbb{E}^{\mathbf{Z}} [ |U_t^{N,i} - V_t^{N,i}|^2 ] \right] \leq CN(r^N)^2 \exp(C\theta^N), \quad (46)$$

where  $\theta^N$  is given in (36). Plugging (46) into (45), we deduce that

$$\sum_j \mathbb{E}^{\mathbf{Z}} \left[ \int_{t_0}^T |DU_s^{N,j,j} - DV_s^{N,j,j}|^2 ds \right] \leq CN(r^N)^2 \theta^N \exp(C\theta^N).$$

Inserting this bound in the right-hand side of (44) and applying Gronwall's lemma once again, we finally end up with:

$$\begin{aligned} & \sup_{t \in [t_0, T]} \mathbb{E}^{\mathbf{Z}} [ |U_t^{N,i} - V_t^{N,i}|^2 ] + \mathbb{E}^{\mathbf{Z}} \left[ \int_{t_0}^T |DU_s^{N,i,i} - DV_s^{N,i,i}|^2 ds \right] \\ & \leq C(r^N)^2 \theta^N \exp(C\theta^N). \end{aligned} \quad (47)$$

This gives (39). By the definition of  $U^{N,i}$  and  $V^{N,i}$  this implies that

$$|u^{N,i}(t_0, \mathbf{Z}) - v^{N,i}(t_0, \mathbf{Z})| \leq Cr^N (\theta^N)^{1/2} \exp(C\theta^N) \quad \text{a.e.}$$

Then choosing  $\mathbf{Z}_i$  with a uniform law on  $\mathbb{T}^d$  implies, by continuity of  $u^{N,i}$  and  $v^{N,i}$  that

$$|u^{N,i}(t_0, \mathbf{x}) - v^{N,i}(t_0, \mathbf{x})| \leq Cr^N (\theta^N)^{1/2} \exp(C\theta^N) \quad \forall \mathbf{x} \in (\mathbb{T}^d)^N,$$

which shows (37).

We now estimate the difference  $X_{i,t} - Y_{i,t}$ , for  $t \in [t_0, T]$  and  $i \in \{1, \dots, N\}$ . In view of the equation satisfied by the processes  $(X_{i,t})_{t \in [t_0, T]}$  and by  $(Y_{i,t})_{t \in [t_0, T]}$ , we have

$$\begin{aligned} |X_{i,t} - Y_{i,t}| & \leq \int_{t_0}^t |D_p H(X_{i,s}, D_{x_i} u^{N,i}(s, \mathbf{X}_s)) - D_p H(Y_{i,s}, D_{x_i} v^{N,i}(s, \mathbf{Y}_s))| ds \\ & \leq C \int_{t_0}^t (1 + \|D_{x_i, x_i}^2 u^{N,i}\|_{\infty}) |X_{i,s} - Y_{i,s}| + \sum_{j \neq i} \|D_{x_i, x_j}^2 u^{N,i}\|_{\infty} |X_{j,s} - Y_{j,s}| ds \\ & \quad + \int_{t_0}^t |D_p H(Y_{i,s}, D_{x_i} u^{N,i}(s, \mathbf{Y}_s)) - D_p H(Y_{i,s}, D_{x_i} v^{N,i}(s, \mathbf{Y}_s))| ds \\ & \leq C \hat{\alpha}^N \int_{t_0}^t |X_{i,s} - Y_{i,s}| ds + C \hat{\beta}^N \sum_{j \neq i} \int_{t_0}^t |X_{j,s} - Y_{j,s}| ds \\ & \quad + C \int_{t_0}^T |DU_s^{N,i,i} - DV_s^{N,i,i}| ds. \end{aligned} \quad (48)$$

Computing as before the inequality satisfied by the sum  $\sum_i |X_{i,t} - Y_{i,t}|$ , using the exchangeability of the  $((X^{N,i}, Y^{N,i}))$  and (47), we obtain (38) thanks to Gronwall inequality.  $\square$

Following exactly the same argument as for Theorem 2.13 in [2], we deduce:

