CONVERGENCE AND IMPLICIT BIAS OF SOFTMAX EXTRAGRADIENT FLOW IN BILINEAR ZERO-SUM GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Many learning problems optimize over probability distributions represented by softmax logits. Although extragradient enjoys strong guarantees for convex–concave games in policy space, its behavior under softmax parameterization is less understood: even bilinear games become nonconvex–nonconcave in logit space, and boundary equilibria correspond to logits at infinity. We study the high-resolution ODE of the extragradient method in softmax-parametrized skew-symmetric bilinear zero-sum games under symmetric self-play. We establish ***global convergence*** under two structural conditions. First, if there is an interior equilibrium, every trajectory converges in logit space. Second, under a condition we call *flat dominance*, we show convergence of every trajectory to the Nash set, with explicit rates. We further characterize the initialization-dependent ***implicit bias*** of the dynamics: which policy is selected? Conditional on policy convergence, we show that the trajectory must converge to a Nash equilibrium (NE). Under flat dominance, almost every such initialization leads to a *pure-strategy* NE. We further prove policy convergence for two structured classes of games and derive an explicit initialization-dependent NE selection rule. Finally, we present numerical experiments illustrating our theoretical results.

1 INTRODUCTION

Bilinear Game over Simplex. Many learning problems optimize over probability distributions: mixed strategies in finite games (von Neumann, 1928), policies in (multi-agent) reinforcement learning (Sutton & Barto, 2018), and preference distributions in alignment (Azar et al., 2024; Munos et al., 2024; Swamy et al., 2024) are representative examples. Particularly within game-theoretic applications, the finite two-player zero-sum game

$$\min_{\boldsymbol{x}\in\Delta_K}\max_{\boldsymbol{y}\in\Delta_K} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} \tag{Bilinear}$$

is one of the simplest globally solvable saddle-point problems: the payoff is bilinear and the feasible sets are simplices. Extragradient and optimistic methods exploit this structure and enjoy strong last-iterate guarantees (Cen et al., 2024; Daskalakis & Panageas, 2019; Daskalakis et al., 2018; Korpelevich, 1977; Liang & Stokes, 2019; Mokhtari et al., 2020; Nemirovski, 2004; Wei et al., 2021). In particular, Wei et al. (2021, Theorems 3, 8) showed linear convergence of OMWU over simplices and of OGDA over arbitrary polytopes for bilinear games.

In practice, however, probabilities are often represented by unconstrained parameters. A basic example is the softmax map $\boldsymbol{p}(\boldsymbol{\theta})$ for $\boldsymbol{\theta}\in\mathbb{R}^K$ defined by $p_i(\boldsymbol{\theta}) = \frac{\exp(\theta_i)}{\sum_{j=1}^K \exp(\theta_j)}$, which is standard in policy optimization (Agarwal et al., 2021; Li et al., 2021; 2023; Mei et al., 2020; 2021), models of human choice (Reverdy & Leonard, 2016), and multi-agent learning (Bernasconi et al., 2025; Fox et al., 2022; Hennes et al., 2020; Leung et al., 2024; Pattathil et al., 2023; Zhang et al., 2022). Composing (Bilinear) with softmax removes the explicit simplex constraint (which has been the primary difficulty in bilinear games literature until recently; see Cen et al. (2024); Wei et al. (2021)), but it also turns a convex–concave problem into a nonconvex–nonconcave one in the logits (as already highlighted in Cen et al. (2024, Section 1.1)). Moreover, softmax reaches only the relative interior of the simplex: a boundary equilibrium has no finite-logit representation, and approaching it requires some logits to diverge. Thus the logit dynamics remain outside the scope of existing theory for

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.

simplex-constrained bilinear games and convex–concave minimax optimization. In particular, it remains unclear whether the logit-space extragradient converges, how quickly it approaches boundary equilibria, or which equilibrium it selects when the Nash set is not a singleton. Equilibrium selection, in particular, remains comparatively underexplored in minimax optimization and has only recently begun to receive attention (Chen et al., 2025; Leal, 2026; Lee & Kim, 2025; Shcherbinin et al., 2026).

This separation between the simple simplex game and the geometry of its logit-space optimization makes softmax-parametrized bilinear games a particularly clean testbed for *structured nonconvex–nonconcave optimization*. The underlying bilinear game on the simplex is well understood (Mokhtari et al., 2020; Wei et al., 2021), yet under the softmax parameterization, the objective is no longer convex–concave in the logits, and boundary equilibria can only be approached as some logit differences diverge. Thus, any qualitatively new behavior can be traced to the interaction between the representation and the update rule rather than to an intrinsically difficult payoff.

Hidden Bilinearity. After fixing the additive-logit gauge,[1] softmax is a diffeomorphism onto the simplex interior, so the logit game is said to have a ***hidden bilinear structure*** (Pattathil et al., 2023; Vlatakis-Gkaragkounis et al., 2019). Prior work has leveraged this observation in two complementary ways. One line aims to redesign the algorithm to cancel or compensate for the representation map, using natural-gradient geometry, Jacobian preconditioning, surrogate losses, or explicit regularization (D'Orazio et al., 2025; Mladenovic et al., 2022; Pattathil et al., 2023; Sakos et al., 2023). Another line *keeps the parameter-space dynamics unchanged and studies how the hidden structure shapes their behavior* (Patel & Vlatakis-Gkaragkounis, 2025; Vlatakis-Gkaragkounis et al., 2019; 2021). We take the latter viewpoint for Euclidean-logit extragradient: rather than neutralizing the softmax parameterization, we characterize its consequences for convergence, boundary rates, and equilibrium selection. Our paper is closest in spirit to Vlatakis-Gkaragkounis et al. (2021), as both works seek global guarantees despite an apparently nonconvex–nonconcave landscape. The settings and conclusions are complementary: their framework treats a broader class of hidden games, while ours focuses on unregularized bilinear softmax games and resolves phenomena created by boundary equilibria and diverging logits. The resulting equilibrium-selection question is an instance of *implicit bias*; to our knowledge, rigorous implicit-bias theory for minimax dynamics remains sparse and has not been discussed by Vlatakis-Gkaragkounis et al. (2021). We defer further discussions to Appendix A.

Our Research Question and Results. We ask:

What happens when Euclidean extragradient is applied directly in softmax-logit space, without changing the geometry or regularizing the game?

We study the first-order high-resolution continuous-time model of extragradient (Grimmer et al., 2022; Lu, 2022), specialized to a skew-symmetric game (e.g., when $\boldsymbol{A}^\top = -\boldsymbol{A}$) and symmetric self-play; we call it the ***Softmax Self-Play Extragradient Flow*** (S2P-EG-Flow). Our focus on symmetric self-play is motivated by its success in RL for two-player zero-sum games (Bai & Jin, 2020; Bai et al., 2020; Liu et al., 2021; Xiong et al., 2022) and preference learning for LLMs (Lee et al., 2026b; Munos et al., 2024; Swamy et al., 2024; Wu et al., 2025; Ye et al., 2024; Zhang et al., 2025a;b), where policies or responses are naturally compared against one another. Our use of a continuous-time model likewise follows a long tradition in optimization, dating to Polyak (1964) and continuing in recent minimax optimization literature (Chae et al., 2024; Feng et al., 2025; Grimmer et al., 2022; Hsieh et al., 2021; Lee et al., 2026a; Lu, 2022; Suh et al., 2023); such models expose an algorithm's geometry, make dynamical-systems tools available, and often guide—though do not automatically imply—discrete-time analysis.

In this paper, we provide a global description of S2P-EG-Flow in the following two aspects:

Global Convergence (Section 3). We identify two *structural* conditions that are sufficient for global convergence. If an interior symmetric Nash equilibrium (SNE) exists, then S2P-EG-Flow converges to a finite logit equilibrium, eventually at an exponential rate (**Theorem 3.1**). If the game satisfies flat dominance (Assumption 1; a flat Nash face uniformly dominates all excluded actions), then S2P-EG-Flow converges to the Nash set, with the dominated probability mass decaying as $\Theta(t^{-1})$ (**Theorem 3.2**). Furthermore, assuming $\boldsymbol{p}(\boldsymbol{\theta}(t)) \to \boldsymbol{q}$, we explicitly characterize the (asymptotic) convergence rate. Specifically, when such $\boldsymbol{q}$ is quasi-strict (Harsanyi, 1973), every excluded

[1] Since softmax is translation invariant, we restrict the logits to $\{\boldsymbol{\theta} \in \mathbb{R}^K : \sum_{i=1}^K \theta_i = 0\}$.

action j with margin $\delta_j = (\boldsymbol{A}\boldsymbol{q})_j > 0$ satisfies $p_j(t) = \frac{|\mathrm{supp}(\boldsymbol{q})|}{K\delta_j}\, t^{-1}(1 + o(1))$ (**Corollary 3.4**), reminiscent of the t^{-1} asymptotics arising in the implicit-bias analysis of logistic regression (Soudry et al., 2018) When such $\boldsymbol{q}$ is non-quasi-strict, we show that there is an instance where the decay rate can be arbitrarily slower than $\Theta(t^{-1})$ (**Proposition 3.1**).

Implicit Bias (Section 4). We begin by proving that if S2P-EG-Flow converges in the policy space ($\boldsymbol{p}(\boldsymbol{\theta}(t)) \to \boldsymbol{q}$), then $\boldsymbol{q}$ is necessarily a SNE (**Theorem 4.1**), Next, we return to flat dominance (Assumption 1) and ask *which Nash equilibrium S2P-EG-Flow selects*. We first show that almost every initialization selects a vertex of the Nash face rather than a mixed equilibrium, conditional on convergence in the policy space (**Theorem 4.2**). We then prove convergence in the policy space in two structured cases. If there are two actions on the Nash face and two outside it, every trajectory converges (**Theorem 4.3**). The trajectory is always convergent also when the submatrix between actions on the Nash face and those outside it has rank one; moreover, its limiting support is characterized by an initialization-dependent explicit rule (**Theorem 4.4**). For both cases, every initialization except on a measure zero set converges to a vertex of the simplex.

More broadly, our analysis characterizes how the geometry induced by the softmax representation shapes convergence, boundary behavior, and equilibrium selection of extragradient in the original bilinear game. Our proofs rely on a rich set of tools from dynamical systems, such as Lyapunov/LaSalle stability, normally hyperbolic invariant manifolds, and asymptotic pseudotrajectories (Benaïm, 1999; Benaïm & Hirsch, 1996; do Carmo, 1994; Hirsch et al., 1977; Mischaikow et al., 1995; Perko, 2001; Prüss et al., 2009; Teschl, 2012; Wiggins, 2003), which we review in **Appendix B**. By outlining the main proof ideas, we illustrate how these tools are utilized to overcome key technical hurdles of our problem, hoping to make them more accessible to the optimization community.

Notation. $\mathrm{Skew}(K)$ denotes the set of $K \times K$ skew-symmetric matrices, and Δ_K and Δ_K° denote the simplex and its relative interior. We write $\|\boldsymbol{A}\|_{\max} := \max_{i,j} |A_{ij}|$, while $\|\boldsymbol{A}\|_2$ denotes the spectral norm. The notation $\boldsymbol{A} > 0$ denotes an element-wise inequality. For a vector field $\boldsymbol{g} : \mathbb{R}^K \to \mathbb{R}^K$, $\mathrm{D}\boldsymbol{g}$ is its Jacobian. For a vector $\boldsymbol{q} \in \Delta_K$, $\mathrm{supp}(\boldsymbol{q}) := \{i \in [K] : q_i > 0\}$.

2 Problem Setup

Simplex-Constrained Bilinear Zero-Sum Symmetric Games. Let $\boldsymbol{A} \in \mathrm{Skew}(K) \setminus \{\boldsymbol{0}\}$ be given, and let $\mathcal{SN} = \mathcal{SN}(\boldsymbol{A})$ denote the set of symmetric Nash equilibria of the *simplex-constrained* bilinear zero-sum symmetric game (Bilinear). Its value is zero as $\boldsymbol{A}^\top = -\boldsymbol{A}$, and in particular, $\mathcal{SN}(\boldsymbol{A}) = \{\boldsymbol{q} \in \Delta_K : \boldsymbol{A}\boldsymbol{q} \geq 0\}$, which is convex.[2]

We also state a structural assumption in advance that we will often consider throughout the paper:

Assumption 1. *$\boldsymbol{A}$ satisfies **flat dominance**: there exists a partition $[K] = \mathcal{I} \dot\cup \mathcal{L}$ with $\mathcal{I}, \mathcal{L} \neq \emptyset$ and $0 < b_- < b_+ < \infty$ such that $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = \boldsymbol{0}$ (flat) and $\boldsymbol{A}_{ji} \in [b_-, b_+]$ for all $j \in \mathcal{L}$, $i \in \mathcal{I}$ (dominance).*

Intuitively, the above assumption requires that every retained action ($i \in \mathcal{I}$) strictly beats every dominated action ($j \in \mathcal{L}$) head-to-head. We emphasize that this is *not* needed for every statement to be presented, and that we will explicitly state only when this is required. We relegate further discussions on the structure of $\mathcal{SN}(\boldsymbol{A})$ to **Appendix K**.

Following standard practice in machine learning for probability distributions, we parametrize using ***softmax***. This induces the following *unconstrained* nonconvex-nonconcave minimax problem:

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^K} \max_{\boldsymbol{\theta}' \in \mathbb{R}^K} \boldsymbol{p}(\boldsymbol{\theta})^\top \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}'), \quad \boldsymbol{p}(\boldsymbol{\theta}) := \left(\frac{\exp(\theta_i)}{\sum_{j=1}^K \exp(\theta_j)} \right)_{i \in [K]}. \tag{1}$$

Softmax Self-Play Extragradient Flow. We study extragradient (Korpelevich, 1977; Mertikopoulos et al., 2019; Nemirovski, 2004) applied directly to (1), whose update rule with stepsize γ is

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t - \gamma F(\boldsymbol{z}_t - \gamma F(\boldsymbol{z}_t)), \tag{EG}$$

[2] Indeed, if $\boldsymbol{A}\boldsymbol{q} \geq 0$, then $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{q} \geq 0$ and $\boldsymbol{q}^\top \boldsymbol{A}\boldsymbol{y} = -(\boldsymbol{A}\boldsymbol{q})^\top \boldsymbol{y} \leq 0$ for every $\boldsymbol{x}, \boldsymbol{y} \in \Delta_K$. Hence, for $i \in \mathrm{supp}(\boldsymbol{q})$ then necessarily $(\boldsymbol{A}\boldsymbol{q})_i = 0$ since $\boldsymbol{q}^\top \boldsymbol{A}\boldsymbol{q} = 0$.

where $F(\boldsymbol{z} := (\boldsymbol{\theta}, \boldsymbol{\phi})) = [\nabla_{\boldsymbol{\theta}} f^\top \ -\nabla_{\boldsymbol{\phi}} f^\top]^\top$ is the coupled gradient for $\min_{\boldsymbol{\theta}} \max_{\boldsymbol{\phi}} f(\boldsymbol{\theta}, \boldsymbol{\phi})$.

Exploiting the skew-symmetry of $\boldsymbol{A}$, we focus on *self-play dynamics*, in which a single strategy plays against itself; this is sufficient for finding an equilibrium because every skew-symmetric game admits a symmetric Nash equilibrium (Swamy et al., 2024). Applying self-play dynamics to (EG) is equivalent to considering extragradient on the set $\{(\boldsymbol{\theta}, \boldsymbol{\phi}) : \boldsymbol{\phi} = \boldsymbol{\theta}\}$, yielding

$$F(\boldsymbol{\theta}, \boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{g}(\boldsymbol{\theta}) \\ \boldsymbol{g}(\boldsymbol{\theta}) \end{pmatrix}, \qquad \boldsymbol{g}(\boldsymbol{\theta}) := \left(\operatorname{diag}(\boldsymbol{p}(\boldsymbol{\theta})) - \boldsymbol{p}(\boldsymbol{\theta})\boldsymbol{p}(\boldsymbol{\theta})^\top\right) \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) = \operatorname{diag}(\boldsymbol{p}(\boldsymbol{\theta}))\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}),$$

where the last equality follows from $\boldsymbol{p}(\boldsymbol{\theta})^\top \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) = 0$.

In this paper, we study its continuous-time counterpart via its high-resolution ODE (Lu, 2022). Deferring the complete derivation to Appendix C (where we also prove well-posedness), we refer to the resulting flow as the ***Softmax Self-Play Extragradient Flow*** (S2P-EG-Flow): for some fixed $\gamma > 0$ and initialization $\boldsymbol{\theta}(0) \in \mathbb{R}^K$,

$$\boxed{\dot{\boldsymbol{\theta}} = -\boldsymbol{g}(\boldsymbol{\theta}) + \frac{\gamma}{2} \mathrm{D}\boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{\theta}), \quad \boldsymbol{g}(\boldsymbol{\theta}) := \boldsymbol{p}(\boldsymbol{\theta}) \odot \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}),} \qquad \text{(S2P-EG-Flow)}$$

where the vector field $\boldsymbol{g}(\boldsymbol{\theta})$ is the gradient of the min-max objective $J(\boldsymbol{\theta}, \boldsymbol{\theta}') := \boldsymbol{p}(\boldsymbol{\theta})^\top \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}')$ with respect to $\boldsymbol{\theta}$, at $\boldsymbol{\theta}' = \boldsymbol{\theta}$. Throughout this paper, we let $\boldsymbol{\theta}(t)$ denote the solution of (S2P-EG-Flow), and $\boldsymbol{p}(t) := \boldsymbol{p}(\boldsymbol{\theta}(t)) \in \Delta_K$ denote the corresponding probability distribution or policy.

3 Global Convergence Behavior of the S2P-EG-Flow

This section studies the global convergence behavior of the S2P-EG-Flow (S2P-EG-Flow). To do so, we first distinguish between three notions of convergence:

Definition 3.1. *We say that the S2P-EG-Flow initialized at $\boldsymbol{\theta}(0) \in \mathbb{R}^K$*

1. ***converges in the logit space*** *if there exists a $\boldsymbol{\theta}^\infty \in \mathbb{R}^K$ such that $\boldsymbol{\theta}(t) \to \boldsymbol{\theta}^\infty$;*
2. ***converges in the policy space*** *if there exists a $\boldsymbol{p}^\infty \in \Delta_K$ such that $\boldsymbol{p}(t) \to \boldsymbol{p}^\infty$;*
3. ***converges in the Nash sense*** *if $\mathrm{DGap}(\boldsymbol{p}(t)) := \min_{\boldsymbol{x} \in \Delta_K} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{p}(t) \to 0$.*[a]

[a] The dual gap $\mathrm{DGap}(\boldsymbol{p})$ can be characterized as $\mathrm{DGap}(\boldsymbol{p}) = \min_k (\boldsymbol{A}\boldsymbol{p})_k \le \boldsymbol{p}^\top \boldsymbol{A}\boldsymbol{p} = 0$, with equality if and only if $\boldsymbol{p} \in \mathcal{SN}(\boldsymbol{A})$; thus $-\mathrm{DGap}(\boldsymbol{p}) \ge 0$ is the *exploitability* of $\boldsymbol{p}$.

3.1 Two Sufficient Conditions for Global Convergence

We characterize when (S2P-EG-Flow) converges. First, if the game has an interior equilibrium, we have the convergence in the logit space:

Theorem 3.1 (Interior SNE $\Rightarrow$ Logit Convergence)**.** *Suppose that $\{\boldsymbol{q} \in \Delta_K^\circ : \boldsymbol{A}\boldsymbol{q} = 0\}$ is nonempty.*[a] *For every $\boldsymbol{\theta}(0) \in \mathbb{R}^K$ and every $\gamma > 0$, there exists a finite $\boldsymbol{\theta}^\infty \in \mathbb{R}^K$ such that*

$$\boldsymbol{\theta}(t) \longrightarrow \boldsymbol{\theta}^\infty, \qquad \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}^\infty) = 0,$$

i.e., $\boldsymbol{p}(\boldsymbol{\theta}^\infty) \in \mathcal{SN}(\boldsymbol{A})$. Moreover, the convergence is eventually exponential: there are trajectory-dependent constants $T, C, \lambda > 0$ such that $\|\boldsymbol{\theta}(t) - \boldsymbol{\theta}^\infty\|_2 \le Ce^{-\lambda(t-T)}$ for $t \ge T$.

[a] Since $\boldsymbol{q}^\top \boldsymbol{A}\boldsymbol{q} = 0$, an equilibrium $\boldsymbol{q} \in \Delta_K^\circ$ should satisfy $\boldsymbol{A}\boldsymbol{q} = \boldsymbol{0}$.

Proof Sketch. The key observation is that we can construct a Lyapunov function that is nonincreasing along the trajectory as follows: $\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}) := \sum_i q_i e^{-\theta_i}$ for *any* interior SNE $\boldsymbol{q} \in \Delta_K^\circ \cap \mathcal{SN}(\boldsymbol{A})$ (**Lemma F.1**). Then, as $H_s := \{\boldsymbol{\theta} \in \mathbb{R}^K : \boldsymbol{1}^\top \boldsymbol{\theta} = \boldsymbol{1}^\top \boldsymbol{\theta}(0)\}$ is invariant (**Lemma E.1, logit conservation**), the LaSalle's invariance principle (Theorem B.3) applied to the sublevel set induced by $\Phi_{\boldsymbol{q}}$ implies that the omega-limit set of the trajectory is contained in the manifold $\mathcal{E} = \{\boldsymbol{\theta} : \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) = 0\}$. Via explicit computation, we can show that $\mathcal{E}$ is a normally attracting manifold of equilibria (**Lemma F.2**), and hence, by the asymptotic-phase theorem, the trajectory converges (eventually exponentially) to a single point of $\mathcal{E}$. The detailed proof is provided in Appendix F.1. □

Next, we consider the case where no interior equilibrium exists, i.e., when all SNEs are on the boundary. In this case, we will assume ***flat dominance*** (Assumption 1) of the game; let $(\mathcal{I}, \mathcal{L})$ be the corresponding partition. With this, we prove that Nash convergence occurs:

Theorem 3.2 (Flat Dominance ⇒ Nash Convergence)**.** *Suppose that Assumption 1 holds and* $0 < \gamma < \frac{2}{3\|\boldsymbol{A}\|_{\max}}$. *For every initial condition, every trajectory converges in the Nash sense, where the dominated mass* $\rho(t) := \sum_{j\in\mathcal{L}} p_j(\boldsymbol{\theta}(t))$ *decays as* $\rho(t) = \Theta(t^{-1})$.

Proof Sketch. We first show that $\liminf_{t\to\infty} \rho(t) = 0$. By Assumption 1, a uniform probability vector $\boldsymbol{c}$ supported on $\mathcal{I}$ is in $\mathcal{SN}(\boldsymbol{A})$. Constructing a Lyapunov potential using $\boldsymbol{c}$ (**Lemma F.1**) and then applying **Lemma E.3** yields the result. We then show that once $\rho(t)$ is sufficiently small, it should decrease to 0 at the rate of $\Theta(t^{-1})$ (**Lemma E.4**). For every $i \in \mathcal{I}$, Assumption 1 implies $-b_+\rho \le (\boldsymbol{Ap})_i \le -b_-\rho$. Once ρ is sufficiently small, $\mathrm{DGap}(\boldsymbol{p}) = \min_{k\in[K]}(\boldsymbol{Ap})_k$ is minimized on $\mathcal{I}$ and thus $b_-\rho \le -\mathrm{DGap}(p) \le b_+\rho$. The detailed proof is provided in Appendix F.2. □

Remark 1 (Nonconvergence of (S2P-EG-Flow))**.** *One may wonder whether convergence in Nash implies convergence in the policy space. The answer is* ***no****; we can construct an instance of* $\boldsymbol{A}$ *(that does not satisfy Assumption 1) in which depending on the choice of initialization, we either have convergence in Nash sense but not in policy (Proposition J.1), or not even convergence in Nash (Proposition J.2). We relegate detailed discussions to Appendix J.*

3.2 Sharp Boundary Asymptotics

We next characterize the rate of convergence conditional on $\boldsymbol{p}(t) \to \boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$. Let

$$\mathcal{S} := \mathrm{supp}(\boldsymbol{q}), \qquad \mathcal{B} := \{j \notin \mathcal{S} : (\boldsymbol{Aq})_j = 0\}, \qquad \mathcal{M} := \{j \notin \mathcal{S} : (\boldsymbol{Aq})_j > 0\}. \tag{2}$$

Here $\mathcal{B}$ is the ***zero-margin boundary set***: these actions vanish in probability but have zero Nash margin. In contrast, $\mathcal{M}$ is the ***positive-margin boundary set***: these actions vanish in probability and have strictly positive margin. We call $\boldsymbol{q}$ ***quasi-strict*** (Harsanyi, 1973) when $\mathcal{B} = \varnothing$. Write $s := |\mathcal{S}|$, $b := |\mathcal{B}|$, $m := |\mathcal{M}|$ and margin $\delta_j := (\boldsymbol{Aq})_j > 0$ for $j \in \mathcal{M}$.

The rate calculation reduces to the asymptotic behavior of the logits on $\mathcal{B}$. The stabilization condition below is automatic in the quasi-strict case, where $\mathcal{B} = \varnothing$. It is also automatic under flat dominance, conditional on the policy convergence assumed here: if $\mathcal{S} = \mathrm{supp}(\boldsymbol{q}) \subseteq \mathcal{I}$, then $\mathcal{B} = \mathcal{I} \setminus \mathcal{S}$, $\mathcal{M} = \mathcal{L}$, and every $\theta_k(t)$, $k \in \mathcal{B}$, converges to a finite limit. For a detailed explanation see Appendix I.4.

Theorem 3.3 (Boundary rates under the summed-logit condition on $\mathcal{B}$)**.** *Suppose* $\boldsymbol{p}(t) \to \boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ *and* $m \ge 1$. *Assume that* $\sum_{k\in\mathcal{B}} \theta_k(t)$ *converges to a finite limit, with the empty sum interpreted as* 0. *Then, we have that*

$$p_j(t) = \frac{s}{(s+m)\delta_j}\frac{1}{t}(1+o(1)), \ \ \forall j \in \mathcal{M}, \qquad \theta_i(t) = \frac{m}{s+m}\log t + O(1), \ \ \forall i \in \mathcal{S}. \tag{3}$$

If, in addition, $\theta_k(t)$ *converges to a finite limit for some* $k \in \mathcal{B}$, *then for a constant* $c_k > 0$,

$$p_k(t) = c_k t^{-m/(s+m)}(1+o(1)). \tag{4}$$

The proof uses the intrinsic clock and logit conservation, and is deferred to Appendix G.1. In the quasi-strict case, $\mathcal{B} = \varnothing$ and $m = K - s$, so Theorem 3.3 immediately yields the following.

Corollary 3.4 (Quasi-strict boundary rates)**.** *Suppose* $\boldsymbol{p}(t) \to \boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ *and* $\boldsymbol{q}$ *is a quasi-strict boundary equilibrium. Then,*

$$p_j(t) = \frac{s}{K\delta_j}\frac{1}{t}(1+o(1)), \qquad \forall j \notin \mathcal{S} \tag{5}$$

We next consider non-quasi-strict limits. In the simplest case, where the limiting equilibrium is pure and there is a single zero-margin boundary action ($|\mathcal{B}| = 1$), its logit stabilizes automatically.

Consequently, the positive-margin boundary actions decay as t^{-1}, while the zero-margin boundary action decays more slowly as $t^{-(K-2)/(K-1)}$; see Appendix G.3. However, this behavior does not extend uniformly to multiple zero-margin boundary actions: their individual decay rates need not be determined by the sizes of the three sets. The following proposition, whose proof is deferred to Appendix G.4, gives an explicit example where such actions can decay at arbitrarily slow, polynomial rates:

Proposition 3.1 (Nonuniversality with multiple zero-margin boundary actions)**.** *For every $\varepsilon \in (0, 1/2)$, there exists a four-action game with $\boldsymbol{p}(t) \to \boldsymbol{e}_1$ and $|\mathcal{S}| = |\mathcal{M}| = 1$, $|\mathcal{B}| = 2$, such that up to relabeling $\mathcal{B} = \{2, 3\}$, we have $p_2(t) = \Theta\big(t^{-(1-\varepsilon)}\big)$, $p_3(t) = \Theta\big(t^{-\varepsilon}\big)$, while $j \in \mathcal{M}$ satisfies $p_j(t) = \Theta(t^{-1})$.*

4 Implicit Bias of the S2P-EG Flow

This section studies the problem of *which* SNE is selected by (S2P-EG-Flow).

4.1 Every Policy Limit Is a SNE

By continuity of softmax, logit convergence implies policy convergence, but policy convergence alone need not imply Nash convergence for an arbitrary dynamical system. The following theorem shows that policy convergence does imply Nash convergence for (S2P-EG-Flow):

Theorem 4.1. *Suppose $0 < \gamma < \frac{2}{\|\boldsymbol{A}\|_2}$. If $\boldsymbol{p}(t) \to \boldsymbol{q} \in \Delta_K$, then $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$.*

Proof Sketch. The proof has two steps. First, we can show that policy convergence implies the limiting logit velocity vanishes: $\dot{\boldsymbol{\theta}} \to \boldsymbol{0}$. Direct computations then show that necessarily $\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0}$, i.e., $(\boldsymbol{A}\boldsymbol{q})_i = 0$ on $\operatorname{supp}(\boldsymbol{q})$. Second, if an excluded action $j \notin \operatorname{supp}(\boldsymbol{q})$ had $(\boldsymbol{A}\boldsymbol{q})_j < 0$, then it must be that $e^{-\theta_j(t)}$ eventually becomes negative (**Lemmas E.2 and E.3**), a contradiction. Hence $\boldsymbol{A}\boldsymbol{q} \geq 0$ and $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$. See Appendix H.1 for the complete argument. □

4.2 Implicit Bias under Flat Dominance

Throughout this subsection, write

$$\mathcal{I} := \bigcup_{\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})} \operatorname{supp}(\boldsymbol{q}), \quad \mathcal{L} := [K] \setminus \mathcal{I}, \quad m := |\mathcal{I}|, \quad l := |\mathcal{L}| \geq 1, \quad \boldsymbol{B} := \boldsymbol{A}_{\mathcal{L},\mathcal{I}} \in \mathbb{R}^{l \times m}. \tag{6}$$

Here $\mathcal{I}$ and $\mathcal{L}$ denote the retained and dominated actions. We assume ***flat dominance*** (Assumption 1) throughout: $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = \boldsymbol{0}$ and $0 < b_- \leq \boldsymbol{B} \leq b_+ < \infty$, entry-wise. Then $\mathcal{SN}(\boldsymbol{A}) = \Delta_{\mathcal{I}}$ (Proposition K.2). Relabeling actions if necessary, take $\mathcal{I} = [m]$ and write $\boldsymbol{p} = (\boldsymbol{p}_{\mathcal{I}}, \boldsymbol{p}_{\mathcal{L}}) \in \Delta_K$.

With $0 < \gamma < \frac{2}{3\|\boldsymbol{A}\|_{\max}}$ (which we will also assume throughout), every trajectory converges to $\Delta_{\mathcal{I}}$ by Theorem 3.2. We then ask: *does the policy converge to a point of this face (i.e., convergence in the policy space), and if so, which point?* If $m = 1$, the face consists of one vertex and there is nothing to select, so henceforth assume $m \geq 2$.

Assuming Policy Convergence. Following the implicit bias literature in deep learning theory (Arora et al., 2019; Chizat & Bach, 2020; Gunasekar et al., 2017; Woodworth et al., 2020), we first present a surprisingly strong sparsity implicit bias result *assuming* policy convergence under Assumption 1:

Theorem 4.2 (Policy Convergence ⇒ Vertex Selection *a.e.*)**.** *Suppose that Assumption 1 holds. Then, there is a Lebesgue-null set $\mathcal{N}_{\text{mix}} \subset \mathcal{O} := \{\boldsymbol{\theta}(0) \in \mathbb{R}^K : \boldsymbol{p}(t) \text{ converges}\}$ such that*

$$\mathcal{O} = \mathcal{N}_{\text{mix}} \,\dot\cup \bigcup_{i \in \mathcal{I}} \mathcal{O}_i, \qquad \mathcal{O}_i := \{\boldsymbol{\theta}(0) \in \mathbb{R}^K : \boldsymbol{p}(t) \to \boldsymbol{e}_i\}. \tag{7}$$

Moreover, each $\mathcal{O}_i$ is nonempty and open in $\mathbb{R}^K$.

Without Assuming Policy Convergence. We next give two structured classes of games in which the policy always converges and the selected equilibrium can be characterized more sharply. The first class is when there are two retained and two dominated actions:

Theorem 4.3 (Implicit Bias When $m = l = 2$)**.** *In addition to Assumption 1, suppose that* $m = l = 2$ *with* $\boldsymbol{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} > 0$*. Then, there is a Lebesgue-null set* $\mathcal{N}_{\mathrm{mix},2} \subset \mathbb{R}^4$ *such that*

- *For any* $\boldsymbol{\theta}(0) \in \mathcal{N}_{\mathrm{mix},2}$*, we have* $\boldsymbol{p}(t) \to \boldsymbol{q} = \left(\frac{\sqrt{bd}}{\sqrt{ac}+\sqrt{bd}}, \frac{\sqrt{ac}}{\sqrt{ac}+\sqrt{bd}}, 0, 0\right)$*;*
- *For any* $\boldsymbol{\theta}(0) \in \mathbb{R}^K \setminus \mathcal{N}_{\mathrm{mix},2}$*, we have* $\boldsymbol{p}(t) \to \boldsymbol{e}_{i^\star}$ *for some initialization-dependent* $i^\star \in \mathcal{I}$*.*

Corollary I.1 implies that a trajectory converging to a vertex satisfies $\|\boldsymbol{p}(t) - \boldsymbol{e}_{i^\star}\|_1 = \Theta(t^{-2/3})$. Unlike the rank-one case (described below), our proof characterizes $\mathcal{N}_{\mathrm{mix},2}$ only implicitly through the stable-manifold geometry of the full dynamics.

The second class is when the cross-interactions are rank-one:

Theorem 4.4 (Implicit Bias When $\mathrm{rank}(\boldsymbol{B}) = 1$)**.** *In addition to Assumption 1, suppose that* $\boldsymbol{B} = \boldsymbol{a}\boldsymbol{b}^\top$ *for some* $\boldsymbol{a} = (a_j)_{j\in\mathcal{L}} \in \mathbb{R}^l_{++}$ *and* $\boldsymbol{b} = (b_i)_{i\in\mathcal{I}} \in \mathbb{R}^m_{++}$*. Then, for any initialization* $\boldsymbol{\theta}(0) \in \mathbb{R}^K$*, we have* $\boldsymbol{p}(t) \to \boldsymbol{q}$*, with* $\boldsymbol{q}$ *characterized as follows:*

$$q_i = \frac{b_i^{-1}\mathbf{1}_{\{i\in\mathcal{M}(\boldsymbol{\theta}(0))\}}}{\sum_{k\in\mathcal{M}(\boldsymbol{\theta}(0))} b_k^{-1}} \quad (i \in \mathcal{I}), \qquad q_j = 0 \quad (j \in \mathcal{L}), \tag{8}$$

where $\mathcal{M}(\boldsymbol{\theta}(0)) := \arg\max_{i\in\mathcal{I}}\{\theta_i(0) + \log b_i\} \subseteq \mathcal{I}$ *is the initialization-dependent support set.*

Proof Sketch. The proof proceeds by directly showing that the ordering of $\exp(-\theta_i(0) - \log b_i)$ is preserved, i.e., only the initial maximizers remain; the full proof is provided in Appendix H.2. □

Thus a mixed limit occurs only on a finite union of affine hyperplanes, i.e., for Lebesgue-almost every initialization in $\mathbb{R}^K$, $\mathcal{M}(\boldsymbol{\theta}(0))$ is a singleton and the selected vertex is determined by the largest initial adjusted score $\theta_i(0) + \log b_i$. In particular, this implies that the attracting basin for $\boldsymbol{e}_i$ is exactly $\left\{\boldsymbol{\theta}(0) \in \mathbb{R}^K : \theta_i(0) + \log b_i > \theta_k(0) + \log b_k,\ \forall k \in \mathcal{I} \setminus \{i\}\right\}$. By **Corollary I.2**, we have that $\|\boldsymbol{p}(t) - \boldsymbol{q}\|_1 = 2\left(\sum_{k\in\mathcal{I}\setminus\mathcal{M}(\boldsymbol{\theta}(0))} p_k(t) + \rho(t)\right)$. Denoting $s = |\mathcal{M}(\boldsymbol{\theta}(0))|$, for $s < m$, the first sum has the slower rate $\|\boldsymbol{p}(t) - \boldsymbol{q}\|_1 = \Theta(t^{-l/(s+l)})$, whose exponent decreases as s grows. At $s = m$, this sum is empty, implying $\|\boldsymbol{p}(t) - \boldsymbol{q}\|_1 = \Theta(t^{-1})$.

Comparison to Chen et al. (2025). *With a full-support Nash equilibrium* and a sufficiently small step size, Chen et al. (2025, Lemmas 1–2 and Theorem 2) show that OMWU converges to the unique KL projection of its initial policy onto the Nash set, $\arg\min_{\boldsymbol{q}\in\mathcal{SNE}(\boldsymbol{A})} D_{\mathrm{KL}}(\boldsymbol{q}\|\boldsymbol{p}(0))$. This projection remains fixed along the auxiliary OMWU iterates.

Our flat-dominance results concern a boundary Nash face, where convergent (S2P-EG-Flow) trajectories in softmax logits select vertices for almost every initialization. Since the equilibrium assumptions differ, this comparison does not isolate the effect of optimization geometry. The distinct selection rules motivate studying how parameterization and game structure jointly determine implicit bias.

4.3 Proof Overview of Theorems 4.2 and 4.3

We now provide an overview of the proofs of Theorems 4.2 and 4.3, both of which follow a similar structure: derivation of the blow-up boundary system governing (S2P-EG-Flow) as $\rho(t) := \sum_{j\in\mathcal{L}} p_j(t) \to 0$ (where $\mathcal{L}$ is the set of *dominated* actions in the flat dominance assumption), and combining it with dynamical systems theory to show the final implicit bias. Due to space constraints, the full details are relegated to the Appendix: **Appendix B** collects the technical tools, **Appendix D** derives the boundary system, and **Appendix I** presents the implicit bias proofs sequentially.

By Theorem 3.2, $\rho(t) \to 0$, under which the policy dynamics slow down near the Nash face. To resolve this slowdown while retaining the relative proportions of the dominated actions, we "blow up" the boundary by setting

$$\tau(t) := \int_0^t \rho(s)\, ds, \qquad x_i(t) := \frac{p_i(t)}{1-\rho(t)}\ (i \in \mathcal{I}), \qquad \lambda_j(t) := \frac{p_j(t)}{\rho(t)}\ (j \in \mathcal{L}). \tag{9}$$

Then $(\rho, \boldsymbol{x}, \boldsymbol{\lambda}) \in (0,1) \times \Delta_{\mathcal{I}}^{\circ} \times \Delta_{\mathcal{L}}^{\circ}$ at finite times, and $\tau(t) \to \infty$ (**Lemma E.4**). Writing $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{\lambda})$, we have the following analytic extension to $\Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$ (**Proposition D.1**):

$$\frac{d\boldsymbol{z}}{d\tau} = \mathcal{B}(\boldsymbol{z}) + \rho E_\gamma(\rho, \boldsymbol{z}), \qquad \mathcal{B}: \begin{aligned} \frac{dx_i}{d\tau} &= x_i\left(x_i(\boldsymbol{B}^\top \boldsymbol{\lambda})_i - \langle \boldsymbol{x}, \boldsymbol{u}\rangle\right), \\ \frac{d\lambda_j}{d\tau} &= -\lambda_j\left(\lambda_j(\boldsymbol{B}\boldsymbol{x})_j + \langle \boldsymbol{\lambda}, \boldsymbol{w}\rangle\right), \end{aligned} \tag{10}$$

where $u_i := x_i(\boldsymbol{B}^\top \boldsymbol{\lambda})_i$, $w_j := -\lambda_j(\boldsymbol{B}\boldsymbol{x})_j$, and E_γ is polynomial. We refer to $\mathcal{B}$ as the ***boundary system***. Since $\rho(\tau) \to 0$, (10) is an *asymptotically autonomous perturbation* of $\mathcal{B}$. For a trajectory $\boldsymbol{y}$, we define its *omega-limit set*, $\omega(\boldsymbol{y})$, as the set of its subsequential limits as time tends to infinity. Applied to the transformed trajectory $\boldsymbol{z}$, the seminal results of Benaïm & Hirsch (1996); Mischaikow et al. (1995) (see **Theorem D.1**) show that $\omega(\boldsymbol{z})$ is nonempty, compact, and invariant under the boundary flow $\mathcal{B}$. Since $\boldsymbol{p} = ((1-\rho)\boldsymbol{x}, \rho\boldsymbol{\lambda})$, $\rho \to 0$, and $\tau(t) \to \infty$, compactness gives the following "blow-down" characterization of the omega-limit set of the trajectory $\boldsymbol{p}$ of (S2P-EG-Flow):

$$\omega(\boldsymbol{p}) = \left\{(\boldsymbol{x}, \boldsymbol{0}) : (\boldsymbol{x}, \boldsymbol{\lambda}) \in \omega(\boldsymbol{z}) \text{ for some } \boldsymbol{\lambda}\right\}.$$

Thus, it now suffices to characterize the omega-limit set of the boundary system $\mathcal{B}$!

For **Theorem 4.2**, assume $\boldsymbol{p}(t) \to \boldsymbol{q}$ and set $S := \operatorname{supp}(\boldsymbol{q})$ and $s := |S|$. We show that the dominated actions' normalized probabilities $\lambda_j = p_j/\rho$ also converge, so $(\boldsymbol{x}, \boldsymbol{\lambda})$ approaches a boundary equilibrium uniquely determined by S (**Lemma I.2(i),(ii)**). Thus each support admits at most one policy limit, despite the continuum of Nash equilibria. Keeping ρ as a state variable gives an exact autonomous system in $(\rho, \boldsymbol{x}, \boldsymbol{\lambda})$, with $s-1$ unstable and $K-s$ stable directions at the corresponding equilibrium. A mixed limit ($s \geq 2$) therefore attracts only a lower-dimensional stable set, whereas a vertex ($s = 1$) attracts a full-dimensional neighborhood. Arguments based on the Stable Manifold Theorem (**Theorem B.4 and Lemma B.1**) show that, in initial policy coordinates, mixed basins are Lebesgue-null and vertex basins are nonempty and open. To obtain the same conclusions for initial logits, write $\boldsymbol{\theta}(0) = \log \boldsymbol{p}(0) + c\boldsymbol{1}$: including the free common shift $c \in \mathbb{R}$ gives a smooth, invertible change of coordinates that preserves nullity and openness.

