J. Math. Anal. Appl. 512 (2022) 126138

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Regular Articles

Rates of convergence for the policy iteration method for Mean Field Games systems

Fabio Camilli [a], Qing Tang [b,*]

[a] SBAI, Sapienza Università di Roma, Roma, Italy
[b] China University of Geosciences (Wuhan), Wuhan, China

ARTICLE INFO

Article history:
Received 2 August 2021
Available online 4 March 2022
Submitted by G. Carlier

Keywords:
Mean Field Games
Policy iteration
Convergence rate

ABSTRACT

Convergence of the policy iteration method for discrete and continuous optimal control problems holds under general assumptions. Moreover, in some circumstances, it is also possible to show a quadratic rate of convergence for the algorithm. For Mean Field Games, convergence of the policy iteration method has been recently proved in [9]. Here, we provide an estimate of its rate of convergence.
© 2022 Elsevier Inc. All rights reserved.

1. Introduction

The policy iteration method, introduced by Bellman [3] and Howard [15], is a general procedure to solve the Hamilton-Jacobi-Bellman (HJB in short) equation, a nonlinear equation arising in discrete and continuous optimal control problems. The solution of the nonlinear HJB equation, which it is well known to suffer from the so-called "curse of dimensionality" (see [3]), is replaced by the solution of a sequence of linear problems, coupled at each step with an optimization problem for the updating of the new policy. It has been proved that, under general assumptions, the algorithm converges to the solution of the original problem (see [2,14,22–24]); moreover, in some cases, it is possible to show a (local) quadratic rate of convergence of the method which explains its very rapid convergence to the solution of the original problem (see [7,17]).

Mean Field Games (MFG in short) theory has been introduced in [16,21] to characterize Nash equilibria for differential games involving a large (infinite) number of agents. The corresponding mathematical formulation leads to the study of a system of PDEs, composed by a HJB equation, characterizing the value function and the optimal control for the agents; a Fokker-Planck (FP in short) equation, governing the distribution of the population when the agents behaves in an optimal way. In the case of a finite horizon problem with periodic boundary conditions, the MFG system reads as

* Corresponding author.
E-mail addresses: fabio.camilli@uniroma1.it (F. Camilli), tangqingthomas@gmail.com (Q. Tang).

https://doi.org/10.1016/j.jmaa.2022.126138
0022-247X/© 2022 Elsevier Inc. All rights reserved.

$$\begin{cases} -\partial_t u - \Delta u + H(Du) = \sigma F[m(t)](x) & \text{in } Q \\ \partial_t m - \Delta m - \operatorname{div}(m H_p(Du)) = 0 & \text{in } Q \\ m(x,0) = m_0(x),\; u(x,T) = u_T(x) & \text{in } \mathbb{T}^d\,, \end{cases} \tag{1.1}$$

where $Q := \mathbb{T}^d \times [0,T]$, $\mathbb{T}^d$ stands for the flat torus $\mathbb{R}^d/\mathbb{Z}^d$, H is a convex Hamiltonian and σ is a positive constant. The difficulty in solving the previous system can prevent the concrete application of the MFG model to real life problems. Indeed, system (1.1) not only involves the resolution of a HJB equation, but the two equations are strongly coupled and the system has a forward-backward structure which does not permit to solve the two equations in parallel. For this reason, strategies other than the simple discretization of (1.1) must be implemented (see for example [8,10]).

In [9], the following version of the policy iteration method for the MFG system (1.1) was considered:

Policy iteration algorithm: Fixed $R > 0$ and given a bounded, measurable vector field $q^{(0)} : \mathbb{T}^d \times [0,T] \to \mathbb{R}^d$ with $\|q^{(0)}\|_{L^\infty(Q)} \le R$, iterate

(i) Solve

$$\begin{cases} \partial_t m^{(n)} - \Delta m^{(n)} - \operatorname{div}(m^{(n)} q^{(n)}) = 0, & \text{in } Q \\ m^{(n)}(x,0) = m_0(x) & \text{in } \mathbb{T}^d. \end{cases}$$

(ii) Solve

$$\begin{cases} -\partial_t u^{(n)} - \Delta u^{(n)} + q^{(n)} D u^{(n)} - L(q^{(n)}) = \sigma F[m^{(n)}(t)](x) & \text{in } Q \\ u^{(n)}(x,T) = u_T(x) & \text{in } \mathbb{T}^d. \end{cases}$$

(iii) Update the policy

$$q^{(n+1)}(x,t) = \arg\max_{|q|\le R} \Big\{ q \cdot D u^{(n)}(x,t) - L(q) \Big\} \qquad \text{in } Q.$$

Here $L(q) = \sup_{p\in\mathbb{R}^d}\{p\cdot q - H(p)\}$ is the Legendre transform of H. Compared to the algorithm for the HJB equation alone, each iteration of the previous method also includes the resolution of the FP equation. The main advantage of the method, in addition to what has already been observed previously for the HJB equation, is that at each iteration the linear HJB and FP equations are completely decoupled and can be quickly solved with different numerical methods. In [9], it was proved that the previous algorithm converges to a solution (u,m) of the MFG system if the Hamiltonian H is Lipschitz continuous or if $H(p) \simeq |p|^\gamma$, $\gamma > 1$.

In this paper, we study rates of convergence for the MFG policy iteration method. We obtain, via purely PDE methods, a linear rate of convergence for the solution to the HJB equation in MFG system. More precisely, the error between two successive iterations of the sequence $\|u^{(n)}\|_{B_1} + \sigma\|m^{(n)}\|_{B_2}$ generated by the algorithm improves linearly with respect to the error of the previous iterations, for sufficiently large n and small σ, where $\|\cdot\|_{B_i}$, $i = 1,2$, denotes some Banach space norm which will be specified later. Without the coupling, for the HJB equation, the policy iteration method is equivalent to the Newton's method (see [23] and also Section 5) and therefore gives a (local) quadratic improvement with respect to the error $|u^{(n)} - u|$ at each step. However, since the policy $q^{(n)} = H_p(Du^{(n-1)})$ enters as a velocity field in the FP equation, its improvement can correspond at most to a linear one for the distribution error $|m^{(n)} - m|$. This also reflects on the HJB equation due to the coupling term $F[m]$ on the right side of this equation. This point is further explained by an interpretation of the policy iteration method for the MFG system as a quasi-Newton's method obtained, in order the eliminate the coupling among the equations, by suppressing off-diagonal elements in Jacobian of the map of which we are calculating the roots.

Despite the previous limitations, however, the policy iteration method retains the advantage of replacing the resolution of a strongly coupled nonlinear system with a sequence of decoupled linear problems. Moreover, in a neighborhood of the solution, the rapid convergence of the value function $u^{(n)}$ is also reflected in an equally rapid convergence of the distribution $m^{(n)}$ (see [9] and [19] for some numerical simulations).

In this paper we restrict the discussion to MFGs with separable Hamiltonians. Recently in [19] the authors considered the convergence rate of policy iteration algorithms for MFGs with non-separable Hamiltonians using contraction fixed point method. The key difference is that here we do not impose the short time assumption, which is essential for the reasoning in [19].

The paper is organized as follows. In Section 2, we introduce some notations and recall the convergence result in [9]. In Section 3 and 4, we prove the convergence rate for the MFG policy iteration method for the evolutive problem and, respectively, for the stationary one. In Section 5, we provide the interpretation of the policy iteration algorithm as a quasi-Newton method.

2. The policy iteration method for the Mean Field Games system

In this section, we recall some results about the convergence of the policy iteration method for the MFG system.

