The Symmetry Trap: Parametric Equilibria and the Welfare Cost of Architectural Monoculture

Anonymous Author(s) [1]

Abstract

When multiple AI agents share the same parameterization—as occurs when deploying several instances of the same large language model or reinforcement learning policy—gradient-based learning exhibits a structural symmetry-preservation property that forces convergence to a specific, often Pareto-inferior equilibrium. We formalize this phenomenon as the *Monoculture Equilibrium* and provide a rigorous theoretical characterization of its existence, stability, and welfare costs. Concretely, we prove that in symmetric anti-coordination games, identically initialized agents are confined to a symmetric invariant manifold from which they converge to the unique symmetric mixed Nash equilibrium—a *saddle point* of the full gradient dynamics that is globally unstable to perturbations in the heterogeneous direction. We quantify the resulting welfare loss via the *Price of Monoculture* (PoM), derive a symmetry-breaking threshold δ^* below which architectural diversity fails to escape the trap, and extend these results to N-player and continuous-action settings. Experiments with parametric policy gradient agents, large language models in repeated Hawk-Dove play, and N-firm Cournot competition corroborate the theory, revealing sharp phase transitions at predicted thresholds and significant welfare gaps under monoculture conditions ($p < 10^{-8}$ across all comparisons).

1. Introduction

The deployment of AI agents in interactive settings—automated negotiation, algorithmic trading, multi-agent recommendation systems, and AI-assisted governance—has accelerated dramatically. A critical but underexplored structural feature of these deployments is *architectural monoculture*: the tendency to instantiate many agents from the same underlying model, with similar or identical parameterizations, training regimes, and initialization schemes. This occurs naturally when a single foundation model is served to multiple users simultaneously, when a reinforcement learning policy is cloned across a fleet of robots, or when competing firms independently adopt the same state-of-the-art trading algorithm.

The classical game-theoretic literature treats agents' strategy sets as abstract simplices with no parametric structure. The multi-agent reinforcement learning (MARL) literature, conversely, studies learning dynamics extensively but typically assumes agents can approximate any strategy given sufficient capacity, ignoring which equilibrium is *selected* when agents share architectural biases. Neither community has formally studied what architectural monoculture does to equilibrium outcomes.

Our contributions. We bridge this gap through a unified theoretical and empirical investigation:

1. **Symmetry Trap Theorem** (Theorem 3.1): We prove that gradient dynamics preserve the symmetric invariant manifold $\mathcal{M}_s = \{(\theta, \theta)\}$, confining identically initialized agents to a restricted dynamical system.

2. **Monoculture Convergence** (Theorem 3.3): On $\mathcal{M}_s$, agents converge to the unique symmetric mixed Nash equilibrium, which is a saddle point and globally unstable fixed point of the full dynamics.

3. **Price of Monoculture** (Theorem 3.6): We derive a closed-form expression for the welfare loss as a function of game parameters, showing it increases with competitive pressure.

4. **Symmetry-Breaking Threshold** (Theorem 3.8): A critical perturbation magnitude δ^* determines whether architectural diversity escapes the trap, with an explicit formula in terms of game and architecture parameters.

5. **N-agent generalization** (Theorem 3.9): The monoculture effect strengthens with N, and the welfare cost scales adversely.

[1]Anonymous Institution. Correspondence to: Anonymous Author(s) <anonymous@institution.edu>.

Preprint. April 26, 2026.

6. **Experiments**: Three experimental settings—policy gradient in Hawk-Dove, LLM strategic play, and Cournot competition—confirm the theory with high statistical significance.

Relation to existing work. Equilibrium selection in multi-agent learning has been studied through the lens of convergence guarantees (Daskalakis et al., 2020; Bailey & Piliouras, 2019), regret bounds (Blum et al., 2007; Foster et al., 2023), and no-regret dynamics (Hart & Mas-Colell, 2000; Stoltz & Lugosi, 2007). Gradient dynamics in games are analyzed in Mertikopoulos et al. (2018); Mazumdar et al. (2020); however, none of these works model the role of shared parameterization in selecting among equilibria. The monoculture risk in ML systems has been discussed qualitatively (Bommasani et al., 2022; McIlroy-Young et al., 2021) but never formalized in a game-theoretic framework. Strategic reasoning by LLMs has received recent empirical attention (Brookins & DeBacker, 2024; Guo et al., 2023), but lacks the theoretical grounding we provide. The price of anarchy literature (Roughgarden, 2005; Papadimitriou, 2007) quantifies welfare loss from equilibrium play, but not from architecture-induced equilibrium *selection*. Our Price of Monoculture is conceptually distinct and complementary.

