A NEURAL SCORE-BASED PROBABILITY FLOWS APPROACH FOR SECOND-ORDER MEAN-FIELD GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Second-order mean-field games (MFGs) model strategic interactions among large populations with stochastic dynamics, but their coupled forward-backward PDE systems are difficult to solve in high dimensions. We develop a unified probability-flow framework for general and potential second-order MFGs using score-based neural ODEs. Building on the flow-velocity formulation based on the Nelson's approach, we rewrite the Fokker-Planck equation as a continuity equation whose velocity depends on the density score. Under suitable regularity conditions, the resulting deterministic probability flow and the controlled diffusion have identical time marginals, yielding a grid-free Lagrangian formulation that avoids spatial discretization. We study an analytical interbank benchmark, synthetic transport including a 1000-dimensional case, nonlinear 12D quadcopter control, and unpaired CelebA image translation task at latent space. The experiments recover analytical solutions in benchmark settings and demonstrate the framework's feasibility for high-dimensional transport and nonlinear control, without claiming that finite training necessarily converges to an MFG equilibrium.

1 INTRODUCTION

Mean-field games (MFGs) characterize Nash equilibria in stochastic differential games with a continuum of infinitesimal agents. Each representative agent responds to an anticipated population distribution while having negligible individual influence on it (Lasry and Lions, 2007; Huang et al., 2006). Given an anticipated population flow, a representative agent solves an optimal-control problem. At Nash equilibrium, the law induced by the optimal response coincides with the anticipated flow. With nondegenerate individual noise, this fixed point is classically characterized by a backward Hamilton-Jacobi-Bellman equation coupled to a forward Fokker-Planck equation, giving a second-order MFG. Such models arise in systemic risk, collective control, and generative modeling (Carmona et al., 2013; Liu et al., 2018; Zhang and Katsoulakis, 2023).

Classical finite-difference and semi-Lagrangian schemes provide accurate low-dimensional solvers, but the cost of a spatial grid grows rapidly with the state dimension (Achdou and Capuzzo-Dolcetta, 2010; Achdou et al., 2013; Carlini and Álvarez, 2014; Lauriere, 2021). Mesh-free neural approaches instead parameterize value functions, controls, densities, or trajectories. Representative examples include variational and primal-dual solvers for high-dimensional potential MFGs (Ruthotto et al., 2020; Lin et al., 2021), trajectory-regularized normalizing flows (Huang et al., 2023), normalizing-flow McKean-Vlasov solvers (Liu et al., 2025), and actor-critic population flows (Zhou et al., 2025b). Particle flow matching has also recently been used for high-dimensional general MFGs (Yu et al., 2026).

A separate probabilistic line characterizes broad classes of stochastic MFGs, under stochastic-maximum-principle assumptions, through McKean-Vlasov forward-backward stochastic differential equations (FBSDEs) (Carmona and Delarue, 2013). This representation avoids a spatial grid but not the intrinsic forward-backward coupling. Neural solvers must therefore learn the backward component and update the population law consistently. Recent approaches combine fictitious play with repeated deep-FBSDE solves (Han et al., 2024), or combine a McKean-Vlasov FBSDE fixed point with process-regularized normalizing flows (Liu et al., 2025).

Schrödinger bridge methods provide another close connection. They steer a reference diffusion between prescribed endpoint laws and admit stochastic-control formulations based on coupled

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

forward-backward potentials and scores (Léonard, 2014; Liu et al., 2022). Both Schrödinger bridges and our method express diffusion through evolving densities and their scores. The problems are nevertheless different. A classical bridge fixes both endpoint marginals and solves an entropy-regularized stochastic transport problem, whereas a general MFG best response has a free response law for a frozen population belief and must additionally satisfy a Nash fixed point.

The closest diffusion-aware flow formulations primarily address mean-field control (MFC) or density-path optimization. Primal-dual saddle point formulation of the second-order MFG is solved by adverserial training (Lin et al., 2021). Score-based Neural ODEs propagate first- and second-order scores for second-order MFC (Zhou et al., 2025a). Variational conditional normalizing flows learn MFC flow maps and scores (Zhao et al., 2025). Parametric density-path optimization represents fixed-endpoint actions by spline paths in parameter space (Hernandez et al., 2025). For non-potential second-order MFGs, monotone-inclusion methods provide a finite-difference alternative (Nurbekyan et al., 2024). These developments establish important high-dimensional and second-order problem solvers, but leave a distinct question: whether a probability-flow reformulation preserves each frozen best-response problem, and hence the Nash fixed points, of a general second-order MFG.

We answer this question using the flow velocity $v = b - \nu\nabla \log \rho$, where b is the stochastic drift, ρ is the response law, and $\nu > 0$ is the diffusivity. This variable, rooted in Nelson's stochastic-mechanics formulation and closely related to probability-flow ODEs in score-based modeling (Nelson, 1966; Song et al., 2021), converts the Fokker-Planck constraint into a continuity equation with the same time marginals. The continuity dynamics can be represented by deterministic characteristics and a Neural ODE (Chen et al., 2018; Grathwohl et al., 2018). Diffusion remains in the reconstructed drift and, for quadratic effort, in Fisher-information and endpoint-entropy terms. The required score is that of the current response law and can be estimated from particles. Unlike an FBSDE solver or primal-dual methods, this formulation does not parameterize a backward process or min-max adversarial training. Unlike classical bridge and MFC planning formulations, it preserves the frozen best-response map needed by a general MFG.

For a general MFG, our computation retains the game's best-response structure. An inner solver alternates a flow-velocity/control network with a ScoreNet fitted to the current response particles. An outer relaxed population update is inspired by fixed-point iteration and fictitious play (Cardaliaguet and Hadikhanloo, 2017; Yu et al., 2024). The potential subclass additionally admits a single population-level variational problem, so its probability-flow action can be optimized without the outer fixed-point loop.

