SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

November 2025, Vol. 68, Iss. 11, 210206:1–210206:17
https://doi.org/10.1007/s11432-025-4646-9

Special Topic: Mean-Field Game and Control of Large Population Systems: From Theory to Practice

Policy iteration method for discounted infinite horizon mean field games: the semi-Lagrangian approach

Qing TANG[1*], Fabio CAMILLI[2] & Yong-Shen ZHOU[3]

[1]*School of Mathematics and Physics, China University of Geosciences (Wuhan), Wuhan 430074, China*
[2]*Department of Engineering and Geology, University of Chieti-Pescara "G. d'Annunzio", Pescara 65127, Italy*
[3]*Department of Basic and Applied Sciences for Engineering, Sapienza University of Rome, Rome 00161, Italy*

Received 4 April 2025/Revised 10 September 2025/Accepted 10 October 2025/Published online 3 November 2025

Abstract We study the policy iteration method for solving discounted infinite-horizon mean field games. At the continuous level, a policy iteration algorithm can be used to establish the existence and uniqueness of solutions for mean field games with a large discount factor λ. At a discrete level, it can be used to compute a solution of the problem. To implement the method, we employ a semi-Lagrangian method, where the Hamilton-Jacobi-Bellman equation is first discretized in time using the dynamic programming principle and then in space by projecting onto a grid. To support our theoretical findings, we present numerical examples in both one and two dimensions.

Keywords mean field games, policy iteration, numerical approximation, semi-Lagrangian method, dynamic programming

Citation Tang Q, Camilli F, Zhou Y-S. Policy iteration method for discounted infinite horizon mean field games: the semi-Lagrangian approach. Sci China Inf Sci, 2025, 68(11): 210206, https://doi.org/10.1007/s11432-025-4646-9

1 Introduction

Mean field games (MFG) theory was introduced in [1,2] as a framework to analyze dynamic games involving an infinite number of interacting agents. The theory provides a mathematical model to investigate and understand the collective behavior of agents who make decisions based on both their individual state and the statistical distribution of the states of other agents. The MFG framework couples a Hamilton-Jacobi-Bellman (HJB) equation, describing the optimal control of a representative agent, with a Fokker-Planck-Kolmogorov (FPK) equation, governing the evolution of the population distribution. For a comprehensive introduction and detailed exposition of the theory and its applications, we refer to [3,4]. For a general survey on numerical methods for solving MFGs, we refer to [5].

In this paper, we consider the discounted stationary MFG system (see [6]),

$$\begin{cases} \text{(i)} & \lambda u - \varepsilon\Delta u + H(x, Du) = f(x, m) + g[m](x), & \text{in } \mathbb{T}^d, \\ \text{(ii)} & -\varepsilon\Delta m - \operatorname{div}(mD_pH(x, Du)) = 0, & \text{in } \mathbb{T}^d, \\ & \displaystyle\int m dx = 1. \end{cases} \tag{1}$$

Here $\lambda, \varepsilon > 0$, $\mathbb{T}^d$ denotes the d-dimensional unit torus and $\mathcal{P}(\mathbb{T}^d)$ is the set of Borel probability measures on $\mathbb{T}^d$. The function $f : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}^+$ defines a local mean field interaction and $g : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}^+$ is a nonlocal interaction term. Throughout this paper we use the shorthand notation $\int = \int_{\mathbb{T}^d}$. This system arises from an infinite-horizon stochastic optimal control problem:

$$u(x) = \inf_q \mathbb{E}\left\{ \int_0^\infty e^{-\lambda t} \left(L(X_t, q_t) + f(X_t, m) + g[m](X_t)\right) dt \middle| X_0 = x \right\}, \; dX_t = -q_t dt + \sqrt{2\varepsilon} dW_t. \tag{2}$$

* Corresponding author (email: tangqingthomas@gmail.com)

In this case, the Hamiltonian is $H(x,p) = \sup_{q\in\mathbb{R}^d}\{pq - L(x,q)\}$. In this setup, an agent optimally controls his trajectory, taking a frozen probability distribution m as given. The associated value function u satisfies the HJB equation (1)(i). The invariant probability measure $\mathcal{L}(X_t)$ which describes the distribution of the agents is the solution to the FPK equation (1)(ii) generated by the policy $q^*(X_t) = D_pH(X_t, Du(X_t))$. A fixed-point condition is imposed, ensuring that the perceived law of motion matches the actual one, i.e., $\mathcal{L}(X_t) = m$. With λ sufficiently large, we show the solution to system (1) is unique by a contraction mapping argument. The existence and uniqueness of a classical solution to this system, for sufficiently small λ and a smoothing coupling g, were established in Section 3A of [6]. Here we extend the result to the case with also a local coupling.

The system (1) can be interpreted as an approximation, for small λ, of the ergodic MFG system:

$$\begin{cases} \text{(i)} \quad -\varepsilon\Delta u + H(x,Du) + \Lambda = f(x,m) + g[m](x), & \text{in } \mathbb{T}^d, \quad \displaystyle\int u dx = 0, \\ \text{(ii)} \quad -\varepsilon\Delta m - \operatorname{div}(mD_pH(x,Du)) = 0, & \text{in } \mathbb{T}^d, \quad \displaystyle\int m dx = 1. \end{cases} \tag{3}$$

This allows us to approximate the solution to the ergodic system by solving the discounted system with a vanishing discount factor λ and, to solve the latter system, we consider a policy iteration algorithm combined with a suitable discretization scheme.

Policy iteration, also known as Howard's algorithm, is a well-established method for solving optimal control problems by alternating between policy evaluation and a greedy update step (see [7, 8]). For MFG systems, this approach was first introduced in [9] and further developed in [10–13], where it was implemented with a finite difference (FD) approximation of the system.

In this work, we explore an alternative approach based on a semi-Lagrangian (SL) discretization of the MFG system. An SL scheme is a numerical method that exploit the dynamic programming principle, computing the value function by tracing backward along the characteristics of the controlled system. This method combines time-stepping with interpolation, providing stability and accuracy even for high-dimensional, nonlinear systems. SL schemes are widely used for solving Hamilton-Jacobi-Bellman equations in optimal control problems (see [14–17]). In particular, Policy iteration algorithms with SL schemes were considered in [8, 18]. For evolutive MFGs, SL schemes have been studied in [19–21] and more recently applied to solve an MFG price formation model in [22]. Lagrange-Galerkin schemes for MFGs have been studied in [23, 24]. SL schemes are used not only to approximate value functions in optimal control problems but also to construct approximate closed loop optimal controls. By contrast, the controls obtained by numerical methods based on the Pontryagin maximum principle (e.g., [25]) are in open loop form. A method of using approximate feedback control in SL schemes was proposed in [19, 20] for first and second order MFGs and in [21] for MFG with space-fractional diffusions.

In [9], policy iteration is formulated at the continuous level as a sequence of linear PDE systems, which are then solved using FD schemes. In contrast, the SL approach begins by discretizing the MFG system in time, and the resulting discrete nonlinear system, which remains an optimal control problem, is then solved via policy iteration. Hence, while the FD scheme follows a linearize-then-discretize strategy, the SL approach follows a discretize-then-linearize strategy. Indeed, while both methods benefit from policy iteration's ability to linearize the HJB equation, one key advantage of SL schemes over FD schemes is their ability to preserve the original optimal control structure of the problem.

2 Preliminaries

We first fix some notations. For $k \in \mathbb{N}$ and $\alpha \in (0,1]$, $\mathcal{C}^k$ denotes the space of k-times continuously differentiable functions on $\mathbb{T}^d$. We write $\mathcal{C}^{k,\alpha}(\mathbb{T}^d)$ for the Hölder space on $\mathbb{T}^d$ with the norm defined by $\|u\|_{\mathcal{C}^{k,\alpha}} = \sum_{|j|\leqslant k}\|\partial_j u\|_{L^\infty} + \sum_{|j|=k}[\partial_j u]_\alpha$ with $[u]_\alpha = \sup_{x,y\in\mathbb{T}^d, x\neq y}\frac{|u(x)-u(y)|}{|x-y|^\alpha}$. For brevity, we use $\mathcal{C}^\alpha$ for $\mathcal{C}^{0,\alpha}$. Unless otherwise specified, r denotes a constant such that $r > d$. We write L^r for the standard Lebesgue space and $W^{1,r}$, $W^{2,r}$ for the usual Sobolev spaces on $\mathbb{T}^d$. We recall the classical Sobolev embedding in Hölder spaces:

$$W^{1,r}(\mathbb{T}^d) \subset \mathcal{C}^{\alpha^*}(\mathbb{T}^d), \quad W^{2,r}(\mathbb{T}^d) \subset \mathcal{C}^{1,\alpha^*}(\mathbb{T}^d), \quad \text{where} \quad \alpha^* = 1 - \frac{d}{r}. \tag{4}$$

We will also recall the Poincaré's inequality on the torus: given $\phi \in W^{1,r}$ such that $\int \phi(x)dx = 0$, then there exists a constant C_{P} depending only on d such that

$$\|\phi\|_{L^r} \leqslant C_{\mathrm{P}} \|D\phi\|_{L^r}. \tag{5}$$

The space $\mathcal{P}(\mathbb{T}^d)$ of probability measures is endowed with the Wasserstein distance: for $m, m' \in \mathcal{P}(\mathbb{T}^d)$, $\mathbf{d}_1(m, m') = \sup_\phi \int_{\mathbb{T}^d} \phi(x) d(m - m')(x)$ where the supremum is taken over all 1-Lipschitz maps $\phi : \mathbb{T}^d \to \mathbb{R}$. Given a map $\mathcal{G} : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$, $\frac{\delta \mathcal{G}}{\delta m} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ denotes the flat derivative of $\mathcal{G}$ if, with the normalization $\int_{\mathbb{T}^d} \frac{\delta \mathcal{G}[m]}{\delta m}(x) \mathrm{d}m(x) = 0$,

$$\mathcal{G}[m'] - \mathcal{G}[m] = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta \mathcal{G}}{\delta m}[(1-s)m + sm'](x) d(m' - m)(x) ds.$$

We now state some assumptions in this paper. Throughout the paper, λ denotes a positive constant.

(A1) Let $\mathbf{I}$ denote the $d \times d$ identity matrix. For all $x \in \mathbb{T}^d$, $p \in \mathbb{R}^d$ and some $C_H > 0$, the Hamiltonian $H(x, p)$ satisfies

$$\begin{aligned} &H(\cdot, \cdot) \in \mathcal{C}^2(\mathbb{T}^d \times \mathbb{R}^d) \text{ and } \frac{1}{C_H}\mathbf{I} \leqslant D_{pp}H(x,p) \leqslant C_H \mathbf{I}, \\ &|D_p H(x,p)| + |D_{px}H(x,p)| \leqslant C_H(|p|+1), \ \ |D_{xx}H(x,p)| \leqslant C_H(|p|^2+1). \end{aligned} \tag{6}$$

(A2) $f : \mathbb{T}^d \times \mathbb{R}^+ \to \mathbb{R}$ is uniformly bounded and Lipschitz continuous in both variables.

(A3) $f'(x, \cdot) \geqslant 0$ for all $x \in \mathbb{T}^d$.

