



# Finite state N-agent and mean field control problems

Alekos Cecchin

## ► To cite this version:

Alekos Cecchin. Finite state N-agent and mean field control problems. ESAIM: Control, Optimisation and Calculus of Variations, 2021, 27, pp.31. 10.1051/cocv/2021032 . hal-03192295

**HAL Id: hal-03192295**

**<https://hal.science/hal-03192295>**

Submitted on 7 Apr 2021

**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.

## FINITE STATE $N$ -AGENT AND MEAN FIELD CONTROL PROBLEMS<sup>\*,\*\*</sup>

ALEKOS CECCHIN<sup>\*\*\*</sup>

**Abstract.** We examine mean field control problems on a finite state space, in continuous time and over a finite time horizon. We characterize the value function of the mean field control problem as the unique viscosity solution of a Hamilton-Jacobi-Bellman equation in the simplex. In absence of any convexity assumption, we exploit this characterization to prove convergence, as  $N$  grows, of the value functions of the centralized  $N$ -agent optimal control problem to the limit mean field control problem value function, with a convergence rate of order  $\frac{1}{\sqrt{N}}$ . Then, assuming convexity, we show that the limit value function is smooth and establish propagation of chaos, *i.e.* convergence of the  $N$ -agent optimal trajectories to the unique limiting optimal trajectory, with an explicit rate.

**Mathematics Subject Classification.** 35B65, 35F21, 49L25, 49M25, 60F15, 60J27, 91A12.

Received November 7, 2020. Accepted March 17, 2021.

### 1. INTRODUCTION

Mean field control problems (MFCP), also called control of McKean-Vlasov equations, can be interpreted as limit of cooperative  $N$ -agent games, as the number of players tends to infinity. Such agents have a common cost to minimize and the minimizers are also called *Pareto* equilibria; alternatively, we could think of a social planner that minimizes an average cost.

We investigate  $N$ -agent optimization and mean field control problems in continuous time over a finite time horizon, with dynamics belonging to a finite state space  $\{1, \dots, d\}$ . More precisely, the  $N$  agents  $\mathbf{X} = (X^1, \dots, X^N)$  follow the dynamics

$$\mathbb{P}(X_{t+h}^k = j | \mathbf{X}_t = \mathbf{x}) = Q_{x_k, j}(t, \beta_k(t, \mathbf{x}), \mu_{\mathbf{x}}^N)h + o(h),$$

---

\* This research benefited from the support of the project ANR-16-CE40-0015-01 on “Mean Field Games”, LABEX Louis Bachelier Finance and Sustainable Growth - project ANR-11-LABX-0019, under the Investments for the Future program (in accordance with Article 8 of the Assignment Agreements for Communication Assistance), ECOREES ANR Project, FDD Chair and Joint Research Initiative FiME in partnership with Europlace Institute of Finance.

\*\* The author is grateful to François Delarue and Charles Bertucci for helpful comments and suggestions, as well as to two anonymous referees for their remarks which helped to improve the paper.

*Keywords and phrases:* Mean field control problem, control of Markov chains, finite state space, cooperative games, social planner, Hamilton-Jacobi-Bellman equation, viscosity solution, finite difference approximation, classical solution, propagation of chaos.

Centre de Mathématiques Appliquées, École Polytechnique, Route de Saclay, 91124 Palaiseau Cedex, France.

\*\*\* Corresponding author: [alekos.cecchin@polytechnique.edu](mailto:alekos.cecchin@polytechnique.edu)

where  $\mu^N$  is the empirical measure and the controls (here in feedback form)  $\beta = (\beta_1, \dots, \beta_N)$  are chosen in order to minimize the common cost

$$J^N(\beta) := \frac{1}{N} \sum_{k=1}^N \mathbb{E} \left[ \int_0^T f(t, X_t^k, \beta_k(t, \mathbf{X}_t), \mu_t^N) dt + g(X_T^k, \mu_T^N) \right].$$

Assuming that controls depend only on the empirical measure, the usual propagation of chaos arguments suggest that the limit of this  $N$ -agent optimization, as  $N$  grows, consists in a single player which evolves according to

$$\mathbb{P}(X_{t+h} = j | X_t = i) = Q_{i,j}(t, \alpha^i(t), \text{Law}(X_t))h + o(h)$$

and aims at minimizing

$$J(\alpha) := \mathbb{E} \left[ \int_0^T f(t, X_t, \alpha(t, X_t), \text{Law}(X_t)) dt + g(X_T, \text{Law}(X_T)) \right].$$

Our goal is to study in detail the  $N$ -agent optimization and the mean field control problem and thus prove convergence of the former to the latter, the main result being to provide an explicit convergence rate.

Such problems are studied so far mainly for continuous state space and diffusion-based dynamics. In such situation, the limiting MFPC can be analyzed in two different ways, by considering either open-loop or (Markovian) feedback controls. In the first case, a version of the Pontryagin principle is derived in [11], which leads to study a forward-backward system of (Itô-type) SDEs of McKean-Vlasov type. While the case of feedback controls is analyzed in [34, 35], where the MFPC is reformulated in a deterministic way as the optimal control of the Fokker-Planck equation, which permits to derive a dynamic programming principle and then a Hamilton-Jacobi-Bellman (HJB) equation for the value function, written in the Wasserstein space of probability measures. See also [5, 31] for previous ideas in this direction and [4, 21] for more general versions of the dynamic programming principle. We refer to Chapter 6 of [12] for a comparison of the two approaches. Mean field control problems arise also in the study of potential mean field games, see *e.g.* [6, 10], where it is shown that the mean field game system represents the necessary conditions for optimality of a suitable mean field control problem.

The question of convergence of the  $N$ -agent optimization to the MFPC, still in the diffusion setting, has been analyzed mainly in two ways. The first consists in showing that the set of (relaxed open-loop) optimizers of the  $N$ -agent optimization is precompact and then that the limit points are supported on optimizers of the MFPC. We remark that the optimizer is non-unique in general, but is unique under additional convexity assumptions; we return to this point below. This strategy is employed first in [29], then in [25] for deterministic dynamics and, more recently, in [22] for more general dynamics with a common noise and in [20] for problems with interaction also through the law of the control. Clearly, a convergence rate for the convergence of the value functions can not be proved using compactness arguments. The other way is to prove convergence via the system of FBSDEs, in case the limit solution is unique. In [11], the value functions are shown to converge, with a suitable convergence rate, assuming that the cost  $f$  is convex in  $(a, x, m)$  and  $g$  is convex in  $(x, m)$ ; see also Section 6.1.3 of [13]. Moreover, a propagation of chaos property is also proved, *i.e.* the prelimit (unique) optimal trajectories are shown to converge to the (unique) limit optimal trajectory, with a convergence rate (actually, such result is not stated in this way, but can be immediately derived from the proof of Theorem 6.1 therein). More recently, this method has been applied to problems with interaction through the law of the control in [32]. In both cases, as a consequence of the convergence of the value functions, an optimal control for the MFPC is shown to be  $\varepsilon_N$ -optimal for the  $N$ -agent optimization, with  $\lim_N \varepsilon_N = 0$ , with the same limitations just explained (no convergence rate in the first case and convexity required in the second).

Here, in the finite state setting, we first analyze the MFPC with feedback controls. We rewrite it as a deterministic control problem for the Fokker-Planck equation and show that its value function  $V$  is the unique viscosity solution of the corresponding HJB equation, stated in the  $d$ -dimensional simplex. Then we examine

the convergence problem: we stress that convergence can be understood both in terms of value functions and of optimal trajectories. Our main result is to show that the value function  $V^N$  of the  $N$ -agent optimization converges to  $V$ , with a convergence rate of order  $1/\sqrt{N}$ , in absence of any convexity assumption. As explained above, a result of this type is not available for diffusion-based models. As a consequence of this convergence, we also prove that any optimal control for the MFCP is  $\frac{C}{\sqrt{N}}$ -optimal for the  $N$ -agent optimization. A similar result is proved also in [27] with different methods and without interpreting the prelimit model as a  $N$ -agent optimization; see Remark 2.12 for the details. In the discrete time setting, the MFCP and the related convergence of the  $N$  agent are investigated in full generality in [33].

Our main novelty consists in the method of the proof of the main result, which is different with respect to the two explained above and we believe can be of interest. This is based on the viscosity solution characterization of  $V$ : indeed, we will see that the ODE satisfied by  $V^N$  can be seen as a finite difference scheme for the HJB equation satisfied by  $V$ . Notably,  $V$  is not differentiable as we do not assume any convexity of the costs, neither in  $a$  nor in  $m$ . Finite difference schemes for viscosity solutions have been investigated by many authors in the last decades. Being impossible to give a complete bibliography on the subject, we would like to mention two papers which inspired our proof of convergence and that established, in particular, a rate of convergence. The first is [8], which studied a semidiscrete approximation scheme for the HJB equation of infinite horizon control problem with discount; see also the book ([1], Sect. VI.1). While the second is [38], which analyzed a finite difference scheme for a general time-dependent Hamilton-Jacobi equation; see also [18].

We also study the propagation of chaos property for the optimal trajectories, in case the limit is unique. If the value function  $V$  is sufficiently smooth, *i.e.* in  $\mathcal{C}^{1,1}$ , then the MFCP is uniquely solvable and we prove that the prelimit (unique) optimal empirical measures converge to the limit deterministic optimal flow, with a suitable convergence rate. We give also sufficient conditions for which  $V \in \mathcal{C}^{1,1}$ : these are the standard convexity assumptions. We remark that also these smoothness and propagation of chaos results seem to be new in the study of MFCP. Moreover, it is worth saying that we do not treat here neither problems with a common noise nor with interaction through the law of the control; these are left to future work.

Finally, we mention that convergence results have been obtained also for the opposite regime of mean field games. In that case, players are non-cooperative in the prelimit  $N$ -player game and the notion of optimality is that of Nash equilibrium, which highly depends on the set of admissible strategies that is considered. This makes the convergence analysis more difficult, especially in case limiting mean field game solutions are non-unique; some references are [9, 19, 23, 28, 30] for diffusion-based models and [2, 3, 14, 17, 26] for finite state space.

The rest of the paper is organized as follows. In Section 2, we collect our main results: after introducing the notations and assumptions that will be in force, we define properly the  $N$ -agent optimization and show the equivalence with the mean field formulation, then we present the MFCP with its well-posedness result (Thm. 2.9) and thus we state the convergence theorems. In Section 3, we examine the MFCP and prove first well-posedness of viscosity solutions and then, under additional convexity assumptions, well-posedness of classical solutions. Notably, we establish a comparison principle (Thm. 3.4) for viscosity solutions on the interior of the simplex – without boundary conditions, by exploiting the invariance of the domain –, which is a new result in the theory. Finally, Section 4 contains the proofs of the convergence results (Thms. 2.10 and 2.13): first the convergence of the value functions *via* viscosity solutions and then, assuming  $V \in \mathcal{C}^{1,1}$ , the propagation of chaos property.

## 2. MAIN RESULTS

### 2.1. Notation

We denote  $\llbracket d \rrbracket = \{1, \dots, d\}$  and let

$$S_d := \left\{ (m_1, \dots, m_d) \in \mathbb{R}^d : m_i \geq 0 \quad \forall i \in \llbracket d \rrbracket, \quad \sum_{i=1}^d m_i = 1 \right\}$$

be the  $(d-1)$ -dimensional simplex, endowed with the euclidean norm  $|\cdot|$  in  $\mathbb{R}^d$ . We denote by  $\langle \cdot, \cdot \rangle$  the scalar product in  $\mathbb{R}^d$  and, for a matrix  $(Q_{i,j})_{i,j \in \llbracket d \rrbracket}$ , we let  $Q_{i,\bullet}$  be the row  $i$ . We denote the elements of the simplex by  $m$ , while  $\mu$  denotes processes with values in the simplex.

The simplex can be viewed as a subset of  $\mathbb{R}^{d-1}$ , by expressing the last coordinate as  $m_d = 1 - \sum_{j=1}^{d-1} m_j$ : denote then

$$\widehat{S}_d := \left\{ (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1} : x_j \geq 0 \quad \forall j = 1, \dots, d-1, \quad \sum_{j=1}^{d-1} x_j \leq 1 \right\}.$$

We express the simplex using the particular local chart  $(m_1, \dots, m_{d-1}, m_d) = (x_1, \dots, x_{d-1}, 1 - \sum_{j=1}^{d-1} x_j)$ . Let  $\text{Int}(\widehat{S}_d)$  be the interior of  $\widehat{S}_d$  in  $\mathbb{R}^{d-1}$ ; when we refer to the interior of simplex, denoted by  $\text{Int}(S_d)$ , we mean the image of  $\text{Int}(\widehat{S}_d)$  under the above chart. *via* the above local chart, a function  $v$  defined on the simplex is equivalently written as a function  $\widehat{v}$  defined on  $\widehat{S}_d$ . Thus we say that  $v \in \mathcal{C}^1(S_d)$  if  $\widehat{v} \in \mathcal{C}^1(\widehat{S}_d)$ , meaning that  $\widehat{v} \in \mathcal{C}^1(\text{Int}(\widehat{S}_d))$  with derivatives that extend continuously up to the boundary. In the interior of the simplex, derivatives are allowed only along directions  $(\delta_j - \delta_i)_{i,j \in \llbracket d \rrbracket}$ , which are denoted as  $\partial_{m_j - m_i} v(m)$ ; we define the vector

$$D^i v(m) := (\partial_{m_j - m_i} v(m))_{j \in \llbracket d \rrbracket}.$$

Note that this is indeed a vector in  $\mathbb{R}^{d-1}$  since  $D_i^i v(m) = 0$ . The derivatives of  $\widehat{v} : \widehat{S}_d \rightarrow \mathbb{R}$  are denoted by  $D_x \widehat{v}(x) = (\partial_{x_j} \widehat{v}(x))_{j=1}^{d-1}$ , with the obvious identity  $\partial_{x_j} \widehat{v}(x) = \partial_{m_j - m_d} v(m)$ ,  $\forall j \in \llbracket d-1 \rrbracket$ , if  $x$  represents  $m$  in local chart. Let  $\mathcal{C}^{1,1}(S_d)$  be the set of  $v \in \mathcal{C}^1(S_d)$  whose derivative  $D_x \widehat{v}$  is Lipschitz-continuous, globally in  $S_d$ .

For  $\mathbf{x} = (x_1, \dots, x_N) \in \llbracket d \rrbracket^N$ , denote the empirical measure

$$\mu_{\mathbf{x}}^N := \frac{1}{N} \sum_{k=1}^N \delta_{x_k}, \quad \text{that is } \mu_{\mathbf{x}}^N[i] = \frac{1}{N} \sum_{k=1}^N \mathbb{1}_{\{x_k=i\}}, \quad i \in \llbracket d \rrbracket. \quad (2.1)$$

It takes values in the discretized simplex  $S_d^N := S_d \cap \frac{1}{N} \mathbb{N}^d$ . We also denote the  $N$ -discretized derivative  $D^{N,i} v = (D_j^{N,i} v)_{j \in \llbracket d \rrbracket}$  of a function  $v : S_d \rightarrow \mathbb{R}$ , with

$$D_j^{N,i} v(m) := N \left[ v \left( m + \frac{1}{N} (\delta_j - \delta_i) \right) - v(m) \right]; \quad (2.2)$$

while for a function  $u : \llbracket d \rrbracket^N \rightarrow \mathbb{R}$  we denote by  $\Delta^k u \in \mathbb{R}^d$  the vector of differences, that is,  $\Delta^k u(\mathbf{x})[j] = u([\mathbf{x}^{-k}, j]) - u(\mathbf{x})$ , whereas, for  $\mathbf{x} \in \llbracket d \rrbracket^N$  and  $j \in \llbracket d \rrbracket$ ,  $[\mathbf{x}^{-k}, j]$  denotes the vector in  $\llbracket d \rrbracket^N$  such that  $[\mathbf{x}^{-k}, j]_l = \begin{cases} x_l & \text{if } l \neq k \\ j & \text{if } l = k. \end{cases}$

## 2.2. Assumptions

Let  $Q_{i,j} : [0, T] \times S_d \times A \rightarrow [0, +\infty)$  be the transition rate,  $f^i : [0, T] \times A \times S_d \rightarrow \mathbb{R}$  the running cost and  $g^i : S_d \rightarrow \mathbb{R}$  the terminal cost, where  $i \in \llbracket d \rrbracket$  represents the state. Let  $F : [0, T] \times A^d \times S_d \rightarrow \mathbb{R}$  and  $G : S_d \rightarrow \mathbb{R}$  be defined by

$$F(t, a^1, \dots, a^d, m) := \sum_{i \in \llbracket d \rrbracket} m_i f^i(t, a^i, m), \quad G(m) = \sum_{i \in \llbracket d \rrbracket} m_i g^i(m). \quad (2.3)$$

We provide three sets of assumptions. The first is the weakest and gives convergence of the value functions, with a convergence rate. Note in particular that we don't assume convexity, neither in  $a$ , nor in  $m$ .