Throughout the paper we work with maps which are periodic in space, i.e. on the torus $\mathbb{T}^d$. This simplification allows us to ignore problems related to boundary conditions or growth conditions on the data. The main ideas of this paper can be extended to consider models in $\mathbb{R}^d$. We denote by $L^r(\mathbb{T}^d)$, $1 \le r \le \infty$, the set of r summable periodic functions and by $W^{k,r}(\mathbb{T}^d)$, $k \in \mathbb{N}$ and $1 \le r \le \infty$, the Sobolev space of periodic functions having r-summable weak derivatives up to order k. For any $r \ge 1$, we denote by $W^{2,1}_r(Q)$ the space of functions f such that $\partial_t^\delta D_x^\beta f \in L^r(Q)$ for all multi-indices β and δ such that $|\beta| + 2\delta \le 2$. All these spaces are endowed with the corresponding standard norm.
Defined $W^{1,0}_s(Q)$ as the space of functions such that the norm

$$\|u\|_{W^{1,0}_s(Q)} := \|u\|_{L^s(Q)} + \sum_{|\beta|=1} \|D_x^\beta u\|_{L^s(Q)}$$

is finite, we denote by $\mathcal{H}^1_s(Q)$ the space of functions $u \in W^{1,0}_s(Q)$ with $\partial_t u \in (W^{1,0}_{s'}(Q))'$, where $\frac{1}{s} + \frac{1}{s'} = 1$, equipped with the natural norm

$$\|u\|_{\mathcal{H}^1_s(Q)} := \|u\|_{W^{1,0}_s(Q)} + \|\partial_t u\|_{(W^{1,0}_{s'}(Q))'} \ .$$

If $s > d+2$, then $\mathcal{H}^1_s(Q)$ is continuously embedded onto $C^{\delta,\delta/2}(Q)$ for some $\delta \in (0,1)$, see [20, Appendix A]. We describe the assumptions on the data of the problem. Concerning the Hamiltonian, we consider two different settings

(H1) H is differentiable, convex and globally Lipschitz continuous, i.e. there exists a constant $R_0 > 0$ such that

$$|D_p H(p)| \le R_0 \qquad \text{for all } p \in \mathbb{R}^d \ .$$

(H2) H is of the form

$$H(p) = |p|^\gamma, \qquad \gamma > 1.$$

Recall that

$$H(p) = p \cdot q - L(q) \quad \text{if and only if} \quad q = D_p H(p) \ .$$

Concerning the coupling cost, we assume that

(F1) F maps continuously $\mathcal{P}(\mathbb{T}^d)$, the set of probability measure on $\mathbb{T}^d$, into a bounded subset of $C^{0,1}(\mathbb{T}^d)$. Moreover

$$\int_{\mathbb{T}^d} (F[m_1] - F[m_2])d(m_1 - m_2) > 0 \qquad \text{if } m_1 \neq m_2,$$

for $x_1, x_2 \in \mathbb{T}^d$, $m_1, m_2 \in \mathcal{P}_1(\mathbb{T}^d)$.

Finally, for the initial and terminal data, we suppose that

(I) $u_T \in W^{2-\frac{2}{r},r}(\mathbb{T}^d)$, $r > d+2$;
$m_0 \in W^{1,s}(\mathbb{T}^d)$, $s > d+2$, $m_0 \geq 0$ and $\int_{\mathbb{T}^d} m_0(x)dx = 1$.

In the following we recall some a priori estimate for the solution of the linear equations involved in policy iteration method (see [9, Lemma 2.1 and 2.2])

Lemma 2.1. *Given $b \in L^\infty(Q;\mathbb{R}^d)$, $f \in L^r(Q)$ and $u_T \in W^{2-\frac{2}{r},r}(\mathbb{T}^d)$ for some $r > 1$, then the problem*

$$\begin{cases} -\partial_t u - \Delta u + b(x,t)Du = f(x,t) & \text{in } Q \\ u(x,T) = u_T(x) & \text{in } \mathbb{T}^d \end{cases}$$

admits a unique solution $u \in W^{2,1}_r(Q)$ and it holds

$$\|u\|_{W^{2,1}_r(Q)} \leq C(\|f\|_{L^r(Q)} + \|u_T\|_{W^{2-\frac{2}{r},r}(\mathbb{T}^d)}),$$

where C depends on the norm of the coefficients as well as on r, d, T. Furthermore, if $r > d+2$ we have $Du \in C^{\alpha,\alpha/2}$ for some $\alpha \in (0,1)$.

Lemma 2.2. *Given a bounded, measurable vector field $g : Q \to \mathbb{R}^d$ and $m_0 \in L^2(\mathbb{T}^d)$, $m_0 \geq 0$, then the problem*

$$\begin{cases} \partial_t m - \Delta m - \operatorname{div}(g(x,t)m) = 0 & \text{in } Q, \\ m(x,0) = m_0(x) & \text{in } \mathbb{T}^d, \end{cases}$$

has a unique non negative solution $m \in \mathcal{H}^1_2(Q)$. Furthermore, if $m_0 \in L^s(\mathbb{T}^d)$, $s \in (1,\infty)$, then $m \in L^\infty(0,T;L^s(\mathbb{T}^d)) \cap \mathcal{H}^1_2(Q)$ and, if $m_0 \in W^{1,s}(\mathbb{T}^d)$, then

$$\|m\|_{\mathcal{H}^1_s(Q)} \leq C$$

for some constant $C = C(\|g\|_{L^\infty(Q;\mathbb{R}^d)}, \|m_0\|_{W^{1,s}(\mathbb{T}^d)})$.

In [9, Theorems 2.3 and 2.5], it has been proved the following convergence result for the policy iteration method.

Theorem 2.3. *Let either **(H1)** or **(H2)**, **(F1)** and **(I)** be in force. Then, for R sufficiently large, the sequence $(u^{(n)}, m^{(n)})$, generated by the policy iteration algorithm, converges to the solution $(u,m) \in W^{2,1}_r(Q) \times \mathcal{H}^1_s(Q)$ of (1.1).*

Remark 2.4. If **(H1)** holds, one can write $H(p) = \sup_{|q|\leq R_0}\{p \cdot q - L(q)\}$. Therefore, in this case, it is sufficient to consider $R = R_0$ in the policy iteration algorithm to get a converging sequence to the solution of (1.1).
If **(H2)** holds and (u, m) is the solution of (1.1), then there exists a constant $\overline{R}$ such that $\|Du(t)\|_{L^\infty(\mathbb{T}^d)} \leq \overline{R}$ for any $t \in [0, T]$ (see [9, Lemma 2.4]). Then one introduces the truncated Hamiltonian defined by

$$H_{\overline{R}}(p) = \begin{cases} |p|^\gamma & \text{if } |p| < \overline{R}\ , \\ (1-\gamma)\overline{R}^\gamma + \gamma\overline{R}^{\gamma-1}|p| & \text{if } |p| \geq \overline{R}\ , \end{cases}$$

and the problem

$$\begin{cases} -\partial_t u - \Delta u + H_{\overline{R}}(Du) = \sigma F[m(t)](x) & \text{in } Q, \\ \partial_t m - \Delta m - \operatorname{div}(mD_pH_{\overline{R}}(Du)) = 0 & \text{in } Q, \\ m(x,0) = m_0(x),\ u(x,T) = u_T(x) & \text{in } \mathbb{T}^d\ . \end{cases} \tag{2.1}$$

Observe that a solution (u, m) of (1.1) is also a solution of (2.1) and $H_{\overline{R}}$ is globally Lipschitz continuous. Since $H_{\overline{R}}$ satisfies assumption **(H1)**, the policy iteration method with $R = \overline{R}$ converges to the solution of (2.1) and therefore also of (1.1).
Note also that, by the Sobolev embedding of $W^{2,1}_r(Q)$ in $C^{1+\alpha,\frac{1+\alpha}{2}}(Q)$ for $r > d + 2$ with

$$\|u\|_{C^{1+\alpha,(1+\alpha)/2}(Q)} \leq C\|u\|_{W^{2,1}_r(Q)}$$

(see [18, Cor. IV.9.1]) and since $q^{(n)} = H_p(Du^{(n-1)})$, it also follows the convergence of policy $q^{(n)}$ to the optimal control $q = H_p(Du)$ in $L^\infty(Q)$ for $n \to \infty$.