Our contributions are summarized as follows. We propose a neural score-based probability-flow approach to second-order MFG. The Fokker-Planck equation constraint can be transformed into the continuity equation while retaining diffusion through the score, Fisher information, and entropy. This reformulation simplify the classical solvers to second-order MFG problem that using forward-backward iteration and adverserial training, which is not stable. We further explore the control-affine dynamics on quadcopter swarm and banking problems, clarifies which learned velocities are physically realizable. This framework is suitable for both potential and non-potential MFG. We demonstrate the efficiency and effectness of our approach on high-dimensional problems such as 1000d transport and image translation generative modeling.

2 Second-order MFG formulation

Let $T > 0$, $d \geq 1$, and $\nu \geq 0$. Denote $\mathcal{P}_2(\mathbb{R}^d)$ as the probability measures with finite second moment. We use $\mathcal{C} := C([0,T]; \mathcal{P}_2(\mathbb{R}^d))$ for population flows, with continuity understood in W_2. Here ρ is absolutely continuous, *i.e.* $\rho \in AC^2([0,T]; \mathcal{P}_2)$.

Wasserstein distance and transport geometry For $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$, the squared Wasserstein-2 distance in the dynamic Benamou-Brenier representation is Benamou and Brenier (2000); Villani et al. (2009)

$$\begin{aligned} W_2^2(\mu_0,\mu_1) = \inf_{\mu,v} \mu_0^1 \int_{\mathbb{R}^d} \|v(x,t)\|^2 \, d\rho(x)\, dt, \\ \text{s.t.} \quad \partial_t \rho + \nabla \cdot (\rho v) = 0, \qquad \rho(0) = \mu_0, \quad \rho(1) = \mu_1. \end{aligned} \tag{1}$$

When v generates a unique characteristic flow X, the solution of the continuity equation admits the Lagrangian representation

$$\frac{\mathrm{d}}{\mathrm{d}t}X(z,t) = v(X(z,t),t), \qquad X(z,0) = z, \qquad \rho(t) = \big(X(\cdot,t)\big)_{\#}\mu_0.$$

This transport geometry motivates the deterministic particle representation used later.

Non-potential MFG The best-response formulation applies to both potential and non-potential games. A MFG is non-potential when its population-dependent cost fields are not the first variations of common population functionals.

Fix an anticipated flow $\widetilde{\rho} \in C([0,T];\mathcal{P}_2(\mathbb{R}^d))$. A representative agent chooses a drift b and solves

$$\begin{aligned} \inf_b \; J(b;\widetilde{\rho}) &= \mathbb{E}\left[\int_0^T \big(K(X,b(X,t)) + F(\widetilde{\rho})(X)\big)\,dt + G(\widetilde{\rho})(X_T)\right], \\ dX &= b(X,t)\,dt + \sqrt{2\nu}\,dW, \qquad X_0 \sim \mu_0. \end{aligned} \tag{2}$$

Assumption 1 (Admissibility and well-posed best response)**.** *The initial law satisfies $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$. The kinetic cost $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is proper, bounded below, lower semicontinuous, convex and coercive in its second argument. The costs $F, G : \mathcal{P}_2(\mathbb{R}^d) \to C(\mathbb{R}^d)$ are continuous. Admissible feedbacks $b(X(t),t)$ is square-integrable with respect to $\mathbb{P} \otimes \mathrm{d}t$.*

The conditions in Assumption 1 ensures the existence and integrability properties used below. The monotonicity condition of F, G and the convexity condition of Hamiltonian $H(x,p)$ with respect to the second variable Lasry and Lions (2007) provide the uniqueness of solution.

An MFG equilibrium is the fixed point

$$\rho^* \in \mathrm{BR}(\rho^*), \tag{3}$$

where the best response

$$\mathrm{BR}(\widetilde{\rho}) := \left\{ \rho^b : b \in \arg\min_{b \in \mathcal{A}} J(b;\widetilde{\rho}) \right\},$$

$\mathcal{A}$ is the admissible feedback class, and $\rho^b(t) = \mathrm{Law}(X^b(t))$ is the law induced by b, with $\rho^b(0) = \mu_0$.

The anticipated flow is fixed during the individual optimization and is identified with the response law only after that optimization Lasry and Lions (2007); Huang et al. (2006).

The law of a controlled diffusion satisfies the Fokker-Planck equation, so the same fixed-flow problem can be written

$$\begin{aligned} \inf_{\rho,b} \quad & \int_0^T \int \big(K(x,b) + F(\widetilde{\rho})(x)\big)\rho(x)\,dx\,dt + \int G(\widetilde{\rho}_T)(x)\rho_T(x)\,dx, \\ \text{s.t.} \quad & \partial_t\rho + \nabla\cdot(\rho b) - \nu\Delta\rho = 0, \qquad \rho(0) = \mu_0. \end{aligned} \tag{4}$$

For $K(x,\cdot)$ continuously differentiable, strictly convex, and superlinear, use the Hamiltonian convention

$$H(x,p) := \sup_{b \in \mathbb{R}^d} \{-b\cdot p - K(x,b)\}. \tag{5}$$

When H is differentiable, the optimizer is $b^* = -\nabla_p H(x,p)$. In particular, $K(x,b) = \frac{1}{2}\|b\|^2$ gives $H(x,p) = \frac{1}{2}\|p\|^2$ and $b^* = -p$. Assuming a classical value function u, the equilibrium system is

$$\begin{cases} -\partial_t u - \nu\Delta u + H(x,\nabla u) = F(\rho)(x), \\ \partial_t\rho - \nabla\cdot\big(\rho\nabla_p H(x,\nabla u)\big) - \nu\Delta\rho = 0, \\ \rho(0) = \mu_0, \qquad u(T,x) = G(\rho_T)(x). \end{cases} \tag{6}$$

This is the classical HJB-Fokker-Planck characterization of the Nash fixed point Lasry and Lions (2007); Huang et al. (2006).

For a non-potential model, this fixed-point system is the equilibrium formulation: the fixed-flow Eulerian problem equation 4 remains an individual best response and is not a simultaneous optimization over the population.

