STABILITY OF UNDERDAMPED LANGEVIN DYNAMICS IN DISTRIBUTIONAL MINIMAX OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Underdamped Langevin methods have received increasing attention in mean-field optimization, motivated by their convergence and discretization properties. In this paper, we show that for distributional minimax optimization, coupling the two players can destabilize underdamped Langevin dynamics. For a bilinear game, we derive an exact friction–coupling threshold separating exponentially stable and unstable underdamped dynamics, while the overdamped counterpart remains stable at every coupling strength. We further show that this instability persists for a bounded smooth interaction. To avoid simultaneous cross-player coupling, we introduce an anchored scheme that freezes the current strategies while running underdamped Langevin dynamics toward their Gibbs responses, and updates the strategies between these runs. For a class of convex-concave functionals of probability laws in the mean-field setting, we establish last-iterate convergence of the proposed scheme, with an $O(1/K)$ decay of the Nikaido–Isoda error. Lastly, we provide a finite-particle, discrete-time implementation with a last-iterate convergence guarantee. More broadly, our results provide a principled framework for understanding and stabilizing underdamped dynamics in distributional minimax optimization.

1 INTRODUCTION

Distributional minimax optimization lifts minimax problems to spaces of probability measures, providing a common framework for continuous games and mean-field formulations of neural network learning (Domingo-Enrich et al., 2020; Kim et al., 2024; Liu et al., 2025). A key appeal of the measure perspective is that it describes training through the evolution of parameter distributions, providing a tractable framework for studying convergence despite nonconvexity in the original parameters (Nitanda & Suzuki, 2017; Chizat & Bach, 2018; Mei et al., 2018; Rotskoff & Vanden-Eijnden, 2018). Within this framework, overdamped mean-field Langevin dynamics (MFLD) provide a natural approach to entropy-regularized optimization, with convergence guarantees established for mean-field minimization and distributional games (Hu et al., 2021; Nitanda et al., 2022; Chizat, 2022; Domingo-Enrich et al., 2020; Kim et al., 2024; Liu et al., 2025).

Underdamped Langevin dynamics have received increasing attention in mean-field optimization, where velocity and friction introduce kinetic transport beyond overdamped dynamics (Kazeykina et al., 2024; Fu & Wilson, 2024). For entropy-regularized minimization, hypocoercive analyses establish convergence of these dynamics, while recent work provides improved mixing and discretization guarantees under suitable assumptions (Villani, 2009; Chen et al., 2024; Fu & Wilson, 2024). However, the minimax setting introduces an additional difficulty: each player's law influences the evolution of the other, creating cross-player feedback absent from single-objective mean-field optimization. Such feedback can interact nontrivially with kinetic effects. Meanwhile, studies of finite-dimensional Euclidean games already show that momentum-based game dynamics can behave differently from single-objective optimization (Gidel et al., 2019; Zhang & Wang, 2021). Recent results for distributional games further reveal a delicate stability picture, with local convergence guarantees and nonconvergent dynamics arising in different regimes (Seo et al., 2026; Mourrat & Pillaud-Vivien, 2026). These results motivate the central question of this work:

When are underdamped Langevin dynamics stable and provably convergent in distributional minimax optimization, and how can they be stabilized under cross-player coupling?

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.

We first study an entropy-regularized bilinear game with Gaussian reference measures. For the directly underdamped descent–ascent dynamics, we decompose each player's evolution into its mean and centered components. The centered components converge to their Gaussian equilibria for every coupling strength, whereas the coupled means are exponentially stable if and only if $\tau\gamma_x\gamma_y > \|B\|_{\mathrm{op}}^2$, where $\tau > 0$ is the regularization temperature, $\gamma_x, \gamma_y > 0$ are the two players' friction coefficients, and B is the bilinear coupling matrix. By contrast, the corresponding overdamped mean dynamics remain stable for every coupling strength. Thus, underdamped dynamics can destabilize the coupled game dynamics even when the corresponding overdamped dynamics are stable. Motivated by this instability, we separate strategy evolution from the computation of Gibbs responses. Building on the anchored best-response framework of Kim et al. (2024), we freeze the current pair of laws at each outer iteration, run underdamped Langevin dynamics toward their linearized Gibbs responses, and then update the strategies through mixtures.

In summary, the contributions of this work are as follows:

- We first derive the exact stability condition for the means of the directly coupled underdamped bilinear dynamics and, under this condition, establish exponential convergence of the position–velocity laws in relative entropy and of the Nikaido–Isoda error (Thm. 1). We then show that the directly coupled underdamped dynamics can be unstable even for bounded functionals (Thm. 2).
- We propose an underdamped anchored best-response scheme that freezes the current strategies during underdamped Langevin runs toward their Gibbs responses, separating the inner underdamped dynamics from the outer strategy updates. For bounded convex-concave functionals in the mean-field setting, we establish last-iterate convergence with an $O(1/K)$ decay of the Nikaido–Isoda error (Thm. 3 and Cor. 1).
- We further characterize how cross-player coupling affects the proposed underdamped anchored scheme. We show that anchoring shifts the coupling-dependent stability condition from the inner underdamped dynamics to the outer update, and derive a convergence condition (Prop. 1).
- We provide a finite-particle, discrete-time implementation of the underdamped anchored scheme and establish a last-iterate convergence guarantee (Thm. 4). Numerical experiments illustrate both the stability transitions and the empirical performance of the proposed algorithm.

2 Related work

Mean-field Langevin dynamics for minimization. Mean-field analyses represent wide two-layer networks by probability measures and relate particle gradient methods to Wasserstein gradient flows (Nitanda & Suzuki, 2017; Mei et al., 2018; Chizat & Bach, 2018; Hu et al., 2021). For entropy-regularized convex functionals, global and exponential convergence guarantees for mean-field Langevin dynamics follow under suitable convexity and uniform log-Sobolev assumptions (Chizat, 2022; Nitanda et al., 2022). Non-asymptotic analyses subsequently covered time discretization, stochastic gradients, and finite-particle approximation (Suzuki et al., 2023a;b). More recent work sharpens functional inequalities and particle approximation in minimization and in structured control/game settings (Chewi et al., 2026; Nitanda et al., 2025; Arnese & Lacker, 2026). Underdamped MFLD augments positions with velocities and is typically analyzed through coupling and hypocoercivity (Villani, 2009; Kazeykina et al., 2024; Chen et al., 2024). Fu & Wilson (2024) proved fast mixing and global total-variation convergence for an exponential-integrator spacetime discretization, and Monmarché (2026) recently established Nesterov-type acceleration for continuous-time mean-field underdamped Langevin dynamics with displacement-convex free energies.

Mean-field Langevin dynamics for minimax optimization. Domingo-Enrich et al. (2020) formulated entropy-regularized continuous zero-sum games over mixed strategies and derived the associated mean-field Langevin descent–ascent (MFL-DA) flow. Global convergence results have mainly required additional structure or modified dynamics, including quasistatic, separated-timescale schemes (Ma & Ying, 2022; Lu, 2023), averaged-gradient, and best-response constructions (Kim et al., 2024; Lu & Monmarché, 2025; Lascu et al., 2025), and strongly convex–strongly concave interactions with finite-particle and time-discretization guarantees (Cai et al., 2025). In the single-loop, discrete-time setting, Liu et al. (2025) established a last-iterate complexity of $\mathcal{O}(\varepsilon^{-1}\log(1/\varepsilon))$ for alternating, two-timescale mean-field Langevin stochastic gradient descent–ascent. The behavior

of the original single-timescale MFL-DA for general smooth games was posed as an open problem by Wang & Chizat (2024). Seo et al. (2026) proved local exponential W_2 convergence on the flat torus and corresponding particle estimates, up to a finite-particle error, on time intervals exponential in N. Wang & Chizat (2026) obtained local convergence in chi-squared divergence for linearly monotone games. Conversely, Mourrat & Pillaud-Vivien (2026) constructed smooth games whose single-timescale MFL-DA flow oscillates under strong coupling and weak regularization. These results distinguish local stability from global convergence, but do not characterize the coupling-dependent effect of underdamped dynamics studied in this paper.

3 Notation and problem setting

3.1 Notation

Let $X = \mathbb{R}^{d_x}$ and $Y = \mathbb{R}^{d_y}$ be the strategy spaces of the two players, where $d_x, d_y \in \mathbb{N}$. Let $\mathcal{P}_2(X)$ and $\mathcal{P}_2(Y)$ denote the sets of probability measures on X and Y, respectively, with finite second moments. We use ρ, σ for the two players' position laws and μ, ν for their joint position–velocity laws. When a law has a density, we use the same symbol for that density. The total variation distance is $\mathrm{TV}(\mu,\nu) := \sup_A |\mu(A) - \nu(A)| = \frac{1}{2}\int |\mathrm{d}\mu - \mathrm{d}\nu|$. If $\mu \ll \nu$, then $D_{\mathrm{KL}}(\mu \mid \nu) := \int \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\mathrm{d}\mu$, $\mathrm{FI}(\mu \mid \nu) := \int \left\|\nabla \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\right\|^2 \mathrm{d}\mu$. For a symmetric positive definite matrix S, we define the modified Fisher information $\mathrm{FI}_S(\mu \mid \nu) := \int \left\|S^{1/2}\nabla \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\right\|^2 \mathrm{d}\mu$. The Gaussian reference measures are $\gamma_x^0(\mathrm{d}x) = (2\pi)^{-d_x/2}e^{-\|x\|^2/2}\mathrm{d}x$, $\gamma_y^0(\mathrm{d}y) = (2\pi)^{-d_y/2}e^{-\|y\|^2/2}\mathrm{d}y$. For the velocity variable we write $g_{\tau,d}(\mathrm{d}v) = (2\pi\tau)^{-d/2}e^{-\|v\|^2/(2\tau)}\mathrm{d}v$. For $L : \mathcal{P}_2(X) \times \mathcal{P}_2(Y) \to \mathbb{R}$, the functional derivatives are denoted by $\frac{\delta L}{\delta \rho}(\rho,\sigma,x)$, $\frac{\delta L}{\delta \sigma}(\rho,\sigma,y)$. The intrinsic derivatives are written as $\nabla_x \frac{\delta L}{\delta \rho}(\rho,\sigma,x), \nabla_y \frac{\delta L}{\delta \sigma}(\rho,\sigma,y)$. We denote by $W_1(\mu,\nu)$ and $W_2(\mu,\nu)$ the 1- and 2-Wasserstein distances between probability measures μ and ν. We write $\|\cdot\|$ for the Euclidean norm, $\|\cdot\|_\infty$ for the supremum norm, and $\|\cdot\|_{\mathrm{op}}$ for the matrix operator norm. We write I_d for the identity matrix and $\lambda_{\max}(\cdot)$ for the largest eigenvalue of a matrix. We write $\mathcal{B}_t^x$ and $\mathcal{B}_t^y$ for independent standard Brownian motions in $\mathbb{R}^{d_x}$ and $\mathbb{R}^{d_y}$, respectively. A detailed summary of the core notation can be found at Tab. 1 of the appendix.

3.2 Problem setting

Fix a temperature $\tau > 0$, throughout this paper, we consider the entropy-regularized game:

$$\min_{\rho\in\mathcal{P}_2(X)} \max_{\sigma\in\mathcal{P}_2(Y)} E_\tau(\rho,\sigma), \qquad E_\tau(\rho,\sigma) := L(\rho,\sigma) + \tau D_{\mathrm{KL}}(\rho \mid \gamma_x^0) - \tau D_{\mathrm{KL}}(\sigma \mid \gamma_y^0). \tag{1}$$

The Nikaido–Isoda error is:

$$\mathrm{NI}_\tau(\rho,\sigma) := \sup_{\bar\sigma\in\mathcal{P}_2(Y)} E_\tau(\rho,\bar\sigma) - \inf_{\bar\rho\in\mathcal{P}_2(X)} E_\tau(\bar\rho,\sigma).$$

This quantity measures the total improvement of the two players and vanishes at a mixed Nash equilibrium (MNE). A pair with $\mathrm{NI}_\tau(\rho,\sigma) \le \varepsilon$ is called an ε-MNE.

To facilitate our analysis, we impose the following assumptions, adapted from the mean-field literature (Kim et al., 2024). We provide concrete examples satisfying these assumptions in Sec. B, including GAN objectives, and zero-sum Markov games.

Assumption 1 (Convexity-concavity in the laws)**.** *The functional L satisfies the following convexity-concavity conditions:*

$$L(\rho',\sigma) - L(\rho,\sigma) \ge \int_X \frac{\delta L}{\delta \rho}(\rho,\sigma,x)(\rho'-\rho)(\mathrm{d}x), \quad L(\rho,\sigma') - L(\rho,\sigma) \le \int_Y \frac{\delta L}{\delta \sigma}(\rho,\sigma,y)(\sigma'-\sigma)(\mathrm{d}y).$$

Assumption 2 (Bounded derivatives)**.** *The first and second functional derivatives of L exist. There are finite constants $C_x, C_y, C_{xx}, C_{xy}, C_{yx}, C_{yy}$ such that, for all laws in $\mathcal{P}_2(X) \times \mathcal{P}_2(Y)$ and all spatial points, $\left\|\frac{\delta L}{\delta \rho}(\rho,\sigma,\cdot)\right\|_\infty \le C_x$, $\left\|\frac{\delta L}{\delta \sigma}(\rho,\sigma,\cdot)\right\|_\infty \le C_y$, and $\left\|\frac{\delta^2 L}{\delta \rho^2}\right\|_\infty \le$*

C_{xx}, $\left\|\frac{\delta^2 L}{\delta\rho\delta\sigma}\right\|_\infty \le C_{xy}$, $\left\|\frac{\delta^2 L}{\delta\sigma\delta\rho}\right\|_\infty \le C_{yx}$, $\left\|\frac{\delta^2 L}{\delta\sigma^2}\right\|_\infty \le C_{yy}$. *The mixed second functional derivatives are symmetric in the usual sense:* $\frac{\delta^2 L}{\delta\rho\delta\sigma}(\rho,\sigma,x,y) = \frac{\delta^2 L}{\delta\sigma\delta\rho}(\rho,\sigma,y,x)$.

Assumption 3 (Regularity of L)**.** *We assume the intrinsic derivatives of L are uniformly bounded and Lipschitz continuous with respect to the inputs and the laws. That is, there exist constants $G_x, L_x, L'_x > 0$ such that $\|\nabla_x \frac{\delta L}{\delta\rho}(\rho,\sigma,x)\| \le G_x$ and*

$$\|\nabla_x \frac{\delta L}{\delta\rho}(\rho,\sigma,x) - \nabla_x \frac{\delta L}{\delta\rho}(\rho',\sigma',x')\| \le L_x\|x-x'\| + L'_x\big(W_1(\rho,\rho') + W_1(\sigma,\sigma')\big).$$

The same properties hold for $\nabla_y \frac{\delta L}{\delta\sigma}(\rho,\sigma,y)$ with $G_y, L_y, L'_y > 0$.

Under Asm. 1 to 3, the game has a unique MNE $(\rho^\star,\sigma^\star)$ Kim et al. (2024). Next, for each pair (ρ,σ), define the proximal Gibbs responses:

$$\widehat{\rho}_{\rho,\sigma}(x) \propto \exp\left(-\frac{1}{\tau}\frac{\delta L}{\delta\rho}(\rho,\sigma,x)\right)\gamma_x^0(x), \qquad \widehat{\sigma}_{\rho,\sigma}(y) \propto \exp\left(\frac{1}{\tau}\frac{\delta L}{\delta\sigma}(\rho,\sigma,y)\right)\gamma_y^0(y). \tag{2}$$

The Gibbs laws in Eqn. 2 solve entropy-regularized linearizations at (ρ,σ). They are best responses when L is affine in each law (Lascu et al., 2025). The bounded perturbations of the Gaussian references also yield uniform log-Sobolev inequalities, a key ingredient in the mean-field Langevin dynamics literature (Chizat, 2022; Suzuki et al., 2023b; Kim et al., 2024). The proof of the following lemma is given in Sec. C.1.

Lemma 1 (Uniform log-Sobolev inequalities)**.** *Under Asm. 2, the Gibbs law $\widehat{\rho}_{\rho,\sigma}$ in Eqn. 2 satisfies the following log-Sobolev inequality (LSI) uniformly in (ρ,σ): $D_{\mathrm{KL}}(q \mid \widehat{\rho}_{\rho,\sigma}) \le \frac{\mathrm{FI}(q|\widehat{\rho}_{\rho,\sigma})}{2\alpha_x}$ with LSI constant $\alpha_x = e^{-2C_x/\tau}$, for every probability density q on X with finite relative Fisher information. The identical statement holds on Y for $\widehat{\sigma}_{\rho,\sigma}$ with LSI constant $\alpha_y = e^{-2C_y/\tau}$.*

4 Instability of direct underdamped dynamics

Let $\mu_t = \mathrm{Law}(X_t,V_t)$ and $\nu_t = \mathrm{Law}(Y_t,W_t)$, with position marginals ρ_t and σ_t. The direct underdamped descent-ascent dynamics are given in Alg. 1.

Algorithm 1: Direct underdamped MFL–DA

Input: Temperature $\tau > 0$, frictions $\gamma_x, \gamma_y > 0$, horizon T, initial phase laws μ_0, ν_0
Run for $t \in [0,T]$:

$$\mathrm{d}X_t = V_t\,\mathrm{d}t, \qquad \mathrm{d}V_t = \left(-\gamma_x V_t - \tau X_t - \nabla_x \frac{\delta L}{\delta\rho}(\rho_t,\sigma_t,X_t)\right)\mathrm{d}t + \sqrt{2\gamma_x\tau}\,\mathrm{d}\mathcal{B}_t^x, \tag{3}$$

$$\mathrm{d}Y_t = W_t\,\mathrm{d}t, \qquad \mathrm{d}W_t = \left(-\gamma_y W_t - \tau Y_t + \nabla_y \frac{\delta L}{\delta\sigma}(\rho_t,\sigma_t,Y_t)\right)\mathrm{d}t + \sqrt{2\gamma_y\tau}\,\mathrm{d}\mathcal{B}_t^y. \tag{4}$$

Output: position–velocity law trajectories $(\mu_t,\nu_t)_{t\ge0}$, with position marginals $(\rho_t,\sigma_t)_{t\ge0}$

We first study a bilinear example in which the stability mechanism of the directly coupled dynamics can be characterized explicitly. Consider:

$$L_B(\rho,\sigma) = \iint x^\top B y\,\rho(\mathrm{d}x)\sigma(\mathrm{d}y), \tag{5}$$

with $B \in \mathbb{R}^{d_x\times d_y}$ denoting the coupling matrix. Write the means for each player as $\bar{x}_\rho = \int x\,\rho(\mathrm{d}x)$, $\bar{y}_\sigma = \int y\,\sigma(\mathrm{d}y)$. Then we can re-express Eq. (5) as $L_B(\rho,\sigma) = \bar{x}_\rho^\top B\bar{y}_\sigma$, so the interaction between the two players depends only on their means. The corresponding Gibbs responses and equilibrium are explicit, as shown in the following lemma. The proof is provided in Sec. C.2.

Lemma 2 (Explicit Gibbs responses and equilibrium)**.** *The bilinear Gibbs responses and the unique equilibrium are $\widehat{\rho}_{\rho,\sigma} = \mathcal{N}(-B\bar{y}_\sigma/\tau, I_{d_x})$, $\widehat{\sigma}_{\rho,\sigma} = \mathcal{N}(B^\top\bar{x}_\rho/\tau, I_{d_y})$, and $\rho^\star = \gamma_x^0$, $\sigma^\star = \gamma_y^0$. The position–velocity equilibrium laws are $\mu^\star = \gamma_x^0 \otimes g_{\tau,d_x}$, $\nu^\star = \gamma_y^0 \otimes g_{\tau,d_y}$.*

To connect the dynamics to the game objective, the next lemma separates the Nikaido–Isoda error into KL divergences from the equilibrium laws and quadratic terms involving the player means. Its proof is deferred to Sec. C.3.

Lemma 3 (Nikaido–Isoda gap decomposition)**.** *Consider Eq.* (5)*, for $\rho \in \mathcal{P}_2(X)$ and $\sigma \in \mathcal{P}_2(Y)$,*

$$\mathrm{NI}_\tau(\rho,\sigma) = \tau D_{\mathrm{KL}}(\rho \mid \gamma_x^0) + \tau D_{\mathrm{KL}}(\sigma \mid \gamma_y^0) + \frac{\|B^\top \bar{x}_\rho\|^2 + \|B\bar{y}_\sigma\|^2}{2\tau}. \tag{6}$$

The mean-dependent terms in Lemma 3 motivate studying the evolution of the player means. Since the bilinear interaction depends on each law only through its mean, taking expectations yields a closed system for the mean trajectories, as shown in the next lemma with proof given in Sec. C.4.

Lemma 4 (Evolution of the mean trajectories)**.** *Consider the dynamics Eq.* (3)*–Eq.* (4) *with L_B defined in Eq.* (5)*, define the means $(a_t, p_t, b_t, q_t) = (\mathbb{E}X_t, \mathbb{E}V_t, \mathbb{E}Y_t, \mathbb{E}W_t)$, then the means satisfy:*

$$\begin{aligned} \dot{a}_t &= p_t, & \dot{p}_t &= -\gamma_x p_t - \tau a_t - B b_t, \\ \dot{b}_t &= q_t, & \dot{q}_t &= -\gamma_y q_t - \tau b_t + B^\top a_t. \end{aligned} \tag{7}$$

Together, Lemmas 3 and 4 show that the player means both contribute to the Nikaido–Isoda error and capture the interaction between the two players. We next separate the mean dynamics from the centered fluctuations, define the centered variables:

$$\widetilde{X}_t := X_t - a_t, \qquad \widetilde{V}_t := V_t - p_t, \qquad \widetilde{Y}_t := Y_t - b_t, \qquad \widetilde{W}_t := W_t - q_t.$$

Subtracting the mean equations Eq. (7) from the bilinear specialization of Eq. (3)–Eq. (4) gives

$$\begin{aligned} \mathrm{d}\widetilde{X}_t &= \widetilde{V}_t\,\mathrm{d}t, & \mathrm{d}\widetilde{V}_t &= (-\gamma_x \widetilde{V}_t - \tau \widetilde{X}_t)\,\mathrm{d}t + \sqrt{2\gamma_x\tau}\,\mathrm{d}\mathcal{B}_t^x, \\ \mathrm{d}\widetilde{Y}_t &= \widetilde{W}_t\,\mathrm{d}t, & \mathrm{d}\widetilde{W}_t &= (-\gamma_y \widetilde{W}_t - \tau \widetilde{Y}_t)\,\mathrm{d}t + \sqrt{2\gamma_y\tau}\,\mathrm{d}\mathcal{B}_t^y. \end{aligned} \tag{8}$$

Notably, the coupling terms Bb_t and $B^\top a_t$ cancel exactly after centering. Thus all cross-player feedback is confined to the coupled mean system Eq. (7), whereas the centered laws evolve as two decoupled quadratic underdamped Langevin processes, independently of B. This decomposition separates the full dynamics into two parts:

$$\underbrace{\text{full dynamics}}_{Eq.\ (3)\text{–}Eq.\ (4)} = \underbrace{\text{coupled mean dynamics}}_{Eq.\ (7)} + \underbrace{\text{centered quadratic Langevin dynamics}}_{Eq.\ (8)}. \tag{9}$$

This dynamical decomposition has an exact counterpart at the level of relative entropy. Let $\widetilde{\mu}_t := \mathrm{Law}(\widetilde{X}_t, \widetilde{V}_t)$, $\widetilde{\nu}_t := \mathrm{Law}(\widetilde{Y}_t, \widetilde{W}_t)$. Since the equilibrium position–velocity laws in Lemma 2 are centered Gaussians, the Gaussian translation identity (Lemma 7) gives

$$\begin{aligned} &D_{\mathrm{KL}}(\mu_t \mid \mu^\star) + D_{\mathrm{KL}}(\nu_t \mid \nu^\star) \\ &= \frac{1}{2}\underbrace{\left(\|a_t\|^2 + \tau^{-1}\|p_t\|^2 + \|b_t\|^2 + \tau^{-1}\|q_t\|^2\right)}_{=:Q_t\ \text{(mean energy)}} + \underbrace{D_{\mathrm{KL}}(\widetilde{\mu}_t \mid \mu^\star) + D_{\mathrm{KL}}(\widetilde{\nu}_t \mid \nu^\star)}_{\text{centered fluctuations}}. \end{aligned} \tag{10}$$

Thus, convergence of the full position–velocity laws separates into two tasks: hypocoercive relaxation of the centered fluctuations and stability of the coupled mean system. The coupled mean system determines whether the dynamics are stable, while the centered processes converge to their Gaussian equilibrium laws by hypocoercivity. Combining these two components lifts mean stability to convergence of the position–velocity laws, yielding the following result.

