NETWORKS AND HETEROGENEOUS MEDIA doi:10.3934/nhm.2012.7.263
©American Institute of Mathematical Sciences
Volume **7**, Number **2**, June **2012** pp. **263–277**

A SEMI-DISCRETE APPROXIMATION FOR A FIRST ORDER MEAN FIELD GAME PROBLEM

Fabio Camilli

"Sapienza", Università di Roma
Dipartimento di Scienze di Base e Applicate per l'Ingegneria
00161 Roma, Italy

Francisco Silva

"Sapienza", Università di Roma
Dipartimento di Matematica Guido Castelnuovo
00185 Rome, Italy

Abstract. In this article we consider a model first order mean field game problem, introduced by J.M. Lasry and P.L. Lions in [18]. Its solution (v, m) can be obtained as the limit of the solutions of the second order mean field game problems, when the *noise* parameter tends to zero (see [18]). We propose a semi-discrete in time approximation of the system and, under natural assumptions, we prove that it is well posed and that it converges to (v, m) when the discretization parameter tends to zero.

1. **Introduction.**

The aim of Mean Field Game theory (MFG), introduced by Lasry and Lions in [16, 17, 18], is to describe analytically when $N \uparrow \infty$ the behavior of the limit of Nash equilibria for N-players stochastic differential games. In the presence of stochastic dynamics of Itô type for the state of each player, the limit system is, at least formally, of the following form (see [18]):

$$- \partial_t v(t,x) - \sigma^2 \Delta v(x,t) + H(x, Dv(x,t))F(x, m(t)), \qquad \text{in } \mathbb{R}^d \times (0,T), \quad (1)$$

$$\partial_t m(t,x) - \sigma^2 \Delta m(x,t) - \operatorname{div}\big(m(t,x)\frac{\partial H}{\partial p}(x, Dv(t,x))\big) = 0, \quad \text{in } \mathbb{R}^d \times (0,T), \quad (2)$$

$$v(x,T) = G(x, m(T)) \text{ for } x \in \mathbb{R}^d, \quad m(0) = m_0 \in \mathcal{P}_1. \quad (3)$$

In the notation above we have: the *noise* parameter $\sigma \in \mathbb{R}$, the Hamiltonian H, which is convex with respect to the second variable, the space $\mathcal{P}_1$ of probability measures on $\mathbb{R}^d$ and the functions $F, G : \mathbb{R}^d \times \mathcal{P}_1 \to \mathbb{R}$, which can be non local operators. It is shown in [18] that, under natural assumptions and in an appropriate sense, system (1)-(3) admits a unique solution (v_σ, m_σ). Let us also point out that a discrete analogous model has been proposed in [13]. The authors prove its well-posedness as well as exponential convergence towards an equilibrium.

Regarding numerical methods, a finite-difference scheme for (1)-(3) is proposed in [2], as well as the corresponding scheme for the stationary version. See also [1]

2000 *Mathematics Subject Classification.* Primary: 91A13; Secondary: 65M25, 49L25.

Key words and phrases. Mean field games, optimal control, convex duality, numerical methods.

The support of the Chair "Modélisation Mathématique et Simulation Numérique", EADS-Polytechnique-INRIA and of the European Union under the "7th Framework Programme CFP7-PEOPLE-2010-ITNE Grant agreement number 264735-SADCO" are gratefully acknowledged.

where a Newton method is used to solve the discretized system in the context of the planning problem. A different approach is explored in [15], based on the fact that, for some particular cases (see [18]), system (1)-(3) characterizes the saddle points of an associated optimization problem. Let us also mention [14] where, in the case of a quadratic Hamiltonian, a change of variable allows to rewrite system (1)-(3) in a simpler manner which in turns permits to design a monotone scheme.

In this work we focus our attention on the *vanishing viscosity* limit of (1)-(3) as $\sigma \downarrow 0$, when $H(x,p) = \frac{1}{2}|p|^2$. This is interpreted as that the noise σ for each player tends to disappear. The limit equations are the so-called *first order mean field system* (see [18])

$$\begin{aligned} &-\partial_t v(t,x) + \tfrac{1}{2}|Dv(t,x)|^2 = F(x,m(t)), \qquad \text{in } \mathbb{R}^d \times (0,T), \\ &\partial_t m(t,x) - \operatorname{div}\big(Dv(t,x)m(t,x)\big) = 0, \qquad \text{in } \mathbb{R}^d \times (0,T), \\ &v(x,T) = G(x,m(T)) \quad \text{for } x \in \mathbb{R}^d, \qquad m(0) = m_0 \in \mathcal{P}_1. \end{aligned} \tag{4}$$

It is well known (see [18])) that, under natural assumptions, system (4) has a unique solution (v,m) and that $(v_\sigma, m_\sigma) \to (v,m)$ in appropriate spaces.

In this work we study a time discretization of (4), which is precisely described in section 4. This is motivated by a forthcoming work in which a fully discrete (i.e. where space discretization is added) semi-lagrangian method is studied, together with some numerical experiments. Our main results in this article is that the resulting discretization is well-posed (see theorem 4.2) and that, as the discretization parameter goes to zero, we have the convergence to the solution (v,m) of (4). This is the first result concerning the approximation of a first order MFG system.

The article is organized as follows: After setting the notation and introducing the fundamental assumptions in section 2, we analyze in section 3 a discrete in time approximation of the first and second equations in (4), separately. This is fundamental in order to prove in section 4 the well-posedness of the semi-discrete scheme, as well as its convergence. Finally, we provide in the appendix a proof of a technical result stated in section 3.

2. **Prerequisites.**

Let us first fix some notation. We denote by $\mathcal{P}_1$ the set of the Borel probability measures m such that $\int_{\mathbb{R}^d} |x| \mathrm{d}m(x) < \infty$. The set $\mathcal{P}_1$ is be endowed with the Kantorovich-Rubinstein distance

$$\mathbf{d}_1(\mu,\nu) = \sup\left\{ \int_{\mathbb{R}^d} \phi(x) \mathrm{d}[\mu-\nu](x) \; ; \; \phi : \mathbb{R}^d \to \mathbb{R} \text{ is 1-Lipschitz} \right\}. \tag{5}$$

Given a measure $\mu \in \mathcal{P}_1$ we denote by supp (μ) its support and given $x \in \mathbb{R}^{d'}$ ($d' \in \mathbb{N}$) we set x_i for its i-th coordinate. For a Lebesgue measurable set $A \subseteq \mathbb{R}^d$ we denote by $\mathcal{L}^d(A)$ its Lebesgue measure. In what follows, in order to simplify the notation, the operator D (resp. D^2) will denote the derivative (resp. the second derivative) with respect to the space variable $x \in \mathbb{R}^d$. We suppose that the functions $F, G : \mathbb{R}^d \times \mathcal{P}_1 \to \mathbb{R}$ and $m_0 \in \mathcal{P}_1$, which are the data of system (4), satisfy the following assumptions:

(H1) F and G are continuous over $\mathbb{R}^d \times \mathcal{P}_1$.

(H2) There exists a constant $C > 0$ such that for any $m \in \mathcal{P}_1$

$$\|F(\cdot,m)\|_{C^2} + \|G(\cdot,m)\|_{C^2} \le C.$$

(H3) The following monotonicity conditions holds true

$$\int_{\mathbb{R}^d} [F(x,m_1) - F(x,m_2)]\,\mathrm{d}[m_1 - m_2](x) > 0 \quad \text{for all } m_1,\ m_2 \in \mathcal{P}_1,\ m_1 \neq m_2, \tag{6}$$

$$\int_{\mathbb{R}^d} [G(x,m_1) - G(x,m_2)]\,\mathrm{d}[m_1 - m_2](x) > 0 \quad \text{for all } m_1,\ m_2 \in \mathcal{P}_1,\ m_1 \neq m_2. \tag{7}$$

(H4) The measure m_0, appearing in the third equation of (4), is absolutely continuous with respect to the Lebesgue measure. Its density, still denoted as m_0, is essentially bounded and satisfies supp $(m_0) \subset B(0,C)$.