Potential MFG A potential MFG is the subclass for which population functionals $\mathcal{F}$ and $\mathcal{G}$ generate the representative-agent cost fields through $\frac{\delta \mathcal{F}}{\delta \rho}(\rho)(x) = F(\rho)(x)$, $\frac{\delta \mathcal{G}}{\delta \rho}(\rho)(x) = G(\rho)(x)$, the equilibrium admit the objective functional

$$\int_0^T \left[\int_\Omega K(x,b)\rho dx + \mathcal{F}(\rho) \right] dt + \mathcal{G}(\rho_T) \tag{7}$$

under the same Fokker-Planck constraint. Its first-order conditions recover the MFG system under the usual regularity hypotheses Monderer and Shapley (1996); Briani and Cardaliaguet (2018).

3 Probability-flow reformulation

In this section, we propose our score-based probability-flow reformulation on the second-order MFG. The score fuction is denoted as $s_\rho = \nabla \log \rho$. Define negative entropy and Fisher information by:

$$\mathcal{H}(\rho) = \int_\Omega \rho \log \rho \, dx, \qquad \mathcal{I}(\rho) = \int_\Omega \|\nabla \log \rho\|^2 \rho \, dx.$$

Assumption 2 (Score and entropy regularity)**.** *For almost every $t \in (0,T)$, we have $\rho > 0$, $\int_0^T \mathcal{I}(\rho)dt < \infty$, and the endpoints are absolutely continuous with finite $\mathcal{H}(\rho(0))$ and $\mathcal{H}(\rho(T))$. The drift belongs to L^2_ρ, the space of vector fields square-integrable with respect to $\rho(x,t)\,dxdt$. Together with the bounded $\int_0^T \mathcal{I}(\rho)dt$, this gives finite energy for the flow velocity defined below. Conversely, any reconstructed drift is required to be admissible in the sense of Assumption 1.*

Define the inverse maps on their respective admissible classes by:

$$\mathsf{v}_\nu(\rho,b) := \begin{cases} b - \nu s_\rho, & \nu > 0, \\ b, & \nu = 0, \end{cases} \qquad \mathsf{b}_\nu(\rho,v) := \begin{cases} v + \nu s_\rho, & \nu > 0, \\ v, & \nu = 0. \end{cases} \tag{8}$$

Proposition 1 (Flow velocity)**.** *Let (ρ, b) satisfy the Fokker-Planck constraint and Assumption 2 when $\nu > 0$. Define $v = \mathsf{v}_\nu(\rho, b)$ by equation 8. Then $v \in L^2_\rho$, $\partial_t \rho + \nabla \cdot (\rho v) = 0$, and consequently $\rho \in AC^2([0,T]; \mathcal{P}_2)$. Conversely, a flow pair with regularity gives a Fokker-Planck pair through $b = \mathsf{b}_\nu(\rho, v)$, provided the reconstructed drift is admissible.*

Proof. When $\nu = 0$, set $v = b$, the forward equations coincide and the claim is immediate without a density or score assumption. Suppose that $\nu > 0$. Since $\rho \nabla \log \rho = \nabla \rho$,

$$\partial_t \rho + \nabla \cdot (\rho b) - \nu \Delta \rho = \partial_t \rho + \nabla \cdot (\rho v).$$

The assumed drift energy and the bounded $\int_0^T \mathcal{I}(\rho)dt$ give $v \in L^2_\rho$. The standard finite-energy characterization of the continuity equation then gives the AC^2 conclusion Villani et al. (2009); Nelson (1966). The same identity proves the converse. □

For the response class with regularity in Assumption 2, denote the objective in equation 4 by $J^{\mathrm{SDE}}(\rho, b; \widetilde{\rho})$ and define its exact probability-flow reparameterization by

$$J^{\mathrm{PF}}(\rho, v; \widetilde{\rho}) := J^{\mathrm{SDE}}\big(\rho, \mathsf{b}_\nu(\rho, v); \widetilde{\rho}\big). \tag{9}$$

Let $\mathfrak{A}^{\mathrm{SDE}}(\widetilde{\rho})$ contain all admissible Fokker-Planck pairs in equation 4, let $\mathfrak{A}^{\mathrm{SDE}}_{\mathrm{reg}}(\widetilde{\rho})$ be its regularity subclass, and let $\mathfrak{A}^{\mathrm{PF}}_{\mathrm{reg}}(\widetilde{\rho})$ contain the flow pairs for which the reconstructed drift is admissible. Write $\mathrm{BR}^{\mathrm{SDE}}$, $\mathrm{BR}^{\mathrm{SDE}}_{\mathrm{reg}}$, and $\mathrm{BR}^{\mathrm{PF}}_{\mathrm{reg}}$ for the density projections of their respective minimizers.

Proposition 2 (Quadratic probability-flow action)**.** *Restrict equation 4 to pairs satisfying Assumption 2, and let $K(x,b) = \frac{1}{2}\|b\|^2$. For $\nu > 0$, it is equivalent on the class to*

$$\begin{aligned} \inf_{\rho, v} \quad & \int_0^T \left[\frac{1}{2} \int \|v\|^2 \rho \, dx + \frac{\nu^2}{2} \mathcal{I}(\rho) + \int F(\widetilde{\rho})(x) \rho(x,t) \, dx \right] dt \\ & + \int G(\widetilde{\rho}_T)(x) \rho(x,T) \, dx + \nu \big(\mathcal{H}(\rho_T) - \mathcal{H}(\mu_0) \big), \\ \text{s.t.} \quad & \partial_t \rho + \nabla \cdot (\rho v) = 0, \qquad \rho(0) = \mu_0. \end{aligned} \tag{10}$$

For $\nu > 0$, minimizers correspond through $b = v + \nu \nabla \log \rho$. For $\nu = 0$, the score, Fisher, and entropy terms are omitted and the result reduces to the ordinary first-order action with $v = b$.

Theorem 1 (Best-response equivalence). *For every frozen $\widetilde{\rho}$, the map in equation 8 is a cost-preserving bijection between the regular SDE and probability-flow feasible sets. Consequently,*

$$\mathrm{BR}^{\mathrm{SDE}}_{\mathrm{reg}}(\widetilde{\rho}) = \mathrm{BR}^{\mathrm{PF}}_{\mathrm{reg}}(\widetilde{\rho}), \tag{11}$$

with identical optimal values and paired suboptimality gaps. If every full-problem minimizer is regular, this set also equals $\mathrm{BR}^{\mathrm{SDE}}(\widetilde{\rho})$.

