Policy iteration method for inverse mean field games

Shanyin Tong

jointly with: Nathan Soedjak Kui Ren

Department of Applied Physics and Applied Mathematics, Columbia University, NY

June 18, 2024

New England Numerical Analysis Day (NENAD), Dartmouth College, Hanover, NH

What are mean field games?

- Study non-cooperative games with a large number of rational agents
- Characterize Nash equilibrium
- Introduced 2006 by Lasry and Lions
- Applications to economics, finance, traffic flow, crowd motions, epidemics control
- Connection to reinforcement learning
- PDEs with a forward-backward structure

Mean field games — crowd motion model

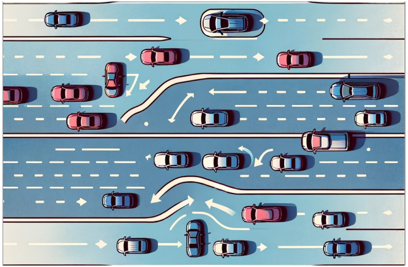
An agent (car) moves according to SDE:

$$dX_s = -q(X_s, s)\, ds + \sqrt{2\varepsilon}\, dW_s$$

Each agent is **rational**: control its velocity by q to minimize cost

$$\inf_q \mathbb{E}\left[\int_0^T \left(\frac{1}{2}|q(X_s, s)|^2 + V(X_s) + F(m(X_s, s))\right) ds + \psi(X_T)\right]$$

- $q(x, s)$: control process
- $m(x, s)$: distribution/density of all agents
- $\frac{1}{2}|q|^2$: kinetic energy of the agent
- $V(x)$: obstacle function
- $F(m)$: interaction cost (e.g., $F(m) = m^2$)
- $\psi(x)$: terminal cost



Mean field games — crowd motion model

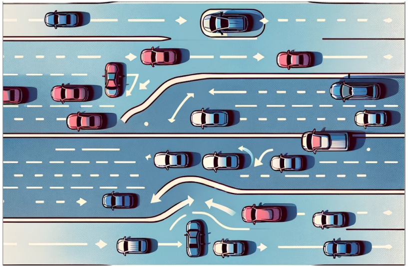
An agent (car) moves according to SDE:

$$dX_s = -q(X_s, s)\, ds + \sqrt{2\varepsilon}\, dW_s, \quad X_t = x$$

Value function $u(x,t)$: min cost of an agent starting at time t at position x

$$u(x,t) := \inf_q \mathbb{E}\left[\int_t^T \left(\frac{1}{2}|q(X_s,s)|^2 + V(X_s) + F(m(X_s,s))\right) ds + \psi(X_T)\right]$$

- $q(x,s)$: control process
- $m(x,s)$: distribution/density of all agents
- $\frac{1}{2}|q|^2$: kinetic energy of the agent
- $V(x)$: obstacle function
- $F(m)$: interaction cost (e.g., $F(m) = m^2$)
- $\psi(x)$: terminal cost



Mean field games — crowd motion model

An agent (car) moves according to SDE:

$$dX_s = -q(X_s, s)\, ds + \sqrt{2\varepsilon}\, dW_s$$

Value function $u(x, t)$: min cost of an agent starting at time t at position x

$$u(x,t) := \inf_q \mathbb{E}\left[\int_t^T \left(\frac{1}{2}|q(X_s,s)|^2 + V(X_s) + F(m(X_s,s))\right) ds + \psi(X_T)\right]$$

- Optimality condition $\Rightarrow$ u satisfies **Hamilton-Jacobi-Bellman (HJB)** equation

$$-\partial_t u - \varepsilon \Delta u + \frac{1}{2}|Du|^2 = V(x) + F(m(x,t)), \quad u(x,T) = \psi(x)$$

- Optimal control $q = Du$: optimizer of Hamiltonian $H(Du) = \sup_q Du \cdot q - \frac{1}{2}|q|^2$
- m evolves according to **Fokker-Planck** equation: $\partial_t m - \varepsilon \Delta m - \mathrm{div}(mDu) = 0$