Remark 1. Assumption (7) is slightly stronger than the corresponding assumption in [7].

Now we define what we mean by a solution of (4).

Definition 2.1. The pair $(v,m) \in W^{1,\infty}_{loc}(\mathbb{R}^d \times [0,T]) \times L^1(\mathbb{R}^d \times (0,T))$ is a solution of (4) if the first equation is satisfied in the viscosity sense, while the second one in sense of distributions.

The following existence and uniqueness result is proved in [18, 19, 7]

Theorem 2.2. *Under* **(H1)-(H4)** *there exists a unique solution* (v,m) *to* (4).

In the proof of the above theorem, as well that in the the proof of our main results, the concept of semi-concavity plays a crucial role. For a complete account of the theory and its applications to the solution of HJB equations, we refer the reader to [6].

Definition 2.3. We say that $w : \mathbb{R}^d \to \mathbb{R}$ is semi-concave with constant $C_{conc} > 0$ if for every $x_1, x_2 \in \mathbb{R}^d$, $\lambda \in (0,1)$ we have

$$w(\lambda x_1 + (1-\lambda)x_2) \geq \lambda w(x_1) + (1-\lambda)w(x_2) - \lambda(1-\lambda)C_{conc}|x_1 - x_2|^2. \tag{8}$$

A function $\hat{w}$ is said to be semi-convex if $-\hat{w}$ is semi-concave.

Recall that for $w : \mathbb{R}^d \to \mathbb{R}$ the super-differential $D^+w(x)$ at $x \in \mathbb{R}^d$ is defined as

$$D^+w(x) := \left\{ p \in \mathbb{R}^d \ ; \ \limsup_{y \to x} \frac{w(y) - w(x) - \langle p, y - x\rangle}{|y - x|} \leq 0 \right\}.$$

We collect in the following lemmas some useful properties of semi-concave functions (see [6]) .

Lemma 2.4. *For a function* $w : \mathbb{R}^d \to \mathbb{R}$, *the following assertions are equivalent:*

(i) *The function* w *is semi-concave, with constant* C_{conc}.

(ii) *For all* $x, y \in \mathbb{R}^d$, *we have*

$$w(x+y) + w(x-y) - 2w(x) \leq C_{conc}|y|^2.$$

(iii) *For all* $x_1,\ x_2 \in \mathbb{R}^d$ *and* $p \in D^+w(x_1)$, $q \in D^+w(x_2)$

$$\langle q - p, x_2 - x_1\rangle \leq C_{conc}|x - y|^2. \tag{9}$$

(iv) *Setting* I_d *for the identity matrix, we have that* $D^2 w \leq C_{conc} I_d$ *in the sense of distributions.*

Lemma 2.5. *Let $w : \mathbb{R}^d \to \mathbb{R}$ be semi-concave. Then:*

(i) *w is locally Lipschitz.*

(ii) *If w_n is a sequence of uniformly semi-concave converging point-wisely to w, then the convergence is locally uniform and $Dw_n(x) \to Dw(x)$ for a.e. $x \in \mathbb{R}^d$.*

3. **The semi-discrete scheme for the "separated" equations.** In order to describe a discrete in time approximation for (4), we first analyze the discretization of both equations separately.

3.1. **Approximation of the Hamilton-Jacobi equation.** For a given $m \in C([0,T];\mathcal{P}_1)$, the first equation in (4), together with the corresponding boundary value at $t = T$, characterizes the *value function* of an associated optimal control problem. In this subsection we recall some basic properties of this problem and some facts about its semi-discrete in time discretization (see [8, 9, 10] for the semi-discrete approximation and the forthcoming book [11] for a complete account of the theory including the fully discrete semi-lagrangian scheme).

For $t \in [0,T]$ set $\mathcal{A}(t) := L^2([t,T];\mathbb{R}^d)$ for the *set of admissible controls.* Given $\alpha \in \mathcal{A}(t)$ its associated *state* $X^{x,t}[\alpha]$ is defined as the unique solution of

$$\dot{X}(s) = -\alpha(s) \quad \text{for } s \in (t,T), \quad X(t) = x. \tag{10}$$

Now, let $f : \mathbb{R}^d \times [0,T] \to \mathbb{R}$, $g : \mathbb{R}^d \to \mathbb{R}$ be continuous, satisfying that

$$\|f(\cdot,t)\|_{C^2} \leq C \quad \forall t \in [0,T], \qquad \|g\|_{C^2} \leq C. \tag{11}$$

Consider the following cost function

$$J(\alpha; x,t) := \int_t^T \left[\tfrac{1}{2}|\alpha(s)|^2 + f(X^{x,t}[\alpha](s),s)\right] ds + g(X^{x,t}[\alpha](T)). \tag{12}$$

We define $w : \mathbb{R}^d \times [0,T] \to \mathbb{R}$ as the value function of the optimal control problem with dynamics (10) and cost (12), i.e.

$$w(x,t) = \inf_{\alpha \in \mathcal{A}(t)} J(\alpha; x,t) \quad \text{for all } (x,t) \in \mathbb{R}^d \times [0,T]. \tag{13}$$

We recall in the next proposition some basic properties of the function w (see e.g. [7] for our model problem and [4, 6] for generalizations).

Proposition 1. *The following statement hold true:*

(i) *The value function is globally Lipschitz w.r.t. the pair (x,t).*

(ii) *The value function is semi-concave w.r.t. x.*

It is well known that w is the unique viscosity solution of (see e.g. [4])

$$\begin{cases} -\partial_t w(x,t) + \frac{1}{2}|Dw(x,t)|^2 = f(x,t), & \text{in } \mathbb{R}^d \times (0,T), \\ w(x,T) = g(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{14}$$

We define the set of optimal controls $\mathcal{A}(x,t)$ as

$$\mathcal{A}(x,t) := \{\alpha \in \mathcal{A}(t)\ ;\ J(\alpha; x,t) = w(x,t)\}. \tag{15}$$

Under our assumptions, classical arguments show that for all $(x,t) \in \mathbb{R}^d \times (0,T)$ the set $\mathcal{A}(x,t)$ is non-empty. Moreover, the next lemma (proved in e.g. [7, Lemma 4.8 and Lemma 4.9]) shows that unicity in $\mathcal{A}(x,t)$ characterizes the existence of $Dw(x,t)$.

Lemma 3.1. *Let* $(x,t)\in\mathbb{R}^d\times[0,T]$. *Then,*

(i) *For any* $\alpha\in\mathcal{A}(x,t)$ *and* $s\in(t,T]$, *we have that* $\mathcal{A}(X^{x,t}[\alpha](s),s)=\left\{\alpha(\cdot)\big|_{(s,T]}\right\}$.

(ii) *If* $\alpha\in\mathcal{A}(x,t)$ *then* α *is* C^1 *in* $[t,T]$ *and satisfies*

$$\dot\alpha(t)=-Df(X^{x,t}[\alpha](s),s),\ \text{for } s\in(t,T);\quad \alpha(T)=Dg(X^{x,t}[\alpha](T)). \tag{16}$$

(iii) *The function* w *is differentiable with respect to* x *at* (x,t) *iff there exists* $\alpha\in\mathcal{A}(t)$ *such that* $\mathcal{A}(x,t)=\{\alpha\}$. *In this case,* $D_xw(x,t)=\alpha(t)$.