Corollary 1 (Nash-equilibrium equivalence). *The two formulations have the same MFG equilibrium densities:*

$$\left\{\rho : \rho \in \mathrm{BR}^{\mathrm{SDE}}_{\mathrm{reg}}(\rho)\right\} = \left\{\rho : \rho \in \mathrm{BR}^{\mathrm{PF}}_{\mathrm{reg}}(\rho)\right\}. \tag{12}$$

Under the additional minimizer condition with regularity in Theorem 1, this fixed-point set also equals the fixed-point set of the full SDE response $\mathrm{BR}^{\mathrm{SDE}}$.

After the reformulation, the MFG equilibrium define in Eq. 3 can also be expressed as $\rho^b(t) = \big(T^b(\cdot, t)\big)_{\#}\mu_0$. Here, T^b is the probability-flow map associated with the controlled SDE.

The proofs are given in Appendix A.

Proposition 3 (Control-affine probability flow). *Fix a population path $\widetilde{\rho}$ and consider*

$$dX = \big[f(X, t; \widetilde{\rho}) + B(X, t)\alpha_t\big]dt + \sqrt{2\nu}\,\sigma(X, t)\,dW. \tag{13}$$

Let $A = \sigma\sigma^\top$ and $(\mathrm{div}\, A)_i = \sum_j \partial_{x_j} A_{ij}$. For every positive and sufficiently regular response density ρ, the corresponding Fokker-Planck equation is equivalent to $\partial_t \rho + \nabla \cdot (\rho v) = 0$, where

$$v = f + B\alpha - \nu\big(A\nabla \log \rho + \mathrm{div}\, A\big). \tag{14}$$

Hence a prescribed velocity v is realizable if and only if

$$v - f + \nu\big(A\nabla \log \rho + \mathrm{div}\, A\big) \in \mathrm{Range}(B) \tag{15}$$

pointwise on the support of ρ.

Proof. Using $\sum_j \partial_{x_j}(A_{ij}\rho) = \rho(\mathrm{div}\, A)_i + \rho(A\nabla \log \rho)_i$ in the Fokker-Planck equation gives equation 14. Rearranging this identity gives equation 15. □

Example 1. *For the Carmona inter-bank model, $f = a(\widetilde{m} - x)$, $B = \sigma = 1$, and $\widetilde{m} = \int y\, d\widetilde{\rho}(y)$. Therefore*

$$v = a(\widetilde{m} - x) + \alpha - \nu\partial_x \log \rho, \qquad \alpha = v - a(\widetilde{m} - x) + \nu\partial_x \log \rho.$$

Example 2. *For the 12-dimensional quadcopter, the implemented actuator-aligned noise has $\sigma = B$. Since $\mathrm{div}(BB^\top) = 0$ for its control geometry,*

$$v = f + B\alpha - \nu BB^\top \nabla \log \rho = f + B\widetilde{\alpha}, \qquad \widetilde{\alpha} = \alpha - \nu B^\top \nabla \log \rho.$$

Thus the learned field remains control-affine and the physical control is $\alpha = \widetilde{\alpha} + \nu B^\top \nabla \log \rho$.

See Appendices B.2 and B.3.

4 Methodology

Fixed-point iteration on particle level Choose a single-valued ideal best-response selection $\mathcal{R}(\widetilde{\rho}) \in \mathrm{BR}^{\mathrm{SDE}}(\widetilde{\rho})$. The outer update is inspired from the fictitious play in game-theory and MFG versions Brown (1951); Cardaliaguet and Hadikhanloo (2017); Perrin et al. (2020):

$$\widetilde{\rho}^{k+1} = (1 - \beta_{k+1})\widetilde{\rho}^k + \beta_{k+1}\mathcal{R}(\widetilde{\rho}^k), \qquad \beta_{k+1} \in (0, 1]. \tag{16}$$

The choice $\beta_{k+1} = 1$ is the naive fixed-point iteration and a classical choice is $\beta_{k+1} = 1/(k+1)$. In the probability-flow formulation under regularity, the ideal selection is instead taken from $\mathrm{BR}^{\mathrm{PF}}_{\mathrm{reg}} = \mathrm{BR}^{\mathrm{SDE}}_{\mathrm{reg}}$. The algorithm below replaces it by a finite particle approximation. At the particle level, the mixture is implemented on complete trajectories, not independently at each time.

To preserve the prescribed mixture weight with a fixed particle budget and $0 < \beta_{k+1} < 1$, draw $N_o \geq 1$ old trajectories $\widetilde{X}^{k,i}_{[0,T]}$ and $N_n \geq 1$ fresh response trajectories $\widehat{X}^{k+1,j}_{[0,T]}$, independently within and between the two batches. For $N_o + N_n = N$, use the weighted empirical path law

$$\widetilde{\rho}^{k+1,N} = \frac{1-\beta_{k+1}}{N_o} \sum_{i=1}^{N_o} \delta_{\widetilde{X}^{k,i}_{[0,T]}} + \frac{\beta_{k+1}}{N_n} \sum_{j=1}^{N_n} \delta_{\widehat{X}^{k+1,j}_{[0,T]}}. \tag{17}$$

Its time marginals are obtained by evaluating every path at t. Conditional on the stated old and response sampling laws, its value against any bounded test functional is an unbiased estimator of the corresponding exact mixture. When $\beta_{k+1} = 1$, only the new batch is used, with $N_n = N$. An equal-weight rule with $N_n = \lfloor \beta_{k+1} N \rfloor$ instead realizes the rounded weight N_n/N.

Neural ODE At particle level the mixture is taken over complete trajectories. Candidate marginals and their characteristic log densities are generated by the augmented Neural ODE

$$\begin{aligned} \dot{X} &= v_\theta(X,t), & X(0) &\sim \mu_0, \\ \dot{\ell} &= -\nabla \cdot v_\theta(X,t), & \ell_0 &= \log \rho(0), \end{aligned} \qquad \ell = \log \rho, \quad \rho = (X)_{\#}\mu_0. \tag{18}$$

This is the instantaneous change-of-variables formula Chen et al. (2018); Grathwohl et al. (2018). Here ρ_0 is the density of μ_0 when it is available.