(A4) g, $\partial_{x_i} g$, $\partial_{x_i x_j} g$ and the measure derivative $\frac{\delta g}{\delta m} : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \to \mathbb{R}$ are all Lipschitz continuous.

(A5) For any $m, m' \in \mathcal{P}(\mathbb{T}^d)$, $\int_{\mathbb{T}^d} (g[m](x) - g[m'](x))(m - m')dx \geqslant 0$.

Assumptions (A3) and (A5) are the Lasry-Lions monotonicity conditions for the local and nonlocal couplings, respectively.

Remark 1. If $g : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ derives from potentials, i.e., there exists $\mathcal{G} : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$, $\frac{\delta \mathcal{G}}{\delta m}(x) = g[m](x)$, then the system can be defined as a potential MFG: the optimal control q^* and distribution m from system (1) can be obtained from considering the mean field control problem $\inf \mathcal{J}(m, q)$ such that

$$\mathcal{J} := \int_0^{+\infty} e^{-\lambda t} \left(\int L(x,q) m dx + \int_0^m f(x,\rho) d\rho + \mathcal{G}[m] \right) dt, \quad \text{s.t. } -\varepsilon \Delta m - \mathrm{div}(mq) = 0, \int m dx = 1.$$

Remark 2. An example of coupling term admitting a potential, which also satisfies the monotonicity condition (A5), is $g[m](x) = x \int x m dx, \mathcal{G}[m] = \frac{1}{2}\left(\int x m dx\right)^2$.

Next, we discuss the well posedness of system (1). We start considering a stability property of the stationary FPK equation.

Lemma 1 ([10, Lemma 4.2])**.** Given $q \in L^\infty(\mathbb{T}^d)$ and $B \in L^r(\mathbb{T}^d)$, let μ be the solution to

$$-\varepsilon \Delta \mu - \mathrm{div}(\mu q) = \mathrm{div}(B), \quad \int \mu dx = 0.$$

Then, there exists a constant C depending only on $\|q\|_{L^\infty}$ and d, such that $\|\mu\|_{W^{1,r}} \leqslant C \|B\|_{L^r}$.

Remark 3. The case $\|\mu\|_{W^{1,2}} \leqslant C\|B\|_{L^2}$ has been obtained in [6, Corollary 1.3]. We need the stronger result from [10, Lemma 4.2] in order to address models with local coupling $f(x, m)$.

We can then obtain the following.

Lemma 2. Let $\iota = \{1, 2\}$, $q_\iota \in L^\infty$, m_ι be the solution to

$$-\varepsilon \Delta m_\iota - \mathrm{div}(m_\iota q_\iota) = 0, \quad \int m_\iota dx = 1. \tag{7}$$

Then, there exists a positive constant $C = C(\|q_\iota\|_{L^\infty}, d)$ such that $\|m_\iota\|_{W^{1,r}} + \|m_\iota\|_{L^\infty} \leqslant C$. Moreover, there exists a constant C depending only on $\|q_\iota\|_{L^\infty}$ and d, such that

$$\|m_1 - m_2\|_{W^{1,r}} \leqslant C \|m_1(q_1 - q_2)\|_{L^r}.$$

We proceed to give a priori estimate of a classical solution to system (1).

Proposition 1. Assume (A1), (A2) and (A4); then for every classical solution (u, m) to system (1) it holds that

$$\|u\|_{L^\infty} \leqslant \frac{\|f\|_{L^\infty} + \|g\|_{L^\infty}}{\lambda}. \tag{8}$$

Moreover, there exists a positive constant $\mathsf{K} = \mathsf{K}(\|f\|_{L^\infty}, \|g\|_{L^\infty}, d)$ and $C(\mathsf{K}) > 0$ depending only on K such that

$$\|Du\|_{L^\infty} \leqslant \mathsf{K}, \tag{9}$$

$$\|u - \int u dx\|_{\mathcal{C}^\alpha} + \|\Delta u\|_{L^\infty} + \|m\|_{W^{1,r}} \leqslant C(\mathsf{K}). \tag{10}$$

Proof. We only show the estimate for $\|u - \int u dx\|_{\mathcal{C}^\alpha}$, as the others can be shown as in [26, Proposition 2.4]. In particular, Eq. (9) can be obtained by using [27, Theorem 1.1]. Let $\tilde{u} = u - \int u dx$. It is clear that $D\tilde{u} = Du$ and $\int \tilde{u} dx = 0$. Using Poincaré's inequality, we obtain $\|\tilde{u}\|_{L^r} \leqslant C\|Du\|_{L^r} \leqslant C\|Du\|_{L^\infty}$. Hence $\|\tilde{u}\|_{W^{1,r}} \leqslant C$. By Sobolev embedding we obtain (10).

Now we consider the existence of solution to system (1).

Proposition 2. Under assumptions (A1), (A2) and (A4), the system (1) has a classical solution.

Proof. Let $\varrho \subset \mathcal{C}^\alpha$, $\int \varrho(x) dx = 1$ and $\varrho > 0$. We consider the map $\Phi : \mathcal{C}^\alpha \to W^{1,r}$, $m = \Phi(\varrho)$ defined as

$$\begin{cases} \text{(i)} \quad \lambda u - \varepsilon \Delta u + H(x, Du) = f(x, \varrho) + g[\varrho](x), & \text{in } \mathbb{T}^d, \\ \text{(ii)} \quad -\varepsilon \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0, & \text{in } \mathbb{T}^d, \quad \int m dx = 1. \end{cases} \tag{11}$$

From Proposition 1, we can obtain a bound on $\|Du\|_{L^\infty}$ with (9) and a bound on $\|m\|_{W^{1,r}}$ with (10). With $r > d$ we have $W^{1,r}$ continuously embedded in $\mathcal{C}^{\alpha'}$ with $\alpha' > \alpha$. Hence the map Φ is a compact map from $\mathcal{C}^\alpha$ to $W^{1,r}$. From (A2) and (A4), $f(\varrho_k)$ and $g[\varrho_k](\cdot)$ are uniformly convergent for any uniformly convergent sequence ϱ_k. From viscosity solution theory, u_k converges uniformly to u. Using a classical semiconcavity argument, Du_k converges a.e. to Du. Since Du_k is uniformly bounded, by Egorov theorem it follows that Du_k converges strongly to Du in L^r. With Lemma 2, m_k converges to m in $W^{1,r}$. Hence the map Φ is continuous from $\mathcal{C}^\alpha$ to $W^{1,r}$. From Schauder fixed point theorem, there exists $m^* \in \mathcal{C}^\alpha$ such that $m^* = \Phi(m^*)$. Replacing ϱ by m^* in 11(i), we obtain a classical solution u^*. From the regularity of Du^*, m^* is in fact a classical solution and (u^*, m^*) is a classical solution to the system (1).

We now turn to the uniqueness of solutions to (1). In particular, we show that the map generated by a policy iteration algorithm satisfies a contraction property for sufficiently large λ. The policy iteration algorithm we consider was originally proposed in [9]. Given $\bar{q}^{(0)}$, iterate for each $n \geqslant 0$.

(i) Generate the distribution from the policy. Solve

$$-\varepsilon \Delta m^{(n)} - \operatorname{div}(m^{(n)} \bar{q}^{(n)}) = 0 \quad \text{in } \mathbb{T}^d, \quad \int m^{(n)} dx = 1. \tag{12}$$

(ii) Policy evaluation. Solve

$$\lambda u^{(n)} - \varepsilon \Delta u^{(n)} + \bar{q}^{(n)} Du^{(n)} - L(x, \bar{q}^{(n)}) = f(x, m^{(n)}) + g[m^{(n)}](x) \quad \text{in } \mathbb{T}^d. \tag{13}$$

(iii) Policy update.

$$q^{(n+1)}(x) = \arg\max_{|q| \leqslant K} \left\{ q Du^{(n)}(x) - L(x, q) \right\}. \tag{14}$$

(iv) Policy smoothing.

$$\bar{q}^{(n+1)} = \gamma_n q^{(n+1)} + (1 - \gamma_n) \bar{q}^{(n)}. \tag{15}$$

The Policy smoothing step was introduced in [13], where $\gamma_n \in (0, 1]$ can be a general relaxation parameter or learning rate.

For $K > 0$ as in (14), we define the truncated Hamiltonian

$$H_K(x, p) = \max_{|q| \leqslant K} \{qp - L(x, q)\}. \tag{16}$$

Note that $H_K(x, p)$ is Lipschitz in p with a constant depending only on K, i.e.,

$$|H_K(x, p_1) - H_K(x, p_2)| \leqslant C_{H_K} |p_1 - p_2|, \quad p_1, p_2 \in \mathbb{R}^d. \tag{17}$$

Theorem 1. Let (A1), (A2) and (A4) hold, $\gamma_n = 1$ for all n. If λ is sufficiently large, then the sequence $(u^{(n)}, m^{(n)}, q^{(n)})$ converges in $W^{1,r} \times W^{1,r} \times L^\infty$ to the solution of the discounted MFG.

Proof. Let K be a constant such that $K > C_H(\mathsf{K}+1)$. Here K is as in (1) and C_H as in assumption (A1), while $C(K)$ denotes a generic constant which may increase from line to line, but always depends only on K and d. Consider the system

$$\begin{cases} \text{(i)} \quad \lambda u - \varepsilon\Delta u + H_K(x, Du) = f(x,m) + g[m](x), & \text{in } \mathbb{T}^d, \\ \text{(ii)} \quad -\varepsilon\Delta m - \operatorname{div}(mD_pH_K(x,Du)) = 0, & \text{in } \mathbb{T}^d, \quad \displaystyle\int m dx = 1. \end{cases} \tag{18}$$

By writing the iterative procedure (12)–(14) as $u^{(n+1)} = \Phi(u^{(n)})$ when $n \geqslant 1$, a solution (u^*, m^*) to system (18) is equivalent to a fixed point $u^* \in W^{1,r}$ of Φ map. By Sobolev embedding, $\|u^{(n+1)} - u^{(n)}\|_{W^{1,r}} = 0$ implies $\|Du^{(n+1)} - Du^{(n)}\|_{L^\infty} = 0$; hence there exists $q^* \in L^\infty$ such that

$$q^*(x) = \arg\max_{|q| \leqslant K} \{qDu^*(x) - L(x,q)\}. \tag{19}$$

Moreover, $\|u^{(n)} - u^*\|_{W^{1,r}} = 0$ implies $q^{(n+1)} = q^{(n)} = q^*$ pointwise. From (12) and $\|q^{(n)}\|_{L^\infty} \leqslant K$, we have $\|m^{(n)}\|_{L^\infty} \leqslant C(K)$. Let $v^{(n+1)} = u^{(n+1)} - u^{(n)}$ and $\mu^{(n+1)} = m^{(n+1)} - m^{(n)}$, then

$$\begin{aligned} \lambda v^{(n)} - \varepsilon\Delta v^{(n+1)} - q^{(n+1)}Dv^{(n+1)} &= \mathsf{F}^{(n+1)}, \\ -\varepsilon\Delta\mu^{(n+1)} - \operatorname{div}(\mu^{(n+1)}q^{(n+1)}) &= \operatorname{div}(m^{(n)}(q^{(n+1)} - q^{(n)})), \end{aligned}$$

where

$$\begin{aligned} \mathsf{F}^{(n+1)} =& - q^{(n+1)}Du^{(n)} + L(x, q^{(n+1)}) + q^{(n)}Du^{(n)} - L(x, q^{(n)}) + f(x, m^{(n+1)}) \\ & - f(x, m^{(n)}) + g[m^{(n+1)}](x) - g[m^{(n)}](x). \end{aligned}$$