**Assumption A.** (A1) The action space  $(A, d)$  is a compact metric space.

(A2) The transition rate  $Q_{i,j}$  is continuous on  $[0, T] \times A \times S_d$  (thus uniformly continuous and bounded) and Lipschitz-continuous in  $(t, m)$ , uniformly in  $a$ :

$$|Q_{i,j}(t, a, m) - Q_{i,j}(s, a, p)| \leq C(|t - s| + |m - p|). \quad (2.4)$$

(A3) The functions  $F$  is continuous on  $[0, T] \times A^d \times S_d$  and

$$|F(t, a, m) - F(s, a, m)| \leq C(|t - s| + |m - p|), \quad (2.5)$$

$$|G(m) - G(p)| \leq C|m - p|. \quad (2.6)$$

We denote then, for  $z \in \mathbb{R}^d$  such that  $z_i = 0$ ,  $a \in A$  and  $m \in S_d$ , the pre-Hamiltonian

$$\mathcal{H}^i(t, a, m, z) := -\langle Q_{i,\bullet}(t, a, m), z \rangle - f^i(t, a, m) \quad (2.7)$$

and the Hamiltonian

$$H^i(t, m, z) := \max_{a \in A} \mathcal{H}^i(t, a, m, z). \quad (2.8)$$

Note that  $\sum_i m_i H^i$  is Lipschitz-continuous in  $(t, m, z)$  if **A** holds.

The second assumption is a linear-convex assumption, very common in control theory, which, together with existence of classical solution to the limiting problem, gives convergence of the optimal trajectories.

**Assumption B.** Assumption **A** holds and, in addition:

(B1)  $A = [0, M]^d$ .

(B2) The transition rate is  $Q_{i,j}(t, m, a) = a_j$ .

(B3) The running cost  $f$  is continuously differentiable in  $A$ ,  $\nabla_a f$  is Lipschitz-continuous with respect to  $m$ , and  $f$  is uniformly convex in  $A$ , *i.e.* there exists  $\lambda > 0$  such that

$$f^i(t, b, m) \geq f^i(t, a, m) + \langle \nabla_a f(t, a, m), (b - a) \rangle + \lambda |b - a|^2. \quad (2.9)$$

Under this assumption, thanks to Proposition 1 in [26], there exists a unique maximizer of  $\mathcal{H}$ , which we denote by  $a^*(t, i, m, z)$ , and, further,  $a^*$  is Lipschitz continuous with respect to  $m$  and  $z$ , *i.e.*

$$|a^*(t, i, m, z) - a^*(t, i, p, w)| \leq C(|m - p| + |z - w|). \quad (2.10)$$

We will consider feedback controls  $\alpha : [0, T] \times \llbracket d \rrbracket \rightarrow A$  (or equivalently  $\alpha : [0, T] \rightarrow A^d$ ), thus, when (B1) holds, we denote  $\alpha_{i,j}(t) := \alpha_j(t, i)$ .

The last is a convexity assumption in the couple  $(\alpha, m)$  that is needed to prove smoothness of the value function of the MFCP.

**Assumption C.** Assumption **B** holds and, in addition:

(C1)  $F(\cdot, a, \cdot) \in \mathcal{C}^{1,1}([0, T] \times S_d)$ , uniformly in  $a$ , and  $G \in \mathcal{C}^{1,1}(S_d)$ .

(C2) The function

$$[0, T] \times [0, +\infty)^{d \times d} \times \text{Int}(S_d) \ni (t, w, m) \rightarrow \sum_i m_i f^i \left( t, \left( \frac{w_{i,j}}{m_i} \right)_{j \neq i}, m \right) \in \mathbb{R}$$

is convex in  $(w, m)$  and  $G$  is convex in  $m$ .

Assumption (C2) may seem strange at this stage, but will be clarified in Section 3.2.

**Remark 2.1.** The reader will notice that not all of the conditions of Assumption A (Lipschitz-continuity in  $(t, m)$ ) are necessary in all the statements in which A is assumed. The same is true also for Assumption B. We choose to make only three Assumptions for the sake of definiteness. We stress again that the main differences among the three assumptions are that in A nothing is convex, while in B we assume convexity in  $a$  and in C in  $(a, m)$ .

We remark that A allows to treat also the case of directed graphs, in which some transitions are forbidden: if  $E = \{e_1, \dots, e_d\}$  is the set of nodes and  $E_+(e_i)$ , for each  $i \in \llbracket d \rrbracket$ , is the subset of  $E \setminus \{e_i\}$  of nodes  $e_j$  for which there exists a directed edge from  $e_i$  to  $e_j$ , then the transition matrix  $Q$  is required to satisfy  $Q_{i,j}(t, a, m) = 0$  whenever  $e_j \notin E_+(e_i)$ .

We also remark that we do not assume that  $f$  splits in a function of  $(i, a)$  plus a function of  $(i, m)$ , as it would be in the case of potential mean field games which are described in Section 3.3.

We conclude this part with a natural example for which B and/or C are satisfied.

**Example 2.2.** As a natural running cost for which (2.9) holds, we could take

$$f^i(t, \alpha, m) := \frac{1}{2} \sum_{j \neq i} \alpha_{i,j}^2 + f_0^i(m). \quad (2.11)$$

In this case, we get

$$H^i(t, m, z) = \sum_{j \neq i} \left\{ -\mathfrak{a}^*(-z_j)z_j - \frac{1}{2} |\mathfrak{a}^*(-z_j)|^2 \right\} - f_0^i(m), \quad a_j^*(t, i, z) = \mathfrak{a}^*(-z_j),$$

where, for  $r \in \mathbb{R}$ ,  $\mathfrak{a}^*(r) := \begin{cases} 0 & r \leq 0 \\ r & 0 \leq r \leq M \\ M & r \geq M \end{cases}$ .

In fact, we remark that our Assumptions B and C are more general since they allow for a running cost which does not split as in (2.11). If  $f$  is as in (2.11), then (C2) is satisfied if the function  $m \rightarrow \sum_i m_i f_0^i(m)$  is convex; indeed, one can easily verify that the function  $(w, m) \rightarrow \sum_i \sum_{j \neq i} \frac{w_{i,j}^2}{m_i}$  is convex in  $(w, m)$ .

### 2.3. $N$ -agent optimization

Consider  $N$  players,  $\mathbf{X} = (X^1, \dots, X^N)$ , such that  $X_t^i \in \llbracket d \rrbracket$ , evolving in continuous time over a finite horizon  $T$ . Agents can choose controls  $\beta = (\beta_1, \dots, \beta_N)$  in feedback form, *i.e.* any  $\beta_k$  is a measurable function of time and state of all players:  $\beta_k : [0, T] \times \llbracket d \rrbracket^N \rightarrow A$ . The dynamics is given as a Markov chain such that

$$\mathbb{P}(X_{t+h}^k = j | \mathbf{X}_t = \mathbf{x}) = Q_{x_k, j}(t, \beta_k(t, \mathbf{x}), \mu_{\mathbf{x}}^N)h + o(h), \quad (2.12)$$

for  $j \neq x_k$ , as  $h \rightarrow 0^+$ . Agents are cooperative and aim at minimizing the common cost

$$J^N(\beta) := \frac{1}{N} \sum_{k=1}^N \mathbb{E} \left[ \int_0^T f(t, X_t^k, \beta_k(t, \mathbf{X}_t), \mu_t^N) dt + g(X_T^k, \mu_T^N) \right]. \quad (2.13)$$

The cost coefficients  $f$  and  $g$  depend on the empirical measure

$$\mu_t^N = \frac{1}{N} \sum_{k=1}^N \delta_{X_t^k} \in S_d^N. \quad (2.14)$$

We denote  $\mu^N = (\mu_1^N, \dots, \mu_d^N)$ , that is,  $\mu_{i,t}^N = \frac{1}{N} \sum_{k=1}^N \mathbb{1}_{\{X_t^k=i\}}$ .

The initial conditions  $(X_0^1, \dots, X_0^N)$  are assumed to be i.i.d with  $\text{Law}(X_0^1) = m_0$ . This can be seen as a single optimization problem for the process  $\mathbf{X}$ , governed by the generator

$$\mathcal{L}_t^{N,\beta} \varphi(\mathbf{x}) = \sum_{k=1}^N \sum_{j \neq x_k} Q_{x_k,j}(t, \beta_k(t, \mathbf{x}), \mu_{\mathbf{x}}^N) [\varphi(\mathbf{x}^{-k}, j) - \varphi(\mathbf{x})], \quad (2.15)$$

for any  $\varphi : \llbracket d \rrbracket^N \rightarrow \mathbb{R}$ . The value function  $v^N$  of this control problem solves the HJB equation

$$-\frac{d}{dt} v^N(t, \mathbf{x}) + \sup_{\mathbf{b} \in A^N} \left\{ -\mathcal{L}_t^{N,\mathbf{b}} v^N(t, \mathbf{x}) - \frac{1}{N} \sum_{k=1}^N f(t, x_k, b_k, \mu_{\mathbf{x}}^N) \right\} = 0,$$

which, by definition of the Hamiltonian in (2.8), gives

$$\begin{aligned} -\frac{d}{dt} v^N(t, \mathbf{x}) + \frac{1}{N} \sum_{k=1}^N H^{x_k}(t, \mu_{\mathbf{x}}^N, N \Delta^k v^N(t, \mathbf{x})) &= 0, \\ v^N(T, \mathbf{x}) &= \frac{1}{N} \sum_{k=1}^N g(x_k, \mu_{\mathbf{x}}^N). \end{aligned} \quad (2.16)$$

This is a system of ODEs indexed by  $\mathbf{x} \in \llbracket d \rrbracket^N$ .

**Proposition 2.3.** *Under Assumption A, there exists a unique solution  $v^N$  to (2.16), which is  $\mathcal{C}^1$  in time. It is the value function of the control problem (2.13)–(2.15) and there exists an optimal feedback control.*

*Proof.* This is a standard verification theorem (see e.g. [24], Thm. III.8.1) and existence of solution is given by the Lipschitz continuity of the Hamiltonian in (2.16), which follows by the continuity of the coefficients and compactness of  $A$ . These properties also yield existence of a maximizer in (2.7) and hence existence of an optimal feedback.  $\square$

**Remark 2.4.** The optimal control is not unique and there might exist non-exchangeable optimizers. We recall that a vector of stochastic processes is said to be exchangeable if its joint law is invariant under permutations.

Under Assumption B the optimal control is unique.

**Remark 2.5.** For problem (2.13)–(2.15), the choice of controls in Markovian feedback form is made for convenience only and is the most natural setup for this type of control problems. We could consider more general open-loop controls: in this case, the strategy vector  $(\pi_t^1, \dots, \pi_t^N)_{t \in [0, T]}$  is a vector of predictable  $A$ -valued processes and the dynamics of the state process  $\mathbf{X}$  can be defined as the solution of the controlled martingale problem related to the generator (2.15), in which  $\beta_k$  is replaced by the stochastic process  $\pi^k$ . Notably, the value function of this more general control problem still solves equation (2.16), which admits a unique solution, and thus these two control problems are equivalent.



### 2.3.1. Mean field formulation

We give another formulation of the  $N$ -agent optimal control, by restricting the class of admissible controls. Nevertheless, we will prove that the two formulations are equivalent, in the sense that the value function is the same and thus the infimum of the cost is the same.

Assume then that the control is the same for any player and is given by a feedback (measurable) Markovian function  $\alpha_N : [0, T] \times \llbracket d \rrbracket \times S_d^N \rightarrow A$ : we make the mean field assumption for which the control depends on the private state and on the state of other players only through the empirical measure  $\mu_t^N$  of the entire system. Namely, we assume that, for any  $k \in \llbracket N \rrbracket$ ,

$$\beta_k(t, \mathbf{x}) = \alpha_N(t, x_k, \mu_{\mathbf{x}}^N). \quad (2.17)$$

Thus we have

$$\mathbb{P}(X_{t+h}^k = j | X_t^k = i, \mu_t^N = m) = Q_{i,j}(t, \alpha_N(t, i, m), m)h + o(h), \quad (2.18)$$

as  $h \rightarrow 0^+$ .

The aim of the players is to choose  $\alpha^N$  in order to minimize the cost in (2.13), which, assuming now (2.17), can be rewritten as

$$\begin{aligned} J^N(\alpha_N) &:= \frac{1}{N} \sum_{k=1}^N \mathbb{E} \left[ \int_0^T f(t, X_t^k, \alpha_N(t, X_t^k, \mu_t^N), \mu_t^N) dt + g(X_T^k, \mu_T^N) \right] \\ &= \frac{1}{N} \sum_{k=1}^N \mathbb{E} \left[ \int_0^T \sum_{i \in \llbracket d \rrbracket} \mathbb{1}_{\{X_t^k = i\}} f(t, i, \alpha_N(t, i, \mu_t^N), \mu_t^N) dt + \sum_{i \in \llbracket d \rrbracket} \mathbb{1}_{\{X_T^k = i\}} g(i, \mu_T^N) \right], \end{aligned}$$

and hence

$$J^N(\alpha_N) = \mathbb{E} \left[ \int_0^T \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N f^i(t, \alpha_N(t, i, \mu_t^N), \mu_t^N) dt + \sum_{i \in \llbracket d \rrbracket} \mu_{i,T}^N g^i(\mu_T^N) \right]. \quad (2.19)$$

Therefore the  $N$ -agent mean field control problem can be seen as a single optimization problem for the empirical measure, which is a time-inhomogeneous Markov chain on  $S_d^N$  such that

$$\mathbb{P} \left( \mu_{t+h}^N = m + \frac{1}{N}(\delta_j - \delta_i) \middle| \mu_t^N = m \right) = N m_i Q_{i,j}(t, \alpha_N(t, i, m), m)h + o(h) \quad (2.20)$$

for any  $m \in S_d^N$  and  $i \neq j \in \llbracket d \rrbracket$ . The control (in feedback form) is now the vector valued measurable function  $\alpha^N(t, \cdot, m) \in A^d$ . We remark that the dynamics remains in  $S_d^N$  because  $m + \frac{1}{N}(\delta_j - \delta_i)$  can be outside  $S_d^N$  only if  $m_i = 0$ , but in such case the transition rate is zero. Clearly, we assume that  $\mu_0 = \frac{1}{N} \sum_{k=1}^N \delta_{X_0^k}$ . The generator of this Markov chain is hence given, for  $v : S_d^N \rightarrow \mathbb{R}$ , by

$$\mathcal{L}_t^{N, \alpha_N} v(m) = N \sum_{i,j \in \llbracket d \rrbracket} m_i Q_{i,j}(t, \alpha_N(t, i, m), m) \left[ v \left( m + \frac{1}{N}(\delta_j - \delta_i) \right) - v(m) \right]. \quad (2.21)$$

The HJB equation for the value function of this problem is then

$$-\frac{d}{dt}V^N(t, m) + \max_{a \in A^d} \left\{ -\mathcal{L}^{N,a}V^N(t, m) - \sum_{i \in \llbracket d \rrbracket} m_i f^i(t, a^i, m) \right\} = 0,$$

which rewrites as

$$\begin{aligned} -\frac{d}{dt}V^N + \sum_{i \in \llbracket d \rrbracket} m_i H^i(t, m, D^{N,i}V^N(t, m)) &= 0 \\ V(T, m) &= \sum_{i \in \llbracket d \rrbracket} m_i g^i(m), \end{aligned} \tag{2.22}$$

that is an ODE indexed by  $m \in S_d^N$ .

**Proposition 2.6.** *Under Assumption A, the HJB equation (2.22) admits a unique solution  $V^N$  which is  $\mathcal{C}^1$  in time;  $V^N$  is the value function of the control problem (2.19)–(2.20), defined as the infimum over feedback controls  $\alpha_N : [0, T] \times S_d^N \rightarrow A^d$ , and there exists an optimal feedback.  $V^N$  satisfies the Lipschitz property*

$$|V^N(t, m) - V^N(s, p)| \leq C(|t - s| + |m - p|) \tag{2.23}$$

for a constant  $C$  independent of  $N$ .

Moreover, the value function  $v^N$  of (2.13)–(2.15) satisfies

$$v^N(t, \mathbf{x}) := V^N(t, \mu_{\mathbf{x}}^N) \tag{2.24}$$

and thus the control problems (2.13)–(2.15) and (2.19)–(2.20) are equivalent, i.e.

$$\inf_{\beta} J^N(\beta) = \inf_{\alpha_N} J^N(\alpha_N). \tag{2.25}$$

*Proof.* The first claim follows by the Lipschitz continuity of  $H$  and a standard verification theorem, an optimal feedback being a measurable function that attains the maximum in (2.7), which exists by compactness of  $A$  and continuity of the coefficients. The Lipschitz continuity of  $V^N$  is proved in Section 4.1, Lemma 4.1. If  $w^N$  is the function defined by the r.h.s. of (2.24), then

$$\begin{aligned} &-\frac{d}{dt}w^N(t, \mathbf{x}) + \frac{1}{N} \sum_{k=1}^N H^{x_k}(t, \mu_{\mathbf{x}}^N, N(w^N(t, j, \mathbf{x}^{-k}) - w^N(t, \mathbf{x}))_{j \in \llbracket d \rrbracket}) \\ &= -\frac{d}{dt}V^N(t, \mu_{\mathbf{x}}^N) + \frac{1}{N} \sum_{k=1}^N \sum_{i \in \llbracket d \rrbracket} \mathbb{1}_{\{x_k=i\}} H^i(t, \mu_{\mathbf{x}}^N, N(V^N(t, \mu_{\mathbf{x}}^N + \frac{1}{N}(\delta_j - \delta_i)) - V^N(t, \mu_{\mathbf{x}}^N))_{j \in \llbracket d \rrbracket}) \\ &= -\frac{d}{dt}V^N(t, \mu_{\mathbf{x}}^N) + \sum_{i \in \llbracket d \rrbracket} \mu_{\mathbf{x}}^N[i] H^i(t, \mu_{\mathbf{x}}^N, D^{N,i}V^N(t, \mu_{\mathbf{x}}^N)) = 0 \end{aligned}$$

because  $V^N$  solves (2.22). Hence, by uniqueness of the solution to (2.16), we have  $w^N = v^N$  and thus (2.24) is satisfied.  $\square$

**Remark 2.7.** We could restrict the class of admissible controls to the set of  $\alpha_N : [0, T] \rightarrow A^d$  that are deterministic functions only of time and of the private state  $i \in \llbracket d \rrbracket$ . This control problem is equivalent to what we consider here because its value function  $W^N$  is clearly Lipschitz in time, and hence absolutely continuous,

and then, by the dynamic programming principle, it is easy to show that  $W^N$  solves (2.22) at any point of differentiability. Thus, by uniqueness of the solution to (2.22) (defined in the sense of Caratheodory in the class of absolutely continuous functions) we get  $V^N = W^N$ , which means that the costs have the same infimum.