**Corollary 3.3.** Fix  $N \geq 1$  and  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$ . For any  $i \in \{1, \dots, N\}$  and  $x \in \mathbb{T}^d$ , let us set

$$w^{N,i}(t_0, x_i, m_0) := \int_{(\mathbb{T}^d)^{(N-1)}} v^{N,i}(t_0, \mathbf{x}) \prod_{j \neq i} m_0(dx_j) \quad \text{where } \mathbf{x} = (x_1, \dots, x_N).$$

Then

$$\|w^{N,i}(t_0, \cdot, m_0) - U^N(t_0, \cdot, m_0)\|_\infty \leq \begin{cases} Cr^N(\theta^N)^{1/2} \exp(C\theta^N) + C\|D_m U^N\|_\infty N^{-1/d} & \text{if } d \geq 3, \\ Cr^N(\theta^N)^{1/2} \exp(C\theta^N) + C\|D_m U^N\|_\infty N^{-1/2} \log(N) & \text{if } d = 2, \\ Cr^N(\theta^N)^{1/2} \exp(C\theta^N) + C\|D_m U^N\|_\infty N^{-1/2} & \text{if } d = 1. \end{cases}$$

where the constant  $C$  does not depend on  $t_0$ ,  $m_0$ ,  $i$  and  $N$  and where  $\theta^N$  and  $\hat{\theta}^N$  are defined before Theorem 3.2.

*Proof.* We use the Lipschitz continuity of  $U^N$  and a result by Fournier and Guillin [8] to deduce that, for  $d \geq 3$  and for any  $x_i \in \mathbb{T}^d$ ,

$$\begin{aligned} & \int_{\mathbb{T}^{d(N-1)}} |u^{N,i}(t, \mathbf{x}) - U^N(t, x_i, m_0)| \prod_{j \neq i} m_0(dx_j) \\ &= \int_{\mathbb{T}^{d(N-1)}} |U^N(t, x_i, m_{\mathbf{x}}^{N,i}) - U^N(t, x_i, m_0)| \prod_{j \neq i} m_0(dx_j) \\ &\leq \|D_m U^N\|_\infty \int_{\mathbb{T}^{d(N-1)}} \mathbf{d}_1(m_{\mathbf{x}}^{N,i}, m_0) \prod_{j \neq i} m_0(dx_j) \leq C\|D_m U^N\|_\infty N^{-1/d}. \end{aligned}$$

If  $d = 1$  (respectively  $d = 2$ ), the  $N^{-1/d}$  has to be replaced by  $N^{-1/2}$  (respectively  $N^{-1/2} \log(N)$ ). Combining Theorem 3.2 with the above inequality, we obtain therefore, for  $d \geq 3$ ,

$$\begin{aligned} & |w^{N,i}(t_0, x_i, m_0) - U^N(t_0, x_i, m_0)| \\ &= \left| \int_{\mathbb{T}^{d(N-1)}} v^{N,i}(t_0, (x_j)) \prod_{j \neq i} m_0(dx_j) - U^N(t_0, x_i, m_0) \right| \\ &= \|v^{N,i} - u^{N,i}\|_\infty + \int_{\mathbb{T}^{dN}} |u^{N,i}(t_0, \mathbf{x}) - U^N(t_0, x_i, m_0)| \prod_{j \neq i} m_0(dx_j) \\ &\leq Cr^N(\theta^N)^{1/2} \exp(C\theta^N) + C\|D_m U^N\|_\infty N^{-1/d}. \end{aligned}$$

As above, the last term is  $N^{-1/2}$  if  $d = 1$  and  $N^{-1/2} \log(N)$  if  $d = 2$ .  $\square$

### 3.3 Putting the estimates together

Here we fix a initial condition  $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$ , where  $m_0$  is a smooth, positive density. Let  $v^{N,i}$  be the solution of the Nash system (28). Following the averaging procedure of Corollary 3.3, we set

$$w^{N,i}(t_0, x, m_0) := \int_{\mathbb{T}^d} \dots \int_{\mathbb{T}^d} v^{N,i}(t_0, \mathbf{x}) \prod_{j \neq i} m_0(dx_j) \quad \text{where } \mathbf{x} = (x_1, \dots, x_N).$$

Let  $u$  be the solution to the MFG system (16).