Score matching In every inner iteration, a score network is trained on the particle batch in the anticipated population. With uniform distribution sampling of time collocation points $t \in (0,T)$ and $X \sim \rho$, the implicit score-matching loss is $\mathcal{L}_{\text{ISM}}(\phi) = \mathbb{E}\left[\nabla_x \cdot s_\phi(X,t) + \frac{1}{2}\|s_\phi(X,t)\|^2\right]$. At high dimension ($d > 10$), the divergence is replaced by the unbiased Hutchinson estimator. For $\xi \sim \mathcal{N}(0, I_d)$,

$$\mathcal{L}_{\text{SSM}}(\phi) = \mathbb{E}\left[\xi^\top \nabla_x \left(s_\phi(X,t)^\top \xi\right) + \frac{1}{2}\|s_\phi(X,t)\|^2\right]. \tag{19}$$

These losses follow score matching and sliced score matching Hyvärinen (2005); Song et al. (2020).

After the score network training, the stochastic drift and the SDe can be reconstructed at the density level by $b_{\theta,\phi} = v_\theta + \nu s_\phi$ for $\nu > 0$, $b_\theta = v_\theta$ for $\nu = 0$.

All numerical examples use ScoreNet to approximate the score. What changes between cases is the treatment of the log-density and entropy terms.

Remark 1. *The score function can also be traced by neural ODE. Along $X = \Phi_\theta(z,t)$, the log-density and score function are*

$$\frac{d}{dt}\nabla \log \rho(X) = -\nabla(\nabla \cdot v_\theta)(X,t) - (\nabla v_\theta(X,t))^\top \nabla \log \rho(X) \tag{20}$$

once the initial condition $\nabla \log \rho_0$ is known Hernandez et al. (2025); Zhou et al. (2025a).

Remark 2. *We summarize the unified particle algorithm used for general and potential MFGs in the Algorithm 1. When $\nu > 0$, ScoreNet s_ϕ is alternately fitted to particles from the current response law and then frozen during each control update. For a general MFG, the method consists of an outer fixed-point iteration and an inner optimal control problem. At outer iteration k, the anticipated population path $\widetilde{\rho}^{k,N}$ is frozen and detached from the Neural ODE computation graph. The inner problem then computes a best response by minimizing the continuity-equation-constrained probability-flow objective $J^{\text{PF}}(\rho, v^\theta; \widetilde{\rho}^k)$. For a potential MFG, the outer anticipated-population iteration is unnecessary. We directly minimize the full population potential $\mathcal{J}^{\text{PF}}$, the probability-flow representation of $\mathcal{J}^{\text{SDE}}$, over the current particle law. This constitutes a single population-level solve, although its numerical implementation may contain many alternating ScoreNet and ControlNet updates.*

Algorithm 1 Probability flow of second-order mean-field games

1: Initialize an empirical anticipated trajectory law $\widetilde{\rho}^{0,N}$.
2: **for** $k = 0, \ldots, K_{\text{out}} - 1$ **do**
3: Given the anticipated population $\widetilde{\rho}^{k,N}$.
4: Initialize the velocity / control network v_θ and ScoreNet s_ϕ.
5: **for** each inner best-response update **do**
6: Generate candidate trajectories under the flow velocity or control.
7: Update ScoreNet s_ϕ on these trajectories using Algorithm 4 in D.
8: Update the control network v_θ using the current score estimate and the batch of particles in the anticipated population.
9: **end for**
10: Generate a fresh approximate response trajectory law $\widehat{\rho}^{k+1,N}$ and update the anticipated population according to equation 17.
11: **end for**
12: Return the population, velocity / control, and score approximations.

5 Experiments

We perform our probability-flow approach for MFG on several examples: three synthetic distribution transport, an analytical one-dimensional MFG, the nonlinear control-affine dynamics and generative modeling. Here we display the qualitative results in majority. The complete formulations, analytical solutions, numerical evaluations, and seed-wise results are given in Appendix B.

Synthetic distribution transport in MFGs. *We test whether one probability-flow solver transports particle laws through an obstacle and congestion geometry, between disconnected and multimodal supports, and in 1000 dimensions.*

For a target μ_1, define the terminal population potential by $\mathcal{G}_{MMD^2}(\rho) = \lambda_G \mathcal{D}^2_{\text{MMD}}(\rho, \mu_1)$, where

$$\mathcal{D}^2_{\text{MMD}}(\rho, \mu_1) = \mathbb{E}_{\rho\otimes\rho} k - 2\mathbb{E}_{\rho\otimes\mu_1} k + \mathbb{E}_{\mu_1\otimes\mu_1} k, \tag{21}$$

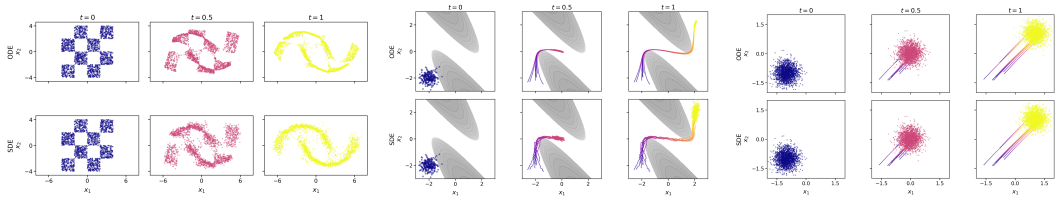
whose first variation in the MFG is

$$\frac{\delta \mathcal{G}_{MMD^2}(T)}{\delta \rho} = G_{MMD^2}(\widetilde{\rho}_T)(x) = 2\lambda_G \left[\int k(x,y)\widetilde{\rho}_T(y)dy - \int k(x,y)\mu_1(y)dy \right]. \tag{22}$$

Here, k is the Gaussian kernel and the details are in B.1.