For **Theorem 4.3**, write $x = p_1/(p_1+p_2)$ and $y = p_3/(p_3+p_4)$. The boundary dynamics at $\rho = 0$ form a planar system on $[0,1]^2$. For all sufficiently large times, entering a sufficiently small neighborhood of $x = 0$ or $x = 1$ forces x to converge to that endpoint. Since $\rho \to 0$, these cases give $\boldsymbol{p} \to \boldsymbol{e}_2$ and $\boldsymbol{p} \to \boldsymbol{e}_1$, respectively. Otherwise, x remains bounded away from both endpoints, as does y (**Lemma I.1**). The trajectory therefore remains in a compact rectangle contained in $(0,1)^2$. We can show that its omega-limit set is nonempty, compact, and internally chain transitive (ICT) for the boundary flow (**Theorem D.1**). This flow is complete on $(0,1)^2$, and its only equilibrium there is a hyperbolic saddle (**Lemma I.2(i),(iii)**). Planar index theory excludes periodic orbits and homoclinic loops, so the planar ICT classification (Benaïm & Hirsch, 1996) implies that the limit set consists only of this saddle (**Corollary B.9**). Thus every policy converges, which gives vertex selection for almost every initialization (**Theorem 4.2**).

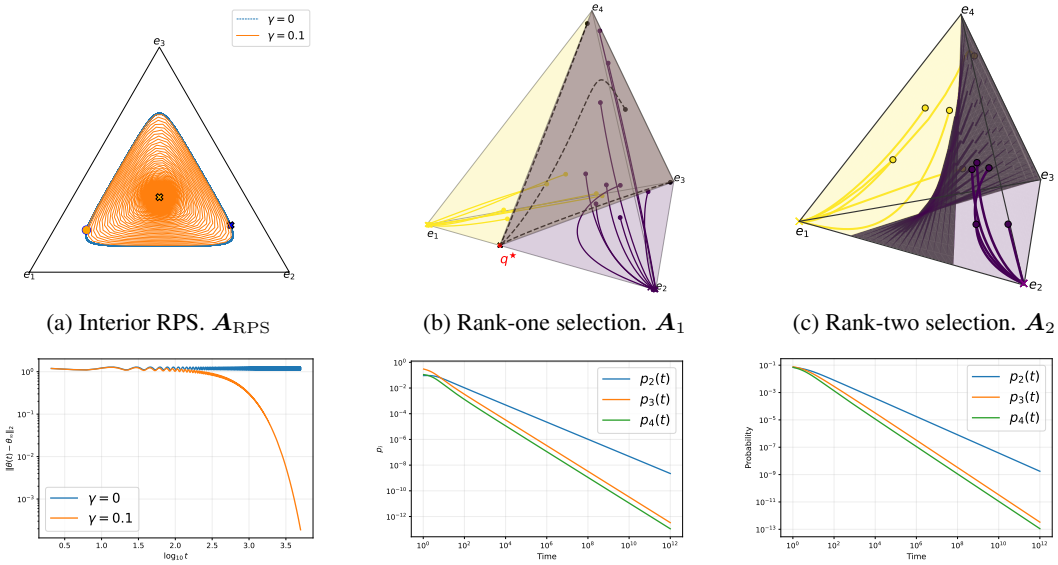
5 Toy Experiments

We illustrate our theoretical results with three toy examples (see **Appendix L** for additional results):

$$\boldsymbol{A}_{\mathrm{RPS}} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{A}_1 = \begin{pmatrix} 0 & 0 & -1 & -2 \\ 0 & 0 & -2 & -4 \\ 1 & 2 & 0 & 1 \\ 2 & 4 & -1 & 0 \end{pmatrix}, \quad \boldsymbol{A}_2 = \begin{pmatrix} 0 & 0 & -1 & -3 \\ 0 & 0 & -4 & -1 \\ 1 & 4 & 0 & 1/2 \\ 3 & 1 & -1/2 & 0 \end{pmatrix},$$

Convergence to an interior equilibrium. For the rock–paper–scissors game $\boldsymbol{A}_{\mathrm{RPS}}$, there is a unique interior SNE $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. **Theorem 3.1** predicts finite-logit convergence for every $\gamma > 0$. In Figure 1a, the first-order flow ($\gamma = 0$) keeps rotating, whereas S2P-EG-Flow with $\gamma = 0.1$ spirals toward the interior SNE. The logit error in Figure 1d also decays exponentially, as predicted.

Selection under flat dominance. Let $\mathcal{I} = \{1, 2\}$ and $\mathcal{L} = \{3, 4\}$. We consider $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ with $\gamma = 0.1$, both of which satisfy Assumption 1. **Theorem 3.2** implies that every trajectory converges to $\Delta_{\mathcal{I}}$. For $\boldsymbol{A}_1$, $B_1 = (\boldsymbol{A}_1)_{\mathcal{L},\mathcal{I}} = (1, 2)^\top (1, 2)$ has rank one. **Theorem 4.4** separates the two vertex basins at $\theta_1(0) - \theta_2(0) = \log 2$; equality yields $\boldsymbol{q}^\star = (2/3, 1/3, 0, 0)$, as shown in Figure 1b. Along a trajectory converging to $\boldsymbol{e}_1$, we have $p_2 = \Theta(t^{-2/3})$, $p_3 = (3t)^{-1}(1 + o(1))$, and $p_4 = (6t)^{-1}(1 + o(1))$ (Figure 1e), aligning with **Theorem 3.3**. For $\boldsymbol{A}_2$, $B_2 = (\boldsymbol{A}_2)_{\mathcal{L},\mathcal{I}} = \left(\begin{smallmatrix} 1 & 4 \\ 3 & 1 \end{smallmatrix}\right)$ has rank two. **Theorem 4.3** predicts selection of $\boldsymbol{e}_1$ or $\boldsymbol{e}_2$ almost everywhere (Figure 1c), but the characterization of such null set is not available; for the figure, we consider finer grids of initializations over the simplex. We note that the null set in this case resembles a ruled surface (do Carmo, 2016, Chapter 3.5A), but we do not prove this rigorously. Along a trajectory converging to $\boldsymbol{e}_1$, we have $p_2 = \Theta(t^{-2/3})$, $p_3 = (3t)^{-1}(1 + o(1))$, and $p_4 = (9t)^{-1}(1 + o(1))$ (Figure 1f).



(a) Interior RPS. $\boldsymbol{A}_{\mathrm{RPS}}$ (b) Rank-one selection. $\boldsymbol{A}_1$ (c) Rank-two selection. $\boldsymbol{A}_2$

(d) Interior RPS logit error. (e) Rank-one boundary rates. (f) Rank-two boundary rates.

Figure 1: **Toy experiments.** Top row: trajectories and basins. Circles/crosses mark initial/final policies. (a) RPS: first-order flow rotates, while S2P-EG-Flow converges to the interior SNE. (b) $\boldsymbol{A}_1$: the black separating set converges to $(2/3, 1/3, 0, 0)$. (c) $\boldsymbol{A}_2$: yellow/purple regions select $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. Bottom row: convergence rates. (d) RPS logit error decays exponentially for $\gamma = 0.1$. (e) Rates along a trajectory to $\boldsymbol{e}_1$ in $\boldsymbol{A}_1$. (f) Rates along a trajectory to $\boldsymbol{e}_1$ in $\boldsymbol{A}_2$.

6 Conclusion

We study the high-resolution modified flow of Euclidean extragradient in softmax logits for bilinear self-play zero-sum games and identify two regimes with global convergence guarantees. If an interior SNE exists, every trajectory converges to a finite equilibrium in logit space at an eventual exponential rate. Under flat dominance (Assumption 1), every trajectory converges in the Nash sense, with the dominated probability mass decaying as $\Theta(t^{-1})$. Under the summed-logit condition, we derive sharp t^{-1} asymptotics for positive-margin actions, while zero-margin actions can exhibit arbitrarily slow polynomial decay. Under flat dominance, the dynamics also exhibit a strong implicit bias on Nash faces: conditional on policy convergence, almost every initialization selects a vertex. Furthermore, we establish policy convergence for two structured classes of games. Together, these results show that despite the nonconvex–nonconcave geometry, softmax parametrized bilinear game admits a precise analysis, revealing how logit-space optimization shapes convergence rates and equilibrium selection. Our work leaves several future directions, including extensions beyond symmetric games, characterization of implicit bias beyond flat dominance, and extending the continuous-time results to discrete-time EG updates.

AI USE STATEMENT

In this work, we used generative AI tools to support research and writing tasks, including refining mathematical formulations and hypotheses, suggesting proof strategies, checking derivations, providing feedback on methodology and experimental design, assisting with result interpretation, retrieving related work, preparing code for experiments and expository figures, and drafting, editing, and translating portions of the manuscript. All AI-assisted research content was critically evaluated by the authors. Mathematical claims and proofs were independently checked, references and related-work suggestions were validated using the original sources, and experimental and interpretive suggestions were evaluated against our own analyses and results. All AI-assisted text was reviewed and revised by the authors. We did not use generative AI to clean or reformat datasets, or perform qualitative or thematic data analysis. We take responsibility for the final content of this work, including all AI-assisted text, claims, results, and artifacts.

REFERENCES

Alekh Agarwal, Sham M. Kakade, Jason D. Lee, and Gaurav Mahajan. On the Theory of Policy Gradient Methods: Optimality, Approximation, and Distribution Shift. *Journal of Machine Learning Research*, 22(98):1–76, 2021. URL http://jmlr.org/papers/v22/19-736.html.

Sanjeev Arora, Nadav Cohen, Wei Hu, and Yuping Luo. Implicit Regularization in Deep Matrix Factorization. In *Advances in Neural Information Processing Systems*, volume 32, pp. 7413–7424. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/hash/c0c783b5fc0d7d808f1d14a6e9c8280d-Abstract.html.

Mohammad Gheshlaghi Azar, Zhaohan Daniel Guo, Bilal Piot, Rémi Munos, Mark Rowland, Michal Valko, and Daniele Calandriello. A General Theoretical Paradigm to Understand Learning from Human Preferences. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 4447–4455. PMLR, 02–04 May 2024. URL https://proceedings.mlr.press/v238/gheshlaghi-azar24a.html.

Yu Bai and Chi Jin. Provable Self-Play Algorithms for Competitive Reinforcement Learning. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 551–560. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/bai20a.html.

Yu Bai, Chi Jin, and Tiancheng Yu. Near-Optimal Reinforcement Learning with Self-Play. In *Advances in Neural Information Processing Systems*, volume 33, pp. 2159–2170. Curran Associates, Inc., 2020. URL https://arxiv.org/abs/2006.12007.

Michel Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pp. 1–68. Springer, 1999.

Michel Benaïm and Morris W. Hirsch. Asymptotic Pseudotrajectories and Chain Recurrent Flows, with Applications. *Journal of Dynamics and Differential Equations*, 8(1):141–176, 1996. doi:10.1007/BF02218617.

Martino Bernasconi, Federico Cacciamani, Simone Fioravanti, Nicola Gatti, and Francesco Trovò. The evolutionary dynamics of soft-max policy gradient in multi-agent settings. *Theoretical Computer Science*, 1027:115011, 2025. ISSN 0304-3975. doi:10.1016/j.tcs.2024.115011.

Shicong Cen, Yuting Wei, and Yuejie Chi. Fast Policy Extragradient Methods for Competitive Games with Entropy Regularization. *Journal of Machine Learning Research*, 25(4):1–48, 2024. URL http://jmlr.org/papers/v25/21-1205.html.

Jiseok Chae, Kyuwon Kim, and Donghwan Kim. Two-timescale Extragradient for Finding Local Minimax Points. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=6CIGhcJYJH.

Shulun Chen, Runlong Zhou, Zihan Zhang, Maryam Fazel, and Simon S. Du. Unregularized Linear Convergence in Zero-Sum Game from Preference Feedback. *arXiv preprint arXiv:2512.24818*, 2025. URL https://arxiv.org/abs/2512.24818.

Lénaïc Chizat and Francis Bach. Implicit Bias of Gradient Descent for Wide Two-layer Neural Networks Trained with the Logistic Loss. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pp. 1305–1338. PMLR, 09–12 Jul 2020. URL https://proceedings.mlr.press/v125/chizat20a.html.

Constantinos Daskalakis and Ioannis Panageas. Last-Iterate Convergence: Zero-Sum Games and Constrained Min-Max Optimization. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 27:1–27:18, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-095-8. doi:10.4230/LIPIcs.ITCS.2019.27.

Constantinos Daskalakis, Andrew Ilyas, Vasilis Syrgkanis, and Haoyang Zeng. Training GANs with Optimism. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=SJJySbbAZ.

Jelena Diakonikolas, Constantinos Daskalakis, and Michael I. Jordan. Efficient Methods for Structured Nonconvex-Nonconcave Min-Max Optimization. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 2746–2754. PMLR, 2021. URL https://proceedings.mlr.press/v130/diakonikolas21a.html.

Duong The Do, Matthew Aldridge, and Duong Tung Nguyen. Projected Variational Quantum Extragradient for Zero-Sum Games. *IEEE Control Systems Letters*, 10:2041–2046, 2026. doi:10.1109/LCSYS.2026.3710797.

Manfredo P. do Carmo. *Differential Forms and Applications*. Universitext. Springer-Verlag, Berlin and Heidelberg, 1994.

Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces*. Dover Books on Mathematics. Dover Publications, Mineola, NY, 2 edition, 2016.

Ryan D'Orazio, Danilo Vucetic, Zichu Liu, Junhyung Lyle Kim, Ioannis Mitliagkas, and Gauthier Gidel. Solving hidden monotone variational inequalities with surrogate losses. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=4ZX2a3OKEV.

Yi Feng, Kaito Fujii, Stratis Skoulakis, Xiao Wang, and Volkan Cevher. Continuous-Time Analysis of Heavy Ball Momentum in Min-Max Games. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 16670–16710. PMLR, 13–19 Jul 2025. URL https://openreview.net/forum?id=7vdg7VICsX.

Roy Fox, Stephen M. Mcaleer, Will Overman, and Ioannis Panageas. Independent Natural Policy Gradient always converges in Markov Potential Games. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 4414–4425. PMLR, 28–30 Mar 2022. URL https://proceedings.mlr.press/v151/fox22a.html.

Ian Gemp and Sridhar Mahadevan. Global Convergence to the Equilibrium of GANs using Variational Inequalities. *arXiv preprint arXiv:1808.01531*, 2018. URL https://arxiv.org/abs/1808.01531.

Gauthier Gidel, David Balduzzi, Wojciech Czarnecki, Marta Garnelo, and Yoram Bachrach. A Limited-Capacity Minimax Theorem for Non-Convex Games or: How I Learned to Stop Worrying about Mixed-Nash and Love Neural Nets. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 2548–2556. PMLR, 13–15 Apr 2021. URL https://proceedings.mlr.press/v130/gidel21a.html.

Alan J. Goldman and Albert W. Tucker. Theory of Linear Programming. In Harold W. Kuhn and Albert W. Tucker (eds.), *Linear Inequalities and Related Systems*, number 38 in Annals of Mathematics Studies, pp. 53–97. Princeton University Press, Princeton, NJ, 1956.

Benjamin Grimmer, Haihao Lu, Pratik Worah, and Vahab Mirrokni. Limiting Behaviors of Nonconvex-Nonconcave Minimax Optimization via Continuous-Time Systems. In *Proceedings of The 33rd International Conference on Algorithmic Learning Theory*, volume 167 of *Proceedings of Machine Learning Research*, pp. 465–487. PMLR, 2022. URL https://proceedings.mlr.press/v167/grimmer22a.html.

Suriya Gunasekar, Blake E Woodworth, Srinadh Bhojanapalli, Behnam Neyshabur, and Nati Srebro. Implicit Regularization in Matrix Factorization. In *Advances in Neural Information Processing Systems*, volume 30, pp. 6152–6160. Curran Associates, Inc., 2017. URL https://papers.nips.cc/paper_files/paper/2017/hash/58191d2a914c6dae66371c9dcdc91b41-Abstract.html.

Suriya Gunasekar, Jason Lee, Daniel Soudry, and Nathan Srebro. Characterizing Implicit Bias in Terms of Optimization Geometry. In Jennifer Dy and Andreas Krause (eds.), *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 1832–1841. PMLR, 10–15 Jul 2018. URL https://proceedings.mlr.press/v80/gunasekar18a.html.

John C. Harsanyi. Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points. *International Journal of Game Theory*, 2(1):1–23, 1973. doi:10.1007/BF01737554.

Daniel Hennes, Dustin Morrill, Shayegan Omidshafiei, Rémi Munos, Julien Perolat, Marc Lanctot, Audrunas Gruslys, Jean-Baptiste Lespiau, Paavo Parmas, Edgar Duèñez-Guzmán, and Karl Tuyls. Neural Replicator Dynamics: Multiagent Learning via Hedging Policy Gradients. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '20, pp. 492–501, Richland, SC, 2020. International Foundation for Autonomous Agents and Multiagent Systems. URL https://dl.acm.org/doi/10.5555/3398761.3398822.

Morris W. Hirsch, Charles C. Pugh, and Michael Shub. *Invariant Manifolds*, volume 583 of *Lecture Notes in Mathematics*. Springer, 1977.

Ya-Ping Hsieh, Panayotis Mertikopoulos, and Volkan Cevher. The Limits of Min-Max Optimization Algorithms: Convergence to Spurious Non-Critical Sets. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 4337–4348. PMLR, 2021. URL https://proceedings.mlr.press/v139/hsieh21a.html.

Chi Jin, Praneeth Netrapalli, and Michael Jordan. What is Local Optimality in Nonconvex-Nonconcave Minimax Optimization? In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 4880–4889. PMLR, 2020. URL https://proceedings.mlr.press/v119/jin20e.html.

Fivos Kalogiannis, Jingming Yan, and Ioannis Panageas. Learning Equilibria in Adversarial Team Markov Games: A Nonconvex-Hidden-Concave Min-Max Optimization Problem. In *Advances in Neural Information Processing Systems*, volume 37, pp. 92832–92890. Curran Associates, Inc., 2024. URL https://openreview.net/forum?id=BrvLTxEx08.

Galina M. Korpelevich. The Extragradient Method for Finding Saddle Points and Other Problems. *Matekon*, 13(4):35–49, 1977.

Yassine Laguel, Yasa Syed, Necdet Serhat Aybat, and Mert Gürbüzbalaban. High-probability complexity bounds for stochastic non-convex minimax optimization. In *Advances in Neural Information Processing Systems*, volume 37, pp. 140937–140969. Curran Associates, Inc., 2024. URL https://openreview.net/forum?id=XMQTNzlgTJ.

J. P. LaSalle. Some Extensions of Liapunov's Second Method. *IRE Transactions on Circuit Theory*, 7(4):520–527, 1960. doi:10.1109/TCT.1960.1086720.

J. P. LaSalle. Stability Theory for Ordinary Differential Equations. *Journal of Differential Equations*, 4(1):57–65, 1968. doi:10.1016/0022-0396(68)90048-X.

Luis Leal. Which Nash Equilibrium? Solver-Dependent Selection on Zero-Sum Nash Polytopes. *arXiv preprint arXiv:2606.28308*, 2026. URL https://arxiv.org/abs/2606.28308.

Hyunwoo Lee, Jeongyeol Kwon, and Dohyun Kwon. Continuous-Time Analysis for Minimax and Bilevel Problems. *arXiv preprint arXiv:2605.20898*, 2026a. URL https://arxiv.org/abs/2605.20898.

John M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer New York, 2 edition, 2012.

Junghyun Lee, Minju Hong, Kwang-Sung Jun, Chulhee Yun, and Se-Young Yun. Provably Efficient Regularized Online RLHF with Generalized Bilinear Preferences. *arXiv preprint arXiv:2602.23116*, 2026b. URL https://arxiv.org/abs/2602.23116.

Kyungjae Lee and Donghwan Kim. Zero-Infinity GAN: Stable Dynamics and Implicit Bias of Extragradient. In *OPT 2025: Optimization for Machine Learning*, 2025. URL https://openreview.net/forum?id=H2da7FTe3W.

Chin-wing Leung, Shuyue Hu, and Ho-fung Leung. The Stochastic Evolutionary Dynamics of Softmax Policy Gradient in Games. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '24, pp. 1101–1109, Richland, SC, 2024. International Foundation for Autonomous Agents and Multiagent Systems. URL https://dl.acm.org/doi/10.5555/3635637.3662966.

Gen Li, Yuting Wei, Yuejie Chi, Yuantao Gu, and Yuxin Chen. Softmax Policy Gradient Methods Can Take Exponential Time to Converge. In *Proceedings of Thirty Fourth Conference on Learning Theory*, volume 134 of *Proceedings of Machine Learning Research*, pp. 3107–3110. PMLR, 15–19 Aug 2021. URL https://proceedings.mlr.press/v134/li21b.html.

Gen Li, Yuting Wei, Yuejie Chi, and Yuxin Chen. Softmax policy gradient methods can take exponential time to converge. *Mathematical Programming*, 201(1):707–802, 2023. doi:10.1007/s10107-022-01920-6.

Tengyuan Liang and James Stokes. Interaction Matters: A Note on Non-asymptotic Local Convergence of Generative Adversarial Networks. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 907–915. PMLR, 16–18 Apr 2019. URL https://proceedings.mlr.press/v89/liang19b.html.

Tianyi Lin, Chi Jin, and Michael Jordan. On Gradient Descent Ascent for Nonconvex-Concave Minimax Problems. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6083–6093. PMLR, 2020. URL https://proceedings.mlr.press/v119/lin20a.html.

Qinghua Liu, Tiancheng Yu, Yu Bai, and Chi Jin. A Sharp Analysis of Model-based Reinforcement Learning with Self-Play. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 7001–7010. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/liu21z.html.

Haihao Lu. An $O(s^r)$-resolution ODE framework for understanding discrete-time algorithms and applications to the linear convergence of minimax problems. *Mathematical Programming*, 194(1): 1061–1112, 2022. doi:10.1007/s10107-021-01669-4.

Aleksandr Mikhailovich Lyapunov. *Obshchaya zadacha ob ustoychivosti dvizheniya [The General Problem of the Stability of Motion]*. Kharkov Mathematical Society, Kharkov, 1892. Doctoral dissertation, in Russian.

Jincheng Mei, Chenjun Xiao, Csaba Szepesvári, and Dale Schuurmans. On the Global Convergence Rates of Softmax Policy Gradient Methods. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6820–6829. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/mei20b.html.

Jincheng Mei, Yue Gao, Bo Dai, Csaba Szepesvári, and Dale Schuurmans. Leveraging Non-uniformity in First-order Non-convex Optimization. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 7555–7564. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/mei21a.html.

Panayotis Mertikopoulos and William H. Sandholm. Learning in Games via Reinforcement and Regularization. *Mathematics of Operations Research*, 41(4):1297–1324, 2016. doi:10.1287/moor.2016.0778.

Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the twenty-ninth annual ACM-SIAM symposium on discrete algorithms*, pp. 2703–2717. SIAM, 2018.

Panayotis Mertikopoulos, Bruno Lecouat, Houssam Zenati, Chuan-Sheng Foo, Vijay Chandrasekhar, and Georgios Piliouras. Optimistic mirror descent in saddle-point problems: Going the extra(-gradient) mile. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=Bkg8jjC9KQ.

Konstantin Mischaikow, Hal Smith, and Horst R. Thieme. Asymptotically Autonomous Semiflows: Chain Recurrence and Lyapunov Functions. *Transactions of the American Mathematical Society*, 347(5):1669–1685, 1995. doi:10.1090/S0002-9947-1995-1290727-7.

Andjela Mladenovic, Iosif Sakos, Gauthier Gidel, and Georgios Piliouras. Generalized Natural Gradient Flows in Hidden Convex-Concave Games and GANs. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=bsycpMi00R1.

Aryan Mokhtari, Asuman Ozdaglar, and Sarath Pattathil. A Unified Analysis of Extra-gradient and Optimistic Gradient Methods for Saddle Point Problems: Proximal Point Approach. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 1497–1507. PMLR, 26–28 Aug 2020. URL https://proceedings.mlr.press/v108/mokhtari20a.html.

Rémi Munos, Michal Valko, Daniele Calandriello, Mohammad Gheshlaghi Azar, Mark Rowland, Zhaohan Daniel Guo, Yunhao Tang, Matthieu Geist, Thomas Mesnard, Côme Fiegel, Andrea Michi, Marco Selvi, Sertan Girgin, Nikola Momchev, Olivier Bachem, Daniel J Mankowitz, Doina Precup, and Bilal Piot. Nash Learning from Human Feedback. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 36743–36768. PMLR, 21–27 Jul 2024. URL https://proceedings.mlr.press/v235/munos24a.html.

Arkadi Nemirovski. Prox-Method with Rate of Convergence O(1/t) for Variational Inequalities with Lipschitz Continuous Monotone Operators and Smooth Convex-Concave Saddle Point Problems. *SIAM Journal on Optimization*, 15(1):229–251, 2004. doi:10.1137/S1052623403425629.

Maher Nouiehed, Maziar Sanjabi, Tianjian Huang, Jason D. Lee, and Meisam Razaviyayn. Solving a Class of Non-Convex Min-Max Games Using Iterative First Order Methods. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/hash/25048eb6a33209cb5a815bff0cf6887c-Abstract.html.

Deep Patel and Emmanouil-Vasileios Vlatakis-Gkaragkounis. Solving Neural Min-Max Games: The Role of Architecture, Initialization & Dynamics. In *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 66297–66351. Curran Associates, Inc., 2025. URL https://openreview.net/forum?id=5xdbWUdM87.

Sarath Pattathil, Kaiqing Zhang, and Asuman Ozdaglar. Symmetric (Optimistic) Natural Policy Gradient for Multi-Agent Learning with Parameter Convergence. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 5641–5685. PMLR, 25–27 Apr 2023. URL https://proceedings.mlr.press/v206/pattathil23a.html.

Lawrence Perko. *Differential Equations and Dynamical Systems*, volume 7 of *Texts in Applied Mathematics*. Springer, 3 edition, 2001.

Boris T. Polyak. Some Methods of Speeding Up the Convergence of Iteration Methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964. doi:10.1016/0041-5553(64)90137-5.

Jan Prüss, Gieri Simonett, and Rico Zacher. On convergence of solutions to equilibria for quasilinear parabolic problems. *Journal of Differential Equations*, 246(10):3902–3931, 2009. doi:10.1016/j.jde.2008.10.034.

Mansi Rankawat, Michael Muehlebach, Simon Lacoste-Julien, and Damien Scieur. From Lyapunov Analysis to Algorithm Design in Two-Sided PL Minimax Optimization. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=UIyXPhhxlf.

Paul Reverdy and Naomi Ehrich Leonard. Parameter Estimation in Softmax Decision-Making Models With Linear Objective Functions. *IEEE Transactions on Automation Science and Engineering*, 13 (1):54–67, 2016. doi:10.1109/TASE.2015.2499244.

Iosif Sakos, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Panayotis Mertikopoulos, and Georgios Piliouras. Exploiting Hidden Structures in Non-Convex Games for Convergence to Nash Equilibrium. In *Advances in Neural Information Processing Systems*, volume 36, pp. 66979–67006. Curran Associates, Inc., 2023. URL https://openreview.net/forum?id=05P1U0jk8r.

Eugene Shcherbinin, Arina Redina, Maxim Kalpin, and Vlad Kochetov. Equilibrium Selection in Multi-Agent Policy Gradients via Opponent-Aware Basin Entry. In *New Frontiers in Game-Theoretic Learning - NExT-Game*, 2026. URL https://openreview.net/forum?id=gxrPZFFD9h.

Daniel Soudry, Elad Hoffer, Mor Shpigel Nacson, Suriya Gunasekar, and Nathan Srebro. The Implicit Bias of Gradient Descent on Separable Data. *Journal of Machine Learning Research*, 19(70):1–57, 2018. URL http://jmlr.org/papers/v19/18-188.html.

Jaewook Suh, Jisun Park, and Ernest Ryu. Continuous-time Analysis of Anchor Acceleration. In *Advances in Neural Information Processing Systems*, volume 36, pp. 32782–32866. Curran Associates, Inc., 2023. URL https://openreview.net/forum?id=rN99gLCBe4.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2 edition, 2018.

Gokul Swamy, Christoph Dann, Rahul Kidambi, Steven Wu, and Alekh Agarwal. A Minimaximalist Approach to Reinforcement Learning from Human Feedback. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 47345–47377. PMLR, 21–27 Jul 2024. URL https://proceedings.mlr.press/v235/swamy24a.html.

Gerald Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012.

Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, and Georgios Piliouras. Poincaré Recurrence, Cycles and Spurious Equilibria in Gradient-Descent-Ascent for Non-Convex Non-Concave Zero-Sum Games. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/hash/6c7cd904122e623ce625613d6af337c4-Abstract.html.

Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, and Georgios Piliouras. Solving Min-Max Optimization with Hidden Structure via Gradient Descent Ascent. In *Advances in Neural Information Processing Systems*, volume 34, pp. 2373–2386. Curran Associates, Inc., 2021. URL https://openreview.net/forum?id=Efqe8E4Bww.

J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, Dec 1928. ISSN 1432-1807. doi:10.1007/BF01448847.

Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Linear Last-iterate Convergence in Constrained Saddle-point Optimization. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=dx11_7vm5_r.

Stephen Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, volume 2 of *Texts in Applied Mathematics*. Springer, 2 edition, 2003.

Blake Woodworth, Suriya Gunasekar, Jason D. Lee, Edward Moroshko, Pedro Savarese, Itay Golan, Daniel Soudry, and Nathan Srebro. Kernel and Rich Regimes in Overparametrized Models. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pp. 3635–3673. PMLR, 09–12 Jul 2020. URL https://proceedings.mlr.press/v125/woodworth20a.html.

Di Wu, Chengshuai Shi, Jing Yang, and Cong Shen. Greedy Sampling Is Provably Efficient for RLHF. In *Advances in Neural Information Processing Systems*, volume 38, pp. 119785–119819. Curran Associates, Inc., 2025. URL https://openreview.net/forum?id=jQH0gdwsuT.

Wei Xiong, Han Zhong, Chengshuai Shi, Cong Shen, and Tong Zhang. A Self-Play Posterior Sampling Algorithm for Zero-Sum Markov Games. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 24496–24523. PMLR, 17–23 Jul 2022. URL https://proceedings.mlr.press/v162/xiong22b.html.

Chenlu Ye, Wei Xiong, Yuheng Zhang, Hanze Dong, Nan Jiang, and Tong Zhang. Online iterative reinforcement learning from human feedback with general preference model. In *Advances in Neural Information Processing Systems*, volume 37, pp. 81773–81807. Curran Associates, Inc., 2024. doi:10.52202/079017-2598. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/94d13c2401fe119e57ba325b6fe526e0-Paper-Conference.pdf.

Runyu Zhang, Jincheng Mei, Bo Dai, Dale Schuurmans, and Na Li. On the Global Convergence Rates of Decentralized Softmax Gradient Play in Markov Potential Games. In *Advances in Neural Information Processing Systems*, volume 35, pp. 1923–1935. Curran Associates, Inc., 2022. URL https://openreview.net/forum?id=X1oVDZIABwF.

Yuheng Zhang, Dian Yu, Tao Ge, Linfeng Song, Zhichen Zeng, Haitao Mi, Nan Jiang, and Dong Yu. Improving LLM General Preference Alignment via Optimistic Online Mirror Descent. In *Advances in Neural Information Processing Systems*, volume 38. Curran Associates, Inc., 2025a. URL https://openreview.net/forum?id=kZstGANG8D.

Yuheng Zhang, Dian Yu, Baolin Peng, Linfeng Song, Ye Tian, Mingyue Huo, Nan Jiang, Haitao Mi, and Dong Yu. Iterative Nash Policy Optimization: Aligning LLMs with General Preferences via No-Regret Learning. In *The Thirteenth International Conference on Learning Representations*, 2025b. URL https://openreview.net/forum?id=Pujt3ADZgI.

CONTENTS

1 Introduction **1**

2 Problem Setup **3**

3 Global Convergence Behavior of the S2P-EG-Flow **4**

3.1 Two Sufficient Conditions for Global Convergence 4

3.2 Sharp Boundary Asymptotics 5

4 Implicit Bias of the S2P-EG Flow **6**

4.1 Every Policy Limit Is a SNE 6

4.2 Implicit Bias under Flat Dominance 6

4.3 Proof Overview of Theorems 4.2 and 4.3 7

5 Toy Experiments **8**

6 Conclusion **9**

A Detailed Comparison with Prior Work **19**

A.1 Hidden Games and Unmodified Parameter-Space Dynamics 19

A.2 Softmax Game Dynamics and Extragradient 19

A.3 Geometry-Aware and Regularized Methods 20

A.4 Implicit Bias and Equilibrium Selection 20

B Dynamical Systems Toolkit **21**

B.1 Solutions, Evolution Maps, and Limit Sets 21

B.2 Lyapunov Functions, Invariance, and Attraction 22

B.3 Local Behavior Near Equilibria 23

B.4 Asymptotic Pseudotrajectories and Chain-Transitive Limit Sets 24

B.5 Planar Dynamics and Index Theory 25

B.6 Deferred Proofs of “Folklore” Results 26

B.6.1 Proof of Lemma B.1: Nullity of a Hyperbolic Stable Set 26

B.6.2 Proof of Theorem B.5: Local Asymptotic Phase 27

B.6.3 Proof of Corollary B.9: A Unique Saddle in a Rectangle 28

C From Extragradient Method to S2P-EG-Flow. **30**

D Boundary System Limit of S2PE-EG-Flow **32**

D.1 Boundary Reduction and Limit Sets 32

D.2 Proof of Lemma D.1: Exact Blow-Up System 34

D.3 Proof of Proposition D.1: Analytic Extension of Blow-Up System 35

E Key Technical Lemmas **36**

F Proofs for Section 3.1: Global Convergence Behavior of the S2P-EG-Flow **39**

F.1 Proof of Theorem 3.1: Interior SNE ⇒ Logit Convergence 39

F.2 Proof of Theorem 3.2: Flat Dominance ⇒ Nash Convergence 41

G Proofs for Section 3.2: Sharp Boundary Asymptotics **43**

G.1 Full boundary-rate reduction 43

G.2 Quasi-strict boundary limits 46

G.3 A pure limit with one zero-margin boundary action 46

G.4 Multiple zero-margin boundary actions: nonuniversal exponents 49

H "Direct" Proofs of Two Implicit Bias Results **53**

H.1 Proof of Theorem 4.1: Every Policy Limit Is a SNE 53

H.2 Proof of Theorem 4.4: Rank-One Cross-Interactions 54

I Implicit Bias under Flat Dominance **56**

I.1 Two Key Technical Lemmas 56

I.2 Proof of Theorem 4.2: Conditional Vertex Selection 57

I.3 Proof of Theorem 4.3: The 2×2 Case 58

I.4 Convergence Rates for the Selected Limits 60

I.5 Proofs of the Technical Lemmas 61

I.5.1 Proof of Lemma I.1: Boundary Estimates 61

I.5.2 Proof of Lemma I.2: Balanced Equilibria 62

J Nonconvergence Results of S2PE Flow **65**

K Geometry of Symmetric Games **69**

L Additional Experiments **72**

L.1 Numerical Experiments for Non-Quasi-Strict Equilibria 72

L.2 Numerical Experiments for Nonconvergence Results 74

A Detailed Comparison with Prior Work

Here we compare our results in detail with prior work. A broad class of nonconvex–nonconcave games arises by composing a structured latent game with a nonlinear parametrization:

$$\min_{\boldsymbol{\theta}} \max_{\boldsymbol{\phi}} \; L\big(F(\boldsymbol{\theta}), G(\boldsymbol{\phi})\big), \tag{11}$$

where $L(\cdot,\cdot)$ may be convex–concave or monotone even though the objective in parameter space $(\boldsymbol{\theta}, \boldsymbol{\phi})$ may not be. Prior work asks whether the represented model class admits an approximate equilibrium (Gidel et al., 2021), analyzes the dynamics of fixed parameter-space algorithms (Vlatakis-Gkaragkounis et al., 2019; 2021), or designs geometry-aware or surrogate updates that exploit the latent structure (D'Orazio et al., 2025; Mladenovic et al., 2022; Sakos et al., 2023). We focus on the second question for a fixed algorithm: Euclidean-logit extragradient under tied softmax self-play.

The main distinction of our setting appears at the simplex boundary: a boundary policy has no finite softmax representation, so convergence requires diverging logits. Much of the related literature either avoids this regime through regularity assumptions or changes the geometry or objective so that the target remains well conditioned.

A.1 Hidden Games and Unmodified Parameter-Space Dynamics

The closest conceptual starting point is the literature on hidden bilinear and hidden convex–concave games. Vlatakis-Gkaragkounis et al. (2019) show that continuous-time GDA after nonlinear coordinate-separable representations of a bilinear zero-sum game can be conservative rather than convergent: safe trajectories may be periodic or Poincaré recurrent, while other representations admit spurious stationary points. Vlatakis-Gkaragkounis et al. (2021) extend this framework to hidden convex–concave objectives and give nonlocal convergence guarantees for vanilla GDA under structural assumptions on the latent game and representation. Strict latent convexity–concavity yields convergence to a von Neumann solution, whereas under mere convexity–concavity the Lyapunov argument gives stability but not convergence; when strictness is absent, they introduce regularization and obtain convergence to a nearby perturbed solution. Their analysis uses initialization-dependent Lyapunov functions and a reachability condition formalized through *safety*.

Our result also uses Lyapunov argument, but through a different mechanism. For an interior SNE $\boldsymbol{q}$,

$$\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}) = \sum_{i=1}^{K} q_i e^{-\theta_i}$$

together with logit-sum conservation gives compactness for every finite initialization, without safety conditions (Lemmas E.1 and F.1). The latent payoff remains bilinear and unregularized; the $O(\gamma)$ extragradient correction produces the Lyapunov decrease needed for convergence. Thus, our **Theorem 3.1** gives convergence to a single finite logit vector, eventually at an exponential rate.

This compactness mechanism breaks at a boundary SNE. Our boundary analysis instead treats this singular regime directly, yielding convergence guarantee under flat dominance and sharp, margin-dependent boundary rates (Theorems 3.2 and 3.3 and Proposition 3.1). Related analyses instead ensure that the representation remains nondegenerate along the trajectory. For example, Patel & Vlatakis-Gkaragkounis (2025) use overparameterization, random initialization, and a path-length bound for AltGDA to keep the Jacobians of wide two-layer networks well-conditioned throughout optimization. Softmax boundary trajectories instead leave this uniformly well-conditioned regime.

A.2 Softmax Game Dynamics and Extragradient

Without the extragradient correction, learning dynamics in zero-sum games can fail to converge in last iterate even for bilinear payoffs. Related work establishes elimination of dominated strategies (Mertikopoulos & Sandholm, 2016), recurrent behavior around interior equilibria (Mertikopoulos et al., 2018), and nonconvergence phenomena for policy-gradient and softmax-parametrized game dynamics (Bernasconi et al., 2025; Hennes et al., 2020). The leading-order term of our modified equation is the corresponding uncorrected softmax flow, but the high-resolution extragradient correction changes its interior behavior: Theorem 3.1 shows that it damps the rotational dynamics and yields last-iterate

convergence. At a flat boundary, by contrast, the leading asymptotics are governed by the vanishing action probabilities, resulting in polynomial rather than exponential convergence.

Vanishing softmax probabilities also produce $O(1/t)$-type convergence in single-agent softmax policy gradient (Mei et al., 2020). Our setting adds two features. First, vanishing actions are separated by their Nash margins: positive-margin actions decay as $O(1/t)$, whereas zero-margin actions can evolve on slower scales. Second, when there are multiple SNE, an equilibrium-selection problem arises. Related nonlinear probability parametrizations similarly show that parameter-space stationarity need not directly characterize the relevant game-space behavior (Do et al., 2026).

Extragradient and optimistic methods applied directly to policy variables provide a complementary comparison. In convex–concave bilinear games, they can achieve last-iterate convergence where GDA cycles (Daskalakis & Panageas, 2019; Daskalakis et al., 2018; Mokhtari et al., 2020; Wei et al., 2021). For games formulated on simplices, boundary equilibria are treated directly as points of the policy space, whereas our dynamics evolve in logit space through the softmax parametrization. Our object is therefore the high-resolution modified equation of extragradient in logit space (Chae et al., 2024; Grimmer et al., 2022; Lu, 2022).

A.3 Geometry-Aware and Regularized Methods

A complementary line of work modifies the update to better reflect the latent convex–concave geometry. Natural-gradient (Mladenovic et al., 2022), Jacobian-preconditioned (Sakos et al., 2023), and surrogate-loss (D'Orazio et al., 2025) methods compensate for the nonlinear representation rather than analyzing a fixed Euclidean parameter update. Our goal is instead to characterize the dynamics of $O(\gamma)$-resolution ODE associated with Euclidean-logit extragradient. Other algorithms exploiting application-specific latent structure include Gemp & Mahadevan (2018); Kalogiannis et al. (2024).

Entropy regularization avoids the same singularity by moving the target into the simplex interior. Entropy-regularized softmax methods can therefore obtain strong convergence guarantees and approximate the unregularized game as the regularization vanishes (Cen et al., 2024; Pattathil et al., 2023). For fixed positive regularization, however, the target remains interior, so the boundary asymptotics studied here do not arise.

Broader nonconvex-nonconcave minimax theory typically targets finite approximate stationarity or convergence under one-sided, two-sided, or related PL structure (Diakonikolas et al., 2021; Jin et al., 2020; Laguel et al., 2024; Lin et al., 2020; Nouiehed et al., 2019; Rankawat et al., 2026). These criteria do not directly characterize our limiting regime. Finite-logit stationarity does not characterize our limiting regime: boundary SNEs are attained only asymptotically, while small logit gradients may also arise simply because some action probabilities are nearly zero. Accordingly, our convergence statements are formulated in policy space; in particular, Theorem 4.1 shows that any convergent policy trajectory can only converge to an SNE.