We give some L^r estimates on $\mathsf{F}^{(n+1)}$, independent of n. First we consider

$$\begin{aligned} & - q^{(n+1)}Du^{(n)} + L(x, q^{(n+1)}) + q^{(n)}Du^{(n)} - L(x, q^{(n)}) \\ &= -q^{(n+1)}Du^{(n)} + L(x, q^{(n+1)}) + q^{(n)}Du^{(n-1)} - L(x, q^{(n)}) + q^{(n)}(Du^{(n)} - Du^{(n-1)}) \\ &= H_K(x, Du^{(n-1)}) - H_K(x, Du^{(n)}) + q^{(n)}(Du^{(n)} - Du^{(n-1)}). \end{aligned}$$

From (17) and $\|q^{(n)}\|_{L^\infty} \leqslant K$, we have

$$\|H_K(\cdot, Du^{(n-1)}) - H_K(\cdot, Du^{(n)})\|_{L^r} \leqslant C_{H_K}\|Dv^{(n)}\|_{L^r}, \quad \|q^{(n)}(Du^{(n)} - Du^{(n-1)})\|_{L^r} \leqslant K\|Dv^{(n)}\|_{L^r}.$$

We can use (A2) and (A4) to obtain

$$\|f(\cdot, m^{(n+1)}) - f(\cdot, m^{(n)})\|_{L^r} + \|g[m^{(n+1)}](\cdot) - g[m^{(n)}](\cdot)\|_{L^r} \leqslant C\|m^{(n+1)} - m^{(n)}\|_{L^r} \leqslant C(K)\|Dv^{(n)}\|_{L^r}.$$

We therefore obtain

$$\begin{aligned} \|\mathsf{F}^{(n+1)}\|_{L^r} &\leqslant \|H_K(\cdot, Du^{(n-1)}) - H_K(\cdot, Du^{(n)})\|_{L^r} + \|q^{(n)}(Du^{(n)} - Du^{(n-1)})\|_{L^r} \\ &\quad + \|f(\cdot, m^{(n+1)}) - f(\cdot, m^{(n)})\|_{L^r} + \|g[m^{(n+1)}](\cdot) - g[m^{(n)}](\cdot)\|_{L^r} \\ &\leqslant C(K)\|Dv^{(n)}\|_{L^r}. \end{aligned}$$

Using the standard estimate of solution for linear elliptic equations (see [28, Chapter 8 Theorem 1 p. 158]),

$$\begin{aligned} \lambda\|v^{(n+1)}\|_{L^r} + \sqrt{\lambda}\|Dv^{(n+1)}\|_{L^r} + \|D^2v^{(n+1)}\|_{L^r} &\leqslant C(K)\|\mathsf{F}^{(n+1)}\|_{L^r} \\ &\leqslant C(K)\|Dv^{(n)}\|_{L^r}. \end{aligned} \tag{20}$$

With sufficiently large λ, we can then obtain $C(K)/\sqrt{\lambda} < 1$. From Banach fixed point theorem, the map $u^{(n+1)} = \Phi(u^{(n)})$ has a unique fixed point u^* in $W^{1,r}$. The sequence $u^{(n)}$ converges to u^* in $W^{1,r}$ and also uniformly by Sobolev embedding. With (17), the convergence of $Du^{(n)}$ in L^r to Du^* implies the convergence of $q^{(n)}$ to q^* in L^r. With (2), $m^{(n)}$ converges to m^* in $W^{1,r}$, where m^* is the solution to the FPK equation (7) with $q_\iota = q^*$. From (13), $Du^{(n)}$ converges uniformly to Du^*. Therefore, $q^{(n)}$ converges uniformly to q^* which solves (19). Finally we recall that with (9) and assumption (A1), in fact $H_K(x, Du^*) = H(x, Du^*)$ and $D_pH_K(x, Du^*) = D_pH(x, Du^*)$.

Now we turn to the case with small discount λ. The results for the ergodic approximation of a stationary MFG system with non-local (smoothing) couplings have already been obtained in Propositions 3.1 and 3.2 of [6]. Our contribution is to extend the result including the local-coupling case.

Proposition 3. Under assumptions (A1)–(A5), there exists $\lambda_0 > 0$ such that for all $0 < \lambda < \lambda_0$, the system (1) has a unique classical solution. Moreover, there exists a constant $\underline{C} > 0$, independent of λ, such that

$$m \geqslant \underline{C}. \tag{21}$$

Proof. To show (21), we observe that the FPK equation can be written as

$$-\varepsilon\Delta m - D_p H(x, Du)Dm - \operatorname{tr}\left(D_{pp}H(x, Du)D^2u\right) m = 0.$$

As $\|\Delta u\|_{L^\infty}$ and $\|Du\|_{L^\infty}$ can be bounded independently of λ, from strong maximum principle it follows a positive lower bound $\underline{C}$ on m independent of λ. For uniqueness, consider two solutions (u_1, m_1) and (u_2, m_2). Using the classical Lasry-Lions monotonicity argument, we obtain

$$\begin{aligned} C\int (m_1 + m_2)|Du_1 - Du_2|^2 dx \leqslant & \int (m_1 - m_2)(f(m_1) - f(m_2))dx \\ & + \int (m_1 - m_2)(g[m_1](x) - g[m_2](x))dx - \lambda\int (m_1 - m_2)(u_1 - u_2)dx. \end{aligned} \tag{22}$$

From $\int m_1 dx = \int m_2 dx = 1$ we have $\int (m_1 - m_2)\left(\int u_1 dx - \int u_2 dx\right) dx = 0$, hence

$$\begin{aligned} -\int (m_1 - m_2)(u_1 - u_2)dx &= -\int (m_1 - m_2)\left(u_1 - \int u_1 dx - u_2 + \int u_2 dx\right) dx \\ &\leqslant \frac{1}{2}\int |m_1 - m_2|^2 dx + \frac{1}{2}\int \left|u_1 - \int u_1 dx - u_2 + \int u_2 dx\right|^2 dx \\ &\leqslant \frac{1}{2}C_m \int |m_1(Du_1 - Du_2)|^2 dx + \frac{1}{2}C_{\mathrm{P}}\int |Du_1 - Du_2|^2 dx \\ &\leqslant \frac{1}{2}C_m\|m_1\|_{L^\infty}\int m_1|Du_1 - Du_2|^2 dx + \frac{1}{2}C_{\mathrm{P}}\int |Du_1 - Du_2|^2 dx, \end{aligned} \tag{23}$$

where C_{P} denotes the constant in (5) from using the Poincaré's inequality. It is important to notice that $C_m\|m_1\|_{L^\infty}$ and C_{P} do not depend on λ. With (21), we can replace the left side of (22) by $2C\underline{C}\int |Du_1 - Du_2|^2 dx$. By choosing λ sufficiently small, we obtain $\int |Du_1 - Du_2|^2 dx \leqslant 0$, therefore $Du_1 = Du_2$ a.e.

To consider the convergence of the system (1) as $\lambda \to 0$, we write

$$\begin{cases} \text{(i)} \quad \lambda u_\lambda - \varepsilon\Delta u_\lambda + H(x, Du_\lambda) = f(x, m_\lambda) + g[m_\lambda](x), & \text{in } \mathbb{T}^d, \\ \text{(ii)} \quad -\varepsilon\Delta m_\lambda - \operatorname{div}(m_\lambda D_p H(x, Du_\lambda)) = 0, & \text{in } \mathbb{T}^d, \quad \int m_\lambda dx = 1. \end{cases} \tag{24}$$

Theorem 2. Let assumptions (A1)–(A5) hold. Let $(\hat{u}, \hat{m}, \Lambda)$ denote the solution to the ergodic problem (3). Then

$$\|Du_\lambda - D\hat{u}\|_{L^2} + \|m_\lambda - \hat{m}\|_{L^2} \leqslant C\lambda^{1/2}. \tag{25}$$

Moreover, as $\lambda \to 0$, $\|Du_\lambda - D\hat{u}\|_{L^r} + \|m_\lambda - \hat{m}\|_{W^{1,r}} + \|u_\lambda - \int u_\lambda - \hat{u}\|_{L^\infty} + \|\lambda u_\lambda - \Lambda\|_{L^\infty} \to 0$.

The estimate (25) is shown exactly the same way as [6, Proposition 3.2]. The main difference in our result is that, with $f(x, m)$ being a local coupling, the L^2 convergence of m_λ is not enough for the uniform convergence of the corrector $u_\lambda - \int u_\lambda$. We only focus on the additional steps in the proof.

Proof. From Egorov theorem, $Du_\lambda - D\hat{u} \to 0$ a.e. with $\|Du_\lambda - D\hat{u}\|_{L^\infty}$ bounded independently of λ implies $\|Du_\lambda - D\hat{u}\|_{L^r} \to 0$ for any $1 < r < \infty$. The convergence $\|m_\lambda - \hat{m}\|_{W^{1,r}} \to 0$ follows then from (2). Moreover, $\|u_\lambda - \int u_\lambda - \hat{u}\|_{W^{1,r}} \to 0$ from Poincaré's inequality and we obtain $\|u_\lambda - \int u_\lambda - \hat{u}\|_{\mathcal{C}^\alpha} \to 0$ from Sobolev embedding.

Remark 4. We have obtained the uniqueness of solution to system (1) with sufficiently small or large λ. However, this does not imply uniqueness of solution to system (1) for an arbitrary λ.

3 Numerical method

We now introduce a policy iteration method based on semi-Lagrangian schemes. To illustrate the main ideas, it is convenient to separate at first between time and space discretization.

3.1 The semi-discrete system

We first consider the discrete in time infinite horizon control problem. Define a time grid $\mathcal{G}_h = \{t_k = kh : k \in \mathbb{Z}^+ \cup \{0\}\}$ with a positive constant h. Consider the discretized problem (see [29]),

$$\begin{cases} u_h(x) := \inf\limits_q \left\{ \sum\limits_{k=0}^{\infty} (1-\lambda h)^k \left(L\left(X_h(t_k), q(t_k)\right) + f(X_h(t_k), m_h)\right) dt \Big| X_h(t_0) = x \right\}, \\ X_h(t_k + h) = X_h(t_k) - hq(t_k) + \sqrt{2d\varepsilon h} \sum\limits_{\iota=1}^{d} \xi_\iota^k, \quad X(t_0) = x, \end{cases} \tag{26}$$

where ξ^k is a sequence of i.i.d random variables such that $\mathbb{P}(\xi_\iota^k = 1) = \mathbb{P}(\xi_\iota^k = -1) = \frac{1}{2d}$ and $\mathbb{P}\left(\bigcup_{\iota,j}^{d} \{\xi_\iota^k \neq 0\} \cap \{\xi_j^k \neq 0\}\right) = 0$.