(iv) *For every* $(x,t)\in\mathbb{R}^d$, *any optimal trajectory* $X(\cdot):[t,T]\to\mathbb{R}^d$ *for problem* (13) *satisfies*

$$\dot X(s)=-Dw(X(s),s)\ \text{ for } s\in(t,T),\quad X(t)=x, \tag{17}$$

i.e. every $\alpha(\cdot)\in\mathcal{A}(x,t)$ *admits a feedback representation as* $\alpha(s)=Dw(X(s),s)$ *where* X *is a solution of* (17). *Conversely, any solution* $X(\cdot):[t,T]\to\mathbb{R}^d$ *of* (17) *is an optimal trajectory for problem* (13). *In particular, equation* (17) *admits a unique solution iff the optimal control problem for* $w(x,t)$ *has a unique solution.*

Remark 2. Proposition 1(i) and lemma 3.1(iii) imply that given $t\in[0,T]$ we have that $\mathcal{A}(x,t)$ is a singleton a.e. in $\mathbb{R}^d$.

Let us now introduce a time discretization of (14). Fix $h>0$, set $N=T/h$ (we assume N to be an integer) and for $n=0,...,N-1$ define the set of *discrete controls* $\mathcal{A}_h(n):=\mathbb{R}^{d\times(N-n)}$. Given $\alpha=(\alpha_k)_{k=n}^{N-1}\in\mathcal{A}_h(n)$, let us define, for $k=n,...,N$, the *discrete states* $X_k^{x,n}[\alpha]$ as the solution of

$$\left\{\begin{array}{rcl} X_{k+1} &=& X_k-h\alpha_k=x-h\sum_{i=n}^k\alpha_i \quad \text{for } k=n,\dots,N-1,\\ X_n &=& x. \end{array}\right. \tag{18}$$

In order to simplify the notation, when the context is clear, we do not write the dependence of the discrete state on (x,n) and α and we will simply write X_k. The *discrete cost function* $J_h(\cdot;x,n):\mathcal{A}_h(n)\to\mathbb{R}$ is defined by

$$J_h(\alpha;x,n)=h\sum_{k=n}^{N-1}\left[\tfrac12|\alpha_k|^2+f(X_k,kh)\right]+g(X_N). \tag{19}$$

The value function w_h is given by

$$(x,n)\in\mathbb{R}^d\times\{0,...,N-1\}\to w_h(x,n):=\inf_{\alpha\in\mathcal{A}_h(n)}J(\alpha;x,n),\quad w_h(x,N)=g(x). \tag{20}$$

We easily obtain that w_h satisfies the following *dynamic programming equation*

$$\left\{\begin{array}{ll} w_h(x,n)=\inf_{\alpha\in\mathbb{R}^d}\left\{w_h(x-h\alpha,n+1)+\frac12h|\alpha|^2\right\}+hf(x,nh), & \text{for } x\in\mathbb{R}^d,\\ & n=1,\dots,N-1,\\ w_h(x,N)=g(x) & \text{for } x\in\mathbb{R}^d. \end{array}\right. \tag{21}$$

We denote by $\hat\alpha(x,n)\in\mathbb{R}^d$ an optimal solution (possibly not unique) of the r.h.s. of (21). The next lemma, proved in the appendix, provide useful properties of the discrete value function.

Lemma 3.2. *The following assertions hold true:*

(i) *The function* w_h *satisfies (recall that* C *is the constant in* (11)*):*

$$|w_h(x,n)|\le C(1+T)\quad \textit{for } x\in\mathbb{R}^d,\ n=1,\dots,N-1. \tag{22}$$

(ii) *There exists a constant $C_1 > 0$ independent of h, such that*

$$|w_h(x_1,n_1) - w_h(x_2,n_2)| \leq C_1\left(|x_1 - x_2| + h|n_1 - n_2|\right).$$

(iii) *$w_h(\cdot,n)$ is semi-concave with an associated constant C_2 independent on h.*

(iv) *The function $w_h(\cdot,n)$ is locally semi-convex at $x - h\hat{\alpha}(x,n)$. More precisely, there exists $C_3 > 0$ (independent of h) such that for y in a compact set, we have*

$$w_h(x - h\hat{\alpha}(x,n) + y, n) + w_h(x - h\hat{\alpha}(x,n) - y, n) - 2w_h(x - h\hat{\alpha}(x,n), n) \geq -\frac{C_3}{h}|y|^2 \quad (23)$$

(v) *The following local Lipschitz property for the derivative along the trajectory holds: We have that*

$$|Dw_h(x - h\hat{\alpha}(x,n), n) - Dw_h(y - h\hat{\alpha}(y,n), n)| \leq \frac{C_3}{h}|x - h\hat{\alpha}(x,n) - y + h\hat{\alpha}(y,n)|, \quad (24)$$

provided that x, y belong to a compact set.

It is clear that for every $(x,n) \in \mathbb{R}^d \times \{0,...,N-1\}$ there exists a least one $\alpha \in \mathcal{A}_h(n)$ with

$$w_h(x,n) = J_h(\alpha; x, n). \quad (25)$$

We denote by $\mathcal{A}_h(x,n)$ for the set of discrete controls $\alpha \in \mathcal{A}_h(n)$ such that (25) holds. Let us now provide a "discrete" version of lemma 3.1.

Lemma 3.3. *Let $(x,n) \in \mathbb{R}^d \times \{0,...,N-1\}$. Then,*

(i) *The set $\mathcal{A}_h(x,n)$ is bounded by a constant independent of h.*

(i) *For any $\alpha \in \mathcal{A}_h(x,n)$ and $k = n+1,...,N-1$, we have $\mathcal{A}_h(X_k^{x,n}[\alpha],k) = \{(\alpha_k,...,\alpha_{N-1})\}$.*

(iii) *The function $w(\cdot,n)$ is differentiable at x iff there exists $\bar{\alpha}(x,n) \in \mathcal{A}_h(n)$ such that $\mathcal{A}_h(x,n) = \{\bar{\alpha}(x,n)\}$. In that case,*

$$D_x w(x,n) = \bar{\alpha}_n(x,n) + hDf(x,nh).$$

In particular, the problem associated with $w_h(x,n)$ has a unique solution a.e. in $\mathbb{R}^d$.

(iv) *For all $\alpha \in \mathcal{A}_h(x,n)$, the following relation hold true:*

$$\alpha_k(X_k,k) = Dw_h(X_k - h\alpha_k(X_k,k), k+1) \quad \textit{for} \quad k = n,...,N-1.$$

Proof. By definition of $\hat{\alpha}(x,n)$, lemma 3.2(ii) yields

$$\begin{aligned}\tfrac{1}{2}h|\hat{\alpha}(x,n)|^2 &= w_h(x,n) - w_h(x - h\hat{\alpha}(x,n), n+1) - hf(x,hn) \\ &\leq C_1 h\left(|\hat{\alpha}(x,n)| + 1\right) + hC,\end{aligned}$$

which easily implies (i). The optimality condition for (25) yields that any $\alpha \in \mathcal{A}_h(x,n)$ satisfies

$$\alpha_k = h\sum_{i=k+1}^{N-1} Df(X_i^{x,n}[\alpha], ih) + Dg(X_N^{x,n}[\alpha]) \quad \text{for } k = n,...,N-1. \quad (26)$$

In particular, for all $k = n,...,N-2$, we get that

$$\alpha_{k+1} - \alpha_k = -hDf(X_{k+1}^{x,n}[\alpha],(k+1)h) = -hDf(X_k^{x,n}[\alpha] - h\alpha_k,(k+1)h). \quad (27)$$

Letting $k = n$ we get that α_{n+1} is uniquely determined by α_n and assertion (ii) follows from a recursive argument. Now, let us prove (iii). For notational convenience we prove the result for $n = 0$. One implication is a straightforward consequence of