**Theorem 3.4.** *We have*

$$\|w^{N,i}(t_0, \cdot, m_0) - u(t_0, \cdot)\|_\infty \leq \begin{cases} CN^{-\frac{1}{d}} K_{N,\alpha}^{15} \exp(CK_{N,\alpha}^6) + C \left(k_N^{R,\alpha}\right)^{\frac{2}{d+2}} & \text{if } d \geq 3 \\ CN^{-\frac{1}{2}} \ln(N) K_{N,\alpha}^{15} \exp(CK_{N,\alpha}^6) + C \left(k_N^{R,\alpha}\right)^{\frac{1}{2}} & \text{if } d = 2 \\ CN^{-\frac{1}{2}} K_{N,\alpha}^{15} \exp(CK_{N,\alpha}^6) + C \left(k_N^{R,\alpha}\right)^{\frac{2}{3}} & \text{if } d = 1. \end{cases} \quad (49)$$

where  $R$  and  $\alpha$  do not depend on  $N$  (but depends on  $m_0$ ).

In particular, if  $K_{N,\alpha} = O((\ln(N))^\theta)$  for some  $\theta \in (0, 1/(6d))$ , then  $w^{N,i}(t_0, \cdot, m_0)$  converges uniformly to  $u(t_0, \cdot)$ .

*Proof.* For  $N \in \mathbb{N}$ , let  $u^N$  be the solution to the MFG system (15). Recalling Proposition 2.3, we have

$$\|u^N(t_0, \cdot) - u(t_0, \cdot)\|_\infty \leq C \left(k_N^{R,\alpha}\right)^{\frac{2}{d+2}},$$

where  $R$  is a bound on the  $C^\alpha$  norm of the  $m^N$  (Proposition 2.3). As

$$U^N(t_0, x, m_0) = u^N(t_0, x),$$

Corollary 3.3 implies that (for  $N \geq 3$ ),

$$\begin{aligned} \|w^{N,i}(t_0, \cdot, m_0) - u^N(t_0, \cdot)\|_\infty &\leq Cr^N (\theta^N)^{1/2} \exp(C\theta^N) + C \|D_m U^N\|_\infty N^{-1/d} \\ &\leq CN^{-\frac{1}{d}} K_{N,\alpha}^{15} \exp(CK_{N,\alpha}^6). \end{aligned}$$

□

We now consider a particular case:

**Corollary 3.5.** *Assume that*

$$F^N(x, m) = F(\cdot, \xi^{\epsilon_N} * m(\cdot)) * \xi^\epsilon(x),$$

where  $\xi^\epsilon$  are as in the example in Proposition 1.2. If one chooses  $\epsilon_N = \ln(N)^{-\beta}$ , with  $\beta \in (0, (6d(2d+15))^{-1})$ , the convergence in (49) is of order  $A(\ln(N))^{-1/B}$  for some constants  $A, B$ .

### 3.4 Convergence of the optimal solutions

We complete the paper by a discussion on the convergence of the optimal solutions and a propagation of chaos.

Let us explain the problem. Let  $m_0 \in \mathcal{P}(\mathbb{T}^d)$  with a smooth, positive density. Let  $(v^{N,i})$  be the solution to the Nash system (28) and, for  $t_0 \in [0, T]$ ,  $m_0 \in \mathcal{P}(\mathbb{T}^d)$ ,  $(u, m)$  be the solution to the MFG system (16) starting at time  $t_0$  from  $m_0$ . Let  $(Z_i)$  be an i.i.d family of  $N$  random variables of law  $m_0$ . We set  $\mathbf{Z} = (Z_1, \dots, Z_N)$ . Let also  $((B_t^i)_{t \in [0, T]})_{i \in \{1, \dots, N\}}$  be a family of  $N$  independent Brownian motions which is also independent of  $(Z_i)$ . We consider the optimal trajectories  $(\mathbf{Y}_t = (Y_{1,t}, \dots, Y_{N,t}))_{t \in [t_0, T]}$  for the  $N$ -player game:

$$\begin{cases} dY_{i,t} = -D_p H(Y_{i,t}, D_{x_i} v^{N,i}(t, \mathbf{Y}_t)) dt + \sqrt{2} dB_t^i, & t \in [t_0, T], \\ Y_{i,t_0} = Z_i \end{cases}$$

and the optimal solution  $(\tilde{\mathbf{X}}_t = (\tilde{X}_{1,t}, \dots, \tilde{X}_{N,t}))_{t \in [t_0, T]}$  to the limit MFG system:

$$\begin{cases} d\tilde{X}_{i,t} = -D_p H(\tilde{X}_{i,t}, D_x u(t, \tilde{X}_{i,t})) dt + \sqrt{2} dB_t^i, & t \in [t_0, T], \\ \tilde{X}_{i,t_0} = Z_i. \end{cases}$$

The next result provides an estimate of the distance between the solutions. To fix the ideas, we work in dimension  $d \geq 3$ .

**Theorem 3.6.** *Under our standing assumptions, for any  $N \geq 1$  and any  $i \in \{1, \dots, N\}$ , we have*

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |Y_{i,t} - \tilde{X}_{i,t}| \right] \leq C \left[ K_{N,\alpha}^{25/2} N^{-1/d} \exp\{CK_{N,\alpha}^6\} + \left(k_N^{R,\alpha}\right)^{\frac{2}{d+2}} \right],$$

where the constant  $C > 0$  is independent of  $t_0$  and  $N$ , but depends on  $m_0$ .

In particular, if  $K_{N,\alpha} = O((\ln(N))^\theta)$  for some  $\theta \in (0, 1/(6d))$ , then the optimal trajectories  $(Y_{i,t})$  converge to the  $(\tilde{X}_{i,t})$  and become asymptotically i.i.d.

In order to illustrate the result, let us come back to our main example:

**Proposition 3.7.** *Assume that  $F^N$  are of the form (12). Then, for  $\epsilon_N = \ln(N)^{-\beta}$  for  $\beta \in (0, [6d(2d+15)]^{-1})$ , we have*

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |Y_{i,t} - \tilde{X}_{i,t}| \right] \leq A(\ln(N))^{-1/B},$$

for some constants  $A, B > 0$ .

*Proof of Theorem 3.6.* Let  $U^N$  be a solution to the master equation (14) and set  $u^{N,i}(t, \mathbf{x}) = U^N(t, x_i, m_{\mathbf{x}}^{N,i})$ . Let  $(X_{i,t})$  be the solution to

$$\begin{cases} dX_{i,t} = -D_p H(X_{i,t}, D_x u^{N,i}(t, \mathbf{X}_t)) dt + \sqrt{2} dB_t^i & t \in [t_0, T], \\ X_{i,t_0} = Z_i, \end{cases}$$

and  $(\hat{X}_{i,t})$  be the solution to

$$\begin{cases} d\hat{X}_{i,t} = -D_p H(\hat{X}_{i,t}, D_x u^N(t, \hat{X}_{i,t})) dt + \sqrt{2} dB_t^i & t \in [t_0, T], \\ \hat{X}_{i,t_0} = Z_i, \end{cases}$$

where  $(u^N, m^N)$  is the solution of the MFG system (15). Note that the  $(\hat{X}_{i,t})$  are i.i.d. with law  $(m^N(t))$ . As, for any  $t \in [t_0, T]$ ,

$$u^N(t, \cdot) = U^N(t, \cdot, m^N(t)),$$

the  $(\hat{X}_{i,t})$  are also solution to

$$\begin{cases} d\hat{X}_{i,t} = -D_p H(\hat{X}_{i,t}, D_x U^N(t, \hat{X}_{i,t}, m^N(t))) dt + \sqrt{2} dB_t^i & t \in [t_0, T], \\ \hat{X}_{i,t_0} = Z_i. \end{cases}$$

The main step of the proof is the following claim:

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |X_{i,t} - \hat{X}_{i,t}| \right] \leq CK_{N,\alpha} N^{-1/d} \exp\{CK_{N,\alpha}\}. \quad (50)$$