Theorem 1. *The mean system Eq.* (7) *is exponentially stable if and only if $\tau\gamma_x\gamma_y > \|B\|_{\mathrm{op}}^2$. Under this condition, there are matrices $S_x, S_y \succ 0$, depending only on the temperature and frictions, and constants $C_{\mathrm{bil}} \geq 1$, $\kappa_{\mathrm{bil}} > 0$, depending only on $(\tau, \gamma_x, \gamma_y, B)$, such that, for any initial phase laws with finite relative entropy and relative Fisher information with respect to $\mu^\star$ and $\nu^\star$ we have*

$$D_{\mathrm{KL}}(\mu_t \mid \mu^\star) + D_{\mathrm{KL}}(\nu_t \mid \nu^\star) \leq C_{\mathrm{bil}} e^{-\kappa_{\mathrm{bil}} t}\mathcal{E}_0, \tag{11}$$

$$\mathrm{NI}_\tau(\rho_t, \sigma_t) \leq (\tau + \|B\|_{\mathrm{op}}^2/\tau) C_{\mathrm{bil}} e^{-\kappa_{\mathrm{bil}} t}\mathcal{E}_0, \tag{12}$$

where $\mathcal{E}_0 := D_{\mathrm{KL}}(\mu_0 \mid \mu^\star) + D_{\mathrm{KL}}(\nu_0 \mid \nu^\star) + \mathrm{FI}_{S_x}(\mu_0 \mid \mu^\star) + \mathrm{FI}_{S_y}(\nu_0 \mid \nu^\star)$.

The proof is given in Sec. C.5. For each singular value s of B, the coupled mean dynamics reduce to a four-dimensional linear system with characteristic polynomial $(z^2+\gamma_x z+\tau)(z^2+\gamma_y z+\tau)+s^2$. The Routh–Hurwitz determinant (Gantmakher, 2000) has the sign of $\tau\gamma_x\gamma_y - s^2$. Taking the largest singular value therefore yields the exact threshold $\tau\gamma_x\gamma_y > \|B\|_{\mathrm{op}}^2$.

Remark 1. *For the same bilinear game, the overdamped mean equations are* $\dot{a} = -\tau a - Bb, \quad \dot{b} = -\tau b + B^\top a$. *For each singular value* s *of* B, *the corresponding eigenvalues are* $-\tau \pm s$i, *so the overdamped mean dynamics are stable for every* B. *In contrast, Thm. 1 shows that the underdamped dynamics require* $\tau\gamma_x\gamma_y > \|B\|_{\mathrm{op}}^2$. *Thus introducing velocity can destabilize a game whose overdamped counterpart remains stable. This motivates freezing both laws during each underdamped response computation, so that the inner dynamics become fixed-target sampling problems, see Sec. 5.*

Remark 2. *The difference between the underdamped and overdamped dynamics can also be seen at the level of the mean energy. The mean energy* Q_t *defined in Eq.* (10) *satisfies, by Eq.* (7),

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_t = -2\gamma_x\tau^{-1}\|p_t\|^2 - 2\gamma_y\tau^{-1}\|q_t\|^2 - 2\tau^{-1}p_t^\top B b_t + 2\tau^{-1}q_t^\top B^\top a_t.$$

The first two terms are nonpositive and represent dissipation due to friction, whereas the last two terms have no fixed sign and do not generally cancel. Consequently, the mean energy Q_t *need not be decreasing despite the presence of friction. For the overdamped mean equations, by contrast, the corresponding cross terms cancel exactly, giving* $\frac{\mathrm{d}}{\mathrm{d}t}\left(\|a_t\|^2+\|b_t\|^2\right) = -2\tau\left(\|a_t\|^2+\|b_t\|^2\right)$. *Thus, introducing velocity changes how cross-player feedback enters the dynamics.*

Recall from Thm. 1 that, in the one-dimensional case with $\gamma_x = \gamma_y = \gamma$, the bilinear dynamics are unstable when $B^2 > \tau\gamma^2$. We next show that instability can also occur for a bounded smooth functional. The proof is deferred to Sec. D.

Theorem 2. *Let* $d_x = d_y = 1$, $\gamma_x = \gamma_y = \gamma > 0$, *and let* $B > 0$ *satisfy* $B^2 > \tau\gamma^2$. *Consider* $L_\epsilon(\rho,\sigma) = \frac{B}{\epsilon^2}\iint \sin(\epsilon x)\sin(\epsilon y)\,\rho(\mathrm{d}x)\sigma(\mathrm{d}y)$. *For every sufficiently small fixed* $\epsilon > 0$, L_ϵ *satisfies Asm. 1 to 3, with constants that may depend on* ϵ. *Let* $(\mu^\star,\nu^\star)$ *be the position–velocity equilibrium, where* $\mu^\star = \nu^\star = \mathcal{N}(0,1)\otimes g_{\tau,1}$. *There exists* $r > 0$ *such that, for every* $\eta > 0$, *there exist initial phase laws* $\mu_0,\nu_0 \in \mathcal{P}_2(\mathbb{R}^2)$ *and a finite time* $T > 0$ *for which the solution of Eq.* (3)*–Eq.* (4) *satisfies*

$$W_2(\mu_0,\mu^\star) + W_2(\nu_0,\nu^\star) < \eta, \qquad W_2(\mu_T,\mu^\star) + W_2(\nu_T,\nu^\star) \geq r.$$

Thus, trajectories starting arbitrarily close to the equilibrium can leave a fixed neighborhood under the original dynamics, showing that the instability persists for a bounded smooth functional.

5 Underdamped mean-field Langevin anchored best response

The bilinear analysis shows that direct underdamped coupling can destabilize the game. This motivates separating the inner Langevin runs from strategy updates. In this section, we propose an underdamped mean-field Langevin anchored best-response (underdamped MFL–ABR) scheme and establish its convergence properties.

5.1 Underdamped MFL–ABR

The algorithm is given in Alg. 2, which is inspired by the anchored best-response framework of Lascu et al. (2025); Kim et al. (2024). We freeze the current pair of laws during each inner Langevin run and update the strategies only at the outer iterations. For a generic frozen pair (ρ,σ), the sampling potentials for both players are

$$\frac{\delta L}{\delta\rho}(\rho,\sigma,x) + \frac{\tau}{2}\|x\|^2, \quad -\frac{\delta L}{\delta\sigma}(\rho,\sigma,y) + \frac{\tau}{2}\|y\|^2.$$

We denote the corresponding position–velocity laws at time t by $\mu_t^{\rho,\sigma} = \mathrm{Law}(X_t,V_t)$ and $\nu_t^{\rho,\sigma} = \mathrm{Law}(Y_t,W_t)$, with invariant laws $\Pi_x^{\rho,\sigma} := \widehat{\rho}_{\rho,\sigma}\otimes g_{\tau,d_x}, \quad \Pi_y^{\rho,\sigma} := \widehat{\sigma}_{\rho,\sigma}\otimes g_{\tau,d_y}$. Thus, with (ρ,σ) frozen, each inner dynamics targets a fixed Gibbs response and has no simultaneous cross-player feedback. At each outer iteration, we run these dynamics for a finite horizon and update the strategies by mixtures of their terminal position laws.

Algorithm 2: Underdamped MFL–ABR

Input: Number of outer iterations $K \geq 1$, temperature $\tau > 0$, frictions $\gamma_x, \gamma_y > 0$, outer step size $\beta \in (0, 1]$, response horizon $T \geq 0$

Initialize $\rho_0 = \gamma_x^0$, $\sigma_0 = \gamma_y^0$.

for $k = 0, \ldots, K-1$ **do**

Freeze (ρ_k, σ_k) and independently restart $(X_0, V_0) \sim \gamma_x^0 \otimes g_{\tau,d_x}$ and $(Y_0, W_0) \sim \gamma_y^0 \otimes g_{\tau,d_y}$.

Run for $t \in [0, T]$:

$$\mathrm{d}X_t = V_t\,\mathrm{d}t, \qquad \mathrm{d}V_t = \big(-\gamma_x V_t - \tau X_t - \nabla_x \frac{\delta L}{\delta \rho}(\rho_k, \sigma_k, X_t)\big)\,\mathrm{d}t + \sqrt{2\gamma_x \tau}\,\mathrm{d}\mathcal{B}_t^x, \quad (13)$$

$$\mathrm{d}Y_t = W_t\,\mathrm{d}t, \quad \mathrm{d}W_t = \big(-\gamma_y W_t - \tau Y_t + \nabla_y \frac{\delta L}{\delta \sigma}(\rho_k, \sigma_k, Y_t)\big)\,\mathrm{d}t + \sqrt{2\gamma_y \tau}\,\mathrm{d}\mathcal{B}_t^y. \quad (14)$$

Set $\mathcal{R}_x^T(\rho_k, \sigma_k) = \mathrm{Law}(X_T)$ and $\mathcal{R}_y^T(\rho_k, \sigma_k) = \mathrm{Law}(Y_T)$.

Update $\rho_{k+1} = (1-\beta)\rho_k + \beta \mathcal{R}_x^T(\rho_k, \sigma_k)$ and $\sigma_{k+1} = (1-\beta)\sigma_k + \beta \mathcal{R}_y^T(\rho_k, \sigma_k)$.

end

Output: Last-iterate strategies (ρ_K, σ_K)

5.2 Convergence of underdamped MFL–ABR

To analyze the convergence of underdamped MFL–ABR, define

$$G(\rho, \sigma) = D_{\mathrm{KL}}(\rho \mid \widehat{\rho}_{\rho,\sigma}) + D_{\mathrm{KL}}(\sigma \mid \widehat{\sigma}_{\rho,\sigma}). \quad (15)$$

The entropy sandwich in Lemma 9 gives

$$D_{\mathrm{KL}}(\rho \mid \rho^\star) + D_{\mathrm{KL}}(\sigma \mid \sigma^\star) \leq \mathrm{NI}_\tau(\rho, \sigma)/\tau \leq G(\rho, \sigma). \quad (16)$$

Thus, G controls both the Nikaido–Isoda error and the distance to equilibrium. We first analyze the frozen inner dynamics. Since the noise in Eqs. (13) and (14) acts only on velocities, convergence to the Gibbs responses requires a hypocoercive argument.

Lemma 5 (Frozen one-player underdamped hypocoercivity)**.** *Under Asm. 2 and 3, fix (ρ, σ). There exists a matrix $S_x \succ 0$, depending only on (τ, γ_x, L_x), such that, for any initial law $\mu_0^{\rho,\sigma}$ with finite relative entropy and relative Fisher information with respect to $\Pi_x^{\rho,\sigma}$, one has*

$$D_{\mathrm{KL}}(\mu_t^{\rho,\sigma} \mid \Pi_x^{\rho,\sigma}) \leq [D_{\mathrm{KL}}(\mu_0^{\rho,\sigma} \mid \Pi_x^{\rho,\sigma}) + \mathrm{FI}_{S_x}(\mu_0^{\rho,\sigma} \mid \Pi_x^{\rho,\sigma})]\, e^{-\lambda_{\mathrm{hyp},x} t}, \quad (17)$$

where $\lambda_{\mathrm{hyp},x} = \tau\gamma_x \min\{\alpha_x, 1\}/[16(\gamma_x^2 + L_x + \tau)]$, and α_x is the LSI constant from Lemma 1. The analogous statement holds for the y-player.

The construction of S_x, S_y and the proof are given in Sec. E.2. For the Gaussian restarts $\gamma_x^0 \otimes g_{\tau,d_x}$ and $\gamma_y^0 \otimes g_{\tau,d_y}$ in Alg. 2, the initial entropy–Fisher energies relative to the corresponding frozen invariant laws are uniformly bounded over the outer iterations. By the data processing inequality, this bound passes to the position marginals. Combining it with the outer update gives the following last-iterate guarantee.

Theorem 3 (Convergence of underdamped MFL–ABR)**.** *For the laws generated by Alg. 2, under Asm. 1 to 3, there is $C_{\mathrm{out}} \geq 1$, depending only on τ and the derivative bounds in Asm. 2, such that for every $\beta \in (0, 1]$, $T \geq 0$, and $k \geq 0$,*

$$\mathrm{NI}_\tau(\rho_k, \sigma_k) \leq \tau e^{-\beta k} G(\rho_0, \sigma_0) + \tau C_{\mathrm{out}}(\beta + \sqrt{C_{\mathrm{kin}}} e^{-\lambda_{\mathrm{kin}} T/2}), \quad (18)$$

where $C_{\mathrm{kin}} = \max\left\{\frac{2C_x}{\tau} + \lambda_{\max}(S_x)\frac{G_x^2}{\tau^2}, \frac{2C_y}{\tau} + \lambda_{\max}(S_y)\frac{G_y^2}{\tau^2}\right\}$, $\lambda_{\mathrm{kin}} = \min\{\lambda_{\mathrm{hyp},x}, \lambda_{\mathrm{hyp},y}\}$, with S_x, S_y and $\lambda_{\mathrm{hyp},x}, \lambda_{\mathrm{hyp},y}$ as in Lemma 5.

The bound in Thm. 3 contains residual errors from the finite outer step size and finite inner response horizon. By decreasing the outer step size while increasing the response horizon, both errors can be driven to zero, yielding the following convergence guarantee.

Corollary 1 (Convergence of underdamped MFL–ABR with decreasing outer step size)**.** *Under the same assumptions as in Thm. 3, in Alg. 2, use* $\beta_k = \frac{2}{k+2}$, $T_k = \frac{2}{\lambda_{\mathrm{kin}}} \log\big(1 + \sqrt{2C_{\mathrm{kin}}}(k+2)\big)$, *then we have* $\mathrm{NI}_\tau(\rho_K, \sigma_K) = O\big(\frac{1}{K}\big)$.

Remark 3. *Consequently, to achieve* $\mathrm{NI}_\tau(\rho_K, \sigma_K) \le \varepsilon$, *it suffices to take* $K = O(\varepsilon^{-1})$ *outer iterations, with inner time* $\sum_{k<K} T_k = O(\lambda_{\mathrm{kin}}^{-1}\varepsilon^{-1}\log(1/\varepsilon))$ *per player.*

The proof of Thm. 3 and Cor. 1 are deferred to Sec. E.3 and E.4.

5.3 Coupling in underdamped MFL–ABR

The preceding analysis shows that anchoring removes the direct restriction in Thm. 1 by freezing the two players during each response computation. This raises a natural question: If cross-player coupling can destabilize the directly coupled dynamics, as shown in Sec. 4, where does its effect appear in MFL–ABR? The following proposition shows that coupling has not disappeared. Rather, it enters through the stability of the outer update and its interaction with the outer step size.

Proposition 1 (Coupling in MFL–ABR)**.** *Consider underdamped MFL–ABR applied to the bilinear game Eq.* (5). *Fix* $\tau, \gamma_x, \gamma_y, T > 0$ *and let* $(\rho_0, \sigma_0) \in \mathcal{P}_2(X) \times \mathcal{P}_2(Y)$. *For the* x*-response, define* $r_x(T) := 1 - \big(e^{T\left(\begin{smallmatrix} 0 & 1 \\ -\tau & -\gamma_x \end{smallmatrix}\right)}\big)_{11}$, *and define* $r_y(T)$ *analogously. Then the outer means* $a_k = \bar{x}_{\rho_k}$ *and* $b_k = \bar{y}_{\sigma_k}$ *satisfy*

$$\begin{pmatrix} a_{k+1} \\ b_{k+1} \end{pmatrix} = \begin{pmatrix} (1-\beta)I & -\beta r_x(T) B/\tau \\ \beta r_y(T) B^\top/\tau & (1-\beta)I \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix}. \tag{19}$$

This mean iteration is exponentially stable iff the outer step size satisfies $\beta < \frac{2}{1 + r_x(T) r_y(T) \|B\|_{\mathrm{op}}^2/\tau^2}$. *Under the same condition, if* $D_{\mathrm{KL}}(\rho_0 \mid \gamma_x^0) + D_{\mathrm{KL}}(\sigma_0 \mid \gamma_y^0) < \infty$, *the outer iteration converges exponentially in* NI_τ.

The proof can be found at Sec. F.1. In contrast to Thm. 1, the frozen response dynamics do not require a condition of the form $\tau\gamma_x\gamma_y > \|B\|_{\mathrm{op}}^2$. Instead, the effect of coupling appears in the outer update through the outer step size condition in Prop. 1. In particular, for any finite $\|B\|_{\mathrm{op}}$, the outer iteration can be stabilized by choosing β sufficiently small. Thus, anchoring moves the coupling-dependent stability condition from the directly coupled underdamped dynamics to the outer update.

Interestingly, the outer step size condition in Prop. 1 does not require $\mathcal{R}_x^T(\rho_k, \sigma_k)$ and $\mathcal{R}_y^T(\rho_k, \sigma_k)$ to be close to their corresponding Gibbs responses $\widehat{\rho}_{\rho_k,\sigma_k}$ and $\widehat{\sigma}_{\rho_k,\sigma_k}$. By the calculation in Sec. F.1:

$$D_{\mathrm{KL}}\big(\mathcal{R}_x^T(\rho_k, \sigma_k) \mid \widehat{\rho}_{\rho_k,\sigma_k}\big) + D_{\mathrm{KL}}\big(\mathcal{R}_y^T(\rho_k, \sigma_k) \mid \widehat{\sigma}_{\rho_k,\sigma_k}\big) = \frac{(1 - r_x(T))^2 \|Bb_k\|^2 + (1 - r_y(T))^2 \|B^\top a_k\|^2}{2\tau^2}. \tag{20}$$

For small T, the quantity in Eq. (20) need not be small even when the outer step size condition in Prop. 1 holds. However, under the outer step size condition in Prop. 1, the mean iteration in Eq. (19) converges exponentially, with $\lim_{k\to\infty}(a_k, b_k) = (0, 0)$, and hence Eq. (20) converges to zero.

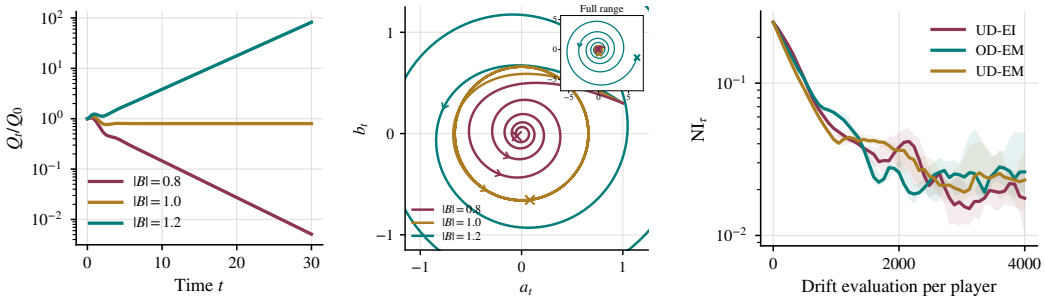
5.4 Finite-particle and discrete-time implementation

Alg. 2 and Thm. 3 describe law-level outer updates with continuous-time inner dynamics. To obtain an implementable algorithm, we represent each player's outer law by N particles and discretize the frozen underdamped dynamics into $M \ge 1$ inner steps with step size $h > 0$. With outer step size parameter $\beta \in (0, 1]$, we run K outer iterations, each replacing $\lfloor \beta N \rfloor \ge 1$ particles per player using the corresponding inner particle cloud. The full algorithm is given in Alg. 3 of Sec. G.

Under Asm. 1 to 3 and sufficiently small h and $\beta > 0$, Thm. 4 establishes a last-iterate convergence guarantee. Writing $\rho_k^{(N)}$ and $\sigma_k^{(N)}$ for the unconditional laws of uniformly selected particles at outer step k, it has the following convergence guarantee:

$$\mathrm{NI}_\tau(\rho_K^{(N)}, \sigma_K^{(N)}) \lesssim \tau\Big[e^{-c\beta K} G(\gamma_x^0, \gamma_y^0) + e^{-\lambda_{\mathrm{kin}} Mh} + \varepsilon_{d_x}(N) + \varepsilon_{d_y}(N) + (1 + d_x + d_y) Mh^3\Big], \tag{21}$$

where $c > 0$ and the implicit constant in $\lesssim$ depend only on problem parameters, and the dimension-dependent rates $\varepsilon_d(N)$ are given in Eq. (122).



(a) Evolution of the energy. (b) Evolution of the trajectory. (c) Particle game.

Figure 1: (a)–(b): Mean dynamics in the bilinear game, with Q_t defined in Eq. (10) and the mean defined in Lemma 4. In (b), crosses mark the terminal state and arrows indicate trajectory direction. (c): Anchored best-response algorithms on the game in Sec. 6.2 at the same drift evaluation budgets. Lines show median estimated NI over ten held-out runs. UD/OD denote underdamped/overdamped samplers. EI/EM denote exponential/Euler–Maruyama discretizations.

6 Numerical Experiments

6.1 Stability Transition in the Bilinear Game

We illustrate the exact stability threshold of the directly coupled underdamped dynamics for the bilinear game in Eq. (5). We set $d_x = d_y = 1$, $\tau = \gamma_x = \gamma_y = 1$, for which Thm. 1 gives the exact stability boundary $|B| = 1$, and initialize the mean state at $(a_0, p_0, b_0, q_0) = (1, 0, 0.3, 0)$. We evaluate the closed mean system in Eq. (7) using exact matrix exponentials. The result is presented in Fig. 1(a)–(b). Consistent with our Thm. 1, the mean energy decays exponentially for $|B| = 0.8$, persists without decay at $|B| = 1$, and grows exponentially for $|B| = 1.2$.

6.2 Bounded Distributional Game

We next consider a bounded mean-field distributional game. Following the numerical experiment of Kim et al. (2024), we consider $L(\rho, \sigma) = \iint \frac{1}{1+e^{-(x-y)^2}} \rho(\mathrm{d}x)\sigma(\mathrm{d}y)$. We use standard-Gaussian reference laws and set the temperature to $\tau = 0.01$. We compare three response samplers within the anchored best response: underdamped exponential discretizations (UD–EI), overdamped Euler–Maruyama (OD–EM), and underdamped Euler–Maruyama (UD–EM). This discretization comparison follows Fu & Wilson (2024), using the EI update with its prescribed noise covariance. Each run uses 1000 particles per player and 4000 drift evaluations per player: 200 outer rounds with 20 inner steps. Parameters are selected in separate tuning runs and then fixed for the test runs. Additional experimental details and parameter search ranges are provided in Sec. H. We evaluate the NI error using Gaussian-smoothed particle laws and their corresponding Gibbs responses, using $\mathrm{NI}_\tau = \tau G$. In Fig. 1(c), we observe that the terminal estimated NI is lower for both underdamped samplers than for the overdamped sampler under the same computational budget, with UD–EI giving the lowest value. This suggests a benefit of using underdamped Langevin dynamics for the inner runs within the anchored best-response framework.

7 Conclusion

In this paper, we show that convergence of underdamped Langevin dynamics to a fixed target does not by itself ensure stability when both players evolve together. The direct bilinear game has an exact threshold, and a bounded smooth interaction can also lead to instability. We next propose an underdamped anchored best response algorithm that decouples both players' updates with convergence guarantees. Important directions include exploring mirror-type outer updates and understanding the implications of these stability mechanisms for large-scale deep network training in the game setting.

AI USE STATEMENT

In this work, we used generative AI tools to aid and polish the writing, including improving the clarity, grammar, notation issue, and presentation of the paper. We also used generative AI tools for literature retrieval, including several recent papers cited in Sec. 2. All references were subsequently checked by the authors for relevance and accuracy. Generative AI tools were also used to assist with the initial implementation of code. The initial AI-assisted code was subsequently refined through multiple rounds of interaction and manual modification by the authors. We did not use generative AI tools to establish research idea or prove the claims in this paper, nor to independently formulate the research contributions or draft the manuscript. We take responsibility for the final content of this work.

ETHICS STATEMENT

This work is primarily theoretical and studies stability and convergence properties of Langevin dynamics for distributional minimax optimization. It does not involve human subjects, data set releases, harmful insights, or other privacy and security issues and ethics issue. So we are not aware of any specific ethical concerns arising from this work.

REPRODUCIBILITY STATEMENT

The assumptions are clearly stated in Sec. 3. The appendix provides the bilinear stability proofs (Sec. C), the bounded functional instability proof (Sec. D), and the anchored best-response convergence proofs (Sec. E). Sec. 6 and Sec. H describe the experimental settings, including the objective, parameter selection, and error evaluation.

REFERENCES

Manuel Arnese and Daniel Lacker. Sharp propagation of chaos for mean field langevin dynamics, control, and games. *arXiv preprint arXiv:2603.10988*, 2026.

Yang Cai, Siddharth Mitra, Xiuyuan Wang, and Andre Wibisono. On the convergence of min-max langevin dynamics and algorithm. In Nika Haghtalab and Ankur Moitra (eds.), *Proceedings of Thirty Eighth Conference on Learning Theory*, volume 291 of *Proceedings of Machine Learning Research*, pp. 677–754. PMLR, 30 Jun–04 Jul 2025. URL https://proceedings.mlr.press/v291/cai25a.html.

Fan Chen, Yiqing Lin, Zhenjie Ren, and Songbo Wang. Uniform-in-time propagation of chaos for kinetic mean field langevin dynamics. *Electronic Journal of Probability*, 29:1–43, 2024.

Sinho Chewi, Atsushi Nitanda, and Matthew S Zhang. Uniform-in-n-log-sobolev inequality for the mean-field langevin dynamics with convex energy. *SIAM Journal on Mathematical Analysis*, 58 (4):3652–3668, 2026.