Danskin theorem (see e.g. [5, Theorem 4.13]). Now, let us assume that $w_h(\cdot,0)$ is differentiable at $x\in\mathbb{R}^d$ and let $\alpha\in\mathcal{A}_h(x,0)$. For $p\in\mathbb{R}^d$ and $\sigma>0$, we have

$$w_h(x+\sigma p,0)\leq \tfrac{1}{2}h|\alpha_0|^2+hf(x+\sigma p,0)+h\sum_{k=1}^{N-1}\left[\tfrac{1}{2}|\alpha_k|^2+f(X_k+\sigma p,kh)\right]+g(X_N+\sigma p),$$

with equality for $p=0$. Subtracting $w_h(x,0)$, dividing by σ and letting $\sigma\downarrow 0$, we get

$$Dw_h(x,0)p\leq hDf(x,0)p+h\sum_{k=1}^{N-1}Df(X_k,kh)p+Dg(X_N)p.$$

Since p is arbitrary we obtain that

$$Dw_h(x,0)=hDf(x,0)+h\sum_{k=1}^{N-1}Df(X_k,kh)+Dg(X_N).$$

Using (26) with $k=0$, we obtain

$$Dw_h(x,0)=hDf(x,0)+\alpha_0,$$

and thus α_0 is unique. Therefore, in view of (27), we obtain that α is unique. Finally, by (ii), we can apply Danskin theorem to obtain that

$$Dw_h(X_k-h\alpha_k(X_k,k),k+1)=h\sum_{i=k+1}^{N-1}Df(X_i^{x,n}[\alpha],ih)+Dg(X_N^{x,n}[\alpha])$$

for $k=n,...,N-1$, and with (26) we obtain (iv). □

Consider the function $\hat{w}_h:\mathbb{R}^d\times[0,T]\to\mathbb{R}$ defined by $\hat{w}_h(x,t):=w_h(x,[t/h])$. For notational convenience, when the context is clear we will write also w_h for $\hat{w}_h$. The following convergence result is well known (see e.g. [4, Chapter V, theorem 1.1], for the stationary case and for more general assumptions).

Theorem 3.4. *As $h\downarrow 0$, w_h converges locally uniformly in $\mathbb{R}^d\times[0,T]$ to the solution w of (14).*

3.2. **Approximation of the continuity equation.** We now discuss the approximation scheme for the continuity equation

$$\begin{cases}\partial_t\mu(x,t)-\operatorname{div}(Dw(x,t)\mu(x,t)) = 0 & \text{for } (x,t)\in\mathbb{R}^d\times(0,T],\\ \mu(x,0) = m_0(x), & \text{for } x\in\mathbb{R}^d.\end{cases}\tag{28}$$

with w being the solution of (14). We recall that μ is said to be a solution of (28) if it satisfies the equation in the distributional sense. Let us first present some well-known results about the existence and uniqueness for a solution of (28).

Consider the multivalued mapping $x\in\mathbb{R}^d\to\mathcal{A}(x,0)$. Standard arguments (see e.g. [3]) easily show that it admits a measurable selection α. Thus, we can define a measurable map $x\in\mathbb{R}^d\to\Phi(x,\cdot)\in C([0,T];\mathbb{R}^d)$ by

$$\Phi(x,t)=x-\int_0^t\alpha(x,s)\mathrm{d}s\quad \text{for all } x\in\mathbb{R}^d.\tag{29}$$

Note that lemma 3.1(iv), implies that given $x\in\mathbb{R}^d$, we have

$$\partial_t\Phi(x,t)=-Dw(\Phi(x,t),t)\quad\text{for } t\in(0,T),\quad \Phi(x,0)=x.\tag{30}$$

Recall that for $\mathbb{T}:\mathbb{R}^d\to\mathbb{R}^d$ and $\mu\in\mathcal{P}_1$, the *push-forward* $\mathbb{T}\sharp\mu\in\mathcal{P}_1$ of μ is defined by $\mathbb{T}\sharp\mu(A):=\mu(\mathbb{T}^{-1}(A))$ for all $A\in\mathcal{B}(\mathbb{R}^d)$. We have (see e.g. [7, theorem 4.18 and lemma 4.14])

Theorem 3.5. *Given a solution w of (13), the map $t\in[0,T]\to\mu(t):=\Phi(\cdot,t)\sharp m_0\in\mathcal{P}_1$ is the unique solution of (28). Moreover, the following assertions hold true:*

(i) *There exists $C_4>0$ such that for all $s\in[0,T]$, $\mu(s)$ is absolutely continuous (with density still denoted by $\mu(s)$), is bounded and has support in $B(0,C_4)$.*

(ii) *For all $t,t'\in[0,T]$, we have that*

$$\mathbf{d}_1(\mu(t),\mu(t'))\le\|Dw(x,t)\|_\infty|t-t'|.$$

Our aim now is to define a discrete in time approximation of $t\in[0,T]\to\mu(t):=\Phi(\cdot,t)\sharp m_0\in\mathcal{P}_1$. Given $h>0$, let $\bar\alpha(\cdot,0)$ be a measurable selection of the optimal controls $\mathcal{A}_h(\cdot,0)$. For $x\in\mathbb{R}^d$, the optimal flux $(\Phi_k(x))_{k=0}^N$ for (20) (where the dependence on h is omitted for simplicity) is defined as

$$\Phi_k(x)=x-h\sum_{i=0}^{k-1}\bar\alpha_i(x,0)\ \text{ for }\quad k=1,...,N-1\quad\text{and}\quad\Phi_0(x)=x. \tag{31}$$

Lemma 3.3 allows us to give the following equivalent form

$$\Phi_{k+1}(x)=\Phi_k(x)-hDw_h(\Phi_{k+1}(x),k+1)\ \text{ for }\quad k=1,...,N-1\quad\text{and}\quad\Phi_0(x)=x. \tag{32}$$

Lemma 3.6. *The flow Φ_k satisfies*

$$|\Phi_k(x)-\Phi_k(y)|^2\ge\left(\frac{1}{1+C_3+2Ch}\right)^k|x-y|^2. \tag{33}$$

Hence the map $x\mapsto\Phi_k(x)$ is invertible in $\Phi_k(\mathbb{R}^d)$ and the inverse is Lipschitz continuous.

Proof. Set $\Phi_k=\Phi_k(x)$ and $\Psi_k=\Phi_k(y)$. Using (32), we have

$$\begin{aligned}|\Phi_{k+1}-\Psi_{k+1}|^2\ &=\ |\Phi_k-\Psi_k|^2+h^2|Dw_h(\Phi_{k+1},k+1)-Dw_h(\Psi_{k+1},k+1)|^2\\ &\quad-2h\langle Dw_h(\Phi_{k+1},k+1)-Dw_h(\Psi_{k+1},k+1),\Phi_k-\Psi_k\rangle.\end{aligned} \tag{34}$$

Now, since $\Phi_k=\Phi_{k+1}+hDw_h(\Phi_{k+1},k+1)$ and $\Psi_k=\Psi_{k+1}+hDw_h(\Psi_{k+1},k+1)$, the sum of the last two terms in (34) is given by

$$\begin{aligned}&-2h\langle Dw_h(\Phi_{k+1},k+1)-Dw_h(\Psi_{k+1},k+1),\Phi_{k+1}-\Psi_{k+1}\rangle\\ &-h^2|Dw_h(\Phi_{k+1},k+1)-Dw_h(\Psi_{k+1},k+1)|^2.\end{aligned}$$

Therefore, by the semi-concavity of $w_h(\cdot,k+1)$ and lemma 3.2(v), we obtain

$$|\Phi_{k+1}-\Psi_{k+1}|^2\ge|\Phi_k-\Psi_k|^2-2Ch|\Phi_{k+1}-\Psi_{k+1}|^2-C_3|\Phi_{k+1}-\Psi_{k+1}|^2,$$

which implies that

$$|\Phi_{k+1}-\Psi_{k+1}|^2\ge\frac{1}{1+C_3+2Ch}|\Phi_k-\Psi_k|^2$$

and (33) follows by iteration. □

We define $\mu_h:\{1,...,N\}\to\mathcal{P}_1$ by

$$\mu_h(k):=\Phi_k(\cdot)\sharp m_0,\quad\text{for all }k=1,...,N. \tag{35}$$

Equivalently,

$$\int_{\mathbb{R}^d} \phi(x) \mathrm{d}\mu_h(k)(x) = \int_{\mathbb{R}^d} \phi(\Phi_k(x)) m_0(x) \mathrm{d}x, \quad \text{for any } \phi \in C(\mathbb{R}^d). \tag{36}$$

We have the following lemma, whose proof follows the lines of lemma 4.14 in [7].