Settings. Three experiments using the MMD^2 as terminal conidtion. The checkerboard-moons is the potential MFG that directly optimizing the $\mathcal{G}^2_{\text{MMD}}$ potential, with $\nu = 0.05$. Obstacle avoidance and 1000d transport are the MFG between agents, thus compute the expectation of the first-variation terminal G_{MMD^2}. Obstacle avoidance has diffusion coefficient $\nu = 0.5$, while the high-dimensional MFG is setted deterministic for simplicity, transporting from $\mathcal{N}(-\mathbf{1}, 0.3^2 I_{1000})$ to $\mathcal{N}(\mathbf{1}, 0.3^2 I_{1000})$, $\nu = 0$. All panels show $t = 0, 0.5, 1$.

Results. Terminal ODE/SDE $\mathcal{G}^2_{MMD}$ values are $0.0260/0.0184$ for S-curve and $0.00345/0.00511$ for checkerboard-moons. The 1000D has the $\mathcal{G}^2_{MMD} = 0.00864$ and coordinate-mean RMSE 0.0432.



(a) 2D potential MFG (b) 2D MFG with obstacles (c) 1000D MFG

Figure 1: Three synthetic cases, without / with obstacles using potential MFG / MFG.

Inter-bank benchmark with analytical sulution. *For the model of Carmona et al. (2013), let $X \in \mathbb{R}$ denote a bank's log reserve and $\widetilde{m} = \int y\, d\widetilde{\rho}(y)$ be the anticipated population mean. We denote $y(x) = \widetilde{m} - x$, $dX = [ay(X) + \alpha]dt + \sqrt{2\nu}\, dW$, and use running cost $\frac{1}{2}\alpha^2 - q\alpha y + \frac{\epsilon}{2}y^2$ and terminal cost $\frac{c}{2}y_T^2$. The direct probability-flow problem is*

$$\begin{aligned} \inf_{\rho, v} \quad & \int_0^T \int_0^1 [\tfrac{1}{2}\alpha^2 - q\alpha y + \tfrac{\epsilon}{2}y^2]\rho\, dx\, dt + \tfrac{c}{2}\int_0^1 y_T^2 \rho_T\, dx, \\ \text{s.t.} \quad & \alpha = v - ay + \nu\partial_x \log\rho, \qquad \partial_t\rho + \partial_x(\rho v) = 0. \end{aligned} \tag{23}$$

We set $T = a = q = c = 1$, $\epsilon = 1.5$, and $\nu = 0.02$. The analytical equilibrium feedback is

$$\alpha^* = (\eta + q)(m - X), \qquad \dot{\eta} = \eta^2 + 2(a+q)\eta + q^2 - \epsilon, \quad \eta_T = c. \tag{24}$$

Figure 2 shows close agreement on particle-occupied states at $t = \{0, 0.5, 1\}$. The ODE and learned-score SDE are consistent at the density level. Other problem details, experimental settins and results are in Appendix B.2

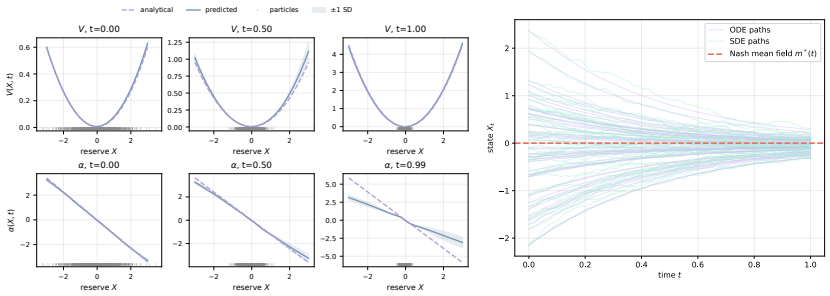


Figure 2: Inter-bank benchmark MFG. Left: learned and analytical controls on the occupied support (grey dash points at the bottom). Right: flow-velocity ODE paths and learned-score reconstructed SDE paths. The dashed line is the Nash mean.

Nonlinear 12D Control-Affine quadcopters MFG. *We examine the 12 dimensional control-affine quadcopters MFG problem as Lin et al. (2021). The state contains position, linear velocity, Euler angles, and angular velocity. Four inputs enter through $\dot{x} = f(x) + B(x)\alpha$, with $B(x) \in \mathbb{R}^{12\times 4}$. We parameterize α_θ directly, use $\nu = 0$ as baseline, and compare with those with a Gaussian congestion interaction potential. Figure 3 shows that congestion changes the transient spread and path curvature while both populations approach the same target. The full dynamics are given in the Appendix B.3.*

Every flight experiment has a twelve-dimensional state

$$x = (p_x, v_x, p_y, v_y, p_z, v_z, \psi, \omega_\psi, \theta, \omega_\theta, \phi, \omega_\phi)^\top \tag{25}$$

and four controls $\alpha = (\alpha_T, \tau_\psi, \tau_\theta, \tau_\phi)^\top$. Thus, 4d in a directory name denotes four actuator/noise channels, not a four-dimensional state. With $m = 0.5$ and $g = 9.81$, the drift is

$$\begin{aligned} \dot{p}_x &= v_x, \quad \dot{v}_x = \alpha_T T_x/m, & \dot{\psi} &= \omega_\psi, \quad \dot{\omega}_\psi = \tau_\psi, \\ \dot{p}_y &= v_y, \quad \dot{v}_y = \alpha_T T_y/m, & \dot{\theta} &= \omega_\theta, \quad \dot{\omega}_\theta = \tau_\theta, \\ \dot{p}_z &= v_z, \quad \dot{v}_z = \alpha_T T_z/m - g, & \dot{\phi} &= \omega_\phi, \quad \dot{\omega}_\phi = \tau_\phi, \end{aligned} \tag{26}$$

where $T_x = \sin\phi\sin\psi + \cos\phi\cos\psi\sin\theta$, $T_y = -\cos\psi\sin\phi + \cos\phi\sin\theta\sin\psi$, and $T_z = \cos\theta\cos\phi$. The horizon is $T = 4$ and the running control cost is $\frac{1}{2}\|\alpha\|^2$. The source positions are independent $\mathcal{N}(-2, 0.5^2)$ variables. The other nine source coordinates are zero in deterministic runs and are independent $\mathcal{N}(0, 0.05^2)$ variables in positive-diffusion runs. The point target is $y_1 = (2, 0, 2, 0, 2, 0)^\top$ in the six position-linear-velocity coordinates, and the terminal objective is