Remark 2.5. Sobolev regularities for solutions to the Fokker-Planck equation in the MFG system have been also considered in [5, Section 10.3].

3. A rate of convergence for the policy iteration method: the finite horizon problem

In this section, we give an estimate of the rate of convergence for the policy iteration method. We replace assumption **(H1)** with

(H3) H is two times differentiable, satisfies **(H1)** and for any $S > 0$, there exists $C_S > 0$ such that

$$H_{pp}(p)q \cdot q \leq C_S|q|^2 \quad \text{for any } |p| \leq S,\ q \in \mathbb{R}^d,$$

and **(F1)** with

(F2) $F : \mathbb{T}^d \times L^s(\mathbb{T}^d) \to L^r(\mathbb{T}^d)$ and for all $t \in (0, T)$,

$$\|F[m_1] - F[m_2]\|_{L^r(\mathbb{T}^d)} \leq C_F\|m_1 - m_2\|_{L^s(\mathbb{T}^d)},$$

for $r, s > d + 2$ and all m_1, $m_2 \in \mathcal{H}^1_s(Q)$.

We prove an estimate for the rate of convergence for the policy iteration method.

Theorem 3.1. *Let either* ***(H2)*** *or* ***(H3)***, ***(F2)*** *and* ***(I)*** *be in force and R as in Theorem 2.3. Then, there exists a constant C, depending only on the data of problem, such that, if $(u^{(n)}, m^{(n)})$ is the sequence generated by the policy iteration method, we have*

$$\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))} \le C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)}, \tag{3.1}$$

$$\|m^{(n+1)} - m^{(n)}\|_{\mathcal{H}^1_2(Q)} \le C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)}, \tag{3.2}$$

$$\begin{aligned}\|u^{(n+1)} - u^{(n)}\|_{W^{2,1}_r(Q)} \le C\big(&\|u^{(n)} - u^{(n-1)}\|^2_{W^{2,1}_r(Q)} \\ &+ \sigma\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))}\big).\end{aligned} \tag{3.3}$$

Proof. Along the proof, the constant C can change from line to line, but it is always independent of n. Set $M^{n+1} = m^{(n+1)} - m^{(n)}$. Then M^{n+1} satisfies the equation

$$\partial_t M^{n+1} - \Delta M^{n+1} - \operatorname{div}(q^{(n+1)} M^{n+1}) = \operatorname{div}((q^{(n+1)} - q^{(n)})m^n), \tag{3.4}$$

with $M^{n+1}(0) = 0$.
We first show (3.1). The proof follows the techniques from [11, Lemma 4.1]. We set

$$f_0 = q^{(n+1)} M^{n+1} + (q^{(n+1)} - q^{(n)})m^{(n)},$$

and, for any τ such that $0 < \tau < t$, we notice

$$\begin{aligned}&\frac{d}{d\tau}|M^{n+1}(\tau)|^s \\ =&s|M^{n+1}(\tau)|^{s-2}M^{n+1}(\tau)\frac{\partial M^{n+1}(\tau)}{\partial\tau} \\ =&s|M^{n+1}(\tau)|^{s-2}M^{n+1}(\tau)\big(\Delta M^{n+1}(x,\tau) + \operatorname{div} f_0(\tau)\big).\end{aligned}$$

Integrating the previous relation in $\mathbb{T}^d$ and observing that

$$\begin{aligned}&\int_{\mathbb{T}^d} s|M^{n+1}(\tau)|^{s-2}M^{n+1}(\tau)\Delta M^{n+1}(\tau)dx \\ &= -s(s-1)\int_{\mathbb{T}^d} |M^{n+1}(\tau)|^{s-2}|DM^{n+1}(\tau)|^2 dx \\ &= -s(s-1)\int_{\mathbb{T}^d} \big(|M^{n+1}(\tau)|^{(s-2)/s}|DM^{n+1}(\tau)|^{2/s}\big)^s dx,\end{aligned}$$

we get

$$\begin{aligned}&\int_{\mathbb{T}^d} \frac{d}{d\tau}|M^{n+1}(\tau)|^s dx \\ &+ s(s-1)\int_{\mathbb{T}^d} \big(|M^{n+1}(\tau)|^{(s-2)/s}|DM^{n+1}(\tau)|^{2/s}\big)^s dx \\ &= s\int_{\mathbb{T}^d} |M^{n+1}(\tau)|^{s-2}M^{n+1}(\tau)\big(\operatorname{div} f_0(\tau)\big)dx\end{aligned}$$

$$
\begin{aligned}
&= -s(s-1)\int_{\mathbb{T}^d} |M^{n+1}(\tau)|^{s-2} DM^{n+1}(\tau) f_0(\tau)dx \\
&\le s(s-1)\int_{\mathbb{T}^d} |M^{n+1}(\tau)|^{s-2}|DM^{n+1}(\tau)||f_0(\tau)|dx.
\end{aligned} \tag{3.5}
$$

By Young inequality with $1/s + 1/s' = 1$, we have

$$
\begin{aligned}
&|\int_{\mathbb{T}^d} |M^{n+1}(\tau)|^{s-2} DM^{n+1}(\tau) f_0(\tau)dx| \\
&\le \int_{\mathbb{T}^d} \big(|M^{n+1}(\tau)|^{(s-2)/2}|DM^{n+1}(\tau)|\big)^{s'} |M^{n+1}(\tau)|^{s'(s-2)/2}dx \\
&+C\|f_0(\tau)\|^s_{L^s(\mathbb{T}^d)}.
\end{aligned}
$$

By

$$
\frac{s'(s-2)}{2}\cdot\frac{1}{1-s'/2} = s
$$

and

$$
\begin{aligned}
&\big(|M^{n+1}(\tau)|^{(s-2)/2}|DM^{n+1}(\tau)|\big)^{s'} |M^{n+1}(\tau)|^{s'(s-2)/2} \\
\le&\frac{1}{2}\Big(\big(|M^{n+1}(\tau)|^{(s-2)/2}|DM^{n+1}(\tau)|\big)^{s'}\Big)^{2/s'} + C\big(|M^{n+1}(\tau)|^{s'(s-2)/2}\big)^{\frac{1}{1-s'/2}},
\end{aligned}
$$

we estimate

$$
\begin{aligned}
&\int_{\mathbb{T}^d} \big(|M^{n+1}(\tau)|^{(s-2)/2}|DM^{n+1}(\tau)|\big)^{s'} |M^{n+1}(\tau)|^{s'(s-2)/2}dx \\
\le&\frac{1}{2}\||M^{n+1}(\tau)|^{(s-2)/s}|DM^{n+1}(\tau)|^{2/s}\|^s_{L^s(\mathbb{T}^d)} + C\|M^{n+1}(\tau)\|^s_{L^s(\mathbb{T}^d)}.
\end{aligned}
$$

Moreover

$$
\begin{aligned}
&\|f_0(\tau)\|^s_{L^s(\mathbb{T}^d)} \\
\le&2^{s-1}\big(\|q^{(n+1)}(\tau)M^{n+1}(\tau)\|^s_{L^s(\mathbb{T}^d)} + \|(q^{(n+1)}(\tau)-q^{(n)}(\tau))m^{(n)}(\tau)\|^s_{L^s(\mathbb{T}^d)}\big) \\
\le&C\big(\|M^{n+1}(\tau)\|^s_{L^s(\mathbb{T}^d)} + \|q^{(n+1)}(\tau)-q^{(n)}(\tau)\|^s_{L^\infty(\mathbb{T}^d)}\big).
\end{aligned}
$$