Lemma 3.7. *There exists $C_5 > 0$ (independent of h) such that:*

(i) *For all $k_1, k_2 \in \{1, ..., N\}$, we have that*

$$\mathbf{d}_1(\mu_h(k_1), \mu_h(k_2)) \leq C_5 h |k_1 - k_2|. \tag{37}$$

(ii) *For all $k = 1, ..., N$, $\mu_h(k)$ is absolutely continuous (with density still denoted by $\mu_h(k)$), is bounded and has support in $B(0, C_5)$.*

Proof. Lemma 3.3(i) implies the existence of $C_6 > 0$ (independent of h) such that

$$|\Phi_{k_1}(x) - \Phi_{k_2}(x)| \leq C_6 |k_1 - k_2| h. \tag{38}$$

For any 1-Lipschitz function $\phi : \mathbb{R}^d \to \mathbb{R}$, by (36)

$$\int_{\mathbb{R}^d} \phi(x)\, d[\mu_h(k_1) - \mu_h(k_2)](x) \leq \int_{\mathbb{R}^d} |\Phi_{k_1}(x) - \Phi_{k_2}(x)| dm_0 \leq C_6 h |k_1 - k_2|.$$

Since supp $(m_0) \subset B(0, C)$, then supp $(\mu_h(k)) \subset B(0, C + NC_6 h) = B(0, C + TC_6)$ for any $k = 1, \ldots, N$. For any Borel set A and $k = 1, ..., N$,

$$\mu_h(k)(A) = m_0(\Upsilon_h(A, k)) \leq \|m_0\|_\infty \mathcal{L}^d(\Upsilon_h(A, k)).$$

Thus, lemma 3.6 yields that $\mu_h(k)$ is absolutely continuous with bounded density. □

Now, we consider the time continuous version for μ_h. Denote by $\Phi_h(x, t)$ the unique solution of

$$\partial_t \Phi_h(x, t) = \bar{\alpha}_{[t/h]}(x, 0), \quad \text{for } t \in (0, T), \quad \Phi_h(x, 0) = x, \tag{39}$$

which is nothing else than the linear interpolation of $\left\{\Phi_{[t/h]}(x) \; ; \; t \in [0, T]\right\}$. Analogously, let us define $\mu_h : [0, T] \to \mathcal{P}_1$ as

$$\mu_h(t) := \Phi_h(\cdot, t) \sharp m_0, \tag{40}$$

We extend to continuous time lemma 3.7. The proof being analogous is omitted.

Lemma 3.8. *There exists a constant $C_7 > 0$ (independent of h) such that:*

(i) *For all $t_1, t_2 \in [0, T]$, we have that*

$$\mathbf{d}_1(\mu_h(t_1), \mu_h(t_2)) \leq C_7 |t_1 - t_2|. \tag{41}$$

(ii) *For all $t \in [0, T]$, $\mu_h(t)$ is absolutely continuous (with density still denoted by $\mu_h(t)$), it is bounded and has support in $B(0, C_7)$.*

We have the following convergence result.

Proposition 2. *As $h \downarrow 0$ we have that $\mu_h \to \mu$ in $C([0, T]; \mathcal{P}_1)$.*

We do not provide the proof of the above result since it is a slight variation of the second part of the proof of theorem 4.3 in the next section. As we will see, the argument is strongly based on the fact that the μ is transported with the optimal controls $\alpha(x, 0) \in \mathcal{A}(x, 0)$.

4. **The semi-discrete scheme for the deterministic MFG system.** In this section we combine the analysis done in the previous one and we study an semi-discrete in time approximation of (4). Note that, in view of theorem 3.5, we seek for a pair (v,m) satisfying

$$\begin{array}{rcl} -\partial_t v(t,x)+\frac{1}{2}|Dv(t,x)|^2 & = & F(x,m(t)), \ \text{ in } \mathbb{R}^d\times(0,T), \\ m(t) & = & \Phi(\cdot,t)\sharp m_0 \quad \text{for } t\in[0,T], \\ v(x,T)=G(x,m(T)) & \text{for} & x\in\mathbb{R}^d, \quad m(0)=m_0\in\mathcal{P}_1, \end{array} \tag{42}$$

where $\Phi(x,\cdot)$ is the given by (29), with $\alpha(x,\cdot)$ being a solution for the optimal control associated with $v(x,0)$. This reflects the coupled nature of (42). We set

$$\mathcal{K}_h=\left\{(m(i))_{i=0}^N : m(i)\in\mathcal{P}_1 \quad \text{for all } i=0,\dots,N\right\},$$

and, for given $h>0$, consider the following discretization of (42)

$$\left\{\begin{array}{rl} v_h(x,k) & = \inf_{\alpha\in\mathbb{R}^d}\left\{v_h(x-h\alpha,k+1)+\frac{1}{2}h|\alpha|^2\right\}+hF(x,m_h(k)); \\ & \qquad\qquad\qquad \text{for } x\in\mathbb{R}^d, k=0,\dots,N-1, \\ m_h(k) & = \Phi_k(\cdot)\sharp m_0 \quad \text{for } k=1,\dots,N, \\ m_h(0) & = m_0\in\mathcal{P}_1, \quad v_h(x,N)=G(x,m_h(N)) \quad \text{for } x\in\mathbb{R}^d. \end{array}\right. \tag{43}$$

where $\Phi_k(\cdot)$ is defined by (31) with discrete control being the solution of the r.h.s. in the first equation of (43). Our aim in this section is to analyze the existence and uniqueness of (v_h,m_h) satisfying (43), as well as its convergence to the solution (v,m) of (42). Let us first fix some notations that will be useful to prove a stability result (see lemma 4.1 below). For $m\in\mathcal{K}_h$ set $(w_h[m](\cdot,k))_{k=0}^N$ for the solution of

$$\left\{\begin{array}{rl} w(x,k) & = \inf_{\alpha\in\mathbb{R}^d}\left\{w(x-h\alpha,k+1)+\frac{1}{2}h|\alpha|^2\right\}+F(x,m(k)); \\ & \qquad\qquad\qquad \text{for } x\in\mathbb{R}^d, k=0,\dots,N-1, \\ w(x,N) & = G(x,m(N)) \quad \text{for } x\in\mathbb{R}^d. \end{array}\right. \tag{44}$$

Next, we set $\alpha_k[m](\cdot)$ for a measurable selection of $\mathcal{A}_h[m](\cdot,0)$, the set of optimal controls associated with $w_h[m](\cdot,0)$. We also set $\mu_h[m](k):=\Phi_k[m](\cdot)\sharp m(0)$, where, for $x\in\mathbb{R}^d$, $\Phi_k[m](x)$ is given by the solution of

$$\left\{\begin{array}{rcl} \Phi_{k+1} & = & \Phi_k-h\alpha_k[m](\Phi_k) \ \text{ for } k=0,\dots,N-1, \\ \Phi_0 & = & x. \end{array}\right. \tag{45}$$