We could further restrict the class of admissible controls, for instance, to the set of piecewise constant (deterministic) functions of time. This setting is the one considered in [27].

If assumption B holds, then the optimal control is unique, since the maximizer in (2.7) is unique (see *e.g.* [16], Thm. 5). The control is the transition rate  $\alpha_N(t, i, m) \in [0, M]^d$ , which we denote as a transition matrix  $\alpha_N = (\alpha_N^{i,j})_{i,j \in \llbracket d \rrbracket} \in [0, M]^{d \times d}$ . In such case, (2.18) becomes simply

$$\mathbb{P}(X_{t+h}^k = j | X_t^k = i, \mu_t^N = m) = \alpha_N^{i,j}(t, m)h + o(h), \quad (2.26)$$

and the dynamics of  $\mu^N$  is then given by

$$\mathbb{P}\left(\mu_{t+h}^N = m + \frac{1}{N}(\delta_j - \delta_i) \middle| \mu_t^N = m\right) = Nm_i \alpha_N^{i,j}(t, m)h + o(h). \quad (2.27)$$

Note that the values of  $\alpha^{i,i}$  never enter in the dynamics. Then it follows:

**Proposition 2.8.** *Under Assumption B, there exists a unique optimal control. It is given by*

$$\alpha_{N,*}^{i,j}(t, m) := a_j^*(t, i, m, D^{N,i}V^N(t, m)). \quad (2.28)$$

## 2.4. Mean field control problem

In the limit, there is a single player which evolves according to

$$\mathbb{P}(X_{t+h} = j | X_t = i) = Q_{i,j}(t, \alpha^i(t), \text{Law}(X_t))h + o(h) \quad (2.29)$$

and  $\text{Law}(X_0) = m_0$ . The control is here a deterministic measurable function  $\alpha : [0, T] \rightarrow A^d$ , or equivalently  $\alpha : [0, T] \times \llbracket d \rrbracket \rightarrow A$ , which is indeed a feedback function of the state  $X_t$ , denoted by  $i \in \llbracket d \rrbracket$ . As a particular case, this set includes the functions  $[0, T] \times S_d \ni (t, \text{Law}(X_t)) \rightarrow \tilde{\alpha}(t, \text{Law}(X_t)) \in A^d$ , since  $t \rightarrow \text{Law}(X_t)$  is a deterministic function of time. The reference player aims at minimizing the cost

$$J(\alpha) := \mathbb{E} \left[ \int_0^T f(t, X_t, \alpha(t, X_t), \text{Law}(X_t)) dt + g(X_T, \text{Law}(X_T)) \right]. \quad (2.30)$$

The problem can be recasted into a deterministic control problem for the dynamics of the law of  $X$ , thanks to the fact that we consider Markovian feedback controls only. Indeed, denoting  $\mu_t = \text{Law}(X_t)$ , *i.e.*  $\mu_t^i = \mathbb{P}(X_t = i)$ , the cost is written as

$$J(\alpha) = \int_0^T \sum_{i \in \llbracket d \rrbracket} f^i(t, \alpha^i(t), \mu_t) \mu_t^i dt + \sum_{i \in \llbracket d \rrbracket} g^i(\mu_T) \mu_T^i, \quad (2.31)$$

where  $\mu$  solves the ODE, indexed by  $i \in \llbracket d \rrbracket$ ,<sup>1</sup>

<sup>1</sup>For the deterministic control problem (2.31)–(2.32) the control  $\alpha : [0, T] \rightarrow A^d$  is indeed in open-loop form, as it is just a measurable function of time.

$$\begin{aligned} \frac{d}{dt}\mu_t^i &= \sum_{j \in \llbracket d \rrbracket} \left( \mu_t^j Q_{j,i}(t, \alpha^j(t), \mu_t) - \mu_t^i Q_{i,j}(t, \alpha^i(t), \mu_t) \right) \\ \mu_0 &= m_0. \end{aligned} \quad (2.32)$$

The HJB equation for the value function  $V$  of this problem is then

$$-\partial_t V(t, m) + \max_{a \in A^d} \left\{ - \sum_{i,j \in \llbracket d \rrbracket} m_i Q_{i,j}(t, a^i, m) \partial_{m_j - m_i} V(t, m) - \sum_{i \in \llbracket d \rrbracket} m_i f^i(t, a^i, m) \right\} = 0,$$

which rewrites as

$$\begin{aligned} -\partial_t V(t, m) + \sum_{i \in \llbracket d \rrbracket} m_i H^i(t, m, D^i V(t, m)) &= 0, \\ V(t, m) &= \sum_{i \in \llbracket d \rrbracket} m_i g^i(m), \end{aligned} \quad (2.33)$$

in which we recall that  $[D^i V(t, m)]_j := \partial_{m_j - m_i} V(t, m)$ .

The peculiarity of this first-order equation is that it is stated in the simplex, which is a bounded domain, but there are no boundary conditions. This is explained by the fact that the simplex is invariant for the dynamics (2.32), and so is its interior. The first order HJB equation has no classical solutions in general, and for this reason viscosity solutions were introduced. These are defined properly in the next section, by viewing the simplex as a subset of  $\mathbb{R}^{d-1}$  instead of  $\mathbb{R}^d$ . Viscosity solutions can be defined either on  $S_d$  or on  $\text{Int}(S_d)$ , depending on the boundary regularity of the test functions involved. One problem with defining viscosity solutions on  $S_d$  is that it is not clear that a classical solution is a viscosity solution on  $S_d$ ; however, this definition is the one we will use to prove convergence of  $V^N$  to  $V$ . Exploiting the fact that  $\text{Int}(S_d)$  is invariant for the state dynamics, which results in a property on the subdifferential of  $H$ , it is possible to show uniqueness of viscosity solutions on  $\text{Int}(S_d)$ ; see the comparison principle below (Thm. 3.4).

If B holds, we denote as above  $\alpha_{i,j}(t) = \alpha_j(t, i)$ , which is the transition rate matrix. In the next section, we prove the following:

**Theorem 2.9.** *Let  $V$  be the value function of the deterministic control problem (2.31)–(2.32):*

1. *if Assumption A holds, then  $V$  is the unique viscosity solution of (2.33) on  $S_d$ , and  $V$  is Lipschitz-continuous in  $(t, m)$ ;*
2. *if Assumption A holds, then  $V$  is the unique viscosity solution of (2.33) on  $\text{Int}(S_d)$ , and, if B holds, there exists an optimal control;*
3. *if Assumption C holds, then  $V \in \mathcal{C}^{1,1}([0, T] \times S_d)$  and is the unique classical solution of (2.33);*
4. *if  $V \in \mathcal{C}^{1,1}(S_d)$  and B holds, then the control given by the feedback*

$$\alpha_*^{i,j}(t, m) := a_j^*(t, i, m, D^i V(t, m)) \quad (2.34)$$

*is the unique optimal control, in the sense that any optimal control  $\alpha : [0, T] \rightarrow [0, M]^{d \times d}$ , with related optimal process  $\mu$  is such that  $\alpha(t) = \alpha_*(t, \mu_t)$  for dt-a.e.  $t \in [0, T]$ .*

## 2.5. Convergence results

We state here the results about the convergence, as  $N \rightarrow \infty$ , of the value function  $V^N$  of the  $N$ -agent optimization (2.19)–(2.20) to the value function  $V$  of the mean field control problem (2.31)–(2.32), with a convergence rate of order  $1/\sqrt{N}$ . We recall that  $V^N$  is the classical solution to ODE (2.22), while  $V$  is the viscosity solution to PDE (2.33). The following is our main result:

**Theorem 2.10.** *Under Assumption A,*

$$\max_{t \in [0, T], m \in S_d^N} |V^N(t, m) - V(t, m)| \leq \frac{C}{\sqrt{N}}. \quad (2.35)$$

The theorem is proved in Section 4.1. As announced in the Introduction, we exploit the characterization of  $V$  as the viscosity solution to (2.33) in order to prove the result. In fact, ODE (2.22) can be seen as a finite difference scheme for the PDE (2.33), even if time is still continuous. Indeed the argument  $D^{N,i}V$  of the Hamiltonian in (2.22) converges, at least formally, to  $D^iV$  appearing in (2.33), as

$$\lim_{N \rightarrow \infty} N \left[ V \left( m + \frac{1}{N} (\delta_j - \delta_i) \right) - V(m) \right] = \partial_{m_j - m_i} V(m). \quad (2.36)$$

This result also permits to construct quasi-optimal controls for the  $N$ -agent optimization, starting from quasi-optimal controls for the MFCEP, with an explicit rate of approximation.

**Theorem 2.11.** *Assume A and fix  $\varepsilon > 0$  and  $N \in \mathbb{N}$ . Let  $\alpha : [0, T] \rightarrow A^d$  be an  $\varepsilon$ -optimal control for the MFCEP. Then*

$$J^N(\alpha) \leq \inf_{\alpha_N} J^N(\alpha_N) + \frac{C}{\sqrt{N}} + \varepsilon. \quad (2.37)$$

This is also proved in Section 4.1. Here,  $J^N(\alpha)$  is understood as applying the control  $\alpha_N(t, m) = \alpha(t)$ , which is independent of  $m$ . Recall that the infimum over controls  $\alpha_N$  is the same as the infimum over controls  $\beta$ , depending on states of all the players, by (2.25), and is also equal to the infimum over controls not depending on  $m$  (like the  $\alpha$  considered), by Remark 2.7.

**Remark 2.12.** In [27], Kolokoltsov proved a result similar to Theorem 2.10, but assuming in addition that  $F$ ,  $G$  and  $Q$  and  $\mathcal{C}^{1,1}$  w.r.t.  $m$  (similarly to Assumption (C1)). He analyzes a mean field  $N$ -optimization like in (2.19)–(2.21), but allowing for controls that are piecewise constant functions of time only, and are the same controls he considers in the limiting deterministic control problem (2.31)–(2.32). However, we explained in Remark 2.7 that considering this smaller class is not restrictive, as the value of the  $N$ -agent optimization is the same as what we treat here, *i.e.* over controls that might depend also on  $m$ . Then, by applying the convergence of the generator (2.21) to the limiting dynamics (2.32), he shows convergence of the value functions with a stronger convergence rate (Thm. 2 therein):

$$\max_{t \in [0, T], m \in S_d^N} |V^N(t, m) - V(t, m)| \leq \frac{C}{N}. \quad (2.38)$$

As a matter of fact, from his method of the proof (basically, the same set of controls is considered for the prelimit and the limit optimization problems), it is also possible to derive estimate (2.35). Indeed, by applying standard arguments in propagation of chaos, we can get a convergence rate of order  $1/\sqrt{N}$ , assuming that the costs and the transition rate are just Lipschitz-continuous w.r.t.  $m$ , and not in  $\mathcal{C}^{1,1}$ . Therefore, what we propose in this paper is a new method for proving the convergence in (2.35), based on the theory of viscosity solutions, which we believe can be of interest.

In case  $V$  is smooth, the optimal control of the MFCP is unique and then, if **B** holds, we are also able to establish a *propagation of chaos* result, that is, we prove convergence of the optimal trajectory of the  $N$ -agent optimization to the unique optimal trajectory of the MFCP, with a suitable convergence rate.

Denote then by  $\alpha^N$  the unique optimal feedback control for the  $N$ -agent optimization defined by (2.28), and by  $\mu^N$  the corresponding optimal process satisfying (2.27). Also, let  $\alpha_*$  be the unique optimal feedback control for the MFCP defined by (2.34) and  $\mu$  the corresponding optimal trajectory given by (2.32). We stress that  $\alpha_N$  and  $\alpha_*$  are functions of  $t$  and  $m$ .

**Theorem 2.13** (Part I). *Under Assumption **B**, if  $V \in C^{1,1}([0, T] \times S_d)$  then*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\mu_t^N - \mu_t| \right] \leq \frac{C}{N^{1/9}}. \quad (2.39)$$

The propagation of chaos result can be stated also for the vector of processes  $\mathbf{X}$  related to the optimal control  $\alpha_N$  and optimal empirical measure  $\mu^N$ , that is,  $\mathbf{X}$  is given by (2.12) assuming (2.17). For  $N$  fixed, denote by  $\widetilde{\mathbf{X}}$  the i.i.d process (given by (2.12)) in which all players choose the same local control  $\alpha(t, i) := \alpha_*(t, i, \mu_t)$  depending only on the private state, *i.e.*  $\beta_k(t, \mathbf{x}) = \alpha(t, x_k)$ . The propagation of chaos consists in proving convergence of  $\mathbf{X}$  to the i.i.d. process  $\widetilde{\mathbf{X}}$ .

**Theorem 2.13** (Part II). *Under Assumption **B** and (C1), if  $V \in C^{1,1}([0, T] \times S_d)$  then*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |X_t^i - \widetilde{X}_t^i| \right] \leq \frac{C}{\sqrt{N}}. \quad (2.40)$$

### 3. MEAN FIELD CONTROL PROBLEM

The aim here is to examine in detail the mean field control problem (2.31)–(2.32) in order to prove Theorem 2.9. We first rewrite the state dynamics and the cost in terms of the local chart  $x \in \mathbb{R}^{d-1}$  (recall that we set  $(m_1, \dots, m_{d-1}, m_d) = (x_1, \dots, x_{d-1}, 1 - \sum_{j=1}^{d-1} x_j)$ ), so that we are allowed to apply standard results about deterministic control problems on Euclidean spaces; see *e.g.* [1, 7, 24].

For  $x \in \widehat{S}_d$ , let us denote  $x^{-d} = 1 - \sum_{j=1}^{d-1} x_j$  and  $\tilde{x} = (x, x^{-d}) \in S_d$ . We simply replace  $m_d$  by  $x^{-d}$ : dynamics (2.32) becomes, for  $i = 1, \dots, d-1$ ,

$$\begin{aligned} \frac{d}{dt} x_t^i &= \sum_{j=1}^{d-1} \left( x_t^j Q_{j,i}(t, \alpha^j(t), \tilde{x}_t) - x_t^i Q_{i,j}(t, \alpha^i(t), \tilde{x}_t) \right) \\ &\quad + x_t^{-d} Q_{d,i}(t, \alpha^d(t), \tilde{x}_t) - x_t^i Q_{i,d}(t, \alpha^i(t), \tilde{x}_t), \end{aligned} \quad (3.1)$$

while the cost (2.31) is written as

$$\widehat{J}(\alpha) := \int_0^T \left( \sum_{j=1}^{d-1} x_t^j f^j(t, \alpha^j(t), \tilde{x}_t) + x_t^{-d} f^d(t, \alpha^d(t), \tilde{x}_t) \right) dt + \sum_{j=1}^{d-1} x_T^j g^j(\tilde{x}_T) + x_T^{-d} g^d(\tilde{x}_T). \quad (3.2)$$

It is clear that the value function  $\widehat{V}$  of (3.1)–(3.2), defined on  $\widehat{S}_d$ , is equal to the value function  $V$  of (2.31)–(2.32), defined on  $S_d$ , by setting  $\widehat{V}(t, x) = V(t, \check{x})$ . The HJB equation for  $\widehat{V}$  in  $\widehat{S}_d$  is then

$$\begin{aligned} -\partial_t \widehat{V} + \sum_{i=1}^{d-1} x_i \widehat{H}^i(t, x, D_x \widehat{V}) + x^{-d} \widehat{H}^d(t, x, D_x \widehat{V}) &= 0 \\ \widehat{V}(T, x) &= \sum_{i=1}^{d-1} x_i g^i(\check{x}) + x^{-d} g^d(x), \end{aligned} \quad (3.3)$$

where the modified Hamiltonian is defined, for  $z \in \mathbb{R}^{d-1}$ , by

$$\widehat{H}^i(t, x, w) = H^i(t, \check{x}, w_1 - w_i, \dots, w_{d-1} - w_i, -w_i), \quad \widehat{H}^d(t, x, w) = H^d(t, \check{x}, w, 0). \quad (3.4)$$

We will use several times the fact that  $\text{Int}(S_d)$  is invariant for the dynamics (2.32), or equivalently that  $\text{Int}(\widehat{S}_d)$  is invariant for (3.1). Indeed, by assumption A follows that  $Q$  is bounded (as it is continuous on a compact set): let us set

$$M := \max_{i,j \in \llbracket d \rrbracket, t \in [0, T], a \in A, m \in S_d} Q_{i,j}(t, a, m) \quad (3.5)$$

and recall that  $Q_{i,j} \geq 0$ . Then (2.32) gives

$$\frac{d}{dt} \mu_t^i \geq 0 - M(d-1) \mu_t^i,$$

which, by Gronwall's inequality, provides

$$\mu_t^i \geq \mu_0^i e^{-TM(d-1)}, \quad (3.6)$$

meaning that  $\mu_t \in \text{Int}(S_d)$  if  $\mu_0 \in \text{Int}(S_d)$ .