Inverse mean field games

- Forward model (MFG): $V \longmapsto (m, u)$

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = V + F(m) & \text{in } \mathbb{T}^d \times (0, T) \\ \partial_t m - \Delta m - \mathrm{div}(mDu) = 0 & \text{in } \mathbb{T}^d \times (0, T) \\ u(x, T) = \psi(x), \ \ m(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

Inverse mean field games

- Forward model (MFG): $V \longmapsto (m, u)$

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = V + F(m) & \text{in } \mathbb{T}^d \times (0, T) \\ \partial_t m - \Delta m - \operatorname{div}(mDu) = 0 & \text{in } \mathbb{T}^d \times (0, T) \\ u(x, T) = \psi(x), \; m(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

- **Inverse problem:**
  data (observations of MFG solution) $\Longrightarrow$ reconstruct obstacle fun V (environment info)

$$u(x, 0) \text{ or } u_t(x, T) \quad \longmapsto \quad V$$

Inverse mean field games

- Forward model (MFG): $V \longmapsto (m, u)$

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = V + F(m) & \text{in } \mathbb{T}^d \times (0,T) \\ \partial_t m - \Delta m - \operatorname{div}(mDu) = 0 & \text{in } \mathbb{T}^d \times (0,T) \\ u(x,T) = \psi(x),\ \ m(x,0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

- **Inverse problem:**
  data (observations of MFG solution) $\Longrightarrow$ reconstruct obstacle fun V (environment info)

$$u(x,0) \text{ or } u_t(x,T) \quad \longmapsto \quad V$$

- **Challenges**:
  - MFG: *forward-backward coupled nonlinear* system $\Rightarrow$ nontrivial & inefficient to solve

Inverse mean field games

- Forward model (MFG): $V \longmapsto (m, u)$

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = V + F(m) & \text{in } \mathbb{T}^d \times (0, T) \\ \partial_t m - \Delta m - \operatorname{div}(mDu) = 0 & \text{in } \mathbb{T}^d \times (0, T) \\ u(x, T) = \psi(x), \;\; m(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

- **Inverse problem:**
  data (observations of MFG solution) $\Longrightarrow$ reconstruct obstacle fun V (environment info)

$$u(x, 0) \text{ or } u_t(x, T) \quad \longmapsto \quad V$$

- **Challenges**:
  - MFG: *forward-backward coupled nonlinear* system $\Rightarrow$ nontrivial & inefficient to solve
    Iterative methods, optimal control, machine learning, reinforcement learning

Inverse mean field games

- Forward model (MFG): $V \longmapsto (m, u)$

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = V + F(m) & \text{in } \mathbb{T}^d \times (0, T) \\ \partial_t m - \Delta m - \mathrm{div}(mDu) = 0 & \text{in } \mathbb{T}^d \times (0, T) \\ u(x, T) = \psi(x),\ m(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

- **Inverse problem:**
  data (observations of MFG solution) $\Longrightarrow$ reconstruct obstacle fun V (environment info)

$$u(x, 0) \text{ or } u_t(x, T) \quad \longmapsto \quad V$$

- **Challenges**:
  - MFG: *forward-backward coupled nonlinear* system $\Rightarrow$ nontrivial & inefficient to solve
    Iterative methods, optimal control, machine learning, reinforcement learning
  - PDE-constrained optimization $\Rightarrow$ adjoint equation with same structure

Inverse mean field games

- Forward model (MFG): $V \longmapsto (m, u)$

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = V + F(m) & \text{in } \mathbb{T}^d \times (0,T) \\ \partial_t m - \Delta m - \operatorname{div}(mDu) = 0 & \text{in } \mathbb{T}^d \times (0,T) \\ u(x,T) = \psi(x), \; m(x,0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

- **Inverse problem:**
  data (observations of MFG solution) $\Longrightarrow$ reconstruct obstacle fun V (environment info)

$$u(x,0) \text{ or } u_t(x,T) \quad \longmapsto \quad V$$

- **Challenges**:
  - MFG: *forward-backward coupled nonlinear* system $\Rightarrow$ nontrivial & inefficient to solve
    Iterative methods, optimal control, machine learning, reinforcement learning
  - PDE-constrained optimization $\Rightarrow$ adjoint equation with same structure
    Direct least-square, primal-dual method, bilevel optimization, Gaussian process