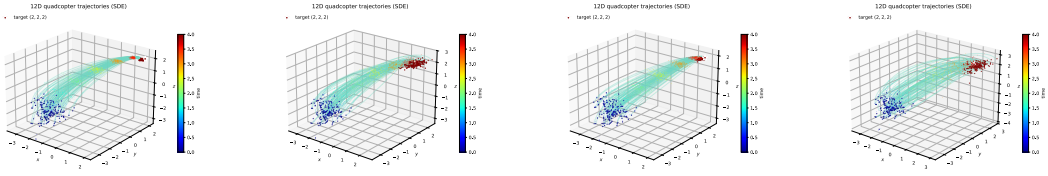
$$\mathcal{G}(\rho_T) = \mathbb{E}_{X_T\sim\rho_T}\|(X_{T,1}, \ldots, X_{T,6}) - y_1\|^2. \tag{27}$$

Let $P_p : \mathbb{R}^{12} \to \mathbb{R}^3$ *denote the position projection,* $P_p x = (p_x, p_y, p_z)^\top$*. The interaction is*

$$W(\widetilde{\rho})(x) = \frac{\gamma}{(2\pi)^{3/2}} \int \exp\left[-\frac{1}{2}\|P_p x - P_p y\|^2\right] \widetilde{\rho}(y)dy, \qquad \gamma \in \{0, 20\}. \tag{28}$$

Actuator-aligned diffusion uses $dX = (f + B\alpha)dt + \sqrt{2\nu}\, B(X)dW^{(4)}$ *and the projected score* $B^\top \nabla \log \rho$*. Positive diffusion has* $\nu = \{0, 0.005\}$*.*

Settings and results. Each row uses 20 outer updates, 100 inner epochs, 150 response particles, 3000 belief particles, 1200 target samples, 80 training steps, 200 evaluation steps, and three seeds. Mean terminal position errors for deterministic baseline in (a), stochastic baseline in (b), deterministic interaction in (c), and stochastic interaction in (d) are respectively $0.355, 0.236, 0.613, 0.394$*.*



(a) Deterministic, $\{\nu = 0, \gamma = 0\}$ (b) Stochastic, $\{\nu = 0.0005, \gamma = 0\}$ (c) Deterministic, $\{\nu = 0, \gamma = 20\}$ (d) Stochastic, $\{\nu = 0.0005, \gamma = 20\}$

Figure 3: Representative one of the probability-flow paths for all four actuator-aligned quadcopter variants. Thin cyan curves follow fixed particles continuously. The colored clouds mark physical time, so the figure is not a terminal-only snapshot.

Generative modeling: latent image translation. *We evaluate our approach on CelebA (Liu et al. 2015). A convolutional VAE encoder* E_ψ *and decoder* D_ψ *are trained once on the official training split and then frozen. We define* Z_0 *as the coordinatewise-normalized posterior mean produced by* $E_\psi(X_0)$ *and solve a PMFG directly in the* $d_z = 128$ *latent space with* $\nu = 0.05$ *and terminal MMD weight* $\lambda_G = 1000$*. The source subset has the CelebA attribute* $\text{Male} = +1$*, whereas the unpaired target subset has* $\text{Male} = -1$*.*

The four rows of Fig. 4 show reconstructed source images, independently sampled reconstructed target images, probability-flow ODE endpoints, and controlled-SDE endpoints reconstructed using the learned score. On 512 test particles, the ODE and SDE attain latent target MMD^2 *values of* 1.316 *and* 1.698*, respectively. FID improves from* 101.11 *for the untranslated source to* 97.38 *and* 97.51*. The displayed ODE reconstructions are visually sharper, but the overall changes remain modest and are limited by VAE reconstruction quality. Accordingly, we regard this experiment as a feasibility study of unpaired latent distribution transport rather than high-fidelity attribute editing. Further implementation details are provided in the Appendix B.4.*

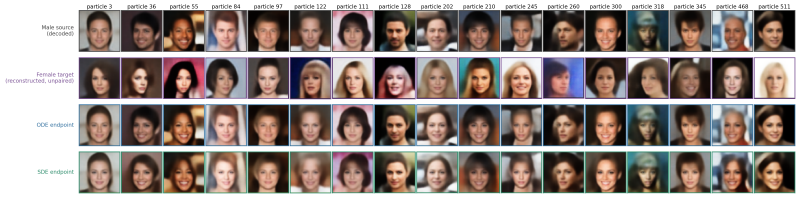


Figure 4: Potential MFG on image translation, from male to female on dataset CelebA.

6 Conclusion

We propose a score-based probability flow approach to solve the MFG and potential MFG in the viewpoint of solving best response. MFG and potential MFG can be solved by optimization methods iteratively. We prove the formulations equivalence bewteen the second-order MFG and its first-order probability flow mathematically. Our approach is examined on various MFG problems, including high-dimensional synthetic data transport, inter-bank systemic risk problem, nonlinear twelve-dimensional quadcopters, distribution matching and generative modeling in image translation.

AI USE STATEMENT

LLM (GPT 5.6) is used for coding assistance and paper polishing.

REPRODUCIBILITY STATEMENT

Experiments were run either on a MacBook Pro with an Apple M1 Pro processor and 16 GB of unified memory or on a machine equipped with an NVIDIA GeForce RTX 4090 GPU with 24 GB of memory. The source code, configuration files, and scripts required to reproduce the reported experiments will be released upon acceptance.