### 3.1. Viscosity solution

The value function is not  $\mathcal{C}^1$  in general and thus does not solve (2.33) in the classical sense, unless the convexity assumption C holds; see the next Subsection. Hence we present the two definitions of viscosity solutions we make use of: one in  $S_d$  and one in  $\text{Int}(S_d)$ . The notion of viscosity solution is the usual one, but we prefer to state them to avoid confusions, because the use of test functions defined on a closed set is not standard. By our definition of  $\mathcal{C}^1(S_d)$ , viscosity solutions can be defined in two equivalent ways, since it is equivalent to define functions on  $S_d$  or on  $\widehat{S}_d$ , by the relation  $\widehat{v}(x) = v(\check{x})$ . Thus we prefer to state only the definitions of solutions to (2.33) on  $S_d$  or  $\text{Int}(S_d)$ ; the definitions of solutions to (3.3) on  $\widehat{S}_d$  or  $\text{Int}(\widehat{S}_d)$  being equivalent.

**Definition 3.1.** A function  $v \in \mathcal{C}([0, T] \times S_d)$  is said to be:

- (i) a *viscosity subsolution* of (2.33) on  $S_d$  (resp. on  $\text{Int}(S_d)$ ) if, for any  $\varphi \in \mathcal{C}^1([0, T] \times S_d)$  (resp. in  $\mathcal{C}^1([0, T] \times \text{Int}(S_d))$ ),

$$-\partial_t \varphi(\bar{t}, \bar{m}) + \sum_{i \in \llbracket d \rrbracket} \bar{m}_i H^i(\bar{t}, \bar{m}, D^i \varphi(\bar{t}, \bar{m})) \leq 0, \quad (3.7)$$

at every  $(\bar{t}, \bar{m}) \in [0, T] \times S_d$  (resp. in  $[0, T] \times \text{Int}(S_d)$ ) which is a local maximum of  $v - \varphi$  on  $[0, T] \times S_d$  (resp. on  $[0, T] \times \text{Int}(S_d)$ );

- (ii) a *viscosity supersolution* of (2.33) on  $S_d$  (resp. on  $\text{Int}(S_d)$ ) if, for any  $\varphi \in \mathcal{C}^1([0, T] \times S_d)$  (resp. in  $\mathcal{C}^1([0, T] \times \text{Int}(S_d))$ ),

$$-\partial_t \varphi(\bar{t}, \bar{m}) + \sum_{i \in \llbracket d \rrbracket} \bar{m}_i H^i(\bar{t}, \bar{m}, D^i \varphi(\bar{t}, \bar{m})) \geq 0, \quad (3.8)$$

at every  $(\bar{t}, \bar{m}) \in [0, T] \times S_d$  (resp. in  $[0, T] \times \text{Int}(S_d)$ ) which is a local minimum of  $v - \varphi$  on  $[0, T] \times S_d$  (resp. on  $[0, T] \times \text{Int}(S_d)$ );

- (iii) a *viscosity solution* of (2.33) on  $S_d$  (resp on  $\text{Int}(S_d)$ ) if it is both a viscosity subsolution and a viscosity supersolution of (2.33) on  $S_d$  (resp on  $\text{Int}(S_d)$ ).

Let us remark that, in the above definition, by solutions to (2.33) we clearly mean solutions to the first line of (2.33). As an example of test functions in  $\mathcal{C}^1([0, T] \times S_d)$ , we may consider continuously differentiable functions defined on an open subset of  $\mathbb{R}^{d-1}$  containing  $\hat{S}_d$ . With this remark, it is straightforward to obtain the following result, which proves point (1) in Theorem 2.9:

**Proposition 3.2.** *Under Assumption A, the value function  $V$  of the MFCP is the unique viscosity solution of (2.33) on  $S_d$  (in  $\mathcal{C}([0, T] \times S_d)$  satisfying the terminal condition). Moreover,  $V$  is globally Lipschitz-continuous in  $[0, T] \times S_d$ .*

*Proof.* The Lipschitz-continuity and the viscosity solution property of  $V$  follow from standard results in deterministic control theory, see Thms. 7.4.10 and 7.4.14 of [7], by considering the dynamics in  $\mathbb{R}^{d-1}$  and using the fact that  $\hat{S}_{d-1}$  is invariant for the dynamics (3.1). Uniqueness of viscosity solutions on  $S_d$  follows by the usual proof of uniqueness, see for instance ([24], Thm. II.9.1), by observing that –if the minimizers are on boundary points– we can use the fact that the test functions constructed in the proof are quadratic and thus defined in the whole  $\mathbb{R}^{d-1}$ , in particular then belonging to  $\mathcal{C}^1([0, T] \times S_d)$ . We believe that there is no need to rewrite the proof here.  $\square$

This notion of viscosity solution on  $S_d$  will be used in the proof of the convergence result Theorem 2.10, see Subsection 4.1. Actually, uniqueness of viscosity solutions on  $S_d$  can also be derived as a consequence of the proof therein. The problem with the definition on the closed set  $S_d$  is that it is not clear whether a classical solution is a viscosity solution on  $S_d$ . Indeed, if the value function  $V$  is smooth, then for sure  $V$  is a viscosity solution on  $\text{Int}(S_d)$ , but if a maximizer of  $V - \psi$  lies on the boundary of  $S_d$  then it is not clear a priori that (3.7) holds. We could prove this fact, but we prefer instead to show uniqueness of viscosity solution on  $\text{Int}(S_d)$ , which implies in particular that  $V \in \mathcal{C}^1([0, T] \times S_d)$  is a viscosity solution on  $S_d$ , but is more general and we believe can have an interest on itself. The following result then proves points (2) and (4) in Theorem 2.9.

**Proposition 3.3.** *Under Assumption A,  $V$  is the unique viscosity solution of (2.33) on  $\text{Int}(S_d)$ , in  $\mathcal{C}([0, T] \times S_d)$  satisfying the terminal condition. Moreover, if B holds:*

1. *there exists an optimal control for the MFCP;*
2. *if  $V \in \mathcal{C}^{1,1}([0, T] \times S_d)$ , then the control given by the feedback (2.34) is the unique optimal control, in the sense that any optimal control  $\alpha : [0, T] \rightarrow [0, M]^{d \times d}$ , with related optimal process  $\mu$  is such that  $\alpha(t) = \alpha_*(t, \mu_t)$  for dt-a.e.  $t \in [0, T]$ .*

*Proof.* Existence of an optimal control follows from ([7], Thm 7.4.5), using the convexity of the cost in  $a$ . The viscosity solution property on  $\text{Int}(S_d)$  is given again by ([7], Thm. 7.4.14), by exploiting the invariance of  $\text{Int}(S_d)$ , see (3.6). Uniqueness of viscosity solutions on  $\text{Int}(S_d)$  is an immediate consequence of the Comparison Principle given below (Thm. 3.4).

Point (2) is a classical verification theorem: if  $V \in \mathcal{C}^{1,1}([0, T] \times S_d)$  then it is the unique classical solution of (2.33), solving the equation also at boundary points. Under assumption B, using in particular the strict convexity of  $F$  in  $a$ ,  $a^*$  is the unique argmin of the pre-Hamiltonian 2.7 and thus the optimal control defined by (2.34) is unique. Note that dynamics (2.32) is well-posed using the feedback  $\alpha_*$  because  $D^i V$  is Lipschitz-continuous.  $\square$



It remains to show the comparison principle for viscosity solutions on  $\text{Int}(S_d)$ ; to this end, it turns out that it is better to consider the dynamics in  $\mathbb{R}^{d-1}$  and thus we state the equivalent result on  $\text{Int}(\widehat{S}_d)$ . In absence of boundary conditions in space, we must rely on the invariance of  $\text{Int}(\widehat{S}_d)$  for the dynamics (3.1). The following result then extends what we presented in ([15], Thm. 6.2) to a more general dynamics, and borrows ideas from the proofs of Theorem 3.8 and Proposition 7.3 in [36]. We stress that we do not require here neither differentiability of the Hamiltonian nor convexity of the costs (in  $a$  or  $m$ ).

**Theorem 3.4** (Comparison Principle on  $\text{Int}(\widehat{S}_d)$ ). *Assume A and let  $u, v \in \mathcal{C}([0, T] \times \widehat{S}_d)$ ,  $u$  be a viscosity subsolution and  $v$  be a viscosity supersolution, respectively, of (3.3) on  $\text{Int}(\widehat{S}_d)$ . If  $u(T, x) \leq v(T, x)$  for any  $x \in \widehat{S}_d$ , then  $u(t, x) \leq v(t, x)$  for any  $t \in [0, T]$  and  $x \in \widehat{S}_d$ .*

*Proof.* The idea is to define a supersolution  $v_h$  that dominates  $u$  at points near the boundary, for any  $h$ , and then use the comparison principle and pass to the limit in  $h$ . The parameter  $h$  is needed to force  $v_h$  to be infinity at the boundary of the simplex. Since the simplex has corners, the distance to the boundary is not a smooth function, so the first step is to construct a nice test function that goes to 0 as  $x$  approaches the boundary. Roughly speaking, we consider the product of the distances to the faces of the simplex, and then take its logarithm.

**Step 1.** Let  $\rho_i(x)$ , for  $x \in \text{Int}(\widehat{S}_d)$ , be the distance of  $x$  from the hyperplane  $\{y \in \mathbb{R}^{d-1} : y_i = 0\}$ , for  $i \in \llbracket d-1 \rrbracket$ , and  $\rho_d(x)$  be the distance from  $\{y \in \mathbb{R}^{d-1} : \sum_{l=1}^{d-1} y_l = 1\}$ . Specifically, for  $x \in \text{Int}(\widehat{S}_d)$ , we have

$$\rho_i(x) = \begin{cases} x_i & i \in \llbracket d-1 \rrbracket, \\ x^{-d}/\sqrt{d-1} & i = d, \end{cases}$$

where we recall that  $x^{-d} = 1 - \sum_{l \in \llbracket d-1 \rrbracket} x_l$ . Clearly,  $\rho_i \in \mathcal{C}^\infty(\widehat{S}_d)$  and the derivatives, for  $j \in \llbracket d-1 \rrbracket$ , are

$$\partial_{x_j} \rho_i(x) = \begin{cases} \delta_{i,j} & i \in \llbracket d-1 \rrbracket, \\ -1/\sqrt{d-1} & i = d. \end{cases}$$

Let us denote, for  $x \in \text{Int}(\widehat{S}_d)$  and  $z \in \mathbb{R}^{d-1}$ , the Hamiltonian in (3.3)

$$\widehat{\mathcal{H}}(t, x, w) = \sum_{i=1}^{d-1} x_i \widehat{H}^i(t, x, w) + x^{-d} \widehat{H}^d(t, x, w), \quad (3.9)$$

where  $\widehat{H}^i$ , for  $i \in \llbracket d \rrbracket$ , are defined by (3.4); note that  $\widehat{\mathcal{H}}$  is convex in  $w$ .

**Step 2.** For any  $h > 0$ , let

$$v_h(t, x) := v(t, x) - h^2 \sum_{i \in \llbracket d \rrbracket} \log(\rho_i(x)) + h(T - t), \quad (t, x) \in [0, T] \times \text{Int}(\widehat{S}_d).$$

We claim that  $v_h$  is a viscosity supersolution of (3.3) on  $\text{Int}(\widehat{S}_d)$ . Let then  $\varphi \in \mathcal{C}^1([0, T] \times \text{Int}(\widehat{S}_d))$ , and  $(\bar{t}, \bar{x}) \in [0, T] \times \text{Int}(\widehat{S}_d)$  be a local minimum of  $v_h - \varphi$  on  $[0, T] \times \text{Int}(\widehat{S}_d)$ . Since  $v$  is a viscosity supersolution of (3.3) on  $\text{Int}(\widehat{S}_d)$ , considering the test function  $\varphi_h \in \mathcal{C}^1([0, T] \times \text{Int}(\widehat{S}_d))$  given by  $\varphi_h(t, x) = \varphi(t, x) + h^2 \sum_{i \in \llbracket d \rrbracket} \log(\rho_i(x)) - h(T - t)$ , we get that  $(\bar{t}, \bar{x})$  is a local minimum of  $v - \varphi_h$  and thus

$$-\partial_t \varphi_h(\bar{t}, \bar{x}) + \widehat{\mathcal{H}}(\bar{t}, \bar{x}, D_x \varphi_h(\bar{t}, \bar{x})) \geq 0,$$

that is,

$$0 \leq -\partial_t \varphi(\bar{t}, \bar{x}) - h + \widehat{\mathcal{H}}\left(\bar{t}, \bar{x}, D_x \varphi(\bar{t}, \bar{x}) + h^2 \sum_{i \in \llbracket d \rrbracket} \frac{D \rho_i(\bar{x})}{\rho_i(\bar{x})}\right). \quad (3.10)$$

We denote  $w = D_x \varphi(\bar{t}, \bar{x})$ ,  $\check{w}^i = (w_1 - w_i, \dots, w_{d-1} - w_i, -w_i) \in \mathbb{R}^d$  for  $i \in \llbracket d-1 \rrbracket$  and  $\check{w}^d = (w, 0)$ ,  $y_j = D \rho_j(\bar{x})$  for  $j \in \llbracket d \rrbracket$  and similarly  $\check{y}_j^i$ , and  $\varpi^i = \check{w}^i + h^2 \sum_{j \in \llbracket d \rrbracket} \frac{\check{y}_j^i}{\rho_j(\bar{x})}$ , for  $i \in \llbracket d \rrbracket$ . We apply the following property, which is an immediate consequence of the definition of the Hamiltonian in (2.8):

$$H^i(t, m, z + \zeta) - H^i(t, m, z) \geq -\langle Q_{i, \bullet}(t, a^*(t, i, m, z), m), \zeta \rangle \quad (3.11)$$

for any  $z, \zeta \in \mathbb{R}^d$  with  $z_i = \zeta_i = 0$ ,  $i \in \llbracket d \rrbracket$ ,  $m \in \text{Int}(S_d)$ ,  $t \in [0, T]$ , where  $a^*(t, i, m, z) \in A$  is an argmax of (2.7), which might not be unique. (This is equivalent to say that  $-Q_{i, \bullet}(t, a^*(t, i, m, z), m)$  belongs to the subdifferential of the convex function  $H^i(t, m, z)$ .)

Choosing a maximizer  $a^*(\bar{t}, i, \bar{x}, \varpi^i)$  for any  $i \in \llbracket d \rrbracket$ , applying (3.11) and (3.4), we obtain (we omit the dependence on  $(\bar{t}, \bar{x})$ )

$$\begin{aligned} & \widehat{\mathcal{H}}\left(w + h^2 \sum_{j \in \llbracket d \rrbracket} \frac{y_j}{\rho_j(\bar{x})}\right) \\ &= \sum_{i \in \llbracket d-1 \rrbracket} \bar{x}_i H^i\left(\check{w}^i + h^2 \sum_{j \in \llbracket d \rrbracket} \frac{\check{y}_j^i}{\rho_j(\bar{x})}\right) + \bar{x}^{-d} H^d\left(\check{w}^d + h^2 \sum_{j \in \llbracket d \rrbracket} \frac{\check{y}_j^d}{\rho_j(\bar{x})}\right) \\ &\leq \sum_{i \in \llbracket d-1 \rrbracket} \bar{x}_i H^i(\check{w}^i) + \bar{x}^{-d} H^d(\check{w}^d) \\ &\quad - h^2 \sum_{i \in \llbracket d-1 \rrbracket} \bar{x}_i \sum_{j \in \llbracket d \rrbracket} \left\langle Q_{i, \bullet}(a^*(i, \varpi^i)), \frac{\check{y}_j^i}{\rho_j(\bar{x})} \right\rangle - h^2 \bar{x}^{-d} \sum_{j \in \llbracket d \rrbracket} \left\langle Q_{d, \bullet}(a^*(d, \varpi^d)), \frac{\check{y}_j^d}{\rho_j(\bar{x})} \right\rangle \\ &= \widehat{\mathcal{H}}(w) - h^2 \sum_{i \in \llbracket d-1 \rrbracket} \bar{x}_i \sum_{j \in \llbracket d-1 \rrbracket} \frac{1}{\rho_j(\bar{x})} \left[ \sum_{k \in \llbracket d-1 \rrbracket} Q_{i, k}(a^*(i, \varpi^i))(\delta_{k, j} - \delta_{i, j}) + Q_{i, d}(a^*(i, \varpi^i))(-\delta_{j, i}) \right] \\ &\quad - h^2 \sum_{i \in \llbracket d-1 \rrbracket} \bar{x}_i \frac{1}{\rho_d(\bar{x})} \left[ \sum_{k \in \llbracket d-1 \rrbracket} Q_{i, k}(a^*(i, \varpi^i))(-1/\sqrt{d-1} + 1/\sqrt{d-1}) + Q_{i, d}(a^*(i, \varpi^i))1/\sqrt{d-1} \right] \\ &\quad - h^2 \bar{x}^{-d} \sum_{j \in \llbracket d-1 \rrbracket} \frac{1}{\rho_j(\bar{x})} \sum_{k \in \llbracket d-1 \rrbracket} Q_{d, k}(a^*(d, \varpi^d))\delta_{k, j} - h^2 \bar{x}^{-d} \frac{1}{\rho_d(\bar{x})} \sum_{k \in \llbracket d-1 \rrbracket} Q_{d, k}(a^*(d, \varpi^d))(-1/\sqrt{d-1}) \\ &\leq \widehat{\mathcal{H}}(w) + h^2 \sum_{i \in \llbracket d-1 \rrbracket} \bar{x}_i \frac{1}{\rho_i(\bar{x})} M(d-1) + h^2 \frac{\bar{x}^{-d}}{\rho_d(\bar{x})\sqrt{d-1}} M(d-1) \\ &= \widehat{\mathcal{H}}(w) + h^2 M d(d-1), \end{aligned}$$

where we used the bound  $0 \leq Q_{i, j} \leq M$ , for any  $i \neq j \in \llbracket d \rrbracket$ , and the definition of  $\rho_i$  in the last two lines. The latter inequality, applied in (3.10), gives

$$-\partial_t \varphi(\bar{t}, \bar{x}) + \widehat{\mathcal{H}}(\bar{t}, \bar{x}, D_x \varphi(\bar{t}, \bar{x})) \geq h - h^2 M d(d-1) \geq 0 \quad \text{if } h \leq \frac{1}{M d(d-1)},$$

which implies that  $v_h$  is a viscosity supersolution of (3.3) on  $\text{Int}(\widehat{S}_d)$  if  $h$  is small enough.