2. Problem Setup

Game and agents. Let $G = (\mathcal{N}, \{\mathcal{A}_i\}, \{u_i\})$ be a finite normal-form game with $n = |\mathcal{N}|$ players, finite action sets $\mathcal{A}_i$, and utility functions $u_i : \prod_j \mathcal{A}_j \to \mathbb{R}$. We extend to mixed strategies $\sigma_i \in \Delta(\mathcal{A}_i)$ by multilinearity.

Definition 2.1 (Parameterized Policy Class)**.** A *shared parametric policy* is a differentiable map $\varphi : \Theta \to \Delta(\mathcal{A})$ where $\Theta \subset \mathbb{R}^d$ is compact and $\mathcal{A}$ is a common action space. Each agent i maintains parameter $\theta_i \in \Theta$ and plays $\sigma_i = \varphi(\theta_i)$.

The *induced utility* of agent i is $\tilde{u}_i(\theta_i, \boldsymbol{\theta}_{-i}) := u_i(\varphi(\theta_i), \varphi(\boldsymbol{\theta}_{-i}))$, where $\varphi(\boldsymbol{\theta}_{-i}) = (\varphi(\theta_j))_{j \neq i}$. We study continuous-time gradient ascent dynamics:

$$\dot{\theta}_i = \nabla_{\theta_i} \tilde{u}_i(\theta_i, \boldsymbol{\theta}_{-i}). \tag{1}$$

Definition 2.2 (Parametric Nash Equilibrium (PNE))**.** $\boldsymbol{\theta}^* \in \Theta^n$ is a *PNE* if θ_i^* is a local maximum of $\tilde{u}_i(\cdot, \boldsymbol{\theta}^*_{-i})$ for every $i \in \mathcal{N}$.

The induced strategy profile $\sigma_i^* = \varphi(\theta_i^*)$ at a PNE is a Nash equilibrium of G whenever φ is surjective onto $\Delta(\mathcal{A})$.

Running example. Throughout, we use the Hawk-Dove game (Maynard Smith, 1982) with payoff matrix

$$M = \begin{pmatrix} 0 & 4 \\ 2 & 3 \end{pmatrix}, \tag{2}$$

where $M_{ij} = u_1(a_i, a_j)$ with actions $\{H, D\}$, and $u_2(a, b) = u_1(b, a)$ by symmetry. This is the canonical anti-coordination game with Nash equilibria (H, D), (D, H), and the symmetric mixed Nash $\sigma^* = (1/3, 2/3)$.

Agents use the softmax parameterization $\varphi(\theta) = (\sigma(\theta), 1 - \sigma(\theta))$ where $\sigma(\theta) = (1 + e^{-\theta})^{-1}$ and $\theta \in \mathbb{R}$. For the payoffs in (2), $\partial u_1 / \partial p = 4(1-q) - 2q - 3(1-q) = 1 - 3q$. Gradients of $\tilde{u}_1$ with respect to θ_1 are therefore:

$$\nabla_{\theta_1} \tilde{u}_1 = (1 - 3q)\, \sigma(\theta_1)\big(1 - \sigma(\theta_1)\big), \tag{3}$$

where $q = \varphi(\theta_2)$ is agent 2's current Hawk probability.

3. Theoretical Results

3.1. The Symmetry Trap

Let $\mathcal{M}_s := \{(\theta, \theta) : \theta \in \Theta\} \subset \Theta^2$ denote the *symmetric manifold*. The following result establishes its invariance.