Lénaïc Chizat. Mean-field langevin dynamics : Exponential convergence and annealing. *Transactions on Machine Learning Research*, 2022. ISSN 2835-8856. URL https://openreview.net/forum?id=BDqzLH1gEm.

Lenaic Chizat and Francis Bach. On the global convergence of gradient descent for over-parameterized models using optimal transport. *Advances in neural information processing systems*, 31, 2018.

Carles Domingo-Enrich, Samy Jelassi, Arthur Mensch, Grant Rotskoff, and Joan Bruna. A mean-field analysis of two-player zero-sum games. *Advances in neural information processing systems*, 33:20215–20226, 2020.

Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability theory and related fields*, 162(3):707–738, 2015.

Qiang Fu and Ashia Camage Wilson. Mean-field underdamped langevin dynamics and its spacetime discretization. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=4qsduFJDEB.

Feliks Ruvimovich Gantmakher. *The theory of matrices*, volume 131. American Mathematical Soc., 2000.

Gauthier Gidel, Reyhane Askari Hemmat, Mohammad Pezeshki, Rémi Le Priol, Gabriel Huang, Simon Lacoste-Julien, and Ioannis Mitliagkas. Negative momentum for improved game dynamics. In *The 22nd international conference on artificial intelligence and statistics*, pp. 1802–1811. PMLR, 2019.

Richard Holley and Daniel W. Stroock. Logarithmic sobolev inequalities and stochastic ising models. *Journal of Statistical Physics*, 46:1159–1194, 1987. URL https://api.semanticscholar.org/CorpusID:59322840.

Kaitong Hu, Zhenjie Ren, David Šiška, and Łukasz Szpruch. Mean-field langevin dynamics and energy landscape of neural networks. In *Annales de l'Institut Henri Poincare (B) Probabilites et statistiques*, volume 57, pp. 2043–2065. Institut Henri Poincaré, 2021.

Anna Kazeykina, Zhenjie Ren, Xiaolu Tan, and Junjian Yang. Ergodicity of the underdamped mean-field langevin dynamics. *The Annals of Applied Probability*, 34(3):3181–3226, 2024.

Juno Kim, Kakei Yamamoto, Kazusato Oko, Zhuoran Yang, and Taiji Suzuki. Symmetric mean-field langevin dynamics for distributional minimax problems. In *International Conference on Learning Representations*, volume 2024, pp. 55056–55098, 2024.

Razvan-Andrei Lascu, Mateusz B Majka, and Łukasz Szpruch. Entropic mean-field min–max problems via best response flow. *Applied Mathematics & Optimization*, 91(2):48, 2025.

Joseph Lehec. Representation formula for the entropy and functional inequalities. In *Annales de l'IHP Probabilités et statistiques*, volume 49, pp. 885–899, 2013.

Zhangyi Liu, Feng Liu, Rui Gao, and Shuang Li. Convergence of mean-field langevin stochastic descent-ascent for distributional minimax optimization. In *Forty-second International Conference on Machine Learning*, 2025.

Yulong Lu. Two-scale gradient descent ascent dynamics finds mixed nash equilibria of continuous games: A mean-field perspective. In *International Conference on Machine Learning*, pp. 22790–22811. PMLR, 2023.

Yulong Lu and Pierre Monmarché. Convergence of time-averaged mean field gradient descent dynamics for continuous multi-player zero-sum games. *arXiv preprint arXiv:2505.07642*, 2025.

Chao Ma and Lexing Ying. Provably convergent quasistatic dynamics for mean-field two-player zero-sum games. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=MP904TiHqJ-.

Song Mei, Andrea Montanari, and Phan-Minh Nguyen. A mean field view of the landscape of two-layer neural networks. *Proceedings of the National Academy of Sciences*, 115(33):E7665–E7671, 2018.

Pierre Monmarché. Nesterov acceleration for the wasserstein minimization of displacement-convex free energies. *arXiv preprint arXiv:2605.13186*, 2026.

Jean-Christophe Mourrat and Loucas Pillaud-Vivien. Oscillating solutions to the mean-field langevin descent-ascent flow. *arXiv preprint arXiv:2604.11134*, 2026.

Atsushi Nitanda and Taiji Suzuki. Stochastic particle gradient descent for infinite ensembles. *arXiv preprint arXiv:1712.05438*, 2017.

Atsushi Nitanda, Denny Wu, and Taiji Suzuki. Convex analysis of the mean field langevin dynamics. In *International Conference on Artificial Intelligence and Statistics*, pp. 9741–9757. PMLR, 2022.

Atsushi Nitanda, Anzelle Lee, Damian Tan Xing Kai, Mizuki Sakaguchi, and Taiji Suzuki. Propagation of chaos for mean-field langevin dynamics and its application to model ensemble. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=SjT6JK4KZv.

Grant M Rotskoff and Eric Vanden-Eijnden. Neural networks as interacting particle systems: Asymptotic convexity of the loss landscape and universal scaling of the approximation error. *stat*, 1050:22, 2018.

Geuntaek Seo, Minseop Shin, Pierre Monmarché, and Beomjun Choi. Local exponential stability of mean-field langevin descent-ascent and associated particle system. *arXiv preprint arXiv:2602.01564*, 2026.

Taiji Suzuki, Atsushi Nitanda, and Denny Wu. Uniform-in-time propagation of chaos for the mean-field gradient langevin dynamics. In *The Eleventh International Conference on Learning Representations*, 2023a. URL https://openreview.net/forum?id=_JScUk9TBUn.

Taiji Suzuki, Denny Wu, and Atsushi Nitanda. Convergence of mean-field langevin dynamics: time-space discretization, stochastic gradient, and variance reduction. *Advances in Neural Information Processing Systems*, 36:15545–15577, 2023b.

Cédric Villani. Hypocoercivity. *Memoirs of the American Mathematical Society*, 202(950), 07 2009. ISSN 0065-9266. doi: 10.1090/s0065-9266-09-00567-5. URL https://cir.nii.ac.jp/crid/1360013171841190784.

Guillaume Wang and Lénaïc Chizat. Open problem: Convergence of single-timescale mean-field langevin descent-ascent for two-player zero-sum games. In *The Thirty Seventh Annual Conference on Learning Theory*, pp. 5345–5350. PMLR, 2024.

Guillaume Wang and Lénaïc Chizat. Local convergence of mean-field langevin dynamics: from gradient flows to linearly monotone games. *arXiv preprint arXiv:2602.11999*, 2026.

Guodong Zhang and Yuanhao Wang. On the suboptimality of negative momentum for minimax optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 2098–2106. PMLR, 2021.

CONTENTS OF THE APPENDIX

The Appendix is organized as follows:

- In Sec. A, we summarize the main symbols and notation used throughout the paper.
- In Sec. B, we provide concrete examples that satisfy the assumptions of this work.
- In Sec. C, we provide the proofs for the direct underdamped descent–ascent dynamics in the bilinear game, including the exact threshold and the full-law convergence result.
- In Sec. D, we prove the instability result for a bounded smooth interaction.
- In Sec. E, we provide the auxiliary results and detailed proofs for the convergence analysis of the underdamped anchored best-response method. In Sec. F, we analyze the exact outer stability in the bilinear game setting.
- In Sec. G, we state the finite-particle, discrete-time algorithm and prove its last-iterate convergence guarantee.
- In Sec. H, we provide further details on the numerical experiments, including the game assumptions, error evaluation, particle updates, and parameter selection, we also provide additional experiments and ablation studies.

A NOTATION

A detailed summary of the core notation can be found at Tab. 1.

Table 1: Core symbols and constants used in this paper. Subscripts $i \in \{x, y\}$ refer to the two players.

Symbol	Type	Description
	Game and probability laws	
$X,\ Y$	Euclidean spaces	Strategy spaces $X = \mathbb{R}^{d_x}$ and $Y = \mathbb{R}^{d_y}$.
$\rho, \sigma;\ \rho_k, \sigma_k$	Probability laws	Generic position laws; outer strategies at round k.
μ_t, ν_t	position–velocity laws	Joint position–velocity laws of the direct dynamics.
$\mu_t^{\rho,\sigma}, \nu_t^{\rho,\sigma}$	position–velocity laws	Joint laws of the frozen inner dynamics; μ_t^k, ν_t^k abbreviate the laws with frozen pair (ρ_k, σ_k).
$\mu^\star, \nu^\star$	position–velocity laws	Equilibrium laws $\rho^\star \otimes g_{\tau,d_x}$ and $\sigma^\star \otimes g_{\tau,d_y}$.
ρ_k^N, σ_k^N	Random measures	Empirical position measures of the two particle clouds at round k.
$\rho_k^{(N)}, \sigma_k^{(N)}$	Position laws	Unconditional laws of uniformly selected particles: $\mathbb{E}\rho_k^N$ and $\mathbb{E}\sigma_k^N$.
$\gamma_x^0, \gamma_y^0;\ g_{\tau,d}$	Gaussian laws	Standard-Gaussian position references; velocity law $\mathcal{N}(0, \tau I_d)$.
$L,\ E_\tau$	Functionals	Unregularized payoff and entropy-regularized game objective, Eq. (1).
τ	> 0	Temperature and coefficient of KL regularization.
$\delta L/\delta\rho,\ \delta L/\delta\sigma$	Scalar fields	First functional derivatives of the payoff in the two laws.
$\nabla_x \frac{\delta L}{\delta \rho},\ \nabla_y \frac{\delta L}{\delta \sigma}$	Vector fields	Spatial gradients of the first functional derivatives.
$\widehat{\rho}_{\rho,\sigma}, \widehat{\sigma}_{\rho,\sigma}$	Probability laws	Gibbs responses to the linearized game at (ρ, σ), Eq. (2).
$\rho^\star, \sigma^\star$	Probability laws	Unique mixed Nash equilibrium of the regularized game.
$\mathrm{NI}_\tau;\ G,\ G_k$	$\mathbb{R}_{\geq 0}$	Equilibrium gap; sum of response KL divergences and its value at round k, Eq. (15).
	Dynamics and convergence bounds	
B	$\mathbb{R}^{d_x \times d_y}$	Coupling matrix in the bilinear game, Eq. (5).
$\mathcal{B}_t^x,\ \mathcal{B}_t^y;\ \mathcal{B}_t$		Brownian motions.
γ_x, γ_y	> 0	Velocity frictions, distinct from the reference laws γ_i^0.
$\beta, \beta_k;\ T, T_k$	$(0, 1];\ \mathbb{R}_{\geq 0}$	Outer step size; inner response horizons.
$\Pi_i^{\rho,\sigma}$	position–velocity law	Invariant law of the frozen sampler: $\widehat{\rho}_{\rho,\sigma} \otimes g_{\tau,d_x}$ for $i = x$, and $\widehat{\sigma}_{\rho,\sigma} \otimes g_{\tau,d_y}$ for $i = y$.
$\mathcal{R}_i^T(\rho, \sigma)$	Position law	Output law of the Gaussian-restarted frozen sampler at time T.
$r_x(T), r_y(T)$	$(0, 2)$	Finite-time bilinear response factors for $T > 0$ in Prop. 1.
$S_i;\ \mathrm{FI}_{S_i}$	Matrix; $\mathbb{R}_{\geq 0}$	Positive-definite matrix and modified Fisher information used in the entropy–Fisher estimate.
$C_i,\ C_{ij}$	$\mathbb{R}_{\geq 0}$	Bounds on first and second functional derivatives, Asm. 2.
$G_i,\ L_i,\ L_i'$	> 0	Intrinsic derivatives bound and spatial/law Lipschitz constants, Asm. 3.
α_i	> 0	Uniform spatial log-Sobolev constant, Lemma 1.
$C_{\mathrm{kin}},\ \lambda_{\mathrm{kin}}$	$\mathbb{R}_{\geq 0};\ > 0$	Uniform initial response-energy bound and common inner convergence rate.
$N,\ h$	$\mathbb{N};\ > 0$	Number of particles per player and inner discretization step size.

B EXAMPLES SATISFYING THE ASSUMPTIONS

We provide examples satisfying Asm. 1 to 3. Write C_b^2 for twice continuously differentiable functions whose values and derivatives through order two are globally bounded. Consider

$$L(\rho,\sigma)=\iint K(x,y)\rho(\mathrm{d}x)\sigma(\mathrm{d}y)-\int b(y)\sigma(\mathrm{d}y),$$

where $K\in C_b^2(X\times Y),\quad b\in C_b^2(Y)$. This functional is affine in each law, with first variations $\frac{\delta L}{\delta\rho}(\rho,\sigma,x)=\int K(x,y)\sigma(\mathrm{d}y)$, and $\frac{\delta L}{\delta\sigma}(\rho,\sigma,y)=\int K(x,y)\rho(\mathrm{d}x)-b(y)$. Its pure second variations vanish and its mixed second variation equals K. The bounded derivatives of K and b imply bounded intrinsic gradients of the functional derivatives. Moreover, their dependence on (ρ,σ) is Lipschitz in W_1 by the Kantorovich–Rubinstein duality. Hence all three assumptions are satisfied. This sufficient condition covers the following examples.

Distributional GAN objectives. Following (Lu, 2023), let ρ be the generated-data law and $f_\sigma(x)=\int\phi(x,y)\sigma(\mathrm{d}y)$ the discriminator. The objective $\mathbb{E}_\rho f_\sigma-\mathbb{E}_{p_{\mathrm{data}}}f_\sigma$ has the preceding form with $K=\phi$ and $b(y)=\mathbb{E}_{p_{\mathrm{data}}}\phi(X,y)$. If $\phi\in C_b^2(X\times Y)$, then $b\in C_b^2(Y)$ as well. This condition is satisfied, for example, by mean-field two-layer discriminators with smooth bounded activation functions.

Fixed-Q subproblems in zero-sum Markov games. In the scheme of Kim et al. (2024), each state s requires solving $L_s(\rho,\sigma)=\iint Q(s,x,y)\rho(\mathrm{d}x)\sigma(\mathrm{d}y)$ with Q held fixed. The preceding criterion applies when $Q(s,\cdot,\cdot)\in C_b^2(X\times Y)$. For a finite state space $\mathcal{S}$, a fixed bounded value function V, and discount $\delta\in[0,1)$, one sufficient construction is

$$Q(s,x,y)=r(s,x,y)+\delta\sum_{s'\in\mathcal{S}}P(s'\mid s,x,y)V(s'),$$

with $r(s,\cdot,\cdot)$ and each transition probability $P(s'\mid s,\cdot,\cdot)$ in C_b^2. This verifies the assumptions for each fixed-Q minimax subproblem.

C PROOFS FOR THE DIRECT UNDERDAMPED DESCENT-ASCENT IN BILINEAR GAME

C.1 PROOF OF LEMMA 1

Proof. The reference γ_x^0 has LSI constant 1. The perturbing potential $\tau^{-1}\delta L/\delta\rho(\rho,\sigma,\cdot)$ has range at most $2C_x/\tau$, so the Holley–Stroock argument (Holley & Stroock, 1987) gives $\alpha_x=e^{-2C_x/\tau}$. The y-player proof is identical. □

C.2 PROOF OF LEMMA 2

Proof. First, we have

$$\begin{aligned}\frac{\delta L_B}{\delta\rho}(\rho,\sigma,x)&=x^\top B\bar{y}_\sigma,\qquad &\frac{\delta L_B}{\delta\sigma}(\rho,\sigma,y)&=y^\top B^\top\bar{x}_\rho,\\ \nabla_x\frac{\delta L_B}{\delta\rho}(\rho,\sigma,x)&=B\bar{y}_\sigma,\qquad &\nabla_y\frac{\delta L_B}{\delta\sigma}(\rho,\sigma,y)&=B^\top\bar{x}_\rho.\end{aligned}\tag{22}$$

Substituting Eq. (22) into Eq. (2), and using the standard Gaussian reference measures, gives

$$\widehat{\rho}_{\rho,\sigma}(\mathrm{d}x)\propto\exp\left(-\frac{1}{\tau}x^\top B\bar{y}_\sigma-\frac{1}{2}\|x\|^2\right)\mathrm{d}x$$

$$\widehat{\sigma}_{\rho,\sigma}(\mathrm{d}y)\propto\exp\left(\frac{1}{\tau}y^\top B^\top\bar{x}_\rho-\frac{1}{2}\|y\|^2\right)\mathrm{d}y.$$

Equivalently:

$$\widehat{\rho}_{\rho,\sigma}=\mathcal{N}\left(-\frac{B\bar{y}_\sigma}{\tau},I_{d_x}\right),\qquad\widehat{\sigma}_{\rho,\sigma}=\mathcal{N}\left(\frac{B^\top\bar{x}_\rho}{\tau},I_{d_y}\right),$$

At a mixed Nash equilibrium $(\rho^\star, \sigma^\star)$, each player is a best response to the other. Hence

$$\rho^\star = \widehat{\rho}_{\rho^\star,\sigma^\star}, \qquad \sigma^\star = \widehat{\sigma}_{\rho^\star,\sigma^\star}.$$

Thus

$$\bar{x}_{\rho^\star} = -\tau^{-1} B \bar{y}_{\sigma^\star}, \qquad \bar{y}_{\sigma^\star} = \tau^{-1} B^\top \bar{x}_{\rho^\star}. \tag{23}$$

Substituting the second identity into the first yields

$$\left(I_{d_x} + \tau^{-2} B B^\top\right) \bar{x}_{\rho^\star} = 0.$$

Since $BB^\top \succeq 0$, the matrix $I_{d_x} + \tau^{-2} BB^\top$ is positive definite. Therefore

$$\bar{x}_{\rho^\star} = 0, \qquad \bar{y}_{\sigma^\star} = 0,$$

where the second equality follows from Eq. (23). Substituting these identities gives

$$\rho^\star = \mathcal{N}(0, I_{d_x}) = \gamma_x^0, \qquad \sigma^\star = \mathcal{N}(0, I_{d_y}) = \gamma_y^0.$$

Thus the mixed Nash equilibrium is unique.

Finally, adjoining the equilibrium velocity laws g_{τ,d_x} and g_{τ,d_y} gives

$$\mu^\star = \gamma_x^0 \otimes g_{\tau,d_x}, \qquad \nu^\star = \gamma_y^0 \otimes g_{\tau,d_y},$$

□

C.3 Proof of Lemma 3

Proof. For fixed ρ, the Gibbs variational formula and the moment-generating function of a standard Gaussian give

$$\begin{aligned}\sup_{\bar{\sigma}\in\mathcal{P}_2(Y)} E_\tau(\rho, \bar{\sigma}) &= \tau D_{\mathrm{KL}}(\rho \mid \gamma_x^0) + \tau \log \int_Y \exp\left(\frac{y^\top B^\top \bar{x}_\rho}{\tau}\right) \gamma_y^0(\mathrm{d}y) \\ &= \tau D_{\mathrm{KL}}(\rho \mid \gamma_x^0) + \frac{\|B^\top \bar{x}_\rho\|^2}{2\tau}.\end{aligned}$$

Similarly, for fixed σ,

$$\begin{aligned}\inf_{\bar{\rho}\in\mathcal{P}_2(X)} E_\tau(\bar{\rho}, \sigma) &= -\tau D_{\mathrm{KL}}(\sigma \mid \gamma_y^0) - \tau \log \int_X \exp\left(-\frac{x^\top B \bar{y}_\sigma}{\tau}\right) \gamma_x^0(\mathrm{d}x) \\ &= -\tau D_{\mathrm{KL}}(\sigma \mid \gamma_y^0) - \frac{\|B\bar{y}_\sigma\|^2}{2\tau}.\end{aligned}$$

Subtracting the two identities proves Eq. (6).

□

C.4 Proof of Lemma 4

Proof. The direct dynamics specialize to

$$\mathrm{d}X_t = V_t \,\mathrm{d}t, \tag{24}$$

$$\mathrm{d}V_t = (-\gamma_x V_t - \tau X_t - B b_t)\,\mathrm{d}t + \sqrt{2\gamma_x \tau}\,\mathrm{d}\mathcal{B}_t^x, \tag{25}$$

$$\mathrm{d}Y_t = W_t \,\mathrm{d}t, \tag{26}$$

$$\mathrm{d}W_t = (-\gamma_y W_t - \tau Y_t + B^\top a_t)\,\mathrm{d}t + \sqrt{2\gamma_y \tau}\,\mathrm{d}\mathcal{B}_t^y. \tag{27}$$

The integral form of Eq. (24) is

$$X_t(\omega) = X_0(\omega) + \int_0^t V_s(\omega)\,\mathrm{d}s.$$

Integrating both sides over the underlying probability space gives

$$\begin{aligned}\int_\Omega X_t(\omega)\,\mathbb{P}(\mathrm{d}\omega) &= \int_\Omega X_0(\omega)\,\mathbb{P}(\mathrm{d}\omega) + \int_\Omega \int_0^t V_s(\omega)\,\mathrm{d}s\,\mathbb{P}(\mathrm{d}\omega) \\ &= \int_\Omega X_0(\omega)\,\mathbb{P}(\mathrm{d}\omega) + \int_0^t \left(\int_\Omega V_s(\omega)\,\mathbb{P}(\mathrm{d}\omega)\right)\mathrm{d}s,\end{aligned}$$

where the second equality follows from Fubini's theorem. Since

$$\int_\Omega X_t(\omega)\,\mathbb{P}(\mathrm{d}\omega) = a_t, \qquad \int_\Omega V_t(\omega)\,\mathbb{P}(\mathrm{d}\omega) = p_t,$$

we obtain $a_t = a_0 + \int_0^t p_s\,\mathrm{d}s$, and hence $\dot{a}_t = p_t$.