We obtain a semi-discretized version of system (1):

$$\begin{cases} \text{(i)} \quad u_h(x) = \inf\limits_q \left\{ (1-\lambda h)\mathcal{A}_h(q)u_h(x) + h\left(L(x,q) + f(x,m_h) + g[m_h](x)\right)\right\}, \\ \text{(ii)} \quad \int \mathcal{A}_h(q^*)\phi(x) dm_h(x) - \int \phi(x) dm_h(x) = 0, \quad \int m_h(x)dx = 1, \\ \text{(iii)} \quad q^* = \arg\min\limits_q \left\{ (1-\lambda h)\mathcal{A}_h(q)u_h(x) + hL(x,q)\right\}. \end{cases} \tag{27}$$

Here $\mathcal{A}_h(q)$ denotes the Markov chain transition operator:

$$\mathcal{A}_h(q)\phi(x) = \sum_{\iota=1}^{d} \frac{1}{2d} \left(\phi(x - hq + \mathbf{e}_\iota \sqrt{2d\varepsilon h}) + \phi(x - hq - \mathbf{e}_\iota \sqrt{2d\varepsilon h}) \right), \tag{28}$$

where $\mathbf{e}_\iota$ denotes a d-dimensional canonical basis, i.e., $\mathbf{e}_\iota = (0, \cdots, \underbrace{1}_{\iota-\text{entry}}, \cdots, 0)$. The HJB equation (27)(i) is derived by dynamic programming principle, for any given $m_h \in \mathcal{P}(\mathbb{T}^d) \cap \mathcal{C}(\mathbb{T}^d)$. The discrete in time FPK equation (ii) is derived as follows. For a given policy q, the evolution of probability density for the flow $X(t_k)$ may be characterized by the measure push-forward:

$$\begin{aligned} \int \phi(x) m_h(t_k + h, x) dx &= \mathbb{E}\left[\int \phi(X(t_k + h)) m_h(t_k, x) dx \Big| X(t_k) = x \right] \\ &= \int \mathcal{A}_h(q)\phi(x) m_h(t_k, x) dx. \end{aligned} \tag{29}$$

The equation (iii) in (27) is the fixed point characterizing the Nash equilibrium: if q^* is the optimal policy for the problem solved by (27)(i), then the probability density m_h is also generated by making the agents adopt q^*.

We define the semi-discretized policy iteration method. Given $\bar{q}^{(0)}$, iterate for each $n \geqslant 0$.

(i) Generate the distribution from the policy. Solve $m_h^{(n)}$ for all ϕ,

$$\int \mathcal{A}_h(\bar{q}^{(n)})\phi(x) dm_h^{(n)}(x) - \int \phi(x) dm_h^{(n)}(x) = 0, \quad \int m_h^{(n)} dx = 1. \tag{30}$$

(ii) Policy evaluation. Solve

$$u_h^{(n)}(x) = (1-\lambda h)\mathcal{A}_h(\bar{q}^{(n)})u_h^{(n)}(x) + hL(x, \bar{q}^{(n)}) + hf(x, m_h^{(n)}) + hg[m_h](x). \tag{31}$$

(iii) Policy update.

$$q^{(n+1)} = \arg\min_q \left\{ (1-\lambda h)\mathcal{A}_h(q)u_h^{(n)}(x) + hL(x,q) \right\}. \tag{32}$$

(iv) Policy smoothing.

$$\bar{q}^{(n+1)} = \gamma_n q^{(n+1)} + (1-\gamma_n)\bar{q}^{(n)}. \tag{33}$$

3.2 Fully-discretized system

3.2.1 *One dimensional case*

Next we consider the policy iteration on the fully discretized system. To highlight the main idea, we first restrict the discussion to the $1d$ problem. We define the time-space grid $\mathcal{G}_{h,i} = \{(t_k, x_i) = (kh, i\rho) : k \in \mathbb{Z}^+ \cup \{0\}, i = 0, 1, \cdots, N-1\}$ with positive constants h and ρ. The index operator

$$[\cdot] = \{(\cdot + N) \bmod N\} \tag{34}$$

will be used to account for the periodic boundary conditions.

The vector U_i gives an approximation of u at x_i. We set $L_i = L(x_i, Q_i^*)$, $f_i = f(x_i, M_i)$ and $g_i = g[M](x_i)$. For approximating $u(x_i - hq_i^{(n)} + \sqrt{2d\varepsilon h})$ and $u(x_i - hq_i^{(n)} - \sqrt{2d\varepsilon h})$ we use an interpolation method. Consider the set of $\mathbb{P}_1$ basis functions $\{\beta_i\}$ defined by

$$\beta_i(x) = \max\left\{1 - \frac{|x - x_i|}{\rho}, 0\right\}. \tag{35}$$

It is clear that $0 \leqslant \beta_i(x) \leqslant 1$, $\sum_i \beta_i(x) = 1$ for all $x \in \mathbb{T}^d$ and $\beta_i(x_j) = \delta_{ij}$ where δ_{ij} denotes the Kronecker delta function. We can then define the interpolation

$$I[\phi](x) = \sum_i \phi(x_i)\beta_i(x). \tag{36}$$

Let $(U_{\rho,h}, M_{\rho,h})$ denote the solution to the fully discrete system with q discretized by Q^*,

$$\begin{cases} U_i = \dfrac{1 - \lambda h}{2}\left(\displaystyle\sum_j \beta_j(x_i - hQ_i^* + \sqrt{2d\varepsilon h})U_j + \sum_j \beta_j(x_i - hQ_i^* - \sqrt{2d\varepsilon h})U_j\right) + h(L_i + f_i + g_i), \\ M_i = \dfrac{1}{2}\left(\displaystyle\sum_j \beta_i(x_j - hQ_j^* + \sqrt{2d\varepsilon h})M_j + \sum_j \beta_i(x_j - hQ_j^* - \sqrt{2d\varepsilon h})M_j\right). \end{cases} \tag{37}$$

In order to characterize the discretization of optimal control Q^*, we have two approaches. The first one is to use system (27)(iii):

$$\mathrm{q}_i^* = \arg\min_{Q_i}\left\{\frac{1 - \lambda h}{2}\left(\sum_j \beta_j(x_i - hQ_i + \sqrt{2d\varepsilon h})U_j + \sum_j \beta_j(x_i - hQ_i - \sqrt{2d\varepsilon h})U_j\right) + L(x_i, Q_i)\right\}. \tag{38}$$

The second approach is to follow the method from [19–21] and use approximate feedback control

$$q_{\mathrm{num}}^* = D_p H(\cdot, D\hat{u}^\delta), \quad \text{where } \hat{u}_\delta = \hat{u} * \eta_\delta \quad \text{and } Q_i^* = q_{\mathrm{num}}^*(x_i). \tag{39}$$

In (39), $\hat{u}$ is the piecewise constant interpolation of U, and η_δ is the mollifier $\frac{1}{\delta}\eta(\frac{x}{\delta})$ for $0 \leqslant \eta \in \mathcal{C}_0^\infty(\mathbb{T}^d)$ with $\int \eta dx = 1$. This method has the advantage in efficiency and it preserves the semiconcavity of the numerical solution. Theoretically, it has been shown in [19–21] that if the system (1) has a unique classical solution (u, m), the SL scheme with the approximate feedback control method is convergent.

Theorem 3. Let assumptions (A1)–(A5) hold. Let $(U_{\rho,h,\delta}, M_{\rho,h,\delta})$ be the solution to the system given by (27) and (39). If $\delta \to 0$, $\frac{h}{\delta^2} \to 0$ and $\rho^2/h \to 0$, then the sequence $(U_{\rho,h,\delta}, M_{\rho,h,\delta})$ converges to (u, m) uniformly.

The proof of Theorem (3) will be very long and similar to [20, Theorem 4.2]; hence we omit it.

Now we consider the policy iteration method for the fully discretized system (27) and (39). The full discretization of (31) becomes

$$\begin{aligned} U_i^{(n)} = {} & \frac{1 - \lambda h}{2}\left(\sum_j \beta_j(x_i - hQ_i^{(n)} + \sqrt{2d\varepsilon h})U_j^{(n)} + \sum_j \beta_j(x_i - hQ_i^{(n)} - \sqrt{2d\varepsilon h})U_j^{(n)}\right) \\ & + h(L_i^{(n)} + f_i^{(n)} + g_i^{(n)}), \end{aligned} \tag{40}$$

where $L_i^{(n)} = L(x_i, Q_i^{(n)})$, $f_i^{(n)} = f(x_i, M_i^{(n)})$ and $g_i^{(n)} = g[M^{(n)}](x_i)$. To fully discretize (30), we test (29) with $\phi(x) = \beta_i(x)$, then

$$M_i^{(n)} = \frac{1}{2}\left(\sum_j \beta_i(x_j - hQ_j^{(n)} + \sqrt{2d\varepsilon h})M_j^{(n)} + \sum_j \beta_i(x_j - hQ_j^{(n)} - \sqrt{2d\varepsilon h})M_j^{(n)}\right). \tag{41}$$

A possible approach to update policy would be to use (32) and (38):

$$\begin{aligned} & q_i^{(n+1)} \\ &= \arg\min_{Q_i}\left\{\frac{1-\lambda h}{2}\left(\sum_j \beta_j(x_i - hQ_i + \sqrt{2d\varepsilon h})U_j^{(n)} + \sum_j \beta_j(x_i - hQ_i - \sqrt{2d\varepsilon h})U_j^{(n)}\right) + hL(x_i, Q_i)\right\}. \end{aligned}$$

In practice, it is more efficient to approximate the feedback control method (39). Let $\hat{u}^{(n)}$ be the piecewise constant interpolation of $U^{(n)}$. Then we set

$$q_{\text{num}}^{(n+1)} = D_p H(\cdot, D\hat{u}_\delta^{(n)}). \tag{42}$$

Let $\mathbf{A}(Q)$ denote the $N \times N$ transition matrix where the (i,j)-element is defined as

$$\mathbf{A}_{ij}(Q) = \frac{1}{2d}\left(\beta_j(x_i - hQ_i + \sqrt{2d\varepsilon h}) + \beta_j(x_i - hQ_i - \sqrt{2d\varepsilon h})\right).$$

Clearly, $\mathbf{A}_{ij}(Q)$ is the discretized form of the operator $\mathcal{A}_h(q)$ introduced in (28). We observe that the transposed matrix $\mathbf{A}^{\mathsf{T}}$ is defined by

$$\mathbf{A}_{ij}^{\mathsf{T}}(Q) = \frac{1}{2d}\left(\beta_i(x_j - hQ_j + \sqrt{2d\varepsilon h}) + \beta_i(x_j - hQ_j - \sqrt{2d\varepsilon h})\right).$$

We now introduce the policy iteration algorithm for solving the fully discretized system. Recall that let $\mathbf{I}$ denote the $d \times d$ identity matrix. Let $\mathbf{L}^{(n)}$ and $\mathbf{f}^{(n)}$ denote d-dimensional vectors with elements $L_i^{(n)}$ and $f_i^{(n)}$. For (45), we choose a reasonably large $K > 0$. At the end of iteration we check the output vector $\check{Q}$ satisfies $\max_i\{|\check{Q}_i|\} < K$.