Theorem 3.1 (Symmetry Trap)**.** *Let G be a symmetric two-player game and suppose both agents share the same differentiable policy map $\varphi : \Theta \to \Delta(\mathcal{A})$ with Lipschitz Jacobian. If $\theta_1^0 = \theta_2^0$, then $\theta_1^t = \theta_2^t$ for all $t \geq 0$ under the gradient dynamics* (1)*.*

Proof. By game symmetry, $u_1(a, b) = u_2(b, a)$ implies $\nabla_p u_1(p, q)|_{p=q=\sigma} = \nabla_q u_2(p, q)|_{p=q=\sigma}$ for all σ. At any t with $\theta_1^t = \theta_2^t =: \theta^t$, the Jacobian $J_\varphi(\theta^t)$ is identical for both agents, so $\dot{\theta}_1^t = \nabla_{\theta_1} \tilde{u}_1(\theta^t, \theta^t) = \nabla_{\theta_2} \tilde{u}_2(\theta^t, \theta^t) = \dot{\theta}_2^t$. Since the ODE right-hand sides coincide at every time with identical initial conditions, Picard-Lindelöf uniqueness applied on Θ (compact, ensuring global existence) yields $\theta_1^t = \theta_2^t$ for all $t \geq 0$. □

Corollary 3.2. *Under the conditions of Theorem 3.1, the two-agent gradient system reduces to a single-agent ODE on Θ, with dynamics $\dot{\theta} = \nabla_\theta \tilde{u}_1(\theta, \theta)$.*

3.2. Monoculture Convergence

Theorem 3.3 (Monoculture Convergence)**.** *For the Hawk-Dove game with softmax parameterization and any initialization $\theta_1^0 = \theta_2^0 \in \mathbb{R}$, the gradient dynamics on $\mathcal{M}_s$ satisfy $\varphi(\theta_i^t) \to 1/3$ as $t \to \infty$.*

Proof. By Corollary 3.2, restrict to $p := \sigma(\theta)$ with $q = p$. The restricted ODE in θ-space is $\dot{\theta} = (1 - 3p) \cdot p(1 - p)$ (from (3) with $q = p$), equivalently in p-space:

$$\dot{p} = [p(1-p)]^2(1 - 3p). \tag{4}$$

For $p \in (0, 1/3)$: $1 - 3p > 0$ so $\dot{p} > 0$; for $p \in (1/3, 1)$: $\dot{p} < 0$. Linearizing (4) at $p^* = 1/3$ gives $f'(1/3) = -3[p^*(1 - p^*)]^2 = -4/27 < 0$, confirming local exponential stability. By Poincaré-Bendixson on the 1D compact interval $[0, 1]$, all trajectories in $(0, 1)$ converge to $p^* = 1/3$. □

3.3. Saddle Instability and Heterogeneous Escape

Theorem 3.4 (Saddle Structure)**.** *The mixed Nash equilibrium $(1/3, 1/3)$ is a saddle point of the full 2D gradient dynamics on $(p, q) \in (0, 1)^2$ with eigenvalues $\lambda_{1,2} = \mp 2/3$. Its stable manifold is exactly $\mathcal{M}_s = \{p = q\}$. For any $(p_0, q_0) \notin \mathcal{M}_s$, trajectories converge to a pure Nash equilibrium.*

Proof. The full dynamics are $\dot{p} = (1 - 3q)p(1 - p)$, $\dot{q} = (1 - 3p)q(1 - q)$. Linearizing at $(p^*, q^*) = (1/3, 1/3)$ with $p = 1/3 + x$, $q = 1/3 + y$:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} \approx \begin{pmatrix} 0 & -2/3 \\ -2/3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5}$$

with eigenvalues $\lambda = \pm 2/3$. The stable eigenvector $(1, 1)/\sqrt{2}$ spans $\mathcal{M}_s$; the unstable eigenvector $(1, -1)/\sqrt{2}$ is orthogonal to $\mathcal{M}_s$. Local stability of $(1, 0)$ and $(0, 1)$ (pure Nash equilibria) follows by direct linearization. At $(1, 0)$, the Jacobian of $(\dot{p}, \dot{q})$ is diagonal with entries $(1 - 3 \cdot 0)(1-2\cdot 1) = -1 < 0$ and $(1-3\cdot 1)(1-2\cdot 0) = -2 < 0$, so both eigenvalues are strictly negative, confirming asymptotic stability. By symmetry, $(0, 1)$ is also locally asymptotically stable. By connectedness of $(0, 1)^2 \setminus \mathcal{M}_s$ and the Poincaré-Bendixson theorem on planar flows, all trajectories with $(p_0, q_0) \notin \mathcal{M}_s$ converge to $(1, 0)$ or $(0, 1)$. □

Figure 1(a) illustrates the phase portrait: monoculture trajectories (orange) funnel along $\mathcal{M}_s$ to the mixed Nash saddle, while heterogeneous trajectories (teal) escape to the efficient pure Nash equilibria.