A.4 Implicit Bias and Equilibrium Selection

Implicit bias broadly concerns the solutions selected by an optimization algorithm when the underlying problem admits multiple solutions (Gunasekar et al., 2018; Soudry et al., 2018). Here, the analogous question is one of equilibrium selection: when $\mathcal{SN}(\boldsymbol{A})$ is non-singleton, which SNE is selected by the unregularized logit dynamics? A related explicit result is Lee & Kim (2025), who study a Zero-Infinity GAN and obtain minimum-norm selection for the trained final layer under a local two-timescale analysis. Our setting instead concerns selection among Nash distributions of an unregularized finite game. Neighboring perspectives include empirical solver-dependent selection (Leal, 2026), algorithmic basin steering (Shcherbinin et al., 2026), and the KL projection selected by unregularized OMWU under suitable full-support conditions (Chen et al., 2025).

For Euclidean-logit extragradient, we obtain explicit selection result in Theorem 4.4, while Proposition J.1 shows that convergence to the Nash set need not imply convergence to a single policy in general. Together with the boundary-rate results, this shows that the softmax parametrization does more than complicate the transfer of convex–concave guarantees to parameter space: its boundary geometry determines the asymptotic scales of the dynamics and, in structured cases, the equilibrium that is selected.

B Dynamical Systems Toolkit

This appendix collects the dynamical-systems tools used in the proofs and organizes them in their order of dependence. We first set up solution maps and autonomous limit sets, then give Lyapunov and local-equilibrium tools, transfer asymptotically autonomous paths to chain-transitive limit sets, and record the planar refinements used later. The final subsection proves the stable-set nullity, local asymptotic phase, and unique-saddle results used above.

We largely follow the exposition of Teschl (2012), with additional references for more advanced and modern topics: Hirsch et al. (1977) for normally hyperbolic invariant manifolds; Perko (2001, Chapter 3.12) and do Carmo (1994, Chapter 2) for planar index theory; and Benaïm (1999); Benaïm & Hirsch (1996); Mischaikow et al. (1995) for asymptotic pseudotrajectories. We also refer to Wiggins (2003) throughout for a comprehensive yet intuitive treatment of many of these topics. Results proved here for convenience are standard and are not claimed as novel.

Vector fields are locally Lipschitz, and C^1 whenever derivatives are used. At boundary points, they are understood to extend smoothly to a neighborhood in a Euclidean chart.

B.1 Solutions, Evolution Maps, and Limit Sets

For the purposes of this paper, it suffices to work on open subsets of $\mathbb{R}^d$; dynamics on a closed simplex are understood through a smooth extension to an open neighborhood. Consider the nonautonomous initial-value problem on $\mathbb{R}^d$,

$$\dot{z} = f(t, z), \qquad z(s) = z_0. \tag{12}$$

We begin with the Picard-Lindelöf-type well-posedness of the solution:

Theorem B.1 (Global Existence under Linear Growth (Teschl, 2012, Theorems 2.2 and 2.17)). *Let $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ be continuous and locally Lipschitz in z, locally uniformly in t. Suppose that for every $T > 0$ there are constants $M_T, L_T \geq 0$ such that*[a]

$$\|f(t, z)\| \leq M_T + L_T\|z\| \qquad (|t| \leq T,\ z \in \mathbb{R}^d).$$

Then, for every $(s, z_0) \in \mathbb{R} \times \mathbb{R}^d$, (12) has a unique solution $z(\cdot; s, z_0) : \mathbb{R} \to \mathbb{R}^d$.

[a] Teschl (2012, Problem 2.22) gives the more general condition $\|f(t, z)\| \leq g_T(\|z\|)$ for some positive $g_T \in C([0, \infty))$ satisfying $\int_0^\infty dr/g_T(r) = \infty$.

In particular, every globally bounded, locally Lipschitz vector field on $\mathbb{R}^d$ satisfies the theorem with $L_T = 0$. Thus its solutions are defined for all positive and negative times.

With this, we now define the evolution process. Denote by $z(t; s, z_0)$ the state at time t of the solution to (12) initialized at z_0 at time s.

Definition B.1 (Evolution Process). *Under the hypotheses of Theorem B.1, define*

$$\Theta_{t,s} : \mathbb{R}^d \to \mathbb{R}^d, \qquad \Theta_{t,s}(z_0) := z(t; s, z_0), \qquad s, t \in \mathbb{R}.$$

*The family $(\Theta_{t,s})_{s,t\in\mathbb{R}}$ is the **evolution process** of (12).*

Uniqueness gives

$$\Theta_{s,s} = \mathrm{id}, \qquad \Theta_{u,t} \circ \Theta_{t,s} = \Theta_{u,s}, \qquad s, t, u \in \mathbb{R}. \tag{13}$$

If $f \in C^1$, then Teschl (2012, Theorem 2.10) gives C^1 dependence on the initial state; together with (13), this shows that every $\Theta_{t,s}$ is a C^1-diffeomorphism with inverse $\Theta_{s,t}$. We use the same notation in finite-horizon arguments on restricted domains, where the maps are understood on the initial states for which the indicated solutions exist.

We now specialize to the autonomous equation

$$\dot{z} = F(z), \qquad z(0) = z_0. \tag{ODE}$$

Its evolution depends only on elapsed time, and thus, we can write

$$\Theta_{t,s}(z) = \Phi_{t-s}(z).$$

With this, we define the following concepts:

Definition B.2 (Semiflows and Flows (Teschl, 2012, Section 6.2)). *A **semiflow** on a state space $\mathcal{M} \subseteq \mathbb{R}^d$ is a continuous family $(\Phi_t)_{t\geq 0}$ satisfying*[a]

$$\Phi_0 = \mathrm{id}_{\mathcal{M}}, \qquad \Phi_{t+r} = \Phi_t \circ \Phi_r \quad (t, r \geq 0).$$

*If $(\Phi_t)_{t\in\mathbb{R}}$ is defined and satisfies the above for all $t, r \in \mathbb{R}$, it is a **flow**. An autonomous vector field is **complete** if all of its solutions are defined on $\mathbb{R}$; it then generates a flow. An **equilibrium** is a point $z_\star$ with $F(z_\star) = 0$.*

[a] Thus a semiflow is a one-parameter *semigroup* of continuous self-maps of $\mathcal{M}$. A complete C^1 vector field instead generates a one-parameter *group* of C^1 diffeomorphisms.

In particular, every globally bounded, locally Lipschitz $F : \mathbb{R}^d \to \mathbb{R}^d$ is complete by Theorem B.1.

Definition B.3 (Trajectories and Orbits (Teschl, 2012, Section 6.3)). *For a semiflow, the **forward trajectory** through z is $t \mapsto \Phi_t(z)$ for $t \geq 0$, and its **forward orbit** is $\mathcal{O}^+(z) := \{\Phi_t(z) : t \geq 0\}$. For a flow, the trajectory and orbit through z are defined analogously using $t \in \mathbb{R}$. A trajectory is **periodic** if it is nonconstant and $\Phi_T(z) = z$ for some $T > 0$.*

Definition B.4 (Invariant Sets (Teschl, 2012, Section 6.3)). *A set $\mathcal{K} \subseteq \mathcal{M}$ is **positively invariant** if $\Phi_t(\mathcal{K}) \subseteq \mathcal{K}$ for every $t \geq 0$. It is **invariant** if $\Phi_t(\mathcal{K}) = \mathcal{K}$ for every $t \geq 0$; for a flow, this is equivalent to $\Phi_t(\mathcal{K}) = \mathcal{K}$ for every $t \in \mathbb{R}$.*

Now we introduce a core concept in dynamical systems:

Definition B.5 (Limit Sets (Teschl, 2012, Section 6.3)). *For a forward path $z : [0, \infty) \to \mathcal{M}$, its **omega-limit set** is*

$$\omega(z) := \bigcap_{T\geq 0} \overline{\{z(t) : t \geq T\}}. \tag{14}$$

*The path is **precompact in** $\mathcal{M}$ if its image has compact closure contained in $\mathcal{M}$. For a semiflow, we use the shorthand $\omega(z_0) := \omega(t \mapsto \Phi_t(z_0))$.*

Proposition B.1 (Lemmas 6.5–6.7 in Teschl (2012)). *If the forward orbit of $z_0 \in \mathcal{M}$ is precompact in $\mathcal{M}$, then $\omega(z_0)$ is nonempty, compact, connected, and invariant, and*

$$\mathrm{dist}\big(\Phi_t(z_0), \omega(z_0)\big) \longrightarrow 0. \tag{15}$$

In particular, $\Phi_t(z_0)$ converges if and only if its omega-limit set is a singleton.

B.2 Lyapunov Functions, Invariance, and Attraction

For $L \in C^1$, its Lie (orbital) derivative along the autonomous field F is

$$\dot{L}(z) := dL_z[F(z)],$$

which is $\nabla L(z)^\top F(z)$ in Euclidean coordinates.

Definition B.6 (Stability and Lyapunov Functions (Teschl, 2012, Sections 6.5–6.6)). *Let $z_\star$ be an equilibrium of (ODE).*

*(i) $z_\star$ is **(Lyapunov) stable** if for every neighborhood $\mathcal{U}$ of $z_\star$, there is a neighborhood $\mathcal{W} \subseteq \mathcal{U}$ such that $\Phi_t(\mathcal{W}) \subseteq \mathcal{U}$ for every $t \geq 0$. It is **asymptotically stable** if it is stable and there is a neighborhood $\mathcal{W}$ such that $\Phi_t(z) \to z_\star$ for every $z \in \mathcal{W}$.*

*(ii) A function $L \in C^1(\mathcal{U}; \mathbb{R}_{\geq 0})$ on a neighborhood $\mathcal{U}$ of $z_\star$ is a **Lyapunov function** if $L(z_\star) = 0$, $L(z) > 0$ for $z \neq z_\star$, and $\dot{L}(z) \leq 0$ on $\mathcal{U}$. It is **strict** if $\dot{L}(z) < 0$ on $\mathcal{U} \setminus \{z_\star\}$.*

We now state the seminal stability result due to Lyapunov (1892):

Theorem B.2 (Lyapunov's Stability Criterion (Teschl, 2012, Theorems 6.13–6.14)). *If there is a Lyapunov function for $z_\star$, then $z_\star$ is stable. If the function is strict, then $z_\star$ is asymptotically stable.*

We also state LaSalle's extension of Lyapunov's stability result, which *identifies* the limiting invariant set (LaSalle, 1960; 1968):

Theorem B.3 (LaSalle Invariance Principle (Wiggins, 2003, Theorem 8.3.1)). *Let $\mathcal{K}$ be compact and positively invariant, and let L be C^1 on a neighborhood of $\mathcal{K}$, with $\dot{L} \leq 0$ on $\mathcal{K}$. Define $\mathcal{E}_L := \{z \in \mathcal{K} : \dot{L}(z) = 0\}$, and let $\mathcal{M} \subseteq \mathcal{E}_L$ be the largest invariant subset of $\mathcal{E}_L$. Then*

$$\operatorname{dist}(\Phi_t(z_0), \mathcal{M}) \longrightarrow 0, \qquad \forall z_0 \in \mathcal{K}.$$

Equivalently, $\omega(z_0) \subseteq \mathcal{M}$ for every $z_0 \in \mathcal{K}$.

B.3 Local Behavior Near Equilibria

Definition B.7 (Stable and Unstable Sets (Teschl, 2012, Section 9.2)). *Whenever the indicated trajectories exist, the **stable** and **unstable sets** of an equilibrium $z_\star$ are*

$$\mathcal{W}^s(z_\star) := \{z : \Phi_t(z) \to z_\star \text{ as } t \to +\infty\}, \quad \mathcal{W}^u(z_\star) := \{z : \Phi_t(z) \to z_\star \text{ as } t \to -\infty\}. \tag{16}$$

Definition B.8 (Hyperbolic Equilibrium (Teschl, 2012, Section 9.2)). *Let $z_\star$ be an equilibrium and set $\boldsymbol{A} := DF(z_\star)$ to be its linearization. Then, it is said to be **hyperbolic** if $\boldsymbol{A}$ has no eigenvalue on the imaginary axis. Its **stable** and **unstable subspaces**, denoted $\mathcal{E}^s_\star$ and $\mathcal{E}^u_\star$, are the real invariant subspaces generated by the generalized eigenvectors[a] associated with eigenvalues of negative and positive real part, respectively. Thus $\mathbb{R}^d = \mathcal{E}^s_\star \oplus \mathcal{E}^u_\star$. Moreover, $z_\star$ is a **sink** if $\mathcal{E}^u_\star = \{0\}$, a **source** if $\mathcal{E}^s_\star = \{0\}$, and a **saddle** otherwise.*

[a] In the complexification, a generalized eigenvector associated with λ is a nonzero vector $\boldsymbol{b}$ satisfying $(\boldsymbol{A} - \lambda \boldsymbol{I})^k \boldsymbol{b} = 0$ for some $k \geq 1$; for a nonreal conjugate pair, the real and imaginary parts generate the corresponding real invariant subspace.

Theorem B.4 (Stable Manifold Theorem (Teschl, 2012, Theorems 9.4–9.5)). *Let F be a C^1 vector field on an open $\mathcal{O} \subseteq \mathbb{R}^n$, and let $z_\star$ be a hyperbolic equilibrium. There exist local embedded C^1 manifolds $\mathcal{W}^s_{\mathrm{loc}}(z_\star)$ and $\mathcal{W}^u_{\mathrm{loc}}(z_\star)$ through $z_\star$, tangent to $\mathcal{E}^s_\star$ and $\mathcal{E}^u_\star$, satisfying*

- *(i) An orbit starting on either manifold remains on it for as long as the orbit stays in the chosen neighborhood. (**local invariance**)*
- *(ii) Trajectories starting on $\mathcal{W}^s_{\mathrm{loc}}(z_\star)$ exist for all forward time and converge exponentially to $z_\star$. The corresponding statement holds in backward time for $\mathcal{W}^u_{\mathrm{loc}}(z_\star)$.*
- *(iii) Every trajectory converging to $z_\star$ in forward time eventually lies in $\mathcal{W}^s_{\mathrm{loc}}(z_\star)$. The corresponding statement holds in backward time for $\mathcal{W}^u_{\mathrm{loc}}(z_\star)$.*

For an open set $\mathcal{O} \subseteq \mathbb{R}^n$, a codimension-one embedded C^1 submanifold $\Sigma \subseteq \mathcal{O}$ is **transverse to** F if $F(z) \notin T_z\Sigma$ for every $z \in \Sigma$; we call such a Σ a **transverse section**.

The next lemma, whose proof is deferred to Appendix B.6.1, shows that the set of points converging to a hyperbolic equilibrium with at least one unstable direction has *measure zero* both in the ambient space and on every transverse section:

Lemma B.1 (Nullity of a Hyperbolic Stable Set). *Let F be a C^1 vector field on an open set $\mathcal{O} \subseteq \mathbb{R}^n$, and let $z_\star$ be a hyperbolic equilibrium with $\dim \mathcal{E}^u_\star \geq 1$. Then $\mathcal{W}^s(z_\star)$ is Lebesgue-null. Furthermore, if Σ is a codimension-one embedded C^1 submanifold transverse to F, then $\mathcal{W}^s(z_\star) \cap \Sigma$ is null for the induced $(n-1)$-dimensional measure on Σ.*

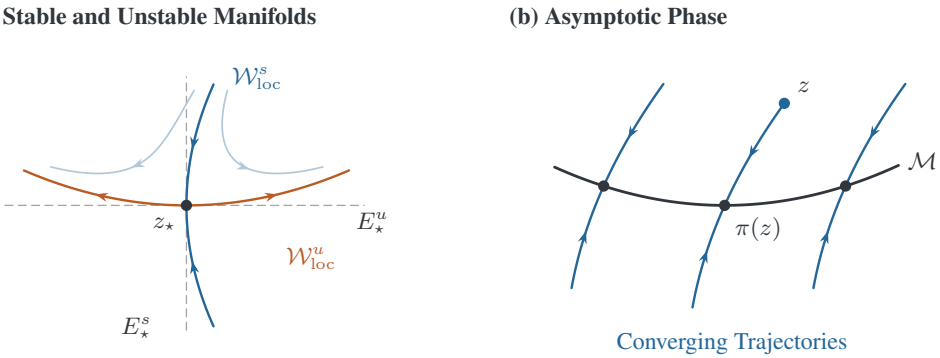


Figure 2: **Theorems B.4 (left) and B.5 (right).** Stable and unstable manifolds at a saddle; trajectories converging to individual equilibria on $\mathcal{M}$, with z approaching $\pi(z)$.

If a positive-dimensional manifold $\mathcal{M}$ consists of equilibria, then $DF(z)$ vanishes on $T_z\mathcal{M}$. Thus its points are not hyperbolic; the relevant condition is contraction normal to $\mathcal{M}$.

Definition B.9 (Normally Attracting Manifold of Equilibria (Hirsch et al., 1977, Section 1, Definition 4))**.** *Let $\mathcal{M}$ be a C^1 manifold consisting of equilibria of $\dot{z} = F(z)$. Working in a local Euclidean chart around $z_\star \in \mathcal{M}$, we say that $\mathcal{M}$ is **normally attracting** near $z_\star$ if, on a neighborhood of $z_\star$ in $\mathcal{M}$, there is a continuous splitting*

$$\mathbb{R}^d = T_z\mathcal{M} \oplus \mathcal{E}^s_z$$

preserved by the linearized flow, meaning $D\Phi_t(z)\mathcal{E}^s_z = \mathcal{E}^s_{\Phi_t(z)}$ for $t \geq 0$, and constants $C, \alpha > 0$ such that

$$\|D\Phi_t(z)v\| \leq Ce^{-\alpha t}\|v\|, \qquad t \geq 0, \quad v \in \mathcal{E}^s_z, \tag{17}$$

uniformly for z in that neighborhood.

Below is a standard local consequence of the stable-foliation theory of Hirsch et al. (1977, Theorem 4.1(a),(e),(h)). For completeness, we provide its full proof in Appendix B.6.2, which utilizes the normal-stability formulation of Prüss et al. (2009, Theorem 2.1 and Eq. (2.27)).

Theorem B.5 (Local Asymptotic Phase)**.** *Let $F \in C^r$, $r \geq 1$, and let $\mathcal{M}$ be a C^r embedded manifold of equilibria that is normally attracting near $z_\star \in \mathcal{M}$. After restricting to a sufficiently small forward-invariant neighborhood $\mathcal{U}$ of $z_\star$, there is a continuous phase map $\pi : \mathcal{U} \to \mathcal{M}$ such that $\pi(z) = z$ on $\mathcal{U} \cap \mathcal{M}$ and*

$$\pi(\Phi_t(z)) = \pi(z), \qquad z \in \mathcal{U}, \quad t \geq 0.$$

Moreover, after shrinking $\mathcal{U}$ if necessary, there are $C_0, \lambda > 0$ such that

$$\|\Phi_t(z) - \pi(z)\| \leq C_0 e^{-\lambda t}, \qquad z \in \mathcal{U}, \quad t \geq 0. \tag{18}$$

Hence every trajectory starting in $\mathcal{U}$ converges exponentially to the single equilibrium $\pi(z) \in \mathcal{M}$.

B.4 Asymptotic Pseudotrajectories and Chain-Transitive Limit Sets

Let $(\mathcal{M}, d)$ be a metric space. A continuous path $z : [0, \infty) \to \mathcal{M}$ is an ***asymptotic pseudotrajectory*** (APT) of a continuous semiflow Φ if, for every $T > 0$,

$$\lim_{t\to\infty} \sup_{0\leq h\leq T} d\big(z(t+h), \Phi_h(z(t))\big) = 0. \tag{19}$$

The following ODE criterion is the form used throughout the paper.

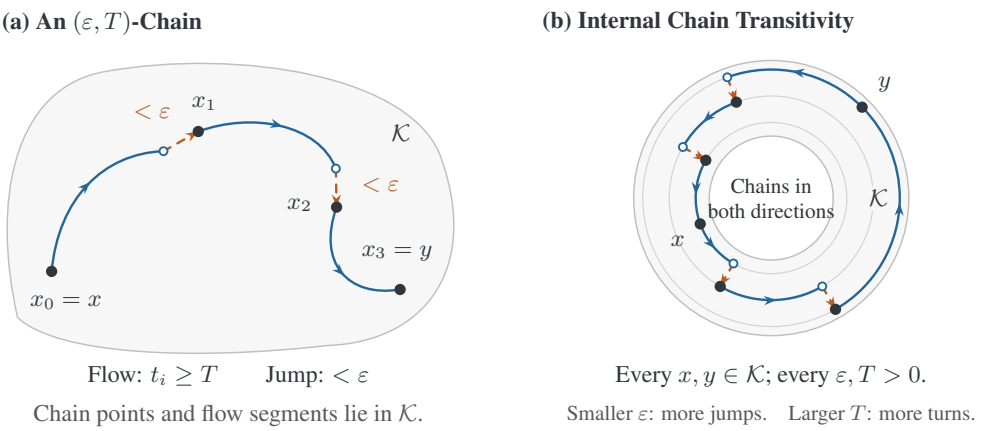


Figure 3: **Definition B.10.** An (ε, T)-chain (left) and an internally chain-transitive annulus (right).

Lemma B.2 (APT Criterion (Benaïm & Hirsch, 1996, Proposition 2.1)). *Suppose that $f \in C([0,\infty) \times \mathbb{R}^d, \mathbb{R}^d)$ satisfies $f(t,\cdot) \to f_\infty$ locally uniformly as $t \to \infty$ for a locally Lipschitz f_∞ that generates a flow Φ_∞. Then every bounded forward solution of $\dot{z} = f(t,z)$ is an APT of Φ_∞.*

Definition B.10 (Internal Chain Transitivity (Benaïm, 1999, Section 5)). *Let $\mathcal{K}$ be nonempty, compact, and invariant. Given $\varepsilon, T > 0$, an (ε, T)-**chain in** $\mathcal{K}$ **from** x **to** y consists of points*

$$x = x_0, x_1, \ldots, x_n = y, \qquad n \geq 1,$$

in $\mathcal{K}$ and times $t_i \geq T$ such that

$$d\big(\Phi_{t_i}(x_i), x_{i+1}\big) < \varepsilon, \qquad 0 \leq i < n.$$

*The set $\mathcal{K}$ is **internally chain transitive** (ICT) if such a chain exists between every ordered pair $x, y \in \mathcal{K}$ for every $\varepsilon, T > 0$.*

The following result is due independently to Mischaikow et al. (1995) and Benaïm & Hirsch (1996).

Theorem B.6 (APT Limit Sets (Benaïm, 1999, Theorem 5.7)). *Let $z : [0,\infty) \to \mathcal{M}$ be a precompact APT of a continuous finite-dimensional semiflow Φ. Then $\omega(z)$ is nonempty, compact, connected, invariant under Φ, and internally chain transitive.*

Corollary B.7 (Strict-Lyapunov Convergence (Benaïm, 1999, Corollary 6.6)). *Let z be a precompact APT of a flow on an invariant open set $\mathcal{O}$. Suppose there is $L \in C^1(\mathcal{O}; \mathbb{R})$ such that $\dot{L}(x) < 0$ at every non-equilibrium $x \in \mathcal{O}$. If the equilibria in $\omega(z)$ are countable, then $z(t)$ converges to an equilibrium. In particular, this holds if $\omega(z)$ lies in an invariant region with a unique equilibrium.*

B.5 Planar Dynamics and Index Theory

Here, we recall the index theory for planar dynamics (dynamical systems over $\mathbb{R}^2$).

Definition B.11 (Homoclinic and Heteroclinic Trajectories (Perko, 2001, Section 3.3)). *A nonconstant complete trajectory γ connecting equilibria a and b satisfies*

$$\lim_{t \to -\infty} \gamma(t) = a, \qquad \lim_{t \to +\infty} \gamma(t) = b.$$

*It is **heteroclinic** if $a \neq b$ and **homoclinic** if $a = b$. A **cyclic orbit chain** is a finite union of equilibria $q_1, \ldots, q_r$ and nonconstant connecting trajectories from q_i to q_{i+1}, where $q_{r+1} = q_1$. This is a **heteroclinic cycle** if each connection is heteroclinic.*

Now we introduce the index theory for planar dynamics. We always assume that all Jordan curves are piecewise C^1 and positively oriented. Whenever we discuss the interior of a Jordan curve $\mathcal{C}$, we assume that the vector field F is defined on an open neighborhood of $\mathcal{C}$ and its bounded interior.

Definition B.12 (Poincaré Index (Perko, 2001, Section 3.12, Definition 1); (do Carmo, 1994, Chapter 2, Remark 1))**.** *Let $\mathcal{C}$ be a Jordan curve and let the planar vector field $F = (F_1, F_2)$ have no zero on $\mathcal{C}$. The **index** of F along $\mathcal{C}$ is the winding number of $F/\|F\| : \mathcal{C} \to \mathbb{S}^1$:*

$$\operatorname{ind}_{\mathcal{C}}(F) := \frac{1}{2\pi} \oint_{\mathcal{C}} \frac{F_1\, dF_2 - F_2\, dF_1}{F_1^2 + F_2^2}.$$

If q is an isolated equilibrium, $\operatorname{ind}(F, q)$ is the index along any small circle that encloses only q.

Proposition B.2 (Planar Index Facts (Perko, 2001, Section 3.12, Theorems 1–3 and 5–6))**.** *Let $\mathcal{C}$ be a Jordan curve on which F does not vanish. If its interior contains only finitely many equilibria $q_1, \dots, q_N$, then*

$$\operatorname{ind}_{\mathcal{C}}(F) = \sum_{r=1}^{N} \operatorname{ind}(F, q_r),$$

with the empty sum interpreted as zero. The index of a periodic orbit is 1, so every periodic orbit encloses an equilibrium. A nondegenerate equilibrium q has index $\operatorname{sign}\det DF(q)$, so a hyperbolic saddle has index -1. Consequently, a periodic orbit cannot enclose a unique hyperbolic saddle.

To deal with a homoclinic loop containing an equilibrium on its boundary, we use the following consequence of planar index theory.

Lemma B.3 (Interior of a Saddle Loop (Perko, 2001, Section 3.12, Remark 1 and Theorems 2 and 6))**.** *Let F be a C^1 planar vector field and let Γ be a simple homoclinic loop based at a hyperbolic saddle $z_\star$; thus $\Gamma \setminus \{z_\star\}$ is one orbit and Γ is a Jordan curve. Then its bounded Jordan domain contains an equilibrium distinct from $z_\star$.*

A simple homoclinic loop based at a hyperbolic saddle admits a Poincaré map on one side. On that side, it can be approximated by Jordan curves of index 1 (Perko, 2001, Section 3.12, Remark 1). If its bounded interior contained no equilibrium, a sufficiently close such curve would enclose either no equilibrium or only $z_\star$. Its index would therefore be 0 or -1, respectively, contradicting the preceding observation (Perko, 2001, Section 3.12, Theorems 1–2 and 6).

The next theorem identifies the recurrent behavior that an ICT set can contain.

Theorem B.8 (Planar ICT Classification (Benaïm, 1999, Theorem 6.12))**.** *Let Φ be a C^1 flow on $\mathbb{R}^2$ with isolated equilibria, and let $\mathcal{K}$ be a compact ICT set. Then every point of $\mathcal{K}$ is either an equilibrium, lies on a periodic trajectory, or belongs to a cyclic orbit chain contained in $\mathcal{K}$.*

Combining everything, we have the following corollary whose proof is deferred to Appendix B.6.3:

Corollary B.9 (A Unique Saddle in a Rectangle)**.** *Let $F \in C^1(\mathcal{R}; \mathbb{R}^2)$ generate a complete flow on $\mathcal{R} = (a_1, b_1) \times (a_2, b_2)$, with $a_i < b_i$. If its unique equilibrium is a hyperbolic saddle $z_\star$, then every nonempty compact ICT set $\mathcal{K} \subset \mathcal{R}$ equals $\{z_\star\}$.*

B.6 Deferred Proofs of "Folklore" Results

B.6.1 Proof of Lemma B.1: Nullity of a Hyperbolic Stable Set

Let us denote $u := \dim \mathcal{E}^u_\star$. By the Stable Manifold Theorem (Theorem B.4), $\dim \mathcal{W}^s_{\mathrm{loc}}(z_\star) = n - u$. Let $\mathcal{D}_N \subseteq \mathcal{O}$ be the open set on which Φ_N is defined and set

$$\mathcal{M}_N := \{z \in \mathcal{D}_N : \Phi_N(z) \in \mathcal{W}^s_{\mathrm{loc}}(z_\star)\} = \Phi_N^{-1}\big(\mathcal{W}^s_{\mathrm{loc}}(z_\star)\big).$$

Every trajectory converging to $z_\star$ lies in $\mathcal{W}^s_{\mathrm{loc}}(z_\star)$ for all sufficiently large times, hence at some nonnegative integer time N. Conversely, every point whose trajectory reaches $\mathcal{W}^s_{\mathrm{loc}}(z_\star)$ converges to $z_\star$. Thus $\mathcal{W}^s(z_\star) = \bigcup_{N=0}^{\infty} \mathcal{M}_N$. Since Φ_N is a local C^1 diffeomorphism on $\mathcal{D}_N$, each $\mathcal{M}_N$ is an $(n-u)$-dimensional embedded submanifold. Hence each $\mathcal{M}_N$ is Lebesgue-null, and so is their countable union.

Now fix N and $z \in \mathcal{M}_N$. By the flow property and local invariance of $\mathcal{W}^s_{\mathrm{loc}}(z_\star)$, for all sufficiently small $h \geq 0$,

$$\Phi_N(\Phi_h(z)) = \Phi_h(\Phi_N(z)) \in \mathcal{W}^s_{\mathrm{loc}}(z_\star).$$

Thus $\Phi_h(z) \in \mathcal{M}_N$ for such h, and differentiating at $h = 0$ gives $F(z) \in T_z\mathcal{M}_N$.

At any $z \in \mathcal{M}_N \cap \Sigma$, transversality of Σ to F gives $F(z) \notin T_z\Sigma$. Since $T_z\Sigma$ has codimension one, this implies

$$T_z\mathcal{M}_N + T_z\Sigma = \mathbb{R}^n.$$

If $n - u = 0$, the intersection is empty: tangency would give $F(z) = 0$, contradicting transversality to Σ. Otherwise, the transverse-intersection theorem (Lee, 2012, Theorem 6.30(b))[3] shows that $\mathcal{M}_N \cap \Sigma$ is a submanifold of Σ of dimension

$$(n-u) + (n-1) - n = n - u - 1 < n - 1.$$

It is therefore null for the induced measure on Σ. Taking the countable union over N proves that $\mathcal{W}^s(z_\star) \cap \Sigma$ is null as well. □

B.6.2 Proof of Theorem B.5: Local Asymptotic Phase

Recall that $F \in C^1$ and $\mathcal{M}$ is a C^1 embedded manifold of equilibria, normally attracting near $z_\star$. Write

$$T := T_{z_\star}\mathcal{M}, \qquad E := \mathcal{E}^s_{z_\star}, \qquad L := DF(z_\star).$$

By Definition B.9, since $z_\star$ is fixed,

$$\mathbb{R}^d = T \oplus E, \qquad D\Phi_t(z_\star)E = E, \qquad \|D\Phi_t(z_\star)|_E\| \leq Ce^{-\alpha t}$$

for some $C, \alpha > 0$ and all $t \geq 0$.

The cases $\dim \mathcal{M} = 0$ and $\dim \mathcal{M} = d$ follow respectively from linearized exponential stability and $F = 0$ locally, using the forward-invariant neighborhood construction below in the former case. Assume henceforth that $0 < \dim \mathcal{M} < d$.

We use the following ODE specialization of Prüss et al. (2009, Theorem 2.1).[4]

Lemma B.4 (Normal stability). *Let F be a C^1 vector field near an equilibrium $z_\star$. Suppose that:*

(i) The set of equilibria near $z_\star$ is a C^1 embedded manifold $\mathcal{M}$ with $0 < \dim \mathcal{M} < d$;
(ii) $\ker DF(z_\star) = T_{z_\star}\mathcal{M}$;
(iii) $\mathbb{R}^d = \ker DF(z_\star) \oplus \operatorname{ran} DF(z_\star)$;
(iv) $\sigma(DF(z_\star)) \setminus \{0\} \subset \{\zeta \in \mathbb{C} : \operatorname{Re}\zeta < 0\}$, where σ denotes the spectrum.

Then $z_\star$ is Lyapunov stable, and every sufficiently nearby initial condition v generates a unique trajectory defined for all $t \geq 0$ and converging to an equilibrium $q(v) \in \mathcal{M}$.

We first verify (ii)–(iv). Since F vanishes on $\mathcal{M}$, differentiation along $\mathcal{M}$ gives $L|_T = 0$. Differentiating the flow equation $\partial_t\Phi_t(z) = F(\Phi_t(z))$ with respect to the initial condition z, and then setting $z = z_\star$, gives

$$\dot{J}(t) = DF(\Phi_t(z_\star))J(t) = LJ(t), \qquad J(0) = I, \;\; J(t) := D\Phi_t(z_\star).$$

Hence $D\Phi_t(z_\star) = e^{tL}$. The assumed invariance of E gives $e^{tL}E = E$; differentiating at $t = 0$ yields $L(E) \subset E$. Thus, with $B := L|_E$,

$$L = 0_T \oplus B, \qquad \|e^{tB}\| \leq Ce^{-\alpha t}.$$

[3] Although this is stated for smooth manifolds, the proof follows through for C^1 manifolds.

[4] In their equation $\dot{u} + A(u)u = \widetilde{F}(u)$, take $X_0 = X_1 = X_\gamma = \mathbb{R}^d$, $p = 2$, $A(u) = I$, and $\widetilde{F}(u) = u + F(u)$. Their coefficient regularity follows from $F \in C^1$, maximal-L^2-regularity holds for $A = I$, and their linearization is $A_0 = -DF(z_\star)$.

Every eigenvalue μ of B therefore satisfies $e^{t\,\mathrm{Re}\,\mu} \le Ce^{-\alpha t}$ for all $t \ge 0$, which implies $\mathrm{Re}\,\mu \le -\alpha < 0$. In particular, B is invertible, so $\ker L = T$ and $\mathrm{ran}\, L = E$. This verifies (ii)–(iv).

For (i), we must exclude other nearby equilibria. Let P^c and P^s be the projections onto T and E. Since $\mathcal{M}$ is embedded with tangent space T at $z_\star$, it is locally a graph

$$z_\star + x + \varphi(x), \qquad x \in T,$$

where φ is C^1, takes values in E, and $\varphi(0) = 0$. Any nearby equilibrium $z_\star + x + y$ with $x \in T$ and $y \in E$ must satisfy $P^s F(z_\star + x + y) = 0$. The derivative of this equation with respect to y at $(0,0)$ is B, which is invertible. The implicit function theorem therefore gives a unique solution y for each small x. Since $y = \varphi(x)$ solves the equation, every nearby equilibrium lies on $\mathcal{M}$, proving (i).

We can now apply Lemma B.4, which gives stability and convergence. To obtain a uniform rate, apply Prüss et al. (2009, Eq. (2.27)) with the splitting and graph above. For all v in a sufficiently small ball about $z_\star$, it gives constants $A, \lambda > 0$, independent of v, such that

$$\|\Phi_t(v) - q(v)\| \le Ae^{-\lambda t}\big\|P^s(v - z_\star) - \varphi(P^c(v - z_\star))\big\|.$$

The factor involving v is continuous, hence bounded on a sufficiently small closed ball about $z_\star$. Consequently, after shrinking the ball, there is $C_0 > 0$ such that

$$\|\Phi_t(v) - q(v)\| \le C_0 e^{-\lambda t}, \qquad t \ge 0, \tag{20}$$

uniformly for initial conditions in that ball.

By Lyapunov stability, choose a smaller open ball $\mathcal{V}$ about $z_\star$ whose trajectories remain in a prescribed small neighborhood, and set

$$\mathcal{U} := \bigcup_{s \ge 0} \Phi_s(\mathcal{V}).$$

Then $\mathcal{U}$ is open (as Φ_s's are local diffeomorphisms), contains $\mathcal{V}$, and is forward invariant.

Every $w \in \mathcal{U}$ has the form $w = \Phi_s(v)$ for some $v \in \mathcal{V}$ and $s \ge 0$. Since $\Phi_t(w) = \Phi_{t+s}(v) \longrightarrow q(v) \in \mathcal{M}$ as $t \to \infty$, the trajectory starting at w has a limit. We therefore define

$$\pi(w) := \lim_{t\to\infty} \Phi_t(w) \quad \overset{(20)}{\Longrightarrow} \quad \|\Phi_t(w) - \pi(w)\| \le C_0 e^{-\lambda(t+s)} \le C_0 e^{-\lambda t}.$$

The estimate shows that Φ_t converges uniformly to π on $\mathcal{U}$ as $t \to \infty$. Since each Φ_t is continuous, π is continuous. Constant equilibrium trajectories give $\pi|_{\mathcal{U}\cap\mathcal{M}} = \mathrm{id}$, while the flow property gives

$$\pi(\Phi_s(w)) = \lim_{t\to\infty} \Phi_{t+s}(w) = \pi(w), \qquad s \ge 0.$$

□

B.6.3 Proof of Corollary B.9: A Unique Saddle in a Rectangle

We first justify applying the planar ICT classification to the flow on $\mathcal{R}$. Define the smooth diffeomorphism $H : \mathcal{R} \to \mathbb{R}^2$ by

$$H(z_1, z_2) := \left(\log \frac{z_1 - a_1}{b_1 - z_1}, \log \frac{z_2 - a_2}{b_2 - z_2}\right).$$

If Φ_t is the flow generated by F, then $\widetilde{\Phi}_t := H \circ \Phi_t \circ H^{-1}$ is a complete C^1 flow on $\mathbb{R}^2$. Its only equilibrium is $H(z_\star)$, which is isolated. Moreover, $H(\mathcal{K})$ is compact and invariant. To verify internal chain transitivity, fix $\varepsilon, T > 0$. Uniform continuity of H on $\mathcal{K}$ gives $\delta > 0$ such that $\|H(z) - H(w)\| < \varepsilon$ whenever $z, w \in \mathcal{K}$ and $\|z - w\| < \delta$. Applying H to a (δ, T)-chain in $\mathcal{K}$ therefore gives an (ε, T)-chain in $H(\mathcal{K})$, with the same flow times. Thus the planar ICT classification applies to $H(\mathcal{K})$ (Theorem B.8). Since H preserves equilibria, periodic trajectories, and connecting trajectories, every point of $\mathcal{K}$ is an equilibrium, lies on a periodic trajectory, or belongs to a cyclic orbit chain in $\mathcal{K}$. The proof idea is shown in Figure 4.

We next exclude periodic trajectories. Suppose that $\mathcal{K}$ contains one, whose orbit is a Jordan curve. Its bounded interior lies in $\mathcal{R}$ because the rectangle is convex, so it encloses either no equilibrium or only $z_\star$. Its index is therefore 0 or -1, whereas a periodic orbit has index 1 (Proposition B.2). This is a contradiction.

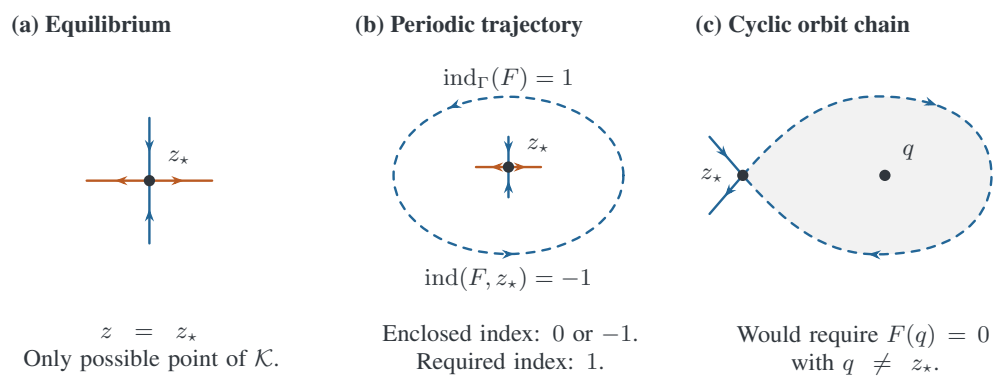


Figure 4: **Corollary B.9: A Unique Saddle in a Rectangle.** Every point of a compact ICT set is an equilibrium, a periodic point, or belongs to a cyclic orbit chain (Theorem B.8). If $z_\star$ is the only equilibrium, (b) a periodic orbit would have index 1 but enclose equilibria of total index 0 or -1, and (c) every connection in a cyclic orbit chain would be homoclinic and require another equilibrium q. Thus only (a) remains. Dashed curves depict the excluded configurations.

Now suppose that $\mathcal{K}$ contains a cyclic orbit chain. Every connection in the chain starts and ends at $z_\star$, since there is no other equilibrium in $\mathcal{R}$. Choose one such nonconstant trajectory. It has no self-intersection by uniqueness of solutions, and both its forward and backward limits are $z_\star$. Adding this common limit identifies the two ends of the trajectory and gives a continuous injective parametrization by a circle. Hence the trajectory together with $z_\star$ is a simple homoclinic loop. Its bounded interior again lies in $\mathcal{R}$ and must contain another equilibrium (Lemma B.3), a contradiction.

The classification therefore leaves only equilibria in $\mathcal{K}$. Since $z_\star$ is the unique equilibrium and $\mathcal{K}$ is nonempty, we conclude that $\mathcal{K} = \{z_\star\}$. □

C FROM EXTRAGRADIENT METHOD TO S2P-EG-FLOW.

In this Appendix, we provide the complete derivation of (S2P-EG-Flow) as well as prove its well-posedness.

Derivation. Our modified flow is taken from Lu (2022, Corollary 1(iii)), specialized to our symmetric self-play setting. For completeness, we provide the derivation below.

Let $\boldsymbol{A} \in \mathbb{R}^{K\times K}$ be skew-symmetric and define $J(\boldsymbol{\theta},\boldsymbol{\phi}) \coloneqq \boldsymbol{p}(\boldsymbol{\theta})^\top \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\phi})$, where $\boldsymbol{p}(\cdot) = \mathrm{softmax}(\cdot)$. Write $\boldsymbol{S}(\boldsymbol{p}) \coloneqq \mathrm{diag}(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^\top$. Since the first player minimizes and the second player maximizes, the two-player saddle operator is

$$F(\boldsymbol{\theta},\boldsymbol{\phi}) \coloneqq \begin{pmatrix} \nabla_{\boldsymbol{\theta}} J(\boldsymbol{\theta},\boldsymbol{\phi}) \\ -\nabla_{\boldsymbol{\phi}} J(\boldsymbol{\theta},\boldsymbol{\phi}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}(\boldsymbol{p}(\boldsymbol{\theta}))\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\phi}) \\ \boldsymbol{S}(\boldsymbol{p}(\boldsymbol{\phi}))\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) \end{pmatrix},$$

where the second equality uses $\boldsymbol{A}^\top = -\boldsymbol{A}$.