Policy iteration method for MFG

- Classical algorithm for optimal control
- First introduced to MFG by [Cacace et al. 2021]
- A quasi-Newton algorithm for root-finding:

$$\begin{cases} -\partial_t u - \Delta u + \frac{1}{2}|Du|^2 = V + F(m) & \text{in } \mathbb{T}^d \times (0, T) \\ \partial_t m - \Delta m - \operatorname{div}(mDu) = 0 & \text{in } \mathbb{T}^d \times (0, T) \\ u(x, T) = \psi(x), \;\; m(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

- Use policy/control

$$q = \arg\max_q \left[q \cdot Du - \frac{1}{2}|q|^2 \right] = Du$$

- Fixed point iteration

Policy iteration method for MFG

Choose $q^{(0)}$, iterate on $k = 0, 1, 2, \dots$

1. Solve *linear* PDE for $m^{(k)}$

$$\begin{cases} \partial_t m^{(k)} - \varepsilon \Delta m^{(k)} - \mathrm{div}(m^{(k)} q^{(k)}) = 0 & \text{in } \mathbb{T}^d \times (0, T) \\ m^{(k)}(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

2. Solve *linear* PDE for $u^{(k)}$

$$\begin{cases} -\partial_t u^{(k)} - \varepsilon \Delta u^{(k)} + q^{(k)} \cdot Du^{(k)} = \frac{1}{2}|q^{(k)}|^2 + V + F(m^{(k)}) & \text{in } \mathbb{T}^d \times (0, T) \\ u^{(k)}(x, T) = \psi(x) & \text{in } \mathbb{T}^d \end{cases}$$

3. Update policy

$$q^{(k+1)} = \arg\max_q \left[q \cdot Du^{(k)} - \frac{1}{2}|q|^2 \right] = Du^{(k)}$$

Policy iteration method for inverse MFG

Choose $q^{(0)}$, iterate on $k = 0, 1, 2, \ldots$

1. Solve *linear* PDE for $m^{(k)}$

$$\begin{cases} \partial_t m^{(k)} - \varepsilon \Delta m^{(k)} - \mathrm{div}(m^{(k)} q^{(k)}) = 0 & \text{in } \mathbb{T}^d \times (0, T) \\ m^{(k)}(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

2. Solve *linear* inverse problem: reconstruct $V^{(k)}$ and $u^{(k)}$ from data $u(x, 0)$ (or $u_t(x, T)$)

$$\begin{cases} -\partial_t u^{(k)} - \varepsilon \Delta u^{(k)} + q^{(k)} \cdot Du^{(k)} = \frac{1}{2}|q^{(k)}|^2 + V^{(k)} + F(m^{(k)}) & \text{in } \mathbb{T}^d \times (0, T) \\ u^{(k)}(x, T) = \psi(x) & \text{in } \mathbb{T}^d \end{cases}$$

3. Update policy

$$q^{(k+1)} = \arg\max_q \left[q \cdot Du^{(k)} - \frac{1}{2}|q|^2 \right] = Du^{(k)}$$

Policy iteration method for inverse MFG

Choose $q^{(0)}$, iterate on $k = 0, 1, 2, \ldots$

1. Solve *linear* PDE for $m^{(k)}$

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

2. Solve *linear* least-square problem: reconstruct $V^{(k)}$ and $u^{(k)}$ from data $h(x)$

$$\begin{aligned} \underset{V,u}{\text{minimize}} \quad & \|u(x,0) - h(x)\|^2_{L^2(\mathbb{T}^d)} \quad \left(\text{or } \|u_t(x,T) - h(x)\|^2_{L^2(\mathbb{T}^d)}\right) \\ \text{subject to} \quad & -\partial_t u - \varepsilon \Delta u + q^{(k)} \cdot Du = \tfrac{1}{2}|q^{(k)}|^2 + V + F(m^{(k)}) \quad \text{in } \mathbb{T}^d \times (0, T) \\ & u(x, T) = \psi(x) \quad \text{in } \mathbb{T}^d \end{aligned}$$