Let us fix  $i \in \{1, \dots, N\}$  and let

$$\rho(t) = \mathbb{E} \left[ \sup_{s \in [t_0, t]} |X_{i,s} - \hat{X}_{i,s}| \right].$$

Then, for any  $t_0 \leq s \leq t \leq T$ , we have

$$\begin{aligned} |X_{i,s} - \hat{X}_{i,s}| &\leq \int_{t_0}^s \left| -D_p H(X_{i,r}, D_x U^N(r, X_{i,r}, m_{\mathbf{X}_r}^{N,i})) + D_p H(\hat{X}_{i,r}, D_x U^N(r, \hat{X}_{i,r}, m^N(r))) \right| dr \\ &\leq C \int_{t_0}^s \left( |X_{i,r} - \hat{X}_{i,r}| + |D_x U^N(r, X_{i,r}, m_{\mathbf{X}_r}^{N,i}) - D_x U^N(r, \hat{X}_{i,r}, m^N(r))| \right) dr, \end{aligned}$$

where we have used the fact that  $D_p H$  is globally Lipschitz continuous. As the map  $(x, m) \rightarrow D_x U^N(r, x, m)$  is Lipschitz continuous with constant  $CK_{N,\alpha}$  (Theorem 2.2), we get

$$|X_{i,s} - \hat{X}_{i,s}| \leq CK_{N,\alpha} \int_{t_0}^s \left( |X_{i,r} - \hat{X}_{i,r}| + \mathbf{d}_1(m_{\mathbf{X}_r}^{N,i}, m_{\hat{\mathbf{X}}_r}^{N,i}) + \mathbf{d}_1(m_{\hat{\mathbf{X}}_r}^{N,i}, m^N(r)) \right) dr,$$

where

$$\mathbf{d}_1(m_{\mathbf{X}_r}^{N,i}, m_{\hat{\mathbf{X}}_r}^{N,i}) \leq \frac{1}{N-1} \sum_{j \neq i} |X_{j,r} - \hat{X}_{j,r}|. \quad (51)$$

Hence

$$|X_{i,s} - \hat{X}_{i,s}| \leq CK_{N,\alpha} \int_{t_0}^s \left( |X_{i,r} - \hat{X}_{i,r}| + \frac{1}{N-1} \sum_{j \neq i} |X_{j,r} - \hat{X}_{j,r}| + \mathbf{d}_1(m_{\hat{\mathbf{X}}_r}^{N,i}, m^N(r)) \right) dr. \quad (52)$$

As the  $(\hat{X}_{i,t})$  are i.i.d. and  $d \geq 3$ , we have from [8] that

$$\mathbb{E} \left[ \mathbf{d}_1(m_{\hat{\mathbf{X}}_r}^{N,i}, m^N(r)) \right] \leq CN^{-1/d}.$$

So, taking the supremum over  $s \in [t_0, t]$  in (52) and then the expectation, gives, since the random variables  $(X_{j,r} - \hat{X}_{j,r})_{j \in \{1, \dots, N\}}$  have the same law:

$$\rho(t) = \mathbb{E} \left[ \sup_{s \in [t_0, t]} |X_{i,s} - \hat{X}_{i,s}| \right] \leq CK_{N,\alpha} \int_{t_0}^t \rho(s) ds + CK_{N,\alpha} N^{-1/d}.$$

Then Gronwall inequality gives (50).

We now complete the proof by recalling that, from Theorem 3.2,

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |Y_{i,t} - X_{i,t}| \right] \leq Cr^N (\theta^N \hat{\theta}^N)^{1/2} \exp(C(\theta^N + \hat{\theta}^N)),$$

where

$$r^N \leq \frac{CK_{N,\alpha}^{12}}{N}, \quad \theta^N \leq CK_{N,\alpha}^6, \quad \hat{\theta}^N \leq CK_{N,\alpha}^3.$$

On the other hand, Corollary 2.4 states that

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |\tilde{X}_t - \hat{X}_t| \right] \leq C \left( k_N^{R,\alpha} \right)^{\frac{2}{d+2}},$$