**Step 3.** As  $\rho_i \leq 1$ , we have  $v_h(t, x) \geq v(t, x)$  for any  $(t, x) \in [0, T] \times \text{Int}(\widehat{S}_d)$ . In particular,  $v_h(T, x) \geq v(T, x) \geq u(T, x)$  for any  $x \in \text{Int}(\widehat{S}_d)$ . We denote  $\rho(x) = \prod_{i=1}^d \rho_i(x)$ . Since  $u$  and  $v$  are bounded, we find that for any  $h > 0$  there exists  $\eta > 0$  (which may depend on  $h$ ) such that  $-h^2 \log \rho(x) \geq \|u\|_\infty + \|v\|_\infty$  if  $\rho(x) \leq \eta$ . We denote  $\Gamma^\eta = \{x \in \widehat{S}_d : \rho(x) = \eta\}$ ,  $\mathcal{O}^\eta = \{x \in \widehat{S}_d : \rho(x) \geq \eta\}$ , and  $\mathcal{O}_c^\eta = \{x \in \widehat{S}_d : \rho(x) \leq \eta\}$ ; note that  $\mathcal{O}^\eta$  is a smooth domain. Thus  $v_h(t, x) \geq u(t, x)$  for any  $t \in [0, T]$  and  $x \in \mathcal{O}_c^\eta$ , in particular for any  $x \in \Gamma^\eta$ . Therefore we can apply the comparison principle (see [24], Thm. II.9.1) in  $[0, T] \times \mathcal{O}^\eta$ , because  $u, v_h \in \mathcal{C}([0, T] \times \mathcal{O}^\eta)$ : we obtain  $u \leq v_h$  on  $[0, T] \times \mathcal{O}^\eta$  and hence  $u \leq v_h$  on the entire  $[0, T] \times \widehat{S}_d$ , since we already have  $u \leq v_h$  on  $[0, T] \times \mathcal{O}_c^\eta$ . Finally, we obtain  $u \leq v$  on  $[0, T] \times \text{Int}(\widehat{S}_d)$  by sending  $h$  to 0, as  $\lim_{h \rightarrow 0} v_h(t, x) = v(t, x)$  for any  $(t, x) \in [0, T] \times \text{Int}(\widehat{S}_d)$ , and then the inequality  $u \leq v$  can be extended up to the boundary of  $\widehat{S}_d$  by continuity.  $\square$

### 3.2. Classical solution

Here we give the sufficient condition for the value function of the MFPCP to belong to  $\mathcal{C}^{1,1}([0, T] \times S^d)$ . This is the the convexity assumption C: we prove hence point (3) in Theorem 2.9.

**Theorem 3.5.** *Under Assumption C, the value function is in  $\mathcal{C}^{1,1}([0, T] \times S^d)$ , and thus it is the unique classical solution to (2.33) (it solves the equation also at boundary points).*

We recall that a function  $v : [0, T] \times S_d \rightarrow \mathbb{R}$  is called *semiconcave*, resp. *semiconvex*, on  $[0, T] \times \text{Int}(S_d)$ , with a constant  $c$ , if

$$\frac{v(t+h, m+p) - 2v(t, x) + v(t-h, m-p)}{|h|^2 + |p|^2} \leq c, \quad \text{resp. } \geq -c, \quad (3.12)$$

for any  $t \in [0, T]$ ,  $m \in \text{Int}(S_d)$ ,  $h$  with  $t \pm h \in [0, T]$ , and  $p$  with  $m \pm p \in \text{Int}(S_d)$ .

*Proof.* We show that the value function is both semiconcave and semiconvex, in time and space, globally in  $\text{Int}(S^d)$ , with a constant  $c$ ; clearly it is equivalent to prove this propeties either for  $V$  defined on  $S_d$  or for  $\widehat{V}$  defined on  $\widehat{S}_d$ . Recall again that  $\text{Int}(S_d)$  is invariant for dynamics (2.32). Hence Corollary 3.3.8 in [7] ensures that  $V \in \mathcal{C}^{1,1}([0, T] \times \text{Int}(S^d))$  and the Lipschitz constant of  $D^i V$  is  $c$ ,  $\forall i \in \llbracket d \rrbracket$ . Thus in particular  $V$  can be extended uniquely to a function in  $\mathcal{C}^{1,1}([0, T] \times S^d)$ , and then it solves (2.33) also at boundary points. The classical solution to (2.33) is unique because any solution is the value function, by a standard application of the verification theorem.

The value function is semiconcave on  $[0, T] \times \text{Int}(S_d)$  by Theorem 7.4.11 in [7], thanks to Assumption (C1); as above, to apply this result set on a Euclidean space, we have just to consider the equivalent formulation of the control problem on  $\text{Int}(\widehat{S}_d)$ . To prove that  $V$  is semiconvex, we rewrite the MFPCP in an equivalent formulation, with a control  $w$ , such that the cost is convex in  $(m, w)$  and the dynamics is linear in  $(m, w)$ . Consider hence the problem of minimizing the cost <sup>2</sup>

$$\tilde{J}(w) := \int_0^T \sum_{i \in \llbracket d \rrbracket} \mu_t^i f^i \left( t, \left( \frac{w_t^{i,j}}{\mu_t^i} \right)_{j \neq i}, \mu_t \right) dt + \sum_{i \in \llbracket d \rrbracket} g^i(\mu_T) \mu_T^i,$$

where the couple  $(\mu, w)$  satisfies the ODE

$$\frac{d}{dt} \mu_t^i = \sum_j (w_t^{j,i} - w_t^{i,j})$$

<sup>2</sup>This reformulation of the MFPCP with the control  $w$  is typically used for studying potential mean field games; see e.g. [6, 10].

and is subject to the constraints  $w_t^{i,j} \geq 0$ ,  $\frac{w_t^{i,j}}{\mu_t^i} \leq M$ . If the initial condition  $\mu_0 \in \text{Int}(S^d)$ , then  $\mu_t \in \text{Int}(S^d)$  for any  $t \in [0, T]$ , i.e.  $\mu_t^i > 0$ , thus the cost is well-defined in this case. This control problem is indeed well-defined only for  $\mu_0 \in \text{Int}(S^d)$  and it is seen to be equivalent to the MFCP (2.32)–(2.31) by setting  $w^{i,j} = \mu^i \alpha^{i,j}$ , meaning that the value function is the same. The advantage in using this new formulation is that the dynamics is now linear in  $w$  and the set of  $(\mu, w)$  satisfying the constraints is convex. Moreover, the running cost is convex in  $(m, w)$  and the terminal cost is convex in  $m$  by Assumption (C2). Thus we can apply ([7], Thm. 7.4.13), which says that  $V(t, m)$  is a convex function of  $m \in \text{Int}(S^d)$ , for any  $t \in [0, T]$ . Then  $V$  is semiconvex in time and space, in  $[0, T] \times \text{Int}(S^d)$ , again by ([7], Thm. 7.4.13), but using the original formulation of the MFCP in the proof therein, for which the coefficients are globally Lipschitz in  $\text{Int}(S^d)$ , yielding thus global semiconvexity, while in the new formulation the cost is only locally Lipschitz in  $m \in \text{Int}(S^d)$ .  $\square$

### 3.3. Further properties and potential mean field games

We collect, for reference, other results concerning the mean field control problem and its relation, in some cases, with a mean field game. These are not used here, but might be useful for future works. They derive directly from the results of Section 7.4 of [7] about deterministic control problems in a Euclidean space. Thus we consider the problem (3.1)–(3.2) defined on  $\hat{S}_d$  whose value function is denoted by  $\hat{V}$ . As before, for a function  $G$  defined on  $S_d$ , we denote by  $\hat{G}$  its version in local chart, i.e.  $\hat{G}(x) = G(\check{x})$ .

**Proposition 3.6.** *Assume A and that  $F(t, a, \cdot), Q_{i,j}(t, a, \cdot), G \in \mathcal{C}^1(S_d)$  and  $\hat{\mathcal{H}}(t, \cdot, \cdot) \in \mathcal{C}^{1,1}(\hat{S}_d \times \mathbb{R}^{d-1})$ , for the Hamiltonian  $\hat{\mathcal{H}}$  defined by (3.9). If  $\alpha$  is an optimal control and  $x$  the corresponding optimal trajectory, then there exists  $w \in \mathcal{C}^1(0, T; \mathbb{R}^{d-1})$  such that*

$$\frac{d}{dt} w_t^j = \partial_{x_j} \hat{\mathcal{H}}(t, x_t, w_t), \quad w_T^j = \partial_{x_j} \hat{G}(x_T); \quad (3.13)$$

and  $w_t$  belongs to the space superdifferential of  $\hat{V}(t, x_t)$  for any time.

Moreover, if  $\hat{\mathcal{H}}(t, x, \cdot)$  is strictly convex for any  $t \in [0, T]$  and  $x \in \text{Int}(\hat{S}_d)$ , and the costs  $F$  and  $G$  are semiconcave w.r.t.  $m$ , then, assuming that the control problem starts at  $(t_0, x_0) \in [0, T] \times \text{Int}(\hat{S}_d)$ ,

- $\hat{V}$  is differentiable in  $(t, x_t)$  for any  $t \in (t_0, T]$ , for any optimal trajectory  $x$ ;
- $\hat{V}$  is differentiable in  $(t_0, x_0)$  if and only if there exists a unique optimal trajectory  $x$ ; in such case the adjoint process  $w$  satisfies  $w_t^j = \partial_{x_j} \hat{V}(t, x_t)$  for any  $t \in [t_0, T]$ .

The Hamiltonian is strictly convex w.r.t.  $w$ , for  $x$  in the interior, in case e.g. the running cost is given by (2.11), as observed in [15]. We recall that the value function is (time-space) Lipschitz continuous, and thus differentiable almost everywhere for the 1-dim Lebesgue measure in time and the  $(d-1)$ -dim. Lebesgue measure in space. Further,  $\hat{V}$  is shown to be (time-space) semiconcave in the proof of Theorem 3.5, in case only B and (C1) hold. The first assertion is instead the Pontryagin principle, which holds also under weaker assumptions.

In case the cost splits as

$$f^i(t, a, m) = \ell^i(t, a) + f_0^i(t, m), \quad (3.14)$$

the Hamiltonian  $H$  in (2.8) splits as  $H^i(t, m, z) = H_0^i(t, a) - f_0^i(t, m)$ , and then (3.13) becomes

$$-\frac{d}{dt} w_t^j + \hat{H}_0^j(t, x_t, w_t) - \hat{H}_0^d(t, x_t, w_t) = \partial_{x_j} \hat{F}_0(t, x_t), \quad w_T = \partial_{x_j} \hat{G}(x_T), \quad j \in \llbracket d-1 \rrbracket, \quad (3.15)$$

where  $F_0(m) = \sum_{i \in \llbracket d \rrbracket} m_i f_0^i(m)$  and  $\widehat{H}$  is defined as in (3.4). The above equation is in fact equivalent to the HJB equation in the MFG system, first analyzed in [26], which is

$$-\frac{d}{dt}u_t^i + H_0^i(t, \mu_t, (u_t^j - u_t^i)_{j \in \llbracket d \rrbracket}) = \mathfrak{f}^i(t, \mu_t), \quad w_T^j = \mathfrak{g}_t^j(m_T), \quad i \in \llbracket d \rrbracket, \quad (3.16)$$

for a given running cost  $\mathfrak{f}$  and terminal cost  $\mathfrak{g}$ , in case B holds, so that the transition rate in the coupled equation for  $\mu$  in (2.32) are given by  $\alpha_t^{i,j} = a_j^*(t, i, (u_t^j - u_t^i)_{j \in \llbracket d \rrbracket})$ . Equivalence holds if the functions  $\mathfrak{f}$  and  $\mathfrak{g}$  are such that

$$\widehat{\mathfrak{f}}^i(t, x) - \widehat{\mathfrak{f}}^d(t, x) = \partial_{x_i} \widehat{F}_0(t, x), \quad \widehat{\mathfrak{g}}^i(t, x) - \widehat{\mathfrak{g}}^d(t, x) = \partial_{x_i} \widehat{G}(t, x), \quad i \in \llbracket d-1 \rrbracket; \quad (3.17)$$

for instance, the latter holds true by defining  $\widehat{\mathfrak{f}}^i(t, x) = \partial_{x_i} \widehat{F}_0(t, x)$  and  $\widehat{\mathfrak{f}}^d(t, x) = 0$ . It is easy to see that (3.15) and (3.16) are equivalent: given a solution  $u$  to (3.16), it suffices to let  $w^j = u^j - u^d$ ,  $j \in \llbracket d-1 \rrbracket$ ; conversely, given  $w$  solution to (3.15), it suffices to solve (3.16) where all the occurrences of  $u^i - u^j$  have been replaced by  $w^i - w^j$  if  $j \in \llbracket d-1 \rrbracket$  and by  $w^i$  if  $j = d$ . We remark that (3.13) can not be interpreted as a mean field game if the cost does not split as in (3.14). In case the cost splits, we have shown that, if an optimal control for the MFCP exists, then it gives rise to a solution of a particular mean field game –with costs determined by (3.17)– and therefore this can provide more informations on the optimal control and on the corresponding optimal trajectory of the MFCP.

In general, a mean field game, in which hence the costs  $(\mathfrak{f}^i)_{i \in \llbracket d \rrbracket}$  and  $(\mathfrak{g}^i)_{i \in \llbracket d \rrbracket}$  are given, is said to be potential if (3.17) holds and the cost  $f_0^i$  and  $g^i$  defining the MFCP do not depend on  $i$  –say they are equal to  $f_0$  and  $g$ – so that  $F_0 = f_0$  and  $G = g$ . Thus the mean field game system represents the necessary conditions for optimality of the deterministic MFCP and we refer to [15] for a detailed study of potential mean field games and corresponding MFCP, in particular for the interpretation of (3.17).

## 4. CONVERGENCE RESULTS

We prove here the main convergence results: Theorems 2.10 and 2.13. Throughout this section,  $V^N$  denotes the value function of the mean field  $N$ -agent optimization (2.19)–(2.20), while  $V$  denotes the value function of the mean field control problem (2.31)–(2.32). We recall that  $V^N$  is the classical solution to ODE (2.22), while  $V$  is the viscosity solution to PDE (2.33).

### 4.1. Convergence of value functions

Here, we prove Theorem 2.10. Assume hence that Assumption A is in force. We exploit here the characterization of  $V$  as the unique viscosity solution to (2.33) on the closed set  $S_d$ ; see Definition 3.1 and Proposition 3.2.

We first need to show that  $V^N$  is time-space Lipschitz-continuous, uniformly in  $N$ ; this is (2.23) in Proposition 2.6. In this point, the compactness of the control set  $A$  is required (Assumption (A1)).

**Lemma 4.1.** *If A holds, then for every  $t, s \in [0, T]$  and  $m, p \in S_d^N$*

$$|V^N(t, m) - V^N(s, p)| \leq C(|t - s| + |m - p|), \quad (4.1)$$

for a constant  $C$  independent of  $N$ .

*Proof.* We represent the dynamics of  $\mu^N$ , given by (2.20), as an SDE with respect to a Poisson random measure, as we did in [16]. We restrict attention to controls  $\alpha : [0, T] \rightarrow A^d$  that are just functions of time and of the private state  $i \in \llbracket d \rrbracket$ . Recall that, by Remark 2.7, the value function over this smaller class is the same  $V^N$ .

Fix  $N \in \mathbb{N}$  and denote by  $\mathcal{N}(dt, d\theta)$  a standard Poisson random measure on  $[0, T] \times [0, M]^{d \times d}$ , with intensity measure  $\nu$  on  $[0, M]^{d \times d}$ , where  $M$  is the maximum of  $Q$  (which is continuous over a compact set). Let  $\nu$  be

defined as the sum of the measures of the intersections with the axes, *i.e.*  $\nu(E) = \sum_{i,j \in \llbracket d \rrbracket} \text{Leb}(E \cap \Theta_{i,j})$ , where  $\Theta_{i,j} := \{\theta \in [0, M]^{d \times d} : \theta_{i',j'} = 0 \text{ for all } (i',j') \neq (i,j)\}$ , which is viewed as a subset of  $\mathbb{R}$ , as  $E \cap \Theta_{i,j}$  is, and  $\text{Leb}$  is the Lebesgue measure on  $\mathbb{R}$ . Hence, the dynamics of  $\mu^N$  is written as

$$d\mu_t^N = \int_{[0,M]^{d \times d}} \frac{1}{N} (\delta_j - \delta_i) \mathbb{1}_{(0, N\mu_{i,t}^N Q_{i,j}(t, \alpha^i(t), \mu_t^N))} \mathcal{N}(dt, d\theta). \quad (4.2)$$

This equation is well-posed because  $\mu^N$  takes values in  $S_d^N$ , which is finite, and then (2.20) follows by equation (2.34) of [16]. Observe that  $\frac{1}{N}(\delta_j - \delta_i)$  represents the increment of  $\mu^N$ , while the indicator function gives the transition rate.