3. Update policy

$$q^{(k+1)} = \arg\max_q \left[q \cdot Du^{(k)} - \frac{1}{2}|q|^2 \right] = Du^{(k)}$$

Policy iteration method for inverse MFG

Choose $q^{(0)}$, iterate on $k = 0, 1, 2, \dots$

1. Solve *linear* PDE for $m^{(k)}$

$$\begin{cases} \partial_t m^{(k)} - \varepsilon \Delta m^{(k)} - \mathrm{div}(m^{(k)} q^{(k)}) = 0 & \text{in } \mathbb{T}^d \times (0,T) \\ m^{(k)}(x,0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

2. Solve *linear* least-square problem: reconstruct $V^{(k)}$ with data $h(x) = u_t(x,T)$

$$V^{(k)} = \arg\min_V \| V - [-\varepsilon \Delta \psi + q^{(k)}(\cdot, T) \cdot \psi - \frac{1}{2}|q^{(k)}(\cdot,T)|^2 - F(m^{(k)}) - h] \|^2_{L^2(\mathbb{T}^d)}$$

Solve *linear* PDE $\begin{cases} -\partial_t u^{(k)} - \varepsilon \Delta u^{(k)} + q^{(k)} \cdot Du^{(k)} = \frac{1}{2}|q^{(k)}|^2 + V^{(k)} + F(m^{(k)}) & \text{in } \mathbb{T}^d \times (0,T) \\ u^{(k)}(x,T) = \psi(x) & \text{in } \mathbb{T}^d \end{cases}$

3. Update policy

$$q^{(k+1)} = \arg\max_q \left[q \cdot Du^{(k)} - \frac{1}{2}|q|^2 \right] = Du^{(k)}$$

Convergence theorem for policy iteration on inverse MFG

Convergence theorem

Under certain regularity assumptions on initial and final conditions ψ, m_0, data h, initial policy q_0 and interaction cost function F, for sufficiently small $T > 0$, the sequence $\{V^{(k)}\}_{k\geq 0}$ generated by the above policy iteration algorithm satisfies

$$V^{(k)} \to V^* \text{ uniformly in } \mathbb{T}^d,$$

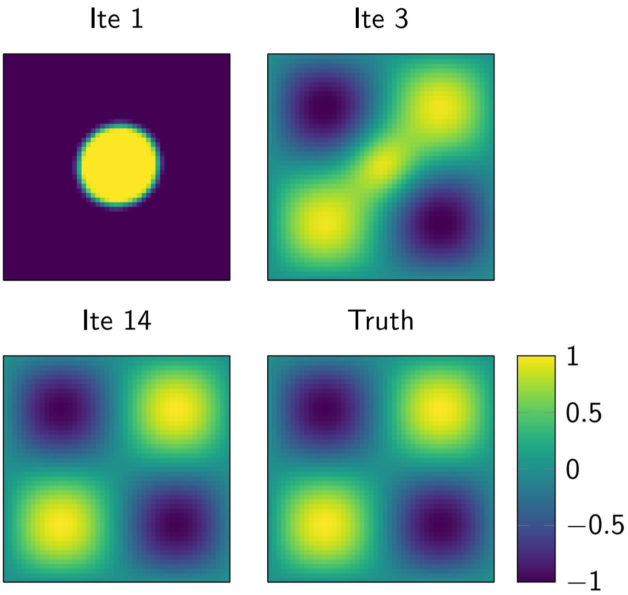
where V^* is a solution to the inverse mean field game problem with data $h(x) = u_t(x, T)$.