Replacing the previous estimate in (3.5), we obtain

$$
\begin{aligned}
&\partial_\tau\|M^{n+1}(\tau)\|^s_{L^s(\mathbb{T}^d)} + \frac{1}{2}s(s-1)\||M^{n+1}(\tau)|^{(s-2)/s}|DM^{n+1}(\tau)|^{2/s}\|^s_{L^s(\mathbb{T}^d)} \\
\le&C\big(\|M^{n+1}(\tau)\|^s_{L^s(\mathbb{T}^d)} + \|q^{(n+1)}(\tau)-q^{(n)}(\tau)\|^s_{L^\infty(\mathbb{T}^d)}\big),
\end{aligned}
$$

and, using Gronwall's inequality, we finally get

$$
\sup_{0<t<T}\|M^{n+1}(t)\|_{L^s(\mathbb{T}^d)} \le C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)}.
$$

We then proceed to show (3.2). Multiplying both sides of (3.4) by M^{n+1} and integrating in $\mathbb{T}^d$, we get

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|M^{n+1}(t)\|^2_{L^2(\mathbb{T}^d)}+\int_{\mathbb{T}^d}|DM^{n+1}(x,t)|^2dx\\ =&-\int_{\mathbb{T}^d}q^{(n+1)}(x,t)M^{n+1}(x,t)DM^{n+1}(x,t)dx\\ &-\int_{\mathbb{T}^d}(q^{(n+1)}(x,t)-q^{(n)}(x,t))m^{(n)}(x,t)DM^{n+1}(x,t)dx.\end{aligned}\tag{3.6}$$

Next we integrate both sides of equation (3.6) on the interval $[0,T]$, and get

$$\begin{aligned}&\|M^{n+1}(T)\|^2_{L^2(\mathbb{T}^d)}-\|M^{n+1}(0)\|^2_{L^2(\mathbb{T}^d)}+\int_0^T\int_{\mathbb{T}^d}|DM^{n+1}|^2dxdt\\ =&-\int_0^T\int_{\mathbb{T}^d}q^{(n+1)}M^{n+1}DM^{n+1}dxdt\\ &-\int_0^T\int_{\mathbb{T}^d}(q^{(n+1)}-q^{(n)})m^{(n)}DM^{n+1}dxdt,\end{aligned}$$

from which, recalling that $M^{n+1}(0)\equiv 0$, we get

$$\begin{aligned}&\int_0^T\int_{\mathbb{T}^d}|DM^{n+1}|^2dxdt\le\int_0^T\int_{\mathbb{T}^d}|q^{(n+1)}M^{n+1}DM^{n+1}|dxdt\\ &+\int_0^T\int_{\mathbb{T}^d}|(q^{(n+1)}-q^{(n)})m^{(n)}DM^{n+1}|dxdt.\end{aligned}\tag{3.7}$$

Using (3.1) and Young inequality, we estimate

$$\begin{aligned}&\int_0^T\int_{\mathbb{T}^d}|q^{(n+1)}M^{n+1}DM^{n+1}|dxdt\\ \le&\frac{1}{4}\int_0^T\int_{\mathbb{T}^d}|DM^{n+1}|^2dxdt+\frac{3}{4}\int_0^T\int_{\mathbb{T}^d}|q^{(n+1)}M^{n+1}|^2dxdt\\ \le&\frac{1}{4}\int_0^T\int_{\mathbb{T}^d}|DM^{n+1}|^2dxdt+C\|q^{(n+1)}-q^{(n)}\|^2_{L^\infty(Q)},\end{aligned}\tag{3.8}$$

and

$$\int_0^T\int_{\mathbb{T}^d}|(q^{(n+1)}-q^{(n)})m^{(n)}DM^{n+1}|dxdt$$

$$\leq \frac{1}{4}\int_0^T\int_{\mathbb{T}^d} |DM^{n+1}|^2 dxdt + \frac{3}{4}\int_0^T\int_{\mathbb{T}^d} |(q^{(n+1)} - q^{(n)})m^{(n)}|^2 dxdt$$

$$\leq \frac{1}{4}\int_0^T\int_{\mathbb{T}^d} |DM^{n+1}|^2 dxdt + C\|q^{(n+1)} - q^{(n)}\|^2_{L^\infty(Q)}. \tag{3.9}$$

Replacing (3.8) and (3.9) in (3.7)

$$\frac{1}{2}\int_0^T\int_{\mathbb{T}^d} |DM^{n+1}|^2 dxdt \leq C\|q^{(n+1)} - q^{(n)}\|^2_{L^\infty(Q)},$$

so that, with (3.1), we get

$$\|m^{(n+1)} - m^{(n)}\|_{W_2^{1,0}(Q)} \leq C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)}. \tag{3.10}$$

For any test function $\phi \in W_2^{1,0}(Q)$, we multiply both sides of (3.4) by ϕ and integrate on Q to obtain

$$\begin{aligned}\int_0^T\int_{\mathbb{T}^d} \partial_t M^{n+1}\phi dxdt = &-\int_0^T\int_{\mathbb{T}^d} DM^{n+1}\cdot D\phi dxdt\\ -\int_0^T\int_{\mathbb{T}^d} D\phi\cdot(q^{(n+1)}M^{n+1})dxdt &- \int_0^T\int_{\mathbb{T}^d} D\phi\cdot((q^{(n+1)} - q^{(n)})m^n)dxdt.\end{aligned} \tag{3.11}$$

We estimate the three terms on the right hand side of (3.11) by

$$\begin{aligned}&\int_0^T\int_{\mathbb{T}^d} DM^{n+1}\cdot D\phi dxdt\\ \leq&\Big(\int_0^T\int_{\mathbb{T}^d} |DM^{n+1}|^2 dxdt\Big)^{\frac{1}{2}}\Big(\int_0^T\int_{\mathbb{T}^d} |D\phi|^2 dxdt\Big)^{\frac{1}{2}}\\ \leq& C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)}\|D\phi\|_{L^2(0,T;L^2(\mathbb{T}^d))},\\ &\int_0^T\int_{\mathbb{T}^d} D\phi\cdot(q^{(n+1)}M^{n+1})dxdt\\ \leq& R\Big(\int_0^T\int_{\mathbb{T}^d} |M^{n+1}|^2 dxdt\Big)^{\frac{1}{2}}\Big(\int_0^T\int_{\mathbb{T}^d} |D\phi|^2 dxdt\Big)^{\frac{1}{2}}\\ \leq& C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)}\|D\phi\|_{L^2(0,T;L^2(\mathbb{T}^d))},\end{aligned}$$

and

$$\int_0^T\int_{\mathbb{T}^d} D\phi\cdot((q^{(n+1)} - q^{(n)})m^n)dxdt$$

$$
\begin{aligned}
&\leq \|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)} \left(\int_0^T \int_{\mathbb{T}^d} |m^n|^2 dxdt \right)^{\frac{1}{2}} \left(\int_0^T \int_{\mathbb{T}^d} |D\phi|^2 dxdt \right)^{\frac{1}{2}} \\
&\leq C \|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)} \|D\phi\|_{L^2(0,T;L^2(\mathbb{T}^d))}.
\end{aligned}
$$

Replacing in (3.11), we obtain

$$
\sup_{\phi \in W_2^{1,0}(Q)} \int_0^T \int_{\mathbb{T}^d} \partial_t M^{n+1} \phi dxdt \leq C \|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)} \|\phi\|_{W_2^{1,0}(Q)},
$$

i.e.

$$
\|\partial_t M^{n+1}\|_{(W_2^{1,0}(Q))'} \leq C \|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)}. \tag{3.12}
$$

From (3.10) and (3.12), we get (3.2).