Let $\boldsymbol{z} \coloneqq (\boldsymbol{\theta},\boldsymbol{\phi}) \in \mathbb{R}^{2K}$. One step of simultaneous Euclidean extragradient with step size γ is

$$\bar{\boldsymbol{z}} = \boldsymbol{z} - \gamma F(\boldsymbol{z}), \qquad \boldsymbol{z}^+ = \boldsymbol{z} - \gamma F(\bar{\boldsymbol{z}}).$$

Taylor expansion of the extragradient update gives

$$\boldsymbol{z}^+ = \boldsymbol{z} - \gamma F(\boldsymbol{z}) + \gamma^2 DF(\boldsymbol{z})F(\boldsymbol{z}) + O(\gamma^3).$$

To obtain a continuous-time system matching this expansion through order γ^2, consider

$$\dot{\boldsymbol{z}} = -F(\boldsymbol{z}) + \gamma \boldsymbol{h}(\boldsymbol{z}).$$

Its flow satisfies

$$\begin{aligned} \boldsymbol{z}(\gamma) &= \boldsymbol{z} + \gamma\dot{\boldsymbol{z}}(0) + \frac{\gamma^2}{2}\ddot{\boldsymbol{z}}(0) + O(\gamma^3), \\ \dot{\boldsymbol{z}}(0) &= -F(\boldsymbol{z}) + \gamma\boldsymbol{h}(\boldsymbol{z}), \\ \ddot{\boldsymbol{z}}(0) &= \left(-DF(\boldsymbol{z}) + \gamma D\boldsymbol{h}(\boldsymbol{z})\right)\left(-F(\boldsymbol{z}) + \gamma\boldsymbol{h}(\boldsymbol{z})\right) \\ &= DF(\boldsymbol{z})F(\boldsymbol{z}) + O(\gamma). \end{aligned}$$

Therefore,

$$\boldsymbol{z}(\gamma) = \boldsymbol{z} - \gamma F(\boldsymbol{z}) + \gamma^2\left(\boldsymbol{h}(\boldsymbol{z}) + \frac{1}{2}DF(\boldsymbol{z})F(\boldsymbol{z})\right) + O(\gamma^3).$$

Matching the two expansions yields

$$\boldsymbol{h}(\boldsymbol{z}) = \frac{1}{2}DF(\boldsymbol{z})F(\boldsymbol{z}),$$

and hence the $O(\gamma)$-resolution modified flow associated with extragradient is

$$\dot{\boldsymbol{z}} = -F(\boldsymbol{z}) + \frac{\gamma}{2}DF(\boldsymbol{z})F(\boldsymbol{z}). \tag{21}$$

Consider the diagonal set $\mathcal{D} \coloneqq \{(\boldsymbol{\theta},\boldsymbol{\phi}) : \boldsymbol{\phi} = \boldsymbol{\theta}\}$. On $\mathcal{D}$, skew-symmetry gives

$$F(\boldsymbol{\theta},\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{g}(\boldsymbol{\theta}) \\ \boldsymbol{g}(\boldsymbol{\theta}) \end{pmatrix}, \qquad \boldsymbol{g}(\boldsymbol{\theta}) \coloneqq \boldsymbol{S}(\boldsymbol{p}(\boldsymbol{\theta}))\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) = \mathrm{diag}(\boldsymbol{p}(\boldsymbol{\theta}))\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}),$$

where the last equality follows from $\boldsymbol{p}(\boldsymbol{\theta})^\top \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) = 0$.

Moreover, differentiating the restriction of F to the diagonal gives, for every $\boldsymbol{v} \in \mathbb{R}^K$,

$$DF(\boldsymbol{\theta},\boldsymbol{\theta})\begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} D\boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{v} \\ D\boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{v} \end{pmatrix}.$$

Taking $\boldsymbol{v} = \boldsymbol{g}(\boldsymbol{\theta})$ shows that both components of Equation (21) have identical velocities whenever $\boldsymbol{\phi} = \boldsymbol{\theta}$. Hence the diagonal is invariant. Therefore, if the two players are initialized identically, they remain identical, and the two-player modified flow reduces exactly to (S2P-EG-Flow). Thus self-play is an invariant symmetry reduction of the full two-player modified flow.

Well-posedness. We prove that the solution of (S2P-EG-Flow) is well-posed and unique at any time.

Proposition C.1 (Global Well-Posedness). *For every skew-symmetric $\boldsymbol{A} \in \mathbb{R}^{K\times K}$, $\gamma > 0$, and $\boldsymbol{\theta}_0 \in \mathbb{R}^K$, there exists a unique global solution $\boldsymbol{\theta} : [0,\infty) \to \mathbb{R}^K$ of (S2P-EG-Flow) satisfying $\boldsymbol{\theta}(0) = \boldsymbol{\theta}_0$.*

Proof. Define the vector field

$$\mathcal{V}_\gamma(\boldsymbol{\theta}) \coloneqq -\boldsymbol{g}(\boldsymbol{\theta}) + \frac{\gamma}{2} D\boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{\theta}).$$

Writing $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{\theta})$, we have

$$D\boldsymbol{g}(\boldsymbol{\theta}) = \operatorname{diag}(\boldsymbol{A}\boldsymbol{p})\boldsymbol{S}(\boldsymbol{p}) + \operatorname{diag}(\boldsymbol{p})\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p}). \tag{22}$$

For every $\boldsymbol{p} \in \Delta_K$, $\|\boldsymbol{p}\|_2 \le 1$ and $\|\operatorname{diag}(\boldsymbol{p})\|_2 \le 1$, hence $\|\boldsymbol{g}(\boldsymbol{\theta})\|_2 \le \|\boldsymbol{A}\|_2$. The next lemma (whose proof is provided at the end) provides a norm bound for $\mathrm{D}\boldsymbol{g}(\boldsymbol{p})$:

Lemma C.1. *For every $\boldsymbol{p} \in \Delta_K$, $\|\mathrm{D}\boldsymbol{g}(\boldsymbol{p})\|_2 \le \|\boldsymbol{A}\|_2$.*

Consequently, the modified-flow vector field is uniformly bounded:

$$\|\mathcal{V}_\gamma(\boldsymbol{\theta})\|_2 \le \|\boldsymbol{A}\|_2 + \frac{\gamma}{2}\|\boldsymbol{A}\|_2^2 =: C_{\boldsymbol{A},\gamma}.$$

Then, the statement immediately follows from Theorem B.1. □

Proof of Lemma C.1. For a unit vector $\boldsymbol{v}$, $\boldsymbol{v}^\top \boldsymbol{S}(\boldsymbol{p})\boldsymbol{v} = \sum_k p_k v_k^2 - \left(\sum_k p_k v_k\right)^2 = \operatorname{Var}_{\boldsymbol{p}}(\boldsymbol{v}) \ge 0$, the variance of the random variable taking value v_k with probability p_k; in particular $\boldsymbol{S}(\boldsymbol{p}) \succeq 0$ and $\|\boldsymbol{S}(\boldsymbol{p})\|_2 = \max_{\|\boldsymbol{v}\|_2=1} \operatorname{Var}_{\boldsymbol{p}}(\boldsymbol{v})$. By Popoviciu's inequality, $\operatorname{Var}_{\boldsymbol{p}}(\boldsymbol{v}) \le \frac{1}{4}\left(\max_k v_k - \min_k v_k\right)^2$, and $\left(\max_k v_k - \min_k v_k\right)^2 \le 2\left((\max_k v_k)^2 + (\min_k v_k)^2\right) \le 2\|\boldsymbol{v}\|_2^2 = 2$, whence $\|\boldsymbol{S}(\boldsymbol{p})\|_2 \le \frac{1}{2}$. Consequently,

$$\begin{aligned}\|\mathrm{D}g(\boldsymbol{p})\|_2 &\le \|\operatorname{diag}(\boldsymbol{A}\boldsymbol{p})\|_2 \, \|\boldsymbol{S}(\boldsymbol{p})\|_2 + \|\operatorname{diag}(\boldsymbol{p})\|_2 \, \|\boldsymbol{A}\|_2 \, \|\boldsymbol{S}(\boldsymbol{p})\|_2 \\ &\le \|\boldsymbol{A}\|_2 \cdot \tfrac{1}{2} + 1 \cdot \|\boldsymbol{A}\|_2 \cdot \tfrac{1}{2} = \|\boldsymbol{A}\|_2 ,\end{aligned}$$

using $\|\operatorname{diag}(\boldsymbol{A}\boldsymbol{p})\|_2 = \|\boldsymbol{A}\boldsymbol{p}\|_\infty \le \|\boldsymbol{A}\boldsymbol{p}\|_2 \le \|\boldsymbol{A}\|_2\|\boldsymbol{p}\|_2 \le \|\boldsymbol{A}\|_2$ and $\|\operatorname{diag}(\boldsymbol{p})\|_2 \le 1$. □

D BOUNDARY SYSTEM LIMIT OF S2PE-EG-FLOW

Fix a nontrivial partition $[K] = \mathcal{I} \dot{\cup} \mathcal{L}$, and write

$$m := |\mathcal{I}|, \qquad l := |\mathcal{L}|, \qquad \boldsymbol{B} := \boldsymbol{A}_{\mathcal{L},\mathcal{I}}.$$

We first derive exact coordinates for this arbitrary partition. Flatness, $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = \mathbf{0}$, then yields the boundary normal form; the additional assumption $\rho(t) \to 0$ lets us identify policy limit sets through the boundary field via the seminal results of Benaïm & Hirsch (1996); Mischaikow et al. (1995).

D.1 BOUNDARY REDUCTION AND LIMIT SETS

Set $\rho(t) := \sum_{j\in\mathcal{L}} p_j(t)$ and recall the "blow-up" reparametrization:

$$\tau(t) := \int_0^t \rho(s)ds, \qquad x_i(t) := \frac{p_i(t)}{1-\rho(t)}\ (i \in \mathcal{I}), \qquad \lambda_j(t) := \frac{p_j(t)}{\rho(t)}\ (j \in \mathcal{L}), \tag{23}$$

which satisfies $(\rho(t), \boldsymbol{x}(t), \boldsymbol{\lambda}(t)) \in (0,1) \times \Delta_{\mathcal{I}}^{\circ} \times \Delta_{\mathcal{L}}^{\circ}$ at every finite time. The following lemma, whose proof is in Appendix D.2, provides the exact blow-up system that is equivalent to (S2P-EG-Flow):

Lemma D.1 (Exact Blow-Up System). *Let $V_k(t) := \dot{\theta}_k(t)$ be the logit velocity, and $\bar{V}_I := \sum_{i\in I} x_i V_i$ and $\bar{V}_L := \sum_{j\in L} \lambda_j V_j$ be the weighted logit velocities. Then, under the above reparametrization, (S2P-EG-Flow) is equivalent to the following blow-up system:*

$$\frac{d\rho}{d\tau} = (1-\rho)\big(\bar{V}_L - \bar{V}_I\big), \qquad \frac{dx_i}{d\tau} = \frac{x_i\big(V_i - \bar{V}_I\big)}{\rho}, \qquad \frac{d\lambda_j}{d\tau} = \frac{\lambda_j\big(V_j - \bar{V}_L\big)}{\rho}. \tag{BlowUp}$$

The following proposition, whose proof is in Appendix D.3, shows that under a structural condition $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = \mathbf{0}$ (flatness), the blow-up system defined over $(0,1) \times \Delta_{\mathcal{I}}^{\circ} \times \Delta_{\mathcal{L}}^{\circ}$ analytically extends to $[0,1] \times \Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$ via a normal form that involves the ***boundary system***:

Proposition D.1 (Analytic Extension of the Blow-Up System). *Suppose $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = \mathbf{0}$, and let us denote $\boldsymbol{B} := \boldsymbol{A}_{\mathcal{L},\mathcal{I}}$. Then, (BlowUp) analytically extends to $[0,1] \times \Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$ as follows: writing $\boldsymbol{z} := (\boldsymbol{x}, \boldsymbol{\lambda})$,*

$$\begin{aligned} \frac{d\boldsymbol{z}}{d\tau} &= \mathcal{B}(\boldsymbol{z}) + \rho E_\gamma(\rho, \boldsymbol{z}), \\ \frac{d\rho}{d\tau} &= -\rho(1-\rho)\big(\kappa(\boldsymbol{z}) + \rho e_\gamma(\rho, \boldsymbol{z})\big), \end{aligned} \tag{24}$$

where E_γ and e_γ are polynomial in $(\rho, \boldsymbol{x}, \boldsymbol{\lambda})$ and affine in γ, and

$$\boldsymbol{u} := \mathrm{diag}(\boldsymbol{x})\boldsymbol{B}^\top\boldsymbol{\lambda}, \qquad \boldsymbol{w} := -\mathrm{diag}(\boldsymbol{\lambda})\boldsymbol{B}\boldsymbol{x}, \qquad \kappa(\boldsymbol{z}) := \sum_{i\in\mathcal{I}} x_i^2(\boldsymbol{B}^\top\boldsymbol{\lambda})_i + \sum_{j\in\mathcal{L}} \lambda_j^2(\boldsymbol{B}\boldsymbol{x})_j. \tag{25}$$

*The **boundary system (field)** $\mathcal{B}(\boldsymbol{z})$ is defined as*

$$\mathcal{B}: \quad \begin{aligned} \frac{dx_i}{d\tau} &= x_i\left(x_i(\boldsymbol{B}^\top\boldsymbol{\lambda})_i - \langle \boldsymbol{x}, \boldsymbol{u}\rangle\right), \\ \frac{d\lambda_j}{d\tau} &= -\lambda_j\left(\lambda_j(\boldsymbol{B}\boldsymbol{x})_j + \langle \boldsymbol{\lambda}, \boldsymbol{w}\rangle\right). \end{aligned} \tag{26}$$

Since $d\tau/dt = \rho$, (24) also gives the following system in "physical time" t:

$$\dot{\boldsymbol{z}} = \rho\mathcal{B}(\boldsymbol{z}) + \rho^2 E_\gamma(\rho, \boldsymbol{z}), \qquad \dot{\rho} = -\rho^2(1-\rho)\big(\kappa(\boldsymbol{z}) + \rho e_\gamma(\rho, \boldsymbol{z})\big). \tag{27}$$

Thus, the reparametrized $\boldsymbol{z}$-system (24) is an *asymptotically autonomous perturbation* of $\mathcal{B}$. The following theorem shows that in order to characterize the omega-limit set of (S2P-EG-Flow), it suffices to do so for $\mathcal{B}$:

Theorem D.1 ($\mathcal{B}$ Suffices!). *In addition to $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = \mathbf{0}$, suppose that $\rho(t) \to 0$ as $t \to \infty$. Let $\Omega_{\mathcal{B}} \coloneqq \omega(\boldsymbol{z})$ be the omega-limit set of the reparametrized trajectory $\boldsymbol{z}(\tau) = (\boldsymbol{x}(\tau), \boldsymbol{\lambda}(\tau))$. Then $\Omega_{\mathcal{B}}$ is nonempty, compact, connected, invariant, and internally chain transitive with respect to the flow generated by $\mathcal{B}$. Moreover, the omega-limit set of* (S2P-EG-Flow) *is*

$$\omega(\boldsymbol{p}) = \{(\boldsymbol{x}, \mathbf{0}) \in \Delta_K : (\boldsymbol{x}, \boldsymbol{\lambda}) \in \Omega_{\mathcal{B}} \text{ for some } \boldsymbol{\lambda} \in \Delta_{\mathcal{L}}\} \subseteq \Delta_{\mathcal{I}}.$$

Essentially, the above theorem converts the implicit bias problem into a classification of the chain-transitive sets of $\mathcal{B}$, which will be the main focus of Appendix I later.

We now present the proof of the above theorem:

Proof of Theorem D.1. The proof has three steps. We first show that the boundary clock diverges. We then use **Lemma B.2** to show that the reparametrized $\boldsymbol{z}$-trajectory is an APT of $\mathcal{B}$, which by **Theorem B.6** gives that $\Omega_{\mathcal{B}}$ is nonempty, compact, connected, invariant, and internally chain transitive for the $\mathcal{B}$-flow. Finally, we identify the omega-limit set of (S2P-EG-Flow) in the original $\boldsymbol{p}$-coordinates by blowing-down the trajectory.

Step 1: $\tau(t)$ diverges. In (27), continuity of κ and e_γ on the compact set $[0,1] \times \Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$ bounds the coefficient of $-\rho^2$. Thus $\dot{\rho} \geq -C\rho^2$ for some $C > 0$, and integration yields

$$\rho(t) \geq \frac{1}{\rho(0)^{-1} + Ct} \quad \Longrightarrow \quad \tau(t) \geq \frac{1}{C} \log\big(1 + C\rho(0)t\big) \longrightarrow \infty.$$

Step 2: The reparametrized trajectory is an APT of $\mathcal{B}$. Here, we verify the hypotheses of **Lemma B.2** for the reparametrized $\boldsymbol{z}$-trajectory, which implies that the reparametrized $\boldsymbol{z}$-trajectory is an APT of $\mathcal{B}$. Then, by Theorem B.6, $\Omega_{\mathcal{B}}$ is nonempty, compact, connected, invariant, and internally chain transitive for the $\mathcal{B}$-flow.

We need to show that

- (i) Locally uniformly convergence of the $\boldsymbol{z}$-ODE of (24) to $\mathcal{B}$ as $\tau \to \infty$.
- (ii) Local Lipschitz continuity of $\mathcal{B}$.
- (iii) Existence of a global flow generated by $\mathcal{B}$ on $\Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$.

Since $\rho(t) > 0$ at every finite time and $\tau(t) \to \infty$, the clock has an inverse $t = t(\tau)$. Moreover, the assumption $\rho(t) \to 0$ gives $\rho(t(\tau)) \to 0$. Hence the $\boldsymbol{z}$-ODE of (24) satisfies

$$\sup_{\boldsymbol{z} \in \Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}} \|\mathcal{B}(\boldsymbol{z}) + \rho(t(\tau)) E_\gamma(\rho(t(\tau)), \boldsymbol{z}) - \mathcal{B}(\boldsymbol{z})\| \longrightarrow 0,$$

verifying (i). ALso, as the fields are polynomial and therefore uniformly Lipschitz on the compact simplex product, (ii) is also verified.

To verify (iii), we first show that boundary trajectories stay in $\Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$. Summing the equations in (26) shows that both coordinate sums remain equal to one if they initially equal one. Along any solution, each coordinate equation has the form $y' = a(\tau)y$, so each coordinate retains its initial sign in both time directions. Thus initially nonnegative coordinates remain nonnegative, and the trajectory stays in the simplex product throughout its existence interval.

Since $\mathcal{B}$ is polynomial on the ambient Euclidean space, local solutions exist uniquely near every point of the simplex product. These solutions remain in a compact set, so the standard continuation argument extends them for all positive and negative times. Consequently, $\mathcal{B}$ generates a global flow on $\Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$, verifying (iii).

Step 3: Blow-down to (S2P-EG-Flow)**.** It remains to prove the displayed identity in the theorem. If $\boldsymbol{q} \in \omega(\boldsymbol{p})$, choose $t_n \to \infty$ such that $\boldsymbol{p}(t_n) \to \boldsymbol{q}$. After passing to a subsequence, compactness gives $\boldsymbol{z}(\tau(t_n)) \to (\boldsymbol{x}, \boldsymbol{\lambda})$ for some $(\boldsymbol{x}, \boldsymbol{\lambda}) \in \Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$. Since $\tau(t_n) \to \infty$, this limit belongs to $\Omega_{\mathcal{B}}$; since $\boldsymbol{p} = ((1-\rho)\boldsymbol{x}, \rho\boldsymbol{\lambda})$ and $\rho(t_n) \to 0$, we have $\boldsymbol{q} = (\boldsymbol{x}, \mathbf{0})$. Conversely, if $(\boldsymbol{x}, \boldsymbol{\lambda}) \in \Omega_{\mathcal{B}}$, choose $\tau_n \to \infty$ such that $\boldsymbol{z}(\tau_n) \to (\boldsymbol{x}, \boldsymbol{\lambda})$ and set $t_n = t(\tau_n)$. Then $t_n \to \infty$, $\rho(t_n) \to 0$, and therefore $\boldsymbol{p}(t_n) \to (\boldsymbol{x}, \mathbf{0})$. This proves both inclusions in the blow-down identity. □

D.2 Proof of Lemma D.1: Exact Blow-Up System

Recall that $V_k(t) := \dot{\theta}_k(t)$ is the coordinate-wise logit velocity. For notational simplicity, we omit (t). Then, for each k, taking the time-derivative on both sides of $p_k = \frac{\exp(\theta_k)}{\sum_{r=1}^K \exp(\theta_r)}$ yields the standard replicator-like equation:

$$\dot{p}_k = p_k \left(\dot{\theta}_k - \sum_{r=1}^{K} p_r \dot{\theta}_r \right) = p_k (V_k - \bar{V}), \tag{28}$$

where we denote $\bar{V} := \sum_{r=1}^K p_r V_r$ as the global average velocity. Note that this holds concurrently with the conservation of logits, $\sum_{k=1}^K \dot{\theta}_k = 0$.

We further note that $\bar{V}$ can be decomposed as $\bar{V} = \sum_{i \in I} p_i V_i + \sum_{j \in L} p_j V_j$. Then, as $p_i = (1-\rho)x_i$ and $p_j = \rho \lambda_j$, we have that

$$\bar{V} = (1-\rho) \underbrace{\sum_{i \in I} x_i V_i}_{\triangleq \bar{V}_I} + \rho \underbrace{\sum_{j \in L} \lambda_j V_j}_{\triangleq \bar{V}_L}. \tag{29}$$

With this, we now derive the ODE for each variable.

ODE for ρ. Taking the time-derivative on both sides of $\rho = \sum_{j \in L} p_j$,

$$\begin{aligned}
\dot{\rho} = \sum_{j \in L} \dot{p}_j &= \sum_{j \in L} p_j (V_j - \bar{V}) && \text{(Eqn. (28))} \\
&= \rho \left(\sum_{j \in L} \lambda_j V_j \right) - \rho \bar{V} \\
&= \rho (\bar{V}_L - \bar{V}) \\
&= \rho (1-\rho)(\bar{V}_L - \bar{V}_I). && \text{(Eqn. (29))}
\end{aligned}$$

ODE for $\boldsymbol{x}$. Applying the quotient rule to $x_i = \frac{p_i}{1-\rho}$:

$$\begin{aligned}
\dot{x}_i &= \frac{\dot{p}_i (1-\rho) - p_i (-\dot{\rho})}{(1-\rho)^2} \\
&= \frac{p_i (V_i - \bar{V})(1-\rho) + p_i \rho (1-\rho)(\bar{V}_L - \bar{V}_I)}{(1-\rho)^2} && \text{(ODE for } \rho) \\
&= x_i \big(V_i - \bar{V} + \rho (\bar{V}_L - \bar{V}_I) \big) && (x_i = \tfrac{p_i}{1-\rho}) \\
&= x_i (V_i - \bar{V}_I). && \text{(Eqn. (29))}
\end{aligned}$$

ODE for $\boldsymbol{\lambda}$. Similarly, applying the quotient rule to $\lambda_j = \frac{p_j}{\rho}$:

$$\begin{aligned}
\dot{\lambda}_j &= \frac{\dot{p}_j \rho - p_j \dot{\rho}}{\rho^2} \\
&= \frac{p_j (V_j - \bar{V}) \rho - p_j \rho (1-\rho)(\bar{V}_L - \bar{V}_I)}{\rho^2} && \text{(ODE for } \rho) \\
&= \lambda_j \big(V_j - \bar{V} - (1-\rho)(\bar{V}_L - \bar{V}_I) \big) && (\lambda_j = \tfrac{p_j}{\rho}) \\
&= \lambda_j (V_j - \bar{V}_L). && \text{(Eqn. (29))}
\end{aligned}$$

Then, we conclude by applying the chain rule for the time variable, $\frac{d}{d\tau} = \frac{dt}{d\tau} \frac{d}{dt} = \frac{1}{\rho} \frac{d}{dt}$. □

D.3 Proof of Proposition D.1: Analytic Extension of Blow-Up System

The proof has three steps. First, we factor one power of ρ from both $\boldsymbol{g}$ and $\mathrm{D}\boldsymbol{g}$, and hence from the logit velocity. Second, we substitute this factorization into the exact identities of Lemma D.1 and evaluate the resulting polynomial field at $\rho = 0$. Finally, we factor the remaining terms that vanish at $\rho = 0$ to obtain E_γ and e_γ.

Step 1: Factorization of the logit velocity. By flatness and skew-symmetry,

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{B}^\top \\ \boldsymbol{B} & \boldsymbol{A}_{\mathcal{L},\mathcal{L}} \end{pmatrix}.$$

Since $\boldsymbol{p} = ((1-\rho)\boldsymbol{x}, \rho\boldsymbol{\lambda})$, its block payoffs are

$$\begin{aligned} (\boldsymbol{A}\boldsymbol{p})_i &= -\rho(\boldsymbol{B}^\top\boldsymbol{\lambda})_i, && i \in \mathcal{I}, \\ (\boldsymbol{A}\boldsymbol{p})_j &= (1-\rho)(\boldsymbol{B}\boldsymbol{x})_j + \rho(\boldsymbol{A}_{\mathcal{L},\mathcal{L}}\boldsymbol{\lambda})_j, && j \in \mathcal{L}. \end{aligned} \tag{30}$$

Consequently, writing $g_k = p_k(\boldsymbol{A}\boldsymbol{p})_k$,

$$\begin{aligned} g_i &= -\rho(1-\rho)x_i(\boldsymbol{B}^\top\boldsymbol{\lambda})_i, \\ g_j &= \rho\lambda_j\big[(1-\rho)(\boldsymbol{B}\boldsymbol{x})_j + \rho(\boldsymbol{A}_{\mathcal{L},\mathcal{L}}\boldsymbol{\lambda})_j\big]. \end{aligned}$$

Thus $\boldsymbol{g} = \rho\widetilde{\boldsymbol{g}}$ for a polynomial vector field $\widetilde{\boldsymbol{g}}$.

We next factor $\mathrm{D}\boldsymbol{g}$. Direct differentiation gives, for every $k, r \in [K]$,

$$(\mathrm{D}\boldsymbol{g})_{kr} = p_k\Big[(\mathbf{1}_{\{k=r\}} - p_r)(\boldsymbol{A}\boldsymbol{p})_k + p_r A_{kr} - p_r(\boldsymbol{A}\boldsymbol{p})_k\Big],$$

and every entry contains a factor ρ. Indeed, if $k \in \mathcal{L}$, the factor comes from $p_k = \rho\lambda_k$. If $k \in \mathcal{I}$, both payoff terms contain ρ by (30); in the remaining term $p_r A_{kr}$, either $r \in \mathcal{I}$ and $A_{kr} = 0$, or $r \in \mathcal{L}$ and $p_r = \rho\lambda_r$. Therefore

$$\mathrm{D}\boldsymbol{g} = \rho M, \qquad \dot{\boldsymbol{\theta}} = \rho\widetilde{\mathcal{V}}, \qquad \widetilde{\mathcal{V}} = -\widetilde{\boldsymbol{g}} + \frac{\gamma}{2}\rho M\widetilde{\boldsymbol{g}},$$

for some polynomials M and $\widetilde{\mathcal{V}}$.

Step 2: Identification of the boundary field. Substitute $V_k = \rho\widetilde{\mathcal{V}}_k$ into Lemma D.1. The apparent factors $1/\rho$ cancel, leaving the polynomial system

$$\frac{dx_i}{d\tau} = x_i\left(\widetilde{\mathcal{V}}_i - \sum_{r\in\mathcal{I}} x_r\widetilde{\mathcal{V}}_r\right), \quad \frac{d\lambda_j}{d\tau} = \lambda_j\left(\widetilde{\mathcal{V}}_j - \sum_{s\in\mathcal{L}} \lambda_s\widetilde{\mathcal{V}}_s\right), \quad \frac{d\rho}{d\tau} = \rho(1-\rho)h_\gamma(\rho, \boldsymbol{z}), \tag{31}$$

where $h_\gamma(\rho, \boldsymbol{z}) := \sum_{j\in\mathcal{L}} \lambda_j\widetilde{\mathcal{V}}_j - \sum_{i\in\mathcal{I}} x_i\widetilde{\mathcal{V}}_i$. This already gives a polynomial extension to the closed simplex product.

At $\rho = 0$, the γ-dependent term in $\widetilde{\mathcal{V}}$ vanishes and

$$\widetilde{\mathcal{V}}_i = x_i(\boldsymbol{B}^\top\boldsymbol{\lambda})_i = u_i, \qquad \widetilde{\mathcal{V}}_j = -\lambda_j(\boldsymbol{B}\boldsymbol{x})_j = w_j,$$

so the first two equations reduce to (26). Moreover,

$$h_\gamma(0, \boldsymbol{z}) = \langle\boldsymbol{\lambda}, \boldsymbol{w}\rangle - \langle\boldsymbol{x}, \boldsymbol{u}\rangle = -\kappa(\boldsymbol{z}).$$

Step 3: Factorization of the remainders. Let $F_\gamma(\rho, \boldsymbol{z})$ denote the tangential $(\boldsymbol{x}, \boldsymbol{\lambda})$-field obtained in Step 2. Both F_γ and h_γ are polynomial, and Step 2 shows that

$$F_\gamma(0, \boldsymbol{z}) = \mathcal{B}(\boldsymbol{z}), \qquad h_\gamma(0, \boldsymbol{z}) = -\kappa(\boldsymbol{z}).$$

Therefore, their respective differences from $\mathcal{B}$ and $-\kappa$ are divisible by ρ. Thus there exist polynomial vector fields E_γ and e_γ such that

$$F_\gamma(\rho, \boldsymbol{z}) = \mathcal{B}(\boldsymbol{z}) + \rho E_\gamma(\rho, \boldsymbol{z}), \qquad h_\gamma(\rho, \boldsymbol{z}) = -\kappa(\boldsymbol{z}) - \rho e_\gamma(\rho, \boldsymbol{z}).$$

All dependence on γ entered linearly through $\widetilde{\mathcal{V}}$, so E_γ and e_γ are affine in γ. Substitution into the system from **Step 2** proves (24).

□

E KEY TECHNICAL LEMMAS

We collect four lemmas for (S2P-EG-Flow) used in the convergence and implicit-bias proofs. The first three apply to general skew-symmetric games; the last gives capture and decay under flat dominance.

Lemma E.1 (Component-wise Logit Conservation). *Let $\mathcal{G}_{\boldsymbol{A}}$ be the undirected graph on $[K]$ with an edge $\{i,j\}$ whenever $\boldsymbol{A}_{ij} \neq 0$. If C is a connected component of $\mathcal{G}_{\boldsymbol{A}}$, then*

$$\sum_{i\in C} \theta_i(t) = \sum_{i\in C} \theta_i(0), \quad \forall t \geq 0.$$

In particular, $\sum_{i\in[K]} \theta_i(t)$ and $\prod_{i\in[K]} e^{-\theta_i(t)}$ are constant.

Proof. Let $\boldsymbol{v} \in \mathbb{R}^K$ satisfy $v_i = v_j$ whenever $A_{ij} \neq 0$. Then skew-symmetry gives

$$\boldsymbol{v}^\top g(\boldsymbol{\theta}) = \sum_{i<j} A_{ij}(v_i - v_j)p_i p_j = 0.$$

Differentiating this identity yields $\boldsymbol{v}^\top \mathrm{D}g(\boldsymbol{\theta}) = \mathbf{0}^\top$. Hence

$$\frac{d}{dt}\boldsymbol{v}^\top\boldsymbol{\theta} = -\boldsymbol{v}^\top g(\boldsymbol{\theta}) + \frac{\gamma}{2}\boldsymbol{v}^\top \mathrm{D}g(\boldsymbol{\theta})g(\boldsymbol{\theta}) = 0.$$

Taking $\boldsymbol{v} = (\mathbb{1}_{[i\in C]})_{i\in[K]}$ proves the first assertion. Summing over all connected components gives conservation of $\sum_{i=1}^K \theta_i(t)$. □

We denote $\boldsymbol{S}(\boldsymbol{p}) = \mathrm{diag}(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^\top$. Recall that $\boldsymbol{g} = \boldsymbol{p} \odot \boldsymbol{A}\boldsymbol{p}$ and $\mathrm{D}\boldsymbol{g} = \mathrm{diag}(\boldsymbol{A}\boldsymbol{p})\boldsymbol{S}(\boldsymbol{p}) + \mathrm{diag}(\boldsymbol{p})\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p})$ depend on $\boldsymbol{\theta}$ only through $\boldsymbol{p}$. Also, viewed as functions of $\boldsymbol{p}$, both $\boldsymbol{g}$ and $D\boldsymbol{g}$ admit polynomial extensions to Δ_K.

Lemma E.2 (Inverse-Weight Dynamics). *For $k \in [K]$, let $u_k(t) := e^{-\theta_k(t)}$ and $Z(t) := \sum_{r=1}^K e^{\theta_r(t)}$. Along (S2P-EG-Flow), we have that $\dot{u}_k = \frac{F_k(\boldsymbol{p})}{Z}$, where*

$$F_k(\boldsymbol{p}) := (\boldsymbol{A}\boldsymbol{p})_k - \frac{\gamma}{2}B_k(\boldsymbol{p}), \quad B_k(\boldsymbol{p}) := (\boldsymbol{A}\boldsymbol{p})_k\big(g_k - \boldsymbol{p}^\top\boldsymbol{g}\big) + \big(\boldsymbol{A}\,\boldsymbol{S}(\boldsymbol{p})\,\boldsymbol{g}\big)_k. \tag{32}$$

Equivalently, the logit field satisfies $\dot{\theta}_k = -p_k F_k(\boldsymbol{p})$. Moreover, if $M := \|\boldsymbol{A}\|_{\max}$, then $|F_k(\boldsymbol{p})| \leq M + 2\gamma M^2$. Finally, $\boldsymbol{g}(\boldsymbol{q}) = \mathbf{0}$ implies $F_k(\boldsymbol{q}) = (\boldsymbol{A}\boldsymbol{q})_k$.

Proof. From $u_k = e^{-\theta_k}$ and (S2P-EG-Flow), $\dot{u}_k = -u_k\dot{\theta}_k = u_k g_k - \frac{\gamma}{2}u_k(\mathrm{D}\boldsymbol{g}\,\boldsymbol{g})_k$. $u_k p_k = 1/Z$ gives $u_k g_k = u_k p_k(\boldsymbol{A}\boldsymbol{p})_k = (\boldsymbol{A}\boldsymbol{p})_k/Z$. For the second-order term, taking the k-th row of $\mathrm{D}\boldsymbol{g} = \mathrm{diag}(\boldsymbol{A}\boldsymbol{p})\boldsymbol{S}(\boldsymbol{p}) + \mathrm{diag}(\boldsymbol{p})\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p})$ and using $(\boldsymbol{S}(\boldsymbol{p})\boldsymbol{g})_k = p_k(g_k - \boldsymbol{p}^\top\boldsymbol{g})$,

$$(\mathrm{D}\boldsymbol{g}\,\boldsymbol{g})_k = (\boldsymbol{A}\boldsymbol{p})_k\big(\boldsymbol{S}(\boldsymbol{p})\boldsymbol{g}\big)_k + p_k\big(\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p})\boldsymbol{g}\big)_k = p_k\,B_k(\boldsymbol{p}),$$

so $u_k(\mathrm{D}\boldsymbol{g}\,\boldsymbol{g})_k = B_k(\boldsymbol{p})/Z$, proving (32). The equation $\dot{\theta}_k = -p_k F_k(\boldsymbol{p})$ follows from $\dot{u}_k = -u_k\dot{\theta}_k$. Finally, note that $|g_k|, |\boldsymbol{p}^\top\boldsymbol{g}| \leq M$ and $\|\boldsymbol{S}(\boldsymbol{p})\boldsymbol{g}\|_1 \leq 2M$. Thus $|B_k| \leq 4M^2$ and $|F_k| \leq M + 2\gamma M^2$. If $\boldsymbol{g}(\boldsymbol{q}) = 0$, then $B_k(\boldsymbol{q})$ vanishes. □

Lemma E.3 (Divergence of the Intrinsic Clock). *Along (S2P-EG-Flow), the intrinsic clock diverges:*

$$R(t) := \int_0^t \frac{ds}{Z(s)} \to \infty, \quad \text{as } t \to \infty. \tag{33}$$

Proof. Suppose that $R(\infty) < \infty$. By Lemma E.2,

$$|\dot{u}_i(t)| = \frac{|F_i(\boldsymbol{p}(t))|}{Z(t)} \leq \frac{M + 2\gamma M^2}{Z(t)}.$$

Hence, for every $t \geq 0$,

$$|u_i(t) - u_i(0)| \leq (M + 2\gamma M^2) \int_0^t \frac{ds}{Z(s)} \leq (M + 2\gamma M^2) R(\infty).$$

Since there are finitely many coordinates, there exists $U < \infty$ such that

$$0 < u_i(t) \leq U, \qquad \forall i \in [K],\ \forall t \geq 0.$$

By conservation of logits (Lemma E.1),

$$Q := \prod_{i=1}^{K} u_i(t) = \exp\left(-\sum_{i=1}^{K} \theta_i(0) \right) > 0$$

is constant. Therefore, for every i,

$$u_i(t) = \frac{Q}{\prod_{k \neq i} u_k(t)} \geq \frac{Q}{U^{K-1}}.$$

It follows that

$$Z(t) = \sum_{i=1}^{K} u_i(t)^{-1} \leq \frac{KU^{K-1}}{Q} \quad \Longrightarrow \quad R(t) = \int_0^t \frac{ds}{Z(s)} \geq \frac{Q}{KU^{K-1}}\, t \longrightarrow \infty,$$

contradicting $R(\infty) < \infty$. □

The final lemma uses the boundary normal form of Proposition D.1 to show that sufficiently small dominated mass keeps decreasing under flat dominance. For the partition in Assumption 1, write $m := |\mathcal{I}|$, $l := |\mathcal{L}|$ (where $\mathcal{L}$ is the set of *dominated* actions), $\boldsymbol{B} = \boldsymbol{A}_{\mathcal{L},\mathcal{I}}$, and recall

$$\rho(t) := \sum_{j \in \mathcal{L}} p_j(t), \qquad \tau(t) := \int_0^t \rho(s)\, ds, \qquad \kappa(\boldsymbol{z}) := \sum_{i \in \mathcal{I}} x_i^2 (\boldsymbol{B}^\top \boldsymbol{\lambda})_i + \sum_{j \in \mathcal{L}} \lambda_j^2 (\boldsymbol{B}\boldsymbol{x})_j.$$

Lemma E.4 (Flat-Dominance Capture). *Under Assumption 1, for every fixed $\gamma > 0$, there are $\varepsilon, c, C > 0$, depending only on $(\boldsymbol{A}, \gamma, \mathcal{I}, \mathcal{L})$, such that*

$$c\rho^2 \leq -\dot{\rho} \leq C\rho^2 \qquad \textit{whenever } 0 < \rho \leq \varepsilon. \tag{34}$$

The region $\{\rho \leq \varepsilon\}$ is forward invariant. Hence, if $\liminf_{t\to\infty} \rho(t) = 0$, then there is $T < \infty$ such that (34) holds for every $t \geq T$, and

$$\frac{1}{\rho(T)^{-1} + C(t-T)} \leq \rho(t) \leq \frac{1}{\rho(T)^{-1} + c(t-T)}, \quad \forall t \geq T. \tag{35}$$

In particular, we have the following explicit rates:

$$\rho(t) = \Theta(t^{-1}), \qquad \tau(t) = \Theta(\log t), \qquad \int_0^\infty \rho(t)^2\, dt < \infty. \tag{36}$$

Proof. We first bound the coefficient of $-\rho^2$ uniformly away from zero, then integrate the resulting differential estimate. Set $\mathcal{K} := \Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$. By flat dominance,

$$\kappa \geq b_- \left(\sum_{i \in \mathcal{I}} x_i^2 + \sum_{j \in \mathcal{L}} \lambda_j^2 \right) \geq b_- \left(\frac{1}{m} + \frac{1}{l} \right) =: \kappa_0 > 0,$$

where the second inequality follows from Cauchy–Schwarz on each simplex. By continuity and compactness, the quantities

$$M_\gamma := \sup_{(\rho, \boldsymbol{z}) \in [0,1] \times \mathcal{K}} |e_\gamma(\rho, \boldsymbol{z})|, \qquad \kappa_1 := \sup_{\boldsymbol{z} \in \mathcal{K}} \kappa(\boldsymbol{z})$$

are finite. Choose $0 < \varepsilon \le 1/2$ so that $\varepsilon M_\gamma \le \kappa_0/2$. For $0 < \rho \le \varepsilon$,

$$\frac{\kappa_0}{2} \le \kappa(\boldsymbol{z}) + \rho e_\gamma(\rho, \boldsymbol{z}) \le \kappa_1 + M_\gamma.$$

Now recall the ρ-ODE (in time t) from (27):

$$\dot{\rho} = -\rho^2(1-\rho)\big(\kappa(\boldsymbol{z}) + \rho e_\gamma(\rho, \boldsymbol{z})\big).$$

Since $1/2 \le 1 - \rho \le 1$, the above yields (34) with $c = \kappa_0/4$ and $C = \kappa_1 + M_\gamma$.

At the boundary $\rho = \varepsilon$, the lower bound gives $\dot{\rho} \le -c\varepsilon^2 < 0$; hence $\{\rho \le \varepsilon\}$ is forward invariant. If $\liminf_{t\to\infty} \rho(t) = 0$, the trajectory enters this region at some finite time T and never leaves it. On this tail $t \ge T$,

$$c \le \frac{d}{dt}\frac{1}{\rho} \le C$$

by (34). Integrating from T to t and inverting gives (35). Those bounds imply $\rho(t) = \Theta(t^{-1})$; integrating them once gives $\tau(t) = \Theta(\log t)$, while integrating their squares gives $\int_0^\infty \rho(t)^2\, dt < \infty$. This proves (36). □

F Proofs for Section 3.1: Global Convergence Behavior of the S2P-EG-Flow

F.1 Proof of Theorem 3.1: Interior SNE ⇒ Logit Convergence

We start by stating the Lyapunov potential construction; the proof of Lemma F.1 is deferred after that of Theorem 3.1.