We make some further observations on the structure of the matrix $\mathbf{A}(Q)$, which is crucial to the implementation of Algorithm 1.

For each point x_i, there exist $j(i), j'(i) \in \mathbb{Z}$ such that $x_{j(i)} \leqslant x_i - hQ_i + \sqrt{2d\varepsilon h} \leqslant x_{j(i)+1}$ and $x_{j'(i)} \leqslant x_i - hQ_i - \sqrt{2d\varepsilon h} \leqslant x_{j'(i)+1}$. Letting $\lfloor\cdot\rfloor$ denote the floor function, $j(i), j'(i)$ are defined as

$$j(i) = \left\lfloor \frac{x_i - hQ_i + \sqrt{2d\varepsilon h}}{\rho} \right\rfloor, \quad j'(i) = \left\lfloor \frac{x_i - hQ_i - \sqrt{2d\varepsilon h}}{\rho} \right\rfloor. \tag{43}$$

With (34), we obtain

$$\begin{aligned} \sum_j \beta_j(x_i - hQ_i + \sqrt{2d\varepsilon h})U_j = & \left(\frac{-hQ_i + \sqrt{2d\varepsilon h}}{\rho} - \left\lfloor \frac{-hQ_i + \sqrt{2d\varepsilon h}}{\rho} \right\rfloor\right) U_{[j(i)+1]} \\ & + \left(1 - \frac{-hQ_i + \sqrt{2d\varepsilon h}}{\rho} + \left\lfloor \frac{-hQ_i + \sqrt{2d\varepsilon h}}{\rho} \right\rfloor\right) U_{[j(i)]}, \\ \sum_j \beta_j(x_i - hQ_i - \sqrt{2d\varepsilon h})U_j = & \left(\frac{-hQ_i - \sqrt{2d\varepsilon h}}{\rho} - \left\lfloor \frac{-hQ_i - \sqrt{2d\varepsilon h}}{\rho} \right\rfloor\right) U_{[j'(i)+1]} \\ & + \left(1 - \frac{-hQ_i - \sqrt{2d\varepsilon h}}{\rho} + \left\lfloor \frac{-hQ_i - \sqrt{2d\varepsilon h}}{\rho} \right\rfloor\right) U_{[j'(i)]}. \end{aligned} \tag{44}$$

Therefore, with $d = 1$ each row of the matrix $\mathbf{A}(Q)$ has at most 5 nonzero entries. The positions of nonzero entries for each row in $\mathbf{A}(Q^{(n)})$ depend on $Q^{(n)}$.

Algorithm 1 Policy Iteration method: semi-Lagrangian schemes.

Data: Initial values $\bar{Q}^{(0)}$, and positive parameters $\lambda, \rho, h, \varepsilon, \delta, K$.
Result: solution (U, M).

while $\|M^{(n+1)} - M^{(n)}\| > 10^{-5}$ **do**

Solve the FPK equation (41): $\left(\mathbf{I} - \mathbf{A}^{\mathsf{T}}(\bar{Q}^{(n)})\right) M^{(n)} = 0$;

Updating the Lagrangian and mean field interaction terms: $L_i^{(n)} = L(x_i, Q_i^{(n)}), \quad f_i^{(n)} = f(x_i, M_i^{(n)}), \quad g_i^{(n)} = g[M^{(n)}](x_i)$;

Solve the HJB equation (40): $\left(\mathbf{I} - (1-\lambda h)\mathbf{A}(\bar{Q}^{(n)})\right) U^{(n)} = h\mathbf{L}^{(n)} + h\mathbf{f}^{(n)} + h\mathbf{g}^{(n)}$;

Update approximate feedback control $q_{\text{num}}^{(n+1)}$ with (42):

$$Q_i^{(n+1)} = \min\{\max\{q_{\text{num}}^{(n+1)}(x_i), -K\}, K\}; \tag{45}$$

Policy smoothing $\bar{Q}^{(n+1)} = \gamma_n Q^{(n+1)} + (1-\gamma_n)\bar{Q}^{(n)}$.

end while

Remark 5. We consider a situation where the construction of matrix $\mathbf{A}(Q)$ is particularly simple and the SL scheme is very similar to an implicit finite difference schemes. If the optimal control q and viscosity ε are small then it might happen

$$\max\{|hQ_i + \sqrt{2d\varepsilon h}|, |hQ_i - \sqrt{2d\varepsilon h}|\} \leqslant \rho, \quad \text{s.t. } x_{[i-1]} \leqslant x_i - hQ_i \pm \sqrt{2d\varepsilon h} \leqslant x_{[i+1]}. \tag{46}$$

We can write the interpolation operators in the form

$$\begin{aligned}
&\sum_j \beta_j(x_i - hQ_i + \sqrt{2d\varepsilon h})U_j \\
&\quad = \frac{(-hQ_i + \sqrt{2d\varepsilon h})_+}{\rho} U_{[i+1]} + \frac{(-hQ_i + \sqrt{2d\varepsilon h})_-}{\rho} U_{[i-1]} + \left(1 - \frac{|-hQ_i + \sqrt{2d\varepsilon h}|}{\rho}\right) U_i, \\
&\sum_j \beta_j(x_i - hQ_i - \sqrt{2d\varepsilon h})U_j \\
&\quad = \frac{(-hQ_i - \sqrt{2d\varepsilon h})_+}{\rho} U_{[i+1]} + \frac{(-hQ_i - \sqrt{2d\varepsilon h})_-}{\rho} U_{[i-1]} + \left(1 - \frac{|-hQ_i - \sqrt{2d\varepsilon h}|}{\rho}\right) U_i.
\end{aligned}$$

In this case, $\mathbf{A}(Q)$ is in fact a tri-diagonal matrix. The central diagonal is $2 - \frac{|-hQ_i+\sqrt{2d\varepsilon h}|}{\rho} - \frac{|-hQ_i-\sqrt{2d\varepsilon h}|}{\rho}$ The upper and lower diagonals are, respectively, $\frac{(-hQ_i+\sqrt{2d\varepsilon h})_+}{\rho} + \frac{(-hQ_i-\sqrt{2d\varepsilon h})_+}{\rho}$, $\frac{(-hQ_i+\sqrt{2d\varepsilon h})_-}{\rho} + \frac{(-hQ_i-\sqrt{2d\varepsilon h})_-}{\rho}$. Under (46), the matrices $\mathbf{I} - (1-\lambda h)\mathbf{A}(Q)$ and $\mathbf{I} - \mathbf{A}^{\mathsf{T}}(Q)$ are tri-diagonal M-matrices. Therefore the scheme is stable. Moreover, Eq. (41) can be understood as a fully discrete Kolmogorov equation. We can write $\left(\mathbf{I} - \mathbf{A}^{\mathsf{T}}(Q)\right) M = 0$ more explicitly as, with $d = 1$,

$$\begin{aligned}
\underbrace{\left(\left|-Q_i + \sqrt{\frac{2\varepsilon}{h}}\right| + \left|-Q_i - \sqrt{\frac{2\varepsilon}{h}}\right|\right) M_i}_{\text{flux out of } x_i} = &\underbrace{\left(\left(-Q_{[i-1]} + \sqrt{\frac{2\varepsilon}{h}}\right)_+ + \left(-Q_{[i-1]} - \sqrt{\frac{2\varepsilon}{h}}\right)_+\right) M_{[i-1]}}_{\text{flux from } x_{[i-1]} \to x_i} \\
&+ \underbrace{\left(\left(-Q_{[i+1]} + \sqrt{\frac{2\varepsilon}{h}}\right)_- + \left(-Q_{[i+1]} - \sqrt{\frac{2\varepsilon}{h}}\right)_-\right) M_{[i+1]}}_{\text{flux from } x_{[i+1]} \to x_i}.
\end{aligned} \tag{47}$$

3.2.2 *Schemes for two dimensional systems*

Next, we sketch the SL for the HJB equation in the $2d$ case. We define the (x, y)-space gird $\mathcal{M}_{i,j} = \{(x_i, y_j) = (i\rho_x, j\rho_y) : i = 0, 1, \cdots, N_x - 1, j = 0, 1, \cdots, N_y - 1\}$ with positive constants ρ_x and ρ_y. Here $\rho_x = x_{i+1} - x_i$ and $\rho_y = y_{j+1} - y_j$ are the grid spacings in the x- and y-directions, respectively. We approximate $u(x_i, y_j)$ and $m(x_i, y_j)$ by $U_{i,j}$ and $M_{i,j}$. The control $q(x_i, y_j)$ is approximated by $Q_{i,j} = (Q_{i,j} \cdot \mathbf{e}_x, Q_{i,j} \cdot \mathbf{e}_y)$. Consider the set of $\mathbb{Q}_1$ basis functions $\beta_{i,j}$ defined by the tensor product of the

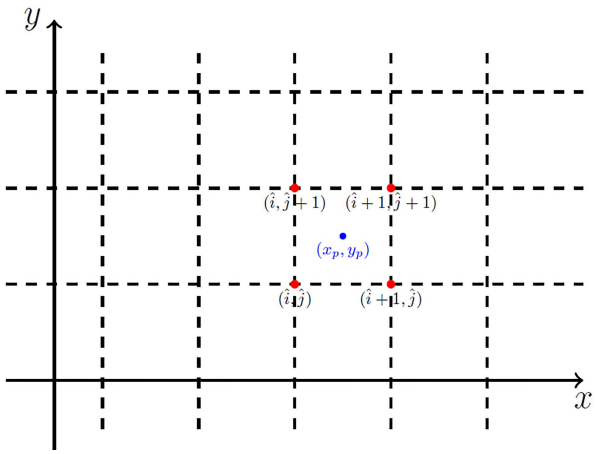

Figure 1 (Color online) Graphic illustration of $2d$ $\mathbb{Q}_1$ interpolation.

1D basis functions: $\beta_i(x) = \max\left\{1 - \frac{|x-x_i|}{\rho_x}, 0\right\}$, $\beta_j(y) = \max\left\{1 - \frac{|y-y_j|}{\rho_y}, 0\right\}$. Using the tensor product of these basis functions, we define the bilinear interpolation as $I[\phi](x,y) = \sum_{i,j} \phi(x_i, y_j)\beta_i(x)\beta_j(y)$.

We now illustrate the method for constructing the $2d$ $\mathbb{Q}_1$ interpolation. Let (x_p, y_p) denote a generic point, enclosed by grid points: $(x_{\hat{i}}, y_{\hat{j}}), (x_{\hat{i}}, y_{\hat{j}+1}), (x_{\hat{i}+1}, y_{\hat{j}}), (x_{\hat{i}+1}, y_{\hat{j}+1})$, as shown in Figure 1.