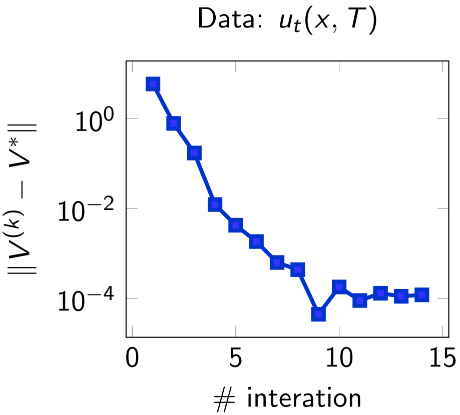
Linear convergence

With same assumptions, for sufficiently small T, there exist constants $C > 0$ and $0 < \lambda < 1$ such that

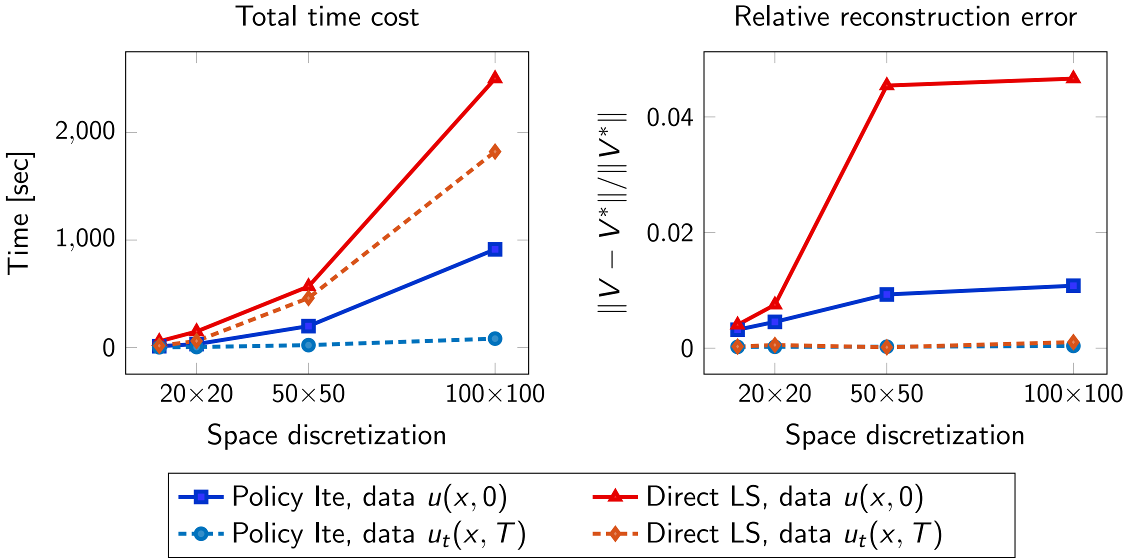
$$\|V^{(k)} - V^*\|_{L^\infty(\mathbb{T}^d)} \leq C\lambda^k, \qquad \forall k \geq 0.$$

Convergence rate for policy iteration





Comparison with direct least-square method



Main takeaways

- Introduce an interesting inverse problem:
  - Inverse mean field game
  - Reconstruct environment info from observations of population & strategy
- Efficient method for solving inverse mean field game:
  - Reformulate a forward-backward coupled nonlinear PDE-constrained opt
  - Only require iterations of linear PDE solves & linear least-square

Current and future plans:

- More complicated setting: non-separable Hamiltonians
- Theoretical study of uniqueness & stability
- Application to real traffic flow models

References

- K. Ren, N Soedjak, and S. Tong, *A policy iteration method for inverse mean field games*, In preparation
- S. Cacace, F. Camilli, and A. Goffi, *A policy iteration method for mean field games*, ESAIM: Control, Optimisation and Calculus of Variations 27 (2021)
- F. Camilli, and Q. Tang, *Rates of convergence for the policy iteration method for mean field games systems*, Journal of Mathematical Analysis and Applications 512.1 (2022)
- M. Laurière, J. Song, and Q. Tang, *Policy iteration method for time-dependent Mean Field Games systems with non-separable Hamiltonians*, Applied Mathematics & Optimization 87.2 (2023)
- Q. Tang, and J. Song, *Learning Optimal Policies in Potential Mean Field Games: Smoothed Policy Iteration Algorithms*, SIAM Journal on Control and Optimization 62.1 (2024)