Let  $\mu^N$  start in  $m$  and  $\rho^N$  start in  $p$ , at a fixed time  $t \in [0, T]$ , with the same control  $\alpha$ . Let  $\varepsilon > 0$  and  $\alpha : [t, T] \rightarrow A^d$  be an  $\varepsilon$ -optimal control for the problem starting at  $(t, m)$ , *i.e.*

$$J^N(t, m, \alpha) \leq \inf_{\beta} J^N(t, m, \beta) + \varepsilon = V^N(t, m) + \varepsilon,$$

with an obvious notation for  $J^N(t, m, \alpha)$ . Then, by Lemma 3 of [16], it follows that

$$\begin{aligned} \mathbb{E}|\mu_s^N - \rho_s^N| &\leq |m - p| + \frac{1}{N} \sum_{i,j} |\delta_j - \delta_i| \mathbb{E} \int_t^s |N\mu_{i,r}^N Q_{i,j}(r, \alpha^i(r), \mu_r^N) - N\rho_{i,t}^N Q_{i,j}(r, \alpha^i(r), \rho_r^N)| dr \\ &\leq |m - p| + C \int_t^s \mathbb{E}|\mu_r^N - \rho_r^N| dr, \end{aligned}$$

by using the Lipschitz-continuity of  $Q$ , and thus Gronwall's lemma yields

$$\sup_{s \in [t, T]} \mathbb{E}|\mu_s^N - \rho_s^N| \leq C|m - p|. \quad (4.3)$$

Then

$$\begin{aligned} V^N(t, p) - V^N(t, m) &\leq J^N(t, p, \alpha) - J^N(t, m, \alpha) + \varepsilon \\ &\leq \mathbb{E} \int_t^T |F(s, \alpha(s), \mu_s^N) - F(s, \alpha(s), \rho_s^N)| ds + \mathbb{E}|G(\mu_T^N) - G(\rho_T^N)| + \varepsilon \\ &\leq C \sup_{s \in [t, T]} \mathbb{E}|\mu_s^N - \rho_s^N| + \varepsilon \leq C|m - p| + \varepsilon, \end{aligned}$$

where we applied the Lipschitz-continuity of  $F$  and  $G$  (defined by (2.3)) and (4.3). Taking the limit as  $\varepsilon$  vanishes, we get  $V^N(t, p) - V^N(t, m) \leq C|m - p|$  and then, changing the role of  $m$  and  $p$  we obtain also the opposite inequality, which provides

$$|V^N(t, p) - V^N(t, m)| \leq C|m - p|. \quad (4.4)$$

To prove the Lipschitz-continuity in time, note that (4.4) implies  $|D^{N,i}V^N(t, m)| \leq C$  for any  $t \in [0, T]$  and  $m \in S_d^N$ , and thus, recalling that  $V^N$  is  $\mathcal{C}^1$  in time, from the HJB equation (2.22) we derive that for any  $t$  and  $m$ ,

$$\left| \frac{d}{dt} V^N(t, m) \right| = \left| \sum_{i \in \llbracket d \rrbracket} m_i H^i(t, m, D^{N,i}V^N(t, m)) \right|$$

$$\leq C(1 + |D^{N,i}V^N(t, m)|) \leq C,$$

where we used the Lipschitz-continuity of  $H$  w.r.t.  $z$ . Therefore

$$\sup_{t \in [0, T], m \in S_d^N} \left| \frac{d}{dt} V^N(t, m) \right| \leq C,$$

which, together with (4.4), yields (4.1).  $\square$

We now turn to the proof of Theorem 2.10. As a first step to prove the convergence, it is required to extend the definition of  $V^N$  outside  $S_d^N$ . One way to do this, similarly to [8], would be to consider the same control problem (2.19)–(2.20), but starting at any point in the simplex, not only on  $S_d^N$ . However, in this way, the dynamics of  $\mu^N$  would go also outside the simplex, and thus we prefer to not follow this strategy. Instead, we use a piecewise constant interpolation of  $V^N$ , and thus we have to pay a price in order to define it carefully at boundary points, so that the maximum in the usual doubling of variable argument (see (4.6) below) is attained.

*Proof of Theorem 2.10.*

We first prove that

$$E_N^+ := \sup_{t \in [0, T], m \in S_d^N} (V(t, m) - V^N(t, m)) \leq \frac{C}{\sqrt{N}}. \quad (4.5)$$

If  $E_N^+ \leq 0$  then (4.5) trivially holds, thus we assume that  $E_N^+ > 0$ . Since  $V^N$  is defined on the grid  $S_d^N$  only, we have to construct a piecewise constant extension  $\tilde{V}^N$  defined on the whole  $S_d$ . Denote by  $n(N, d)$  the number of elements in  $S_d^N$ , and cover the simplex by closed cells  $(\Gamma_k)_{k=1}^{n(N, d)}$  centered in points of  $S_d^N$  and invariant by translations. Note that the cells centered at points on the boundary cover also points outside the simplex. Define  $\tilde{V}^N(p) = V^N(p_k)$  if  $p \in \text{Int}(\Gamma_k)$  and  $p_k$  is the center of  $\Gamma_k$ . It remains to define  $\tilde{V}^N$  at the boundaries of  $\Gamma_k$ .

**Step 1.** We exploit the usual argument of doubling the variables, which prompts us to consider the function, to be defined on  $[0, T]^2 \times S_d^2$ ,

$$\Phi(t, s, m, p) := V(t, m) - \tilde{V}^N(s, p) - \frac{|t - s|^2}{2\varepsilon} - \frac{|m - p|^2}{2\varepsilon} - \frac{2T - t - s}{4T} E_N^+, \quad (4.6)$$

where  $\varepsilon$  is a parameter to be fixed later in terms of  $N$ . Then the value of  $\tilde{V}^N$  at the boundaries of the cells has to be chosen such that the above function admits a maximum. Let us give first the idea of our construction. The strategy is to define first  $\tilde{V}^N$  constant in any closed cell, then perform the maximization in any closed cell –in which a maximum of  $\Phi$  exists– and thus take the maximum of the values obtained, so that a maximum point for  $\Phi$  exists. If this maximum point lies in the interior of a cell, then there is no problem and the value of  $\tilde{V}^N$  at the boundaries does not matter. The critical situation is when the maximum point belongs to the boundary of a cell. In such situation, we have to define carefully  $\tilde{V}^N$  at the boundary of the neighboring cells in order to verify equality (4.7) below. It is required because we want to exploit the ODE (2.22), and this is indeed the main reason for considering a piecewise constant interpolation. We give below an example of our construction in case  $d = 2$ , the generalization to  $d > 2$  being not difficult.

Now, more precisely, for a cell  $\Gamma_k$  centered at  $p_k \in S_d^N$ ,  $k \in \{1, \dots, n(N, d)\}$ , let  $\tilde{V}_k^N(p) := V^N(p_k)$  for any  $p \in \Gamma_k$ , thus also on the boundary of  $\Gamma_k$ . Then define  $\Phi_k$  as is (4.6), but with  $p \in \Gamma_k$  and  $\tilde{V}^N$  therein replaced by  $\tilde{V}_k^N$ , and let

$$\gamma_k := \max_{[0, T]^2 \times S_d \times (\Gamma_k \cap S_d)} \Phi_k(t, s, m, p).$$

Such a maximum exists and is attained at a point  $(\bar{t}_k, \bar{s}_k, \bar{m}_k, \bar{p}_k)$ , which might be non-unique and such that  $\bar{p}_k$  belongs to the boundary of the cell  $\Gamma_k$ : we then let

$$\gamma = \max_{k=1, \dots, n(N, d)} \gamma_k =: \gamma_{\bar{k}}, \quad (\bar{t}, \bar{s}, \bar{m}, \bar{p}) := (\bar{t}_{\bar{k}}, \bar{s}_{\bar{k}}, \bar{m}_{\bar{k}}, \bar{p}_{\bar{k}}).$$

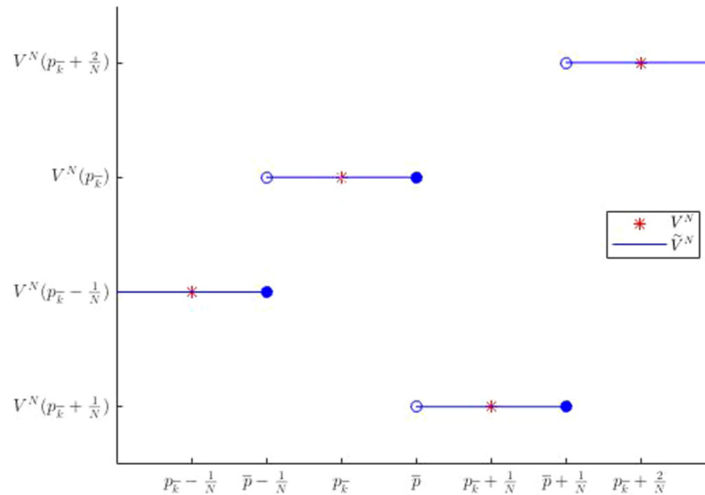
We can now define  $\tilde{V}^N$  such that  $\tilde{V}^N(\bar{s}, \bar{p}_{\bar{k}}) = \tilde{V}^N(\bar{p}) := \tilde{V}_{\bar{k}}^N(\bar{p}) = V^N(\bar{s}, p_{\bar{k}})$ , the last equality holding by the definition above, where we recall  $p_{\bar{k}} \in S_d^N$  is the center of the cell  $\Gamma_{\bar{k}}$ . Note that  $\bar{p} = \bar{p}_{\bar{k}}$  may belong to the boundary of  $\Gamma_{\bar{k}}$ , in which case we have defined  $\tilde{V}^N$  at the boundary of  $\Gamma_{\bar{k}}$ . In addition, we require that

$$V^N\left(\bar{s}, p_{\bar{k}} + \frac{1}{N}(\delta_j - \delta_i)\right) = \tilde{V}^N\left(\bar{s}, \bar{p} + \frac{1}{N}(\delta_j - \delta_i)\right) \quad (4.7)$$

for any  $i, j \in \llbracket d \rrbracket$ . Note that this is automatically satisfied if  $\bar{p} \in \text{Int}(S_d)$ , while it gives the definition of  $\tilde{V}^N$  on a part of the boundary of the cells bordering  $\Gamma_{\bar{k}}$ , if  $\bar{p}$  belongs to the boundary of  $\Gamma_{\bar{k}}$ . Lastly,  $\tilde{V}^N$  can be defined arbitrarily right or left continuous at the other boundary points.

Hence this construction guarantees that  $\Phi$  in (4.6) admits a maximum on  $[0, T]^2 \times S_d^2$  at  $(\bar{t}, \bar{s}, \bar{m}, \bar{p})$ . We stress that  $\tilde{V}^N$  is defined also outside the simplex, so that the RHS of (4.7) is meaningful in case  $p_{\bar{k}} + \frac{1}{N}(\delta_j - \delta_i)$  belongs to the boundary of the simplex. Note that  $\tilde{V}^N$  is not continuous in space, but it remains Lipschitz in time, uniformly in space and w.r.t.  $N$ .

Before proceeding with the rest of the proof, let us give an example of the above construction in case  $d = 2$ . In this case the simplex is one dimensional and, if projected on  $[0, 1]$ , we have  $S_2^N = \{\frac{k}{N} : k = 0, 1, \dots, N\}$ , the increments are  $\pm \frac{1}{N}$  and the cells are  $\Gamma_k = [\frac{k}{N} - \frac{1}{2N}, \frac{k}{N} + \frac{1}{2N}]$ , for  $k = 0, \dots, N$ . Let us assume that we are in the critical situation, that is,  $\bar{p}$  belongs to the boundary of  $\Gamma_{\bar{k}}$ ; for instance, we can assume that  $\bar{p} = \bar{p}_{\bar{k}} = \frac{\bar{k}}{N} + \frac{1}{2N}$ . The following picture provides then the construction of  $\tilde{V}^N$  (where we omit the time dependence):



Namely, if  $\tilde{V}^N$  is left continuous at  $\bar{p}$  then it is defined to be left continuous also at  $\bar{p} \pm \frac{1}{N}$ , so that  $V^N(p_{\bar{k}} \pm \frac{1}{N}) = \tilde{V}^N(\bar{p} \pm \frac{1}{N})$ . This latter property is crucial in the proof, see (4.10) below. We remark again that  $\tilde{V}^N$  can be define arbitrarily right or left continuous at the other boundary points. It is not difficult to generalize this construction to  $d > 2$ .



**Step 2.** The inequality  $\Phi(\bar{t}, \bar{s}, \bar{m}, \bar{p}) \geq \Phi(\bar{t}, \bar{s}, \bar{p}, \bar{p})$  gives

$$\frac{|\bar{m} - \bar{p}|^2}{2\varepsilon} \leq V(\bar{t}, \bar{m}) - V(\bar{t}, \bar{p}),$$

which, using the Lipschitz continuity of  $V$ , yields

$$|\bar{m} - \bar{p}| \leq C\varepsilon. \quad (4.8)$$

Similarly, the inequality  $\Phi(\bar{t}, \bar{s}, \bar{m}, \bar{p}) \geq \Phi(\bar{s}, \bar{s}, \bar{m}, \bar{p})$  provides

$$\frac{|\bar{t} - \bar{s}|^2}{2\varepsilon} \leq V(\bar{t}, \bar{m}) - V(\bar{s}, \bar{m}) + \frac{\bar{t} - \bar{s}}{4T} E_N^+,$$

which, as  $V$  is Lipschitz-continuous and  $E_N^+$  is bounded, gives

$$|\bar{t} - \bar{s}| \leq C\varepsilon. \quad (4.9)$$

We stress that (4.8) and (4.9) hold also in case  $(\bar{t}, \bar{s}, \bar{m}, \bar{p})$  belong to the (time and space) boundary of  $[0, T]^2 \times S_d^2$ . We fix now  $\varepsilon = \frac{1}{\sqrt{N}}$ .

**Step 3.** In order to prove (4.5), we consider the three cases: either  $\bar{t} = T$ , or  $\bar{s} = T$ , or  $\bar{t}, \bar{s} < T$ .

**First case:**  $\bar{t} = T$  and  $\bar{s} \in [0, T]$ . The inequality  $\Phi(\bar{t}, \bar{s}, \bar{m}, \bar{p}) \geq \Phi(t, t, m, m)$ , for any  $(t, m) \in [0, T] \times S_d^N$ , exploiting the Lipschitz-continuity in time of  $\tilde{V}^N$  and in  $S_d$  of  $G$ , gives

$$\begin{aligned} V(t, m) - \tilde{V}^N(t, m) &\leq G(\bar{m}) - G(\bar{p}) + G(\bar{p}) - \tilde{V}^N(\bar{s}, \bar{p}) + \frac{2T - 2t}{4T} E_N^+ \\ &\leq C|\bar{m} - \bar{p}| + C|T - \bar{s}| + \frac{1}{2} E_N^+. \end{aligned}$$

Taking the supremum of the l.h.s. and applying (4.8) and (4.9), we have

$$E_N^+ \leq C\varepsilon + \frac{1}{2} E_N^+ \leq \frac{C}{\sqrt{N}} + \frac{1}{2} E_N^+,$$

which yields (4.5).

**Second case:**  $\bar{s} = T$  and  $\bar{t} \in [0, T]$ . This can be treated as the first case, by using the Lipschitz-continuity in time of  $V$ .

**Third case:**  $\bar{t} \in [0, T)$  and  $\bar{s} \in [0, T)$ . Let  $p_{\bar{k}}$  be, as above, the point in  $S_d^N$  such that  $\tilde{V}^N(\bar{s}, p_{\bar{k}}) = \tilde{V}^N(\bar{s}, \bar{p})$ . Here we use the piecewise constant construction of  $\tilde{V}^N$  and in particular (4.7). Hence, from (2.22) we obtain

$$-\frac{d}{dt} \tilde{V}^N(\bar{s}, \bar{p}) - \sum_i p_{\bar{k}}^i \left\{ \langle Q_{i, \bullet}(\bar{s}, a_i^*, p_{\bar{k}}), D^{N, i} \tilde{V}^N(\bar{s}, \bar{p}) \rangle + f^i(\bar{s}, a_i^*, p_{\bar{k}}) \right\} = 0, \quad (4.10)$$

where  $a_i^* = a_i^*(\bar{s}, p_{\bar{k}})$  is a point in  $A$  which attains the maximum in (2.7). We recall that  $\tilde{V}^N$  is defined also in case  $\bar{p} + \frac{1}{N}(\delta_j - \delta_i)$  is outside the simplex, which could happen if  $p_{\bar{k}} + \frac{1}{N}(\delta_j - \delta_i)$  belongs to the boundary; while  $p_{\bar{k}} + \frac{1}{N}(\delta_j - \delta_i)$  can be outside  $S_d$  only if  $p_{\bar{k}}^i = 0$ , in which case the term in (4.10) is zero.