REFERENCES

Yves Achdou and Italo Capuzzo-Dolcetta. Mean field games: numerical methods. SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010. doi: 10.1137/090758477.*

Yves Achdou, Fabio Camilli, and Italo Capuzzo-Dolcetta. Mean field games: convergence of a finite difference method. SIAM Journal on Numerical Analysis*, 51(5):2585–2612, 2013. doi: 10.1137/120882421.*

Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. Numerische Mathematik*, 84(3):375–393, 2000.*

Ariela Briani and Pierre Cardaliaguet. Stable solutions in potential mean field game systems. Nonlinear Differential Equations and Applications NoDEA*, 25(1):1, 2018.*

George W Brown. Iterative solution of games by fictitious play. Act. Anal. Prod Allocation*, 13(1): 374, 1951.*

Pierre Cardaliaguet and Saeed Hadikhanloo. Learning in mean field games: the fictitious play. ESAIM: Control, Optimisation and Calculus of Variations*, 23(2):569–591, 2017.*

Elisabetta Carlini and Francisco José Silva Álvarez. A semi-lagrangian scheme for a degenerate second order mean field game system. arXiv preprint arXiv:1404.5932*, 2014.*

René Carmona and François Delarue. Probabilistic analysis of mean-field games. SIAM Journal on Control and Optimization*, 51(4):2705–2734, 2013. doi: 10.1137/120883499.*

René Carmona, Jean-Pierre Fouque, and Li-Hsien Sun. Mean field games and systemic risk. arXiv preprint arXiv:1308.2172*, 2013.*

Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. Advances in neural information processing systems*, 31, 2018.*

Will Grathwohl, Ricky TQ Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. Ffjord: Free-form continuous dynamics for scalable reversible generative models. arXiv preprint arXiv:1810.01367*, 2018.*

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. Journal of Machine Learning Research*, 13(25):723–773, 2012.*

Jiequn Han, Ruimeng Hu, and Jihao Long. Learning high-dimensional mckean-vlasov forward-backward stochastic differential equations with general distribution dependence. SIAM Journal on Numerical Analysis*, 62(1):1–24, 2024.*

Sebastian Gutierrez Hernandez, Peng Chen, and Haomin Zhou. Pdpo: Parametric density path optimization. arXiv preprint arXiv:2505.18473*, 2025.*

Han Huang, Jiajia Yu, Jie Chen, and Rongjie Lai. Bridging mean-field games and normalizing flows with trajectory regularization. Journal of Computational Physics*, 487:112155, 2023. doi: 10.1016/j.jcp.2023.112155.*

Minyi Huang, Roland P Malhamé, and Peter E Caines. Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle. Communications in Information and Systems*, 2006.*

Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. Journal of Machine Learning Research*, 6(24):695–709, 2005.*

Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. Japanese journal of mathematics*, 2 (1):229–260, 2007.*

Mathieu Lauriere. Numerical methods for mean field games and mean field type control. arXiv preprint arXiv:2106.06231*, 2021.*

Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. Discrete and Continuous Dynamical Systems*, 34(4):1533–1574, 2014. doi: 10.3934/dcds.2014.34.1533.*

Alex Tong Lin, Samy Wu Fung, Wuchen Li, Levon Nurbekyan, and Stanley J Osher. Alternating the population and control neural networks to solve high-dimensional stochastic mean-field games. Proceedings of the National Academy of Sciences*, 118(31):e2024713118, 2021. doi: 10.1073/pnas.2024713118.*

Guan-Horng Liu, Tianrong Chen, Oswin So, and Evangelos Theodorou. Deep generalized schrödinger bridge. Advances in Neural Information Processing Systems*, 35:9374–9388, 2022.*

Jinwei Liu, Lu Ren, Wang Yao, and Xiao Zhang. A constraint-preserving neural network approach for solving mean-field games equilibrium. arXiv preprint arXiv:2501.17450*, 2025.*

Zhiyu Liu, Bo Wu, and Hai Lin. A mean field game approach to swarming robots control. In 2018 Annual American Control Conference (ACC)*, pages 4293–4298. IEEE, 2018.*

Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In Proceedings of International Conference on Computer Vision (ICCV)*, December 2015.*

Dov Monderer and Lloyd S Shapley. Potential games. Games and economic behavior*, 14(1):124–143, 1996.*

Edward Nelson. Derivation of the schrödinger equation from newtonian mechanics. Physical review*, 150(4):1079, 1966.*

Levon Nurbekyan, Siting Liu, and Yat Tin Chow. Monotone inclusion methods for a class of second-order non-potential mean-field games. arXiv preprint arXiv:2403.20290*, 2024.*

Sarah Perrin, Julien Pérolat, Mathieu Laurière, Matthieu Geist, Romuald Elie, and Olivier Pietquin. Fictitious play for mean field games: Continuous time analysis and applications. Advances in neural information processing systems*, 33:13199–13213, 2020.*

Lars Ruthotto, Stanley J Osher, Wuchen Li, Levon Nurbekyan, and Samy Wu Fung. A machine learning framework for solving high-dimensional mean field game and mean field control problems. Proceedings of the National Academy of Sciences*, 117(17):9183–9193, 2020.*

Yang Song, Sahaj Garg, Jiaxin Shi, and Stefano Ermon. Sliced score matching: A scalable approach to density and score estimation. In Uncertainty in artificial intelligence*, pages 574–584. PMLR, 2020.*

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In International Conference on Learning Representations*, 2021.*

Cédric Villani et al. Optimal transport: old and new*, volume 338. Springer, 2009.*

Jiajia Yu, Xiuyuan Cheng, Jian-Guo Liu, and Hongkai Zhao. Convergence analysis and acceleration of fictitious play for general mean-field games via the best response. arXiv preprint arXiv:2411.07989*, 2024.*

Jiajia Yu, Junghwan Lee, Yao Xie, and Xiuyuan Cheng. High-dimensional mean-field games by particle-based flow matching. In International Conference on Learning Representations*, 2026.*

Benjamin J Zhang and Markos A Katsoulakis. A mean-field games laboratory for generative modeling. arXiv preprint arXiv:2304.13534*, 2023.*

Jiaxi Zhao, Mo Zhou, Xinzhe Zuo, and Wuchen Li. Variational conditional normalizing flows for computing second-order mean field control problems. arXiv preprint arXiv:2503.19580*, 2025.*

Mo Zhou, Stanley Osher, and Wuchen Li. Score-based neural ordinary differential equations for computing mean field control problems. Journal of Computational Physics*, 542:114369, 2025a. doi: 10.1016/j.jcp.2025.114369.*

Mo Zhou, Haosheng Zhou, and Ruimeng Hu. Learning mean-field games through mean-field actor-critic flow. arXiv preprint arXiv:2510.12180*, 2025b.*