Next, the integral form of Eq. (25) is

$$V_t(\omega) = V_0(\omega) + \int_0^t (-\gamma_x V_s(\omega) - \tau X_s(\omega) - Bb_s)\,\mathrm{d}s + \sqrt{2\gamma_x\tau}\int_0^t \mathrm{d}\mathcal{B}_s^x(\omega).$$

Integrating both sides over the probability space and using Fubini's theorem gives

$$\begin{aligned}\int_\Omega V_t(\omega)\,\mathbb{P}(\mathrm{d}\omega) = \int_\Omega V_0(\omega)\,\mathbb{P}(\mathrm{d}\omega) &\\ + \int_0^t \int_\Omega (-\gamma_x V_s(\omega) - \tau X_s(\omega) - Bb_s)\,\mathbb{P}(\mathrm{d}\omega)\,\mathrm{d}s &\\ + \sqrt{2\gamma_x\tau}\int_\Omega \left(\int_0^t \mathrm{d}\mathcal{B}_s^x\right)(\omega)\mathbb{P}(\mathrm{d}\omega).&\end{aligned}$$

The stochastic integral has zero mean: $\int_\Omega \left(\int_0^t \mathrm{d}\mathcal{B}_s^x\right)(\omega)\mathbb{P}(\mathrm{d}\omega) = 0$. Moreover,

$$\int_\Omega (-\gamma_x V_s(\omega) - \tau X_s(\omega) - Bb_s)\,\mathbb{P}(\mathrm{d}\omega) = -\gamma_x p_s - \tau a_s - Bb_s,$$

since Bb_s is deterministic and $\int_\Omega 1\,\mathbb{P}(\mathrm{d}\omega) = 1$. Therefore,

$$p_t = p_0 + \int_0^t (-\gamma_x p_s - \tau a_s - Bb_s)\,\mathrm{d}s,$$

which yields

$$\dot{p}_t = -\gamma_x p_t - \tau a_t - Bb_t.$$

The equations for b_t and q_t follow analogously from Eq. (26) and Eq. (27). □

C.5 Proof of Thm. 1

For the proof, set $\Gamma_x = \gamma_x/\sqrt{\tau}$ and take the coefficients from Lemma 8. Define

$$S_x = \begin{pmatrix} c_{\Gamma_x} & b_{\Gamma_x}\sqrt{\tau} \\ b_{\Gamma_x}\sqrt{\tau} & a_{\Gamma_x}\tau \end{pmatrix} \otimes I_{d_x}, \tag{28}$$

and define S_y identically with x replaced by y. Abbreviate the initial energy by

$$\mathcal{E}_0^{\mathrm{bil}} = D_{\mathrm{KL}}(\mu_0 \mid \mu^\star) + D_{\mathrm{KL}}(\nu_0 \mid \nu^\star) + \mathrm{FI}_{S_x}(\mu_0 \mid \mu^\star) + \mathrm{FI}_{S_y}(\nu_0 \mid \nu^\star). \tag{29}$$

Proof. By Lemma 4, we get

$$\begin{aligned}\dot{a}_t &= p_t, & \dot{p}_t &= -\gamma_x p_t - \tau a_t - Bb_t, \\ \dot{b}_t &= q_t, & \dot{q}_t &= -\gamma_y q_t - \tau b_t + B^\top a_t.\end{aligned} \tag{30}$$

Define

$$\begin{aligned}\widetilde{X}_t &:= X_t - a_t, & \widetilde{V}_t &:= V_t - p_t, \\ \widetilde{Y}_t &:= Y_t - b_t, & \widetilde{W}_t &:= W_t - q_t.\end{aligned}$$

Subtracting Eq. (30) from the original SDEs yields

$$\begin{aligned}
\mathrm{d}\widetilde{X}_t &= \widetilde{V}_t\,\mathrm{d}t,\\
\mathrm{d}\widetilde{V}_t &= (-\gamma_x\widetilde{V}_t - \tau\widetilde{X}_t)\mathrm{d}t + \sqrt{2\gamma_x\tau}\,\mathrm{d}\mathcal{B}^x_t,\\
\mathrm{d}\widetilde{Y}_t &= \widetilde{W}_t\,\mathrm{d}t,\\
\mathrm{d}\widetilde{W}_t &= (-\gamma_y\widetilde{W}_t - \tau\widetilde{Y}_t)\mathrm{d}t + \sqrt{2\gamma_y\tau}\,\mathrm{d}\mathcal{B}^y_t.
\end{aligned} \tag{31}$$

Thus the two centered marginal evolutions are decoupled. Write

$$\widetilde{\mu}_t := \mathrm{Law}(\widetilde{X}_t, \widetilde{V}_t), \qquad \widetilde{\nu}_t := \mathrm{Law}(\widetilde{Y}_t, \widetilde{W}_t).$$

Define

$$\begin{aligned}
\mathcal{H}^0_x &:= D_{\mathrm{KL}}(\widetilde{\mu}_0 \mid \mu^\star) + \mathrm{FI}_{S_x}(\widetilde{\mu}_0 \mid \mu^\star),\\
\mathcal{H}^0_y &:= D_{\mathrm{KL}}(\widetilde{\nu}_0 \mid \nu^\star) + \mathrm{FI}_{S_y}(\widetilde{\nu}_0 \mid \nu^\star).
\end{aligned}$$

Then Lemma 8 gives

$$\begin{aligned}
D_{\mathrm{KL}}(\widetilde{\mu}_t \mid \mu^\star) &\le \mathcal{H}^0_x e^{-\lambda^{\mathrm{hyp}}_{\Gamma_x,\tau}t},\\
D_{\mathrm{KL}}(\widetilde{\nu}_t \mid \nu^\star) &\le \mathcal{H}^0_y e^{-\lambda^{\mathrm{hyp}}_{\Gamma_y,\tau}t}.
\end{aligned} \tag{32}$$

Consequently, with $\kappa_c := \min\{\lambda^{\mathrm{hyp}}_{\Gamma_x,\tau}, \lambda^{\mathrm{hyp}}_{\Gamma_y,\tau}\}$, one has

$$D_{\mathrm{KL}}(\widetilde{\mu}_t \mid \mu^\star) + D_{\mathrm{KL}}(\widetilde{\nu}_t \mid \nu^\star) \le e^{-\kappa_c t}\left(\mathcal{H}^0_x + \mathcal{H}^0_y\right). \tag{33}$$

Next, we diagonalize the deterministic mean interaction. Let $B = U\Sigma V^\top$ be a full singular value decomposition, and let $s_1, \dots, s_r > 0$ be the nonzero singular values. Introduce

$$\alpha := U^\top a, \qquad \pi := U^\top p, \qquad \beta := V^\top b, \qquad \chi := V^\top q.$$

For each coupled singular mode,

$$\begin{aligned}
\ddot{\alpha}_i + \gamma_x\dot{\alpha}_i + \tau\alpha_i + s_i\beta_i &= 0,\\
\ddot{\beta}_i + \gamma_y\dot{\beta}_i + \tau\beta_i - s_i\alpha_i &= 0.
\end{aligned} \tag{34}$$

If the singular value is zero, then the corresponding coordinates decouple and satisfy

$$\begin{aligned}
\ddot{\alpha}_i + \gamma_x\dot{\alpha}_i + \tau\alpha_i &= 0,\\
\ddot{\beta}_i + \gamma_y\dot{\beta}_i + \tau\beta_i &= 0.
\end{aligned}$$

Since $\gamma_x, \gamma_y, \tau > 0$, all roots of their characteristic polynomials have strictly negative real parts, and hence these uncoupled modes are exponentially stable. For a coupled mode with $s_i > 0$, the characteristic polynomial is

$$\begin{aligned}
P_s(\lambda) &= (\lambda^2 + \gamma_x\lambda + \tau)(\lambda^2 + \gamma_y\lambda + \tau) + s^2\\
&= \lambda^4 + (\gamma_x + \gamma_y)\lambda^3 + (\gamma_x\gamma_y + 2\tau)\lambda^2 + \tau(\gamma_x + \gamma_y)\lambda + (\tau^2 + s^2).
\end{aligned} \tag{35}$$

For a monic quartic $z^4 + a_1z^3 + a_2z^2 + a_3z + a_4$, the Routh–Hurwitz criterion (Gantmakher, 2000) is

$$a_j > 0, \qquad a_1a_2 - a_3 > 0, \qquad a_1a_2a_3 - a_3^2 - a_1^2a_4 > 0. \tag{36}$$

For Eq. (35), all coefficients are positive and

$$a_1a_2 - a_3 = (\gamma_x + \gamma_y)(\gamma_x\gamma_y + \tau) > 0.$$

The final Routh–Hurwitz determinant is

$$a_1a_2a_3 - a_3^2 - a_1^2a_4 = (\gamma_x + \gamma_y)^2(\tau\gamma_x\gamma_y - s^2). \tag{37}$$

Hence the singular mode s is exponentially stable if and only if $s^2 < \tau\gamma_x\gamma_y$. Taking the largest singular value proves the threshold given in the theorem.

Next, we will prove Eqs. (11) and (12). Define

$$\mathcal{A}_B := \begin{pmatrix} 0 & I_{d_x} & 0 & 0 \\ -\tau I_{d_x} & -\gamma_x I_{d_x} & -B & 0 \\ 0 & 0 & 0 & I_{d_y} \\ B^\top & 0 & -\tau I_{d_y} & -\gamma_y I_{d_y} \end{pmatrix}, \tag{38}$$

and

$$H := \operatorname{diag}\left(I_{d_x}, \tau^{-1} I_{d_x}, I_{d_y}, \tau^{-1} I_{d_y}\right). \tag{39}$$

Writing $\mathbf{m}_t := (a_t, p_t, b_t, q_t)^\top$, we have $\mathbf{m}_t = e^{t\mathcal{A}_B}\mathbf{m}_0$.

Note that $\lambda \in \operatorname{spec}(\mathcal{A}_B)$ if and only if there exists $(a, b) \neq (0, 0)$ such that

$$\begin{pmatrix} (\lambda^2 + \gamma_x\lambda + \tau) I_{d_x} & B \\ -B^\top & (\lambda^2 + \gamma_y\lambda + \tau) I_{d_y} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0. \tag{40}$$

Substituting $B = U\Sigma V^\top$, since U and V are orthogonal, Eq. (40) is equivalent to

$$\begin{pmatrix} (\lambda^2 + \gamma_x\lambda + \tau) I_{d_x} & \Sigma \\ -\Sigma^\top & (\lambda^2 + \gamma_y\lambda + \tau) I_{d_y} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0. \tag{41}$$

For each nonzero singular value $s_i > 0$, the corresponding coordinates satisfy

$$\begin{pmatrix} \lambda^2 + \gamma_x\lambda + \tau & s_i \\ -s_i & \lambda^2 + \gamma_y\lambda + \tau \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = 0.$$

Thus a nontrivial solution requires

$$\det \begin{pmatrix} \lambda^2 + \gamma_x\lambda + \tau & s_i \\ -s_i & \lambda^2 + \gamma_y\lambda + \tau \end{pmatrix} = (\lambda^2 + \gamma_x\lambda + \tau)(\lambda^2 + \gamma_y\lambda + \tau) + s_i^2 = 0\,, \tag{42}$$

which is the same as the characteristic polynomial in Eq. (35).

So, under the stable condition of the theorem, we have $-\max\{\operatorname{Re}\lambda : \lambda \in \operatorname{spec}(\mathcal{A}_B)\} > 0$.

Denote by:

$$\begin{aligned} \omega_{\text{mean}} &:= -\max\left\{\operatorname{Re}\lambda : \lambda \in \operatorname{spec}(\mathcal{A}_B)\right\} > 0. \\ C_{\text{mean}} &:= \sup_{t\geq 0} e^{\omega_{\text{mean}} t} \left\| H^{1/2} e^{t\mathcal{A}_B} H^{-1/2} \right\|_{\text{op}}^2, \\ \mathcal{Q}_t &:= \|a_t\|^2 + \tau^{-1}\|p_t\|^2 + \|b_t\|^2 + \tau^{-1}\|q_t\|^2, \end{aligned} \tag{43}$$

Clearly we have:

$$\mathcal{Q}_t \leq C_{\text{mean}} e^{-\omega_{\text{mean}} t} \mathcal{Q}_0. \tag{44}$$

Below we show that C_{mean} defined above is finite.

By the Jordan normal form, there exist an invertible matrix S and a block diagonal Jordan matrix J such that $\mathcal{A}_B = SJS^{-1}$. Write $J = \operatorname{diag}\left(J_{\lambda_1}, \ldots, J_{\lambda_m}\right)$, where each Jordan block has the form $J_{\lambda_j} = \lambda_j I_{r_j} + N_j$, with

$$N_j = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}, \quad N_j^{r_j} = 0.$$

For each Jordan block,

$$e^{tJ_{\lambda_j}} = e^{\lambda_j t} e^{tN_j} = e^{\lambda_j t} \sum_{k=0}^{r_j - 1} \frac{t^k N_j^k}{k!}.$$

Hence, by the triangle inequality for the operator norm,

$$\left\| e^{tJ_{\lambda_j}} \right\|_{\text{op}} \leq e^{\operatorname{Re}(\lambda_j) t} \sum_{k=0}^{r_j - 1} \frac{t^k}{k!} \left\| N_j^k \right\|_{\text{op}}. \tag{45}$$

Let $r_{\max} := \max_{1\le j\le m} r_j$, then

$$t^k \le 1 + t^{r_{\max}-1}, \qquad t \ge 0, \qquad 0 \le k \le r_j - 1,$$

Thus, we obtain

$$\left\| e^{tJ_{\lambda_j}} \right\|_{\mathrm{op}} \le C_j \left(1 + t^{r_{\max}-1}\right) e^{\mathrm{Re}(\lambda_j)t}, \qquad t \ge 0, \tag{46}$$

where $C_j := \sum_{k=0}^{r_j-1} \frac{\|N_j^k\|_{\mathrm{op}}}{k!}$. Because J is block diagonal, i.e.,$e^{tJ} = \mathrm{diag}\left(e^{tJ_{\lambda_1}}, \ldots, e^{tJ_{\lambda_m}}\right)$.

Therefore

$$\left\| e^{tJ} \right\|_{\mathrm{op}} = \max_{1\le j\le m} \left\| e^{tJ_{\lambda_j}} \right\|_{\mathrm{op}}. \tag{47}$$

Define $C_J := \max_{1\le j\le m} C_j$. By the definition of ω_{mean},

$$\mathrm{Re}(\lambda_j) \le -\omega_{\mathrm{mean}}, \qquad 1 \le j \le m, \tag{48}$$

so

$$\left\| e^{tJ} \right\|_{\mathrm{op}} \le C_J \left(1 + t^{r_{\max}-1}\right) e^{-\omega_{\mathrm{mean}}t}, \qquad t \ge 0. \tag{49}$$

Since $e^{t\mathcal{A}_B} = Se^{tJ}S^{-1}$, submultiplicativity of the operator norm gives

$$\left\| e^{t\mathcal{A}_B} \right\|_{\mathrm{op}} \le \|S\|_{\mathrm{op}} \left\|S^{-1}\right\|_{\mathrm{op}} C_J \left(1 + t^{r_{\max}-1}\right) e^{-\omega_{\mathrm{mean}}t}, \qquad t \ge 0, \tag{50}$$

For every $\varepsilon > 0$, define

$$C_{\mathrm{poly}}(\varepsilon) := \sup_{t\ge 0} \left(1 + t^{r_{\max}-1}\right) e^{-\varepsilon t}.$$

Since a polynomial multiplied by a strictly decaying exponential is bounded,

$$C_{\mathrm{poly}}(\varepsilon) < \infty.$$

Thus

$$1 + t^{r_{\max}-1} \le C_{\mathrm{poly}}(\varepsilon) e^{\varepsilon t}, \qquad t \ge 0,$$

and Eq. (50) yields

$$\left\| e^{t\mathcal{A}_B} \right\|_{\mathrm{op}} \le \|S\|_{\mathrm{op}} \left\|S^{-1}\right\|_{\mathrm{op}} C_J C_{\mathrm{poly}}(\varepsilon) e^{-(\omega_{\mathrm{mean}}-\varepsilon)t}, \qquad t \ge 0. \tag{51}$$

Taking $\varepsilon = \frac{\omega_{\mathrm{mean}}}{2}$, we obtain

$$\left\| e^{t\mathcal{A}_B} \right\|_{\mathrm{op}} \le \|S\|_{\mathrm{op}} \left\|S^{-1}\right\|_{\mathrm{op}} C_J C_{\mathrm{poly}} \left(\frac{\omega_{\mathrm{mean}}}{2}\right) e^{-\omega_{\mathrm{mean}}t/2}, \qquad t \ge 0. \tag{52}$$

Recall the definition of H in Eq. (39). By submultiplicativity,

$$\begin{aligned} \left\| H^{1/2} e^{t\mathcal{A}_B} H^{-1/2} \right\|_{\mathrm{op}} &\le \left\| H^{1/2} \right\|_{\mathrm{op}} \left\| H^{-1/2} \right\|_{\mathrm{op}} \left\| e^{t\mathcal{A}_B} \right\|_{\mathrm{op}} \\ &\le C_H e^{-\omega_{\mathrm{mean}}t/2}, \end{aligned}$$

where $C_H := \left\| H^{1/2} \right\|_{\mathrm{op}} \left\| H^{-1/2} \right\|_{\mathrm{op}} \|S\|_{\mathrm{op}} \left\|S^{-1}\right\|_{\mathrm{op}} C_J C_{\mathrm{poly}}\left(\frac{\omega_{\mathrm{mean}}}{2}\right) < \infty$. Consequently,

$$\left\| H^{1/2} e^{t\mathcal{A}_B} H^{-1/2} \right\|_{\mathrm{op}}^2 \le C_H^2 e^{-\omega_{\mathrm{mean}}t}, \qquad t \ge 0.$$

The preceding part shows that C_{mean} is finite, i.e., $C_{\mathrm{mean}} \le C_H^2 < \infty$.

Next, we will turn these mean modes into solutions of the full dynamics. Applying Lemma 7 with $\Sigma_x = \mathrm{diag}(I_{d_x}, \tau I_{d_x}), \quad \Sigma_y = \mathrm{diag}(I_{d_y}, \tau I_{d_y})$, gives

$$D_{\mathrm{KL}}(\mu_t \mid \mu^\star) = D_{\mathrm{KL}}(\widetilde{\mu}_t \mid \mu^\star) + \frac{1}{2}\|a_t\|^2 + \frac{1}{2\tau}\|p_t\|^2, \tag{53}$$

$$D_{\mathrm{KL}}(\nu_t \mid \nu^\star) = D_{\mathrm{KL}}(\widetilde{\nu}_t \mid \nu^\star) + \frac{1}{2}\|b_t\|^2 + \frac{1}{2\tau}\|q_t\|^2. \tag{54}$$

Define $C_{\mathrm{bil}} := \max\{1, C_{\mathrm{mean}}\}, \qquad \kappa_{\mathrm{bil}} := \min\{\kappa_c, \omega_{\mathrm{mean}}\}$. Combining Eq. (33), Eq. (44), and Eq. (53)–Eq. (54), we obtain

$$\begin{aligned}
&D_{\mathrm{KL}}(\mu_t \mid \mu^\star) + D_{\mathrm{KL}}(\nu_t \mid \nu^\star) \\
&= D_{\mathrm{KL}}(\widetilde{\mu}_t \mid \mu^\star) + D_{\mathrm{KL}}(\widetilde{\nu}_t \mid \nu^\star) + \frac{1}{2}\mathcal{Q}_t \\
&\le e^{-\kappa_c t}\left\{\mathcal{H}_x^0 + \mathcal{H}_y^0\right\} + \frac{1}{2}C_{\mathrm{mean}} e^{-\omega_{\mathrm{mean}} t}\mathcal{Q}_0 \\
&\le C_{\mathrm{bil}} e^{-\kappa_{\mathrm{bil}} t}\left\{\mathcal{H}_x^0 + \mathcal{H}_y^0 + \frac{1}{2}\mathcal{Q}_0\right\} \\
&= C_{\mathrm{bil}} e^{-\kappa_{\mathrm{bil}} t}\Big\{D_{\mathrm{KL}}(\mu_0 \mid \mu^\star) + D_{\mathrm{KL}}(\nu_0 \mid \nu^\star) + \mathrm{FI}_{S_x}(\widetilde{\mu}_0 \mid \mu^\star) + \mathrm{FI}_{S_y}(\widetilde{\nu}_0 \mid \nu^\star)\Big\} \\
&\le C_{\mathrm{bil}} e^{-\kappa_{\mathrm{bil}} t}\mathcal{E}_0^{\mathrm{bil}}.
\end{aligned} \tag{55}$$

Finally, by the data-processing inequality,

$$D_{\mathrm{KL}}(\rho_t \mid \gamma_x^0) \le D_{\mathrm{KL}}(\mu_t \mid \mu^\star), \qquad D_{\mathrm{KL}}(\sigma_t \mid \gamma_y^0) \le D_{\mathrm{KL}}(\nu_t \mid \nu^\star).$$

Therefore, by Lemmas 3 and 6 and Eq. (11),

$$\begin{aligned}
\mathrm{NI}_\tau(\rho_t, \sigma_t) &\le \left(\tau + \frac{\|B\|_{\mathrm{op}}^2}{\tau}\right)\left\{D_{\mathrm{KL}}(\mu_t \mid \mu^\star) + D_{\mathrm{KL}}(\nu_t \mid \nu^\star)\right\} \\
&\le \left(\tau + \frac{\|B\|_{\mathrm{op}}^2}{\tau}\right) C_{\mathrm{bil}} e^{-\kappa_{\mathrm{bil}} t}\mathcal{E}_0^{\mathrm{bil}},
\end{aligned}$$

which is Eq. (12). □

C.6 Auxiliary lemma

Lemma 6 (KL upper bound for the Nikaido–Isoda gap)**.** *Under the same setting as Lemma 3,*

$$\mathrm{NI}_\tau(\rho, \sigma) \le \left(\tau + \frac{\|B\|_{\mathrm{op}}^2}{\tau}\right)\left[D_{\mathrm{KL}}(\rho \mid \gamma_x^0) + D_{\mathrm{KL}}(\sigma \mid \gamma_y^0)\right]. \tag{56}$$

Proof. The entropy variational inequality and the Gaussian moment-generating function give

$$D_{\mathrm{KL}}(\rho \mid \gamma_x^0) \ge \sup_{u \in \mathbb{R}^{d_x}}\left\{u^\top \bar{x}_\rho - \frac{1}{2}\|u\|^2\right\} = \frac{1}{2}\|\bar{x}_\rho\|^2,$$

and similarly,

$$D_{\mathrm{KL}}(\sigma \mid \gamma_y^0) \ge \frac{1}{2}\|\bar{y}_\sigma\|^2.$$

Hence

$$\|B^\top \bar{x}_\rho\|^2 \le 2\|B\|_{\mathrm{op}}^2 D_{\mathrm{KL}}(\rho \mid \gamma_x^0), \qquad \|B\bar{y}_\sigma\|^2 \le 2\|B\|_{\mathrm{op}}^2 D_{\mathrm{KL}}(\sigma \mid \gamma_y^0).$$

Substituting these bounds into Eq. (6) proves Eq. (56). □

Lemma 7 (Gaussian translation identity)**.** *Let $\Pi = \mathcal{N}(0, \Sigma)$ be a centered Gaussian measure with $\Sigma \succ 0$, let $Z \sim \zeta$ have finite second moment and $D_{\mathrm{KL}}(\zeta \mid \Pi) < \infty$, with mean $m = \mathbb{E}[Z]$, and define $\widetilde{\zeta} := \mathrm{Law}(Z - m)$. Then*

$$D_{\mathrm{KL}}(\zeta \mid \Pi) = D_{\mathrm{KL}}(\widetilde{\zeta} \mid \Pi) + \frac{1}{2}m^\top \Sigma^{-1} m. \tag{57}$$

Moreover, let $S \succ 0$. If $\mathrm{FI}_S(\zeta \mid \Pi) < \infty$, then

$$\mathrm{FI}_S(\zeta \mid \Pi) = \mathrm{FI}_S(\widetilde{\zeta} \mid \Pi) + m^\top \Sigma^{-1} S \Sigma^{-1} m. \tag{58}$$

In particular,

$$\mathrm{FI}_S(\widetilde{\zeta} \mid \Pi) \le \mathrm{FI}_S(\zeta \mid \Pi).$$

Proof. Let π denote the density of Π, and let f denote the density of ζ. Since $\widetilde{\zeta}$ is the translate of ζ by $-m$, its density is $\widetilde{f}(z) = f(z+m)$.

By definition,

$$D_{\mathrm{KL}}(\zeta \mid \Pi) = \int f(x) \log \frac{f(x)}{\pi(x)} \, \mathrm{d}x.$$

Making the change of variables $x = z + m$, so that $\mathrm{d}x = \mathrm{d}z$ and $f(z+m) = \widetilde{f}(z)$, gives

$$D_{\mathrm{KL}}(\zeta \mid \Pi) = \int \widetilde{f}(z) \log \frac{\widetilde{f}(z)}{\pi(z+m)} \, \mathrm{d}z.$$

On the other hand,

$$D_{\mathrm{KL}}(\widetilde{\zeta} \mid \Pi) = \int \widetilde{f}(z) \log \frac{\widetilde{f}(z)}{\pi(z)} \, \mathrm{d}z.$$

Subtracting the two identities yields

$$D_{\mathrm{KL}}(\zeta \mid \Pi) - D_{\mathrm{KL}}(\widetilde{\zeta} \mid \Pi) = \int \log \frac{\pi(z)}{\pi(z+m)} \, \widetilde{\zeta}(\mathrm{d}z).$$

Since $\Pi = \mathcal{N}(0, \Sigma)$, we have $\pi(z) \propto \exp\left(-\frac{1}{2} z^\top \Sigma^{-1} z\right)$. Hence

$$\log \frac{\pi(z)}{\pi(z+m)} = -\frac{1}{2} z^\top \Sigma^{-1} z + \frac{1}{2}(z+m)^\top \Sigma^{-1}(z+m) = z^\top \Sigma^{-1} m + \frac{1}{2} m^\top \Sigma^{-1} m.$$

Therefore,

$$D_{\mathrm{KL}}(\zeta \mid \Pi) - D_{\mathrm{KL}}(\widetilde{\zeta} \mid \Pi) = \left(\int z \, \widetilde{\zeta}(\mathrm{d}z) \right)^\top \Sigma^{-1} m + \frac{1}{2} m^\top \Sigma^{-1} m = \frac{1}{2} m^\top \Sigma^{-1} m.$$

where the last equation is due to the fact that $\widetilde{\zeta}$ has mean zero.