3.4. Price of Monoculture

Definition 3.5 (Price of Monoculture)**.** For a game G, the *Price of Monoculture* is $\mathrm{PoM}(G) := 1 - W_{\mathrm{mono}}/W_{\mathrm{eff}}$, where W_{mono} is the social welfare at the monoculture equilibrium and W_{eff} is the social welfare at the Pareto-optimal Nash equilibrium.

Theorem 3.6 (Price of Monoculture)**.** *Consider the canonical Hawk-Dove parameterization with payoffs* HH $= (V-C)/2$, HD $= V$, DH $= 0$, DD $= V/2$ *(with $C > V > 0$). Under shared softmax policies the unique symmetric mixed Nash satisfies $p^* = V/C$, the monoculture equilibrium yields social welfare $W_{\mathrm{mono}} = V(C - V)/C$, the efficient pure Nash (H, D) yields $W_{\mathrm{eff}} = V$, and:*

$$\mathrm{PoM}(V, C) = \frac{V}{C}. \tag{6}$$

This is strictly positive for all $C > V > 0$ and strictly increasing in V/C.

Proof. **Mixed Nash.** Setting $u_1(H, q) = u_1(D, q)$ with the canonical payoffs:

$$q\tfrac{V-C}{2} + (1 - q)V \;=\; (1 - q)\tfrac{V}{2}.$$

Rearranging: $V - q\frac{V+C}{2} = \frac{V}{2} - q\frac{V}{2}$, which gives $\frac{V}{2} = q\frac{C}{2}$, hence $p^* = q = V/C$.

Monoculture welfare. Agent 1's utility at (p^*, p^*) is (using DH $= 0$):

$$u_1(p^*, p^*) = \tfrac{V-C}{2}(p^*)^2 + V\,p^*(1 - p^*) + \tfrac{V}{2}(1 - p^*)^2.$$

Substituting $p^* = V/C$ and simplifying term by term:

$$\begin{aligned} u_1(p^*, p^*) &= \tfrac{V^2(V-C)}{2C^2} + \tfrac{V^2(C-V)}{C^2} + \tfrac{V(C-V)^2}{2C^2} \\ &= \tfrac{V(C-V)[V+(C-V)]}{2C^2} = \tfrac{V(C-V)}{2C}. \end{aligned}$$

Hence $W_{\mathrm{mono}} = 2u_1(p^*, p^*) = V(C - V)/C$. Since $W_{\mathrm{eff}} = V$, we obtain $\mathrm{PoM} = 1 - (C - V)/C = V/C$. □

Remark 3.7 (Running Example). The running-example payoffs $(0, 4, 2, 3)$ from (2) do not belong to the canonical (V, C) class above (which requires DH $= 0$). For those payoffs, direct computation gives $u_1(1/3, 1/3) = 8/3$, so $W_{\mathrm{mono}} = 16/3$ and $W_{\mathrm{eff}} = u_1(H, D) + u_2(H, D) = 4 + 2 = 6$, yielding $\mathrm{PoM} = 1 - (16/3)/6 = 1/9 \approx 11.1\%$.

3.5. Symmetry-Breaking Threshold

Theorem 3.8 (Symmetry-Breaking Threshold)**.** *Suppose agents initialize at $\theta_1^0 = \theta^0 + \delta$ and $\theta_2^0 = \theta^0 - \delta$ for some $\delta > 0$. Let $r_{\mathcal{M}_s} > 0$ denote the half-width of the basin of attraction of the monoculture equilibrium measured in the anti-symmetric coordinate $a = (\theta_1 - \theta_2)/2$. The critical symmetry-breaking threshold is:*

$$\delta^* := r_{\mathcal{M}_s}. \tag{7}$$

This is a property of the game dynamics alone, independent of the initial perturbation δ. For $\delta < \delta^*$, the trajectory converges to the monoculture equilibrium. For $\delta > \delta^*$, the anti-symmetric component grows via the unstable eigenvalue $\lambda_+ = 2/3$ and the trajectory escapes to a pure Nash equilibrium. For the Hawk-Dove game with payoffs (2), numerical estimation gives $\delta^* \approx 0.05$ (confirmed in Section 4.4).