Note that the existence of a solution for (43) is equivalent to find a fixed point of $m\in\mathcal{K}_h\to\mu_h[m]\in\mathcal{K}_h$. The following stability result holds true:

Lemma 4.1. *Let $h>0$ be fixed and m^j, m in $\mathcal{K}_h$ such that $m^j(k)\to m(k)$ in $\mathcal{P}_1$ as $j\uparrow\infty$ for every $k=0,\dots,N$. Then, for $k=1,\dots,N$, we have that $\mu_h[m^j](k)$ converges to $\mu_h[m](k)$ in $\mathcal{P}_1$.*

Proof. Lemma 3.3(iii) implies that for each j there exists A_j, with $\mathcal{L}^d(A_j)=0$, such that for $x\in A_j^c$, problem $w_h[m^j](x,0)$ admits a unique solution. We denote, by A_m the respective set for $w_h[m](x,0)$ and we set $A:=\cup_{j=1}^{\infty}A_j\cup A_m$, which satisfies $\mathcal{L}^d(A)=0$. Let us fix $x\in A^c$. We have that $\alpha[m^j](x,k)$ are uniformly bounded in j and k. Therefore, up to subsequences, we obtain that $\alpha[m^j](x,k)\to\alpha(x,k)$, for some $\alpha(x,k)\in\mathbb{R}^d$. Since $w_h[m^j](x,0)\to w_h[m](x,0)$ we obtain, by uniqueness of the optimal control outside A_m, that $\alpha(x,k)=\alpha[m](x,k)$. Now, by (36) we have that

$$d_1(\mu_h[m^j](k),\mu_h[m](k)) \leq \sum_{i=0}^{k-1}\int_{\mathbb{R}^d} h\left|\alpha_i[m^j](x)-\alpha_i[m](x)\right| m_0(x)\mathrm{d}x.$$

Since $m_0\in L^\infty(\mathbb{R}^d)$ and the controls are bounded, we conclude by Lebesgue theorem. □

Theorem 4.2. *Given $h>0$ small enough, there exists a unique solution for system* (43).

Proof. Existence: Let C_5 as in lemma 3.7. Define the convex set

$$\mathcal{C}=\{m\in\mathcal{K}_h:\ m(0)=m_0,\ \ \mathrm{supp}(m(k))\in B(0,C_5)\ \text{ for } k=1,...,n\},$$

which is a compact subset of $\mathcal{K}_h$. By lemma 4.1, the map $m\in\mathcal{C}\to\mu_h[m]\in\mathcal{C}$ is continuous and the existence follows from Schauder fixed point theorem.

Uniqueness: Let (v^1,m^1) and (v^2,m^2) be solutions of (43). We denote respectively by $\bar\alpha^1(\cdot)$, $\bar\alpha^2\in\mathcal{A}_h(0)(\cdot)$ for the respective measurable selections of optimal controls associated to $v^1(\cdot,0)$ and $v^2(\cdot,0)$ and by $\Phi_k^1(\cdot)$, $\Phi_k^2(\cdot)$ the corresponding discrete flows. First note that for every continuous $\varphi:\mathbb{R}^d\times\mathcal{P}_1\to\mathbb{R}$ and $k=1,...,N$, we have the identity

$$\int_{\mathbb{R}^d}\left[\varphi(x,m^1(k))-\varphi(x,m^2(k))\right]\mathrm{d}\left[m^1-m^2\right](k)=I_1(\varphi)+I_2(\varphi), \tag{46}$$

where

$$\begin{array}{rcl} I_1(\varphi) &:=& \int_{\mathbb{R}^d}\left[\varphi(\Phi_k^1(x),m^1(k))-\varphi(\Phi_k^2(x),m^1(k))\right]m_0(x)\mathrm{d}x,\\ I_2(\varphi) &:=& \int_{\mathbb{R}^d}\left[\varphi(\Phi_k^2(x),m^2(k))-\varphi(\Phi_k^1(x),m^2(k))\right]m_0(x)\mathrm{d}x. \end{array} \tag{47}$$

Now, let us set the following notations,

$$\begin{array}{r} \delta_1F(k):=F(\Phi_k^1(x),m^1(k))-F(\Phi_k^2(x),m^1(k)),\\ \delta_2F(k):=F(\Phi_k^2(x),m^2(k))-F(\Phi_k^1(x),m^2(k)),\\ \delta_1G(N):=G(\Phi_N^1(x),m^1(N))-G(\Phi_N^2(x),m^1(N)),\\ \delta_2G(N):=G(\Phi_N^2(x),m^2(N))-G(\Phi_N^1(x),m^2(N)). \end{array}$$

Since for $x\in\mathbb{R}^d$, $\bar\alpha^1(x)$ is the minimum for $v^1(x,0)$, we have

$$\begin{array}{rcl} v^1(x,0) &=& \sum_{k=0}^{N-1}\left(\frac{1}{2}h|\bar\alpha_k^1(x)|^2+hF(\Phi_k^1(x),m^1(k))\right)+G(\Phi_N^1(x),m^1(N)),\\ &\leq& \sum_{k=0}^{N-1}\left(\frac{1}{2}h|\bar\alpha_k^2(x)|^2+hF(\Phi_k^2(x),m^1(k))\right)+G(\Phi_N^2(x),m^1(N)). \end{array} \tag{48}$$

Therefore, we have the following inequality

$$h\sum_{k=0}^{N-1}\delta_1F(k)+\delta_1G(N)\leq\tfrac{1}{2}h\sum_{k=0}^{N-1}\left(|\bar\alpha_k^2(x)|^2-|\bar\alpha_k^1(x)|^2\right). \tag{49}$$

Analogously, since $\bar\alpha^2(x)$ is the minimum for $v^2(x,0)$, we have

$$h\sum_{k=0}^{N-1}\delta_2F(k)+\delta_2G(N)\leq\tfrac{1}{2}h\sum_{k=0}^{N-1}\left(|\bar\alpha_k^1(x)|^2-|\bar\alpha_k^2(x)|^2\right). \tag{50}$$

Adding (49) with (50) and integrating with respect to m_0 gives

$$h\sum_{k=0}^{N-1}\int_{\mathbb{R}^d}[\delta_1F(k)+\delta_2F(k)]m_0(x)\mathrm{d}x+\int_{\mathbb{R}^d}[\delta_1G(N)+\delta_2G(N)]m_0(x)\mathrm{d}x\leq 0,$$

which, by (46) and (47), implies that

$$h\sum_{k=0}^{N-1}\int_{\mathbb{R}^d}\left(F(x,m^1(k))-F(x,m^2(k))\right)\mathrm{d}[m^1-m^2](k)$$
$$+\int_{\mathbb{R}^d}\left(G(x,m^1(N))-G(x,m^2(N))\right)\mathrm{d}[m^1-m^2](N)\leq 0.$$

Therefore, **(H3)** gives $m^1(\cdot)=m^2(\cdot)$ which in turns yields $v^1(\cdot,\cdot)=v^2(\cdot,\cdot)$. □

Let us define the function $\hat{v}_h(\cdot,t):=v_h(\cdot,[\frac{t}{h}])$ and $\bar{m}_h(t):=\Phi_h(\cdot,t)\sharp m_0$, where for $x\in\mathbb{R}^d$, $\Phi_h(x,t)$ is defined as in (39) with $\bar{\alpha}(\cdot)$ being a measurable selection of the optimal controls associated to $v_h(\cdot,0)$. We write still v_h, m_h for $\hat{v}_h$ and $\hat{m}_h$ respectively. We now prove a convergence result.