We now prove the estimate (3.3) for the HJB equation. The function $U^{n+1} = u^{(n+1)} - u^{(n)}$ satisfies the equation

$$
-\partial_t U^{n+1} - \Delta U^{n+1} + q^{(n+1)} D U^{n+1} = \mathcal{F}(x,t)
$$

with $U^{n+1}(x,T) = 0$, where

$$
\mathcal{F}(x,t) = \sigma F[m^{(n+1)}] - \sigma F[m^{(n)}] + q^{(n)} D u^{(n)} - L(q^{(n)}) - (q^{(n+1)} D u^{(n)} - L(q^{(n+1)})). \tag{3.13}
$$

Hence, recalling that $q^{(n+1)} = H_p(Du^{(n)})$ is bounded, we have the estimate (see Lemma 2.1)

$$
\|U^{n+1}\|_{W_r^{2,1}(Q)} \leq \|\mathcal{F}\|_{L^r(Q)}. \tag{3.14}
$$

To estimate $\|\mathcal{F}\|_{L^r(Q)}$, first observe that, since

$$
\begin{aligned}
&q^{(n+1)} D u^{(n)} - L(q^{(n+1)}) = \sup_{q \in \mathbb{R}^d} \left\{ q \cdot D u^{(n)} - L(q) \right\} \\
&\geq q^{(n)} D u^{(n)} - L(q^{(n)}),
\end{aligned}
$$

then we have

$$
\mathcal{F}(x,t) \leq \sigma(F[m^{(n+1)}](x) - F[m^{(n)}](x)). \tag{3.15}
$$

Moreover, by $q^{(n)} = H_p(Du^{(n-1)})$ and

$$
H(Du^{(n-1)}) = q^{(n)} D u^{(n-1)} - L(q^{(n)}),
$$

we have

$$
\begin{aligned}
&q^{(n)} D u^{(n)} - L(q^{(n)}) - (q^{(n+1)} D u^{(n)} - L(q^{(n+1)})) = \\
&q^{(n)} D u^{(n)} + H(Du^{(n-1)}) - q^{(n)} D u^{(n-1)} - H(Du^{(n)}) = \\
&H(Du^{(n-1)}) - H(Du^{(n)}) + H_p(Du^{(n-1)})(Du^{(n)} - Du^{(n-1)}) = \\
&-\frac{1}{2} H_{pp}(\theta D u^{(n)} + (1-\theta) D u^{(n-1)})(Du^{(n)} - Du^{(n-1)}) \cdot (Du^{(n)} - Du^{(n-1)})
\end{aligned}
$$

for some $\theta \in (0,1)$. Therefore, either by **(H3)** or by **(H2)**, recalling that $\|Du^{(n)}\|_\infty$ is uniformly bounded in $n \in \mathbb{N}$, we get

$$\mathcal{F}(x,t) \geq \sigma(F[m^{(n+1)}] - F[m^{(n)}]) - C|Du^{(n)}(x,t) - Du^{(n-1)}(x,t)|^2. \tag{3.16}$$

Recall that $r > d+2$. From (3.15), (3.16) and Minkowski inequality we obtain

$$\begin{aligned}\|\mathcal{F}\|_{L^r(Q)} \leq & C\|Du^{(n)} - Du^{(n-1)}\|^2_{L^{2r}(Q)} + \sigma\|F[m^{(n+1)}] - F[m^{(n)}]\|_{L^r(Q)} \\ \leq & C\big(\|u^{(n)} - u^{(n-1)}\|^2_{W^{2,1}_r(Q)} + \sigma\|F[m^{(n+1)}] - F[m^{(n)}]\|_{L^r(Q)}\big)\end{aligned}$$

From **(F2)**, for each t we have

$$\|F[m^{(n+1)}](t) - F[m^{(n)}](t)\|_{L^r(\mathbb{T}^d)} \leq C\|m^{(n+1)}(t) - m^{(n)}(t)\|_{L^s(\mathbb{T}^d)},$$

so that

$$\begin{aligned}\|F[m^{(n+1)}] - F[m^{(n)}]\|_{L^r(Q)} \leq & C\|F[m^{(n+1)}] - F[m^{(n)}]\|_{L^\infty(0,T;L^r(\mathbb{T}^d))} \\ \leq & C\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))}.\end{aligned}$$

Then we can get (3.3) from (3.14). $\square$

A key difficulty for estimating convergence rate using Theorem 3.1 is that we cannot control the constants C in (3.2) and (3.3). We do not have additional information other than they depend on the data of the problem and not on n. To address this difficulty, we introduce an additional assumption on the smallness of σ. It is not needed for the convergence of the policy iteration method but allows us to get a linear convergence rate to the solution of MFG system in the policy iteration. This type of assumption also plays a key role in [12] for considering MFGs of aggregation.

Corollary 3.2. *Under the same assumptions of Theorem 3.1, there exist constants $\sigma_0 > 0$ and $0 < C^* < 1$, such that for sufficiently large n and $\forall \sigma < \sigma_0$,*

$$\|u^{(n+1)} - u^{(n)}\|_{W^{2,1}_r(Q)} + \sigma\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))} \leq C^*\|u^{(n)} - u^{(n-1)}\|_{W^{2,1}_r(Q)}. \tag{3.17}$$

Proof. First note that, by parabolic Sobolev embedding theorem ([18], Corollary IV.9.1 p.342) and the fact that $r > d+2$, we have

$$\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)} \leq C\|u^{(n)} - u^{(n-1)}\|_{W^{2,1}_r(Q)}.$$

By (3.3) and (3.2) we have

$$\begin{aligned}&\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))} \leq C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(Q)} \\ &\leq C_1\|u^{(n)} - u^{(n-1)}\|_{W^{2,1}_r(Q)}.\end{aligned}$$

From (3.3) we have

$$\begin{aligned}&\|u^{(n+1)} - u^{(n)}\|_{W^{2,1}_r(Q)} \leq C_2\big(\|u^{(n)} - u^{(n-1)}\|^2_{W^{2,1}_r(Q)} \\ &+ \sigma\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))}\big).\end{aligned}$$

Here C_1 and C_2 are always independent of n and σ. Then for sufficiently small σ we have $C_2C_1\sigma < \frac{C^*}{4}$ and

$$\begin{aligned}&\|u^{(n+1)} - u^{(n)}\|_{W_r^{2,1}(Q)} + \sigma\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))}\\ \le& C_2\|u^{(n)} - u^{(n-1)}\|^2_{W_r^{2,1}(Q)} + 2C_2C_1\sigma\|u^{(n)} - u^{(n-1)}\|_{W_r^{2,1}(Q)}.\end{aligned}$$

Since $u^{(n)}$ converges in $W_r^{2,1}(Q)$ and C_2 is independent of n, we have for sufficiently large n,

$$C_2\|u^{(n)} - u^{(n-1)}\|_{W_r^{2,1}(Q)} < \frac{C^*}{2},$$

so that

$$\begin{aligned}&\|u^{(n+1)} - u^{(n)}\|_{W_r^{2,1}(Q)} + \sigma\|m^{(n+1)} - m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))}\\ &\le C^*\|u^{(n)} - u^{(n-1)}\|_{W_r^{2,1}(Q)} \quad \square\end{aligned}$$

Remark 3.3. Assumption **(F2)** is satisfied for example if

$$F[m] = \int_{\mathbb{T}^d} K(x,y)m(y,t)dy,$$

for some bounded kernel $K : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{R}$. We have

$$\begin{aligned}\|F[m_1(t)] - F[m_2(t)]\|_{L^r(\mathbb{T}^d)} &\le C\|F[m_1(t)] - F[m_2(t)]\|_{L^\infty(\mathbb{T}^d)}\\ &\le C\|K\|_{L^\infty(\mathbb{T}^d\times\mathbb{T}^d)}\|m_1(t) - m_2(t)\|_{L^s(\mathbb{T}^d)}.\end{aligned}$$

If $K(x,y) = K(x-y)$, it is sufficient to assume that $K \in L^\zeta(\mathbb{T}^d)$, with $1/s + 1/\zeta = 1/r + 1$. Indeed, by Young's convolution inequality, we have in this case

$$\begin{aligned}&\|F[m_1(t)] - F[m_2(t)]_{C(0,T;L^s(\mathbb{T}^d))}]\|_{L^r(\mathbb{T}^d)}\\ &\le \|K\|_{L^\zeta(\mathbb{T}^d\times\mathbb{T}^d)}\|m_1(t) - m_2(t)\|_{L^s(\mathbb{T}^d)}.\end{aligned}$$

Note that estimate (3.17) also holds for the case of a local coupling, i.e. $F = F(x,m)$, assuming F to be Lipschitz continuous in m, uniformly in x and $r \le s$.