Lemma F.1 (Anchored Lyapunov potential)**.** *Let $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$, set $S := \operatorname{supp}(\boldsymbol{q})$, and define $\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}) := \sum_i q_i e^{-\theta_i}$. If $0 < \gamma < \frac{2}{3\|\boldsymbol{A}\|_{\max}}$ and $\beta := 1 - \frac{3}{2}\gamma\|\boldsymbol{A}\|_{\max} > 0$, then*

$$\frac{d}{dt}\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}(t)) \le -\frac{\beta}{Z(t)}\sum_{j\notin S} p_j(t)\,(\boldsymbol{A}\boldsymbol{q})_j - \frac{\gamma}{2Z(t)}\sum_{i=1}^{K} q_i p_i(t)\big((\boldsymbol{A}\boldsymbol{p}(t))_i\big)^2 \;\le\; 0, \tag{37}$$

where $Z(t) := \sum_k e^{\theta_k(t)}$.[a] *In particular, $\Phi_{\boldsymbol{q}}$ is nonincreasing. If $\boldsymbol{q} \in \Delta_K^\circ$, then $\boldsymbol{A}\boldsymbol{q} = 0$ and*

$$\frac{d}{dt}\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}(t)) = -\frac{\gamma}{2Z(t)}\sum_{i=1}^{K} q_i p_i(t)\big((\boldsymbol{A}\boldsymbol{p}(t))_i\big)^2 \le 0, \tag{38}$$

holds for every $\gamma > 0$.

[a] Although for this proof we only use the interior case, the first part is later utilized to prove Theorem 3.2 under the assumption of flat dominance.

Take $\boldsymbol{c} \in \Delta_K^\circ$ with $\boldsymbol{A}\boldsymbol{c} = 0$. Set $s := \mathbf{1}^\top\boldsymbol{\theta}(0)$ and $\Phi_{\boldsymbol{c}}(\boldsymbol{\theta}) := \sum_i c_i e^{-\theta_i}$. By Lemma E.1 (logit conservation), the affine hyperplane $H_s := \{\boldsymbol{\theta} : \mathbf{1}^\top\boldsymbol{\theta} = s\}$ is invariant. By (38), for every $\gamma > 0$,

$$\frac{d}{dt}\Phi_{\boldsymbol{c}}(\boldsymbol{\theta}(t)) = -\frac{\gamma}{2Z(\boldsymbol{\theta}(t))}\sum_{i=1}^{K} c_i p_i(\boldsymbol{\theta}(t))\big(\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}(t))\big)_i^2 \le 0,$$

with equality if and only if $\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}(t)) = \mathbf{0}$. Consider the positively invariant sublevel set

$$\mathcal{K} := \{\boldsymbol{\theta} \in H_s : \Phi_{\boldsymbol{c}}(\boldsymbol{\theta}) \le \Phi_{\boldsymbol{c}}(\boldsymbol{\theta}(0))\}.$$

It is compact. Indeed, with $M := \Phi_{\boldsymbol{c}}(\boldsymbol{\theta}(0))$, every $\boldsymbol{\theta} \in \mathcal{K}$ satisfies $\theta_i \ge \log(c_i/M)$, while $\sum_i \theta_i = s$ supplies an upper bound for every coordinate.

Let $\mathcal{E} := \{\boldsymbol{\theta} \in H_s : \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) = 0\}$. Since

$$\{\boldsymbol{\theta} \in \mathcal{K} : \dot{\Phi}_c(\boldsymbol{\theta}) = 0\} = \mathcal{K} \cap \mathcal{E},$$

and every point of $\mathcal{E}$ is an equilibrium, $\mathcal{K} \cap \mathcal{E}$ is invariant and hence is the largest invariant subset of the zero-derivative set. Therefore, by LaSalle's invariance principle (Theorem B.3),

$$\omega(\boldsymbol{\theta}(0)) \subseteq \mathcal{K} \cap \mathcal{E}, \tag{39}$$

In particular, we can find a subsequence $\{t_n\}$ and a $\boldsymbol{\theta}^\star \in \omega(\boldsymbol{\theta}(0)) \cap \mathcal{E}$ such that $\boldsymbol{\theta}(t_n) \to \boldsymbol{\theta}^\star$.

We now provide a crucial characterization of $\mathcal{E}$ and $\boldsymbol{\theta}^\star$ (whose proof we provide at the end):

Lemma F.2. *$\mathcal{E}$ (as defined in (39)) is a normally attracting manifold of equilibria near $\boldsymbol{\theta}^\star$.*

By the local asymptotic-phase theorem (Theorem B.5), there is a forward-invariant stable neighborhood $\mathcal{U} \subseteq H_s$ of $\boldsymbol{\theta}^\star$ such that every solution in $\mathcal{U}$ converges exponentially to one point of $\mathcal{E}$. In particular, since $\boldsymbol{\theta}(t_n) \to \boldsymbol{\theta}^\star$ as $n \to \infty$, the trajectory indeed enters $\mathcal{U}$ at some finite time, and therefore, there exists a $\boldsymbol{\theta}^\infty \in \mathcal{E}$ (not necessarily $\boldsymbol{\theta}^\star$ at this point) such that

$$\boldsymbol{\theta}(t) \longrightarrow \boldsymbol{\theta}^\infty \in \mathcal{E} \qquad (t \to \infty).$$

The exponential estimate in (18) gives trajectory-dependent constants $T, C, \lambda > 0$ such that

$$\|\boldsymbol{\theta}(t) - \boldsymbol{\theta}^\infty\|_2 \le Ce^{-\lambda(t-T)} \qquad (t \ge T).$$

Finally, $\boldsymbol{\theta}^\star = \boldsymbol{\theta}^\infty \in \mathcal{E}$ means $\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}^\infty) = 0$, so $\boldsymbol{p}(\boldsymbol{\theta}^\infty) \in \mathcal{SN}(\boldsymbol{A}) \cap \Delta_K^\circ$. □

Below, we provide the proofs of Lemma F.1 and Lemma F.2.

Proof of Lemma F.1. Fix $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ and write $S := \operatorname{supp}(\boldsymbol{q})$. Recall the complementary-slackness property (104) of Lemma K.1: $(\boldsymbol{A}\boldsymbol{q})_i = 0$ for $i \in S$, and $(\boldsymbol{A}\boldsymbol{q})_j \geq 0$ for $j \notin S$. Throughout, we use the elementary but crucial identity $e^{-\theta_k} p_k = 1/Z$ for every k, and abbreviate $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{\theta})$, $g = \boldsymbol{g}(\boldsymbol{\theta})$, $Z = Z(\boldsymbol{\theta})$.

Step 1: the first-order term. Since $\nabla \Phi_{\boldsymbol{q}}(\boldsymbol{\theta}) = -\big(q_k e^{-\theta_k}\big)_{k\in[K]}$ and $g_k = p_k(\boldsymbol{A}\boldsymbol{p})_k$,

$$-\nabla\Phi_{\boldsymbol{q}}^{\top} g = \sum_k q_k e^{-\theta_k} p_k (\boldsymbol{A}\boldsymbol{p})_k = \frac{1}{Z}\,\boldsymbol{q}^{\top}\boldsymbol{A}\boldsymbol{p} = -\frac{1}{Z}\,\boldsymbol{p}^{\top}\boldsymbol{A}\boldsymbol{q} = -\frac{1}{Z}\sum_{j\notin S} p_j (\boldsymbol{A}\boldsymbol{q})_j =: -\frac{\tilde{P}}{Z}, \quad (40)$$

where we used skew-symmetry and (104); note $\tilde{P} \geq 0$.

Step 2: the second-order term. Recall $\mathrm{D}g = \operatorname{diag}(\boldsymbol{A}\boldsymbol{p})\boldsymbol{S}(\boldsymbol{p}) + \operatorname{diag}(\boldsymbol{p})\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p})$ and $(\boldsymbol{S}(\boldsymbol{p})g)_k = p_k(g_k - \boldsymbol{p}^{\top} g)$. Hence

$$\begin{aligned}\nabla\Phi_{\boldsymbol{q}}^{\top}\mathrm{D}g\, g &= -\sum_k q_k e^{-\theta_k}\Big[(\boldsymbol{A}\boldsymbol{p})_k\big(\boldsymbol{S}(\boldsymbol{p})g\big)_k + p_k\big(\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p})g\big)_k\Big] \\ &= -\frac{1}{Z}\sum_k q_k (\boldsymbol{A}\boldsymbol{p})_k\big(g_k - \boldsymbol{p}^{\top} g\big) - \frac{1}{Z}\,\boldsymbol{q}^{\top}\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p})g.\end{aligned}$$

For the first sum, $g_k = p_k(\boldsymbol{A}\boldsymbol{p})_k$ and (40) give

$$\sum_k q_k (\boldsymbol{A}\boldsymbol{p})_k\big(g_k - \boldsymbol{p}^{\top} g\big) = \sum_k q_k p_k (\boldsymbol{A}\boldsymbol{p})_k^2 - (\boldsymbol{p}^{\top} g)\,\boldsymbol{q}^{\top}\boldsymbol{A}\boldsymbol{p} = \sum_k q_k p_k (\boldsymbol{A}\boldsymbol{p})_k^2 + (\boldsymbol{p}^{\top} g)\tilde{P}.$$

For the second, skew-symmetry, (104), and $(\boldsymbol{S}(\boldsymbol{p})g)_j = p_j(g_j - \boldsymbol{p}^{\top} g)$ yield

$$\boldsymbol{q}^{\top}\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p})g = -(\boldsymbol{A}\boldsymbol{q})^{\top}\boldsymbol{S}(\boldsymbol{p})g = -\sum_{j\notin S}(\boldsymbol{A}\boldsymbol{q})_j p_j\big(g_j - \boldsymbol{p}^{\top} g\big) = -\sum_{j\notin S}(\boldsymbol{A}\boldsymbol{q})_j p_j^2 (\boldsymbol{A}\boldsymbol{p})_j + (\boldsymbol{p}^{\top} g)\tilde{P}.$$

Combining, we obtain the *exact identity*

$$\frac{d}{dt}\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}(t)) = -\frac{\tilde{P}}{Z} - \frac{\gamma}{2Z}\Big[\sum_k q_k p_k (\boldsymbol{A}\boldsymbol{p})_k^2 + 2(\boldsymbol{p}^{\top} g)\tilde{P} - \sum_{j\notin S}(\boldsymbol{A}\boldsymbol{q})_j p_j^2 (\boldsymbol{A}\boldsymbol{p})_j\Big]. \quad (41)$$

Step 3: absorbing the cross terms. Put $M := \|\boldsymbol{A}\|_{\max}$. Since $\boldsymbol{p} \in \Delta_K$, $|(\boldsymbol{A}\boldsymbol{p})_k| \leq M$ for every k, and hence $|\boldsymbol{p}^{\top} g| = |\sum_k p_k^2 (\boldsymbol{A}\boldsymbol{p})_k| \leq M$. Moreover,

$$\Big|\sum_{j\notin S}(\boldsymbol{A}\boldsymbol{q})_j p_j^2 (\boldsymbol{A}\boldsymbol{p})_j\Big| \leq M\sum_{j\notin S}(\boldsymbol{A}\boldsymbol{q})_j p_j^2 \leq M\,\tilde{P}.$$

Thus the two cross terms can reduce the bracket in (41) by at most $3M\tilde{P}$. With $\beta := 1 - 3\gamma M/2$, we obtain

$$\frac{d}{dt}\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}(t)) \leq -\frac{\beta\tilde{P}}{Z} - \frac{\gamma}{2Z}\sum_k q_k p_k (\boldsymbol{A}\boldsymbol{p})_k^2.$$

This is (37) whenever $0 < \gamma < 2/(3M)$.

Step 4: the interior case. If $S = [K]$, then (104) reads $\boldsymbol{A}\boldsymbol{q} = \boldsymbol{0}$ identically, so $\tilde{P} \equiv 0$ and the last sum in (41) is empty. The exact identity then collapses to

$$\frac{d}{dt}\Phi_{\boldsymbol{q}}(\boldsymbol{\theta}(t)) = -\frac{\gamma}{2Z}\sum_k q_k p_k (\boldsymbol{A}\boldsymbol{p})_k^2 \leq 0$$

for every $\gamma > 0$, with equality if and only if $\boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}(t)) = \boldsymbol{0}$. This proves (38). □

Proof of Lemma F.2. Recall that $H_s := \{\boldsymbol{\theta} \in \mathbb{R}^K : \mathbf{1}^\top \boldsymbol{\theta} = \mathbf{1}^\top \boldsymbol{\theta}(0)\}$ and $\mathcal{E} := \{\boldsymbol{\theta} \in H_s : \boldsymbol{A}\boldsymbol{p}(\boldsymbol{\theta}) = 0\}$. Let $\mathcal{V}_\gamma(\boldsymbol{\theta}) := -\boldsymbol{g}(\boldsymbol{\theta}) + \frac{\gamma}{2}\mathrm{D}\boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{\theta})$. The softmax map restricted to H_s is a real-analytic diffeomorphism onto Δ_K°, with inverse $\boldsymbol{p} \mapsto \log \boldsymbol{p} + \frac{1}{K}(s - \mathbf{1}^\top \log \boldsymbol{p})\mathbf{1}$. Moreover, $\mathcal{E}_p := \Delta_K^\circ \cap \ker \boldsymbol{A}$ is a relatively open subset of the affine space $\{\boldsymbol{p} : \boldsymbol{A}\boldsymbol{p} = 0,\ \mathbf{1}^\top \boldsymbol{p} = 1\}$. Hence its preimage $\mathcal{E}$ is an embedded real-analytic submanifold of H_s.

Fix $\boldsymbol{\theta} \in \mathcal{E}$, and abbreviate $\boldsymbol{p} := \boldsymbol{p}(\boldsymbol{\theta})$, $\boldsymbol{D}_{\boldsymbol{p}} := \mathrm{diag}(\boldsymbol{p})$, and $\boldsymbol{G}_{\boldsymbol{p}} := \boldsymbol{D}_{\boldsymbol{p}}\boldsymbol{A}\boldsymbol{D}_{\boldsymbol{p}}$. The derivative of softmax is $\boldsymbol{S}(\boldsymbol{p}) = \boldsymbol{D}_{\boldsymbol{p}} - \boldsymbol{p}\boldsymbol{p}^\top$. Since $\boldsymbol{A}\boldsymbol{p} = 0$, we have $\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p}) = \boldsymbol{A}\boldsymbol{D}_{\boldsymbol{p}}$, and therefore

$$T_{\boldsymbol{\theta}}\mathcal{E} = \{\boldsymbol{v} \in \mathbf{1}^\perp : \boldsymbol{A}\boldsymbol{D}_{\boldsymbol{p}}\boldsymbol{v} = 0\} = \ker(\boldsymbol{G}_{\boldsymbol{p}}) \cap \mathbf{1}^\perp.$$

At every point of $\mathcal{E}$, $\boldsymbol{g}(\boldsymbol{\theta}) = 0$. Consequently,

$$\mathrm{D}\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{G}_{\boldsymbol{p}}, \qquad \mathrm{D}\mathcal{V}_\gamma(\boldsymbol{\theta}) = -\boldsymbol{G}_{\boldsymbol{p}} + \frac{\gamma}{2}\boldsymbol{G}_{\boldsymbol{p}}^2. \tag{42}$$

The matrix $\boldsymbol{G}_{\boldsymbol{p}}$ is skew-symmetric and satisfies $\boldsymbol{G}_{\boldsymbol{p}}\mathbf{1} = 0$. Thus, setting $\mathcal{R}_{\boldsymbol{\theta}} := \mathrm{range}(\boldsymbol{G}_{\boldsymbol{p}})$, we obtain the orthogonal decomposition

$$\mathbf{1}^\perp = T_{\boldsymbol{\theta}}\mathcal{E} \oplus \mathcal{R}_{\boldsymbol{\theta}}.$$

Since $\boldsymbol{D}_{\boldsymbol{p}}$ is invertible, $\mathrm{rank}(\boldsymbol{G}_{\boldsymbol{p}}) = \mathrm{rank}(\boldsymbol{A})$. As $\mathcal{E}$ is an embedded submanifold of H_s, $\mathcal{R}_{\boldsymbol{\theta}}$ varies smoothly with $\boldsymbol{\theta}$.

It remains to prove uniform contraction on $\mathcal{R}_{\boldsymbol{\theta}}$. Let $r := \mathrm{rank}(\boldsymbol{A}) > 0$, and let $\sigma_r(\boldsymbol{G}_{\boldsymbol{p}})$ denote the smallest nonzero singular value of $\boldsymbol{G}_{\boldsymbol{p}}$. Since $\mathrm{rank}(\boldsymbol{G}_{\boldsymbol{p}}) = r$ and singular values depend continuously on the matrix entries, $\sigma_r(\boldsymbol{G}_{\boldsymbol{p}}) > 0$ varies continuously with $\boldsymbol{\theta}$. Choose a neighborhood $\mathcal{U}$ of $\boldsymbol{\theta}^\star$ in $\mathcal{E}$ whose closure is compact and contained in $\mathcal{E}$, and set

$$\sigma_\star := \min_{\boldsymbol{\theta} \in \overline{\mathcal{U}}} \sigma_r(\boldsymbol{G}_{\boldsymbol{p}}) > 0.$$

Because $\boldsymbol{G}_{\boldsymbol{p}}$ is skew-symmetric, $\mathcal{R}_{\boldsymbol{\theta}} = \mathrm{range}(\boldsymbol{G}_{\boldsymbol{p}}) = \ker(\boldsymbol{G}_{\boldsymbol{p}})^\perp$. Hence $\|\boldsymbol{G}_{\boldsymbol{p}}\boldsymbol{v}\|_2 \geq \sigma_\star \|\boldsymbol{v}\|_2$ for every $\boldsymbol{\theta} \in \overline{\mathcal{U}}$ and $\boldsymbol{v} \in \mathcal{R}_{\boldsymbol{\theta}}$.

Fix $\boldsymbol{\theta} \in \overline{\mathcal{U}}$ and $\boldsymbol{v} \in \mathcal{R}_{\boldsymbol{\theta}}$, and let $\boldsymbol{w}(t) := e^{t\mathrm{D}\mathcal{V}(\boldsymbol{\theta})}\boldsymbol{v}$. Since $\mathrm{D}\mathcal{V}(\boldsymbol{\theta})$ is a polynomial in $\boldsymbol{G}_{\boldsymbol{p}}$, the subspace $\mathcal{R}_{\boldsymbol{\theta}}$ is invariant, so $\boldsymbol{w}(t) \in \mathcal{R}_{\boldsymbol{\theta}}$. Using $\boldsymbol{G}_{\boldsymbol{p}}^\top = -\boldsymbol{G}_{\boldsymbol{p}}$, we obtain

$$\begin{aligned}\frac{d}{dt}\|\boldsymbol{w}(t)\|_2^2 &= -2\boldsymbol{w}(t)^\top \boldsymbol{G}_{\boldsymbol{p}}\boldsymbol{w}(t) + \gamma \boldsymbol{w}(t)^\top \boldsymbol{G}_{\boldsymbol{p}}^2 \boldsymbol{w}(t) \\ &= -\gamma \|\boldsymbol{G}_{\boldsymbol{p}}\boldsymbol{w}(t)\|_2^2 \leq -\gamma\sigma_\star^2 \|\boldsymbol{w}(t)\|_2^2.\end{aligned}$$

Grönwall's inequality therefore gives

$$\left\| e^{t\mathrm{D}\mathcal{V}(\boldsymbol{\theta})}\boldsymbol{v} \right\|_2 \leq e^{-\gamma\sigma_\star^2 t/2}\|\boldsymbol{v}\|_2. \tag{43}$$

Thus $\mathrm{D}\mathcal{V}(\boldsymbol{\theta})$ vanishes on $T_{\boldsymbol{\theta}}\mathcal{E}$ and contracts the complementary invariant subspaces $\mathcal{R}_{\boldsymbol{\theta}}$ uniformly near $\boldsymbol{\theta}^\star$. Hence $\mathcal{E}$ is normally attracting near $\boldsymbol{\theta}^\star$. □

F.2 Proof of Theorem 3.2: Flat Dominance ⇒ Nash Convergence

Let us denote $\rho(t) := \sum_{j \in \mathcal{L}} p_j(t) \in (0, 1)$.

Step 1: recurrent approach to the Nash face. Let $\boldsymbol{c} \in \Delta_{\mathcal{I}}$ be uniform on $\mathcal{I}$, zero-padded on $\mathcal{L}$: $c_i = 1/m$ for $i \in \mathcal{I}$. By flat dominance,

$$(\boldsymbol{A}\boldsymbol{c})_i = 0 \quad (i \in \mathcal{I}), \qquad (\boldsymbol{A}\boldsymbol{c})_j = (\boldsymbol{B}\boldsymbol{c}_{\mathcal{I}})_j \geq b_- \quad (j \in \mathcal{L}).$$

Thus $\boldsymbol{c} \in \mathcal{SN}(\boldsymbol{A})$. Set

$$\beta := 1 - \frac{3\gamma\|\boldsymbol{A}\|_{\max}}{2} > 0, \qquad R(t) := \int_0^t \frac{ds}{Z(s)}.$$

The first part of the anchored Lyapunov lemma (Eqn. (37) of Lemma F.1) gives

$$\dot{\Phi}_{\boldsymbol{c}}(\boldsymbol{\theta}(t)) \le -\frac{\beta}{Z(t)} \sum_{j \in \mathcal{L}} p_j(t)(\boldsymbol{A}\boldsymbol{c})_j \le -\frac{\beta b_-}{Z(t)} \rho(t).$$

Since $\Phi_{\boldsymbol{c}} \ge 0$, integration yields

$$\int_0^\infty \rho(t)\, dR(t) = \int_0^\infty \frac{\rho(t)}{Z(t)}\, dt \le \frac{\Phi_{\boldsymbol{c}}(\boldsymbol{\theta}(0))}{\beta b_-} < \infty. \tag{44}$$

By Lemma E.3, $R(t) \to \infty$. Therefore

$$\liminf_{t \to \infty} \rho(t) = 0; \tag{45}$$

otherwise ρ would be bounded below by a positive constant on a tail, contradicting (44). Lemma E.4 now gives a finite time T after which the capture region is forward invariant and the two-sided bounds (34) hold. In particular, $\rho(t) = \Theta(t^{-1})$.

Step 2: distance and exploitability. By Proposition K.2, $\mathcal{SN}(\boldsymbol{A}) = \Delta_{\mathcal{I}}$. For every $\boldsymbol{q} \in \Delta_{\mathcal{I}}$,

$$\|\boldsymbol{p} - \boldsymbol{q}\|_1 \ge \sum_{j \in \mathcal{L}} p_j + \left| \sum_{i \in I} (p_i - q_i) \right| = 2\rho.$$

Equality is attained at $\boldsymbol{q} = (\frac{\boldsymbol{p}_{\mathcal{I}}}{1-\rho}, \boldsymbol{0})$ supported on $\mathcal{I}$, thus

$$\mathrm{dist}_1(\boldsymbol{p}, \mathcal{SN}(\boldsymbol{A})) = 2\rho. \tag{46}$$

Since $\rho(t) \to 0$, choose $T_1 \ge T$ so that $\rho(t) \le b_-/[2(b_- + \|\boldsymbol{A}\|_{\max})]$ for $t \ge T_1$. Then (30) implies

$$(\boldsymbol{A}\boldsymbol{p})_j \ge (1-\rho)b_- - \rho\|\boldsymbol{A}\|_{\max} \ge \frac{b_-}{2} > 0 \qquad (j \in L),$$

whereas

$$(\boldsymbol{A}\boldsymbol{p})_i = -\rho(\boldsymbol{B}^\top \boldsymbol{\lambda})_i < 0 \qquad (i \in I).$$

Hence

$$-\mathrm{DGap}(\boldsymbol{p}) = \rho \max_{i \in \mathcal{I}} (\boldsymbol{B}^\top \boldsymbol{\lambda})_i,$$

and therefore

$$b_- \rho \le -\mathrm{DGap}(\boldsymbol{p}) \le b_+ \rho.$$

Since $\rho(t) \to 0$, every trajectory converges in the Nash sense. □

G Proofs for Section 3.2: Sharp Boundary Asymptotics

Here we prove the results in Section 3.2 and provide additional arguments. We first establish a strengthened form of the boundary-rate reduction: under convergence of the sum of the zero-margin boundary logits, the time change R, the softmax normalization, and the asymptotics of the supported and positive-margin coordinates are all determined, with explicit constants. The quasi-strict result then follows immediately because the zero-margin boundary set is empty. We next consider a simple non-quasi-strict setting, where the limiting equilibrium is pure and there is a single zero-margin boundary action, in which that logit can itself be shown to converge. Finally, we give an explicit example with multiple zero-margin boundary actions showing that convergence of the sum of the zero-margin boundary logits does not, in general, imply convergence of the individual zero-margin boundary logits.

G.1 Full boundary-rate reduction

Recall partition (2),

$$\mathcal{S} := \operatorname{supp}(\boldsymbol{q}), \qquad \mathcal{B} := \{j \notin \mathcal{S} : (\boldsymbol{A}\boldsymbol{q})_j = 0\}, \qquad \mathcal{M} := \{j \notin \mathcal{S} : (\boldsymbol{A}\boldsymbol{q})_j > 0\},$$

and write $s := |\mathcal{S}|$, $b := |\mathcal{B}|$, $m := |\mathcal{M}|$, and $\delta_j := (\boldsymbol{A}\boldsymbol{q})_j > 0$ for $j \in \mathcal{M}$. Following Lemma E.2, we abbreviate the inverse weights by $u_n(t) := e^{-\theta_n(t)}$, and set

$$Z(t) := \sum_{n=1}^{K} e^{\theta_n(t)}, \qquad R(t) := \int_0^t \frac{dr}{Z(r)}.$$

We call R the *intrinsic clock*: it measures time in units of the inverse partition function, so that the inverse-weight dynamics become $du_n = F_n(\boldsymbol{p})\, dR$, independent of the global softmax scale Z. Lemma E.3 gives $R(t) \to \infty$.

By logit conservation (Lemma E.1), the following quantity is constant:

$$Q := \prod_{n=1}^{K} u_n(t) = e^{-\sum_{n=1}^{K} \theta_n(0)}.$$

The following proposition isolates the clock–normalization reduction used to obtain the boundary rates.

Proposition G.1 (Clock–normalization reduction)**.** *Suppose $\boldsymbol{p}(t) \to \boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ and $m \geq 1$. Assume*

$$\sum_{k \in \mathcal{B}} \theta_k(t) \longrightarrow \beta_{\mathcal{B}} \in \mathbb{R}, \tag{47}$$

with the empty sum interpreted as 0, and define

$$C_{\boldsymbol{q}} := \frac{e^{-\beta_{\mathcal{B}}} \prod_{j \in \mathcal{M}} \delta_j}{Q \prod_{i \in \mathcal{S}} q_i}.$$

Then

$$\frac{u_j(t)}{R(t)} \longrightarrow \delta_j, \qquad j \in \mathcal{M},$$

and

$$\frac{Z(t)^s}{R(t)^m} \longrightarrow C_{\boldsymbol{q}}.$$

Proof. By Lemma E.2,

$$\dot{u}_n(t) = \frac{1}{Z(t)} F_n(\boldsymbol{p}(t)),$$

where

$$F_n(\boldsymbol{p}) := (\boldsymbol{A}\boldsymbol{p})_n - \frac{\gamma}{2} B_n(\boldsymbol{p}),$$

with

$$B_n(\boldsymbol{p}) := (\boldsymbol{A}\boldsymbol{p})_n\big(g_n(\boldsymbol{p}) - \boldsymbol{p}^\top g(\boldsymbol{p})\big) + (\boldsymbol{A}\,\boldsymbol{S}(\boldsymbol{p})\,g(\boldsymbol{p}))_n.$$

Since $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$, complementary slackness gives $g(\boldsymbol{q}) = \boldsymbol{q} \odot \boldsymbol{A}\boldsymbol{q} = \boldsymbol{0}$. Hence $B_n(\boldsymbol{p}(t)) \to 0$ by continuity. For every $j \in \mathcal{M}$, $(\boldsymbol{A}\boldsymbol{p}(t))_j \to (\boldsymbol{A}\boldsymbol{q})_j = \delta_j$, and therefore

$$F_j(\boldsymbol{p}(t)) \longrightarrow \delta_j.$$

By Lemma E.3, $R(t) \to \infty$. Since $dR = dt/Z(t)$,

$$u_j(t) = u_j(0) + \int_0^t F_j(\boldsymbol{p}(r))\, dR(r).$$

Write $F_j(\boldsymbol{p}(t)) = \delta_j + \varepsilon_j(t)$ with $\varepsilon_j(t) \to 0$. Then

$$\frac{u_j(t)}{R(t)} = \frac{u_j(0)}{R(t)} + \delta_j + \frac{1}{R(t)} \int_0^t \varepsilon_j(r)\, dR(r).$$

For any fixed $t_0 < t$,

$$\left| \int_0^t \varepsilon_j(r)\, dR(r) \right| \le \left| \int_0^{t_0} \varepsilon_j(r)\, dR(r) \right| + \sup_{r \ge t_0} |\varepsilon_j(r)|\, [R(t) - R(t_0)].$$

After division by $R(t)$, first taking $t \to \infty$ and then $t_0 \to \infty$, the error term vanishes. This gives $u_j(t)/R(t) \to \delta_j$.

For $i \in \mathcal{S}$, $u_i(t) = 1/(Z(t)p_i(t))$, so the conserved product Q decomposes as

$$Q = \left(\prod_{i \in \mathcal{S}} \frac{1}{Z(t)p_i(t)} \right) \left(\prod_{k \in \mathcal{B}} u_k(t) \right) \prod_{j \in \mathcal{M}} u_j(t).$$

The middle factor is $\exp\big(-\sum_{k \in \mathcal{B}} \theta_k(t)\big)$, which converges to $e^{-\beta_{\mathcal{B}}}$ by (47). Using $p_i(t) \to q_i > 0$ for $i \in \mathcal{S}$ and $u_j(t) = \delta_j R(t)(1 + o(1))$ for $j \in \mathcal{M}$,

$$Q = Z(t)^{-s} R(t)^m \, \frac{e^{-\beta_{\mathcal{B}}} \prod_{j \in \mathcal{M}} \delta_j}{\prod_{i \in \mathcal{S}} q_i} \, (1 + o(1)),$$

which gives $Z(t)^s / R(t)^m \to C_{\boldsymbol{q}}$.

From this ratio, $Z(t) = C_{\boldsymbol{q}}^{1/s} R(t)^{m/s}(1 + o(1))$, so, since $\dot{R} = 1/Z$,

$$\frac{d}{dt} R(t)^{(s+m)/s} = \frac{s+m}{s} \cdot \frac{R(t)^{m/s}}{Z(t)} \longrightarrow \frac{s+m}{s} C_{\boldsymbol{q}}^{-1/s}.$$

Therefore $R(t)^{(s+m)/s} = \frac{s+m}{s} C_{\boldsymbol{q}}^{-1/s}\, t\,(1 + o(1))$, and

$$R(t) = C_{\boldsymbol{q}}^{-1/(s+m)} \left(\frac{s+m}{s} \right)^{s/(s+m)} t^{s/(s+m)} (1 + o(1)).$$

Substituting into $Z(t) = C_{\boldsymbol{q}}^{1/s} R(t)^{m/s}(1 + o(1))$ and using $\frac{1}{s} - \frac{m}{s(s+m)} = \frac{1}{s+m}$,

$$Z(t) = C_{\boldsymbol{q}}^{1/(s+m)} \left(\frac{s+m}{s} \right)^{m/(s+m)} t^{m/(s+m)} (1 + o(1)).$$

Multiplying $R(t)$ and $Z(t)$, by cancellation, we have

$$\frac{R(t)Z(t)}{t} \longrightarrow \frac{s+m}{s}.$$

□

Using the above result, we now derive our boundary rates:

Corollary G.1 (Boundary rates)**.** *Under the assumptions of Proposition G.1, for every $j \in \mathcal{M}$,*

$$p_j(t) = \frac{s}{(s+m)\delta_j}\frac{1}{t}(1+o(1)). \tag{48}$$

If, in addition, $u_k(t) = e^{-\theta_k(t)} \to \ell_k \in (0,\infty)$ for some $k \in \mathcal{B}$, then there exists $c_k > 0$ such that

$$p_k(t) = c_k t^{-m/(s+m)}(1+o(1)). \tag{49}$$

Proof. For $j \in \mathcal{M}$, $p_j(t) = 1/(u_j(t)Z(t))$, and Proposition G.1 together with $R(t)Z(t)/t \to (s+m)/s$ gives

$$t\,p_j(t) = \frac{t}{R(t)Z(t)} \cdot \frac{R(t)}{u_j(t)} \longrightarrow \frac{s}{(s+m)\delta_j}.$$

Rearranging, we have (48).

If $u_k(t) \to \ell_k \in (0,\infty)$ for some $k \in \mathcal{B}$, then plugging in the expansion of $Z(t)$ into $p_k(t) = 1/(u_k(t)Z(t))$ gives

$$p_k(t) = \frac{1}{\ell_k}\, C_{\boldsymbol{q}}^{-1/(s+m)} \left(\frac{s+m}{s}\right)^{-m/(s+m)} \cdot t^{-m/(s+m)}(1+o(1)),$$

which is (49). □

We now state our second corollary, which gives relative rates of the logits.

Corollary G.2 (Refined logits and relative rates)**.** *Under the assumptions of Proposition G.1,*

$$\theta_i(t) = \frac{m}{s+m}\log t + \frac{1}{s+m}\log C_{\boldsymbol{q}} + \frac{m}{s+m}\log\frac{s+m}{s} + \log q_i + o(1), \qquad i \in \mathcal{S},$$

and

$$\theta_j(t) = -\frac{s}{s+m}\log t + \frac{1}{s+m}\log C_{\boldsymbol{q}} - \frac{s}{s+m}\log\frac{s+m}{s} - \log\delta_j + o(1), \qquad j \in \mathcal{M}.$$

In particular, for all $j, j' \in \mathcal{M}$,

$$\theta_j(t) - \theta_{j'}(t) \longrightarrow \log\frac{\delta_{j'}}{\delta_j}, \qquad \frac{p_j(t)}{p_{j'}(t)} \longrightarrow \frac{\delta_{j'}}{\delta_j}.$$

Proof. For $i \in \mathcal{S}$, since $p_i(t) = e^{\theta_i(t)}/Z(t)$ and $p_i(t) \to q_i > 0$, we have

$$\theta_i(t) = \log Z(t) + \log p_i(t) = \log Z(t) + \log q_i + o(1).$$

Plugging in expansion of $Z(t)$ gives the desired result. For $j \in \mathcal{M}$, $u_j(t)/R(t) \to \delta_j$ gives $\theta_j(t) = -\log u_j(t) = -\log\delta_j - \log R(t) + o(1)$, so the expansion of $R(t)$ yields the corollary statement. Subtracting two instances of the $\mathcal{M}$-logit expansion gives the first relative limit, and $p_j/p_{j'} = e^{\theta_j - \theta_{j'}}$ gives the second. □

Summed versus coordinatewise convergence. The conserved product Q depends on $\mathcal{B}$ only through $\prod_{k\in\mathcal{B}} u_k(t) = \exp\bigl(-\sum_{k\in\mathcal{B}}\theta_k(t)\bigr)$. Convergence of this sum therefore determines the clock R and the normalization Z, and with them the rates on $\mathcal{S}$ and $\mathcal{M}$. A rate for a single action $k \in \mathcal{B}$ is a stronger statement: $p_k = 1/(u_k Z)$ requires $u_k(t) \to \ell_k \in (0,\infty)$, which is the extra hypothesis in (49). Proposition G.2 below exhibits an example for which the sum is constant while no individual u_k converges.

G.2 Quasi-strict boundary limits

For quasi-strict limits, we have $\mathcal{B} = \varnothing$, so the condition (47) holds automatically with $\beta_{\mathcal{B}} = 0$. We state the logit expansions for quasi-strict cases which were omitted in the main text.

Corollary G.3 (Refined quasi-strict logits). *Suppose $\boldsymbol{p}(t) \to \boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$, where $\boldsymbol{q}$ is a quasi-strict boundary equilibrium. Let $\mathcal{S} := \operatorname{supp}(\boldsymbol{q})$, $s := |\mathcal{S}| < K$, and $\delta_j := (\boldsymbol{A}\boldsymbol{q})_j > 0$ for $j \notin \mathcal{S}$. Define*

$$C := e^{\sum_{n=1}^{K} \theta_n(0)} \left(\prod_{i \in \mathcal{S}} q_i \right)^{-1} \prod_{j \notin \mathcal{S}} \delta_j .$$

Then

$$\theta_i(t) = \frac{K-s}{K} \log t + \frac{1}{K} \log C + \frac{K-s}{K} \log \frac{K}{s} + \log q_i + o(1), \qquad i \in \mathcal{S},$$

and

$$\theta_j(t) = -\frac{s}{K} \log t + \frac{1}{K} \log C - \frac{s}{K} \log \frac{K}{s} - \log \delta_j + o(1), \qquad j \notin \mathcal{S}.$$

Consequently, for every $j, j' \notin \mathcal{S}$,

$$\theta_j(t) - \theta_{j'}(t) \longrightarrow \log \frac{\delta_{j'}}{\delta_j}, \qquad \frac{p_j(t)}{p_{j'}(t)} \longrightarrow \frac{\delta_{j'}}{\delta_j}.$$

Proof. Quasi-strictness gives $\mathcal{B} = \varnothing$, $\mathcal{M} = [K] \setminus \mathcal{S}$, and $m = K - s \geq 1$, so $\beta_{\mathcal{B}} = 0$ and $e^{-\beta_{\mathcal{B}}} = 1$. Hence

$$C_{\boldsymbol{q}} = \frac{\prod_{j \notin \mathcal{S}} \delta_j}{Q \prod_{i \in \mathcal{S}} q_i} = e^{\sum_n \theta_n(0)} \left(\prod_{i \in \mathcal{S}} q_i \right)^{-1} \prod_{j \notin \mathcal{S}} \delta_j = C,$$

so the claimed expansions are exactly those of Corollary G.2 with $s + m = K$. The corresponding probability rates $p_j(t) = \frac{s}{K\delta_j} \frac{1}{t}(1 + o(1))$ for $j \notin \mathcal{S}$ are (48) of Corollary G.1. □

Full-support limits. The preceding result concerns boundary equilibria. If $\mathcal{S} = [K]$, then $\boldsymbol{A}\boldsymbol{q} = \boldsymbol{0}$ by complementary slackness and there are no vanishing coordinates. The logarithmic boundary scaling considered here is absent; Theorem 3.1 instead gives convergence to a finite logit equilibrium, with eventual exponential convergence.

G.3 A pure limit with one zero-margin boundary action

We next consider a simple non-quasi-strict setting. Here the limiting equilibrium is pure, and there is one zero-margin boundary action. The rate calculation itself is already contained in Proposition G.1; the additional task is to establish convergence of that logit, which is what makes the hypothesis (47) and the coordinatewise statement (49) possible.

Theorem G.4 (Vertex with one zero-margin boundary action). *Suppose $K \geq 3$ and $\boldsymbol{p}(t) \to \boldsymbol{e}_1 \in \mathcal{SN}(\boldsymbol{A})$ with $|\mathcal{B}| = 1$. Up to relabeling, $\mathcal{B} = \{2\}$ and $\mathcal{M} = \{3, \dots, K\}$. Then $\theta_2(t)$ converges to a finite limit, and there exists $c_2 > 0$ such that*

$$p_j(t) = \frac{1}{(K-1)\delta_j} \frac{1}{t}(1 + o(1)), \qquad j \in \mathcal{M},$$

$$p_2(t) = c_2 t^{-(K-2)/(K-1)}(1 + o(1)).$$

Proof. Since $R(t) \to \infty$ and is strictly increasing, we use $r := R(t)$ as the time variable. With a slight abuse of notation, we write $u_n(r) := u_n(R^{-1}(r))$ and $\boldsymbol{p}(r) := \boldsymbol{p}(R^{-1}(r))$. Then $\dot{u}_n = F_n(\boldsymbol{p})/Z$ becomes

$$\frac{du_n}{dr} = F_n(\boldsymbol{p}).$$

For every $j \in \mathcal{M}$, $F_j(\boldsymbol{p}(r)) \to \delta_j$, so the same averaging argument as in the proof of Proposition G.1 gives

$$\frac{u_j(r)}{r} \longrightarrow \delta_j.$$

Since $\mathcal{M}$ is finite, there exists a common $r_0 > 0$ such that

$$\frac{\delta_j}{2} r \le u_j(r) \le 2\delta_j r, \qquad j \in \mathcal{M}, \quad r \ge r_0.$$

Let $\rho_{\mathcal{M}}(\boldsymbol{p}) := \sum_{j\in\mathcal{M}} p_j$ and $M := \|\boldsymbol{A}\|_{\max}$. The pure-limit assumption is crucial here. Since $2 \in \mathcal{B}$, we have $A_{21} = (\boldsymbol{A}\boldsymbol{e}_1)_2 = 0$, while skew-symmetry gives $A_{22} = 0$. Hence

$$(\boldsymbol{A}\boldsymbol{p})_2 = \sum_{j\in\mathcal{M}} A_{2j} p_j,$$

and therefore

$$|(\boldsymbol{A}\boldsymbol{p})_2| \le M \rho_{\mathcal{M}}(\boldsymbol{p}).$$

Since $g_n(\boldsymbol{p}) = p_n(\boldsymbol{A}\boldsymbol{p})_n$ and $|(\boldsymbol{A}\boldsymbol{p})_n| \le M$ on Δ_K, we have $|g_n(\boldsymbol{p})| \le M p_n$ and $|\boldsymbol{p}^\top g(\boldsymbol{p})| \le M$. Hence

$$\big|(\boldsymbol{A}\boldsymbol{p})_2 \big(g_2 - \boldsymbol{p}^\top g\big)\big| \le 2M^2 \rho_{\mathcal{M}}(\boldsymbol{p}).$$

Also, recall that

$$\boldsymbol{S}(\boldsymbol{p}) = \operatorname{diag}(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^\top, \qquad \big(\boldsymbol{S}(\boldsymbol{p})\boldsymbol{g}\big)_j = p_j\big(g_j - \boldsymbol{p}^\top \boldsymbol{g}\big).$$

Since $A_{21} = A_{22} = 0$,

$$(\boldsymbol{A}\,\boldsymbol{S}(\boldsymbol{p})\,g)_2 = \sum_{j\in\mathcal{M}} A_{2j} p_j \big(g_j - \boldsymbol{p}^\top g\big),$$

so $|(\boldsymbol{A}\,\boldsymbol{S}(\boldsymbol{p})\,g)_2| \le 2M^2 \rho_{\mathcal{M}}(\boldsymbol{p})$. Thus

$$|F_2(\boldsymbol{p})| \le C_A \rho_{\mathcal{M}}(\boldsymbol{p}), \qquad C_A := M + 2\gamma M^2.$$

The conserved product Q gives

$$u_1 = \frac{Q}{u_2 \prod_{j\in\mathcal{M}} u_j}.$$

Since $p_j/p_1 = u_1/u_j$ and $p_1 \le 1$,

$$p_j \le \frac{Q}{u_2 u_j \prod_{e\in\mathcal{M}} u_e}.$$

Using the lower bounds on $u_j(r)$ and summing over $j \in \mathcal{M}$,

$$\rho_{\mathcal{M}}(r) \le \frac{C_{\mathcal{M}}}{u_2(r) r^{m+1}},$$

where

$$C_{\mathcal{M}} := \frac{2^{m+1} Q}{\prod_{e\in\mathcal{M}} \delta_e} \sum_{j\in\mathcal{M}} \frac{1}{\delta_j}.$$

Combining the r-time dynamics of u_2 with the bounds on F_2 and $\rho_{\mathcal{M}}$,

$$\left|\frac{d}{dr} u_2(r)^2\right| = 2u_2(r)|F_2(\boldsymbol{p}(r))| \le \frac{2C_A C_{\mathcal{M}}}{r^{m+1}}.$$

Because $m = K - 2 \ge 1$, the right-hand side is integrable. Hence $u_2(r)^2$ converges to some $L \ge 0$, with

$$|u_2(r)^2 - L| \le \frac{2C_A C_{\mathcal{M}}}{m\, r^m}$$

for all sufficiently large r.