Since  $\tilde{V}^N - \varphi$  has a minimum at  $(\bar{s}, \bar{p})$ , where

$$\varphi(s, p) = V(\bar{t}, \bar{m}) - \frac{|s - \bar{t}|^2}{2\varepsilon} - \frac{|p - \bar{m}|^2}{2\varepsilon} - \frac{2T - s - \bar{t}}{4T} E_N^+,$$

we get  $\frac{d}{dt}\tilde{V}^N(\bar{s}, \bar{p}) \geq \partial_t \varphi(\bar{s}, \bar{p})$ , with equality if  $\bar{s} \neq 0$ , and thus

$$\frac{\bar{s} - \bar{t}}{\varepsilon} - \frac{1}{4T} E_N^+ - \sum_i p_k^i \left\{ \langle Q_{i, \bullet}(\bar{s}, a_i^*, p_k^-), D^{N,i} V^N(\bar{s}, \bar{p}) \rangle + f^i(\bar{s}, a_i^*, p_k^-) \right\} \geq 0. \quad (4.11)$$

On the other hand, as  $V - \psi$  has a maximum at  $(\bar{t}, \bar{m})$ , where

$$\psi(t, m) = \tilde{V}^N(\bar{s}, \bar{p}) + \frac{|\bar{s} - t|^2}{2\varepsilon} + \frac{|\bar{p} - m|^2}{2\varepsilon} + \frac{2T - \bar{s} - t}{4T} E_N^+,$$

since  $V$  is a viscosity subsolution of (2.33) on the entire  $S_d$  (see again Def. 3.1) and  $\psi$  is indeed defined on  $\mathbb{R}^d$ , we get

$$-\frac{\bar{t} - \bar{s}}{\varepsilon} + \frac{1}{4T} E_N^+ + \sum_i \bar{m}_i H^i \left( \bar{t}, \bar{m}, \left( \frac{\langle \bar{m} - \bar{p}, \delta_j - \delta_i \rangle}{\varepsilon} \right)_{j \in \llbracket d \rrbracket} \right) \leq 0,$$

which, by definition of the Hamiltonian, gives

$$\begin{aligned} & -\frac{\bar{t} - \bar{s}}{\varepsilon} + \frac{1}{4T} E_N^+ \\ & - \sum_i \bar{m}_i \left\{ \left\langle Q_{i, \bullet}(\bar{t}, a_i^*, \bar{m}), \left( \frac{\langle \bar{m} - \bar{p}, \delta_j - \delta_i \rangle}{\varepsilon} \right)_{j \in \llbracket d \rrbracket} \right\rangle + f^i(\bar{t}, a_i^*, \bar{m}) \right\} \leq 0. \end{aligned} \quad (4.12)$$

The inequality  $\Phi(\bar{t}, \bar{s}, \bar{m}, \bar{p}) \geq \Phi(\bar{t}, \bar{s}, \bar{m}, \bar{p} + \frac{1}{N}(\delta_j - \delta_i))$  gives, for any  $i$  and  $j$  in  $\llbracket d \rrbracket$ ,

$$\begin{aligned} \tilde{V}^N(\bar{s}, \bar{p} + \frac{1}{N}(\delta_j - \delta_i)) & \geq \tilde{V}^N(\bar{s}, \bar{p}) + \frac{|\bar{m} - \bar{p}|^2}{2\varepsilon} - \frac{|\bar{m} - \bar{p} - \frac{1}{N}(\delta_j - \delta_i)|^2}{2\varepsilon} \\ & \geq \tilde{V}^N(\bar{s}, \bar{p}) + \frac{1}{N\varepsilon} \left\langle \delta_j - \delta_i, \bar{m} - \bar{p} - \frac{1}{2N}(\delta_j - \delta_i) \right\rangle, \end{aligned}$$

which, applied in (4.11), yields, as  $Q_{i,j} \geq 0$ ,

$$\frac{\bar{s} - \bar{t}}{\varepsilon} - \frac{1}{4T} E_N^+ \geq \sum_i p_k^i \left\{ \left\langle Q_{i, \bullet}(\bar{s}, a_i^*, p_k^-), \left( \frac{\langle \bar{m} - \bar{p}, \delta_j - \delta_i \rangle}{\varepsilon} - \frac{1}{N\varepsilon} \right)_{j \in \llbracket d \rrbracket} \right\rangle + f^i(\bar{s}, a_i^*, p_k^-) \right\}. \quad (4.13)$$

Summing (4.12) and (4.13), using Assumption A (*i.e.* the boundedness of  $Q$  and the Lipschitz-continuity of  $Q$  and  $F$  w.r.t.  $(t, m)$ ), we obtain

$$\begin{aligned} \frac{2}{4T} E_N^+ & \leq \sum_i \bar{m}_i \left\{ \left\langle Q_{i, \bullet}(\bar{s}, a_i^*, \bar{m}), \left( \frac{\langle \bar{m} - \bar{p}, \delta_j - \delta_i \rangle}{\varepsilon} \right)_{j \in \llbracket d \rrbracket} \right\rangle + f^i(\bar{s}, a_i^*, \bar{m}) \right\} \\ & \quad - \sum_i p_k^i \left\{ \left\langle Q_{i, \bullet}(\bar{t}, a_i^*, p_k^-), \left( \frac{\langle \bar{m} - \bar{p}, \delta_j - \delta_i \rangle}{\varepsilon} - \frac{1}{N\varepsilon} \right)_{j \in \llbracket d \rrbracket} \right\rangle + f^i(\bar{t}, a_i^*, p_k^-) \right\} \\ & \leq C|\bar{t} - \bar{s}| + C|\bar{m} - p_k^-| + C \frac{|\bar{m} - \bar{p}|^2}{2\varepsilon} + \frac{C}{N\varepsilon}, \end{aligned}$$

which provides, using (4.8), (4.9) and that  $|\bar{p} - p_{\bar{\epsilon}}| \leq \frac{1}{N}$  by construction,

$$E_N^+ \leq C \left( |\bar{t} - \bar{s}| + |\bar{m} - \bar{p}| + \frac{1}{N} + \frac{|\bar{m} - \bar{p}|^2}{2\varepsilon} + \frac{1}{N\varepsilon} \right) \leq C \left( 3\varepsilon + \frac{1}{N} + \frac{1}{N\varepsilon} \right) \leq \frac{C}{\sqrt{N}},$$

recalling that  $\varepsilon = \frac{1}{\sqrt{N}}$ .

**Step 4.** The opposite inequality,

$$E_N^- := \sup_{t \in [0, T], m \in S_d^N} (V^N(t, m) - V(t, m)) \leq \frac{C}{\sqrt{N}}, \quad (4.14)$$

can be proved in the same way, by changing the roles of  $\tilde{V}^N$  and  $V$ , and using instead the viscosity supersolution property of  $V$ . Finally, (4.5) and (4.14) give (2.35).  $\square$

**Remark 4.2.** Similarly to [8], see also ([1], Sect. VI.1), it might be possible to establish a stronger convergence rate of order  $1/N$ , if  $V^N$  is semiconcave uniformly in  $N$ :

$$V^N(t, m + p) - 2V^N(t, m) + V^N(t, m - p) \leq c|p|^2,$$

for a constant  $c$  independent of  $N$ . Such estimate should hold if the costs  $F$  and  $G$  are semiconcave in space. This is left to future work.

As a consequence, we can also prove Theorem 2.11. In the proof as well as in the next section, we use several times the following basic estimate: if  $\xi = (\xi_1, \dots, \xi_N)$  is a vector of  $N$  i.i.d. random variables with values in  $\llbracket d \rrbracket$ , such that  $\text{Law}(\xi_1) = m \in S_d$ , then

$$\mathbb{E}|\mu_{\xi}^N - m| \leq \frac{1}{\sqrt{N}} \frac{\sqrt{d}}{2}. \quad (4.15)$$

Indeed,  $N\mu_{\xi}^N = (N\mu_{\xi}^N[i])_{i \in \llbracket d \rrbracket}$  has a multinomial distribution with parameters  $N$  and  $m$ , and thus the variance is  $\mathbb{E}|N\mu_{\xi}^N[i] - Nm_i|^2 = Nm_i(1 - m_i)$ . Hence (4.15) follows by Cauchy-Schwarz inequality and recalling that we are using the Euclidean norm.

*Proof of Theorem 2.11.* Let  $\alpha^N(t, i, m)$  be an optimal feedback control for the  $N$ -agent optimization and  $\alpha(t, i)$  be  $\varepsilon$ -optimal for the MFCP, *i.e.*

$$V(0, m_0) \leq J(\alpha) \leq V(0, m_0) + \varepsilon. \quad (4.16)$$

Recall that the dynamics start at time 0 and the initial condition  $\mathbf{X}_0$  is assumed to be i.i.d. with  $\text{Law}(X_0^1) = m_0$ . Since  $J^N(\alpha^N) = \mathbb{E}[V^N(0, \mu_0^N)]$ , we have

$$\begin{aligned} 0 &\leq J^N(\alpha) - J^N(\alpha^N) \\ &\leq |J^N(\alpha) - J(\alpha)| + |J(\alpha) - V(0, m_0)| + |V(0, m_0) - \mathbb{E}[V(0, \mu_0^N)]| + |\mathbb{E}[V(0, \mu_0^N)] - \mathbb{E}[V^N(0, \mu_0^N)]| \\ &\leq |J^N(\alpha) - J(\alpha)| + \varepsilon + C\mathbb{E}|\mu_0^N - m_0| + \frac{C}{\sqrt{N}}, \end{aligned}$$

where we used (2.35), (4.16), and the Lipschitz-continuity of  $V$ . As  $\mu_0^N$  is the empirical measure of an i.i.d sequence, (4.15) permits to bound  $\mathbb{E}|\mu_0^N - m_0|$ . Thus, in order to prove (2.37), it remains to show

$$|J^N(\alpha) - J(\alpha)| \leq \frac{C}{\sqrt{N}}. \quad (4.17)$$

This follows by standard arguments in propagation of chaos.

Let  $\tilde{\rho}^N$  be the process given by dynamics (2.20), when using the decentralized control  $\alpha(t, i)$ . Assumption (A3) gives (denoting  $\alpha = (\alpha^1, \dots, \alpha^d)$  and using definition (2.3))

$$\begin{aligned} |J^N(\alpha) - J(\alpha)| &\leq \mathbb{E} \int_0^T |F(t, \alpha(t), \tilde{\rho}_t^N) - F(t, \alpha(t), \mu_t)| dt + \mathbb{E} |G(\tilde{\rho}_t^N) - G(\mu_t)| \\ &\leq C \sup_{t \in [0, T]} \mathbb{E} |\tilde{\rho}_t^N - \mu_t|. \end{aligned}$$

Hence, to prove (4.17), we have to show

$$\sup_{t \in [0, T]} \mathbb{E} |\tilde{\rho}_t^N - \mu_t| \leq \frac{C}{\sqrt{N}}. \quad (4.18)$$

We use the SDE representation of the dynamics introduced in (4.2), thus  $\tilde{\rho}^N$  solves the SDE

$$d\tilde{\rho}_t^N = \int_{[0, M]^{d \times d}} \frac{1}{N} (\delta_j - \delta_i) \mathbb{1}_{(0, N\tilde{\rho}_{i,t}^N Q_{i,j}(t, \alpha^i(t), \tilde{\rho}_t^N))} \mathcal{N}(dt, d\theta).$$

This process is not the empirical measure of a vector of independent processes because of the dependence of  $Q$  on  $m$ , thus we introduce also the process  $\tilde{\mu}^N$  in which the empirical measure is replaced by the limit deterministic flow  $\mu$ :  $\tilde{\mu}^N$  solves

$$d\tilde{\mu}_t^N = \int_{[0, M]^{d \times d}} \frac{1}{N} (\delta_j - \delta_i) \mathbb{1}_{(0, N\mu_{i,t} Q_{i,j}(t, \alpha^i(t), \mu_t))} \mathcal{N}(dt, d\theta).$$

and represents the empirical measure associated to  $N$  i.i.d copies of the limit process  $\tilde{X}$  satisfying (2.29) and such that  $\text{Law}(\tilde{X}_t) = \mu_t$ . Again (4.15) gives

$$\sup_{t \in [0, T]} \mathbb{E} |\tilde{\mu}_t^N - \mu_t| \leq \frac{C}{\sqrt{N}}. \quad (4.19)$$

Thus it remains to estimate the distance between  $\tilde{\rho}^N$  and  $\tilde{\mu}^N$ . As in the proof of Lemma 4.1, the representation in terms of SDEs yields, by exploiting the Lipschitz-continuity of  $Q$  w.r.t. the variable  $m$ ,

$$\begin{aligned} \mathbb{E} |\tilde{\rho}_t^N - \tilde{\mu}_t^N| &\leq \frac{1}{N} \sum_{i,j} |\delta_j - \delta_i| \mathbb{E} \int_0^t |N\tilde{\rho}_{i,s}^N Q_{i,j}(s, \alpha^i(s), \tilde{\rho}_s^N) - N\mu_{i,t} Q_{i,j}(s, \alpha^i(s), \mu_s)| ds \\ &\leq C \int_0^t \mathbb{E} |\tilde{\rho}_s^N - \mu_s| ds \leq C \int_0^t \mathbb{E} |\tilde{\rho}_s^N - \tilde{\mu}_s^N| ds + C \sup_{t \in [0, T]} \mathbb{E} |\tilde{\mu}_t^N - \mu_t| \\ &\leq C \int_0^t \mathbb{E} |\tilde{\rho}_s^N - \tilde{\mu}_s^N| ds + \frac{C}{\sqrt{N}}. \end{aligned}$$

Therefore Gronwall's inequality gives  $\sup_{t \in [0, T]} \mathbb{E} |\tilde{\rho}_t^N - \tilde{\mu}_t^N| \leq \frac{C}{\sqrt{N}}$ , which, together with (4.19), provides (4.18) that concludes the proof.  $\square$

## 4.2. Propagation of chaos

Here we prove Theorem 2.13. Throughout this subsection, we hence assume that  $V \in \mathcal{C}^{1,1}([0, T] \times S_d)$  and that Assumptions B and (C1) are in force.

We first show the following simple result.

**Lemma 4.3.** *There exists a constant  $C$  such that for any  $i \in \llbracket d \rrbracket$*

$$\max_{t \in [0, T], m \in S_d^N} |D^{N,i}V(t, m) - D^iV(t, m)| \leq \frac{C}{N}. \quad (4.20)$$

*Proof.* By definition, for each  $i, j \in \llbracket d \rrbracket$  and any  $m \in S_d^N$  (we omit the time in the notation), using Taylor's formula and the Lipschitz-continuity of  $D^iV$ , we obtain

$$\begin{aligned} |D^{N,i}V(m) - D^iV(m)| &= \left| N \left[ V \left( m + \frac{1}{N}(\delta_j - \delta_i) \right) - V(m) \right] - \partial_{m_j - m_i} V(m) \right| \\ &= \left| \int_0^1 \partial_{m_j - m_i} V \left( m + \frac{1}{N}s(\delta_j - \delta_i) \right) ds - \partial_{m_j - m_i} V(m) \right| \\ &\leq \frac{C}{N} \int_0^1 s |\delta_j - \delta_i| ds \leq \frac{C}{N}. \end{aligned}$$

$\square$

As in the statement of Theorem 2.13, let  $\alpha^N$  be the unique optimal feedback control for the  $N$ -agent optimization defined by (2.28), and  $\mu^N$  be the corresponding optimal process satisfying (2.27). Also, let  $\alpha_*$  be the unique optimal feedback control for the MFCP defined by (2.34) and  $\mu$  the corresponding optimal trajectory given by (2.32). We stress that  $\alpha_N$  and  $\alpha_*$  are functions of  $t$  and  $m$ . Since Assumption (C1) is in force, as explained in Remark 2.12, we can assume here that the convergence of the value functions holds with a stronger rate, given by (2.38).

As an intermediate step, we consider the process  $\rho^N$  satisfying (2.27) with the limiting feedback control  $\alpha_*$ ; such intermediate process is needed to prove convergence and describes the empirical measure of a standard mean field interacting particle system, in the sense that the transition rate function  $\alpha_*$  is the same for any  $N$  (and depends on the empirical distribution  $\rho^N$ ), while in  $\mu^N$  the rate  $\alpha^N$  depends on  $N$ . Thus, we first show the proximity of  $\mu^N$  and  $\rho^N$  and then prove convergence of  $\rho^N$  to  $\mu$ . We assume that  $\rho^N$  and  $\mu^N$  start at time zero in the same point in  $S_d^N$ . The proof of the following proposition is the point where assumption (B3) is required.