Remark 3.4. The results can be generalized to a Hamiltonian H dependent on $x \in \mathbb{T}^d$, where the assumptions (H1), (H2), (H3) are replaced respectively by

($\widetilde{H1}$) H is differentiable, convex and globally Lipschitz continuous, i.e. there exists a constant $R_0 > 0$ such that

$$|D_pH(x,p)| \le R_0 \qquad \text{for all } p \in \mathbb{R}^d\ ,$$

uniformly in x.

($\widetilde{H2}$) H is of the form

$$H(x,p) = h(x)|p|^\gamma, \qquad \gamma > 1,$$

where $0 < h_0 < h(x) < h_1$, h_0 and h_1 are two constants.

($\widetilde{H}3$) H is two times differentiable, satisfies ($\widetilde{H}1$) and for any $S>0$, there exists $C_S>0$ such that

$$H_{pp}(x,p)q\cdot q\le C_S|q|^2 \quad \text{for any } |p|\le S,\ q\in\mathbb{R}^d,$$

uniformly in x.

When either ($\widetilde{H}1$) or ($\widetilde{H}2$) holds, the uniform boundedness of Du has been shown in [13]. Moreover, either by ($\widetilde{H}2$) or by ($\widetilde{H}3$), we get (3.16).

Remark 3.5. We can also generalize the results to include the case

$$u(x,T)=\sigma' u_T[m],$$

where σ' is a positive constant, i.e. the final cost depends on the agents state distribution, assuming the regularizing assumption

$$\|u_T[m_1]-u_T[m_2]\|_{W^{2-\frac{2}{r},r}(\mathbb{T}^d)}\le C\|m_1(T)-m_2(T)\|_{L^s(\mathbb{T}^d)}. \tag{3.18}$$

The function $U^{n+1}=u^{(n+1)}-u^{(n)}$ satisfies the equation

$$-\partial_t U^{n+1}-\Delta U^{n+1}+q^{(n+1)}DU^{n+1}=\mathcal{F}(x,t),$$

with $U^{n+1}(x,T)=\sigma'(u_T[m^{(n+1)}]-\sigma' u_T[m^{(n)}])$, $\mathcal{F}$ defined as (3.13). Then, using Lemma 2.1 and (3.18) to estimate

$$\begin{aligned}\|u_T[m^{(n+1)}]-u_T[m^{(n)}]\|_{W^{2-\frac{2}{r},r}(\mathbb{T}^d)}\le & C\sigma'\|m^{(n+1)}(T)-m^{(n)}(T)\|_{L^s(\mathbb{T}^d)}\\ \le & C\sigma'\|m^{(n+1)}-m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))}\end{aligned}$$

we will have

$$\begin{aligned}\|u^{(n+1)}-u^{(n)}\|_{W^{2,1}_r(Q)}\le C\big(&\|u^{(n)}-u^{(n-1)}\|^2_{W^{2,1}_r(Q)}\\ &+(\sigma+\sigma')\|m^{(n+1)}-m^{(n)}\|_{C(0,T;L^s(\mathbb{T}^d))}\big).\end{aligned} \tag{3.19}$$

Hence, we get a linear rate of convergence if we assume both σ and σ' sufficiently small.

Remark 3.6. Assume that F is independent of m, i.e. $F[m](x)=F(x)$. In this case, by Proposition 3.1, we recover two well known properties of the policy iteration method for the Hamilton-Jacobi-Bellman equation. Firstly, by (3.15), we have that $U^{n+1}=u^{(n+1)}-u^{(n)}$ satisfies

$$-\partial_t U^{n+1}-\Delta U^{n+1}+q^{(n+1)}DU^{n+1}\le 0$$

with $U^{n+1}(x,T)\equiv 0$. Therefore, by comparison principle, $U^{n+1}\le 0$, hence the policy iteration method generates a decreasing sequence $u^{(n)}$. Moreover, by estimate (3.3), we get a (local) quadratic rate of convergence for the method (a similar estimate is proved in [17] via probabilistic techniques).

4. A rate of convergence for the policy iteration method: the ergodic problem

In this section, we prove a rate of convergence for the policy iteration method for the ergodic MFG system

$$\begin{cases} -\Delta u + H(Du) + \lambda = \sigma F[m] & \text{in } \mathbb{T}^d \\ -\Delta m - \operatorname{div}(mH_p(Du)) = 0 & \text{in } \mathbb{T}^d \\ \int_{\mathbb{T}^d} m(x)dx = 1, \quad m \geq 0, \quad \int_{\mathbb{T}^d} u(x)dx = 0 \ . \end{cases} \tag{4.1}$$

Policy iteration algorithm: For fixed $R > 0$ and given a bounded, measurable function $q^{(0)}$ such that $\|q^{(0)}\|_{L^\infty(\mathbb{T}^d)} \leq R$, a policy iteration method for (4.1) is given by

(i) Solve

$$\begin{cases} -\Delta m^{(n)} - \operatorname{div}(m^{(n)} q^{(n)}) = 0, & \text{in } \mathbb{T}^d \\ \int_{\mathbb{T}^d} m^{(n)}(x)dx = 1, \quad m^{(n)} \geq 0. \end{cases}$$

(ii) Solve

$$\begin{cases} -\Delta u^{(n)} + q^{(n)} \cdot Du^{(n)} - L(q^{(n)}) + \lambda^{(n)} = \sigma F[m^{(n)}] & \text{in } \mathbb{T}^d \\ \int_{\mathbb{T}^d} u^{(n)}(x)dx = 0. \end{cases}$$

(iii) Update the policy

$$q^{(n+1)}(x,t) = \arg\max_{|q|\leq R} \left\{ q \cdot Du^{(n)}(x) - L(q) \right\} \quad \text{in } \mathbb{T}^d.$$

In [9, Section 4], the following convergence theorem is proved

Theorem 4.1. *Let either **(H1)** or **(H2)** and **(F1)** be in force and R sufficiently large. Then, the sequence $(u^{(n)}, \lambda^{(n)}, m^{(n)})$, generated by the policy iteration algorithm converges to the solution $(u, \lambda, m) \in W^{2,r}(\mathbb{T}^d) \times \mathbb{R} \times W^{1,s}(\mathbb{T}^d)$ of (4.1), uniformly in $\mathbb{T}^d$.*

For the proof of the convergence estimate, we need a preliminary lemma.