Now we show $L > 0$ through contradiction. Suppose $L = 0$. Then

$$u_2(r)^2 \le \frac{2C_A C_{\mathcal{M}}}{m\, r^m},$$

while the upper bound on $u_j(r)$ gives

$$\prod_{j\in\mathcal{M}} u_j(r) \le 2^m \left(\prod_{j\in\mathcal{M}} \delta_j\right) r^m.$$

Therefore,

$$u_2(r)^2 \prod_{j\in\mathcal{M}} u_j(r) \le \frac{2^{m+1} C_A C_{\mathcal{M}}}{m} \prod_{j\in\mathcal{M}} \delta_j.$$

But

$$\frac{p_2}{p_1} = \frac{u_1}{u_2} = \frac{Q}{u_2^2 \prod_{j\in\mathcal{M}} u_j},$$

so p_2/p_1 would remain bounded below by a positive constant for all sufficiently large r, which contradicts $p_2(t)/p_1(t) \to 0$. Hence

$$u_2(t) \longrightarrow \ell_2 := \sqrt{L} \in (0, \infty),$$

and taking logarithms gives $\theta_2(t) \to -\log \ell_2 =: \beta_{\mathcal{B}}$.

We may therefore apply Proposition G.1 with $s = 1$, $m = K - 2$, $q_1 = 1$, and $e^{-\beta_{\mathcal{B}}} = \ell_2$. Writing

$$C_* := \frac{\ell_2 \prod_{j\in\mathcal{M}} \delta_j}{Q},$$

we have $C_{\boldsymbol{q}} = C_*$, and therefore $R(t)Z(t)/t \to K - 1$, together with the $\mathcal{M}$-rate (48). For the coordinate in $\mathcal{B}$, the explicit prefactor in the proof of Corollary G.1 with $k = 2$ gives

$$p_2(t) = \frac{1}{\ell_2}(K-1)^{-(K-2)/(K-1)} C_*^{-1/(K-1)} t^{-(K-2)/(K-1)}(1 + o(1)),$$

which is the claimed rate for action 2 with

$$c_2 := \frac{Q^{1/(K-1)}}{(K-1)^{(K-2)/(K-1)} \ell_2^{K/(K-1)} \left(\prod_{j\in\mathcal{M}} \delta_j\right)^{1/(K-1)}} > 0.$$

Since $(K-2)/(K-1) < 1$,

$$\sum_{j\in\mathcal{M}} p_j(t) = \mathcal{O}(t^{-1}) = o\left(t^{-(K-2)/(K-1)}\right),$$

so

$$1 - p_1(t) = c_2 t^{-(K-2)/(K-1)}(1 + o(1)), \qquad \|\boldsymbol{p}(t) - \boldsymbol{e}_1\|_1 = 2(1 - p_1(t)).$$

The expansions

$$\theta_1(t) = \frac{K-2}{K-1} \log t + \frac{1}{K-1} \log C_* + \frac{K-2}{K-1} \log(K-1) + o(1)$$

and

$$\theta_j(t) = -\frac{1}{K-1} \log t + \frac{1}{K-1} \log C_* - \frac{1}{K-1} \log(K-1) - \log \delta_j + o(1), \qquad j \in \mathcal{M},$$

are those of Corollary G.2 with $s = 1$, $m = K - 2$, and $q_1 = 1$. □

The exponent $(K-2)/(K-1)$ is the specialization of the general $\mathcal{B}$-rate $m/(s+m)$ to $s = 1$ and $m = K - 2$. The pure-limit assumption is essential: $(\boldsymbol{A}\boldsymbol{e}_1)_2 = A_{21} = 0$ makes the drift of the zero-margin boundary coordinate controlled by the vanishing $\mathcal{M}$-mass, whereas for mixed-support limits zero margin may arise through cancellation across supported coordinates.

G.4 Multiple zero-margin boundary actions: nonuniversal exponents

The previous theorem shows that, for a pure limit with a single zero-margin boundary action, the corresponding logit stabilizes. More generally, however, the boundary-rate reduction only requires convergence of the sum $\sum_{k\in\mathcal{B}}\theta_k(t)$. When $|\mathcal{B}| > 1$, convergence of this sum need not imply convergence of the individual zero-margin boundary logits. The following construction shows that this distinction can lead to nonuniversal decay exponents.

Proposition G.2 (Explicit construction with multiple zero-margin boundary actions)**.** *Fix $\gamma > 0$ and choose $0 < c < a$ such that $\gamma a < 6\sqrt{3}$. Let*

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 & -a \\ 0 & 0 & c & 0 \\ 0 & -c & 0 & 0 \\ a & 0 & 0 & 0 \end{pmatrix}. \tag{50}$$

For every initialization of the form $\boldsymbol{\theta}(0) = (x_0, y_0, -y_0, -x_0)$, the corresponding trajectory satisfies $\boldsymbol{p}(t) \to \boldsymbol{e}_1$. For this limit,

$$\mathcal{S} = \{1\}, \qquad \mathcal{B} = \{2, 3\}, \qquad \mathcal{M} = \{4\}.$$

Let $\kappa := c/a \in (0, 1)$ and $b_0 := y_0 + \kappa x_0$. Then

$$p_2(t) = e^{b_0}(2at)^{-(1+\kappa)/2}(1 + o(1)),$$

$$p_3(t) = e^{-b_0}(2at)^{-(1-\kappa)/2}(1 + o(1)),$$

and

$$p_4(t) = (2at)^{-1}(1 + o(1)).$$

Moreover, the sum of the zero-margin boundary logits is exactly stabilized,

$$\sum_{k\in\mathcal{B}} \theta_k(t) = \theta_2(t) + \theta_3(t) \equiv 0,$$

while the individual zero-margin boundary logits satisfy

$$\theta_2(t) \longrightarrow -\infty, \qquad \theta_3(t) \longrightarrow +\infty.$$

Proof. The only nonzero interactions in (50) are between actions 1 and 4, and between actions 2 and 3. Hence the interaction graph has exactly two connected components, $\{1, 4\}$ and $\{2, 3\}$. By componentwise logit conservation (Lemma E.1),

$$\theta_1(t) + \theta_4(t) = \theta_1(0) + \theta_4(0) = 0, \qquad \theta_2(t) + \theta_3(t) = \theta_2(0) + \theta_3(0) = 0.$$

Therefore the trajectory remains on the invariant plane

$$\boldsymbol{\theta}(t) = (x(t), y(t), -y(t), -x(t)).$$

On this plane,

$$Z = 2\cosh x + 2\cosh y, \qquad p_1 p_4 = p_2 p_3 = Z^{-2}.$$

Set

$$f := \frac{a}{Z^2} = a p_1 p_4, \qquad h := \frac{c}{Z^2} = c p_2 p_3.$$

Then direct calculation gives

$$\boldsymbol{g} = (-f, h, -h, f).$$

To reduce (S2P-EG-Flow) to the (x, y)-coordinates, we evaluate $\mathrm{D}\boldsymbol{g}\,\boldsymbol{g}$. For the softmax parametrization,

$$\nabla_{\boldsymbol{\theta}} \log p_i = \boldsymbol{e}_i - \boldsymbol{p}.$$

Since $f = a p_1 p_4$ and $g_1 + g_4 = 0$,

$$\begin{aligned} \big(\nabla_{\boldsymbol{\theta}} \log f\big)^\top \boldsymbol{g} &= \big(\nabla_{\boldsymbol{\theta}} \log p_1 + \nabla_{\boldsymbol{\theta}} \log p_4\big)^\top \boldsymbol{g} \\ &= g_1 + g_4 - 2\boldsymbol{p}^\top \boldsymbol{g} = -2\boldsymbol{p}^\top \boldsymbol{g}. \end{aligned}$$

Hence

$$\big(\nabla_{\boldsymbol{\theta}} f\big)^{\top} \boldsymbol{g} = -2f\, \boldsymbol{p}^{\top} \boldsymbol{g}.$$

Similarly, since $h = cp_2p_3$ and $g_2 + g_3 = 0$,

$$\big(\nabla_{\boldsymbol{\theta}} h\big)^{\top} \boldsymbol{g} = -2h\, \boldsymbol{p}^{\top} \boldsymbol{g}.$$

Differentiating the components of $\boldsymbol{g} = (-f, h, -h, f)$ along the direction $\boldsymbol{g}$ therefore gives

$$(\mathrm{D}_{\boldsymbol{g}}\, \boldsymbol{g})_1 = 2f\, \boldsymbol{p}^{\top} \boldsymbol{g}, \qquad (\mathrm{D}_{\boldsymbol{g}}\, \boldsymbol{g})_4 = -2f\, \boldsymbol{p}^{\top} \boldsymbol{g},$$

and

$$(\mathrm{D}_{\boldsymbol{g}}\, \boldsymbol{g})_2 = -2h\, \boldsymbol{p}^{\top} \boldsymbol{g}, \qquad (\mathrm{D}_{\boldsymbol{g}}\, \boldsymbol{g})_3 = 2h\, \boldsymbol{p}^{\top} \boldsymbol{g}.$$

Substituting into (S2P-EG-Flow) therefore yields

$$\dot{x} = f\big(1 + \gamma \boldsymbol{p}^{\top} \boldsymbol{g}\big), \qquad \dot{y} = -h\big(1 + \gamma \boldsymbol{p}^{\top} \boldsymbol{g}\big).$$

We next show that the common factor $1 + \gamma \boldsymbol{p}^{\top} \boldsymbol{g}$ remains strictly positive. If $u, v \geq 0$ and $u + v = r$, then

$$uv|u - v| \leq \frac{r^3}{6\sqrt{3}}.$$

For $r > 0$, this follows by writing $u = rz$, $v = r(1 - z)$ and maximizing $z(1 - z)|2z - 1|$ over $z \in [0, 1]$. In the present construction,

$$\boldsymbol{p}^{\top} \boldsymbol{g} = ap_1p_4(p_4 - p_1) + cp_2p_3(p_2 - p_3).$$

Let $r := p_1 + p_4$, so that $p_2 + p_3 = 1 - r$. Using the preceding cubic bound, $0 < c < a$, and $r^3 + (1 - r)^3 \leq 1$, we obtain

$$\boldsymbol{p}^{\top} \boldsymbol{g} \geq -\frac{ar^3 + c(1 - r)^3}{6\sqrt{3}} \geq -\frac{a}{6\sqrt{3}}.$$

Therefore

$$1 + \gamma \boldsymbol{p}^{\top} \boldsymbol{g} \geq 1 - \frac{\gamma a}{6\sqrt{3}} > 0. \tag{51}$$

Since $f > 0$ and (51), we have $\dot{x} > 0$. Dividing the two coordinate equations therefore gives

$$\frac{dy}{dx} = -\frac{h}{f} = -\frac{c}{a} = -\kappa.$$

Hence

$$y(t) = b_0 - \kappa x(t), \qquad b_0 := y_0 + \kappa x_0. \tag{52}$$

We claim that $x(t) \to \infty$. Suppose that x is bounded above. Since x is increasing and $y = b_0 - \kappa x$, both x and y would remain bounded, and therefore $Z = 2\cosh x + 2\cosh y$ would be bounded above. Thus $f = a/Z^2$ would be bounded away from zero. Together with (51), this would give a positive lower bound on $\dot{x}$, contradicting the boundedness of x. Hence $x(t) \to \infty$.

Since $y = b_0 - \kappa x$ and $0 < \kappa < 1$,

$$\frac{p_2}{p_1} = e^{y - x} = e^{b_0 - (1+\kappa)x} \longrightarrow 0,$$

$$\frac{p_3}{p_1} = e^{-y - x} = e^{-b_0 - (1-\kappa)x} \longrightarrow 0, \qquad \frac{p_4}{p_1} = e^{-2x} \longrightarrow 0.$$

Thus $\boldsymbol{p}(t) \to \boldsymbol{e}_1$. Moreover,

$$\boldsymbol{A}\boldsymbol{e}_1 = (0, 0, 0, a)^{\top},$$

so

$$\mathcal{S} = \{1\}, \qquad \mathcal{B} = \{2, 3\}, \qquad \mathcal{M} = \{4\},$$

with $\delta_4 = a$.

We next derive the sharp rates. From (52) and $0 < \kappa < 1$,

$$Z(t) = e^{x(t)} \left(1 + e^{-2x(t)} + e^{b_0 - (1+\kappa)x(t)} + e^{-b_0 - (1-\kappa)x(t)}\right) = e^{x(t)}(1 + o(1)).$$

Since $\boldsymbol{p}(t) \to \boldsymbol{e}_1$ and $\boldsymbol{g}(\boldsymbol{e}_1) = \mathbf{0}$, continuity gives $\boldsymbol{p}(t)^\top \boldsymbol{g}(\boldsymbol{p}(t)) \to 0$. Therefore

$$\dot{x}(t) = ae^{-2x(t)}(1 + o(1)).$$

Consequently,

$$\frac{d}{dt}e^{2x(t)} = 2e^{2x(t)}\dot{x}(t) \longrightarrow 2a.$$

Integrating this asymptotic relation gives

$$\frac{e^{2x(t)}}{t} \longrightarrow 2a,$$

equivalently,

$$e^{2x(t)} = 2at(1 + o(1)). \tag{53}$$

Using $y = b_0 - \kappa x$ and $Z = e^x(1 + o(1))$,

$$p_2 = e^{b_0}e^{-(1+\kappa)x}(1 + o(1)), \qquad p_3 = e^{-b_0}e^{-(1-\kappa)x}(1 + o(1)).$$

Together with (53), this gives

$$p_2(t) = e^{b_0}(2at)^{-(1+\kappa)/2}(1 + o(1)),$$
$$p_3(t) = e^{-b_0}(2at)^{-(1-\kappa)/2}(1 + o(1)).$$

Likewise,

$$p_4(t) = e^{-2x(t)}(1 + o(1)) = (2at)^{-1}(1 + o(1)).$$

Finally, the zero-margin boundary logits satisfy

$$\theta_2(t) = y(t) = b_0 - \kappa x(t) \longrightarrow -\infty, \qquad \theta_3(t) = -y(t) \longrightarrow +\infty,$$

while componentwise logit conservation gives

$$\sum_{k\in\mathcal{B}} \theta_k(t) = \theta_2(t) + \theta_3(t) \equiv 0.$$

Thus the sum of the zero-margin boundary logits is exactly stabilized even though neither zero-margin boundary logit converges individually. In particular, the summed-logit condition on $\mathcal{B}$ governing the $\mathcal{M}$-rate remains valid, whereas the coordinatewise condition needed for individual rates of actions in $\mathcal{B}$ fails.

Moreover,

$$\frac{1-\kappa}{2} < \frac{1+\kappa}{2}, \qquad \frac{1-\kappa}{2} < 1,$$

so p_3 is the slowest vanishing coordinate. Hence

$$1 - p_1(t) = p_2(t) + p_3(t) + p_4(t) = p_3(t)(1 + o(1)).$$

□

Remark 2 (Consistency with Proposition G.1)**.** *Here $s = m = 1$, while*

$$\prod_{k\in\mathcal{B}} e^{-\theta_k(t)} = e^{-(\theta_2(t)+\theta_3(t))} \equiv 1$$

and

$$Q = e^{-(x_0+y_0-y_0-x_0)} = 1.$$

Since $q_1 = 1$ and $\delta_4 = a$, the constant in Proposition G.1 reduces to $C_{\boldsymbol{q}} = a$. That proposition therefore predicts

$$Z(t) = (2at)^{1/2}(1 + o(1)),$$

which agrees with $e^{2x(t)} = 2at(1 + o(1))$ and $Z = e^x(1 + o(1))$. Likewise, Corollary G.1 gives

$$p_4(t) = (2at)^{-1}(1 + o(1)).$$

The coordinatewise conclusion (49) does not apply: indeed,

$$u_2(t) = e^{-\theta_2(t)} \longrightarrow \infty, \qquad u_3(t) = e^{-\theta_3(t)} \longrightarrow 0.$$

Deduction of Proposition 3.1. Given $\varepsilon \in (0, 1/2)$, set $\kappa := 1 - 2\varepsilon \in (0, 1)$. Choose $a > 0$ such that $\gamma a < 6\sqrt{3}$ and set $c := \kappa a$, so that (50) is admissible. With $\mathcal{B} = \{k_1, k_2\} := \{2, 3\}$,

$$\frac{1+\kappa}{2} = 1 - \varepsilon, \qquad \frac{1-\kappa}{2} = \varepsilon.$$

Therefore

$$p_{k_1}(t) = \Theta\big(t^{-(1-\varepsilon)}\big), \qquad p_{k_2}(t) = \Theta\big(t^{-\varepsilon}\big),$$

while

$$p_4(t) = \Theta(t^{-1}).$$

The support counts $(|\mathcal{S}|, |\mathcal{B}|, |\mathcal{M}|) = (1, 2, 1)$ remain fixed for every $\varepsilon \in (0, 1/2)$, while taking ε close to 0 makes the slower exponent of an action in $\mathcal{B}$ arbitrarily small. This proves Proposition 3.1.

H "DIRECT" PROOFS OF TWO IMPLICIT BIAS RESULTS

H.1 PROOF OF THEOREM 4.1: EVERY POLICY LIMIT IS A SNE

Let $\mathcal{S} \coloneqq \operatorname{supp}(\boldsymbol{q})$ and write

$$\mathcal{V}_\gamma(\boldsymbol{p}) \coloneqq -\boldsymbol{g}(\boldsymbol{p}) + \frac{\gamma}{2}\mathrm{D}\boldsymbol{g}(\boldsymbol{p})\boldsymbol{g}(\boldsymbol{p}).$$

Due to continuity, $\dot{\boldsymbol{\theta}}(t) \to \mathcal{V}_\gamma(\boldsymbol{q})$. Also recall that

$$D\boldsymbol{g}(\boldsymbol{\theta}) = \operatorname{diag}(\boldsymbol{A}\boldsymbol{p})\boldsymbol{S}(\boldsymbol{p}) + \operatorname{diag}(\boldsymbol{p})\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p}). \tag{54}$$

Step 1: $(\boldsymbol{A}\boldsymbol{q})_i = 0$ for $i \in \mathcal{S}$. We first prove that $\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0}$. We begin by showing that $\mathcal{V}_\gamma(\boldsymbol{q}) = \boldsymbol{0}$.

If $j \notin \mathcal{S}$, then $g_j(\boldsymbol{q}) = 0$. In addition, the j-th rows of both $\boldsymbol{S}(\boldsymbol{q})$ and $\operatorname{diag}(\boldsymbol{q})$ vanish; hence, the jth row of $\mathrm{D}\boldsymbol{g}(\boldsymbol{q})$ is zero and $\mathcal{V}_{\gamma,j}(\boldsymbol{q}) = 0$.

Now consider $i \in \mathcal{S}$. If there is only one such i, then by the logit conservation (Lemma E.1) and $\mathcal{V}_{\gamma,j}(\boldsymbol{q}) = 0$ for all $j \notin \mathcal{S}$, it immediately follows that $\mathcal{V}_{\gamma,i}(\boldsymbol{q}) = 0$. We therefore assume that $|\mathcal{S}| \geq 2$ from now on. For $i, k \in \mathcal{S}$,

$$\theta_i(t) - \theta_k(t) = \log \frac{p_i(t)}{p_k(t)} \longrightarrow \log \frac{q_i}{q_k}.$$

The derivative of the left-hand side converges to $\mathcal{V}_{\gamma,i}(\boldsymbol{q}) - \mathcal{V}_{\gamma,k}(\boldsymbol{q})$. A convergent differentiable function whose derivative has a limit must have derivative converging to zero. Hence

$$\mathcal{V}_{\gamma,i}(\boldsymbol{q}) = \mathcal{V}_{\gamma,k}(\boldsymbol{q}), \qquad i, k \in \mathcal{S}.$$

Thus, all coordinates of $\mathcal{V}_\gamma(\boldsymbol{q})$ on $\mathcal{S}$ have a common value, say c.

By Lemma E.1, $\sum_{r=1}^K \theta_r(t)$ is constant, and therefore $\sum_{r=1}^K \dot{\theta}_r(t) = 0$, $t \geq 0$.

Passing to the limit as $t \to \infty$ gives

$$0 = \sum_{r=1}^K \mathcal{V}_{\gamma,r}(\boldsymbol{q}) = \sum_{i \in \mathcal{S}} \mathcal{V}_{\gamma,i}(\boldsymbol{q}) = |\mathcal{S}|\, c,$$

where we used $\mathcal{V}_{\gamma,j}(\boldsymbol{q}) = 0$ for every $j \notin \mathcal{S}$. Hence $c = 0$, and therefore

$$\mathcal{V}_\gamma(\boldsymbol{q}) = \boldsymbol{0}.$$

We now use this stationarity of the modified flow to show that $\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0}$. By the definition of $\mathcal{V}_\gamma$,

$$\boldsymbol{0} = \mathcal{V}_\gamma(\boldsymbol{q}) = -\left(\boldsymbol{I} - \frac{\gamma}{2}\mathrm{D}\boldsymbol{g}(\boldsymbol{q})\right)\boldsymbol{g}(\boldsymbol{q}).$$

By Lemma C.1 and the assumed step-size condition,

$$\left\|\frac{\gamma}{2}\mathrm{D}\boldsymbol{g}(\boldsymbol{q})\right\|_2 \leq \frac{\gamma}{2}\|\boldsymbol{A}\|_2 < 1.$$

Hence $\boldsymbol{I} - \frac{\gamma}{2}\mathrm{D}\boldsymbol{g}(\boldsymbol{q})$ is invertible, and therefore $\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0}$. Since $q_i > 0$ for every $i \in \mathcal{S}$ and $g_i(\boldsymbol{q}) = q_i(\boldsymbol{A}\boldsymbol{q})_i$, it follows that

$$(\boldsymbol{A}\boldsymbol{q})_i = 0, \qquad i \in \mathcal{S}.$$

Step 2: $(\boldsymbol{A}\boldsymbol{q})_j \geq 0$ for $j \notin \mathcal{S}$. We argue by contradiction using the inverse-weight dynamics of Lemma E.2. Suppose that $(\boldsymbol{A}\boldsymbol{q})_j < 0$ for some $j \notin \mathcal{S}$. Since $\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0}$, Lemma E.2 gives

$$F_j(\boldsymbol{p}(t)) \longrightarrow F_j(\boldsymbol{q}) = (\boldsymbol{A}\boldsymbol{q})_j < 0.$$

Hence there exist $T, \eta > 0$ such that

$$F_j(\boldsymbol{p}(t)) \leq -\eta, \qquad t \geq T.$$

Using $\dot{u}_j = F_j(\boldsymbol{p})/Z$ and $R(t) = \int_0^t Z(s)^{-1}\,ds$, we obtain

$$\begin{aligned} u_j(t) - u_j(T) &= \int_T^t \frac{F_j(\boldsymbol{p}(s))}{Z(s)}\,ds \\ &= \int_T^t F_j(\boldsymbol{p}(s))\,dR(s) \\ &\le -\eta\big(R(t) - R(T)\big). \end{aligned}$$

Therefore,

$$u_j(t) \le u_j(T) - \eta\big(R(t) - R(T)\big).$$

By Lemma E.3, $R(t) \to \infty$, so the right-hand side eventually becomes negative, contradicting $u_j(t) = e^{-\theta_j(t)} > 0$. Thus

$$(\boldsymbol{A}\boldsymbol{q})_j \ge 0, \qquad j \notin \mathcal{S}.$$

Together with $(\boldsymbol{A}\boldsymbol{q})_i = 0$ for $i \in \mathcal{S}$, this gives $\boldsymbol{A}\boldsymbol{q} \ge \boldsymbol{0}$, and hence $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$. □

H.2 Proof of Theorem 4.4: Rank-One Cross-Interactions

Recall that $\boldsymbol{B} = \boldsymbol{a}\boldsymbol{b}^\top$, where $a_j, b_i > 0$ for every $j \in \mathcal{L}$ and $i \in \mathcal{I}$. Write $Z(t) \coloneqq \sum_{r=1}^K e^{\theta_r(t)}$ and define

$$h(t) \coloneqq \boldsymbol{a}^\top \boldsymbol{p}_{\mathcal{L}}(t), \qquad r_i(t) \coloneqq \frac{e^{-\theta_i(t)}}{b_i} \quad (i \in \mathcal{I}). \tag{55}$$

Since $r_i(0) = \exp(-\theta_i(0) - \log b_i)$, the *initialization-dependent support set* can be written as

$$\mathcal{M}(\boldsymbol{\theta}(0)) = \arg\min_{i \in \mathcal{I}} r_i(0). \tag{56}$$

Step 1: The ordering of $r_i(t)$ is preserved. We first derive an equation for $r_i - r_k$, $i, k \in \mathcal{I}$. By flat dominance and skew-symmetry, $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = 0$ and $\boldsymbol{A}_{\mathcal{I},\mathcal{L}} = -\boldsymbol{b}\boldsymbol{a}^\top$, so, for every $i \in \mathcal{I}$,

$$(\boldsymbol{A}\boldsymbol{p})_i = -b_i h, \qquad g_i = -b_i p_i h.$$

Let $\boldsymbol{y} \coloneqq \boldsymbol{S}(\boldsymbol{p})\boldsymbol{g}$. Then $(\boldsymbol{A}\boldsymbol{y})_i = -b_i \boldsymbol{a}^\top \boldsymbol{y}_{\mathcal{L}}$, and Lemma E.2 gives

$$\dot{r}_i = \frac{F_i(\boldsymbol{p})}{b_i Z} = -\frac{h}{Z} + \frac{\gamma}{2Z}\big[h(g_i - \boldsymbol{p}^\top \boldsymbol{g}) + \boldsymbol{a}^\top \boldsymbol{y}_{\mathcal{L}}\big].$$

Subtracting the equations for $i, k \in \mathcal{I}$, we obtain

$$\begin{aligned} \frac{d}{dt}(r_i - r_k) &= \frac{\gamma h}{2Z}(g_i - g_k) && (57) \\ &= \frac{\gamma h^2}{2Z}(b_k p_k - b_i p_i) && (58) \\ &= \frac{\gamma h^2}{2Z^2 r_i r_k}(r_i - r_k) && (b_i p_i = (Z r_i)^{-1}) \\ &= \frac{\gamma}{2} b_i b_k p_i p_k h^2 (r_i - r_k), && (59) \end{aligned}$$

Thus,

$$r_i(t) - r_k(t) = \big(r_i(0) - r_k(0)\big) \exp\left(\frac{\gamma}{2}\int_0^t b_i b_k p_i(s) p_k(s) h(s)^2\,ds\right). \tag{60}$$

In particular, $r_i(0) = r_k(0)$ implies $r_i(t) = r_k(t)$ for every $t \ge 0$. If $r_i(0) < r_k(0)$, then the nonnegativity of the integrand gives

$$r_k(t) - r_i(t) \ge r_k(0) - r_i(0) > 0, \qquad t \ge 0.$$

Step 2: Initialization-dependent implicit bias. We first show that $Z(t) \to \infty$. By logit conservation (Lemma E.1),

$$\prod_{r=1}^{K} p_r(t) = \frac{\exp(\sum_{r=1}^{K} \theta_r(t))}{Z(t)^K} = \frac{\exp(\sum_{r=1}^{K} \theta_r(0))}{Z(t)^K}.$$

By Theorem 3.2, $p_j(t) \to 0$ for every $j \in \mathcal{L}$. Since $\mathcal{L} \neq \varnothing$ and all probabilities are bounded by one, the product on the left tends to zero. Hence, it must be that $Z(t) \to \infty$.

Now fix $k \in \mathcal{I} \setminus \mathcal{M}(\boldsymbol{\theta}(0))$ and choose $i \in \mathcal{M}(\boldsymbol{\theta}(0))$. By **Step 1** and $r_i(t) > 0$,

$$r_k(t) \geq r_k(t) - r_i(t) \geq r_k(0) - r_i(0) > 0.$$

It follows that

$$0 < p_k(t) = \frac{1}{b_k Z(t) r_k(t)} \leq \frac{1}{b_k Z(t)(r_k(0) - r_i(0))} \longrightarrow 0.$$

Together with $p_j(t) \to 0$ for $j \in \mathcal{L}$, this gives $\sum_{i \in \mathcal{M}(\boldsymbol{\theta}(0))} p_i(t) \to 1$.

For $i, k \in \mathcal{M}(\boldsymbol{\theta}(0))$, we have $r_i(0) = r_k(0)$, so **Step 1** gives $r_i(t) = r_k(t)$, or equivalently $b_i p_i(t) = b_k p_k(t)$, for every $t \geq 0$. Therefore, for each $i \in \mathcal{M}(\boldsymbol{\theta}(0))$,

$$p_i(t) = \frac{b_i^{-1}}{\sum_{k \in \mathcal{M}(\boldsymbol{\theta}(0))} b_k^{-1}} \sum_{k \in \mathcal{M}(\boldsymbol{\theta}(0))} p_k(t) \longrightarrow \frac{b_i^{-1}}{\sum_{k \in \mathcal{M}(\boldsymbol{\theta}(0))} b_k^{-1}} =: q_i.$$

In particular, $\boldsymbol{p}(t) \to \boldsymbol{e}_i$ if and only if $S = \{i\}$, so the basin of attraction of $\boldsymbol{e}_i$ is exactly

$$\left\{\boldsymbol{\theta}(0) \in \mathbb{R}^K : \theta_i(0) + \log b_i > \theta_k(0) + \log b_k, \quad \forall k \in \mathcal{I} \setminus \{i\}\right\}.$$

If the limit is mixed, then $|\mathcal{M}(\boldsymbol{\theta}(0))| \geq 2$, and hence $\theta_i(0) + \log b_i = \theta_k(0) + \log b_k$ for some distinct $i, k \in \mathcal{I}$. Such initializations lie in a finite union of affine hyperplanes and therefore form a Lebesgue-null set. □

I Implicit Bias under Flat Dominance

We first state the boundary estimates and balanced-equilibrium properties used in the proofs of Theorems 4.2 and 4.3. We then prove these two results in Appendices I.2 and I.3 and derive rates for the selected limits in Appendix I.4. The technical lemmas are proved in Appendix I.5.

Notation. As in Section 4.2, we assume Assumption 1, and denote $m := |\mathcal{I}| \geq 2$, $l := |\mathcal{L}| \geq 1$, and

$$\boldsymbol{B} := \boldsymbol{A}_{\mathcal{L},\mathcal{I}} \in \mathbb{R}^{l\times m}, \qquad b_- \leq B_{ji} \leq b_+ \quad (j \in \mathcal{L},\ i \in \mathcal{I}).$$

We also impose $0 < \gamma < 2/(3\|\boldsymbol{A}\|_{\max})$. By Theorem 3.2 and Lemma E.4,

$$\rho(t) := \sum_{j\in\mathcal{L}} p_j(t) \longrightarrow 0, \qquad \tau(t) := \int_0^t \rho(s)\, ds \longrightarrow \infty.$$

Recall the blow-up coordinates from Appendix D:

$$x_i := \frac{p_i}{1-\rho} \quad (i \in \mathcal{I}), \qquad \lambda_j := \frac{p_j}{\rho} \quad (j \in \mathcal{L}), \qquad \boldsymbol{z} := (\boldsymbol{x}, \boldsymbol{\lambda}).$$

For notational convenience, write

$$\begin{aligned} u_i(\boldsymbol{x},\boldsymbol{\lambda}) &:= x_i(\boldsymbol{B}^\top\boldsymbol{\lambda})_i, & U(\boldsymbol{x},\boldsymbol{\lambda}) &:= \langle \boldsymbol{x}, \boldsymbol{u}\rangle = \sum_{i\in\mathcal{I}} x_i^2(\boldsymbol{B}^\top\boldsymbol{\lambda})_i, \\ w_j(\boldsymbol{x},\boldsymbol{\lambda}) &:= -\lambda_j(\boldsymbol{B}\boldsymbol{x})_j, & V(\boldsymbol{x},\boldsymbol{\lambda}) &:= -\langle \boldsymbol{\lambda}, \boldsymbol{w}\rangle = \sum_{j\in\mathcal{L}} \lambda_j^2(\boldsymbol{B}\boldsymbol{x})_j, \\ \kappa(\boldsymbol{x},\boldsymbol{\lambda}) &:= U(\boldsymbol{x},\boldsymbol{\lambda}) + V(\boldsymbol{x},\boldsymbol{\lambda}). \end{aligned} \tag{61}$$

Thus the boundary system $\mathcal{B}$ in Proposition D.1 is

$$\frac{dx_i}{d\tau} = x_i(u_i - U), \qquad \frac{d\lambda_j}{d\tau} = \lambda_j(w_j + V). \tag{62}$$

Throughout, prime denotes differentiation with respect to boundary time τ.

I.1 Two Key Technical Lemmas

The first lemma bounds the error in the logarithmic derivatives uniformly, even when x_i or λ_j is small. It also shows that each λ_j remains bounded away from zero, although the total probability assigned to dominated actions vanishes.

Lemma I.1 (Boundary Estimates). *There are polynomials $r_{i,\gamma}$ and $s_{j,\gamma}$ such that every full-flow trajectory satisfies*

$$\frac{d}{d\tau}\log x_i = u_i - U + \rho r_{i,\gamma}(\rho, \boldsymbol{z}), \qquad \frac{d}{d\tau}\log \lambda_j = w_j + V + \rho s_{j,\gamma}(\rho, \boldsymbol{z}). \tag{63}$$

The polynomial right-hand sides extend to $[0,1]\times\Delta_{\mathcal{I}}\times\Delta_{\mathcal{L}}$, where the remainders are uniformly bounded and

$$b_-\left(\frac{1}{m} + \frac{1}{l}\right) \leq \kappa \leq 2b_+. \tag{64}$$

Moreover, with $\eta := b_-/(4lb_+)$, every full-flow trajectory satisfies $\min_{j\in\mathcal{L}} \lambda_j(\tau) \geq \eta$ for all sufficiently large τ.

For a nonempty $S \subseteq \mathcal{I}$, we call a pair $(\boldsymbol{q},\boldsymbol{\lambda}) \in \mathrm{relint}(\Delta_S)\times\Delta_{\mathcal{L}}^\circ$ ***balanced*** if $q_i(\boldsymbol{B}^\top\boldsymbol{\lambda})_i$ is the same for all $i \in S$ and $\lambda_j(\boldsymbol{B}\boldsymbol{q})_j$ is the same for all $j \in \mathcal{L}$. The equality is within each group; the two common values may differ. By (62), *this is equivalent to being an equilibrium of $\mathcal{B}$ on that face.*

The following lemma collects the three properties of these equilibria that we need: there is one for each support, every convergent policy approaches one in blow-up coordinates, and its stability is determined by the support size. To state stability at boundary points, we use the polynomial extension to the ambient affine space

$$\mathcal{A} := \{(\boldsymbol{x},\boldsymbol{\lambda}) \in \mathbb{R}^m \times \mathbb{R}^l : \mathbf{1}^\top\boldsymbol{x} = \mathbf{1}^\top\boldsymbol{\lambda} = 1\}.$$

Lemma I.2 (Balanced Equilibria). *The boundary field $\mathcal{B}$ has the following properties.*

(i) ***Existence and uniqueness.*** *For every nonempty $S \subseteq \mathcal{I}$, with $s := |S|$, there is a unique balanced equilibrium* $\boxed{\boldsymbol{z}^S := (\boldsymbol{q}^S, \boldsymbol{\lambda}^S) \in \operatorname{relint}(\Delta_S) \times \Delta_{\mathcal{L}}^\circ}$ *with*

$$q_i^S (\boldsymbol{B}^\top \boldsymbol{\lambda}^S)_i = \underbrace{\frac{l}{s}\beta_S}_{=:\alpha_S} \quad (i \in S), \qquad \lambda_j^S (\boldsymbol{B}\boldsymbol{q}^S)_j = \underbrace{\left(\sum_{j \in \mathcal{L}} \frac{1}{(\boldsymbol{B}\boldsymbol{q}^S)_j} \right)^{-1}}_{=:\beta_S} \quad (j \in \mathcal{L}). \tag{65}$$

Here $\boldsymbol{q}^S$ has zero coordinates outside S.

(ii) ***Policy limits.*** *If $\boldsymbol{p}(t) \to \boldsymbol{q}$ and $S := \operatorname{supp}(\boldsymbol{q})$, then $S \subseteq \mathcal{I}$ and*

$$(\boldsymbol{x}(t), \boldsymbol{\lambda}(t)) \longrightarrow (\boldsymbol{q}^S, \boldsymbol{\lambda}^S), \qquad \boldsymbol{q} = (\boldsymbol{q}^S, \boldsymbol{0}).$$

(iii) ***Stability.*** *In the polynomial affine extension of $\mathcal{B}$ on $\mathcal{A}$, the linearization at $\boldsymbol{z}^S$ has $s-1$ eigenvalues with positive real part and $m - s + l - 1$ with negative real part, counted with multiplicity. Thus $\boldsymbol{z}^S$ is hyperbolic, and it is a sink exactly when $s = 1$.*

In particular, part (ii) shows thatalthough the Nash face is a continuum, there is at most one policy limit for each support set $S \subseteq \mathcal{I}$. Part (iii) then distinguishes these limits: a mixed support gives at least one unstable direction, whereas a singleton gives only stable directions.

I.2 Proof of Theorem 4.2: Conditional Vertex Selection

Policy convergence alone does not identify the limit of the blow-up state: all points $(0, \boldsymbol{q}, \boldsymbol{\lambda})$ (in the blow-up coordinates) with the same $\boldsymbol{q}$ map to the same policy $(\boldsymbol{q}, \boldsymbol{0})$, as illustrated in Figure 5 below:

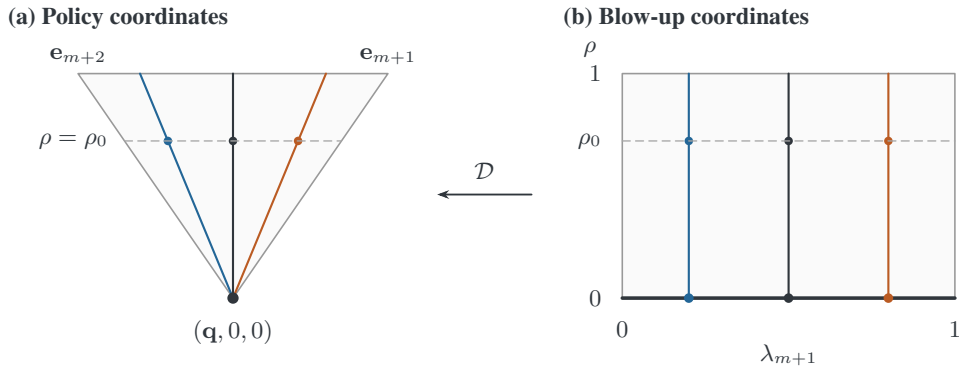


Figure 5: **Blow-up Coordinates.** At fixed $\mathbf{x} = \mathbf{q}$ with $l = 2$, the entire edge $\rho = 0$ maps to the single policy $(\mathbf{q}, 0, 0)$. The colored lines are coordinate rays, and $\mathcal{D}(\rho, \boldsymbol{x}, \boldsymbol{\lambda}) := ((1-\rho)\boldsymbol{x}, \rho\boldsymbol{\lambda})$.

Step 1: Convergence to a hyperbolic equilibrium. By Lemma I.2(ii), however, $\boldsymbol{\lambda}$ also converges. We can therefore apply the Stable Manifold Theorem (Theorem B.4) at a specific equilibrium of the full system. To do so, we retain ρ as a state variable. By Proposition D.1, the exact field is

$$\mathcal{F}_\gamma(\rho, \boldsymbol{z}) := \begin{pmatrix} -\rho(1-\rho)\big(\kappa(\boldsymbol{z}) + \rho e_\gamma(\rho, \boldsymbol{z})\big) \\ \mathcal{B}(\boldsymbol{z}) + \rho E_\gamma(\rho, \boldsymbol{z}) \end{pmatrix}, \qquad \frac{d}{d\tau}(\rho, \boldsymbol{z}) = \mathcal{F}_\gamma(\rho, \boldsymbol{z}). \tag{66}$$

Since $\mathcal{F}_\gamma$ is polynomial in affine simplex coordinates, it extends smoothly to an open neighborhood of $[0,1] \times \Delta_{\mathcal{I}} \times \Delta_{\mathcal{L}}$ in the $(K-1)$-dimensional affine space $\mathbb{R} \times \mathcal{A}$. For a nonempty $S \subseteq \mathcal{I}$, set $\boldsymbol{\xi}_S := (0, \boldsymbol{z}^S)$ and $s := |S|$. This is an equilibrium of $\mathcal{F}_\gamma$, with linearization

$$D\mathcal{F}_\gamma(\boldsymbol{\xi}_S) = \begin{pmatrix} -\kappa(\boldsymbol{z}^S) & 0 \\ E_\gamma(0, \boldsymbol{z}^S) & D\mathcal{B}(\boldsymbol{z}^S) \end{pmatrix}. \tag{67}$$

The matrix is block triangular, and thus, its eigenvalues are those of $D\mathcal{B}(\boldsymbol{z}^S)$ together with $-\kappa(\boldsymbol{z}^S) = -(\alpha_S + \beta_S) < 0$, where the equality follows from Lemma I.2(i). Then, Lemma I.2(iii) implies that $\boldsymbol{\xi}_S$ is *hyperbolic*, with $s-1$ unstable and $K-s$ stable directions.