Theorem 4.3. *As $h\downarrow 0$, v_h converges locally uniformly to v and m_h converge to m in $C([0,T];\mathcal{P}_1)$.*

Proof. Lemma 3.8 and Ascoli theorem implies that m_h converges in $C([0,T];\mathcal{P}_1)$, up to some subsequence, to some measure $\bar{m}$. Now, let us define

$$v^*(x,t):=\limsup_{(x',t')\to(x,t),h\downarrow 0}v_h(x',t'),\quad v_*(x,t):=\liminf_{(x',t')\to(x,t),h\downarrow 0}v_h(x',t').$$

We prove now that v^* is a viscosity subsolution of

$$\begin{array}{rcll}-\partial_t v(x,t)+\frac{1}{2}|Dv(x,t)|^2&=&F(x,\bar{m}(t))&\text{for }(x,t)\in\mathbb{R}^d\times(0,T),\\ v(x,T)&=&G(x,\bar{m}(t))&\text{for }x\in\mathbb{R}^d.\end{array}\tag{51}$$

Let $(\bar{x},\bar{t})\in\mathbb{R}^d\times(0,T)$, $\bar{B}$ a ball of radius $r>0$ centered at $(\bar{x},\bar{t})$ and $\phi\in C^1(\mathbb{R}^d\times(0,T))$ such that $(v^*-\phi)(\bar{x},\bar{t})=\max_{(x,t)\in\bar{B}}(v^*-\phi)(x,t)$. Classical arguments imply the existence of (x_h,t_h), with $(x_h,t_h)\to(\bar{x},\bar{t})$ as $h\downarrow 0$, such that

$$(v_h-\phi)(x_h,t_h)=\max_{(x,t)\in\bar{B}}(v_h-\phi)(x,t).\tag{52}$$

Note that lemma 3.3(i) implies that we can consider the infimum in (21) over a compact set $A\subseteq\mathbb{R}^d$ (independent of h). Setting $n_h=[t_h/h]$, for h small enough and $a\in A$, condition (52) implies that

$$v_h(x_h-ha,t_h+h)-\phi(x_h-ha,t_h+h)\leq v_h(x_h,t_h)-\phi(x_h,t_h).\tag{53}$$

Letting $t'_h=h[t_h/h]$, by definition of $v_h(x_h,t_h)=v_h(x_h,t'_h)$, inequality (53) gives

$$\frac{1}{h}\left[\phi(x_h,t_h)-\phi(x_h-ha,t_h+h)\right]\leq\tfrac{1}{2}|a|^2+F(x_h,m_h(t'_h)).$$

By passing to the limit in the above inequality, and the convergence of m_h to $\bar{m}$ in $C([0,T];\mathcal{P}_1)$ we get that

$$-\partial_t\phi(\bar{x},\bar{t})+\langle D\phi(\bar{x},\bar{t}),a\rangle-\tfrac{1}{2}|a|^2\leq F(\bar{x},\bar{m}(\bar{t}))\quad\text{for all}\quad a\in A.$$

By taking the supremum with respect to $a\in A$, we get that v^* is a viscosity subsolution of (51). Now we prove that v_* is a viscosity super-solution of (51). Let $(\hat{x},\hat{t})\in\mathbb{R}^d\times(0,T)$, $\bar{B}$ a ball of radius $r>0$ centered at $(\hat{x},\hat{t})$ and $\phi\in C^1(\mathbb{R}^d\times(0,T))$ such that $(v^*-\phi)(\hat{x},\hat{t})=\min_{(x,t)\in\hat{B}}(v^*-\phi)(x,t)$. There exists a sequence (x_h,t_h), with $(x_h,t_h)\to(\hat{x},\hat{t})$ as $h\downarrow 0$, such that

$$(v_h-\phi)(x_h,t_h)=\min_{(x,t)\in\bar{B}}(v_h-\phi)(x,t).\tag{54}$$

Therefore for h small enough,

$$v_h(x_h,t_h)-\phi(x_h,t_h) \le v_h(x_h-h\hat{\alpha}(x_h,n_h),t_h+h)-\phi(x_h-h\hat{\alpha}(x_h,n_h),t_h+h), \quad (55)$$

where we recall that $\hat{\alpha}(x,n)$ is defined below (21). Letting $t'_h = h[t_h/h]$, by definition of $v_h(x_h,t'_h)$, we get

$$\frac{1}{h}\left[\phi(x_h - h\hat{\alpha}(x_h,n_h), t_h+h) - \phi(x_h,t_h)\right] \le -\tfrac{1}{2}\left|\hat{\alpha}(x_h,n_h)\right|^2 - F(x_h, m_h(t'_h)).$$

Up to subsequences, $\hat{\alpha}(x_h,n_h)$ converge to some $\bar{a}$. By passing to the limit in the above inequality we get that

$$F(\hat{x},\bar{m}(\hat{t})) \le -\partial_t\phi(\hat{x},\hat{t}) + \langle D\phi(\hat{x},\hat{t}), \bar{a}\rangle - \tfrac{1}{2}\left|\bar{a}\right|^2,$$

and by taking the supremum with respect to a we obtain the result. Finally, by a classical comparison argument, v_h converges locally uniformly to the unique viscosity solution $v[\bar{m}]$ of (51).

In order to conclude the proof we have to show that $\bar{m} = \Phi(\cdot,0)\sharp m_0 = m$. Given $h_n \downarrow 0$, note that there exists a set $A \subseteq \mathbb{R}^d$, with $\mathcal{L}^d(\mathbb{R}^d \setminus A) = 0$, such that for all $x \in A$ we have $\mathcal{A}(x,0) = \{\alpha(x,\cdot)\}$ and $\mathcal{A}_{h_n}(x,\cdot) = \{\alpha^n(x)\}$, for some $\alpha(x,\cdot) \in \mathcal{A}(0)$ and $\alpha^n(x) \in \mathcal{A}_{h_n}(0)$. Lemma 3.3(i) gives that $t \in [0,T] \to \alpha^n(x,t) := \alpha^n_{[t/h]}(x) \in \mathcal{A}(0)$ are bounded in $L^\infty([0,T];\mathbb{R}^d)$ (in particular they are bounded in $L^2([0,T];\mathbb{R}^d)$). Thus, up to subsequence, $\alpha^n(x,\cdot)$ converges weakly in $L^2([0,T);\mathbb{R}^d)$ to some $\bar{\alpha}(x,\cdot)$ and thus its associated flow

$$\Phi^n(x,\cdot) := x - h\int_0^{\cdot} \alpha^n(x,s)\,\mathrm{d}s$$

converge uniformly to some $\bar{\Phi}(x,\cdot)$. Now, since $v_h(x,0) \to v(x,0)$, by the uniqueness of the optimal control for $x \in A$, we obtain that $\bar{\alpha}(x,\cdot) = \alpha(x,\cdot)$ and

$$\bar{\Phi}(x,\cdot) = \Phi(x,\cdot) = x - \int_0^{\cdot} \alpha(x,s)\mathrm{d}s, \quad \text{for all } x \in A.$$

Using that m_0 has compact support, by the dominated convergence theorem, the convergence for $x \in A$ of $\Phi n(x,\cdot)$ to $\bar{\Phi}(x,\cdot) = \Phi(x,\cdot)$ yields that for every 1-Lipschitz $\varphi : \mathbb{R}^d \to \mathbb{R}$ and $t \in [0,T]$,

$$\int_{\mathbb{R}^d} \varphi(x)\mathrm{d}[m_h(t) - m(t)](x) \le \int_{\mathbb{R}^d} |\Phi_n(x,t) - \bar{\Phi}(x,t)| m_0(x)\mathrm{d}x \to 0 \quad \text{as } h \downarrow 0.$$

Therefore $\bar{m} = m$, which completes the proof. □

5. **Appendix.**

Proof of lemma 3.2. Proof of (i): Since g is bounded by C by (11), we have that $g(x-h\alpha)+\frac{1}{2}h|\alpha|^2$ is bounded by the parabolas $-C+\frac{1}{2}h|\alpha|^2$ and $C+\frac{1}{2}h|\alpha|^2$. Thus, using that f is also bounded by C, equation (21) implies

$$w(x,N-1) \le C(1+h),$$

and by a recurrence argument

$$|w_h(x,n)| \le C(1+(N-n)h),$$

from which the result follows.