Lemma 4.2. *Let $f \in W^{-1,s}(\mathbb{T}^d) = (W^{1,s'}(\mathbb{T}^d))'$, $s' = s/(s-1)$, and $q \in L^\infty(\mathbb{T}^d)$. If M satisfies*

$$\begin{cases} -\Delta M - \operatorname{div}(qM) = f, & \text{in } \mathbb{T}^d \\ \int_{\mathbb{T}^d} M dx = 0, \end{cases} \tag{4.2}$$

then

$$\|M\|_{W^{1,s}(\mathbb{T}^d)} \leq C\|f\|_{W^{-1,s}(\mathbb{T}^d)}. \tag{4.3}$$

Proof. From [6, Prop. 1.2.3])

$$\begin{aligned} \|M\|_{W^{1,s}(\mathbb{T}^d)} \leq & C\|\operatorname{div}(qM) + f\|_{W^{-1,s}(\mathbb{T}^d)} \\ \leq & C\big(\|q\|_{L^\infty(\mathbb{T}^d)}\|M\|_{L^s(\mathbb{T}^d)} + \|f\|_{W^{-1,s}(\mathbb{T}^d)}\big) \\ \leq & C(\|M\|_{L^s(\mathbb{T}^d)} + \|f\|_{W^{-1,s}(\mathbb{T}^d)}). \end{aligned} \tag{4.4}$$

We claim this leads to (4.3). Following the argument in [25, pag.6], we assume by contradiction that there exists a sequence M_k of solutions to (4.2) such that

$$\|M_k\|_{W^{1,s}(\mathbb{T}^d)} = 1,\ \|AM_k\|_{W^{-1,s}(\mathbb{T}^d)} \to 0,\ \int\limits_{\mathbb{T}^d} M_k dx = 0,$$

where

$$A := -\Delta \cdot -\mathrm{div}(q\cdot).$$

By Rellich-Kondrachov Theorem, $W^{1,s}(\mathbb{T}^d)$ is compactly embedded in $L^s(\mathbb{T}^d)$ for $1 < s < \infty$. Then from Banach-Alaoglu theorem there is a subsequence, again denoted by M_k, which converges weakly in $W^{1,s}(\mathbb{T}^d)$ and strongly in $L^s(\mathbb{T}^d)$. We have

$$\|M_j - M_k\|_{W^{1,s}(\mathbb{T}^d)} \le C(\|AM_j - AM_k\|_{W^{-1,s}(\mathbb{T}^d)} + \|M_j - M_k\|_{L^s(\mathbb{T}^d)}).$$

Then we have that M_k converges to $\bar{M}$ in $W^{1,s}(\mathbb{T}^d)$ and

$$A\bar{M} = 0,\ \int\limits_{\mathbb{T}^d} \bar{M} dx = 0.$$

By Theorem 4.2 and Lemma 4.3 in [4], the previous problem has only a trivial solution and therefore a contradiction since $\|\bar{M}\|_{W^{1,s}(\mathbb{T}^d)} = 1$. □

Theorem 4.3. *Let either **(H2)** or **(H3)**, **(F2)** be in force, $r, s > d$ and R as in Theorem 4.1. Then, there exists a constant C, depending only on the data of problem, such that, if $(u^{(n)}, \lambda^{(n)}, m^{(n)})$ is the sequence generated by the policy iteration method, we have*

$$\|m^{(n+1)} - m^{(n)}\|_{W^{1,s}(\mathbb{T}^d)} \le C\|q^{(n+1)} - q^{(n)}\|_{L^\infty(\mathbb{T}^d)}, \tag{4.5}$$

$$\begin{aligned} \|u^{(n+1)} - u^{(n)}\|_{W^{2,r}(\mathbb{T}^d)} + |\lambda^{(n+1)} - \lambda^{(n)}| \le C\big(&\|u^{(n)} - u^{(n-1)}\|^2_{W^{2,r}(\mathbb{T}^d)} \\ &+\sigma\|m^{(n+1)} - m^{(n)}\|_{W^{1,s}(\mathbb{T}^d)}\big), \end{aligned} \tag{4.6}$$

Proof. Along the proof, the constant C can change from line to line, but it is always independent of n. Set $M^{n+1} = m^{(n+1)} - m^{(n)}$. Then M^{n+1} satisfies the equation

$$-\Delta M^{n+1} - \mathrm{div}(q^{(n+1)} M^{n+1}) = \mathrm{div}((q^{(n+1)} - q^{(n)})m^{(n)}),$$

with $\int_{\mathbb{T}^d} M^{n+1} dx = 0$. Hence, by (4.3) we have

$$\begin{aligned} \|M^{n+1}\|_{W^{1,s}(\mathbb{T}^d)} \le& C\|\mathrm{div}((q^{(n+1)} - q^{(n)})m^{(n)})\|_{W^{-1,s}(\mathbb{T}^d)} \\ \le& C\|(q^{(n+1)} - q^{(n)})m^{(n)}\|_{L^s(\mathbb{T}^d)} \\ \le& C\|(q^{(n+1)} - q^{(n)})\|_{L^\infty(\mathbb{T}^d)}\|m^{(n)}\|_{L^s(\mathbb{T}^d)} \\ \le& C\|(q^{(n+1)} - q^{(n)})\|_{L^\infty(\mathbb{T}^d)} \end{aligned}$$

and therefore (4.5).

The couple $U^{n+1} = u^{(n+1)} - u^{(n)}$, $\Lambda^{n+1} = \lambda^{(n+1)} - \lambda^{(n)}$ satisfies the equation

$$-\Delta U^{n+1} + q^{(n+1)} DU^{n+1} + \Lambda^{n+1} = \mathcal{F}(x)$$

with $\int U^{n+1}(x)dx = 0$, where

$$\mathcal{F}(x) = \sigma(F[m^{(n+1)}] - F[m^{(n)}]) + q^{(n)}Du^{(n)} - L(q^{(n)}) - (q^{(n+1)}Du^{(n)} - L(q^{(n+1)})).$$

Exploiting the results in [4] (Theorem 6.1, Pag. 196), we get

$$|\Lambda^{n+1}| \leq \|\mathcal{F}(x)\|_{L^r(\mathbb{T}^d)} \tag{4.7}$$

$$\|U^{n+1}\|_{W^{2,r}(\mathbb{T}^d)} \leq C\|\mathcal{F}(x)\|_{L^r(\mathbb{T}^d)} \tag{4.8}$$

Repeating similar computations to the one of the corresponding estimate in Theorem 3.1, we obtain

$$\|\mathcal{F}(x)\|_{L^r(\mathbb{T}^d)} \leq C\big(\|Du^{(n)} - Du^{(n-1)}\|^2_{L^\infty(\mathbb{T}^d)} + \sigma\|F[m^{(n+1)}] - F[m^{(n)}]\|_{L^r(\mathbb{T}^d)}\big).$$

Replacing the previous estimate in (4.7)-(4.8) and exploiting **(F2)** and $r, s > d$, we get (4.6). $\square$

Remark 4.4. For $s = 2$, estimate (4.5) is a special case of Lemma 3.8 from [1].

Arguing as in Corollary 3.2, we can obtain the rate of convergence

Corollary 4.5. *Under the same assumptions of Theorem 4.3, there exist constants σ_1 and $0 < C^{**} < 1$, such that for $\sigma < \sigma_1$*

$$\begin{aligned}&\|u^{(n+1)} - u^{(n)}\|_{W^{2,r}(\mathbb{T}^d)} + |\lambda^{(n+1)} - \lambda^{(n)}| + \sigma\|m^{(n+1)} - m^{(n)}\|_{W^{1,s}(\mathbb{T}^d)}\\ \leq & C^{**}\|u^{(n)} - u^{(n-1)}\|^2_{W^{2,r}(\mathbb{T}^d)}.\end{aligned}$$

5. An interpretation of the policy iteration method for the MFG system

The following computations only hold at a formal level and, for simplicity, we assume that F is local coupling. Define the map

$$\mathcal{F} : (u, m) \to \begin{pmatrix} -\partial_t u - \Delta u + H(Du) - \sigma F(m) \\ \partial_t m - \Delta m - \operatorname{div}(mH_p(Du)) \\ u(T) - u_T(x) \\ m(0) - m_0(x) \end{pmatrix}.$$

Then system (1.1) is equivalent to find the roots of $\mathcal{F}$ and the corresponding Newton's iterations can be written as

$$J\mathcal{F}(u^{(n-1)}, m^{(n-1)})((u^{(n)}, m^{(n)}) - (u^{(n-1)}, m^{(n-1)})) = -\mathcal{F}(u^{(n-1)}, m^{(n-1)}). \tag{5.1}$$