For the Fisher information identity, recall that $\nabla \log \pi(z) = -\Sigma^{-1} z$. Since $\widetilde{f}(z) = f(z+m)$, we have

$$\nabla \log \frac{f(z+m)}{\pi(z+m)} = \nabla \log \widetilde{f}(z) - \nabla \log \pi(z+m) = \nabla \log \frac{\widetilde{f}(z)}{\pi(z)} + \Sigma^{-1} m.$$

Therefore, by the change of variables $x = z + m$,

$$\begin{aligned}
\mathrm{FI}_S(\zeta \mid \Pi) &= \int \left\| S^{1/2} \nabla \log \frac{f(x)}{\pi(x)} \right\|^2 \zeta(\mathrm{d}x) \\
&= \int \left\| S^{1/2} \left(\nabla \log \frac{\widetilde{f}(z)}{\pi(z)} + \Sigma^{-1} m \right) \right\|^2 \widetilde{\zeta}(\mathrm{d}z) \\
&= \mathrm{FI}_S(\widetilde{\zeta} \mid \Pi) + 2 m^\top \Sigma^{-1} S \int \nabla \log \frac{\widetilde{f}(z)}{\pi(z)} \, \widetilde{\zeta}(\mathrm{d}z) + m^\top \Sigma^{-1} S \Sigma^{-1} m.
\end{aligned}$$

It remains to show that the cross term vanishes. Indeed,

$$\begin{aligned}
\int \nabla \log \frac{\widetilde{f}(z)}{\pi(z)} \, \widetilde{\zeta}(\mathrm{d}z) &= \int \left(\nabla \log \widetilde{f}(z) - \nabla \log \pi(z) \right) \widetilde{f}(z) \, \mathrm{d}z \\
&= \int \frac{\nabla \widetilde{f}(z)}{\widetilde{f}(z)} \widetilde{f}(z) \, \mathrm{d}z - \int \nabla \log \pi(z) \, \widetilde{f}(z) \, \mathrm{d}z \\
&= \int \nabla \widetilde{f}(z) \, \mathrm{d}z - \int \nabla \log \pi(z) \, \widetilde{\zeta}(\mathrm{d}z) \\
&= - \int \nabla \log \pi(z) \, \widetilde{\zeta}(\mathrm{d}z) \\
&= \int \Sigma^{-1} z \, \widetilde{\zeta}(\mathrm{d}z) \\
&= \Sigma^{-1} \int z \, \widetilde{\zeta}(\mathrm{d}z) \\
&= 0.
\end{aligned}$$

where we used $\int \nabla \widetilde{f}(z)\,\mathrm{d}z = 0, \quad \nabla \log \pi(z) = -\Sigma^{-1}z$, and the fact that $\widetilde{\zeta}$ is centered. Consequently,

$$\mathrm{FI}_S(\zeta \mid \Pi) = \mathrm{FI}_S(\widetilde{\zeta} \mid \Pi) + m^\top \Sigma^{-1} S \Sigma^{-1} m,$$

which proves Eq. (58). Since $S \succ 0$, we have $m^\top \Sigma^{-1} S \Sigma^{-1} m \geq 0$, and hence

$$\mathrm{FI}_S(\widetilde{\zeta} \mid \Pi) \leq \mathrm{FI}_S(\zeta \mid \Pi).$$

□

Lemma 8 (Centered underdamped hypocoercivity)**.** *Fix $\gamma, \tau > 0$, and let $\mu_t = \mathrm{Law}(X_t, V_t)$ solve*

$$\begin{aligned} \mathrm{d}X_t &= V_t\,\mathrm{d}t, \\ \mathrm{d}V_t &= (-\gamma V_t - \tau X_t)\,\mathrm{d}t + \sqrt{2\gamma\tau}\,\mathrm{d}\mathcal{B}_t \end{aligned} \tag{59}$$

on $\mathbb{R}^d \times \mathbb{R}^d$, with μ_0 of finite second moment and finite relative entropy and relative Fisher information with respect to $\Pi_{\tau,d}$ defined below. Set

$$\Pi_{\tau,d} := \mathcal{N}(0, I_d) \otimes g_{\tau,d}, \qquad \Gamma := \frac{\gamma}{\sqrt{\tau}}.$$

Then $\Pi_{\tau,d}$ is the unique invariant law. Moreover, there exist $a_\Gamma, b_\Gamma, c_\Gamma, \kappa_\Gamma > 0$, depending only on Γ, such that $a_\Gamma c_\Gamma > b_\Gamma^2$, and with

$$\bar{S}_\Gamma := \begin{pmatrix} c_\Gamma & b_\Gamma \\ b_\Gamma & a_\Gamma \end{pmatrix} \otimes I_d, \qquad S_{\Gamma,\tau} := \begin{pmatrix} c_\Gamma & b_\Gamma\sqrt{\tau} \\ b_\Gamma\sqrt{\tau} & a_\Gamma\tau \end{pmatrix} \otimes I_d,$$

one has $\bar{S}_\Gamma \succ 0, \quad S_{\Gamma,\tau} \succ 0$. Then for every $t \geq 0$,

$$D_{\mathrm{KL}}(\mu_t \mid \Pi_{\tau,d}) \leq \left[D_{\mathrm{KL}}(\mu_0 \mid \Pi_{\tau,d}) + \mathrm{FI}_{S_{\Gamma,\tau}}(\mu_0 \mid \Pi_{\tau,d})\right] e^{-\lambda^{\mathrm{hyp}}_{\Gamma,\tau} t}. \tag{60}$$

where $\lambda^{\mathrm{hyp}}_{\Gamma,\tau} := \sqrt{\tau}\,\kappa_\Gamma \min\left\{1, \frac{1}{2\lambda_{\max}(\bar{S}_\Gamma)}\right\} > 0$.

Proof. The covariance of $\Pi_{\tau,d}$ solves the Lyapunov equation for this stable linear Langevin process, so $\Pi_{\tau,d}$ is its unique invariant law. Rescale $s = \sqrt{\tau}\,t$, $p = v/\sqrt{\tau}$, and $\Gamma = \gamma/\sqrt{\tau}$. The resulting process is

$$\mathrm{d}\widetilde{X}_s = P_s\,\mathrm{d}s, \qquad \mathrm{d}P_s = (-\Gamma P_s - \widetilde{X}_s)\,\mathrm{d}s + \sqrt{2\Gamma}\,\mathrm{d}\widetilde{\mathcal{B}}_s, \tag{61}$$

with standard Gaussian invariant law $\widetilde{\Pi} = \mathcal{N}(0, I_{2d})$ on $\mathbb{R}^{2d}$. Apply the direct entropy–Fisher computation in Sec. E.2 with normalized potential $\widetilde{U}(x) = \|x\|^2/2$, hence $\Theta = 1$ and $\bar{\alpha} = 1$. Take

$$\eta_\Gamma = \frac{\Gamma}{16(\Gamma^2+1)}, \quad a_\Gamma = 2\eta_\Gamma, \quad b_\Gamma = c_\Gamma = \eta_\Gamma, \quad \kappa_\Gamma = \frac{\Gamma}{16(\Gamma^2+1)}.$$

Then $a_\Gamma c_\Gamma - b_\Gamma^2 = \eta_\Gamma^2 > 0$ and Eq. (83) bounds the dissipation matrix below by $\kappa_\Gamma I$. If $\widetilde{m}_s$ denotes the rescaled law, this gives

$$D_{\mathrm{KL}}(\widetilde{m}_s \mid \widetilde{\Pi}) + \mathrm{FI}_{\bar{S}_\Gamma}(\widetilde{m}_s \mid \widetilde{\Pi}) \leq e^{-\kappa_\Gamma \min\{1, (2\lambda_{\max}(\bar{S}_\Gamma))^{-1}\} s} [D_{\mathrm{KL}}(\widetilde{m}_0 \mid \widetilde{\Pi}) + \mathrm{FI}_{\bar{S}_\Gamma}(\widetilde{m}_0 \mid \widetilde{\Pi})].$$

The change of coordinates preserves relative entropy and transforms the Fisher matrix into $S_{\Gamma,\tau}$. Therefore,

$$D_{\mathrm{KL}}(m_t \mid \Pi_{\tau,d}) + \mathrm{FI}_{S_{\Gamma,\tau}}(m_t \mid \Pi_{\tau,d}) \leq e^{-\lambda^{\mathrm{hyp}}_{\Gamma,\tau} t} \left[D_{\mathrm{KL}}(m_0 \mid \Pi_{\tau,d}) + \mathrm{FI}_{S_{\Gamma,\tau}}(m_0 \mid \Pi_{\tau,d})\right]. \tag{62}$$

Dropping the terminal Fisher information proves Eq. (60). □

D PROOF OF THM. 2

Proof. We first verify the assumptions and identify the equilibrium. Write $\phi_\epsilon(x) = \sin(\epsilon x)/\epsilon$. The interaction is affine in each law. Its first and mixed second functional derivatives are bounded by B/ϵ^2, and its pure second functional derivatives vanish. Its intrinsic x-derivative is $B\mathbb{E}_\sigma[\phi_\epsilon]\cos(\epsilon x)$, which is bounded by B/ϵ and B-Lipschitz in space and in the opposing law for W_1. The y-derivative is analogous. Thus all three assumptions hold. Denote by $\gamma^0 = \mathcal{N}(0,1)$, since ϕ_ϵ is odd function, both Gibbs responses equal to γ^0 at (γ^0, γ^0). By uniqueness, this is the equilibrium, with $\mu^\star = \nu^\star = \mathcal{N}(0,1) \otimes g_{\tau,1}$.

Since $B^2 > \tau\gamma^2$, the equation $\lambda^2 + \gamma\lambda + \tau - \mathrm{i}B = 0$ has a root $\lambda = \alpha + \mathrm{i}\omega$ with $\alpha > 0$. For $h > 0$, start from Gaussian laws obtained by translating the equilibrium position–velocity distributions:

$$\mu_0 = \mathcal{N}((h, \alpha h), \mathrm{diag}(1,\tau)), \qquad \nu_0 = \mathcal{N}((0, \omega h), \mathrm{diag}(1,\tau)).$$

Couple it with two stationary linear Langevin processes $(\bar{X}_t, \bar{V}_t)$ and $(\bar{Y}_t, \bar{W}_t)$ having laws $\mu^\star$ and $\nu^\star$, respectively, using the same Brownian noises and the deterministic initial shifts above. The Brownian noises then cancel in the differences.

Choose $C \geq 1$ and $c > 0$, depending only on (τ, γ), such that $\|e^{t\left(\begin{smallmatrix} 0 & 1 \\ -\tau & -\gamma \end{smallmatrix}\right)}\|_{\mathrm{op}} \leq Ce^{-ct}$. Variation of constants bounds the sum of the two L^∞ position–velocity displacements by the deterministic quantity

$$D(t) = C\big(\sqrt{1+\alpha^2} + |\omega|\big)he^{-ct} + CB\int_0^t e^{-c(t-s)}\big(|\mathbb{E}\phi_\epsilon(X_s)| + |\mathbb{E}\phi_\epsilon(Y_s)|\big)\,\mathrm{d}s.$$

Here the L^∞ norm is over the coupling probability space. While $D(t) \leq 1$, set $\Delta X_t = X_t - \bar{X}_t$. Since $\bar{X}_t$ is a stationary symmetric Gaussian process and ϕ_ϵ is odd, we have $\mathbb{E}\bar{X}_t = \mathbb{E}\phi_\epsilon(\bar{X}_t) = 0$, so

$$\begin{aligned}|\mathbb{E}\phi_\epsilon(X_t) - \mathbb{E}X_t| &= \left|\mathbb{E}\left[\Delta X_t \int_0^1 \big(\cos(\epsilon(\bar{X}_t + s\Delta X_t)) - 1\big)\,\mathrm{d}s\right]\right| \\ &\leq \frac{\epsilon^2}{2}\|\Delta X_t\|_\infty \mathbb{E}(|\bar{X}_t| + 1)^2 \leq 2\epsilon^2\|\Delta X_t\|_\infty.\end{aligned}$$

The same bound holds for Y. Also, $|\mathbb{E}\phi_\epsilon(X_t)| + |\mathbb{E}\phi_\epsilon(Y_t)| \leq D(t)$ and $|1 - \mathbb{E}\cos(\epsilon X_t)|, |1 - \mathbb{E}\cos(\epsilon Y_t)| \leq 2\epsilon^2$. Thus the complex mean $z(t) = \mathbb{E}X_t + \mathrm{i}\mathbb{E}Y_t$ satisfies

$$z'' + \gamma z' + (\tau - \mathrm{i}B)z = e(t), \qquad |e(t)| \leq 4B\epsilon^2 D(t).$$

Moreover, $D' \leq (-c + 2CB\epsilon^2)D + \sqrt{2}CB|z|$.

Set:

$$z_+ = \frac{z' + (\gamma+\lambda)z}{2\lambda + \gamma}, \qquad z_- = z - z_+.$$

Then $z_+' = \lambda z_+ + e/(2\lambda+\gamma)$ and $z_-' = -(\gamma+\lambda)z_- - e/(2\lambda+\gamma)$, with $z_+(0) = h$ and $z_-(0) = 0$. Choose

$$K > \max\left\{1,\ C(\sqrt{1+\alpha^2} + |\omega|),\ \frac{3\sqrt{2}CB}{c+\alpha}\right\}, \qquad 2CB\epsilon^2 \leq \frac{c}{2}, \qquad \frac{4BK\epsilon^2}{|2\lambda+\gamma|} \leq \frac{\alpha}{2}.$$

We next show that $|z_-| \leq |z_+|/2$ and $D \leq K|z_+|$ are preserved until. Initially both inequalities are strict. Inside this region, $D \leq 1/2$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}|z_+| \geq \frac{\alpha}{2}|z_+|, \qquad \frac{\mathrm{d}}{\mathrm{d}t}|z_-| \leq -(\gamma+\alpha)|z_-| + \frac{\alpha}{2}|z_+|, \qquad D' \leq -\frac{c}{2}D + \sqrt{2}CB(|z_+| + |z_-|).$$

At the respective boundaries these imply

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(|z_-| - |z_+|/2\big) \leq -\left(\frac{\gamma}{2} + \frac{\alpha}{4}\right)|z_+| < 0,$$

$$(D - K|z_+|)' \leq \frac{3\sqrt{2}CB - (c+\alpha)K}{2}|z_+| < 0.$$

Thus neither boundary can be crossed. For $0 < h < 1/(2K)$, the lower bound on the growth of $|z_+|$ yields a finite hitting time T. At that time,

$$W_2(\mu_T, \mu^\star) + W_2(\nu_T, \nu^\star) \geq |\mathbb{E}X_T| + |\mathbb{E}Y_T| \geq |z(T)| \geq |z_+(T)| - |z_-(T)| \geq \frac{1}{4K}.$$

In contrast, the initial sum of W_2 distances is $(\sqrt{1+\alpha^2} + |\omega|)h$, which can be made smaller than any $\eta > 0$. Taking $r = 1/(4K)$ proves the claim. □

E Proofs for the anchored method

E.1 Auxiliary lemma

Lemma 9 (Entropy sandwich; Lascu et al., 2025, Lemma 3.4)**.** *Under Asm. 1, provided the Gibbs responses and an MNE exist, the following inequalities hold:*

$$\tau D_{\mathrm{KL}}(\rho \mid \rho^\star) + \tau D_{\mathrm{KL}}(\sigma \mid \sigma^\star) \leq \mathrm{NI}_\tau(\rho, \sigma) \leq \tau D_{\mathrm{KL}}(\rho \mid \widehat{\rho}_{\rho,\sigma}) + \tau D_{\mathrm{KL}}(\sigma \mid \widehat{\sigma}_{\rho,\sigma}) \tag{63}$$

This is exactly Lascu et al. (2025, Lemma 3.4), with their regularization coefficient $\sigma^2/2$ replaced by τ. The upper bound is an equality when L is affine in each law.

Lemma 10. *Set $C_m := \max\{C_{xy}, C_{yx}\}$, under Asm. 2,*

$$\mathrm{TV}(\widehat{\rho}_{\rho,\sigma}, \widehat{\rho}_{\rho',\sigma'}) \leq \frac{C_{xx}\mathrm{TV}(\rho,\rho') + C_m\mathrm{TV}(\sigma,\sigma')}{\tau}, \tag{64}$$

$$\mathrm{TV}(\widehat{\sigma}_{\rho,\sigma}, \widehat{\sigma}_{\rho',\sigma'}) \leq \frac{C_m\mathrm{TV}(\rho,\rho') + C_{yy}\mathrm{TV}(\sigma,\sigma')}{\tau}. \tag{65}$$

In particular, valid constants for the sum metric are $\ell_x = (C_{xx} + C_m)/\tau$ and $\ell_y = (C_{yy} + C_m)/\tau$.

Proof. Let $f_0 = \delta L/\delta\rho(\rho, \sigma, \cdot)$, $f_1 = \delta L/\delta\rho(\rho', \sigma', \cdot)$, and $d = f_1 - f_0$. By chain rule and Asm. 2:

$$\|d\|_\infty \leq 2C_{xx}\mathrm{TV}(\rho,\rho') + 2C_m\mathrm{TV}(\sigma,\sigma').$$

Let $\pi_t \propto e^{-(f_0+td)/\tau}\gamma_x^0, \quad t \in [0,1]$. Then

$$\partial_t \pi_t = -\tau^{-1}(d - \mathbb{E}_{\pi_t} d)\pi_t.$$

Since $\mathbb{E}|d - \mathbb{E}d| \leq \sqrt{\mathrm{Var}(d)} \leq \|d\|_\infty$,

$$\mathrm{TV}(\pi_0, \pi_1) \leq \frac{1}{2\tau}\int_0^1 \mathbb{E}_{\pi_t}|d - \mathbb{E}_{\pi_t} d| \,\mathrm{d}t \leq \frac{\|d\|_\infty}{2\tau}.$$

The endpoints are the two desired Gibbs laws. The same calculation with the opposite sign in the exponential proves the y-player bound. □

Lemma 11. *Assume Asm. 2 and 3. Then the following bounds hold.*

First, for every (ρ, σ),

$$\left\|\log \frac{\mathrm{d}\widehat{\rho}_{\rho,\sigma}}{\mathrm{d}\gamma_x^0}\right\|_\infty \leq \frac{2C_x}{\tau}, \qquad \left\|\log \frac{\mathrm{d}\widehat{\sigma}_{\rho,\sigma}}{\mathrm{d}\gamma_y^0}\right\|_\infty \leq \frac{2C_y}{\tau}. \tag{66}$$

Second, for every (ρ, σ) and every $T \geq 0$,

$$\left\|\log \frac{\mathrm{d}\mathcal{R}_x^T(\rho,\sigma)}{\mathrm{d}\widehat{\rho}_{\rho,\sigma}}\right\|_\infty \leq \frac{2C_x}{\tau}, \qquad \left\|\log \frac{\mathrm{d}\mathcal{R}_y^T(\rho,\sigma)}{\mathrm{d}\widehat{\sigma}_{\rho,\sigma}}\right\|_\infty \leq \frac{2C_y}{\tau}. \tag{67}$$

Consequently, for every $k \geq 0$,

$$\left\|\log \frac{\mathrm{d}\rho_k}{\mathrm{d}\gamma_x^0}\right\|_\infty \leq \frac{4C_x}{\tau}, \qquad \left\|\log \frac{\mathrm{d}\sigma_k}{\mathrm{d}\gamma_y^0}\right\|_\infty \leq \frac{4C_y}{\tau}. \tag{68}$$

Moreover, for $s \in [0,1]$, define

$$\rho(s) = (1-\beta s)\rho_k + \beta s\, \mathcal{R}_x^T(\rho_k, \sigma_k), \qquad \sigma(s) = (1-\beta s)\sigma_k + \beta s\, \mathcal{R}_y^T(\rho_k, \sigma_k),$$
$$\widehat{\rho}(s) := \widehat{\rho}_{\rho(s),\sigma(s)}, \qquad \widehat{\sigma}(s) := \widehat{\sigma}_{\rho(s),\sigma(s)}.$$

Then

$$\left\| \log \frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)} \right\|_\infty \le R_x := \frac{6C_x}{\tau}, \qquad \left\| \log \frac{\mathrm{d}\sigma(s)}{\mathrm{d}\widehat{\sigma}(s)} \right\|_\infty \le R_y := \frac{6C_y}{\tau}. \tag{69}$$

Proof. We prove the assertions for the x-player. The proof for the y-player is identical.

By Eq. (2),

$$\frac{\mathrm{d}\widehat{\rho}_{\rho,\sigma}}{\mathrm{d}\gamma_x^0}(x) = Z_x(\rho,\sigma)^{-1} \exp\left[-\frac{1}{\tau}\frac{\delta L}{\delta \rho}(\rho,\sigma,x)\right],$$

where

$$Z_x(\rho,\sigma) = \int_X \exp\left[-\frac{1}{\tau}\frac{\delta L}{\delta \rho}(\rho,\sigma,x)\right] \gamma_x^0(\mathrm{d}x).$$

Since $\left\| \frac{\delta L}{\delta \rho}(\rho,\sigma,\cdot) \right\|_\infty \le C_x$, we have $e^{-C_x/\tau} \le Z_x(\rho,\sigma) \le e^{C_x/\tau}$. It follows that $e^{-2C_x/\tau} \le \frac{\mathrm{d}\widehat{\rho}_{\rho,\sigma}}{\mathrm{d}\gamma_x^0}(x) \le e^{2C_x/\tau}$, and hence $\left\| \log \frac{\mathrm{d}\widehat{\rho}_{\rho,\sigma}}{\mathrm{d}\gamma_x^0} \right\|_\infty \le \frac{2C_x}{\tau}$.

Fix (ρ,σ) and abbreviate

$$U(x) := \frac{\delta L}{\delta \rho}(\rho,\sigma,x) + \frac{\tau}{2}\|x\|^2, \qquad \widehat{\rho} := \widehat{\rho}_{\rho,\sigma}, \qquad \Pi := \Pi_x^{\rho,\sigma} = \widehat{\rho} \otimes g_{\tau,d_x}$$

The Fokker–Planck equation of Eq. (13) is

$$\partial_t \mu_t^{\rho,\sigma} = -\nabla_x \cdot (v \mu_t^{\rho,\sigma}) + \nabla_v \cdot \left((\gamma_x v + \nabla U(x)) \mu_t^{\rho,\sigma} \right) + \gamma_x \tau \Delta_v \mu_t^{\rho,\sigma}. \tag{70}$$

Let $h_t := \frac{\mu_t^{\rho,\sigma}}{\Pi}$. Since

$$\nabla_x \log \Pi(x,v) = -\frac{1}{\tau}\nabla U(x), \qquad \nabla_v \log \Pi(x,v) = -\frac{1}{\tau} v,$$

we have

$$\nabla_x \Pi = -\frac{1}{\tau}\nabla U(x)\Pi, \qquad \nabla_v \Pi = -\frac{1}{\tau} v \Pi, \qquad \Delta_v \Pi = \left(-\frac{d_x}{\tau} + \frac{\|v\|^2}{\tau^2} \right) \Pi.$$

Substituting $\mu_t^{\rho,\sigma} = h_t \Pi$ into Eq. (70) and applying the product rule gives

$$\begin{aligned} \Pi\, \partial_t h_t = &-v \cdot \nabla_x h_t\, \Pi - h_t\, v \cdot \nabla_x \Pi \\ &+ (\gamma_x v + \nabla U(x)) \cdot \nabla_v h_t\, \Pi + h_t\, \nabla_v \cdot \left((\gamma_x v + \nabla U(x)) \Pi \right) \\ &+ \gamma_x \tau \left(\Pi \Delta_v h_t + 2 \nabla_v h_t \cdot \nabla_v \Pi + h_t \Delta_v \Pi \right). \end{aligned} \tag{71}$$

Since

$$\nabla_v \cdot \left((\gamma_x v + \nabla U(x)) \Pi \right) = \gamma_x d_x \Pi - \frac{\gamma_x}{\tau}\|v\|^2 \Pi - \frac{1}{\tau}\nabla U(x) \cdot v\, \Pi,$$

all zeroth-order terms in h_t cancel. Moreover, since $2\gamma_x \tau \nabla_v h_t \cdot \nabla_v \Pi = -2\gamma_x v \cdot \nabla_v h_t\, \Pi$, we get

$$\begin{aligned} \Pi\, \partial_t h_t &= -v \cdot \nabla_x h_t\, \Pi + (\gamma_x v + \nabla U(x)) \cdot \nabla_v h_t\, \Pi - 2\gamma_x v \cdot \nabla_v h_t\, \Pi + \gamma_x \tau \Delta_v h_t\, \Pi \\ &= \Pi \left[-v \cdot \nabla_x h_t + \left(\nabla U(x) - \gamma_x v \right) \cdot \nabla_v h_t + \gamma_x \tau \Delta_v h_t \right]. \end{aligned} \tag{72}$$

Since $\Pi > 0$, we conclude that

$$\partial_t h_t = -v \cdot \nabla_x h_t + \left(\nabla U(x) - \gamma_x v \right) \cdot \nabla_v h_t + \gamma_x \tau \Delta_v h_t. \tag{73}$$

We now identify Eq. (73) as a Kolmogorov backward equation.

Consider the reversed underdamped dynamics

$$\begin{aligned} \mathrm{d}\bar{X}_t &= -\bar{V}_t\,\mathrm{d}t, \\ \mathrm{d}\bar{V}_t &= \big(\nabla U(\bar{X}_t) - \gamma_x \bar{V}_t\big)\mathrm{d}t + \sqrt{2\gamma_x\tau}\,\mathrm{d}\bar{\mathcal{B}}_t. \end{aligned} \tag{74}$$

Its infinitesimal generator is

$$\mathcal{L}_x^{\mathrm{rev}} f(x,v) := -v\cdot\nabla_x f(x,v) + \big(\nabla U(x) - \gamma_x v\big)\cdot\nabla_v f(x,v) + \gamma_x\tau\Delta_v f(x,v).$$

Let its Markov semigroup be

$$P_t^{\mathrm{rev}} f(x,v) := \mathbb{E}_{x,v}\big[f(\bar{X}_t,\bar{V}_t)\big].$$

Kolmogorov's backward equation gives

$$\partial_t P_t^{\mathrm{rev}} f = \mathcal{L}_x^{\mathrm{rev}} P_t^{\mathrm{rev}} f, \qquad P_0^{\mathrm{rev}} f = f.$$

We also have

$$h_t(x,v) = P_t^{\mathrm{rev}} h_0(x,v) = \mathbb{E}_{x,v}\big[h_0(\bar{X}_t,\bar{V}_t)\big]. \tag{75}$$

At time zero:

$$h_0(x,v) = \frac{\mathrm{d}\mu_0^{\rho,\sigma}}{\mathrm{d}\Pi}(x,v) = \frac{\mathrm{d}(\gamma_x^0\otimes g_{\tau,d_x})}{\mathrm{d}(\widehat{\rho}\otimes g_{\tau,d_x})}(x,v) = \frac{\mathrm{d}\gamma_x^0}{\mathrm{d}\widehat{\rho}}(x).$$

Recall in the first part of the proof we have shown $e^{-2C_x/\tau} \le h_0(x,v) \le e^{2C_x/\tau}$. Hence Eq. (75) yields

$$e^{-2C_x/\tau} \le h_t(x,v) \le e^{2C_x/\tau}. \tag{76}$$

Because $\mu_t^{\rho,\sigma}(\mathrm{d}x\,\mathrm{d}v) = h_t(x,v)\widehat{\rho}(\mathrm{d}x)g_{\tau,d_x}(\mathrm{d}v)$, the spatial marginal satisfies

$$\frac{\mathrm{d}\mathcal{R}_x^t(\rho,\sigma)}{\mathrm{d}\widehat{\rho}}(x) = \int_{\mathbb{R}^{d_x}} h_t(x,v)\, g_{\tau,d_x}(\mathrm{d}v).$$

Integrating Eq. (76) with respect to g_{τ,d_x} gives

$$e^{-2C_x/\tau} \le \frac{\mathrm{d}\mathcal{R}_x^t(\rho,\sigma)}{\mathrm{d}\widehat{\rho}}(x) \le e^{2C_x/\tau}.$$

Taking $t=T$ proves $\left\|\log\frac{\mathrm{d}\mathcal{R}_x^T(\rho,\sigma)}{\mathrm{d}\widehat{\rho}_{\rho,\sigma}}\right\|_\infty \le \frac{2C_x}{\tau}$.