Proof of (ii): Using expression (20) and assumption (11) we obtain

$$\begin{array}{rcl} |w_h(x,n)-w_h(y,n)| & \leq & \sup_{\alpha\in\mathcal{A}_h(n)} |J(\alpha;x,n)-J(\alpha;y,n)|\,, \\ & \leq & \sup_{\alpha\in\mathcal{A}_h(n)} \left\{ h\sum_{k=n}^{N-1} |f(X_k^{x,n}[\alpha],kh)-f(X_k^{y,n}[\alpha],kh)| \right. \\ & & \left. +\,|g(X_N^{x,n}[\alpha])-g(X_N^{y,n}[\alpha])| \right\}, \\ & \leq & C(Nh+1)|x-y| = C(T+1)|x-y|. \end{array}$$

On the other hand, let $x\in\mathbb{R}^d$ fixed and $n_1,n_2\in\{1,...,N\}$ with $n_1<n_2$. Let $(\alpha)_{k=n_1}^{N-1}\in\mathcal{A}_h(n_1)$ be optimal for $J(\cdot;x,n_1)$ and let $(X_k)_{k=n_1}^{N}$ be its associated state. We have

$$\begin{array}{rl} |w_h(x,n_1)-w_h(x,n_2)| & \leq |w_h(x,n_1)-w_h(X_{n_2},n_2)|+|w_h(X_{n_2},n_2)-w_h(x,n_2)|, \\ & \leq h\sum_{k=n_1}^{n_2-1}\left[\frac{1}{2}|\alpha_k|^2+f(X_k,kh)\right]+C(T+1)|X_{n_2}-x|, \\ & = h\sum_{k=n_1}^{n_2-1}\left[\frac{1}{2}|\alpha_k|^2+f(X_k,kh)\right]+C(T+1)h|\sum_{k=n_1}^{n_2-1}\alpha_k|, \end{array}$$

and the results follows the boundedness uniform on x in lemma 3.3(i).

Proof of (iii): By lemma 2.4(iv) $w_h(\cdot,N)$ is semi-concave. Now, suppose that the result is true for $n=k+1$ and call C_{k+1} the semi-concavity constant of $w_h(\cdot,k+1)$. We have, denoting by $\alpha_h(x,k)$ the optimal vector associated to $w_h(x,k)$,

$$w_h(x,k)=w_h(x-h\alpha_h(x,k),k+1)+\tfrac{1}{2}h\,|\alpha_h(x,k)|^2+hf(x,kh),$$
$$w_h(x\pm y,k)\leq w_h(x-h\alpha_h(x,n)\pm y,k+1)+\tfrac{1}{2}h\,|\alpha_h(x,k)|^2+hf(x\pm y,kh).$$

Therefore, using the semi-concavity of $w_h(\cdot,k+1)$ and of $f(\cdot,kh)$, we easily obtain

$$w_h(x-y,k)+w_h(x+y,k)-2w_h(x,k)\leq C_{k+1}|y|^2+hC|y|^2,$$

where C is the constant in (11). By lemma 2.4(ii) this implies the result for $n=k$ with a semi-concavity constant $C_k\leq C_{k+1}+hC$. Thus, iterating we get that for every n, the function $w_k(\cdot,n)$ is semi-concave with a constant $C_n\leq C(1+Nh)=C(T+1)$.

Proof of (iv): The proof is a straightforward adaptation of [12, Proposition 5.2].

Proof of (v): The proof is a straightforward adaptation of [12, Theorem 5.3]. □

REFERENCES

[1] Y. Achdou, F. Camilli and I. Capuzzo Dolcetta, *Mean field games: Numerical methods for the planning problem,* SIAM J. of Control & Optimization, **50** (2012), 77–109.

[2] Y. Achdou and I. Capuzzo-Dolcetta, *Mean field games: Numerical methods,* SIAM J. Numer. Anal., **48** (2010), 1136–1162.

[3] J.-P. Aubin and H. Frankowska, "Set-Valued Analysis," Systems & Control: Foundations & Applications, **2**, Birkhäuser Boston, Inc., Boston, MA, 1990.

[4] M. Bardi and I. Capuzzo-Dolcetta, "Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations," With appendices by Maurizio Falcone and Pierpaolo Soravia, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 1997.

[5] J. Bonnans and A. Shapiro, "Perturbation Analysis of Optimization Problems," Springer Series in Operations Research, Springer-Verlag, New York, 2000.

[6] P. Cannarsa and C. Sinestrari, "Semiconcave functions, Hamilton-Jacobi equations, and Optimal Control," Progress in Nonlinear Differential Equations and their Applications, **58**, Birkhäuser Boston, Inc., Boston, MA, 2004.

[7] Pierre Cardaliaguet, "Notes on Mean Field Games: From P.-L. Lions' Lectures at Collège de France," Lecture Notes given at Tor Vergata, 2010.

[8] I. Capuzzo-Dolcetta, *On a discrete approximation of the Hamilton-Jacobi equation of dynamic programming,* Appl. Math. Optim., **10** (1983), 367–377.

[9] I. Capuzzo-Dolcetta and M. Falcone, *Discrete dynamic programming and viscosity solutions of the Bellman equation,* Analyse Non Linéaire (Perpignan, 1987), Ann. Inst. H. Poincaré Anal. Non Linéaire, **6** (1989), 161–183.

[10] I. Capuzzo-Dolcetta and H. Ishii, *Approximate solutions of the Bellman equation of deterministic control theory,* Appl. Math. Optim., **11** (1984), 161–181.

[11] M. Falcone and R. Ferretti, "Semi-Lagrangian Approximation Schemes for Linear and Hamilton-Jacobi Equations," MOS-SIAM Series on Optimization, to appear.

[12] D. A. Gomes, *Viscosity solution methods and the discrete Aubry-Mather problem,* Discrete Contin. Dyn. Syst., **13** (2005), 103–116.

[13] D. A. Gomes, J. Mohr and R. R. Souza, *Discrete time, finite state space mean field games,* J. Math. Pures Appl. (9), **93** (2010), 308–328.

[14] O. Guéant, *Mean field games equations with quadratic hamiltonian: a specific approach,* arXiv:1106.3269v1, 2011.

[15] A. Lachapelle, J. Salomon and G. Turinici, *Computation of mean field equilibria in economics,* Math. Models Methods Appl. Sci., **20** (2010), 567–588.

[16] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. I. Le cas stationnaire,* C. R. Math. Acad. Sci. Paris, **343** (2006), 619–625.

[17] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. II. Horizon fini et contrôle optimal,* C. R. Math. Acad. Sci. Paris, **343** (2006), 679–684.

[18] J.-M. Lasry and P.-L. Lions, *Mean field games,* Jpn. J. Math., **2** (2007), 229–260.

[19] P.-L. Lions, Cours du Collège de France. Available from: http://www.college-de-france.fr.

Received November 2011; revised March 2012.

E-mail address: camilli@dmmm.uniroma1.it
E-mail address: fsilva@mat.uniroma1.it