Let $\mathcal{P} := (0,1) \times \Delta^\circ_{\mathcal{I}} \times \Delta^\circ_{\mathcal{L}}$ be the $(K-1)$-dimensional set of physical blow-up states. By Lemma I.2(ii) and the blow-down formula $\boldsymbol{p} = ((1-\rho)\boldsymbol{x}, \rho\boldsymbol{\lambda})$,

$$\boldsymbol{p}(t) \to (\boldsymbol{q}^S, \boldsymbol{0}) \quad \Longleftrightarrow \quad (\rho(\tau), \boldsymbol{z}(\tau)) \to \boldsymbol{\xi}_S. \tag{68}$$

The two limits refer to the same forward trajectory because $\tau(t) \to \infty$ as $t \to \infty$ and vice-versa.

If $s \geq 2$, then $\boldsymbol{\xi}_S$ has at least one unstable direction, so Lemma B.1 implies that

$$\mathcal{W}^s_{\mathcal{F}_\gamma}(\boldsymbol{\xi}_S) \cap \mathcal{P} \quad \text{is Lebesgue-null,}$$

where we recall that $\mathcal{W}^s_{\mathcal{F}_\gamma}(\boldsymbol{\xi}_S)$ is the stable set of $\boldsymbol{\xi}_S$ under $\mathcal{F}_\gamma$ (Definition B.7). Taking the union over the finitely many supports S's with $s \geq 2$ shows that the basin of all mixed policy limits is null in $\mathcal{P}$. We denote this basin by $\widetilde{\mathcal{N}}_{\text{mix}}$.

Step 2: Charcterizing basins in policy space. If $s = 1$, i.e., $S = \{i\}$ for some $i \in \mathcal{I}$, then $\boldsymbol{q}^S = \boldsymbol{e}_i$, and all $K-1$ directions are stable. By Theorem B.4, the local stable manifold at $\boldsymbol{\xi}_{\{i\}}$ therefore has the same dimension as the ambient affine space $\mathbb{R} \times \mathcal{A}$. Hence the local stable manifold contains an open neighborhood $\mathcal{U}_i$ of $\boldsymbol{\xi}_{\{i\}}$ in the ambient affine space, by the inverse function theorem. Since $\boldsymbol{\xi}_{\{i\}} \in \overline{\mathcal{P}}$, the intersection $\mathcal{U}_i \cap \mathcal{P}$ is nonempty and is contained in the basin of $\boldsymbol{e}_i$, say $\widetilde{\mathcal{O}}_i$.

We now prove that $\widetilde{\mathcal{O}}_i$ is open in $\mathcal{P}$. To do so, let us fix any initial state in $\widetilde{\mathcal{O}}_i$. By (68), its trajectory enters $\mathcal{U}_i$ at some finite boundary time T. Continuous dependence on initial conditions gives an open neighborhood of this initial state in $\mathcal{P}$ whose trajectories all lie in $\mathcal{U}_i$ at time T. Every such trajectory converges to $\boldsymbol{\xi}_{\{i\}}$, so the corresponding policy converges to $\boldsymbol{e}_i$. Thus this neighborhood is contained in the basin, proving openness.

Step 3: Transferring to the logit space. Define $\Psi : \mathcal{P} \times \mathbb{R} \to \mathbb{R}^K$ by

$$\Psi(\rho, \boldsymbol{x}, \boldsymbol{\lambda}, c) := \log \mathcal{D}(\rho, \boldsymbol{x}, \boldsymbol{\lambda}) + c\boldsymbol{1}, \qquad \mathcal{D}(\rho, \boldsymbol{x}, \boldsymbol{\lambda}) := ((1-\rho)\boldsymbol{x}, \rho\boldsymbol{\lambda}). \tag{69}$$

Ψ is a smooth diffeomorphism from $\mathcal{P} \times \mathbb{R}$ onto $\mathbb{R}^K$. Indeed, its inverse takes $\boldsymbol{p} = \text{softmax}(\boldsymbol{\theta}) \in \Delta^\circ_K$, the corresponding blow-up coordinates, and $c = \log \sum_r e^{\theta_r} \in \mathbb{R}$. Adding a constant to every initial logit leaves the policy trajectory unchanged. Hence a basin $\mathcal{A} \subseteq \mathcal{P}$ corresponds exactly to the basin $\Psi(\mathcal{A} \times \mathbb{R}) \subseteq \mathbb{R}^K$ in the logit space. Moreover, if $\mathcal{A}$ is Lebesgue-null in $\mathcal{P}$, then $\mathcal{A} \times \mathbb{R}$ is Lebesgue-null by Fubini's theorem, and its image under Ψ is Lebesgue-null because Ψ is locally Lipschitz. The diffeomorphism also preserves openness and nonemptiness.

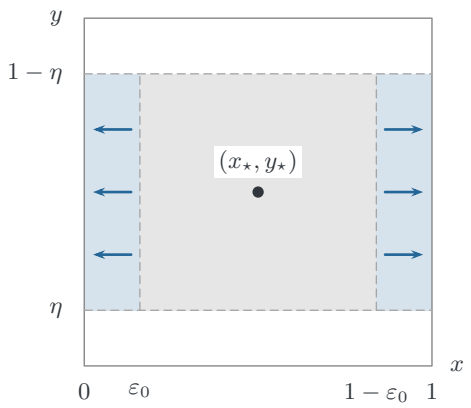
Thus, the mixed basin $\mathcal{N} := \Psi(\widetilde{\mathcal{N}}_{\text{mix}} \times \mathbb{R})$ is null in $\mathbb{R}^K$, and each vertex basin $\mathcal{O}_i := \Psi(\widetilde{\mathcal{O}}_i \times \mathbb{R})$ is nonempty and open for each $i \in \mathcal{I}$. Finally, since every convergent policy has nonempty support in $\mathcal{I}$, its initial condition belongs to exactly one of these basins, establishing the desired disjoint partition of the set $\{\boldsymbol{\theta}(0) \in \mathbb{R}^K : \boldsymbol{p}(t) \text{ converges}\}$. This proves Theorem 4.2. □

I.3 Proof of Theorem 4.3: The 2×2 Case

It suffices to show convergence in policy space for all initializations. Indeed, after showing this, Theorem 4.2 gives the desired selection result. The proof divides into two cases. If the normalized probability of either retained action becomes sufficiently small at a sufficiently large time, it converges to zero. Otherwise, the limit set stays in the interior of the square, where planar index theory implies that the only possible ICT limit set is the singleton containing the unique saddle. In both cases, we have convergence in policy space.

Write $\boldsymbol{B} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $\boldsymbol{x} = (x, 1-x)$, and $\boldsymbol{\lambda} = (y, 1-y)$. In these coordinates, the boundary system is

$$\begin{aligned} x' &= x(1-x)\big[x(ay + c(1-y)) - (1-x)(by + d(1-y))\big], \\ y' &= y(1-y)\big[(1-y)(cx + d(1-x)) - y(ax + b(1-x))\big]. \end{aligned} \tag{70}$$



Case 1: enter a blue strip
Then $x \to 0$ or $x \to 1$.
The policy converges to a vertex.

Case 2: stay in the interior rectangle
The limit set is compact and ICT.
The unique interior saddle gives
$\Omega = \{(x_\star, y_\star)\}$.

Figure 6: **The 2×2 Case.** For $\tau \geq \tau_0$, we have $y \in [\eta, 1-\eta]$. If the trajectory enters either blue strip, the policy converges to a vertex; otherwise, the limit set lies in the gray interior rectangle and consists only of the mixed saddle. Arrows indicate the sign of $dx/d\tau$ in the strips.

By Lemma I.2(i),(iii), this system has a unique equilibrium $(x_\star, y_\star)$ in $(0,1)^2$, which is a hyperbolic saddle. The balance equations give $axy = d(1-x)(1-y)$ and $cx(1-y) = b(1-x)y$, so this equilibrium is

$$x_\star = \frac{\sqrt{bd}}{\sqrt{ac} + \sqrt{bd}}, \qquad y_\star = \frac{\sqrt{cd}}{\sqrt{ab} + \sqrt{cd}}. \tag{71}$$

To apply Corollary B.9, we also need the boundary flow to be complete on $(0,1)^2$. The field in (70) is polynomial, and the factors $x(1-x)$ and $y(1-y)$ make each edge of $[0,1]^2$ invariant. Thus solutions in the closed square exist for all positive and negative times. By uniqueness of solutions, a trajectory starting in the open square cannot meet an edge at a finite time. Hence the boundary flow is complete on $(0,1)^2$, as required.

We now fix any initial condition for the full flow. Choose $\varepsilon_0 \in (0, 1/2)$ small enough that $\varepsilon_0 b_+ - (1-\varepsilon_0)^2 b_- \leq -b_-/2$. Let $C_\gamma \geq 1$ be a uniform bound on $|r_{i,\gamma}|$, $i \in \mathcal{I}$ (where $r_{i,\gamma}$ is the polynomial residual from the $\log x_i$ ODE in Lemma I.1). Since $\rho(\tau) \to 0$, Lemma I.1 gives $\tau_0 < \infty$ such that

$$C_\gamma \rho(\tau) \leq \frac{b_-}{4}, \qquad \min_{j \in \mathcal{L}} \lambda_j(\tau) \geq \eta \qquad (\tau \geq \tau_0). \tag{72}$$

We use this choice in both cases below; the corresponding regions are illustrated in Figure 6.

Case 1: $x_i \leq \varepsilon_0$ at some $\tau \geq \tau_0$. Suppose that $x_i(\tau_1) \leq \varepsilon_0$ for some retained action i and some $\tau_1 \geq \tau_0$. Let k denote the other action in $\mathcal{I}$, so $x_k = 1 - x_i$. Whenever $x_i \leq \varepsilon_0$, we have $u_i \leq \varepsilon_0 b_+$ and $U \geq x_k^2 (\boldsymbol{B}^\top \boldsymbol{\lambda})_k \geq (1-\varepsilon_0)^2 b_-$. Thus Lemma I.1 and (72) imply

$$(\log x_i)' \leq \varepsilon_0 b_+ - (1-\varepsilon_0)^2 b_- + C_\gamma \rho \leq -\frac{b_-}{4}.$$

For $\tau \geq \tau_0$, the derivative of x_i is strictly negative at $x_i = \varepsilon_0$, so x_i cannot cross this level upward. Hence $x_i(\tau) \leq \varepsilon_0$ for every $\tau \geq \tau_1$, and integrating the inequality gives

$$x_i(\tau) \leq x_i(\tau_1) e^{-b_-(\tau - \tau_1)/4} \qquad (\tau \geq \tau_1).$$

Thus $x_i \to 0$ and $x_k \to 1$ as $\tau \to \infty$. Since $\tau(t) \to \infty$ and $\rho(t) \to 0$, the policy converges to $\boldsymbol{e}_k$.

Case 2: Both normalized probabilities remain above ε_0. Otherwise, $x_i(\tau) > \varepsilon_0$ for both $i \in \mathcal{I}$ and every $\tau \geq \tau_0$. Together with (72), this gives

$$(x(\tau), y(\tau)) \in \mathcal{R} \triangleq [\varepsilon_0, 1-\varepsilon_0] \times [\eta, 1-\eta] \subset (0,1)^2 \qquad (\tau \geq \tau_0).$$

By Theorem D.1, the limit set $\Omega := \omega(x, y)$ is nonempty, compact, and internally chain transitive for the boundary flow. It is contained in the rectangle $\mathcal{R}$, hence in $(0,1)^2$. We have already verified that

the boundary flow is complete on $(0,1)^2$ and has a unique equilibrium there, the hyperbolic saddle $(x_\star, y_\star)$. Thus Corollary B.9 yields

$$\Omega = \{(x_\star, y_\star)\}. \tag{73}$$

Since the trajectory remains in a compact rectangle, (73) implies $(x(\tau), y(\tau)) \to (x_\star, y_\star)$ as $\tau \to \infty$. Together with $\tau(t) \to \infty$ and $\rho(t) \to 0$, this gives $\boldsymbol{p}(t) \to (x_\star, 1 - x_\star, 0, 0)$. □

I.4 Convergence Rates for the Selected Limits

Once a policy limit is identified, Theorem 3.3 yields the extinction rates provided the zero-margin boundary logits stabilize. Under flat dominance, this condition is automatically satisfied, as every omitted retained logit converges to a finite limit. For rank-one games, exact ratios among the selected actions further determine the total policy error.

Corollary I.1 (Rates at a Policy Limit). *For a nonempty $S \subseteq \mathcal{I}$, suppose $\boldsymbol{p}(t) \to (\boldsymbol{q}^S, \boldsymbol{0})$ and put $s := |S|$. Then every omitted retained logit $\theta_k(t)$, $k \in \mathcal{I} \setminus S$, converges to a finite limit. Hence the tied-logit stabilization condition of Theorem 3.3 holds. Moreover, for every $j \in \mathcal{L}$ and every $k \in \mathcal{I} \setminus S$,*

$$p_j(t) = \frac{s}{(s+l)(\boldsymbol{B}\boldsymbol{q}^S)_j}\, t^{-1}(1 + o(1)), \tag{74}$$

$$p_k(t) = c_k\, t^{-l/(s+l)}(1 + o(1)) \quad \textit{for some } c_k > 0. \tag{75}$$

In particular, if the selected limit is a vertex $\boldsymbol{e}_i$, then

$$\|\boldsymbol{p}(t) - \boldsymbol{e}_i\|_1 = 2\left(\sum_{k \in \mathcal{I}\setminus\{i\}} c_k\right) t^{-l/(l+1)}(1 + o(1)). \tag{76}$$

Proof. To apply Theorem 3.3, it suffices to show that each zero-margin boundary logit θ_k, $k \in \mathcal{I} \setminus S$, converges to a finite limit. By Lemma I.2(ii), $(\boldsymbol{x}, \boldsymbol{\lambda}) \to (\boldsymbol{q}^S, \boldsymbol{\lambda}^S)$. Hence Lemma I.1 and the polynomial extension of $\dot{\boldsymbol{\theta}}/\rho$ give

$$\frac{d}{d\tau}\log x_k \to -\alpha_S, \qquad \frac{d\theta_k}{d\tau} = x_k(\boldsymbol{B}^\top \boldsymbol{\lambda})_k + \mathcal{O}(\rho).$$

Since $\alpha_S > 0$, x_k decays at least exponentially in boundary time. Moreover, Lemma E.4 gives

$$\int_0^\infty \rho(\tau)\, d\tau = \int_0^\infty \rho(t)^2\, dt < \infty.$$

The bound

$$\left|\frac{d\theta_k}{d\tau}\right| \le C(x_k + \rho)$$

therefore makes $d\theta_k/d\tau$ absolutely integrable. Thus each θ_k converges, as does their sum. If $S = \mathcal{I}$, there are no such coordinates, and the required sum is zero.

At $(\boldsymbol{q}^S, \boldsymbol{0})$, the Nash margin is zero on $\mathcal{I}$ and equals $\boldsymbol{B}\boldsymbol{q}^S > 0$ on $\mathcal{L}$. We can therefore apply Theorem 3.3 with supported set S, positive-margin boundary set $\mathcal{L}$, and zero-margin boundary set $\mathcal{I} \setminus S$. This gives (74) and (75). For a vertex limit, $\mathcal{I} \setminus \{i\}$ is nonempty because $m \ge 2$. Its $t^{-l/(l+1)}$ mass decays more slowly than the t^{-1} dominated mass. Thus $\|\boldsymbol{p} - \boldsymbol{e}_i\|_1 = 2\sum_{r \ne i} p_r$ gives (76). □

Corollary I.2 (Rank-One Rates). *Under the assumptions of Theorem 4.4, let* $S := \mathcal{M}(\boldsymbol{\theta}(0))$, $s := |S|$, *and* $H_\star := \sum_{i\in S} b_i^{-1}$. *For the limit* $\boldsymbol{q}^\star$ *in* (8),

$$(\boldsymbol{B}\boldsymbol{q}_{\mathcal{I}}^\star)_j = \frac{sa_j}{H_\star}.$$

Consequently,

$$p_j(t) = \frac{H_\star}{(s+l)a_j}\, t^{-1}(1+o(1)) \qquad (j \in \mathcal{L}). \tag{77}$$

If $S \subsetneq \mathcal{I}$, *then*

$$\|\boldsymbol{p}(t) - \boldsymbol{q}^\star\|_1 = \Theta\Big(t^{-l/(s+l)}\Big).$$

If $S = \mathcal{I}$, *then* $\|\boldsymbol{p}(t) - \boldsymbol{q}^\star\|_1 = \Theta(t^{-1})$.

Proof. The identity for $\boldsymbol{B}\boldsymbol{q}_{\mathcal{I}}^\star$ and Corollary I.1 give (77) and the rates of every coordinate outside S. To control the total error, use the exact ratios from Appendix H.2: $p_i(t) = q_i^\star \sum_{r\in S} p_r(t)$ for every $i \in S$. All supported coordinates are therefore below their limiting values by the same relative factor, and

$$\|\boldsymbol{p}(t) - \boldsymbol{q}^\star\|_1 = 2\sum_{r\notin S} p_r(t) = 2\left(\sum_{k\in\mathcal{I}\setminus S} p_k(t) + \rho(t)\right). \tag{78}$$

If $S \subsetneq \mathcal{I}$, the positive $t^{-l/(s+l)}$ contributions from $\mathcal{I}\setminus S$ dominate the t^{-1} dominated mass. If $S = \mathcal{I}$, the right-hand side equals $2\rho = \Theta(t^{-1})$. This proves both rates. □

I.5 Proofs of the Technical Lemmas

I.5.1 Proof of Lemma I.1: Boundary Estimates

Step 1: Logarithmic derivatives and the bounds on κ. Let $\widetilde{\mathcal{V}} := \dot{\boldsymbol{\theta}}/\rho = d\boldsymbol{\theta}/d\tau$. As shown in the proof of Proposition D.1, $\widetilde{\mathcal{V}}$ extends polynomially in $(\rho, \boldsymbol{x}, \boldsymbol{\lambda})$, with

$$\widetilde{\mathcal{V}}_i(0,\boldsymbol{z}) = u_i(\boldsymbol{z}) \quad (i\in\mathcal{I}), \qquad \widetilde{\mathcal{V}}_j(0,\boldsymbol{z}) = w_j(\boldsymbol{z}) \quad (j\in\mathcal{L}).$$

By Lemma D.1,

$$\frac{d}{d\tau}\log x_i = \widetilde{\mathcal{V}}_i - \sum_{k\in\mathcal{I}} x_k\widetilde{\mathcal{V}}_k, \qquad \frac{d}{d\tau}\log\lambda_j = \widetilde{\mathcal{V}}_j - \sum_{k\in\mathcal{L}}\lambda_k\widetilde{\mathcal{V}}_k.$$

The right-hand sides are polynomial and equal $u_i - U$ and $w_j + V$, respectively, at $\rho = 0$. Subtracting these values leaves polynomials with zero constant term in ρ, so each difference is divisible by ρ. This proves (63). The resulting polynomials $r_{i,\gamma}$ and $s_{j,\gamma}$ are uniformly bounded on the compact set $[0,1]\times\Delta_{\mathcal{I}}\times\Delta_{\mathcal{L}}$. In particular, this bound holds even when x_i or λ_j is close to zero.

Moreover, $(\boldsymbol{B}^\top\boldsymbol{\lambda})_i$ and $(\boldsymbol{B}\boldsymbol{x})_j$ both lie in $[b_-, b_+]$ by flat dominance (Assumption 1). Since the coordinates of each vector $\boldsymbol{x}$ and $\boldsymbol{\lambda}$ are nonnegative and sum to one, Cauchy–Schwarz gives

$$\frac{b_-}{m} \le b_-\sum_{i\in\mathcal{I}} x_i^2 \le U \le b_+, \qquad \frac{b_-}{l} \le b_-\sum_{j\in\mathcal{L}}\lambda_j^2 \le V \le b_+. \tag{79}$$

Adding these inequalities and using $\kappa = U + V$ proves (64).

Step 2: A lower bound on each λ_j. We now show that each λ_j is at least $\eta = b_-/(4lb_+)$ for all sufficiently large τ. By (63) and the definition of w_j,

$$\frac{d}{d\tau}\log\lambda_j = V - \lambda_j(\boldsymbol{B}\boldsymbol{x})_j + \rho s_{j,\gamma}(\rho,\boldsymbol{z}).$$

Since $s_{j,\gamma}$ is uniformly bounded and $\rho(\tau)\to 0$, there is $\tau_0 < \infty$ such that

$$\big|\rho(\tau)s_{j,\gamma}(\rho(\tau),\boldsymbol{z}(\tau))\big| \le \frac{b_-}{4l} \qquad (j\in\mathcal{L},\ \tau\ge\tau_0).$$

Using $V \geq b_-/l$ from **Step 1** and $(\boldsymbol{B}\boldsymbol{x})_j \leq b_+$, we obtain, whenever $\lambda_j \leq \eta$ and $\tau \geq \tau_0$,

$$\frac{d}{d\tau} \log \lambda_j \geq \frac{b_-}{l} - \eta b_+ - \frac{b_-}{4l} = \frac{b_-}{2l} > 0.$$

Fix $j \in \mathcal{L}$ with $\lambda_j(\tau_0) < \eta$. If λ_j never reached η, integrating this inequality would give

$$\lambda_j(\tau) \geq \lambda_j(\tau_0) \exp\left(\frac{b_-}{2l}(\tau - \tau_0)\right) \qquad (\tau \geq \tau_0).$$

Since $\lambda_j(\tau_0) > 0$ and $\tau(t) \to \infty$, this contradicts $\lambda_j(\tau) < \eta$ for all $\tau \geq \tau_0$. Thus every coordinate below η at τ_0 reaches it in finite boundary time. At $\lambda_j = \eta$ and $\tau \geq \tau_0$, the derivative of λ_j is strictly positive, so it cannot cross below η. Since $\mathcal{L}$ is finite, there is a common time after which $\lambda_j \geq \eta$ for every $j \in \mathcal{L}$, proving the claim. □

I.5.2 Proof of Lemma I.2: Balanced Equilibria

Part (i): Existence and uniqueness. Fix a nonempty set $S \subseteq \mathcal{I}$ and let $s := |S|$. In this part, we identify vectors on Δ_S with their S-coordinates. We first express $\boldsymbol{\lambda}$ in terms of $\boldsymbol{x}$ at an equilibrium, and then show that the remaining equations determine $\boldsymbol{x}$ uniquely. Let

$$\boldsymbol{B}_S := \boldsymbol{B}_{\mathcal{L},S}, \qquad E_S := \{\boldsymbol{\varphi} \in \mathbb{R}^S : \mathbf{1}^\top \boldsymbol{\varphi} = 0\}.$$

For $\boldsymbol{x} \in \operatorname{relint}(\Delta_S)$, define

$$\delta_j^S(\boldsymbol{x}) := (\boldsymbol{B}_S \boldsymbol{x})_j, \qquad C_S(\boldsymbol{x}) := \sum_{j \in \mathcal{L}} \frac{1}{\delta_j^S(\boldsymbol{x})}. \tag{80}$$

At an equilibrium with $\boldsymbol{\lambda} \in \Delta_{\mathcal{L}}^\circ$, the boundary system (62) implies that $\lambda_j \delta_j^S(\boldsymbol{x}) = \beta$ for every $j \in \mathcal{L}$, for some $\beta > 0$. Using $\sum_{j \in \mathcal{L}} \lambda_j = 1$ gives

$$\lambda_j = \frac{\delta_j^S(\boldsymbol{x})^{-1}}{C_S(\boldsymbol{x})}, \qquad \beta = C_S(\boldsymbol{x})^{-1}. \tag{81}$$

Substituting (81) into the boundary system (62) when $\boldsymbol{x}$ is an equilibrium gives

$$x_i \sum_{j \in \mathcal{L}} \frac{B_{ji}}{\delta_j^S(\boldsymbol{x})} = \frac{l}{s} \qquad (i \in S). \tag{82}$$

Indeed, $x_i' = x_i(u_i - U) = 0$ and $x_i > 0$ imply $u_i = U$ for every $i \in S$. Since $\lambda_j = \beta/\delta_j^S(\boldsymbol{x})$, the left-hand side of (82) equals $u_i/\beta = U/\beta$, independently of i. Summing these s equal quantities gives

$$\sum_{i \in S} x_i \sum_{j \in \mathcal{L}} \frac{B_{ji}}{\delta_j^S(\boldsymbol{x})} = \sum_{j \in \mathcal{L}} \frac{\sum_{i \in S} B_{ji} x_i}{\delta_j^S(\boldsymbol{x})} = l.$$

Thus $sU/\beta = l$, proving (82). Conversely, (82) and (81) make u_i constant on S and w_j constant on $\mathcal{L}$, so both equations in (62) vanish.

We now prove that (82) has a unique solution, which we do by expressing it as the stationarity condition for a strictly convex function. For $\boldsymbol{\varphi} \in \mathbb{R}^S$, write $\boldsymbol{x} = \operatorname{softmax}(\boldsymbol{\varphi})$ and define

$$W_S(\boldsymbol{\varphi}) := \sum_{j \in \mathcal{L}} \log\left(\sum_{i \in S} B_{ji} e^{\varphi_i}\right) - \frac{l}{s} \sum_{i \in S} \varphi_i. \tag{83}$$

Differentiating in φ_i gives

$$\frac{\partial W_S}{\partial \varphi_i} = x_i \sum_{j \in \mathcal{L}} \frac{B_{ji}}{\delta_j^S(\boldsymbol{x})} - \frac{l}{s}, \qquad i \in S. \tag{84}$$

The coordinates of the gradient in (84) sum to zero, so $\nabla W_S(\boldsymbol{\varphi}) \in E_S$. Thus $\boldsymbol{\varphi} \in E_S$ is a stationary point of $W_S|_{E_S}$ if and only if $\nabla W_S(\boldsymbol{\varphi}) = \mathbf{0}$, which is equivalent to (82). Moreover, softmax restricted to E_S is a diffeomorphism onto $\operatorname{relint}(\Delta_S)$, so this parametrization is one-to-one.

We now verify strict convexity and existence of a minimizer on E_S.

Suppose first that $s \geq 2$, and set

$$\pi_{ji}(\boldsymbol{\varphi}) := \frac{B_{ji}e^{\varphi_i}}{\sum_{k\in S} B_{jk}e^{\varphi_k}} \qquad (j \in \mathcal{L},\ i \in S).$$

For any $\boldsymbol{a} \in E_S$, the Hessian satisfies

$$\boldsymbol{a}^\top \nabla^2 W_S(\boldsymbol{\varphi})\boldsymbol{a} = \sum_{j\in\mathcal{L}} \left[\sum_{i\in S} \pi_{ji} a_i^2 - \left(\sum_{i\in S} \pi_{ji} a_i \right)^2 \right]. \tag{85}$$

Each bracket is the variance of $(a_i)_{i\in S}$ under the strictly positive weights $(\pi_{ji})_{i\in S}$. It vanishes only when all a_i are equal, which for $\boldsymbol{a} \in E_S$ implies $\boldsymbol{a} = \mathbf{0}$. Hence the Hessian is positive definite on E_S, and $W_S|_{E_S}$ is strictly convex. For existence of a minimizer on E_S, observe that

$$W_S(\boldsymbol{\varphi}) \geq l \log b_- + l \max_{i\in S} \varphi_i \qquad (\boldsymbol{\varphi} \in E_S).$$

Since $\sum_{i\in S} \varphi_i = 0$, we also have $\|\boldsymbol{\varphi}\|_\infty \leq (s-1)\max_{i\in S} \varphi_i$. Therefore $W_S(\boldsymbol{\varphi}) \to \infty$ as $\|\boldsymbol{\varphi}\|_\infty \to \infty$ within E_S; that is, $W_S|_{E_S}$ is coercive. Continuity and coercivity give a minimizer, and strict convexity makes it unique. Denote it by $\boldsymbol{\varphi}^S$. Finally, if $s = 1$, then $E_S = \{0\}$, and $\boldsymbol{\varphi}^S = 0$ is its unique minimizer.

Taking $\boldsymbol{q}^S = \mathrm{softmax}(\boldsymbol{\varphi}^S)$ and defining $\boldsymbol{\lambda}^S$ by (81) gives the unique balanced equilibrium, with zero coordinates of $\boldsymbol{q}^S$ outside S. Finally, (81) gives $\beta_S = C_S(\boldsymbol{q}^S)^{-1}$, and (82) gives $\alpha_S = (l/s)\beta_S$. This proves **Part (i)**.

Part (ii): Policy limits. Suppose $\boldsymbol{p}(t) \to \boldsymbol{q}$ and let $S := \mathrm{supp}(\boldsymbol{q})$. Since $\rho(t) \to 0$, we have $S \subseteq \mathcal{I}$. In this part, we identify $\boldsymbol{q}$ with its $\mathcal{I}$-coordinates, so $\boldsymbol{x} \to \boldsymbol{q}$. It remains to prove convergence of $\boldsymbol{\lambda}$. By Theorem D.1 and Lemma I.1, the boundary limit set $\Omega := \omega(\boldsymbol{x}, \boldsymbol{\lambda})$ is nonempty, compact, and invariant, with

$$\Omega \subseteq \{\boldsymbol{q}\} \times \{\boldsymbol{\lambda} \in \Delta_{\mathcal{L}} : \min_{j\in\mathcal{L}} \lambda_j \geq \eta\}.$$

We show that $\boldsymbol{\lambda}$ has the same value at every point of Ω. Let

$$d_j := (\boldsymbol{B}\boldsymbol{q})_j > 0, \qquad \beta := \left(\sum_{j\in\mathcal{L}} d_j^{-1} \right)^{-1}, \qquad \bar{\lambda}_j := \frac{\beta}{d_j}.$$

Along any boundary trajectory in Ω, we have $\boldsymbol{x} \equiv \boldsymbol{q}$. Thus, the boundary system (62) becomes $\lambda_j' = \lambda_j(V_{\boldsymbol{q}} - \lambda_j d_j)$, where $V_{\boldsymbol{q}} := \sum_{j\in\mathcal{L}} d_j \lambda_j^2$. To show that $\Omega = \{(\boldsymbol{q}, \bar{\boldsymbol{\lambda}})\}$, consider

$$H(\boldsymbol{\lambda}) := D_{\mathrm{KL}}(\bar{\boldsymbol{\lambda}} \| \boldsymbol{\lambda}) = \sum_{j\in\mathcal{L}} \bar{\lambda}_j \log \frac{\bar{\lambda}_j}{\lambda_j}.$$

This function is nonnegative and vanishes only at $\bar{\boldsymbol{\lambda}}$. Along a boundary trajectory in Ω, differentiation gives

$$H' = -\sum_{j\in\mathcal{L}} \bar{\lambda}_j (V_{\boldsymbol{q}} - \lambda_j d_j) = \beta - V_{\boldsymbol{q}} = -\sum_{j\in\mathcal{L}} d_j (\lambda_j - \bar{\lambda}_j)^2. \tag{86}$$

Here we used $d_j \bar{\lambda}_j = \beta$ and $\sum_{j\in\mathcal{L}} \lambda_j = \sum_{j\in\mathcal{L}} \bar{\lambda}_j = 1$.

Since Ω is compact and every $\lambda_j \geq \eta$ on Ω, H is continuous and attains a maximum at some point $(\boldsymbol{q}, \widehat{\boldsymbol{\lambda}}) \in \Omega$. If $\widehat{\boldsymbol{\lambda}} \neq \bar{\boldsymbol{\lambda}}$, then (86) gives $H' < 0$ at $(\boldsymbol{q}, \widehat{\boldsymbol{\lambda}})$. For sufficiently small negative times, the boundary trajectory through this point remains in Ω by invariance and has larger H, contradicting maximality. Hence $\widehat{\boldsymbol{\lambda}} = \bar{\boldsymbol{\lambda}}$, so the maximum of H on Ω is zero. Since $H \geq 0$ and vanishes only at $\bar{\boldsymbol{\lambda}}$, this proves $\Omega = \{(\boldsymbol{q}, \bar{\boldsymbol{\lambda}})\}$.

The trajectory lies in the compact simplex product, so this singleton limit set implies $(\boldsymbol{x}, \boldsymbol{\lambda}) \to (\boldsymbol{q}, \bar{\boldsymbol{\lambda}})$. Invariance of Ω makes $(\boldsymbol{q}, \bar{\boldsymbol{\lambda}})$ an equilibrium. By the uniqueness proved in **Part (i)**, it equals $\boldsymbol{z}^S$, proving **Part (ii)**.

Part (iii): Stability. Fix a nonempty $S \subseteq \mathcal{I}$ and put $s := |S|$, $\boldsymbol{B}_S := \boldsymbol{B}_{\mathcal{L},S}$, and write

$$E_S := \{\boldsymbol{a} \in \mathbb{R}^S : \mathbf{1}^\top \boldsymbol{a} = 0\}, \qquad E_l := \{\boldsymbol{b} \in \mathbb{R}^l : \mathbf{1}^\top \boldsymbol{b} = 0\}.$$

In this part, $\boldsymbol{q}^S$ denotes its vector of S-coordinates.

We first compute the eigenvalues for perturbations within $\operatorname{relint}(\Delta_S) \times \Delta^\circ_{\mathcal{L}}$. We then consider the remaining coordinates x_r, $r \in \mathcal{I} \setminus S$. Define

$$\boldsymbol{R}_S := \operatorname{diag}(\boldsymbol{q}^S)\boldsymbol{B}_S^\top \operatorname{diag}(\boldsymbol{\lambda}^S) \in \mathbb{R}^{s\times l}. \tag{87}$$

At the balanced equilibrium, **Part (i)** gives

$$\boldsymbol{R}_S \mathbf{1}_l = \alpha_S \mathbf{1}_s, \qquad \boldsymbol{R}_S^\top \mathbf{1}_s = \beta_S \mathbf{1}_l, \qquad s\alpha_S = l\beta_S.$$

Consequently, $\boldsymbol{R}_S$ maps E_l into E_S, and $\boldsymbol{R}_S^\top$ maps E_S into E_l. To compute the Jacobian on this face, use centered logits $\boldsymbol{x} = \operatorname{softmax}(\boldsymbol{\varphi})$ on S and $\boldsymbol{\lambda} = \operatorname{softmax}(\boldsymbol{\psi})$, with $(\boldsymbol{\varphi}, \boldsymbol{\psi}) \in E_S \times E_l$. For a perturbation $(\boldsymbol{a}, \boldsymbol{b}) \in E_S \times E_l$, differentiating softmax at the centered logits corresponding to $(\boldsymbol{q}^S, \boldsymbol{\lambda}^S)$ gives

$$\delta x_i = q_i^S\big(a_i - (\boldsymbol{q}^S)^\top \boldsymbol{a}\big), \qquad \delta\lambda_j = \lambda_j^S\big(b_j - (\boldsymbol{\lambda}^S)^\top \boldsymbol{b}\big).$$

Let $c := (\boldsymbol{q}^S)^\top \boldsymbol{a} + (\boldsymbol{\lambda}^S)^\top \boldsymbol{b}$. Substituting these derivatives into the definitions of $\boldsymbol{u}$ and $\boldsymbol{w}$ and again using **Part (i)** gives

$$\delta\boldsymbol{u} = \alpha_S \boldsymbol{a} + \boldsymbol{R}_S \boldsymbol{b} - \alpha_S c \mathbf{1}_s, \qquad \delta\boldsymbol{w} = -\boldsymbol{R}_S^\top \boldsymbol{a} - \beta_S \boldsymbol{b} + \beta_S c \mathbf{1}_l.$$

In centered logits, the boundary system (62) is equivalent to

$$\varphi_i' = u_i - \frac{1}{s}\sum_{k\in S} u_k, \qquad \psi_j' = w_j - \frac{1}{l}\sum_{k\in\mathcal{L}} w_k.$$

Thus the terms proportional to $\mathbf{1}_s$ and $\mathbf{1}_l$ cancel when we subtract the respective means. Since $\boldsymbol{R}_S\boldsymbol{b} \in E_S$ and $\boldsymbol{R}_S^\top \boldsymbol{a} \in E_l$, the Jacobian on $E_S \times E_l$ is

$$\boldsymbol{J}_S = \begin{pmatrix} \alpha_S \boldsymbol{I}_{E_S} & \boldsymbol{R}_S|_{E_l} \\ -\boldsymbol{R}_S^\top|_{E_S} & -\beta_S \boldsymbol{I}_{E_l} \end{pmatrix}. \tag{88}$$

It remains to count the eigenvalues of (88). For a nonzero $\boldsymbol{b} \in E_l$, weighted Cauchy–Schwarz gives

$$\|\boldsymbol{R}_S\boldsymbol{b}\|^2 = \sum_{i\in S}\left(\sum_{j\in\mathcal{L}} (R_S)_{ij} b_j\right)^2 \le \alpha_S \sum_{i\in S}\sum_{j\in\mathcal{L}} (R_S)_{ij} b_j^2 = \alpha_S\beta_S\|\boldsymbol{b}\|^2.$$

Equality would require all b_j to be equal, since every entry of $\boldsymbol{R}_S$ is positive. This is impossible for nonzero $\boldsymbol{b} \in E_l$. Compactness of the unit sphere therefore gives $\|\boldsymbol{R}_S|_{E_l}\|^2 < \alpha_S\beta_S$ whenever $E_l \neq \{0\}$. Hence every singular value σ of the restricted map satisfies $\sigma^2 < \alpha_S\beta_S$. Choosing orthonormal bases of E_S and E_l from a singular-value decomposition of $\boldsymbol{R}_S|_{E_l}$ splits (88) into $\min\{s-1, l-1\}$ paired blocks

$$\begin{pmatrix} \alpha_S & \sigma \\ -\sigma & -\beta_S \end{pmatrix}, \qquad \det = \sigma^2 - \alpha_S\beta_S < 0.$$

Here the singular values include zero values when the restricted map is rank deficient. Each paired block has one positive and one negative eigenvalue. Every unpaired direction in E_S has eigenvalue $\alpha_S > 0$, and every unpaired direction in E_l has eigenvalue $-\beta_S < 0$. Thus the Jacobian restricted to the face has exactly $s-1$ positive and $l-1$ negative eigenvalues. This also covers $s = 1$ or $l = 1$: there are no paired blocks, and the remaining directions have the corresponding diagonal eigenvalues. The centered softmax parametrizations are diffeomorphisms, so the Jacobians in these coordinates and in simplex coordinates are similar at the equilibrium. Hence they have the same eigenvalues.

For every $r \in \mathcal{I} \setminus S$, use x_r as a transverse coordinate. At $\boldsymbol{z}^S$, the equation

$$\frac{dx_r}{d\tau} = x_r\big(x_r(\boldsymbol{B}^\top\boldsymbol{\lambda})_r - U\big)$$

has derivative $-U(\boldsymbol{z}^S) = -\alpha_S$ with respect to x_r. Its derivatives with respect to all other coordinates vanish because $x_r = 0$ at $\boldsymbol{z}^S$. Ordering these $m-s$ coordinates after the face coordinates therefore makes the full Jacobian block triangular, with lower diagonal block $-\alpha_S \boldsymbol{I}_{m-s}$. Consequently, the linearization on $T\mathcal{A}$ has exactly $s-1$ positive and $(l-1)+(m-s) = m-s+l-1$ negative eigenvalues. It is therefore hyperbolic, and it is a sink exactly when $s = 1$. This proves **Part (iii)**. □

J NONCONVERGENCE RESULTS OF S2PE FLOW

Consider the following payoff matrix with Nash set

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \end{pmatrix}, \quad \mathcal{SN}(\boldsymbol{A}) = \left\{ (x, 1-x, 0, 0)^\top : \frac{1}{3} \le x \le \frac{2}{3} \right\}. \tag{89}$$

For $0 < \eta < 1/2$ and $0 < \rho_0 < 1$, consider the following initialization for Equation (S2P-EG-Flow):

$$\boldsymbol{p}(0) = \big((1-\rho_0)(1/2+\eta),\ (1-\rho_0)(1/2-\eta),\ \rho_0/2,\ \rho_0/2\big)^\top, \qquad \theta_i(0) = \log p_i(0). \tag{90}$$

Thus, η controls the imbalance between the first two coordinates at initialization.

We now state the two nonconvergence results:

Proposition J.1 (Nash Convergence without Policy Convergence). *There is a constant $\varepsilon_* > 0$, independent of γ, such that for every $\gamma > 0$ and $0 < \rho_0 \le \frac{\varepsilon_*}{1+\gamma}$, the initialization in Eqn. (90) with $\eta = 1/20$ satisfies*

$$\rho(t) := p_3(t) + p_4(t) = \Theta(t^{-1}), \qquad \mathrm{DGap}(\boldsymbol{p}(t)) \to 0, \tag{91}$$

but $\boldsymbol{p}(t)$ does not converge. Its omega-limit set is a nondegenerate closed subinterval of $\mathcal{SN}(\boldsymbol{A})$.

Proposition J.2 (Failure of Convergence in the Nash Sense). *Under the same assumptions as Proposition J.1, setting $\eta = 1/5$ results in the following:*

$$\rho(t) = \Theta(t^{-1}), \qquad \liminf_{t\to\infty} \mathrm{DGap}(\boldsymbol{p}(t)) < 0 = \limsup_{t\to\infty} \mathrm{DGap}(\boldsymbol{p}(t)). \tag{92}$$

Note that in the setup of Proposition J.2, this implies that convergence in the policy space is also not possible for every $\gamma > 0$: if $\boldsymbol{p}(t)$ is convergent, then $\mathrm{DGap}(\boldsymbol{p}(t))$ should also converge.

We prove Propositions J.1 and J.2 together in four steps. We first isolate the planar boundary motion and construct annuli filled by its periodic orbits. We then control the accumulated perturbation and use Theorem D.1 to identify the limiting orbit. Finally, we project this orbit onto the policy face: the smaller initialization gives an interval inside the Nash segment, whereas the larger one gives an interval extending beyond it. Figure 7 illustrates the construction.