Therefore

$$\mathbb{E} \left[ \sup_{t \in [t_0, T]} |Y_{i,t} - \tilde{X}_{i,t}| \right] \leq C \left[ K_{N,\alpha} N^{-1/d} \exp\{CK_{N,\alpha}\} + K_{N,\alpha}^{33/2} N^{-1} \exp(CK_{N,\alpha}^6) + \left( k_N^{R,\alpha} \right)^{\frac{2}{d+2}} \right].$$

□

## 4 Appendix

In the appendix, we state an estimate for equations of the form:

$$\begin{cases} \partial_t w - \Delta w + V(t, x) \cdot Dw = f & \text{in } [0, T] \times \mathbb{T}^d, \\ w(0, x) = w_0(x) & \text{in } \mathbb{T}^d, \end{cases} \quad (53)$$

where  $V$  is a fixed bounded vector field.

**Proposition 4.1.** *If  $w$  is a solution to the above equation with  $w_0 \in C^{1+\alpha}$ , then*

$$\sup_{t \in [0, T]} \|w(t)\|_{1+\alpha} \leq C [\|w_0\|_{1+\alpha} + \|f\|_\infty],$$

where  $C$  depends on  $\|V\|_\infty$ ,  $T$ ,  $\alpha$  and  $d$  only.

This kind of estimate is standard in the literature: for instance Theorem IV.9.1 of [17] (and its Corollary) states that  $Dw$  is bounded in  $C^{\beta/2, \beta}$  for any  $\beta \in (0, 1)$ . However the bound might depend on the vector field  $V$  and not only on its norm. We only check this is not the case.

*Proof.* Let us first check that the result holds for an homogenous initial datum. More precisely, we prove in a first step that, if  $w$  solve (53) with  $w(0, \cdot) = 0$ , then  $\|Dw\|_\infty \leq C\|f\|_\infty$ , where the constant  $C$  depends on  $\|V\|_\infty$ ,  $T$  and  $d$  only. For this we argue by contradiction and assume for a while that there exists  $V_n$  and  $f_n$ , bounded in  $L^\infty$ , and  $w_n$  such that

$$\begin{cases} \partial_t w_n - \Delta w_n + V_n \cdot Dw_n = f_n & \text{in } [0, T] \times \mathbb{T}^d \\ w_n(0, x) = 0 & \text{in } \mathbb{T}^d. \end{cases}$$

with  $k_n := \|Dw_n\|_\infty \rightarrow +\infty$ . We set  $\tilde{w}_n := w_n/k_n$ ,  $\tilde{f}_n := f_n/k_n$ . Then  $\tilde{w}_n$  solves the heat equation with a right-hand side  $\tilde{f}_n - V_n \cdot D\tilde{w}_n$  which is bounded in  $L^\infty$ . By standard estimates on the heat potential (see (3.2) of Chapter 3 in [17]),  $\partial_t \tilde{w}_n$  and  $D^2 \tilde{w}_n$  are bounded in  $L^p$  for any  $p$  independently of  $n$ . Then a Sobolev type inequality (Lemma II.3.3 in [17]) implies that  $D\tilde{w}_n$  is bounded in  $C^{\beta/2, \beta}$  independently of  $n$  for any  $\beta \in (0, 1)$ . On the other hand,  $(\tilde{f}_n)$  tends to 0 in  $L^2$  and, by standard energy estimates,  $(D\tilde{w}_n)$  tends to 0 in  $L^2$ . This is in contradiction with the fact that  $\|D\tilde{w}_n\|_\infty = 1$  and that  $D\tilde{w}_n$  is bounded in  $C^{\beta/2, \beta}$ . So we have proved that there exists a constant  $C$ , depending on  $\|V\|_\infty$ ,  $d$  and  $T$  only, such that the solution to (53) with  $w(0, \cdot) = 0$  satisfies  $\|Dw\|_\infty \leq C\|f\|_\infty$ . Using the same argument on the the heat potential as above yields to

$$\|Dw\|_{C^{\beta/2, \beta}} \leq C_\beta \|f - V \cdot Dw\|_\infty \leq C_\beta \|f\|_\infty,$$

where  $C_\beta$  depends on  $\|V\|$ ,  $d$ ,  $T$  and  $\beta$  only.