Combining the two density bounds also gives (Eq. (67)), i.e.,

$$e^{-4C_x/\tau} \le \frac{\mathrm{d}\mathcal{R}_x^T(\rho,\sigma)}{\mathrm{d}\gamma_x^0} \le e^{4C_x/\tau}.$$

Equivalently, we have $e^{-4C_x/\tau}\gamma_x^0 \le \mathcal{R}_x^T(\rho_k,\sigma_k) \le e^{4C_x/\tau}\gamma_x^0$. The initialization $\rho_0 = \gamma_x^0$ satisfies the same two-sided bound. Since convex combinations preserve two-sided measure inequalities, the outer update implies inductively that

$$e^{-4C_x/\tau}\gamma_x^0 \le \rho_k \le e^{4C_x/\tau}\gamma_x^0, \qquad k\ge 0.$$

Consequently, we prove the third claim (Eq. (68)), i.e., $\left\|\log\frac{\mathrm{d}\rho_k}{\mathrm{d}\gamma_x^0}\right\|_\infty \le \frac{4C_x}{\tau}$.

For $s\in[0,1]$, the coefficients $1-\beta s$ and βs are nonnegative and sum to one. Hence the interpolation

$$\rho(s) = (1-\beta s)\rho_k + \beta s\,\mathcal{R}_x^T(\rho_k,\sigma_k)$$

also satisfies

$$e^{-4C_x/\tau}\gamma_x^0 \le \rho(s) \le e^{4C_x/\tau}\gamma_x^0.$$

On the other hand, the first step applied to the pair $(\rho(s),\sigma(s))$ gives

$$e^{-2C_x/\tau}\gamma_x^0 \le \widehat{\rho}(s) \le e^{2C_x/\tau}\gamma_x^0.$$

Dividing these two-sided density bounds yields

$$e^{-6C_x/\tau} \le \frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)} \le e^{6C_x/\tau},$$

and therefore $\left\|\log\frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)}\right\|_\infty \le \frac{6C_x}{\tau}$. The y-player proof is identical. □

E.2 Proof of Lemma 5

For an explicit x-player matrix, set $\Gamma_x = \gamma_x/\sqrt{\tau}$, $\Theta_x = (L_x + \tau)/\tau$, and $\eta_x = \Gamma_x\sqrt{\Theta_x}/[16(\Gamma_x^2 + \Theta_x)]$. Then

$$S_x = \eta_x \begin{pmatrix} \Theta_x^{-1} & \sqrt{\tau/\Theta_x} \\ \sqrt{\tau/\Theta_x} & 2\tau \end{pmatrix} \otimes I_{d_x}. \tag{77}$$

The y-player matrix is defined identically with x replaced by y.

Proof. We prove the x-player statement, defining $U(x) := \frac{\delta L}{\delta \rho}(\rho, \sigma, x) + \frac{\tau}{2}\|x\|^2$ and writing $\gamma = \gamma_x$ and $\Pi = \Pi_x^{\rho,\sigma}$. This is the fixed-target specialization of the entropy–Fisher method used by Fu & Wilson (2024, Theorem 3.1) while we extend it for arbitrary positive friction. Set

$$s = \sqrt{\tau}\, t, \qquad p = v/\sqrt{\tau}, \qquad \widetilde{U} = U/\tau, \qquad \Gamma = \gamma/\sqrt{\tau}, \qquad \Theta = (L_x + \tau)/\tau \geq 1.$$

Brownian scaling gives

$$\begin{aligned} \mathrm{d}\widetilde{X}_s &= \widetilde{P}_s\, \mathrm{d}s, \\ \mathrm{d}\widetilde{P}_s &= -\big(\Gamma \widetilde{P}_s + \nabla \widetilde{U}(\widetilde{X}_s)\big)\, \mathrm{d}s + \sqrt{2\Gamma}\, \mathrm{d}\widetilde{\mathcal{B}}_s. \end{aligned}$$

Its invariant law is $\widetilde{\Pi} = \widehat{\rho}_{\rho,\sigma} \otimes g_{1,d_x}$. By Asm. 3, $\nabla\widetilde{U}$ is globally Θ-Lipschitz. Its weak Jacobian is therefore an almost-everywhere defined symmetric Hessian, with $\|\nabla^2\widetilde{U}(x)\|_{\mathrm{op}} \leq \Theta$, for Lebesgue-almost every x. Tensorization provides an LSI with constant $\bar{\alpha} = \min\{\alpha_x, 1\}$. By the approximation argument of Chen et al. (2024, Theorem 2.1) and Fu & Wilson (2024, Theorem 3.1), it suffices to consider smooth potentials and positive smooth densities.

Choose

$$\eta := \frac{\Gamma\sqrt{\Theta}}{16(\Gamma^2 + \Theta)}, \qquad a := 2\eta, \qquad b := \frac{\eta}{\sqrt{\Theta}}, \qquad c := \frac{\eta}{\Theta}, \tag{78}$$

and let $\bar{S} = \begin{pmatrix} c & b \\ b & a \end{pmatrix} \otimes I_{d_x}$. Then $ac - b^2 = \eta^2/\Theta > 0$. Since $\Gamma^2 + \Theta \geq 2\Gamma\sqrt{\Theta}$, we also have

$$\eta \leq \frac{1}{32}, \qquad \lambda_{\max}(\bar{S}) \leq \eta(2 + \Theta^{-1}) \leq \frac{3}{32}. \tag{79}$$

Write $\widetilde{\mu}_s^{\rho,\sigma} = \mathrm{Law}(\widetilde{X}_s, \widetilde{P}_s)$, $u_s = \log(\mathrm{d}\widetilde{\mu}_s^{\rho,\sigma}/\mathrm{d}\widetilde{\Pi})$, and $g_s = \nabla_{x,p} u_s$. We consider

$$\widetilde{\mathcal{H}}(s) := D_{\mathrm{KL}}(\widetilde{\mu}_s^{\rho,\sigma} \mid \widetilde{\Pi}) + \mathrm{FI}_{\bar{S}}(\widetilde{\mu}_s^{\rho,\sigma} \mid \widetilde{\Pi}).$$

The density ratio evolves under

$$\mathcal{L}^{\mathrm{rev}} = -p \cdot \nabla_x + (\nabla\widetilde{U} - \Gamma p) \cdot \nabla_p + \Gamma\Delta_p.$$

In particular, $\partial_s u_s = \mathcal{L}^{\mathrm{rev}} u_s + \Gamma\|\nabla_p u_s\|^2$. Differentiating the entropy and modified Fisher information and integrating by parts gives

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}s} D_{\mathrm{KL}}(\widetilde{\mu}_s^{\rho,\sigma} \mid \widetilde{\Pi}) = -\Gamma \int \|\nabla_p u_s\|^2\, \mathrm{d}\widetilde{\mu}_s^{\rho,\sigma}, \\ &\frac{\mathrm{d}}{\mathrm{d}s} \mathrm{FI}_{\bar{S}}(\widetilde{\mu}_s^{\rho,\sigma} \mid \widetilde{\Pi}) = 2\int g_s^{\mathsf{T}} \bar{S} \begin{pmatrix} 0 & \nabla^2\widetilde{U} \\ -I & -\Gamma I \end{pmatrix} g_s\, \mathrm{d}\widetilde{\mu}_s^{\rho,\sigma} - 2\Gamma \sum_{j=1}^{d_x} \int (\partial_{p_j} g_s)^{\mathsf{T}} \bar{S} (\partial_{p_j} g_s)\, \mathrm{d}\widetilde{\mu}_s^{\rho,\sigma}. \end{aligned} \tag{80}$$

Here the matrix in the first integral is the transpose Jacobian of the drift in $\mathcal{L}^{\mathrm{rev}}$. Consequently,

$$\widetilde{\mathcal{H}}'(s) \leq -\int g_s^{\mathsf{T}} \mathsf{K}(x) g_s\, \mathrm{d}\widetilde{\mu}_s^{\rho,\sigma}, \tag{81}$$

where, writing $H(x) = \nabla^2\widetilde{U}(x)$ for the Hessian,

$$\mathsf{K}(x) = \begin{pmatrix} 2bI & (a + \Gamma b)I - cH(x) \\ (a + \Gamma b)I - cH(x) & (\Gamma + 2\Gamma a)I - 2bH(x) \end{pmatrix}. \tag{82}$$

The last term in Eq. (80) is nonpositive because $\bar{S} \succ 0$.

We now prove a uniform lower bound for $\mathsf{K}(x)$. Let $r = \Gamma/\sqrt{\Theta}$. For any $\xi, \zeta \in \mathbb{R}^{d_x}$,

$$\begin{aligned}\begin{pmatrix}\xi\\ \zeta\end{pmatrix}^{\mathsf{T}} \mathsf{K}(x) \begin{pmatrix}\xi\\ \zeta\end{pmatrix} &\geq \frac{2\eta}{\sqrt{\Theta}}\|\xi\|^2 - 2\eta(3+r)\|\xi\|\|\zeta\| \\ &\quad + (\Gamma + 4\Gamma\eta - 2\eta\sqrt{\Theta})\|\zeta\|^2 \\ &\geq \frac{\eta}{\sqrt{\Theta}}\|\xi\|^2 + \big[\Gamma - \eta\sqrt{\Theta}(r^2+2r+11)\big]\|\zeta\|^2.\end{aligned}$$

The first inequality uses $\|H(x)\|_{\mathrm{op}} \leq \Theta$. The second follows from Young's inequality. Since

$$r^2 + 2r + 11 \leq 12(r^2+1), \qquad \eta\sqrt{\Theta} = \frac{\Gamma}{16(r^2+1)},$$

the coefficient of $\|\zeta\|^2$ is at least $\Gamma/4$. Furthermore, $\eta/\sqrt{\Theta} = \Gamma/[16(\Gamma^2+\Theta)] \leq \Gamma/4$. Thus

$$\mathsf{K}(x) \succeq \kappa I_{2d_x}, \qquad \kappa := \frac{\Gamma}{16(\Gamma^2+\Theta)} > 0. \tag{83}$$

Combining Eq. (81) with the LSI and $\mathrm{FI}_{\bar{S}} \leq \lambda_{\max}(\bar{S})\mathrm{FI}$ gives

$$\begin{aligned}\widetilde{\mathcal{H}}'(s) &\leq -\kappa \mathrm{FI}(\widetilde{\mu}_s^{\rho,\sigma} \mid \widetilde{\Pi}) \\ &\leq -\kappa\bar{\alpha} D_{\mathrm{KL}}(\widetilde{\mu}_s^{\rho,\sigma} \mid \widetilde{\Pi}) - \frac{\kappa}{2\lambda_{\max}(\bar{S})}\mathrm{FI}_{\bar{S}}(\widetilde{\mu}_s^{\rho,\sigma} \mid \widetilde{\Pi}) \\ &\leq -\kappa \min\left\{\bar{\alpha}, \frac{1}{2\lambda_{\max}(\bar{S})}\right\}\widetilde{\mathcal{H}}(s).\end{aligned}$$

Gronwall's inequality yields exponential decay of $\widetilde{\mathcal{H}}$ with the exponent given above.

To return to the original coordinates, set

$$T_\tau(x,v) = (x, v/\sqrt{\tau}), \qquad M_\tau = \mathrm{diag}(I_{d_x}, \sqrt{\tau} I_{d_x}), \qquad S = M_\tau^{\mathsf{T}} \bar{S} M_\tau.$$

Then $\widetilde{\mu}^{\rho,\sigma}_{\sqrt{\tau}t} = (T_\tau)_{\#}\mu_t^{\rho,\sigma}$ and $\widetilde{\Pi} = (T_\tau)_{\#}\Pi$. Relative entropy is invariant under this invertible transformation, and the chain rule gives

$$\mathrm{FI}_{\bar{S}}(\widetilde{\mu}^{\rho,\sigma}_{\sqrt{\tau}t} \mid \widetilde{\Pi}) = \mathrm{FI}_S(\mu_t^{\rho,\sigma} \mid \Pi).$$

Hence

$$D_{\mathrm{KL}}(\mu_t^{\rho,\sigma} \mid \Pi) + \mathrm{FI}_S(\mu_t^{\rho,\sigma} \mid \Pi) \leq e^{-\lambda_{\mathrm{hyp},x} t}\big[D_{\mathrm{KL}}(\mu_0^{\rho,\sigma} \mid \Pi) + \mathrm{FI}_S(\mu_0^{\rho,\sigma} \mid \Pi)\big], \tag{84}$$

where $\lambda_{\mathrm{hyp},x} = \sqrt{\tau}\,\kappa \min\left\{\bar{\alpha}, \frac{1}{2\lambda_{\max}(\bar{S})}\right\}$. Dropping the terminal Fisher-information term proves the lemma. For the present coefficient choice, Eq. (79) implies that the minimum equals $\bar{\alpha}$, so the rate can also be written as

$$\lambda_{\mathrm{hyp},x} = \frac{\tau\bar{\alpha}\gamma_x}{16(\gamma_x^2 + L_x + \tau)}. \tag{85}$$

The y-player proof is identical. □

E.3 Proof of Thm. 3

Proof. Our outer-loop analysis follows the anchored best-response framework introduced by Kim et al. (2024), and the proof is tailored to our underdamped finite-time responses (Lemma 5).

Write $G_k = G(\rho_k, \sigma_k)$ and fix an outer iteration k. The argument applies to any $\beta_k \in (0,1]$ and $T_k \geq 0$. We suppress these subscripts and write

$$\widehat{\rho}_k := \widehat{\rho}_{\rho_k,\sigma_k}, \qquad \widehat{\sigma}_k := \widehat{\sigma}_{\rho_k,\sigma_k}.$$

Let μ_t^k be the frozen x-position–velocity law started from $\mu_0^k = \gamma_x^0 \otimes g_{\tau,d_x}$ with frozen pair (ρ_k, σ_k), and let $\Pi_k^x := \Pi_x^{\rho_k,\sigma_k}$. By Lemma 5,

$$D_{\mathrm{KL}}(\mu_T^k \mid \Pi_k^x) \leq \big[D_{\mathrm{KL}}(\mu_0^k \mid \Pi_k^x) + \mathrm{FI}_{S_x}(\mu_0^k \mid \Pi_k^x)\big]\, e^{-\lambda_{\mathrm{hyp},x} T}. \tag{86}$$

By Eq. (66), $\|\log(\mathrm{d}\gamma_x^0/\mathrm{d}\widehat{\rho}_k)\|_\infty \le 2C_x/\tau$. Therefore

$$D_{\mathrm{KL}}(\mu_0^k \mid \Pi_k^x) = D_{\mathrm{KL}}(\gamma_x^0 \mid \widehat{\rho}_k) \le \frac{2C_x}{\tau}. \tag{87}$$

Also,

$$\nabla_x \log \frac{\mathrm{d}\gamma_x^0}{\mathrm{d}\widehat{\rho}_k}(x) = \tau^{-1}\nabla_x \frac{\delta L}{\delta \rho}(\rho_k, \sigma_k, x), \qquad \nabla_v \log \frac{\mathrm{d}g_{\tau,d_x}}{\mathrm{d}g_{\tau,d_x}} = 0,$$

and Asm. 3 gives

$$\mathrm{FI}_{S_x}(\mu_0^k \mid \Pi_k^x) \le \lambda_{\max}(S_x)\frac{G_x^2}{\tau^2}. \tag{88}$$

Combining Eq. (86)–Eq. (88) and taking the position marginal, the data processing inequality gives:

$$D_{\mathrm{KL}}(\mathcal{R}_x^T(\rho_k, \sigma_k) \mid \widehat{\rho}_k) \le \left(\frac{2C_x}{\tau} + \lambda_{\max}(S_x)\frac{G_x^2}{\tau^2}\right) e^{-\lambda_{\mathrm{hyp},x}T}. \tag{89}$$

The same argument for the y sampler gives

$$D_{\mathrm{KL}}(\mathcal{R}_y^T(\rho_k, \sigma_k) \mid \widehat{\sigma}_k) \le \left(\frac{2C_y}{\tau} + \lambda_{\max}(S_y)\frac{G_y^2}{\tau^2}\right) e^{-\lambda_{\mathrm{hyp},y}T}. \tag{90}$$

With C_{kin} and λ_{kin} as stated in the theorem, we have:

$$D_{\mathrm{KL}}(\mathcal{R}_x^T(\rho_k, \sigma_k) \mid \widehat{\rho}_k) + D_{\mathrm{KL}}(\mathcal{R}_y^T(\rho_k, \sigma_k) \mid \widehat{\sigma}_k) \le 2C_{\mathrm{kin}}e^{-\lambda_{\mathrm{kin}}T}. \tag{91}$$

Define

$$\eta_k := \mathrm{TV}(\mathcal{R}_x^T(\rho_k, \sigma_k), \widehat{\rho}_k) + \mathrm{TV}(\mathcal{R}_y^T(\rho_k, \sigma_k), \widehat{\sigma}_k).$$

Pinsker's inequality and Cauchy–Schwarz imply

$$\eta_k \le \sqrt{2C_{\mathrm{kin}}}e^{-\lambda_{\mathrm{kin}}T/2} =: \eta_T. \tag{92}$$

For signed measures below, we use $\|\theta\|_{\mathrm{TV}} := \sup_A |\theta(A)|$ when θ has total mass zero. In particular $\|\mu - \nu\|_{\mathrm{TV}} = \mathrm{TV}(\mu, \nu)$ for probability measures μ, ν. For $s \in [0, 1]$, set

$$\rho(s) := (1 - \beta s)\rho_k + \beta s\mathcal{R}_x^T(\rho_k, \sigma_k), \qquad \sigma(s) := (1 - \beta s)\sigma_k + \beta s\mathcal{R}_y^T(\rho_k, \sigma_k). \tag{93}$$

Then $(\rho(0), \sigma(0)) = (\rho_k, \sigma_k)$ and $(\rho(1), \sigma(1)) = (\rho_{k+1}, \sigma_{k+1})$. Put

$$\widehat{\rho}(s) := \widehat{\rho}_{\rho(s),\sigma(s)}, \qquad \widehat{\sigma}(s) := \widehat{\sigma}_{\rho(s),\sigma(s)}.$$

Lemma 10 gives

$$\mathrm{TV}(\widehat{\rho}(s), \widehat{\rho}_k) \le \ell_x\{\mathrm{TV}(\rho(s), \rho_k) + \mathrm{TV}(\sigma(s), \sigma_k)\} \le 2\beta s\ell_x, \tag{94}$$

$$\mathrm{TV}(\widehat{\sigma}(s), \widehat{\sigma}_k) \le 2\beta s\ell_y. \tag{95}$$

Let

$$\Delta_\rho(s) := \widehat{\rho}(s) - \rho(s), \qquad \Delta_\sigma(s) := \widehat{\sigma}(s) - \sigma(s).$$

Since $\dot{\rho}(s) = \beta(\mathcal{R}_x^T(\rho_k, \sigma_k) - \rho_k)$ and $\dot{\sigma}(s) = \beta(\mathcal{R}_y^T(\rho_k, \sigma_k) - \sigma_k)$,

$$\dot{\rho}(s) = \beta\Delta_\rho(s) + \beta r_x(s), \qquad \dot{\sigma}(s) = \beta\Delta_\sigma(s) + \beta r_y(s), \tag{96}$$

where

$$r_x(s) := \mathcal{R}_x^T(\rho_k, \sigma_k) - \rho_k - \widehat{\rho}(s) + \rho(s), \qquad r_y(s) := \mathcal{R}_y^T(\rho_k, \sigma_k) - \sigma_k - \widehat{\sigma}(s) + \sigma(s).$$

Using Eq. (94) and $\mathrm{TV}(\rho(s), \rho_k) \le \beta s\mathrm{TV}(\mathcal{R}_x^T(\rho_k, \sigma_k), \rho_k) \le \beta s$, we get

$$\|r_x(s)\|_{\mathrm{TV}} \le \mathrm{TV}(\mathcal{R}_x^T(\rho_k, \sigma_k), \widehat{\rho}_k) + \mathrm{TV}(\widehat{\rho}_k, \widehat{\rho}(s)) + \mathrm{TV}(\rho(s), \rho_k) \le \eta_k + \beta s(1 + 2\ell_x). \tag{97}$$

Similarly,

$$\|r_y(s)\|_{\mathrm{TV}} \le \eta_k + \beta s(1 + 2\ell_y). \tag{98}$$

For the x term,

$$\frac{\mathrm{d}}{\mathrm{d}s}D_{\mathrm{KL}}(\rho(s) \mid \widehat{\rho}(s)) = \int_X \log\left(\frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)}\right)\dot{\rho}(s)(\mathrm{d}x) - \int_X \partial_s \log\widehat{\rho}(s, x)\rho(s)(\mathrm{d}x). \tag{99}$$

Using Eq. (96),

$$\begin{aligned}
&\int_X \log\left(\frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)}\right)\dot{\rho}(s)(\mathrm{d}x)\\
&\quad=\beta\int_X \log\left(\frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)}\right)(\widehat{\rho}(s)-\rho(s))(\mathrm{d}x)+\beta\int_X \log\left(\frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)}\right)r_x(s)(\mathrm{d}x)\\
&\quad=-\beta\left\{D_{\mathrm{KL}}(\rho(s)\mid\widehat{\rho}(s))+D_{\mathrm{KL}}(\widehat{\rho}(s)\mid\rho(s))\right\}+\beta\int_X \log\left(\frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)}\right)r_x(s)(\mathrm{d}x). \qquad (100)
\end{aligned}$$

The envelope Eq. (69) and Eq. (97) imply

$$\int_X \log\left(\frac{\mathrm{d}\rho(s)}{\mathrm{d}\widehat{\rho}(s)}\right)\dot{\rho}(s)(\mathrm{d}x)\le-\beta D_{\mathrm{KL}}(\rho(s)\mid\widehat{\rho}(s))+C\beta(\eta_k+\beta s), \qquad (101)$$

where C depends only on C_x, τ and ℓ_x. The same calculation gives

$$\int_Y \log\left(\frac{\mathrm{d}\sigma(s)}{\mathrm{d}\widehat{\sigma}(s)}\right)\dot{\sigma}(s)(\mathrm{d}y)\le-\beta D_{\mathrm{KL}}(\sigma(s)\mid\widehat{\sigma}(s))+C\beta(\eta_k+\beta s), \qquad (102)$$

with C depending only on C_y, τ and ℓ_y.