Step 1: Reduce the dynamics to three coordinates. We separate the leading boundary field from the smaller extragradient correction. Use the blow-up coordinates of Lemma D.1 with $\mathcal{I} = \{1, 2\}$ and $\mathcal{L} = \{3, 4\}$:

$$\rho = p_3 + p_4, \qquad x = \frac{p_1}{1-\rho}, \qquad y = \frac{p_3}{\rho}.$$

Thus ρ is the mass outside the face, and x, y describe the proportions within the two pairs of actions. These coordinates exist for every finite time. To describe the motion around $(x, y) = (1/2, 1/2)$, we introduce

$$a = x + y - 1, \qquad b = x - y, \qquad x = \frac{1+a+b}{2}, \quad y = \frac{1+a-b}{2}.$$

The initialization is $(a(0), b(0)) = (\eta, \eta)$. Direct computations yield

$$\boldsymbol{A}\boldsymbol{p} = \big(\rho(1-3y), \rho(3y-2), (1-\rho)(3x-1), (1-\rho)(2-3x)\big)^\top,$$
$$\boldsymbol{g}(\boldsymbol{p}) = \rho(1-\rho)\big(x(1-3y), (1-x)(3y-2), y(3x-1), (1-y)(2-3x)\big)^\top.$$

At $\rho = 0$, $\boldsymbol{S}(\boldsymbol{p})$ is supported on its first two rows and columns. Recalling that

$$D\boldsymbol{g}(\boldsymbol{\theta}) = \mathrm{diag}(\boldsymbol{A}\boldsymbol{p})\boldsymbol{S}(\boldsymbol{p}) + \mathrm{diag}(\boldsymbol{p})\boldsymbol{A}\boldsymbol{S}(\boldsymbol{p}). \tag{93}$$

Note that both terms vanish since $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = 0$ and $(\boldsymbol{A}\boldsymbol{p})_{\mathcal{I}} = 0$. $D\boldsymbol{g}$ is polynomial in (ρ, x, y), so it has a factor ρ. Consequently, $D\boldsymbol{g}\,\boldsymbol{g}$ has a factor $\rho^2(1-\rho)$.

Let us define the polynomial vector $\boldsymbol{r}$ by

$$\boldsymbol{r}(\rho, x, y) \coloneqq \frac{1}{2\rho^2(1-\rho)} D\boldsymbol{g}(\boldsymbol{p})\boldsymbol{g}(\boldsymbol{p}).$$

By the exact blow-up system as derived in Lemma D.1, we have

$$\dot{a} = \rho(1-\rho)\left[-\frac{b}{2}(1+3a^2-b^2) + \gamma\rho E_a\right],$$
$$\dot{b} = \rho(1-\rho)\left[a(1-a^2) + \gamma\rho E_b\right], \tag{94}$$
$$\dot{\rho} = \rho^2(1-\rho)^2[-\kappa(a,b) + \gamma\rho E_\rho], \qquad \kappa(a,b) \coloneqq \frac{1+7a^2-5b^2}{2}, \tag{95}$$

where the polynomials E_a, E_b, E_ρ are defined as

$$E_a = x(1-x)(r_1-r_2) + y(1-y)(r_3-r_4),$$
$$E_b = x(1-x)(r_1-r_2) - y(1-y)(r_3-r_4),$$
$$E_\rho = yr_3 + (1-y)r_4 - xr_1 - (1-x)r_2.$$

Note that they are (trivially) bounded on any fixed compact set.

Step 2: Closed curves for the boundary motion. We seek annuli filled by periodic orbits, with $\kappa > 0$ to ensure decay of the mass outside the face. The boundary field of Proposition D.1, written in (a, b)-coordinates, is

$$\mathcal{B}(a,b) = \left(-\frac{b}{2}(1+3a^2-b^2),\ a(1-a^2)\right).$$

We note that it admits a *first integral*

$$R(a,b) = \frac{(3+a^2-b^2)^2}{2+2a^2-b^2}. \tag{96}$$

Indeed,

$$\nabla R = \frac{2(3+a^2-b^2)}{(2+2a^2-b^2)^2}\big(2a(a^2-1),\ -b(1+3a^2-b^2)\big) \Longrightarrow \nabla R \cdot \mathcal{B} = 0.$$

For $0 < a^2+b^2 \le 1/6$, both $1-a^2$ and $1+3a^2-b^2$ are at least $5/6$. Hence

$$a\nabla_a R + b\nabla_b R < 0, \qquad \frac{a\mathcal{B}_2 - b\mathcal{B}_1}{a^2+b^2} \ge \frac{5}{12}, \qquad \frac{1}{12} \le \kappa \le \frac{13}{12}. \tag{97}$$

The first inequality means that R decreases along every ray from the origin; the second means that the boundary motion turns counterclockwise.

The two choices below place the initial level strictly between r_- and r_+. The smaller annulus will project inside the Nash segment; every level in the larger annulus will project beyond it.

Annulus	η	$R(\eta,\eta)$	(r_-, r_+)	Containing disk
$\mathcal{A}_\mathrm{s}$	$1/20$	$400/89$	$(449/100, 899/200)$	$a^2+b^2 < 1/36$
$\mathcal{A}_\mathrm{l}$	$1/5$	$75/17$	$(22/5, 53/12)$	$a^2+b^2 < 1/6$

For each row, with $d = 1/36$ or $1/6$ respectively, define the closed region

$$\mathcal{A} := \{(a,b) : a^2+b^2 \le d,\ r_- \le R(a,b) \le r_+\}.$$

In each row, $2\eta^2 < d$ and $r_- < R(\eta,\eta) < r_+ < 9/2$, so the initial point lies in the interior of $\mathcal{A}$. We can check via direct computations that these annuli indeed exist. On $a^2+b^2 = d \le 1/6$, $R = \frac{(3-d+2a^2)^2}{2-d+3a^2}$. Its derivative with respect to a^2 has the sign of $-1-d+6a^2 \le -1+5d < 0$. Thus its maximum on the circle is $(3-d)^2/(2-d)$, which is $11449/2556 < 449/100$ for $d = 1/36$ and $289/66 < 22/5$ for $d = 1/6$.

Since $R(0,0) = 9/2 > r_+$, strict radial decrease makes every level $R = r$ with $r \in [r_-, r_+]$ meet every ray exactly once, inside its containing disk. The radial derivative is nonzero, so these levels are smooth closed curves. The field $\mathcal{B}$ is tangent to them and has positive angular speed, so its trajectories on them are periodic.

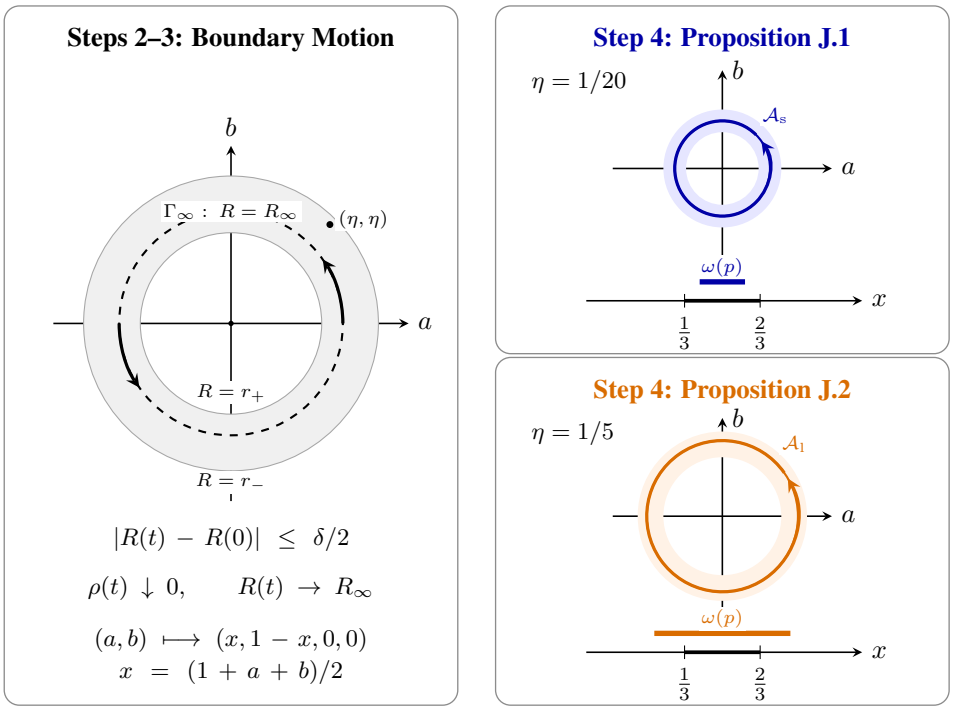


Figure 7: Trapping and policy limits for Propositions J.1 and J.2. Black intervals mark the Nash segment. Schematic.

Step 3: Trap the trajectory and identify its boundary limit set. Along the full trajectory, the first integral satisfies

$$\dot{R} = \gamma\rho^2(1-\rho)(\nabla_a R\, E_a + \nabla_b R\, E_b). \tag{98}$$

We will show that its total drift is too small for the trajectory to leave the chosen annulus. Choose $M \ge 1$ bounding the following two continuous functions on $(\rho, a, b) \in [0, 1/2] \times (\mathcal{A}_s \cup \mathcal{A}_1)$:

$$|E_\rho|, \qquad |\nabla_a R\, E_a + \nabla_b R\, E_b|.$$

Such an M exists by compactness and is independent of γ. Each initial level in the table is at distance at least $\delta = 11/17800$ from either boundary level. Set

$$\varepsilon_* := \frac{\delta}{192M}. \tag{99}$$

Fix either initialization in the table and $0 < \rho_0 \le \varepsilon_*/(1+\gamma)$. As long as (a, b) stays in its annulus and $\rho \le \rho_0$, Eqns. (95) and (98) give

$$-2\rho^2 \le \dot{\rho} \le -\frac{\rho^2}{96}, \qquad |\dot{R}| \le M\gamma\rho^2. \tag{100}$$

The inequality is derived from $1/12 \le \kappa \le 13/12$, $|\gamma\rho E_\rho| < 1/24$, and $1/4 \le (1-\rho)^2 \le 1$. The coordinates remain interior at every finite time, so $\rho > 0$; its strict decrease prevents it from crossing above ρ_0. Integration yields

$$\int_0^t \rho(s)^2\, ds \le 96\rho_0, \qquad |R(a(t), b(t)) - R(\eta, \eta)| \le 96M\gamma\rho_0 \le \frac{\delta}{2}. \tag{101}$$

If (a, b) first reached the boundary of its annulus at time t_e, then $|R(a(t_e), b(t_e)) - R(\eta, \eta)| \ge \delta$, contradicting Eqn (101). Thus the trajectory remains in the interior of the annulus and these bounds hold for every $t \ge 0$. Moreover,

$$\frac{1}{\rho_0^{-1} + 2t} \le \rho(t) \le \frac{1}{\rho_0^{-1} + t/96}. \tag{102}$$

Therefore $\rho = \Theta(t^{-1})$, $\int_0^\infty \rho\, dt = \infty$, and $\int_0^\infty \rho^2\, dt < \infty$. The bound on $|\dot{R}|$ also gives a limit $R(t) \to R_\infty$ strictly between r_- and r_+. Thus the boundary clock $\tau(t) = \int_0^t \rho(s)\, ds$ diverges, while the accumulated drift of R is finite.

It remains to identify the boundary limit set. Write

$$\Omega := \omega(a, b), \qquad \Gamma_\infty := \{(a, b) \in \mathcal{A} : R(a, b) = R_\infty\}.$$

Since $R(t) \to R_\infty$, we have $\Omega \subseteq \Gamma_\infty$. As $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = 0$ and $\rho \to 0$, Theorem D.1 applies in boundary time and gives that Ω is nonempty and invariant under $\mathcal{B}$. The diverging clock does not change the accumulation set, and the change from (x, y) to (a, b) is invertible. Starting at any point of Ω, the boundary flow stays in Ω and traverses the entire periodic curve Γ_∞ from Step 2. Hence $\Omega = \Gamma_\infty$.

Step 4: Project the limiting curve and distinguish the two cases. Since $\rho \to 0$, the policy limit set is

$$\omega(\boldsymbol{p}) = \{(x, 1 - x, 0, 0)^\top : x = (1 + a + b)/2,\ (a, b) \in \Gamma_\infty\}.$$

It is a closed interval, as the continuous image of a compact connected curve into a line. The curve meets both the positive and negative diagonals $a = b$, giving $x > 1/2$ and $x < 1/2$, respectively; thus the interval is nondegenerate.

If $\eta = 1/20$, the trajectory stays in $\mathcal{A}_{\mathrm{s}}$, so

$$|a + b| \le \sqrt{2}/6 < 1/3, \qquad 1/3 < x < 2/3.$$

The last two entries of $\boldsymbol{A}\boldsymbol{p}$ are then nonnegative, and the first two are at least -2ρ. Hence $-2\rho \le \mathrm{DGap}(\boldsymbol{p}) \le 0$. Together with the nondegenerate policy limit interval, this proves Proposition J.1.

If $\eta = 1/5$, the trajectory stays in $\mathcal{A}_{\mathrm{l}}$. The limiting curve meets the positive diagonal $a = b > 0$, where

$$R(a, a) = \frac{9}{2 + a^2} \le \frac{53}{12}, \qquad x = \frac{1}{2} + a \ge \frac{1}{2} + \sqrt{\frac{2}{53}} > \frac{2}{3}.$$

The corresponding policy accumulation point has fourth payoff coordinate $2 - 3x$. Continuity of DGap therefore gives

$$\liminf_{t \to \infty} \mathrm{DGap}(\boldsymbol{p}(t)) \le -\left(3\sqrt{\frac{2}{53}} - \frac{1}{2}\right).$$

The curve also meets $a + b = 0$, so $(1/2, 1/2, 0, 0)^\top$ is a policy accumulation point. This policy is Nash; continuity and $\mathrm{DGap} \le 0$ give $\limsup_{t \to \infty} \mathrm{DGap}(\boldsymbol{p}(t)) = 0$. This proves Proposition J.2.

K GEOMETRY OF SYMMETRIC GAMES

We show that for a generic game, implicit bias becomes a trivial question because the equilibrium is *unique*.

Let $\boldsymbol{A} \in \mathrm{Skew}(K)$ be *principally nondegenerate* if, for every nonempty $S \subseteq [K]$,

$$\mathrm{nullity}(\boldsymbol{A}_{S,S}) = \begin{cases} 0, & |S| \text{ even}, \\ 1, & |S| \text{ odd}. \end{cases} \tag{103}$$

Throughout, let $\mathcal{S}(\boldsymbol{A}) := \bigcup_{\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})} \mathrm{supp}(\boldsymbol{q})$ denote the **maximal support** of the game.

Lemma K.1 (Complementary slackness)**.** *For every $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$,*

$$(\boldsymbol{A}\boldsymbol{q})_i = 0 \quad (i \in \mathrm{supp}(\boldsymbol{q})), \qquad (\boldsymbol{A}\boldsymbol{q})_j \geq 0 \quad (j \notin \mathrm{supp}(\boldsymbol{q})). \tag{104}$$

Proof. Since $\boldsymbol{q} \geq \boldsymbol{0}$, $\boldsymbol{A}\boldsymbol{q} \geq \boldsymbol{0}$, and $\boldsymbol{q}^\top \boldsymbol{A}\boldsymbol{q} = 0$ by skew-symmetry, every summand of $\sum_k q_k (\boldsymbol{A}\boldsymbol{q})_k = 0$ is nonnegative, hence zero. □

Proposition K.1 (Structure of $\mathcal{SN}(\boldsymbol{A})$)**.** *The following hold:*

1. *Every $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ satisfies $(\boldsymbol{A}\boldsymbol{q})_i = 0$ for all $i \in \mathcal{S}(\boldsymbol{A})$, and $\dim \mathcal{SN}(\boldsymbol{A}) = \mathrm{nullity}(\boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})}) - 1$;*
2. *$\mathrm{ri}\mathcal{SN}(\boldsymbol{A}) = \{\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A}) : \mathrm{supp}(\boldsymbol{q}) = \mathcal{S}(\boldsymbol{A}),\ (\boldsymbol{A}\boldsymbol{q})_j > 0\ \forall j \notin \mathcal{S}(\boldsymbol{A})\}$;*
3. *$\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ is quasi-strict if and only if $\boldsymbol{q} \in \mathrm{ri}\mathcal{SN}(\boldsymbol{A})$; consequently, the set of non-quasi-strict SNEs are exactly $\mathrm{rbd}\mathcal{SN}(\boldsymbol{A})$.*

Proof. We first show that $(\boldsymbol{A}\boldsymbol{q})_{\mathcal{S}(\boldsymbol{A})} = \boldsymbol{0}$. By the definition of $\mathcal{S}(\boldsymbol{A})$, every $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ satisfies $\boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})^c} = \boldsymbol{0}$. Fix $i \in \mathcal{S}(\boldsymbol{A})$. There exists $\boldsymbol{r} \in \mathcal{SN}(\boldsymbol{A})$ such that $r_i > 0$. Since $\boldsymbol{r}^\top \boldsymbol{A}\boldsymbol{q} = 0 = \boldsymbol{q}^\top \boldsymbol{A}\boldsymbol{r}$, $(\boldsymbol{A}\boldsymbol{q})_i = 0$ for every $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$. Moreover,

$$\boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})} \boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})} = (\boldsymbol{A}\boldsymbol{q})_{\mathcal{S}(\boldsymbol{A})} = \boldsymbol{0}.$$

Next, we establish the existence of a quasi-strict SNE with support $\mathcal{S}(\boldsymbol{A})$. Consider the primal-dual pair of linear programs

$$\min_{\boldsymbol{x} \in \Delta_K,\, v \in \mathbb{R}} v \quad \text{subject to} \quad \boldsymbol{A}^\top \boldsymbol{x} \leq v\boldsymbol{1},$$

and

$$\max_{\boldsymbol{y} \in \Delta_K,\, w \in \mathbb{R}} w \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{y} \geq w\boldsymbol{1}.$$

Since $\boldsymbol{A}^\top = -\boldsymbol{A}$, the common optimal value is zero. Thus, any optimal primal and dual variables $\boldsymbol{x}$ and $\boldsymbol{y}$ are in $\mathcal{SN}(\boldsymbol{A})$. By the Goldman–Tucker theorem for linear programs (Goldman & Tucker, 1956), there exists an optimal primal-dual pair satisfying

$$x_i + (\boldsymbol{A}\boldsymbol{y})_i > 0, \qquad y_i + (\boldsymbol{A}\boldsymbol{x})_i > 0, \qquad i \in [K]$$

from the strict complementarity. Indeed, the complementary pairs are $(y_i, v - (\boldsymbol{A}^\top \boldsymbol{x})_i)$ and $(x_i, (\boldsymbol{A}\boldsymbol{y})_i - w)$ and we have $v^\star = w^\star = 0$. Define

$$\boldsymbol{q}^\circ := \frac{\boldsymbol{x} + \boldsymbol{y}}{2}.$$

By convexity of $\mathcal{SN}(\boldsymbol{A})$, we have $\boldsymbol{q}^\circ \in \mathcal{SN}(\boldsymbol{A})$. If $q_i^\circ = 0$, then $x_i = y_i = 0$, so strict complementarity yields

$$(\boldsymbol{A}\boldsymbol{x})_i > 0, \qquad (\boldsymbol{A}\boldsymbol{y})_i > 0.$$

Consequently,

$$(\boldsymbol{A}\boldsymbol{q}^\circ)_i > 0.$$

On the other hand, if $q_i^\circ > 0$, then from $q_i^\circ (\boldsymbol{A}\boldsymbol{q}^\circ)_i = 0$ we have

$$(\boldsymbol{A}\boldsymbol{q}^\circ)_i = 0.$$

Thus $\boldsymbol{q}^\circ$ is quasi-strict.

It remains to identify its support. Fix $i \in \mathcal{S}(\boldsymbol{A})$. There exists $\boldsymbol{r} \in \mathcal{SN}(\boldsymbol{A})$ such that $r_i > 0$, yielding $(\boldsymbol{A}\boldsymbol{q}^\circ)_i = 0$. Since $\boldsymbol{q}^\circ$ is quasi-strict, $q_i^\circ = 0$ would imply $(\boldsymbol{A}\boldsymbol{q}^\circ)_i > 0$, which is impossible. Hence $q_i^\circ > 0$ for every $i \in \mathcal{S}(\boldsymbol{A})$. Since every SNE is supported on $\mathcal{S}(\boldsymbol{A})$,

$$\operatorname{supp}(\boldsymbol{q}^\circ) = \mathcal{S}(\boldsymbol{A}).$$

Quasi-strictness therefore also gives

$$(\boldsymbol{A}\boldsymbol{q}^\circ)_j > 0 \qquad \text{for every } j \notin \mathcal{S}(\boldsymbol{A}).$$

We next determine the affine hull. It is immediate that

$$\operatorname{aff}\mathcal{SN}(\boldsymbol{A}) \subseteq \left\{\boldsymbol{q} \in \mathbb{R}^K : \boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})^c} = \boldsymbol{0},\ \boldsymbol{1}^\top \boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})} = 1,\ \boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})}\boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})} = \boldsymbol{0}\right\}.$$

Let $\boldsymbol{h} \in \mathbb{R}^{|\mathcal{S}(\boldsymbol{A})|}$ satisfy

$$\boldsymbol{1}^\top \boldsymbol{h} = 0, \qquad \boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})}\boldsymbol{h} = \boldsymbol{0}. \tag{105}$$

Because $q_i^\circ > 0$ for every $i \in \mathcal{S}(\boldsymbol{A})$ and $(\boldsymbol{A}\boldsymbol{q}^\circ)_j > 0$ for every $j \notin \mathcal{S}(\boldsymbol{A})$, there exists $\varepsilon > 0$ such that

$$\boldsymbol{q}^\circ_{\mathcal{S}(\boldsymbol{A})} \pm \varepsilon \boldsymbol{h} > \boldsymbol{0}, \qquad \boldsymbol{A}_{\mathcal{S}(\boldsymbol{A})^c,\mathcal{S}(\boldsymbol{A})}(\boldsymbol{q}^\circ_{\mathcal{S}(\boldsymbol{A})} \pm \varepsilon \boldsymbol{h}) > \boldsymbol{0}.$$

Therefore, $\boldsymbol{q}^\circ_{\mathcal{S}(\boldsymbol{A})} \pm \varepsilon \boldsymbol{h}$ belong to $\mathcal{SN}(\boldsymbol{A})$. Hence every $\boldsymbol{h}$ satisfying (105) is a translation direction of $\operatorname{aff}\mathcal{SN}(\boldsymbol{A})$. This proves

$$\operatorname{aff}\mathcal{SN}(\boldsymbol{A}) = \left\{\boldsymbol{q} \in \mathbb{R}^K : \boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})^c} = \boldsymbol{0},\ \boldsymbol{1}^\top \boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})} = 1,\ \boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})}\boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})} = \boldsymbol{0}\right\}.$$

$\mathcal{SN}(\boldsymbol{A})$ is defined by the inequalities

$$q_i \geq 0 \quad \text{for } i \in \mathcal{S}(\boldsymbol{A}), \qquad (\boldsymbol{A}\boldsymbol{q})_j \geq 0 \quad \text{for } j \notin \mathcal{S}(\boldsymbol{A}). \tag{106}$$

The equilibrium $\boldsymbol{q}^\circ$ strictly satisfies (106). If $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ satisfies (106) as equalities, then $\boldsymbol{q}_\varepsilon := \boldsymbol{q} + \varepsilon(\boldsymbol{q} - \boldsymbol{q}^\circ) \in \operatorname{aff}\mathcal{SN}(\boldsymbol{A})$ but $\boldsymbol{q}_\varepsilon$ does not satisfy (106). Hence

$$\operatorname{ri}\mathcal{SN}(\boldsymbol{A}) = \{\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A}) : q_i > 0\ \forall i \in \mathcal{S}(\boldsymbol{A}),\ (\boldsymbol{A}\boldsymbol{q})_j > 0\ \forall j \notin \mathcal{S}(\boldsymbol{A})\}.$$

This proves item 2.

Next, we compute the dimension of $\mathcal{SN}(\boldsymbol{A})$. Since $\boldsymbol{q}^\circ_{\mathcal{S}(\boldsymbol{A})} \in \ker(\boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})})$ has coordinate sum $1 \neq 0$, the kernel is *not* contained in the hyperplane $\{\boldsymbol{1}^\top \boldsymbol{q} = 0\}$, so intersecting with it lowers the dimension by exactly one:

$$\dim \mathcal{SN}(\boldsymbol{A}) = \dim \operatorname{ri}\mathcal{SN}(\boldsymbol{A}) = \operatorname{nullity}(\boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})}) - 1.$$

This proves item 1. It remains to prove item 3:

($\Leftarrow$) If $\boldsymbol{q} \in \operatorname{ri}\mathcal{SN}(\boldsymbol{A})$, then $\operatorname{supp}(\boldsymbol{q}) = \mathcal{S}(\boldsymbol{A})$ and $(\boldsymbol{A}\boldsymbol{q})_j > 0$ for every $j \notin \mathcal{S}(\boldsymbol{A}) = \operatorname{supp}(\boldsymbol{q})$; that is, $\boldsymbol{q}$ is quasi-strict.

($\Rightarrow$) Let $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ be quasi-strict. If there were some $i \in \mathcal{S}(\boldsymbol{A}) \setminus \operatorname{supp}(\boldsymbol{q})$, then $q_i = 0$ and $(\boldsymbol{A}\boldsymbol{q})_i = 0$ by item 1, contradicting quasi-strictness. Hence $\operatorname{supp}(\boldsymbol{q}) = \mathcal{S}(\boldsymbol{A})$, and quasi-strictness gives $(\boldsymbol{A}\boldsymbol{q})_j > 0$ for every $j \notin \mathcal{S}(\boldsymbol{A})$; that is, $\boldsymbol{q} \in \operatorname{ri}\mathcal{SN}(\boldsymbol{A})$. □

Theorem K.1 (Generic uniqueness and quasi-strictness)**.** *The set of principally nondegenerate matrices has its complement in* $\operatorname{Skew}(K)$ *of Lebesgue measure zero. For every principally nondegenerate matrix* $\boldsymbol{A}$*, the symmetric Nash set is a singleton,* $\mathcal{SN}(\boldsymbol{A}) = \{\boldsymbol{q}_{\boldsymbol{A}}\}$*. Moreover,* $|\operatorname{supp}(\boldsymbol{q}_{\boldsymbol{A}})|$ *is odd and* $\boldsymbol{q}_{\boldsymbol{A}}$ *is quasi-strict.*

Proof. For brevity, denote the set of principally nondegenerate matrices as G.

Fix a nonempty $\bar{S} \subseteq [K]$. If $|\bar{S}|$ is even, $\det(\boldsymbol{A}_{\bar{S},\bar{S}})$ is a nonzero polynomial. Hence $\boldsymbol{A}_{\bar{S},\bar{S}}$ is nonsingular outside a proper algebraic subset. If $|\bar{S}| = 1$, then the nullity is one. If $|\bar{S}| > 2$ is odd, skew-symmetry implies $\operatorname{rank}(\boldsymbol{A}_{\bar{S},\bar{S}}) \leq |\bar{S}| - 1$. The condition $\operatorname{rank}(\boldsymbol{A}_{\bar{S},\bar{S}}) < |\bar{S}| - 1$ is again a

proper algebraic subset. Since there are only finitely many nonempty $\bar{S} \subseteq [K]$, $G^{\complement}$ has Lebesgue measure zero.

Now fix $\boldsymbol{A} \in \mathcal{G}$. The Nash set is nonempty by the minimax theorem. From Proposition K.1, we have

$$\boldsymbol{A}_{\mathcal{S}(\boldsymbol{A}),\mathcal{S}(\boldsymbol{A})}\boldsymbol{q}_{\mathcal{S}(\boldsymbol{A})} = \mathbf{0}.$$

Hence, $|\mathcal{S}(\boldsymbol{A})|$ is odd, $\mathrm{nullity}(\boldsymbol{A}_{S(\boldsymbol{A}),S(\boldsymbol{A})}) = 1$, and $\dim \mathcal{SN}(\boldsymbol{A}) = 0$. Since the Nash set is nonempty and convex, it is a singleton. Its unique point belongs to the relative interior of the singleton and is therefore quasi-strict by Proposition K.1. □

Proposition K.2 (Nash set under flat dominance)**.** *Suppose Assumption 1 holds for a partition $[K] = \mathcal{I} \dot{\cup} \mathcal{L}$. Then $\mathcal{SN}(\boldsymbol{A}) = \Delta_{\mathcal{I}}$, and in particular, $S(\boldsymbol{A}) = \mathcal{I}$.*

Proof. First take any $\boldsymbol{q} \in \Delta_{\mathcal{I}}$. Since $\boldsymbol{A}_{\mathcal{I},\mathcal{I}} = 0$,

$$(\boldsymbol{A}\boldsymbol{q})_{\mathcal{I}} = 0, \qquad (\boldsymbol{A}\boldsymbol{q})_{\mathcal{L}} = \boldsymbol{A}_{\mathcal{L},\mathcal{I}}\boldsymbol{q}_{\mathcal{I}} \geq b_{-}\mathbf{1} > 0.$$

Hence $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$, so $\Delta_{\mathcal{I}} \subseteq \mathcal{SN}(\boldsymbol{A})$.

Conversely, let $\boldsymbol{q} \in \mathcal{SN}(\boldsymbol{A})$ and choose any $\boldsymbol{c} \in \mathrm{ri}\,\Delta_{\mathcal{I}}$. From the first part, $\boldsymbol{c} \in \mathcal{SN}(\boldsymbol{A})$ and

$$(\boldsymbol{A}\boldsymbol{c})_j > 0, \qquad j \in \mathcal{L}.$$

Since $\boldsymbol{q}^{\top}\boldsymbol{A}\boldsymbol{c} = 0$, $q_j = 0$ for every $j \in \mathcal{L}$, thus $\boldsymbol{q} \in \Delta_{\mathcal{I}}$. Therefore $\mathcal{SN}(\boldsymbol{A}) = \Delta_{\mathcal{I}}$. □

L Additional Experiments

L.1 Numerical Experiments for Non-Quasi-Strict Equilibria

We numerically illustrate the asymptotic rates of Theorem G.4 on a family with $|\mathcal{B}| = 1$. To integrate over long physical times we use the logarithmic reparametrization $\sigma = \log(1+t)$, equivalently $t = e^{\sigma} - 1$; the trajectory in logit space is unchanged.

Test-game family. For $K \geq 3$, we consider

$$A^{(K)}_{12} = A^{(K)}_{21} = 0, \qquad A^{(K)}_{1j} = -1, \qquad A^{(K)}_{2j} = 1, \qquad j = 3, \dots, K,$$

with the remaining entries determined by skew-symmetry and $A^{(K)}_{ij} = 0$ for $i, j \geq 3$. Equivalently,

$$\boldsymbol{A}^{(K)} = \begin{pmatrix} 0 & 0 & -\mathbf{1}^{\top}_{K-2} \\ 0 & 0 & \mathbf{1}^{\top}_{K-2} \\ \mathbf{1}_{K-2} & -\mathbf{1}_{K-2} & \mathbf{0}_{(K-2)\times(K-2)} \end{pmatrix}.$$

For $\boldsymbol{q} = (q_1, \dots, q_K)$,

$$\boldsymbol{A}^{(K)}\boldsymbol{q} = \begin{pmatrix} -r \\ r \\ q_1 - q_2 \\ \vdots \\ q_1 - q_2 \end{pmatrix}, \qquad r := \sum_{j=3}^{K} q_j.$$

Therefore

$$\mathcal{SN}(\boldsymbol{A}^{(K)}) = \left\{(x, 1-x, 0, \dots, 0) : \tfrac{1}{2} \leq x \leq 1\right\}.$$

At the target $\boldsymbol{q}^{\star} = \boldsymbol{e}_1$, one has $\boldsymbol{A}^{(K)}\boldsymbol{q}^{\star} = (0, 0, 1, \dots, 1)^{\top}$, so (2) specializes to

$$\mathcal{S} = \{1\}, \qquad \mathcal{B} = \{2\}, \qquad \mathcal{M} = \{3, \dots, K\}.$$

We take $K \in \{3, 4, 5\}$.

Prediction. Theorem G.4 gives

$$p_2(t) = \Theta\left(t^{-(K-2)/(K-1)}\right), \qquad p_j(t) = \Theta(t^{-1}) \quad (j \in \mathcal{M}).$$

Figures 8 to 10 plot $\log p_j$ against $\log t$ and $-\mathrm{d}\log p_j/\mathrm{d}\log t$. The predicted limiting values of the latter are $(K-2)/(K-1)$ on $\mathcal{B}$ and 1 on $\mathcal{M}$.

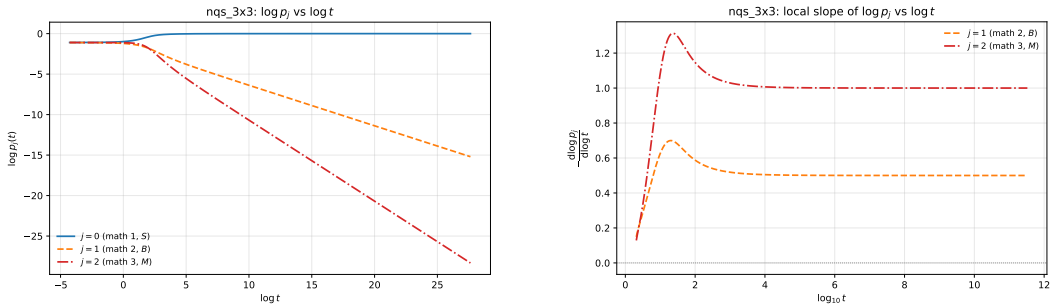


Figure 8: $K = 3$. Left: $\log p_j$ against $\log t$. Right: $-\mathrm{d}\log p_j/\mathrm{d}\log t$. Predicted limits are $1/2$ on $\mathcal{B}$ and 1 on $\mathcal{M}$.

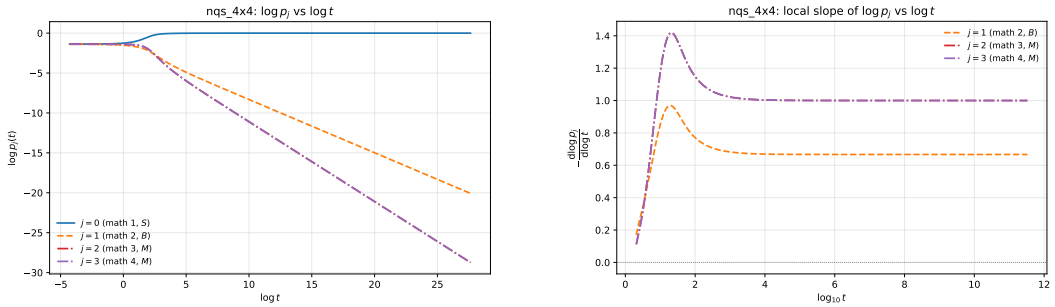


Figure 9: $K = 4$. Left: $\log p_j$ against $\log t$. Right: $-\mathrm{d} \log p_j / \mathrm{d} \log t$. Predicted limits are $2/3$ on $\mathcal{B}$ and 1 on $\mathcal{M}$.

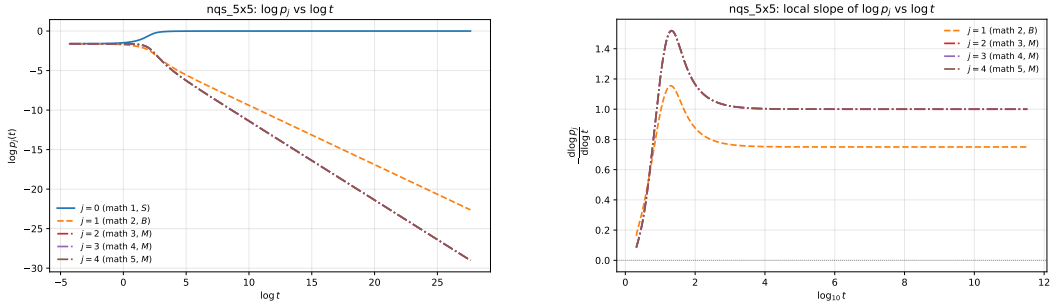


Figure 10: $K = 5$. Left: $\log p_j$ against $\log t$. Right: $-\mathrm{d} \log p_j / \mathrm{d} \log t$. Predicted limits are $3/4$ on $\mathcal{B}$ and 1 on $\mathcal{M}$.

L.2 Numerical Experiments for Nonconvergence Results

We provide numerical evidence for Propositions J.1 and J.2. We set $\rho_0 = 10^{-6}$ in Equation (90) and run Equation (S2P-EG-Flow) for $\gamma = 0.1$. We use the adaptive DOP853 solver implemented in `scipy.integrate.solve_ivp`, with physical time over $t \in [0, 10^{14}]$, relative tolerance 10^{-10}, and absolute tolerance 10^{-12}. For plotting, we evaluate the numerical solution at N_{plot} output times across the intervals $[0, 10^6]$ and $[10^6, 10^{14}]$. Within each interval, the output times are uniformly spaced in $\log(1 + t)$, and the point counts are allocated proportionally to the intervals' logarithmic widths.

When $\eta = 1/20$, the trajectory indeed converges in the Nash sense does not in the policy space. When η is increased to $1/5$, the trajectory does not even converge in the Nash sense (Figure 11). In both cases, the trajectory approaches the face $p_3 = p_4 = 0$.

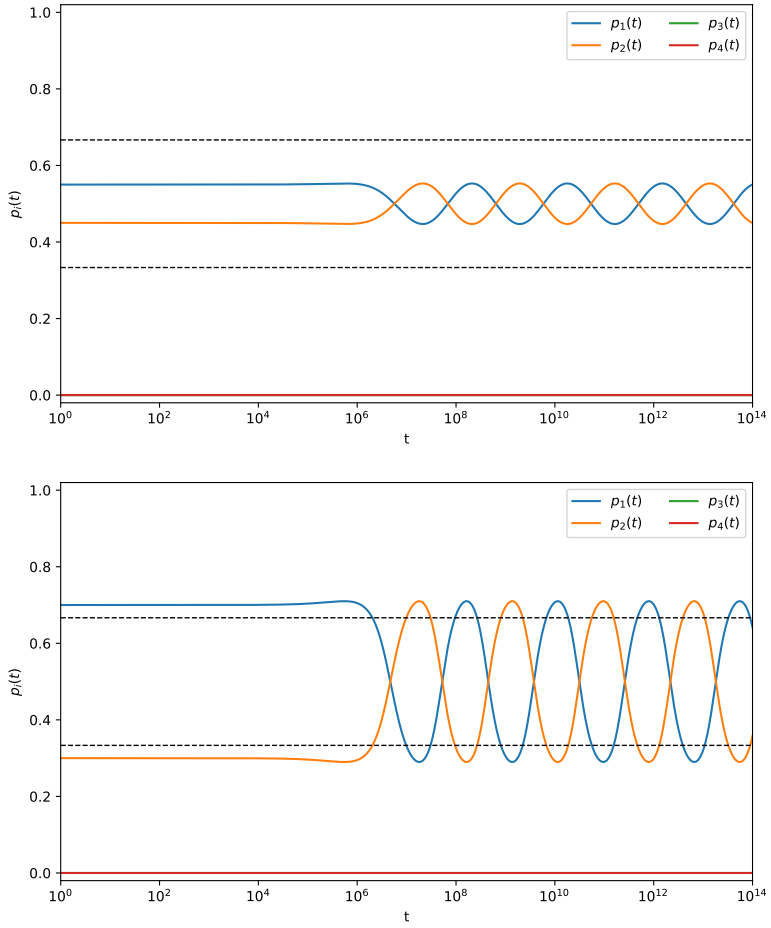


Figure 11: Evolution of $p(t)$ under the modified flow for the game (89) with $\rho_0 = 10^{-6}$. The interval for p_1 and p_2 to be in the Nash set is surrounded by the black lines. Top: $\eta = 1/20$. Bottom: $\eta = 1/5$.

In addition, we numerically test the trajectory outside of regime, by setting $\rho_0 = 0.4$ in Equation (90) and leaving every other setting intact. The same qualitative behavior happens: both trajectories fail to converge in policy space, where the convergence in the Nash sense depends on the value η.

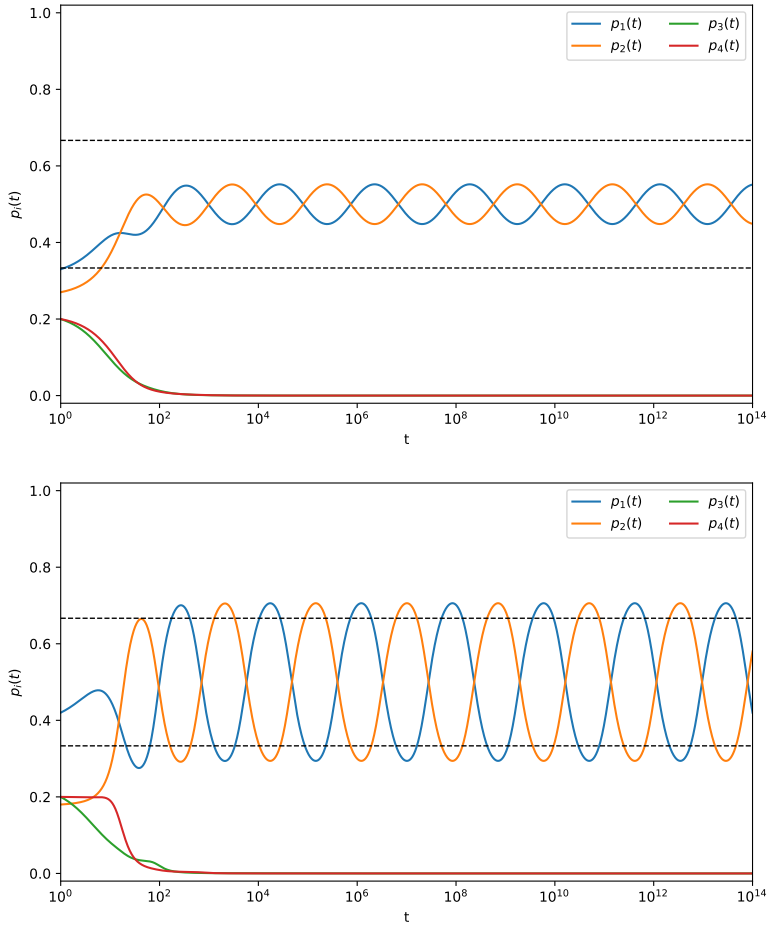


Figure 12: Evolution of $p(t)$ under the modified flow for the game (89) with $\rho_0 = 0.4$. The interval for p_1 and p_2 to be in the Nash set is surrounded by the black lines. Top: $\eta = 1/20$. Bottom: $\eta = 1/5$.