**Proposition 4.4.** *We have*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\mu_s^N - \rho_s^N| \right] \leq \frac{C}{\sqrt{N}}. \quad (4.21)$$

*Proof. Step 1.* We compute the limit value function  $V$  along  $\mu^N$ . Dynkin formula and then the HJB equation (2.33) give

$$\begin{aligned} &\mathbb{E}[V(T, \mu_T^N)] - \mathbb{E}[V(0, \mu_0^N)] \\ &= \mathbb{E} \left[ \int_0^T \left( \partial_t V(t, \mu_t^N) + N \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N \alpha_N^{i,j}(t, \mu_t^N) \left[ V \left( t, \mu_t^N + \frac{1}{N}(\delta_j - \delta_i) \right) - V(t, \mu_t^N) \right] \right) dt \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[ \int_0^T \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N \left( H^i(t, \mu_t^N, D^i V(t, \mu_t^N)) + \langle \alpha_N^i(t, \mu_t^N), D^{N,i} V(t, \mu_t^N) \rangle \right) dt \right] \\
&= \mathbb{E} \left[ \int_0^T \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N \left( \mathcal{H}^i(t, a^*(t, i, \mu_t^N), D^i V(t, \mu_t^N)), \mu_t^N, D^i V(t, \mu_t^N) \right) - \mathcal{H}^i(t, \alpha_N^i(t, \mu_t^N), \mu_t^N, D^i V(t, \mu_t^N)) \right. \\
&\quad \left. - f^i(t, \alpha_N^i(t, \mu_t^N), \mu_t^N) + \langle \alpha_N^i(t, \mu_t^N), D^{N,i} V(t, \mu_t^N) - D^i V(t, \mu_t^N) \rangle \right) dt \right].
\end{aligned}$$

Since  $a^*(t, i, m, z)$  maximizes the pre-Hamiltonian  $\mathcal{H}^i(t, m, z)$  in (2.7), using (2.9), we obtain, for any  $a \in [0, M]^d$ ,

$$\mathcal{H}^i(t, a^*(t, i, m, z), m, z) - \mathcal{H}^i(t, a, m, z) \geq \lambda |a^*(t, i, m, z) - a|^2, \quad (4.22)$$

This inequality, together with (4.20) and the uniform boundedness of  $\alpha^N$ , yields

$$\begin{aligned}
&\mathbb{E} \left[ \int_0^T \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N f^i(t, \alpha_N^i(t, \mu_t^N), \mu_t^N) dt + \sum_{i \in \llbracket d \rrbracket} g^i(\mu_t^N) \mu_{i,T}^N \right] - \mathbb{E}[V(0, \mu_0^N)] \\
&\geq \mathbb{E} \left[ \int_0^T \lambda \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N |\alpha_N^i(t, \mu_t^N) - a^*(t, i, \mu_t^N, D^i V(t, \mu_t^N))|^2 dt \right] - \frac{C}{N}.
\end{aligned}$$

Observe that on the l.h.s. there is exactly  $J^N(\alpha_N)$ , as defined by (2.19), which is equal to  $\mathbb{E}[V^N(0, \mu_0^N)]$ . Therefore we obtain

$$\mathbb{E} \left[ \int_0^T \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N |\alpha_N^i(t, \mu_t^N) - a^*(t, i, \mu_t^N, D^i V(t, \mu_t^N))|^2 dt \right] \leq \frac{C}{N} + C \max_{m \in S_d^N} |V^N(0, m) - V(0, m)|,$$

which, applying (2.38), gives

$$\mathbb{E} \left[ \int_0^T \sum_{i \in \llbracket d \rrbracket} \mu_{i,t}^N |\alpha_N^i(t, \mu_t^N) - a^*(t, i, \mu_t^N, D^i V(t, \mu_t^N))|^2 dt \right] \leq \frac{C}{N}. \quad (4.23)$$

**Step 2.** Considering now the process  $\rho^N$ , and applying the SDE representation as in the proof of Lemma 4.1, we obtain for any  $t > 0$

$$\begin{aligned}
\varphi(t) &:= \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\mu_s^N - \rho_s^N| \right] \leq \mathbb{E} \left[ \int_0^t \sum_{i,j} \frac{1}{N} (\delta_j - \delta_i) \left| N \mu_{i,s}^N \alpha_N^{i,j}(s, \mu_s^N) - N \rho_{i,s}^N \alpha_*^{i,j}(s, \rho_s^N) \right| ds \right] \\
&\leq 2 \mathbb{E} \left[ \int_0^t \sum_{i,j} \left( \mu_{i,s}^N \left| \alpha_N^{i,j}(s, \mu_s^N) - \alpha_*^{i,j}(s, \mu_s^N) \right| + \mu_{i,s}^N \left| \alpha_*^{i,j}(s, \mu_s^N) - \alpha_*^{i,j}(s, \rho_s^N) \right| \right. \right. \\
&\quad \left. \left. + \alpha_*^{i,j}(s, \rho_s^N) |\mu_{i,s}^N - \rho_{i,s}^N| \right) ds \right]
\end{aligned}$$

and applying Jensen's inequality, the boundedness and Lipschitz-continuity of the feedback function  $\alpha_*$  (verified as  $V$  is  $\mathcal{C}^{1,1}$  and  $\alpha_*$  is given by (2.34)), and then (4.23), we get

$$\begin{aligned}\varphi(t) &\leq C \sqrt{\mathbb{E} \left[ \int_0^t \sum_{i \in \llbracket d \rrbracket} \mu_{i,s}^N |\alpha_i^N(s, \mu_s^N) - a^*(t, i, \mu_s^N, D^i V(t, \mu_s^N))|^2 ds \right]} + C \mathbb{E} \left[ \int_0^t |\mu_s^N - \rho_s^N| ds \right] \\ &\leq \frac{C}{\sqrt{N}} + C \int_0^t \varphi(s) ds,\end{aligned}$$

which, by Gronwall's lemma, provides (4.21).  $\square$

We introduce also the process  $\mathbf{Y}$  related to the empirical measure  $\rho^N$ , in which the transition rate function is given in terms of the limiting optimal feedback  $\alpha_*(t, i, m)$ , independent of  $N$ . In order to prove the propagation of chaos result (Thm. 2.13) we first show the proximity of  $\mathbf{X}$  and  $\mathbf{Y}$  and then prove propagation of chaos for the process  $\mathbf{Y}$ , which is a standard mean field interacting particle system. The proof is pretty standard and the arguments are the same that we used in [17], based on a probabilistic representation of the dynamics, thus we are not going to give all the details below. We recall that the initial conditions are fixed and i.i.d.. The following is the counterpart of Proposition 4.4 for the  $N$  trajectories.

**Lemma 4.5.** *For any  $N \in \mathbb{N}$  and  $k \in \llbracket N \rrbracket$ , it holds*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_s^k - Y_s^k| \right] \leq \frac{C}{\sqrt{N}}. \quad (4.24)$$

*Proof.* By exploiting the representation of  $\mathbf{X}$  and  $\mathbf{Y}$  in terms of SDEs driven by Poisson random measures (similar for the representation for  $\mu^N$  used above, see [17] for the details), we obtain

$$\begin{aligned}\varphi(t) &:= \mathbb{E} \left[ \sup_{s \in [0, t]} |X_s^k - Y_s^k| \right] \leq C \mathbb{E} \int_0^t (|\alpha^N(s, X_s^k, \mu_s^N) - \alpha_*(s, Y_s^k, \rho_s^N)| + |X_s^k - Y_s^k|) ds \\ &\leq C \mathbb{E} \int_0^t (|\alpha^N(s, X_s^k, \mu_s^N) - \alpha_*(s, X_s^k, \rho_s^N)| + |X_s^k - Y_s^k|) ds,\end{aligned}$$

where we used that any bounded function is Lipschitz with respect to  $i \in \llbracket d \rrbracket$  finite. Using the exchangeability of the processes  $(\mathbf{X}, \mathbf{Y})$ , we can rewrite

$$\begin{aligned}\mathbb{E} \int_0^t |\alpha^N(s, X_s^k, \mu_s^N) - \alpha_*(s, X_s^k, \rho_s^N)| ds &= \frac{1}{N} \sum_{l=0}^N \mathbb{E} \int_0^t |\alpha^N(s, X_s^l, \mu_s^N) - \alpha_*(s, X_s^l, \rho_s^N)| ds \\ &= \frac{1}{N} \sum_{l=0}^N \mathbb{E} \int_0^t \sum_{i=1}^d \mathbb{1}_{\{X_s^l = i\}} |\alpha^N(s, i, \mu_s^N) - \alpha_*(s, i, \rho_s^N)| ds \\ &= \mathbb{E} \int_0^t \sum_{i=1}^d \mu_{i,s}^N |\alpha^N(s, i, \mu_s^N) - \alpha_*(s, i, \rho_s^N)| ds \leq \frac{C}{\sqrt{N}},\end{aligned}$$

the latter bound following from the proof of the previous proposition. Therefore we derive

$$\varphi(t) \leq \frac{C}{\sqrt{N}} + \int_0^t \varphi(s) ds,$$

which gives (4.24) by Gronwall's lemma.  $\square$

We are finally in the position to prove Theorem 2.13. Recall that  $\widetilde{\mathbf{X}}$  is the decoupled process, made by independent copies of the limit dynamics, in which player  $\widetilde{X}^k$  uses the decentralized strategy  $\alpha_*(t, \widetilde{X}_t^k, \mu(t))$ , where  $\alpha_*$  is the optimal control and  $\mu$  is the optimal trajectory for the MFCP, and we have  $\text{Law}(\widetilde{X}_t^k) = \mu(t)$  for any  $t \in [0, T]$  and  $k \in \llbracket N \rrbracket$ .

*Proof of Theorem 2.13.* Thanks to (4.24), claim (2.40) is proved if we show that

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |Y_t^k - \widetilde{X}_t^k| \right] \leq \frac{C}{\sqrt{N}}. \quad (4.25)$$

Let  $\widetilde{\mu}_t^N$  be the empirical measure related to the i.i.d. process  $\widetilde{\mathbf{X}}$ . Again (4.15) gives

$$\sup_{t \in [0, T]} \mathbb{E} |\widetilde{\mu}_t^N - \mu_t| \leq \frac{C}{\sqrt{N}}. \quad (4.26)$$

The limit feedback  $\alpha_*$  is Lipschitz w.r.t.  $m$  because so is  $DV$ . We use this fact and also the inequality

$$|\mu_{\mathbf{x}}^N - \mu_{\mathbf{y}}^N| \leq C \frac{1}{N} \sum_{k=1}^N |x_k - y_k| \quad (4.27)$$

which follows from the definition of the 1-Wasserstein distance, which is equivalent to the Euclidean distance in finite dimension. As in the proof of the previous lemma, we have

$$\begin{aligned} \varphi(t) &:= \mathbb{E} \left[ \sup_{s \in [0, t]} |Y_s^k - \widetilde{X}_s^k| \right] \\ &\leq \mathbb{E} \left[ \int_0^t |\alpha_*(s, Y_s^k, \rho_s^N) - \alpha_*(s, Y_s^k, \mu_s)| \, ds + C \int_0^t |Y_s^k - \widetilde{X}_s^k| \, ds \right] \\ &\leq C \mathbb{E} \int_0^t |\rho_s^N - \mu_s| \, ds + C \mathbb{E} \int_0^t |Y_s^k - \widetilde{X}_s^k| \, ds \\ &\leq C \mathbb{E} \int_0^t |\rho_s^N - \widetilde{\mu}_s^N| \, ds + C \sup_{t \in [0, T]} \mathbb{E} |\widetilde{\mu}_t^N - \mu_t| + C \mathbb{E} \int_0^t |Y_s^k - \widetilde{X}_s^k| \, ds \\ &\leq \frac{1}{N} \sum_{l=1}^N C \mathbb{E} \int_0^t |Y_s^l - \widetilde{X}_s^l| \, ds + \frac{C}{\sqrt{N}} + C \int_0^t \varphi(s) \, ds \\ &\leq \frac{C}{\sqrt{N}} + C \int_0^t \varphi(s) \, ds, \end{aligned}$$

where in the last inequality we used the exchangeability of the processes. Thus Gronwall's inequality yields (4.25).

Finally, observe that, if (C1) is not in force –and thus just (2.35) is satisfied and not (2.38)– then (4.23) holds with  $N$  replaced by  $\sqrt{N}$  and hence, as a consequence of the above proofs, estimates (4.21) and (4.25) hold with



$\sqrt{N}$  replaced by  $N^{1/4}$ . Thus (4.27) gives

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\rho_t^N - \tilde{\mu}_t^N| \right] \leq \frac{C}{N^{1/4}}.$$

Therefore (2.39) follows from (4.21) (with  $\sqrt{N}$  replaced by  $N^{1/4}$  therein), the above estimate, and the inequality

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\tilde{\mu}_t - m_t| \right] \leq CN^{-1/9},$$

which can be found in [37]. □

## REFERENCES

- [1] M. Bardi and I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhäuser, Boston (1997).
- [2] E. Bayraktar, A. Cecchin, A. Cohen and F. Delarue. Finite state mean field games with Wright-Fisher common noise as limits of  $N$ -player weighted games. Preprint 2020, [arXiv: 2012.04845](#).
- [3] E. Bayraktar and A. Cohen, Analysis of a finite state many player game using its master equation. *SIAM J. Control Optim.* **56** (2018) 3538–3568.
- [4] E. Bayraktar, A. Cosso and H. Pham, Randomized dynamic programming principle and Feynman-Kac representation for optimal control of McKean-Vlasov dynamics. *Trans. Am. Math. Soc.* **370** (2018) 2115–2160.
- [5] A. Bensoussan, J. Frehse and S. Yam, Mean field games and mean field type control theory. Vol. 101 of *Springer Briefs in mathematics*. Springer-Verlag New York (2013).
- [6] A. Briani and P. Cardaliaguet, Stable solutions in potential mean field game systems. *Nonlinear Differ. Equ. Appl.* **25** (2018) 1.
- [7] P. Cannarsa and C. Sinestrari, Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control. Birkhauser, Boston (2004).
- [8] I. Capuzzo-Dolcetta and H. Ishii, Approximate solutions of the Bellman equation of deterministic control theory. *Appl. Math. Optim.* **11** (1984) 161–181.
- [9] P. Cardaliaguet, F. Delarue, J.-M. Lasry and P.-L. Lions, The master equation and the convergence problem in mean field games. Vol. 201 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ (2019).
- [10] P. Cardaliaguet, J. Graber, A. Porretta and D. Tonon, Second order mean field games with degenerate diffusion and local coupling. *Nonlinear Differ. Equ. Appl.* **22** (2015) 1287–1317.
- [11] R. Carmona and F. Delarue, Forward-backward stochastic differential equations and controlled McKean Vlasov dynamics. *Ann. Probab.* **43** (2015) 2647–2700.
- [12] R. Carmona and F. Delarue, Probabilistic theory of mean field games with applications: Vol. I, mean field fbsdes, control, and games. Vol. 83 of *Probability Theory and Stochastic Modelling*. Springer (2018).
- [13] R. Carmona and F. Delarue, Probabilistic theory of mean field games with applications: Vol. II, mean field games with common noise and master equations. Vol. 84 of *Probability Theory and Stochastic Modelling*. Springer (2018).
- [14] A. Cecchin, P. Dai Pra, M. Fischer and G. Pelino, On the convergence problem in mean field games: a two state model without uniqueness. *SIAM J. Control Optim.* **57** (2019) 2443–2466.
- [15] A. Cecchin and F. Delarue, Selection by vanishing common noise for potential finite state mean field games. Preprint 2020, [arXiv: 2005.12153](#).
- [16] A. Cecchin and M. Fischer, Probabilistic approach to finite state mean field games. *Appl. Math. Optim.* **81** (2020) 253–300.
- [17] A. Cecchin and G. Pelino, Convergence, fluctuations and large deviations for finite state mean field games via the master equation. *Stoch. Process. Appl.* **129** (2019) 4510–4555.
- [18] M.G. Crandall and P.L. Lions, Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comput.* **43** (1984) 1–19.
- [19] F. Delarue and R. Foguen Tchuendom, Selection of equilibria in a linear quadratic mean-field game. *Stochastic Process. Appl.* **130** (2020) 1000–1040.
- [20] M.F. Djete, Extended mean field control problem: a propagation of chaos result. Preprint 2020. [arXiv: 2006.12996](#).
- [21] F.M. Djete, D. Possamai and X. Tan, McKean-Vlasov optimal control: the dynamic programming principle. Preprint 2020, [arXiv: 1907.08860](#).
- [22] F.M. Djete, D. Possamai and X. Tan, McKean-Vlasov optimal control: limit theory and equivalence between different formulations. Preprint 2020, [arXiv: 2001.00925](#).
- [23] M. Fischer, On the connection between symmetric  $N$ -player games and mean field games. *Ann. Appl. Prob.* **127** (2017) 757–810.

- [24] W.H. Fleming and H.M. Soner, Controlled Markov processes and viscosity solutions. Vol. 25 of *Applications of Mathematics*. Springer, New York, 2nd edition (2006).
- [25] M. Fornasier, S. Lisini, C. Orrieri and G. Savaré, Mean-field optimal control as Gamma-limit of finite agent controls. *Eur. J. Appl. Math.* **30** (2019) 1153–1186.
- [26] D.A. Gomes, J. Mohr and R.R. Souza, Continuous time finite state mean field games. *Appl. Math. Optim.* **68** (2013) 99–143.
- [27] V.N. Kolokoltsov, Nonlinear Markov Games on a Finite State Space (Mean-field and Binary Interactions). *Int. J. Stat. Prob.* **1** (2012) 77–91.
- [28] D. Lacker, A general characterization of the mean field limit for stochastic differential games. *Prob. Theory Related Fields* **165** (2016) 581–648.
- [29] D. Lacker, Limit theory for controlled McKean-Vlasov dynamics. *SIAM J. Control Optim.* **55** (2017) 1641–1672.
- [30] D. Lacker, On the convergence of closed-loop Nash equilibria to the mean field game limit. *Ann. Appl. Prob.* **30** (2020) 1693–1761.
- [31] M. Laurière and O. Pironneau, Dynamic programming for mean field type control. *Comptes Rendus Math.* **352** (2014) 707–713.
- [32] M. Laurière and L. Tangpi, Convergence of large population games to mean field games with interaction through the controls. Preprint 2020, [arXiv: 2004.08351](https://arxiv.org/abs/2004.08351).
- [33] M. Motte and H. Pham, Mean-field Markov decision processes with common noise and open-loop controls. Preprint 2019, [arXiv: 1912.07883](https://arxiv.org/abs/1912.07883).
- [34] H. Pham and X. Wei, Bellman equation and viscosity solutions for mean field stochastic control problem. *ESAIM: COCV* **24** (2018) 437–461.
- [35] H. Pham and X. Wei, Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics. *SIAM J. Control Optim.* **55** (2017) 1069–1101.
- [36] A. Porretta and M. Ricciardi, Mean field games under invariance conditions for the state space. *Commun. Partial Differ. Equ.* **45** (2020) 146–190.
- [37] S.T. Rachev and L. Rüschendorf, Mass Transportation Problems. Springer-Verlag (1998).
- [38] P.E. Souganidis, Approximation schemes for viscosity solutions of Hamilton-Jacobi equations. *J. Differ. Equ.* **59** (1985) 1–43.