We now remove the assumption that  $w_0 = 0$ . We rewrite  $w$  as the sum  $w = w_1 + w_2$  where  $w_1$  solves the heat equation with initial condition  $w_0$  and  $w_2$  solves equation (53) with right-hand side  $f - V \cdot Dw_1$  and initial condition  $w_2(0, \cdot) = 0$ . By maximum principle, we have

$$\sup_{t \in [0, T]} \|Dw_1(t)\|_\alpha \leq C \|Dw_0\|_\alpha.$$

By the first step of the proof, we also have, for any  $\beta \in (0, 1)$ ,

$$\|Dw_2\|_{C^{\beta/2, \beta}} \leq C_\beta \|f - V \cdot Dw_1\|_\infty \leq C_\beta (\|f\|_\infty + \|Dw_0\|_\alpha).$$

Choosing  $\beta = \alpha$  then gives the result.  $\square$

## References

- [1] Rainer Buckdahn, Juan Li, Shige Peng, and Catherine Rainer. Mean-field stochastic differential equations and associated pdes. *arXiv preprint arXiv:1407.1215*, 2014.
- [2] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. The master equation and the convergence problem in mean field games. *arXiv preprint arXiv:1509.02505*, 2015.
- [3] R Carmona and F Delarue. Probabilistic theory of mean field games,(expected), 2016.
- [4] René Carmona and François Delarue. Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization*, 51(4):2705–2734, 2013.
- [5] Jean-François Chassagneux, Dan Crisan, and François Delarue. Classical solutions to the master equation for large population equilibria. *arXiv preprint arXiv:1411.3009*, 2014.
- [6] Ermal Feleqi. The derivation of ergodic mean field game equations for several populations of players. *Dynamic Games and Applications*, 3(4):523–536, 2013.
- [7] Markus Fischer. On the connection between symmetric  $n$ -player games and mean field games. *arXiv preprint arXiv:1405.1345*, 2014.
- [8] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015.
- [9] Wilfrid Gangbo and Andrzej Świech. Existence of a solution to an equation arising from the theory of mean field games. *Journal of Differential Equations*, 259(11):6573–6643, 2015.
- [10] D Gomes, E Pimentel, and Vardan Voskanyan. Regularity theory for mean-field game systems. *SpringerBriefs in mathematics*, 2016.
- [11] Minyi Huang, Peter E Caines, and Roland P Malhamé. An invariance principle in large population stochastic dynamic games. *Journal of Systems Science and Complexity*, 20(2):162–172, 2007.
- [12] Minyi Huang, Peter E Caines, and Roland P Malhamé. Large-population cost-coupled lqg problems with nonuniform agents: individual-mass behavior and decentralized  $\varepsilon$ -nash equilibria. *IEEE transactions on automatic control*, 52(9):1560–1571, 2007.
- [13] Minyi Huang, Peter E Caines, and Roland P Malhamé. The nash certainty equivalence principle and mckean-vlasov systems: an invariance principle and entry adaptation. In *Decision and Control, 2007 46th IEEE Conference on*, pages 121–126. IEEE, 2007.
- [14] Minyi Huang, Peter E Caines, and Roland P Malhamé. The nce (mean field) principle with locality dependent cost interactions. *IEEE Transactions on Automatic Control*, 55(12):2799–2805, 2010.
- [15] Minyi Huang, Roland P Malhamé, Peter E Caines, et al. Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.
- [16] Daniel Lacker. A general characterization of the mean field limit for stochastic differential games. *Probability Theory and Related Fields*, 165(3-4):581–648, 2016.



- [17] Olga Aleksandrovna Ladyzhenskaia, Vsevolod Alekseevich Solonnikov, and Nina N Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.
- [18] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. i—le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.
- [19] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. ii—horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10):679–684, 2006.
- [20] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.
- [21] Pierre-Louis Lions. Cours au collège de france. 2006-2012.
- [22] Alessio Porretta. Weak solutions to fokker-planck equations and mean field games. *Arch. Ration. Mech. Anal*, 216(1):1–62, 2015.