Since

$$\log\widehat{\rho}(s,x)=-\tau^{-1}\frac{\delta L}{\delta\rho}(\rho(s),\sigma(s),x)+\log\gamma_x^0(x)-\log Z_x(s)$$

and $\int_X \partial_s\log\widehat{\rho}(s,x)\widehat{\rho}(s)(\mathrm{d}x)=0$, the constant $\partial_s\log Z_x(s)$ drops against the zero-mass signed measure $\rho(s)-\widehat{\rho}(s)$ and

$$-\int_X \partial_s\log\widehat{\rho}(s,x)\rho(s)(\mathrm{d}x)=\tau^{-1}\int_X \partial_s\frac{\delta L}{\delta\rho}(\rho(s),\sigma(s),x)(\rho(s)-\widehat{\rho}(s))(\mathrm{d}x). \qquad (103)$$

Likewise,

$$-\int_Y \partial_s\log\widehat{\sigma}(s,y)\sigma(s)(\mathrm{d}y)=-\tau^{-1}\int_Y \partial_s\frac{\delta L}{\delta\sigma}(\rho(s),\sigma(s),y)(\sigma(s)-\widehat{\sigma}(s))(\mathrm{d}y). \qquad (104)$$

For signed measures θ,θ' on X and ξ,ξ' on Y, define

$$\begin{aligned}
\mathsf{D}^2_{\rho\rho,s}L[\theta,\theta']&:=\int_X\int_X\frac{\delta^2 L}{\delta\rho^2}(\rho(s),\sigma(s),x,x')\theta(\mathrm{d}x)\theta'(\mathrm{d}x'),\\
\mathsf{D}^2_{\rho\sigma,s}L[\theta,\xi]&:=\int_X\int_Y\frac{\delta^2 L}{\delta\rho\delta\sigma}(\rho(s),\sigma(s),x,y)\theta(\mathrm{d}x)\xi(\mathrm{d}y),\\
\mathsf{D}^2_{\sigma\rho,s}L[\xi,\theta]&:=\int_Y\int_X\frac{\delta^2 L}{\delta\sigma\delta\rho}(\rho(s),\sigma(s),y,x)\xi(\mathrm{d}y)\theta(\mathrm{d}x),
\end{aligned}$$

and define $\mathsf{D}^2_{\sigma\sigma,s}L$ analogously. The chain rule gives

$$\begin{aligned}
&-\int_X \partial_s\log\widehat{\rho}(s,x)\rho(s)(\mathrm{d}x)-\int_Y \partial_s\log\widehat{\sigma}(s,y)\sigma(s)(\mathrm{d}y)\\
&=\frac{1}{\tau}\mathsf{D}^2_{\rho\rho,s}L[\rho(s)-\widehat{\rho}(s),\dot{\rho}(s)]+\frac{1}{\tau}\mathsf{D}^2_{\rho\sigma,s}L[\rho(s)-\widehat{\rho}(s),\dot{\sigma}(s)]\\
&\quad-\frac{1}{\tau}\mathsf{D}^2_{\sigma\rho,s}L[\sigma(s)-\widehat{\sigma}(s),\dot{\rho}(s)]-\frac{1}{\tau}\mathsf{D}^2_{\sigma\sigma,s}L[\sigma(s)-\widehat{\sigma}(s),\dot{\sigma}(s)]. \qquad (105)
\end{aligned}$$

Using $\rho(s)-\widehat{\rho}(s)=-\Delta_\rho(s)$, $\sigma(s)-\widehat{\sigma}(s)=-\Delta_\sigma(s)$ and Eq. (96), the exact best-response part of Eq. (105) is

$$\begin{aligned}
&-\frac{\beta}{\tau}\mathsf{D}^2_{\rho\rho,s}L[\Delta_\rho(s),\Delta_\rho(s)]-\frac{\beta}{\tau}\mathsf{D}^2_{\rho\sigma,s}L[\Delta_\rho(s),\Delta_\sigma(s)]\\
&\quad+\frac{\beta}{\tau}\mathsf{D}^2_{\sigma\rho,s}L[\Delta_\sigma(s),\Delta_\rho(s)]+\frac{\beta}{\tau}\mathsf{D}^2_{\sigma\sigma,s}L[\Delta_\sigma(s),\Delta_\sigma(s)]. \qquad (106)
\end{aligned}$$

The signed measures $\Delta_\rho(s)$ and $\Delta_\sigma(s)$ have total mass zero. Because L is twice differentiable, convex in ρ, and concave in σ,

$$\mathsf{D}^2_{\rho\rho,s} L[\Delta_\rho(s), \Delta_\rho(s)] \geq 0, \qquad \mathsf{D}^2_{\sigma\sigma,s} L[\Delta_\sigma(s), \Delta_\sigma(s)] \leq 0.$$

The two mixed terms cancel exactly by Asm. 2:

$$\mathsf{D}^2_{\rho\sigma,s} L[\Delta_\rho(s), \Delta_\sigma(s)] = \mathsf{D}^2_{\sigma\rho,s} L[\Delta_\sigma(s), \Delta_\rho(s)].$$

Hence the exact best-response contribution is non-positive.

The residual terms in Eq. (105) contain at least one factor $r_x(s)$ or $r_y(s)$. Using the bounded second functional derivatives in Asm. 2, $\|\Delta_\rho(s)\|_{\mathrm{TV}} \leq 1$, $\|\Delta_\sigma(s)\|_{\mathrm{TV}} \leq 1$, and Eq. (97)–Eq. (98), we obtain, after absorbing universal numerical factors from the signed-measure convention into C,

$$-\int_X \partial_s \log \widehat{\rho}(s,x)\rho(s)(\mathrm{d}x) - \int_Y \partial_s \log \widehat{\sigma}(s,y)\sigma(s)(\mathrm{d}y) \leq C\beta(\eta_k + \beta s), \tag{107}$$

where C depends only on $\tau, C_{xx}, C_m, C_{yy}, \ell_x, \ell_y$, and $C_m := \max\{C_{xy}, C_{yx}\}$.

Define

$$\begin{aligned} A_x &= \frac{12C_x + 4(C_{xx} + C_m)}{\tau}, \qquad A_y = \frac{12C_y + 4(C_{yy} + C_m)}{\tau}, \\ C_0 &= 1 + A_x(1 + 2\ell_x) + A_y(1 + 2\ell_y). \end{aligned} \tag{108}$$

The residual terms satisfy

$$\beta\big[A_x\|r_x(s)\|_{\mathrm{TV}} + A_y\|r_y(s)\|_{\mathrm{TV}}\big] \leq C_0\beta\eta_k + C_0\beta^2 s,$$

using Eqs. (97) and (98) and $\|\theta\|_{\mathrm{variation}} = 2\|\theta\|_{\mathrm{TV}}$ for zero-mass signed measures. Thus the following bounds hold with $C = C_0$.

Combining Eq. (99), Eq. (101), Eq. (102), and Eq. (107), we get, for all $s \in [0,1]$,

$$\frac{\mathrm{d}}{\mathrm{d}s} G(\rho(s), \sigma(s)) \leq -\beta G(\rho(s), \sigma(s)) + C_0\beta\eta_k + C_0\beta^2 s. \tag{109}$$

Gronwall's inequality on $[0,1]$ gives

$$G_{k+1} \leq e^{-\beta} G_k + C_0\beta\eta_k + C_0\beta^2. \tag{110}$$

Iterating Eq. (110) and using Eq. (92),

$$G_k \leq e^{-\beta k} G_0 + \frac{C_0\beta\eta_T + C_0\beta^2}{1 - e^{-\beta}}. \tag{111}$$

For $\beta \in (0,1]$, $1 - e^{-\beta} \geq \beta/2$. Taking $C_{\mathrm{out}} = 2\sqrt{2}C_0 \geq 1$ gives

$$G_k \leq e^{-\beta k} G_0 + C_{\mathrm{out}} \left(\sqrt{C_{\mathrm{kin}}} e^{-\lambda_{\mathrm{kin}} T/2} + \beta\right). \tag{112}$$

Lemma 9 gives

$$\mathrm{NI}_\tau(\rho_k, \sigma_k) \leq \tau\left[D_{\mathrm{KL}}(\rho_k \mid \widehat{\rho}_{\rho_k,\sigma_k}) + D_{\mathrm{KL}}(\sigma_k \mid \widehat{\sigma}_{\rho_k,\sigma_k})\right].$$

Substituting Eq. (112) proves Eq. (18). □

E.4 Proof of Cor. 1

Proof. We write $G_k := G(\rho_k, \sigma_k)$ for the kth step. The Gaussian-restart density envelopes in Lemma 11 are uniform in the horizon and are preserved by any convex outer mixture. Consequently, Eq. (110) applies independently at every step, with the same C_0 in Eq. (108):

$$G_{k+1} \leq e^{-\beta_k} G_k + C_0\beta_k\eta_k + C_0\beta_k^2, \qquad \eta_k \leq \sqrt{2C_{\mathrm{kin}}} e^{-\lambda_{\mathrm{kin}} T_k/2}. \tag{113}$$

Writing $a = \sqrt{2C_{\mathrm{kin}}}$, the schedule in Cor. 1 gives $\eta_k \leq a/[1 + a(k+2)] \leq 1/(k+2)$. Hence

$$G_{k+1} \leq e^{-2/(k+2)} G_k + \frac{6C_0}{(k+2)^2}.$$

For $t > 0$, $\log(1 + 1/t) \le 1/t$, so $e^{-2/t} \le [t/(t+1)]^2$. Multiplying by $(k+3)^2$ and using $(k+3)/(k+2) \le 3/2$ yields

$$(k+3)^2 G_{k+1} \le (k+2)^2 G_k + \frac{27}{2} C_0.$$

Summing from $k = 0$ to $K-1$ proves Eq. (114), using $\mathrm{NI}_\tau \le \tau G$.

$$\mathrm{NI}_\tau(\rho_K, \sigma_K) \le \tau G_K \le \tau \frac{4G_0 + (27/2)C_0 K}{(K+2)^2}. \tag{114}$$

The accumulated response time in the inner loop per player has the explicit bound

$$\sum_{k=0}^{K-1} T_k \le \frac{2K}{\lambda_{\mathrm{kin}}} \log\bigl(1 + \sqrt{2C_{\mathrm{kin}}}(K+1)\bigr). \tag{115}$$

Given ϵ, then $K = O(\varepsilon^{-1})$ suffices by Eq. (114), yielding total response time in inner loop $O(\lambda_{\mathrm{kin}}^{-1} \varepsilon^{-1} \log(1/\varepsilon))$. □

E.5 Acceleration of underdamped anchor best responses

We show that underdamped MFL–ABR can admit a better sampling time guarantee than its overdamped counterpart. This parallels the acceleration mechanism of Fu & Wilson (2024, Theorem 3.1), who obtain a faster underdamped mean-field convergence rate through a square-root friction choice. Indeed, maximizing Eq. (85) over the friction gives $\gamma_i = \sqrt{L_i + \tau}, \quad i \in \{x, y\}$, and hence $\lambda_{\mathrm{kin}} = \frac{\tau}{32} \min_{i \in \{x,y\}} \frac{\alpha_i}{\sqrt{L_i + \tau}}$, where $\alpha_x, \alpha_y \le 1$.

For comparison, consider the same anchored best-response scheme with overdamped Langevin dynamics as the inner samplers. For the x-response,

$$\mathrm{d}Z_t = -\bigl(\tau Z_t + \nabla_x \frac{\delta L}{\delta \rho}(\rho, \sigma, Z_t)\bigr)\,\mathrm{d}t + \sqrt{2\tau}\,\mathrm{d}\mathcal{B}_t^x, \qquad Z_0 \sim \gamma_x^0.$$

The LSI estimate of Kim et al. (2024, Proposition D.1) gives

$$D_{\mathrm{KL}}(\mathrm{Law}(Z_T) \mid \widehat{\rho}_{\rho,\sigma}) \le \frac{2C_x}{\tau} e^{-2\tau\alpha_x T},$$

and analogously for the y-response. Thus $\lambda_{\mathrm{od}} = 2\tau \min\{\alpha_x, \alpha_y\}$. The overdamped responses satisfy the same density-ratio envelopes used in Sec. E.3, so the same outer-loop argument applies.

Let $C_* = \max\{1, C_{\mathrm{kin}}, 2C_x/\tau, 2C_y/\tau\}$. For $j \in \{\mathrm{kin}, \mathrm{od}\}$, the response time $T_j(\varepsilon) = \frac{2}{\lambda_j} \log\Bigl(1 + \frac{3\tau C_{\mathrm{out}}\sqrt{C_*}}{\varepsilon}\Bigr)$ is sufficient to obtain the same NI guarantee with the same outer parameters. Consequently, $\frac{T_{\mathrm{kin}}(\varepsilon)}{T_{\mathrm{od}}(\varepsilon)} = \frac{\lambda_{\mathrm{od}}}{\lambda_{\mathrm{kin}}}$. Therefore, whenever $\lambda_{\mathrm{kin}} > \lambda_{\mathrm{od}}$, our bounds yield a smaller sufficient sampling time for underdamped MFL–ABR than for overdamped MFL–ABR to achieve the same NI accuracy.

F Proof for Sec. 5.3

F.1 Proof of Prop. 1

Proof. Let $\mathsf{A}_x = \left(\begin{smallmatrix} 0 & 1 \\ -\tau & -\gamma_x \end{smallmatrix}\right)$, $h_x(T) = (e^{T\mathsf{A}_x})_{11}$ and $r_x(T) = 1 - h_x(T)$, with the same notation for the y player. With opponent mean b fixed, the x-sampler's mean solves $\ddot{m} + \gamma_x \dot{m} + \tau m = -Bb$, with $m(0) = \dot{m}(0) = 0$. Subtracting its forced equilibrium $-Bb/\tau$ gives $m(T) = -[1 - h_x(T)]Bb/\tau$. The y response similarly has mean $[1 - h_y(T)]B^\top a/\tau$. A mixture has the corresponding mixture of means, proving Eq. (19).

The function h_x solves $\ddot{h}_x + \gamma_x \dot{h}_x + \tau h_x = 0$ with $(h_x(0), \dot{h}_x(0)) = (1, 0)$. For $T > 0$,

$$\tau h_x(T)^2 + \dot{h}_x(T)^2 = \tau - 2\gamma_x \int_0^T \dot{h}_x(t)^2\,\mathrm{d}t < \tau,$$

since $\ddot{h}_x(0) = -\tau \neq 0$. Thus $|h_x(T)| < 1$ and $r_x(T) \in (0, 2)$; the same holds for $r_y(T)$.

An SVD of B decomposes the outer matrix into uncoupled eigenvalues $1 - \beta$ and, for each positive singular value s, the pair

$$1 - \beta \pm \mathrm{i}\,\beta\sqrt{r_x(T)r_y(T)}\, s/\tau.$$

The squared modulus of the latter is $(1 - \beta)^2 + \beta^2 r_x(T)r_y(T)s^2/\tau^2$. It is strictly less than one exactly when $\beta[1 + r_x(T)r_y(T)s^2/\tau^2] < 2$. Taking the largest singular value proves the stated equivalence for mean stability.

To prove Eq. (20), fix the outer means and restart from the reference Gaussian. The constant opponent force shifts only the mean of the linear inner Langevin process. For the x-sampler its covariance remains $\Sigma_x = \mathrm{diag}(I_{d_x}, \tau I_{d_x})$, since

$$(\mathsf{A}_x \otimes I_{d_x})\Sigma_x + \Sigma_x(\mathsf{A}_x^\top \otimes I_{d_x}) + \mathrm{diag}(0, 2\gamma_x\tau I_{d_x}) = 0.$$

The y-sampler is identical. Hence $\mathcal{R}_x^T = \mathcal{N}(-r_x(T)Bb/\tau, I_{d_x})$ and $\mathcal{R}_y^T = \mathcal{N}(r_y(T)B^\top a/\tau, I_{d_y})$. The KL divergence between two Gaussians with the same identity covariance is half the squared difference of their means, proving Eq. (20).

Fix $T > 0$ and abbreviate $r_x = r_x(T)$, $r_y = r_y(T)$. Since the frozen force depends only on the outer means, these response formulas apply at every iteration. Define

$$e_k = \frac{\|a_k\|^2}{r_x} + \frac{\|b_k\|^2}{r_y}, \qquad H_k = D_{\mathrm{KL}}(\rho_k \mid \gamma_x^0) + D_{\mathrm{KL}}(\sigma_k \mid \gamma_y^0), \qquad q^2 = (1-\beta)^2 + \frac{\beta^2 r_x r_y \|B\|_{\mathrm{op}}^2}{\tau^2}. \tag{116}$$

Scaling the mean coordinates by $r_x^{-1/2}$ and $r_y^{-1/2}$ transforms Eq. (19) into

$$(1 - \beta)I + \frac{\beta\sqrt{r_x r_y}}{\tau}\begin{pmatrix} 0 & -B \\ B^\top & 0 \end{pmatrix}.$$

The second matrix is skew-symmetric. Its cross terms therefore vanish in the squared Euclidean norm, yielding $e_{k+1} \leq q^2 e_k$ and $e_k \leq q^{2k} e_0$. The response entropies relative to the reference laws satisfy

$$D_{\mathrm{KL}}(\mathcal{R}_x^T(\rho_k, \sigma_k) \mid \gamma_x^0) + D_{\mathrm{KL}}(\mathcal{R}_y^T(\rho_k, \sigma_k) \mid \gamma_y^0) = \frac{r_x^2\|Bb_k\|^2 + r_y^2\|B^\top a_k\|^2}{2\tau^2} \leq D_T e_0 q^{2k},$$

Here $D_T := \frac{\|B\|_{\mathrm{op}}^2 r_x r_y (r_x + r_y)}{2\tau^2}$. Convexity of KL in its first argument now gives $H_{k+1} \leq (1 - \beta)H_k + \beta D_T e_0 q^{2k}$. Iteration proves the explicit bound

$$H_K \leq (1 - \beta)^K H_0 + \beta D_T e_0 \sum_{j=0}^{K-1} (1 - \beta)^{K-1-j} q^{2j}, \qquad \mathrm{NI}_\tau(\rho_K, \sigma_K) \leq (\tau + \|B\|_{\mathrm{op}}^2/\tau) H_K. \tag{117}$$

Under the outer step size condition in Prop. 1, both $(1 - \beta)$ and q^2 are less than one. For any $\theta \in (\max\{1-\beta, q^2\}, 1)$, the sum in Eq. (117) is bounded by $C\theta^K$ for some constant $C > 0$. Hence the outer iteration converges exponentially in NI_τ. □

G UNDERDAMPED ABR IN THE FINITE-PARTICLE, DISCRETE-TIME SETTING

G.1 ALGORITHM AND MAIN THEOREM

We give the finite-particle, discrete-time implementation in Alg. 3 and prove its last-iterate guarantee.

Algorithm 3: Finite-particle version of underdamped MFL–ABR

Input: Particle number $N \geq 2$, outer rounds K, inner steps $M \geq 1$, step size $h > 0$, temperature $\tau > 0$, frictions $\gamma_x, \gamma_y > 0$, and $\beta \in (0, 1]$ with $m = \lfloor \beta N \rfloor \geq 1$

Independently initialize $X_0^j \sim \gamma_x^0$, $Y_0^j \sim \gamma_y^0$, $j = 1, \dots, N$.

for $k = 0, \dots, K-1$ **do**

Set $u_k^N = (\rho_k^N, \sigma_k^N) = (N^{-1} \sum_j \delta_{X_k^j}, N^{-1} \sum_j \delta_{Y_k^j})$.

foreach *player* $i \in \{x, y\}$ **do**

Independently restart $(z_0^j, v_0^j) \sim \gamma_i^0 \otimes g_{\tau,d_i}$, $j = 1, \dots, N$, with (ρ_k^N, σ_k^N) frozen and force

$$F_k(z) = \begin{cases} \tau z + \nabla_z \frac{\delta L}{\delta \rho}(\rho_k^N, \sigma_k^N, z), & i = x, \\ \tau z - \nabla_z \frac{\delta L}{\delta \sigma}(\rho_k^N, \sigma_k^N, z), & i = y. \end{cases}$$

Set $d_h = e^{-\gamma_i h}$, $a_h = (1 - d_h)/\gamma_i$, $b_h = (h - a_h)/\gamma_i$.

for $\ell = 0, \dots, M-1$ **do**

For all j, draw centered Gaussian pairs, independently across particles, inner steps, players, and outer rounds:

$$\begin{pmatrix} \xi_\ell^{z,j} \\ \xi_\ell^{v,j} \end{pmatrix} \sim \mathcal{N}(0, \Sigma_{h,i} \otimes I_{d_i}),$$

$$\Sigma_{h,i} = \begin{pmatrix} \frac{2\tau}{\gamma_i}\left(h - 2a_h + \frac{1 - d_h^2}{2\gamma_i}\right) & \frac{\tau(1-d_h)^2}{\gamma_i} \\ \frac{\tau(1-d_h)^2}{\gamma_i} & \tau(1 - d_h^2) \end{pmatrix}. \tag{118}$$

For each j, update:

$$z_{\ell+1}^j = z_\ell^j + a_h v_\ell^j - b_h F_k(z_\ell^j) + \xi_\ell^{z,j}, \qquad v_{\ell+1}^j = d_h v_\ell^j - a_h F_k(z_\ell^j) + \xi_\ell^{v,j}. \tag{119}$$

end

Denote the terminal positions by $\widetilde{X}_k^j$ for $i = x$ and $\widetilde{Y}_k^j$ for $i = y$.

end

Independently of all previous randomness, choose uniform m-element subsets $I_k^x, I_k^y \subset \{1, \dots, N\}$ and independent indices $J_k^{x,j}, J_k^{y,j} \sim \mathrm{Unif}\{1, \dots, N\}$.

For each j, update

$$X_{k+1}^j = \begin{cases} \widetilde{X}_k^{J_k^{x,j}}, & j \in I_k^x, \\ X_k^j, & j \notin I_k^x, \end{cases} \qquad Y_{k+1}^j = \begin{cases} \widetilde{Y}_k^{J_k^{y,j}}, & j \in I_k^y, \\ Y_k^j, & j \notin I_k^y. \end{cases} \tag{120}$$

end

Output: $(X_K^j)_{j=1}^N$ and $(Y_K^j)_{j=1}^N$

Theorem 4 (Finite-particle last-iterate convergence)**.** *Under Assumptions 1–3, fix $\gamma_x, \gamma_y > 0$. There exist $C, h_0, \beta_0 > 0$, depending only on the dimensions, temperature, frictions, and assumption constants, such that Alg. 3 with $0 < h \leq h_0$ and $0 < \beta_N := \lfloor \beta N \rfloor / N \leq \beta_0$ satisfies, for every $K \geq 0$,*

$$\begin{aligned} &D_{\mathrm{KL}}(\rho_K^{(N)} \mid \rho^\star) + D_{\mathrm{KL}}(\sigma_K^{(N)} \mid \sigma^\star) \leq \frac{\mathrm{NI}_\tau(\rho_K^{(N)}, \sigma_K^{(N)})}{\tau} \\ &\leq C\Big[e^{-\beta_N K/A^2} G_0 + e^{-\lambda_{\mathrm{kin}} T} + \varepsilon_{d_x}(N) + \varepsilon_{d_y}(N) + (1 + d_x + d_y) T h^2\Big]. \end{aligned} \tag{121}$$

Here $T = Mh$, $\rho_K^{(N)} = \mathbb{E}\rho_K^N$ and $\sigma_K^{(N)} = \mathbb{E}\sigma_K^N$ are the unconditional laws of uniformly selected output particles, $G_0 = G(\gamma_x^0, \gamma_y^0)$, λ_{kin} is as in Thm. 3, and $A = e^{4 \max\{C_x, C_y\}/\tau}$,

$$\varepsilon_d(N) = \begin{cases} (1 + \log N)/N, & d = 1, \\ (1 + \log N)^3/N, & d = 2, \\ (1 + \log N) N^{-2/d}, & d \geq 3. \end{cases} \tag{122}$$

Proof. Set $D = (1 + d_x + d_y)Th^2$ and let q_k be the auxiliary mixture of exact finite-time responses in Lemma 12. Its components and the exact responses have reference densities in $[A^{-1}, A]$ by Lemma 11. Combining Lemmas 12 and 13 gives

$$\mathbb{E}G(q_{k+1}) \le e^{-\beta_N/A^2}\mathbb{E}G(q_k) + C\beta_N \left[e^{-\lambda_{\mathrm{kin}}T} + \varepsilon_{d_x}(N) + \varepsilon_{d_y}(N) + D\right].$$

Since $1 - e^{-\beta_N/A^2} \ge \beta_N/(2A^2)$, summation yields

$$\mathbb{E}G(q_K) \le e^{-\beta_N K/A^2}G_0 + C\left[e^{-\lambda_{\mathrm{kin}}T} + \varepsilon_{d_x}(N) + \varepsilon_{d_y}(N) + D\right]. \tag{123}$$

To transfer this estimate to $(\rho_K^{(N)}, \sigma_K^{(N)})$, first consider the x player and write $\bar{\rho} = \mathbb{E}q_K^x$. By Lemma 11, $w = \mathrm{d}\bar{\rho}/\mathrm{d}\rho^\star \in [A^{-2}, A^2]$. The entropy variational inequality and $\chi^2(\bar{\rho} \mid \rho^\star) \le 2A^2 D_{\mathrm{KL}}(\bar{\rho} \mid \rho^\star)$ give

$$D_{\mathrm{KL}}(\rho_K^{(N)} \mid \rho^\star) \le 2D_{\mathrm{KL}}(\rho_K^{(N)} \mid \bar{\rho}) + \log \int w^2 \mathrm{d}\rho^\star \le 2D_{\mathrm{KL}}(\rho_K^{(N)} \mid \bar{\rho}) + 2A^2 D_{\mathrm{KL}}(\bar{\rho} \mid \rho^\star).$$

The χ^2 bound follows from $w\log w - w + 1 \ge (w-1)^2/(2A^2)$ on $[A^{-2}, A^2]$. Summing this bound and its y-player analogue, using convexity and Lemmas 9 and 12, gives

$$D_{\mathrm{KL}}(\rho_K^{(N)} \mid \rho^\star) + D_{\mathrm{KL}}(\sigma_K^{(N)} \mid \sigma^\star) \le CD + 2A^2\mathbb{E}G(q_K).$$

The entropy comparison in Lemma 13 and Eq. (123) now prove Eq. (121). □

G.2 Auxiliary lemmas

For $u = (u^x, u^y)$, let $\mathcal{P}(u) = (\widehat{\rho}_u, \widehat{\sigma}_u)$ denote the pair of Gibbs responses, and let $\mathcal{R}^{T,h}(u)$ and $\mathcal{R}^T(u)$ denote the numerical and exact response laws. We view the corresponding density ratios with respect to the Gaussian references as elements of $\mathcal{H} = L^2(\gamma_x^0) \oplus L^2(\gamma_y^0)$.