The fully discretization of the HJB equation (27)(i) becomes (see [23, 24])

$$\begin{aligned} U_{i,j} =& \frac{1-\lambda h}{4} \sum_{k,l} \beta_k(x_i - hQ^*_{i,j}\cdot \mathbf{e}_x + \sqrt{4\varepsilon h})\beta_l(y_j - hQ^*_{i,j}\cdot \mathbf{e}_y)U_{k,l} \\ &+ \frac{1-\lambda h}{4} \sum_{k,l} \beta_k(x_i - hQ^*_{i,j}\cdot \mathbf{e}_x - \sqrt{4\varepsilon h})\beta_l(y_j - hQ^*_{i,j}\cdot \mathbf{e}_y)U_{k,l} \\ &+ \frac{1-\lambda h}{4} \sum_{k,l} \beta_k(x_i - hQ^*_{i,j}\cdot \mathbf{e}_x)\beta_l(y_j - hQ^*_{i,j}\cdot \mathbf{e}_y + \sqrt{4\varepsilon h})U_{k,l} \\ &+ \frac{1-\lambda h}{4} \sum_{k,l} \beta_k(x_i - hQ^*_{i,j}\cdot \mathbf{e}_x)\beta_l(y_j - hQ^*_{i,j}\cdot \mathbf{e}_y - \sqrt{4\varepsilon h})U_{k,l} + h(L_{i,j} + f_{i,j} + g_{i,j}), \end{aligned} \tag{48}$$

where $L_{i,j} = L(x_i, y_j, Q^*_{i,j})$, $f_{i,j} = f(x_i, y_j, M_{i,j})$ and $g_{i,j} = g[M](x_i, y_j)$.

Taking $\phi(x,y) = \beta_{i,j}(x,y)$ in (29), the fully discretized FPK equation is given by the dual formulation:

$$\begin{aligned} M_{i,j} =& \frac{1}{4} \sum_{k,l} \beta_i(x_k - hQ^*_{k,l}\cdot \mathbf{e}_x + \sqrt{4\varepsilon h})\beta_j(y_l - hQ^*_{k,l}\cdot \mathbf{e}_y)M_{k,l} \\ &+ \frac{1}{4} \sum_{k,l} \beta_i(x_k - hQ^*_{k,l}\cdot \mathbf{e}_x - \sqrt{4\varepsilon h})\beta_j(y_l - hQ^*_{k,l}\cdot \mathbf{e}_y)M_{k,l} \\ &+ \frac{1}{4} \sum_{k,l} \beta_i(x_k - hQ^*_{k,l}\cdot \mathbf{e}_x)\beta_j(y_l - hQ^*_{k,l}\cdot \mathbf{e}_y + \sqrt{4\varepsilon h})M_{k,l} \\ &+ \frac{1}{4} \sum_{k,l} \beta_i(x_k - hQ^*_{k,l}\cdot \mathbf{e}_x)\beta_j(y_l - hQ^*_{k,l}\cdot \mathbf{e}_y - \sqrt{4\varepsilon h})M_{k,l}. \end{aligned} \tag{49}$$

We now show that even though our methodology and Algorithm 1 were introduced in the one dimensional case, they can be easily adapted to study a dimensional problem. The $N_x \times N_y$ matrix U can be reshaped into the $N_x \times N_y$ dimensional vector $\tilde{U}$:

$$u(x_i, y_j) \to \{u(x_0, y_0), u(x_0, y_1), \ldots, u(x_0, y_{N_y-1}), \ldots, u(x_{N_x-1}, y_{N_y-1})\}, \tag{50}$$

and similar M into $\tilde{M}$. For each given Q^*, we can construct a $(N_x \times N_y) \times (N_x \times N_y)$ matrix $\mathbf{A}(Q^*)$. Let $\mathbf{L}$, $\mathbf{f}$ and $\mathbf{g}$ denote the matrices with entries $L_{i,j}$, $f_{i,j}$, $g_{i,j}$. $\tilde{\mathbf{L}}$, $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}$ are their reshaped vectors

following the same method as in (50). We can then write (48) and (49) in the matrix form

$$\left(\mathbf{I}-(1-\lambda h)\mathbf{A}(\tilde{Q}^*)\right)\tilde{U}=h\tilde{\mathbf{L}}+h\tilde{\mathbf{f}}+h\tilde{\mathbf{g}} \quad \text{and} \quad \left(\mathbf{I}-\mathbf{A}^{\mathsf{T}}(\tilde{Q}^*)\right)\tilde{M}=0.$$

Remark 6. For the discretization of the HJB equation in $2d$ we followed the area-weighting method as in [23, 24]. For the FPK equation, we use the dual formulation. In practice, this involves a matrix transposition as in the $1d$ case. The comparison with the schemes in [23, 24], and a rigorous convergence analysis of our schemes in $2d$ will be addressed in our future work. The effectiveness of our schemes will be shown in the numerical section.

3.3 Alternative approach: finite difference schemes

In this section, we use a finite difference scheme to implement the policy iteration method, with a space grid in $1d$: $\mathcal{G}_i=\{x_i=i\rho : i=0,1,\cdots,Nx-1\}$. U_i and M_i approximate u and m at x_i. Similarly, f_i and g_i approximate $f(x_i,M_i)$ and $g[M](x_i)$. We introduce the finite difference operators:

$$(\Delta_\rho U)_i=\frac{U_{[i-1]}-2U_i+U_{[i+1]}}{\rho^2},\quad (DU)_i=\frac{(U_{[i+1]}-U_i)}{\rho},\quad [\nabla U]_i=\left(\underbrace{(DU)_i}_{\text{forward}},\underbrace{(DU)_{[i-1]}}_{\text{backward}}\right)^{\mathsf{T}}.$$

Discrete Hamiltonian (see [5, Chapter 4.2]): Let $\mathsf{H}:\mathbb{T}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, $(x,p_1,p_2)\mapsto\mathsf{H}(x,p_1,p_2)$ be a discrete Hamiltonian, assumed to satisfy the following properties.

- Monotonicity: for each $x\in\mathbb{T}$, H is nonincreasing in p_1 and nondecreasing in p_2.
- Consistency: for every $x\in\mathbb{T}$, $p\in\mathbb{R}$, $\mathsf{H}(x,p,p)=H(x,p)$.
- Differentiability: for each $x\in\mathbb{T}$, H is almost everywhere differentiable in p_1 and p_2.
- Convexity: for every $x\in\mathbb{T}$, $(p_1,p_2)\mapsto\mathsf{H}(x,p_1,p_2)$ is convex.

We can then discretize the HJB equation (1)(i) by

$$\lambda U_i-\varepsilon(\Delta_\rho U)_i+\mathsf{H}(x_i,[\nabla U]_i)=f_i+g_i. \tag{51}$$

We use a double-sided discretization of the drift q: $[Q]_i=(Q_{i,F},Q_{i,B})$. This terminology is to reflect that $Q_{i,F}$ and $Q_{i,B}$ are often characterized by the forward and backward difference operators of U_i. We discretize $-\varepsilon\Delta u+qDu$ by

$$-\varepsilon(\Delta_\rho U)_i+(Q_{i,F},Q_{i,B})\cdot[\nabla U]_i,$$

which we can transform into a vector form $\mathbf{D}(Q)U$ with a sparse matrix $\mathbf{D}(Q)$. Let $q=D_pH(x,Du)$ and from the monotonicity of H, we can discretize $D_pH(x,Du)$ by

$$(Q_{i,F},Q_{i,B})=\left(\underbrace{D_{p_1}\mathsf{H}(x_i,[\nabla U]_i)}_{\leqslant 0},\underbrace{D_{p_2}\mathsf{H}(x_i,[\nabla U]_i)}_{\geqslant 0}\right). \tag{52}$$

It is clear that under the upwind form (52), $(\lambda\mathbf{I}+\mathbf{D}(Q))$ and $\mathbf{D}^{\mathsf{T}}(Q)$ are all M-matrices. The discretization of the FPK equation with drift q is $\mathbf{D}^{\mathsf{T}}(Q)M=0$ with $\sum M_i=1$, which is the matrix form for

$$-\varepsilon(\Delta_\rho M)_i-\frac{Q_{[i+1],B}M_{[i+1]}-Q_{i,B}M_i}{\rho}-\frac{Q_{i,F}M_i-Q_{[i-1],F}M_{[i-1]}}{\rho}=0,\quad \sum M_i=1. \tag{53}$$

In order to use the policy iteration method with $[Q]_i$, we need to consider the numerical approximation of the Lagrangian.

Discrete Lagrangian: Let $\mathsf{L}:\mathbb{T}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, $(x,q_1,q_2)\mapsto\mathsf{L}(x,q_1,q_2)$ be a discrete Lagrangian

$$\mathsf{L}(x,q_1,q_2)=\sup_{(p_1,p_2)}\left\{(q_1,q_2)\cdot(p_1,p_2)-\mathsf{H}(x,p_1,p_2)\right\}. \tag{54}$$

With (54) and using the upwind form (52), we can rewrite the finite difference approximation (51) as

$$\lambda U_i-\varepsilon(\Delta_\rho U)_i+(Q_{i,F},Q_{i,B})\cdot[\nabla U]_i-\mathsf{L}(x_i,Q_{i,F},Q_{i,B})=f_i+g_i.$$

Remark 7. For example, let $H(x,Du)=\frac{1}{2}|Du|^2$. We can take $\mathsf{H}(x,p_1,p_2)=\frac{1}{2}|((p_1)_-^2+((p_2)_+)^2|$.

$$\mathsf{H}(x_i,[\nabla U]_i)=\frac{1}{2}\left(((DU)_i)_-^2+((DU)_{i-1})_+^2\right),\quad \mathsf{L}(x_i,[Q]_i)=\frac{1}{2}\left((Q_{i,F})^2+(Q_{i,B})^2\right), \tag{55}$$

$$(D_{p_1}\mathsf{H}(x_i,[\nabla U]_i),D_{p_2}\mathsf{H}(x_i,[\nabla U]_i))=(-((DU)_i)_-,((DU)_{i-1})_+).$$

The main idea of implementing policy iteration with FD schemes is presented in Algorithm 2.

Algorithm 2 Policy iteration with finite difference schemes.

Data: initial values $Q^{(0)}$, and parameters $\lambda,\rho,h,\varepsilon$.
Result: solution (U,M).

while $\|M^{(n+1)}-M^{(n)}\|>10^{-5}$ **do**
- Solve the FPK equation: $\left(\mathbf{D}^{\mathsf{T}}(Q^{(n)})\right)M^{(n)}=0$;
- Update Lagrangian and the mean field interaction terms $\mathsf{L}_i^{(n)}=\mathsf{L}(x_i,[Q^{(n)}]_i)$, $f_i^{(n)}=f(x_i,M_i^{(n)})$, $g_i^{(n)}=g[M^{(n)}](x_i)$;
- Solve the HJB equation: $\left(\lambda\mathbf{I}+\mathbf{D}(Q^{(n)})\right)U^{(n)}=\mathbf{L}^{(n)}+\mathbf{f}^{(n)}+\mathbf{g}^{(n)}$;
- Update $Q^{(n+1)}$:

$$[Q^{(n+1)}]_i=\left(D_{p_1}\mathsf{H}(x_i,[\nabla U^{(n)}]_i),D_{p_2}\mathsf{H}(x_i,[\nabla U^{(n)}]_i)\right); \tag{56}$$

- Policy smoothing $\bar{Q}^{(n+1)}=\gamma_n Q^{(n+1)}+(1-\gamma_n)\bar{Q}^{(n)}$;

end while

$\mathbf{L}^{(n)}$, $\mathbf{f}^{(n)}$ and $\mathbf{g}^{(n)}$ are vectors with elements $\mathsf{L}_i^{(n)}$, $f_i^{(n)}$ and $g_i^{(n)}$.