The Jacobian of $\mathcal{F}$ is given by

$$J\mathcal{F}(u, m)(\cdot, \cdot) = \begin{pmatrix} -\partial_t \cdot - \Delta \cdot + H_p(Du)\cdot & -\sigma F'(m)\cdot \\ -\operatorname{div}(mH_{pp}D\cdot) & \partial_t \cdot - \Delta \cdot - \operatorname{div}(H_p(Du)\cdot) \\ \cdot|_{t=T} - u_T(x) & 0 \\ 0 & \cdot|_{t=0} - m_0(x) \end{pmatrix},$$

where $F' = \frac{dF}{dm}$. Replacing in (5.1), we obtain for the first component of (5.1)

$$
\begin{aligned}
&-\partial_t(u^{(n)}-u^{(n-1)})-\Delta(u^{(n)}-u^{(n-1)})+H_p(Du^{(n-1)})(u^{(n)}-u^{(n-1)})\\
&-\sigma F'(m^{(n-1)})(m^{(n)}-m^{(n-1)})\\
&=\partial_t u^{(n-1)}+\Delta u^{(n-1)}-H(Du^{(n-1)})+\sigma F(m^{(n-1)})
\end{aligned}
$$

Now recalling that $q^{(n)}=H_p(Du^{(n-1)})$ and $H(Du^{(n-1)})=q^{(n)}Du^{(n-1)}-L(q^{(n)})$, the previous equation is equivalent to

$$
\begin{aligned}
&-\partial_t u^{(n)}-\Delta u^{(n)}+q^{(n)}Du^{(n)}-L(q^{(n)})=\sigma F(m^{(n-1)})\\
&+\sigma F'(m^{(n-1)})(m^{(n)}-m^{(n-1)})
\end{aligned}
\tag{5.2}
$$

with the final condition $u^{(n)}(x,T)=u_T(x)$. Note that, if $F=F(x)$ and therefore $F'\equiv 0$, we see by (5.2) that the policy iteration method is a Newton's method applied to the HJB equation. With similar computation we get

$$
\begin{aligned}
&\partial_t m^{(n)}-\Delta m^{(n)}-\operatorname{div}(m^{(n)}q^{(n)})\\
&=\operatorname{div}(m^{(n-1)}H_{pp}(Du^{(n-1)})(m^{(n)}-m^{(n-1)}))
\end{aligned}
\tag{5.3}
$$

with the initial condition $m^{(n)}(x,0)=m_0(x)$.

By Theorem 2.3, the terms on the right side of (5.2)-(5.3), which correspond to the off-diagonal entries of the Jacobian $J\mathcal{F}$, are infinitesimal. In the policy iteration method, we suppress these terms from the beginning, in order to remove the coupling between the two equations. In this sense, the policy iteration method can be interpreted as a quasi-Newton method since, instead of the full Jacobian of $\mathcal{F}$, we only use an approximation of it. In any case, after some iterations, the influence of the neglected terms is vanishing and the policy iteration method approximately behaves as a Newton's method, explaining the rapid convergence observed experimentally (see [9, Section 6]).

Some numerical examples have been considered in [19] for comparing the policy iteration method and the Newton method for solving MFGs. In many of these the policy iteration method turns out to be more efficient in terms of computing time.

References

[1] Y. Achdou, M. Laurière, P.L. Lions, Optimal control of conditioned processes with feedback controls, J. Math. Pures Appl. (9) 148 (2021) 308–341.
[2] A. Alla, M. Falcone, D. Kalise, An efficient policy iteration algorithm for dynamic programming equations, SIAM J. Sci. Comput. 37 (1) (2015) A181–A200.
[3] R. Bellman, Dynamic Programming, Princeton Univ. Press, Princeton, 1957.
[4] A. Bensoussan, Perturbation Methods in Optimal Control, Wiley/Gauthier-Villars Series in Modern Applied Mathematics, John Wiley & Sons, Ltd., Chichester, Gauthier-Villars, Montrouge, 1988.
[5] A. Bensoussan, J. Frehse, P. Yam, Mean Field Games and Mean Field Type Control Theory, Springer Briefs in Mathematics, Springer, New York, 2013.
[6] V.I. Bogachev, N.V. Krylov, M. Röckner, S.V. Shaposhnikov, Fokker-Planck-Kolmogorov Equations, Mathematical Surveys and Monographs, vol. 207, American Mathematical Society, Providence, RI, 2015.
[7] O. Bokanowski, S. Maroso, H. Zidani, Some convergence results for Howard's algorithm, SIAM J. Numer. Anal. 47 (4) (2009) 3001–3026.
[8] L.M. Briceño-Arias, D. Kalise, F.J. Silva, Proximal methods for stationary mean field games with local couplings, SIAM J. Control Optim. 56 (2) (2018) 801–836.
[9] S. Cacace, F. Camilli, A. Goffi, A policy iteration method for Mean Field Games, ESAIM Control Optim. Calc. Var. 27 (2021) 85, 19 pp.
[10] P. Cardaliaguet, S. Hadikhanloo, Learning in mean field games: the fictitious play, ESAIM Control Optim. Calc. Var. 23 (2) (2017) 569–591.
[11] N.V. Krylov, An analytic approach to SPDEs, in: Stochastic Partial Differential Equations: Six Perspectives, in: Math. Surveys Monogr., vol. 64, Amer. Math. Soc., Providence, RI, 1999, pp. 185–242.
[12] M. Cirant, D. Ghilli, Existence and non-existence for time-dependent mean field games with strong aggregation, Math. Ann. (2021), https://doi.org/10.1007/s00208-021-02217-3.

[13] M. Cirant, A. Goffi, Lipschitz regularity for viscous Hamilton-Jacobi equations with Lp terms, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 37 (4) (2020) 757–784.
[14] W.H. Fleming, Some Markovian optimization problems, J. Math. Mech. 12 (1963) 131–140.
[15] R. Howard, Dynamic Programming and Markov Processes, MIT Press, Cambridge, 1960.
[16] M. Huang, P.E. Caines, R.P. Malhame, Large-population cost-coupled LQG problems with non uniform agents: individual-mass behaviour and decentralized ϵ-Nash equilibria, IEEE Trans. Autom. Control 52 (2007) 1560–1571.
[17] B. Kerimkulov, D. Šiška, L. Szpruch, Exponential convergence and stability of Howards's policy improvement algorithm for controlled diffusions, SIAM J. Control Optim. 53 (2020) 1314–1340.
[18] O.A. Ladyzenskaja, V.A. Solonnikov, N.N. Ural'ceva, Linear and Quasilinear Equations of Parabolic Type, Translated from the Russian by S. Smith, Translations of Mathematical Monographs, vol. 23, American Mathematical Society, Providence, R.I., 1968.
[19] M. Laurière, J. Song, Q. Tang, Policy iteration method for time-dependent mean field games systems with non-separable Hamiltonians, arXiv:2110.02552, 2021.
[20] G. Metafune, D. Pallara, A. Rhandi, Global properties of transition probabilities of singular diffusions, Teor. Veroâtn. Primen. 54 (2009) 116–148.
[21] J.-M. Lasry, P.-L. Lions, Mean field games, Jpn. J. Math. 2 (2007) 229–260.
[22] M.L. Puterman, On the convergence of policy iteration for controlled diffusions, J. Optim. Theory Appl. 33 (1) (1981) 137–144.
[23] M.L. Puterman, S.L. Brumelle, On the convergence of policy iteration in stationary dynamic programming, Math. Oper. Res. 4 (1979) 60–69.
[24] M.S. Santos, J. Rust, Convergence properties of policy iteration, SIAM J. Control Optim. 42 (6) (2004) 2094–2115.
[25] M. Schechter, On L^p estimates and regularity, I, Am. J. Math. 85 (1963) 1–13.