Lemma 12. *Under the conditions of Thm. 4, define random laws*

$$q_0 = (\gamma_x^0, \gamma_y^0), \qquad q_{k+1} = (1 - \beta_N)q_k + \beta_N \mathcal{R}^T(u_k^N). \tag{124}$$

Then, uniformly in k,

$$D_{\mathrm{KL}}(\rho_k^{(N)} \mid \mathbb{E}q_k^x) + D_{\mathrm{KL}}(\sigma_k^{(N)} \mid \mathbb{E}q_k^y) \le C(1 + d_x + d_y)Th^2, \tag{125}$$

$$\mathbb{E}\|\mathcal{R}^T(u_k^N) - \mathcal{P}(q_k)\|_{\mathcal{H}}^2 \le C\left[e^{-\lambda_{\mathrm{kin}}T} + \varepsilon_{d_x}(N) + \varepsilon_{d_y}(N) + (1 + d_x + d_y)Th^2\right]. \tag{126}$$

Proof. For one frozen response, write $U(z) = \tau\|z\|^2/2 + f(z)$. Here $f(z) = \frac{\delta L}{\delta \rho}(u^x, u^y, z)$ for the x player and $f(z) = -\frac{\delta L}{\delta \sigma}(u^x, u^y, z)$ for the y player. Set $(d, \gamma, G, L) = (d_i, \gamma_i, G_i, L_i)$ for the chosen player $i \in \{x, y\}$. The assumptions give $\|\nabla f\| \le G$ and $\mathrm{Lip}(\nabla U) \le \tau + L$. The linear part of Eq. (119) corresponding to the quadratic potential is represented by

$$A_h = \begin{pmatrix} 1 - \tau b_h & a_h \\ -\tau a_h & d_h \end{pmatrix} = I_2 + hA_0 + O(h^2), \qquad A_0 = \begin{pmatrix} 0 & 1 \\ -\tau & -\gamma \end{pmatrix}.$$

Since A_0 is Hurwitz, a quadratic Lyapunov norm gives $\|A_h^\ell\| \le Ce^{-c\ell h}$ for all sufficiently small h. The recursion is a Gaussian term with uniformly bounded covariance (the one-step noise covariance is $O(h)$), plus a remainder bounded pathwise by $CGh\sum_{j\ge 0} e^{-cjh} \le C$. Hence, for some $\theta > 0$,

$$\sup_{\ell,u,h\le h_0} \mathbb{E}(\|z_\ell\|^2 + \|v_\ell\|^2) \le C(1 + d), \qquad \sup_{\ell,u,h\le h_0} \mathbb{E}e^{\theta(\|z_\ell\|^2 + \|v_\ell\|^2)} \le C. \tag{127}$$

The continuous interpolation with frozen force on $[\ell h, (\ell+1)h]$ satisfies $\mathrm{d}z_t = v_t\mathrm{d}t$ and $\mathrm{d}v_t = -(\gamma v_t + \nabla U(z_{\ell h}))\mathrm{d}t + \sqrt{2\gamma\tau}\mathrm{d}\mathcal{B}_t$. Its explicit velocity formula gives $\sup_t \mathbb{E}\|v_t\|^2 \le C(1 + d)$, hence

$$\mathbb{E}\|z_t - z_{\ell h}\|^2 \le C(1 + d)(t - \ell h)^2.$$

Let $\mathbb{P}_{h,T}, \mathbb{Q}_T$ be the numerical and exact path laws. Following the time-discretization analysis of Fu & Wilson (2024), apply Lehec (2013, Proposition 1), conditional on the common initial state, with

Brownian control $(\nabla U(z_t) - \nabla U(z_{\ell h}))/\sqrt{2\gamma\tau}$. Data processing inequality through the exact SDE solution map and averaging over initialization give

$$D_{\mathrm{KL}}(\mathbb{P}_{h,T} \mid \mathbb{Q}_T) \le \frac{1}{4\gamma\tau} \sum_{\ell=0}^{M-1} \int_{\ell h}^{(\ell+1)h} \mathbb{E}_{\mathbb{P}_{h,T}} \|\nabla U(z_t) - \nabla U(z_{\ell h})\|^2 \mathrm{d}t \le C(1+d)Th^2. \quad (128)$$

Taking the terminal marginal and applying the data processing inequality gives: $\sum_i D_{\mathrm{KL}}(\mathcal{R}_i^{T,h}(u) \mid \mathcal{R}_i^T(u)) \le CD$, where $D = (1 + d_x + d_y)Th^2$.

Let $s_0 = q_0$ and $s_{k+1} = (1-\beta_N)s_k + \beta_N \mathcal{R}^{T,h}(u_k^N)$. For the x player, set $\delta_k = \rho_k^N - s_k^x$. The y-player argument is identical. Conditional on the previous step history $\mathcal{F}_k$, Eq. (120) gives the signed-measure recursion

$$\delta_{k+1} = (1-\beta_N)\delta_k + \zeta_{k+1}, \qquad \mathbb{E}[\zeta_{k+1} \mid \mathcal{F}_k] = 0. \quad (129)$$

For a finite Borel partition $\mathcal{A}$, set $a_B = \rho_k^N(B)$ and $r_B = \mathcal{R}_x^{T,h}(u_k^N)(B)$. Uniform removal without replacement and independent resampling from the terminal response particles give

$$\begin{aligned} &\sum_{B\in\mathcal{A}} \mathbb{E}[\zeta_{k+1}(B)^2 \mid \mathcal{F}_k] \\ &= \frac{m}{N^2}\left[\frac{N-m}{N-1}\left(1 - \sum_B a_B^2\right) + \left(1 + \frac{m-1}{N}\right)\left(1 - \sum_B r_B^2\right)\right] \le \frac{2\beta_N}{N}. \end{aligned} \quad (130)$$

Indeed, the removal and resampling counts are conditionally independent, and two resampled particles share the same proposal index with probability $1/N$. Using $\sum_B \mathbb{E}\delta_0(B)^2 \le 1/N$, martingale orthogonality gives

$$\sum_B \mathbb{E}\delta_k(B)^2 \le \frac{2}{N}, \qquad \left\| \sum_{B\in\mathcal{A}} |\delta_k(B)| \right\|_{L^2(\mathbb{P})} \le \min\{2, \sqrt{2|\mathcal{A}|/N}\}. \quad (131)$$

The deterministic dyadic transport bound of Fournier & Guillin (2015, Lemma 5), applied after rescaling to the laws clipped to $[-a,a]^{d_x}$, combines with Eq. (131), Minkowski's inequality, and the exponential tails in Eq. (127) to give

$$\mathbb{E}W_1(\rho_k^N, s_k^x)^2 \le C\left[a^2\left(\sum_{\ell\ge 1} 2^{-\ell}\min\{1, 2^{d_x\ell/2}/\sqrt{N}\}\right)^2 + e^{-ca^2}\right] \le C\varepsilon_{d_x}(N). \quad (132)$$

For the last inequality take $a^2 = C(1+\log N)$ and split the geometric sum at $\lfloor (\log_2 N)/d_x \rfloor$. The required exponential moments hold pathwise for s_k^x and in expectation for ρ_k^N, by their mixture recursions.

Eq. (129) implies $(\rho_k^{(N)}, \sigma_k^{(N)}) = \mathbb{E}s_k$. Joint convexity of relative entropy gives $\sum_i D_{\mathrm{KL}}(s_k^i \mid q_k^i) \le CD$ pathwise and thus Eq. (125). By Lemma 11, all q_k^i and exact responses have reference densities in $[A^{-1}, A]$. Bounded perturbation and Talagrand therefore yield $\sum_i W_1(s_k^i, q_k^i)^2 \le CD$. Together with Eq. (132), this gives

$$\mathbb{E}\big(W_1(u_k^{N,x}, q_k^x)^2 + W_1(u_k^{N,y}, q_k^y)^2\big) \le C(\varepsilon_{d_x}(N) + \varepsilon_{d_y}(N) + D).$$

For signed potentials $f_x(u) = \delta L/\delta\rho$, $f_y(u) = -\delta L/\delta\sigma$, let $i \in \{x,y\}$ and interpolate between $f_i(u)$ and $f_i(u')$. Writing $g = f_i(u') - f_i(u)$ and π_t for the interpolated Gibbs density, $\partial_t\pi_t = -\tau^{-1}(g - \mathbb{E}_{\pi_t}g)\pi_t$ and Poincaré inequality give

$$\|\partial_t\pi_t\|_2^2 \le A\tau^{-2}\operatorname{Var}_{\pi_t}(g) \le A^3\tau^{-2}(L_i')^2\big(W_1(u^x, u'^x) + W_1(u^y, u'^y)\big)^2.$$

Hence

$$\|\mathcal{P}(u) - \mathcal{P}(u')\|_{\mathcal{H}} \le CW_1(u^x, u'^x) + W_1(u^y, u'^y).$$

On $[A^{-1}, A]$, $\|f-g\|_2^2 \le 2AD_{\mathrm{KL}}(f \mid g)$. Thus Lemma 5 gives $\|\mathcal{R}^T(u) - \mathcal{P}(u)\|_{\mathcal{H}}^2 \le 4AC_{\mathrm{kin}}e^{-\lambda_{\mathrm{kin}}T}$. These two estimates prove Eq. (126). □

Lemma 13. *Under the same assumptions of Thm. 4, there exist $C, \beta_0 > 0$ such that, for pairs u, v with reference densities in $[A^{-1}, A]$ and $0 < b \le \beta_0$,*

$$G((1-b)u + bv) \le e^{-b/A^2} G(u) + Cb\|v - \mathcal{P}(u)\|_{\mathcal{H}}^2. \tag{133}$$

Moreover, for any pair (ρ, σ) with finite relative entropy to the Gaussian references,

$$\frac{\mathrm{NI}_\tau(\rho,\sigma)}{\tau} \le G(\rho,\sigma) \le C\big[D_{\mathrm{KL}}(\rho \mid \rho^\star) + D_{\mathrm{KL}}(\sigma \mid \sigma^\star)\big]. \tag{134}$$

Proof. For reference densities $f, g \in [A^{-1}, A]$,

$$\frac{1}{2A}\|f-g\|_2^2 \le D_{\mathrm{KL}}(f \mid g) \le \frac{A}{2}\|f-g\|_2^2, \qquad \|\log(f/g)\|_2 \le A\|f-g\|_2. \tag{135}$$

The bounded second functional derivatives also give $\|\mathcal{P}(u) - \mathcal{P}(u')\|_{\mathcal{H}} \le C\|u - u'\|_{\mathcal{H}}$. Set $u_s = u + sb(v-u)$, $d_s = u_s - \mathcal{P}(u_s)$, $e = v - \mathcal{P}(u)$, and $r_s = v - u + d_s$. The preceding Lipschitz bound implies, for sufficiently small b,

$$\|r_s\|_{\mathcal{H}} \le \|e\|_{\mathcal{H}} + Cb(\|d_s\|_{\mathcal{H}} + \|r_s\|_{\mathcal{H}}) \le 2\|e\|_{\mathcal{H}} + Cb\|d_s\|_{\mathcal{H}}.$$

Since $\dot{u}_s = b(-d_s + r_s)$, the exact-response contribution to $\mathrm{d}G(u_s)/\mathrm{d}s$ is at most $-b\|d_s\|_{\mathcal{H}}^2/A$, the mixed second-variation terms cancel and the pure ones are nonpositive, as in Eq. (106). The residual contribution is at most $Cb\|d_s\|_{\mathcal{H}}\|r_s\|_{\mathcal{H}}$ by Eq. (135) and the bounded second functional derivatives. Young's inequality and a smaller choice of β_0 give

$$\frac{\mathrm{d}}{\mathrm{d}s}G(u_s) \le -\frac{b}{2A}\|d_s\|_{\mathcal{H}}^2 + Cb\|e\|_{\mathcal{H}}^2 \le -\frac{b}{A^2}G(u_s) + Cb\|e\|_{\mathcal{H}}^2.$$

Integrating over $s \in [0,1]$ proves Eq. (133).

For Eq. (134), let

$$g_x = \frac{\delta L}{\delta \rho}(\rho,\sigma,\cdot) - \frac{\delta L}{\delta \rho}(\rho^\star,\sigma^\star,\cdot), \qquad g_y = \frac{\delta L}{\delta \sigma}(\rho,\sigma,\cdot) - \frac{\delta L}{\delta \sigma}(\rho^\star,\sigma^\star,\cdot).$$

The bounded second functional derivatives and Pinsker's inequality imply

$$\|g_x\|_\infty^2 + \|g_y\|_\infty^2 \le C\big(D_{\mathrm{KL}}(\rho \mid \rho^\star) + D_{\mathrm{KL}}(\sigma \mid \sigma^\star)\big).$$

Since the Gibbs responses satisfy $\frac{d\widehat{\rho}_{\rho,\sigma}}{d\rho^\star} \propto e^{-g_x/\tau}$, $\frac{d\widehat{\sigma}_{\rho,\sigma}}{d\sigma^\star} \propto e^{-g_y/\tau}$, we have

$$\log \mathbb{E}_{\rho^\star} e^{-g_x/\tau} \le -\frac{\mathbb{E}_{\rho^\star} g_x}{\tau} + \frac{\|g_x\|_\infty^2}{2\tau^2},$$

and the analogous inequality holds for the y-player. Therefore,

$$\begin{aligned} G(\rho,\sigma) &\le D_{\mathrm{KL}}(\rho \mid \rho^\star) + D_{\mathrm{KL}}(\sigma \mid \sigma^\star) \\ &\quad + \frac{2}{\tau}\Big(\mathrm{TV}(\rho,\rho^\star)\|g_x\|_\infty + \mathrm{TV}(\sigma,\sigma^\star)\|g_y\|_\infty\Big) + \frac{1}{2\tau^2}(\|g_x\|_\infty^2 + \|g_y\|_\infty^2). \end{aligned}$$

Using Pinsker's inequality and the previous bound on g_x, g_y, we obtain

$$G(\rho,\sigma) \le C\big(D_{\mathrm{KL}}(\rho \mid \rho^\star) + D_{\mathrm{KL}}(\sigma \mid \sigma^\star)\big).$$

The first inequality in Eq. (134) follows from Lemma 9. □

H Further experimental details and results

H.1 Game objective and assumption

We first verify that the sigmoid game satisfies the assumptions used in our analysis. Write $q(r) = (1 + e^{-r^2})^{-1}$, so $q'(r) = \frac{2re^{-r^2}}{(1+e^{-r^2})^2}$, $\frac{1}{2} \le q(r) < 1$. The first derivatives are bounded by $\sqrt{2/e}$, and the absolute second spatial derivatives are bounded by $1/2 + 4/e$. Thus the boundedness and spatial regularity assumptions hold. The Gibbs potentials have oscillation at most $1/2$, so bounded perturbation of the standard Gaussian gives a uniform spatial LSI constant $\alpha \ge e^{-1/(2\tau)}$.

H.2 Evaluation

Let $\gamma^0 = \mathcal{N}(0,1)$. For a pair of densities, define the Gibbs responses by

$$\widehat{\rho}_\sigma(\mathrm{d}x) \propto e^{-\int \frac{1}{1+e^{-(x-y)^2}}\sigma(\mathrm{d}y)/\tau}\gamma^0(\mathrm{d}x), \qquad \widehat{\sigma}_\rho(\mathrm{d}y) \propto e^{\int \frac{1}{1+e^{-(x-y)^2}}\rho(\mathrm{d}x)/\tau}\gamma^0(\mathrm{d}y). \tag{136}$$

Bilinearity cancels the payoff pairings in the two response KL terms, yielding

$$\mathrm{NI}_\tau(\rho,\sigma) = \tau\big\{D_{\mathrm{KL}}(\rho \mid \widehat{\rho}_\sigma) + D_{\mathrm{KL}}(\sigma \mid \widehat{\sigma}_\rho)\big\} = \tau G(\rho,\sigma). \tag{137}$$

For particles $X_1,\dots,X_N$, we obtain a Gaussian-smoothed law by convolving the empirical law $\frac{1}{N}\sum_{i=1}^N \delta_{X_i}$ with $\mathcal{N}(0,0.1^2)$. The max-player density is constructed analogously. Write $\rho_N = N^{-1}\sum_{i=1}^N \delta_{X_i}$ and $\sigma_N = N^{-1}\sum_{i=1}^N \delta_{Y_i}$ for the unsmoothed empirical laws. We compute the Gibbs responses and the KL terms in Eq. (137) using these smoothed laws. The numerical evaluation uses Gaussian convolution and uniform-grid quadrature with spacing 0.01 on $[-12,12]$, with the domain expanded when necessary. In addition, we evaluate $W_1(\rho_N,\rho^\star) + W_1(\sigma_N,\sigma^\star)$ by quantile integration. The reference equilibrium is obtained numerically by solving the coupled Gibbs equations on the quadrature grid. The uncertainty bands in the main figure show pointwise interquartile ranges across the ten runs. The resulting terminal errors are reported in Tab. 3.

H.3 Particle updates and parameter selection

All outer positions initially follow $\gamma^0 = \mathcal{N}(0,1)$, and underdamped velocities independently follow $\mathcal{N}(0,\tau)$. The initial particles are shared across implementations for each seed. All three methods use the Gaussian restarts, frozen outer particle systems, and fixed-count outer replacement in Alg. 3, changing only the inner loop. The forces used in the inner loop are

$$F_x(z) = \tau z + \frac{1}{N}\sum_{j=1}^N q'(z-Y_j), \qquad F_y(z) = \tau z - \frac{1}{N}\sum_{j=1}^N q'(z-X_j). \tag{138}$$

The following updates apply to either player, with its corresponding force F. OD–EM uses

$$z' = z - hF(z) + \sqrt{2\tau h}\,\xi.$$

UD–EI uses the frozen-force exponential integrator of Fu & Wilson (2024), with update Eq. (119) and covariance Eq. (118). UD–EM uses the pre-step velocity in the position update:

$$z' = z + hv, \qquad v' = (1-\gamma h)v - hF(z) + \sqrt{2\gamma\tau h}\,\xi.$$

For the EM updates, $\xi \sim \mathcal{N}(0,1)$ independently across steps, particles, and players. One inner step costs one drift evaluation per player for every implementation. Thus, 200 outer rounds with 20 inner steps give 4000 drift evaluations per player, with $N = 1000$ particles per player. We search over 48 parameter configurations for each method, using the grids in Tab. 2.

Table 2: Parameter grids used for tuning the inner samplers.

Method	Parameter	Search grid
UD–EI, UD–EM	h	$\{0.3, 0.6, 1.2, 2.4, 4.8, 9.6\}$
	γ	$\{0.05, 0.15, 0.45, 1.35\}$
	β	$\{0.01, 0.03\}$
OD–EM	h	$\{0.03, 0.06, 0.12, 0.24, 0.3, 0.48, 0.96, 1.92, 3.84, 7.68, 15.36, 30.72\}$
	β	$\{0.01, 0.03, 0.05, 0.15\}$

Parameters are selected by minimizing the mean estimated NI over force evaluations $\{3200, 3400, 3600, 3800, 4000\}$, averaged over four seeds. The selected parameters are then fixed and evaluated on ten test seeds.

Table 3: Terminal errors on the bounded sigmoid game at 4000 drift evaluations per player. W_1 is computed from the unsmoothed empirical distributions.

Sampler	NI	W_1
UD–EI	0.0176	0.6137
OD–EM	0.0262	1.0976
UD–EM	0.0231	0.8893

H.4 Additional experiment

Ablation on the bilinear game. We follow the setting in Sec. 6.1, retaining $\gamma_x = \gamma_y = 1$, the initial mean state $(1, 0, 0.3, 0)$, and exact matrix-exponential evaluation on $t \in [0, 30]$. We consider two additional temperatures: $\tau = 0.25$ with $|B| \in \{0.4, 0.5, 0.6\}$, and $\tau = 4$ with $|B| \in \{1.6, 2, 2.4\}$. Fig. 2 plots the normalized mean energy and phase trajectories. In both settings, the mean energy decays below and grows above the predicted boundary $|B| = \sqrt{\tau}$. At the boundary, the mean trajectories oscillate while the energy does not decay to zero.

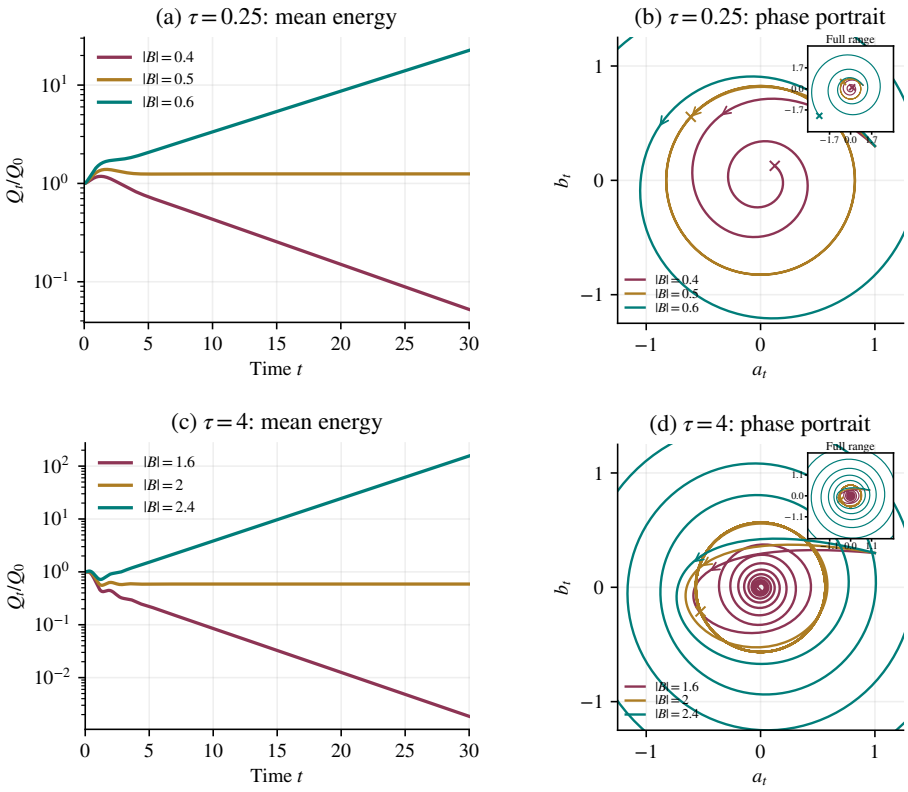


Figure 2: Bilinear-game ablation at two temperatures. Top row: $\tau = 0.25$; bottom row: $\tau = 4$. Left: Q_t/Q_0 with Q_t from Eqn. 10; right: mean phase trajectories (a_t, b_t), with crosses marking terminal states and arrows indicating direction. The phase plots include full-range insets.

Ablation on the smoothing bandwidth. The particle experiment in Sec. 6.2 uses a Gaussian bandwidth $b = 0.1$ to estimate NI. We vary $b \in \{0.05, 0.1, 0.2\}$ while keeping the terminal particles fixed. As shown in Tab. 4, UD–EI has the lowest median NI at every bandwidth, followed by UD–EM and OD–EM, consistent with the unsmoothed W_1 results in Tab. 3.

Table 4: Bandwidth ablation at 4000 drift evaluations per player. Entries are median of the estimated NI over ten runs. Lower is better.

Sampler	$b = 0.05$	$b = 0.1$	$b = 0.2$
UD–EI	**0.0180**	**0.0176**	**0.0171**
OD–EM	0.0265	0.0262	0.0258
UD–EM	0.0235	0.0231	0.0228

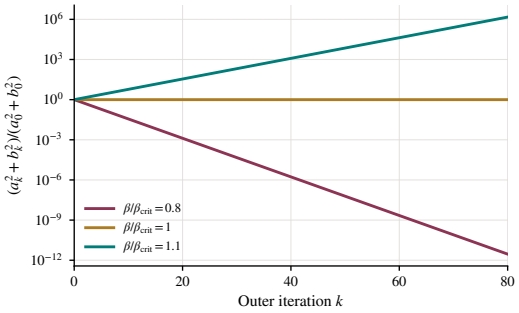


Figure 3: Effect of outer step size in the bilinear game with $B = 1$, $\tau = \gamma_x = \gamma_y = 1$, and $T = 3$. The critical value β_{crit} separates decay from growth.

Ablation on the outer step size. We use the bilinear setting in Sec. 6.1 with $B = 1$, $\tau = \gamma_x = \gamma_y = 1$, and initial means $(a_0, b_0) = (1, 0.3)$. For response horizon $T = 3$, Prop. 1 gives $\beta_{\text{crit}} \approx 0.88$. We vary $\beta/\beta_{\text{crit}} \in \{0.8, 1, 1.1\}$ in the exact outer mean map Eqn. 19. As shown in Fig. 3, the mean energy decays, remains constant, or grows, respectively, in agreement with the predicted stability boundary.