We compare the methodology of discretizing policies between the Lagrangian (SL) and Eulerian (FD) points of views. With SL, the discretized policy q is defined as a vector in $\mathbb{T}^d$ and updated in an implicit way. With FD, the discretized policy q needs to be split into the forward-backward parts, hence defined on $\mathbb{T}^{2d}$. Even though we are considering in (1) a stationary MFG system, with the SL method we still need to introduce a time step h. For some particular cases such that Eq. (46) holds in the SL schemes, the policy evaluation and distribution generation steps are similar for SL and FD methods. Both consist of solving linear systems with tri-diagonal matrices. The matrix structures with FD and SL schemes for solving a single HJB equation have also been discussed in [8].

Remark 8. The FD scheme (57) of the FPK equation in $1d$ also has the Markov chain interpretation, with M being an invariant measure on $\mathcal{G}_i$:

$$\underbrace{\left(Q_{i,B}-Q_{i,F}+\frac{2\varepsilon}{\rho}\right)M_i}_{\text{flux out of }x_i}=\underbrace{\left(Q_{[i+1],B}+\frac{\varepsilon}{\rho}\right)M_{[i+1]}}_{\text{flux from }x_{[i+1]}\to x_i}+\underbrace{\left(-Q_{[i-1],F}+\frac{\varepsilon}{\rho}\right)M_{[i-1]}}_{\text{flux from }x_{[i-1]}\to x_i},\quad \sum M_i=1. \tag{57}$$

It is interesting to compare (57) with (47). We have $\sqrt{\frac{2\varepsilon}{h}}=\frac{2\varepsilon}{\rho}$ if we choose $\rho=\sqrt{2\varepsilon h}$.

4 Numerical examples

We first consider some $1d$ examples with $N_x=500$, with $H(x,Du)=\frac{1}{2}|Du|^2, V(x)=\sin(2\pi x)+\cos(4\pi x)$, as shown in Figures 2–5.

- Test 1: $\varepsilon=0.3$, $f(x,m)=V(x)+m^2$, $g=0$;
- Test 2: $\varepsilon=0.01$, $f(x,m)=V(x)+m^2$, $g=0$;
- Test 3: $\varepsilon=0.01$, $f(x,m)=V(x)$, $g=0$;
- Test 4: $\varepsilon=0.01$, $f(x,m)=V(x)$, $g=10x\int xmdx$.

We solve the discounted MFG (1) with different values of λ using Algorithm 1 (SL schemes). We set the initial guess $q^{(0)}=0$ on grid. We solve the ergodic problem with FD schemes, following [9]. We observe that the solution to the discounted problem becomes closer to the solution of the ergodic problem as λ decreases. Test 1 shows our result is consistent with [9, Figure 1]. Comparing test 2 and test 3, we observe that due to the m^2 term, concentration is penalized in test 2 and more mass is moved to the local minima of u on the left. Similar observation has been made with ergodic MFGs in [30].

Test 4 is an example of potential MFG with $\mathcal{J}=\int_0^{+\infty}e^{-\lambda t}\left(\int\frac{1}{2}q^2mdx+5\left(\int xmdx\right)^2\right)dt$. Comparing with test 3, in test 4 more mass is moved to the local minima on the left. In test 4 there is a high concentration of mass, compared to test 2.

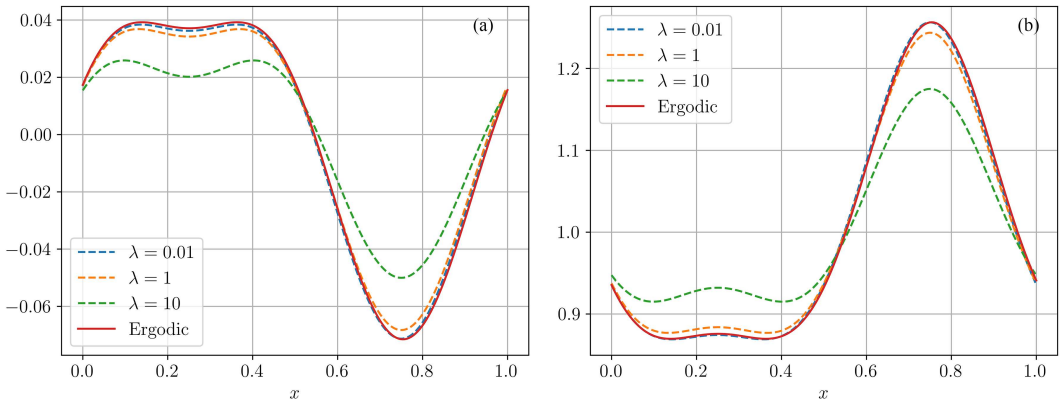

Figure 2 (Color online) Test 1: (a) the corrector $\tilde{u}$ and (b) the density m.

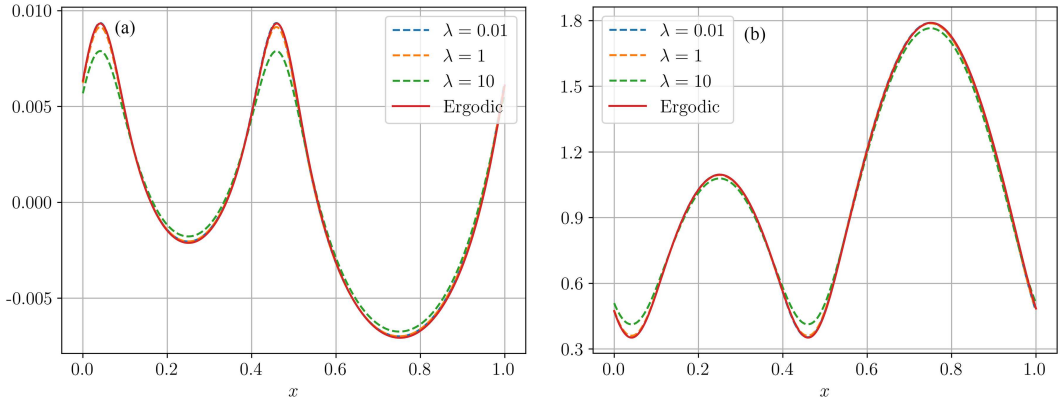

Figure 3 (Color online) Test 2: (a) the corrector $\tilde{u}$ and (b) the density m.

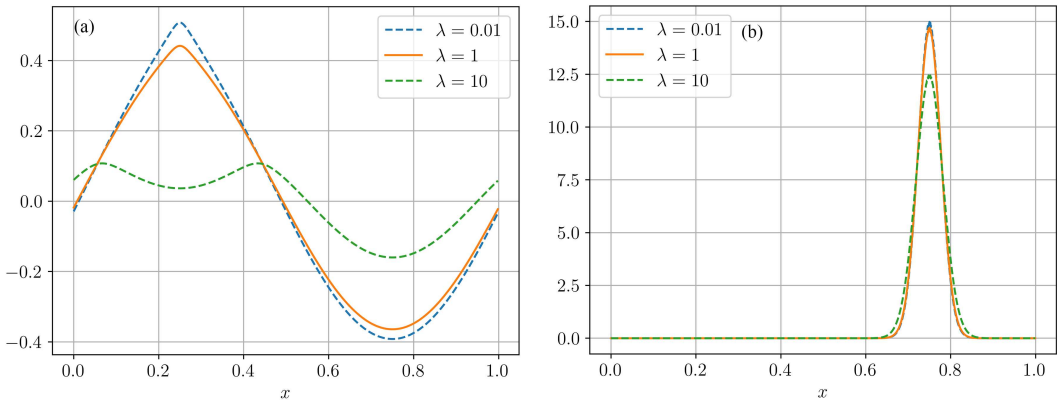

Figure 4 (Color online) Test 3: (a) the corrector $\tilde{u}$ and (b) the density m.

We consider $2d$ examples from section 9 of [31], as shown in Figures 6 and 7. We take $N_x = N_y = 50$. Let $\lambda = 0.01$, $V(x,y) = \sin(2\pi x) + \cos(4\pi x) + \sin(2\pi y)$ and $f(x,y,m) = V(x,y) + m^2$. We plot the solutions with $\varepsilon = 1$ (Test 5) and $\varepsilon = 0.01$ (test 6). We observe that our results for the $2d$ discounted MFG examples are graphically very similar to the solution of the ergodic counterparts (see [31, Figures 27 and 28]).

Finally we consider an example in dimension one with an explicit solution (to the ergodic problem).

- Test 7: $\lambda = 10^{-5}$, $\varepsilon = 0.5$, $f(x,m) = 2\pi^2(-\sin(2\pi x) + (\cos(2\pi x))^2) - 2\sin(2\pi x) + \ln(m) + 1$, $g = 0$.

The corresponding explicit solution to the ergodic system (3) is given by (see [30, Section 5.2])

$$u(x) = -\sin(2\pi x), \quad m(x) = \frac{e^{2\sin(2\pi x)}}{\int_0^1 e^{2\sin(2\pi y)}\, dy}.$$

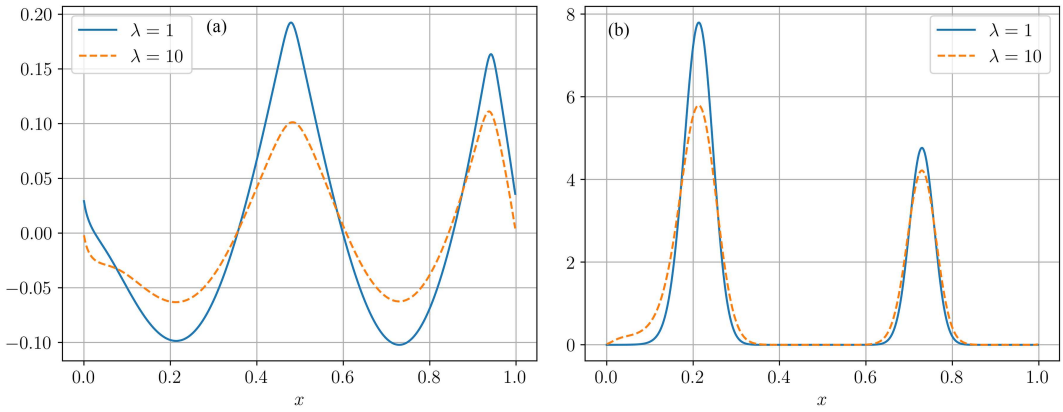

Figure 5 (Color online) Test 4: (a) the corrector $\tilde{u}$ and (b) the density m.

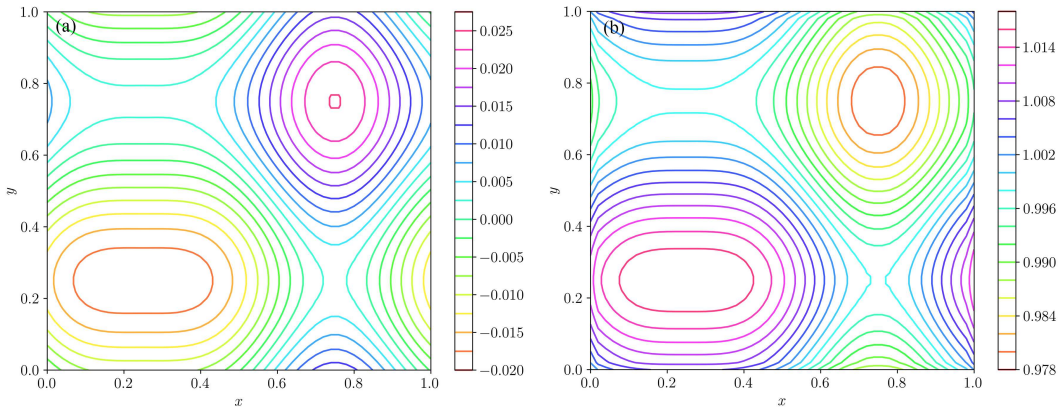

Figure 6 (Color online) Test 5 $\varepsilon = 1$: (a) the corrector $\tilde{u}$ and (b) the density m.

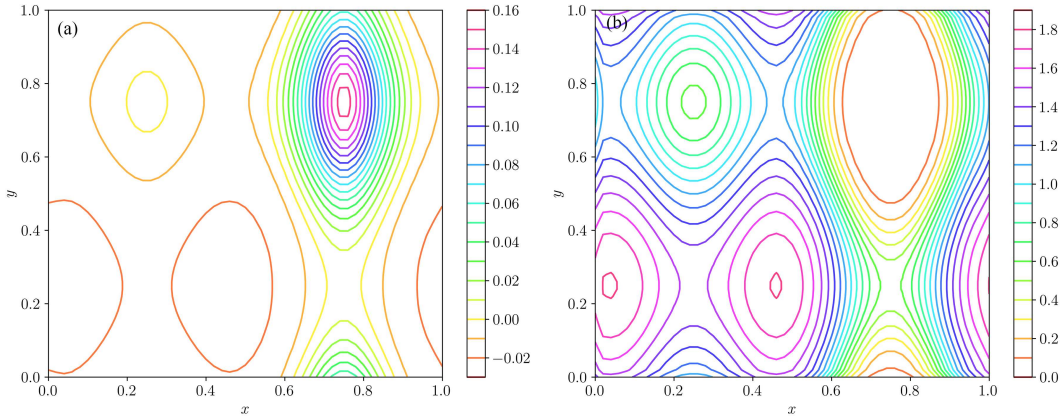

Figure 7 (Color online) Test 6 $\varepsilon = 0.01$: (a) the corrector $\tilde{u}$ and (b) the density m.

We use Table 1 to show the performance of solving with value iteration and policy iteration algorithms, using the SL scheme for discretization. By value iteration, we solve the HJB equation with a standard value iteration algorithm, while the FPK equation is solved in the same way as Algorithm 1. We also show the performance of the finite difference scheme with policy iteration (Algorithm 2).

Each iteration is defined as updating a new distribution. Value iteration is double looped: the HJB equation is fully solved at each iteration. Our policy iteration algorithms are single looped.

We use Figure 8 to show the comparison between the numerical solutions obtained from the SL method (with policy iteration) under different sizes of grids, using the explicit solution to the ergodic problem as the benchmark.

Table 1 Test 7: performance of various methods under grid refinement (N_x), number of iterations (Its), averaged CPU time per iteration (Av.CPU/It), and total CPU time.

	N_x	Its	Av.CPU/It (s)	Total CPU (s)
Value iteration SL	100	10	0.0162	0.1624
Policy iteration FD	100	20	0.0016	0.0321
Policy iteration SL	100	20	0.0007	0.0134
Value iteration SL	200	10	0.0287	0.2867
Policy iteration FD	200	20	0.0030	0.0608
Policy iteration SL	200	20	0.0015	0.0307
Value iteration SL	500	10	0.1850	1.8498
Policy iteration FD	500	20	0.0127	0.2533
Policy iteration SL	500	20	0.0096	0.1928
Value iteration SL	1000	10	0.4701	4.7012
Policy iteration FD	1000	20	0.0443	0.8862
Policy iteration SL	1000	20	0.0388	0.7767
Value iteration SL	2000	10	3.9131	39.131
Policy iteration FD	2000	20	0.2135	4.2695
Policy iteration SL	2000	20	0.3761	3.7611

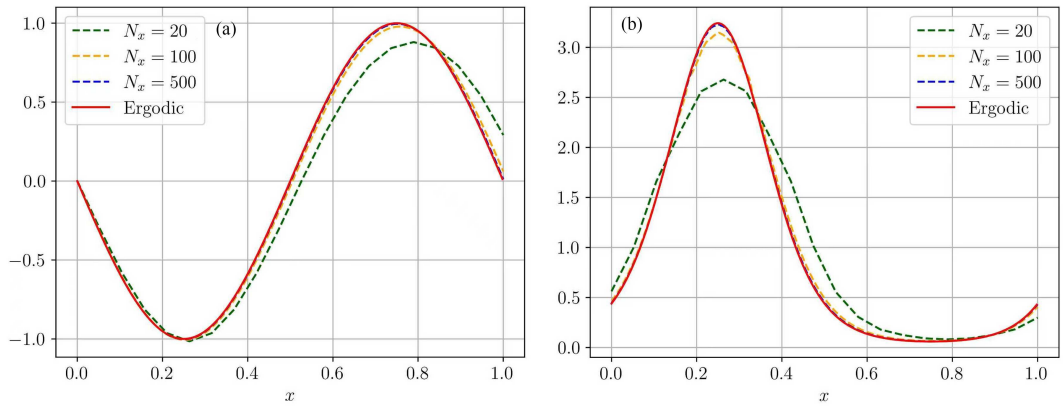

Figure 8 (Color online) Test 7: (a) the corrector $\tilde{u}$ and (b) the density m.

These numerical results in this paper were obtained using CPU-based computations on an AMD Ryzen 7 7840H processor (8 cores, 16 threads, 3.8 GHz). The code was written in Python.

Finally, we observe that policy iteration is a powerful tool for improving the efficiency of numerical resolution. Its parallelization potential, which is particularly relevant for large-scale applications, has been explored in the domain decomposition framework proposed in [8]. We plan to explore these directions in the MFG framework in our future work.

Acknowledgements This work of Qing TANG was supported by China Scholarship Council (Grant No. 202406410221). This work of Yong-Shen ZHOU was supported by China Scholarship Council (Grant No. 202406410016).

References

1 Huang M, Caines P E, Malhame R P. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ε-Nash equilibria. IEEE Trans Automat Contr, 2007, 52: 1560–1571
2 Lasry J M, Lions P L. Mean field games. Jpn J Math, 2007, 2: 229–260
3 Cardaliaguet P, Delarue F, Lasry J M, et al. The Master Equation and the Convergence Problem in Mean Field Games. Princeton: Princeton University Press, 2019
4 Carmona R, Delarue F. Probabilistic Theory of Mean Field Games With Applications I: Mean field FBSDEs, Control, and Games. Cham: Springer, 2018
5 Achdou Y, Cardaliaguet P, Delarue F, et al. Mean Field Games: Cetraro, Italy 2019. Cham: Springer, 2021
6 Cardaliaguet P, Porretta A. Long time behavior of the master equation in mean field game theory. Analysis & PDE, 2019, 12: 1397–1453
7 Bokanowski O, Maroso S, Zidani H. Some convergence results for Howard's algorithm. SIAM J Numer Anal, 2009, 47: 3001–3026
8 Festa A. Domain decomposition based parallel Howards algorithm. Math Comput Simul, 2018, 147: 121139
9 Cacace S, Camilli F, Goffi A. A policy iteration method for mean field games. ESAIM-COCV, 2021, 27: 85
10 Camilli F, Tang Q. Rates of convergence for the policy iteration method for Mean Field Games systems. J Math Anal Appl, 2022, 512: 126138
11 Camilli F, Laurière M, Tang Q. Learning equilibria in Cournot mean field games of controls. SIAM J Control Optim, 2025, 63: 1407–1431

12 Laurière M, Song J, Tang Q. Policy iteration method for time-dependent mean field games systems with non-separable hamiltonians. Appl Math Optim, 2023, 87: 17
13 Tang Q, Song J. Learning optimal policies in potential mean field games: smoothed policy iteration algorithms. SIAM J Control Optim, 2024, 62: 351–375
14 Achdou Y, Camilli F, Capuzzo Dolcetta I. Homogenization of Hamilton-Jacobi equations: numerical methods. Math Model Methods Appl Sci, 2008, 18: 1115–1143
15 Bardi M, Dolcetta I C. Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Boston: Springer, 1997
16 Falcone M, Ferretti R. Semi-Lagrangian Approximation Schemes for Linear and Hamilton Jacobi Equations. Philadelphia: Society for Industrial and Applied Mathematics, 2013
17 Chen Z, Forsyth P A. A semi-Lagrangian approach for natural gas storage valuation and optimal operation. SIAM J Sci Comput, 2008, 30: 339–368
18 Alla A, Falcone M, Kalise D. An efficient policy iteration algorithm for dynamic programming equations. SIAM J Sci Comput, 2015, 37: A181–A200
19 Carlini E, Silva F J. A fully discrete semi-Lagrangian scheme for a first order mean field game problem. SIAM J Numer Anal, 2014, 52: 45–67
20 Carlini E, Silva F J. A semi-Lagrangian scheme for a degenerate second order mean field game system. Discrete Cont Dyn Syst-A, 2015, 35: 4269–4292
21 Chowdhury I, Ersland O, Jakobsen E R. On numerical approximations of fractional and nonlocal mean field games. Found Comput Math, 2023, 23: 1381–1431
22 Ashrafyan Y, Gomes D. A fully-discrete semi-Lagrangian scheme for a price formation MFG model. Dyn Games Appl, 2025, 15: 503–533
23 Calzola E, Carlini E, Silva F J. A high-order scheme for mean field games. J Comput Appl Math, 2024, 445: 115769
24 Carlini E, Silva F J, Zorkot A. A Lagrange-Galerkin scheme for first order mean field game systems. SIAM J Numer Anal, 2024, 62: 167–198
25 Xu F, Fu Q, Shen T. PMP-based numerical solution for mean field game problem of general nonlinear system. Automatica, 2022, 146: 110655
26 Berry J, Ley O, Silva F J. Approximation and perturbations of stable solutions to a stationary mean field game system. J de Mathématiques Pures Appliquées, 2025, 194: 103666
27 Ley O, Nguyen V D. Lipschitz regularity results for nonlinear strictly elliptic equations and applications. J Differ Equ, 2017, 263: 4324–4354
28 Krylov N. Lectures on Elliptic and Parabolic Equations in Sobolev Spaces. Providence: American Mathematical Society, 2008
29 Camilli F, Falcone M. An approximation scheme for the optimal control of diffusion processes. ESAIM-M2AN, 1995, 29: 97–122
30 Carmona R, Laurière M. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games I: the ergodic case. SIAM J Numer Anal, 2021, 59: 1455–1485
31 Cacace S, Camilli F. A generalized Newton method for homogenization of Hamilton–Jacobi equations. SIAM J Sci Comput, 2016, 38: A3589–A3617