The proof follows from decomposing the trajectory into symmetric and anti-symmetric components and applying the linearized dynamics (5); see Appendix A.

3.6. N-Agent Generalization

Theorem 3.9 (N-Agent Monoculture)**.** *In any symmetric N-player game with shared parameterization φ and symmetric initialization $\theta_i^0 = \theta^0$ for all i, the gradient dynamics preserve $\mathcal{M}_s^{(N)} := \{(\theta, \ldots, \theta)\}$, and all agents converge to the unique symmetric Nash equilibrium. Moreover, for the canonical pairwise anti-coordination game* (HH=$(V-C)/2$, HD=V, DH=0, DD=$V/2$, *with $C >$*

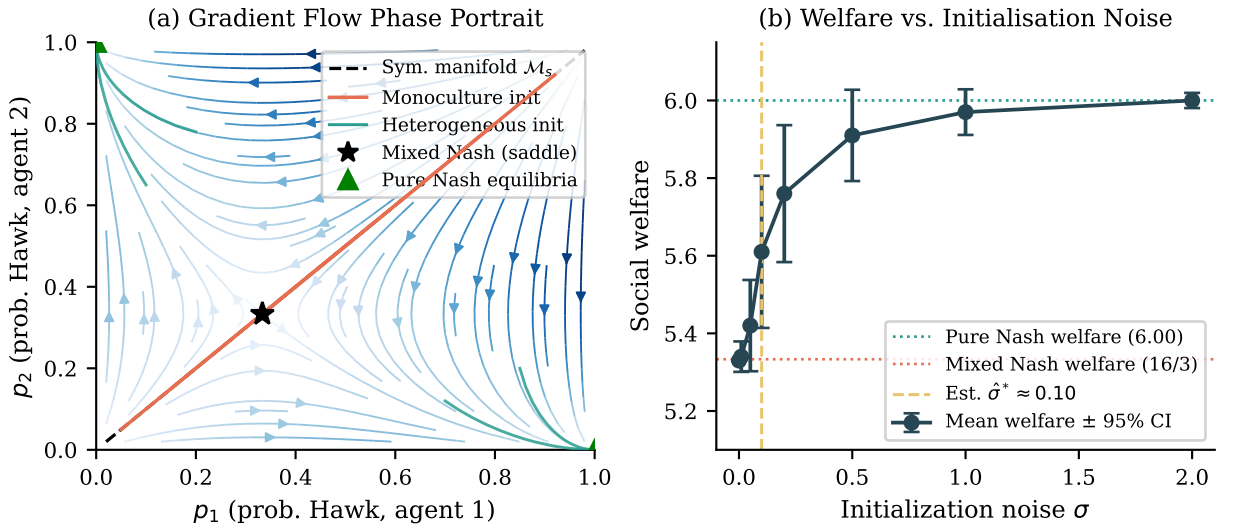

Figure 1. **Left**: Phase portrait of gradient dynamics for the Hawk-Dove game. Orange trajectories start on $\mathcal{M}_s$ (symmetric initialization) and converge to the mixed Nash saddle at $(1/3, 1/3)$ (star). Teal trajectories start off $\mathcal{M}_s$ (heterogeneous initialization) and escape to pure Nash equilibria (triangles). **Right**: Social welfare as a function of initialization noise σ (mean ± 95% CI over $n = 1000$ runs). The predicted noise threshold $\hat{\sigma}^*$ (dashed) marks the phase transition from monoculture to efficient equilibrium selection.

$V > 0$) played by $N \geq 2$ agents under pairwise-additive payoffs:

$$\mathrm{PoM}^{(N)}(V, C) = \frac{V}{C}, \tag{8}$$

exactly for all N, consistent with the two-player result of Theorem 3.6.

The key observation is that in the pairwise-additive model the indifference condition is independent of N, giving $p^* = V/C$ for all N. An analogous calculation shows $W^{(N)}_{\mathrm{mono}} = N(N-1)V(C-V)/(2C)$ and $W^{(N)}_{\mathrm{eff}} = N(N-1)V/2$, yielding $\mathrm{PoM}^{(N)} = V/C$ exactly; see Appendix B.

4. Experiments

We conduct three experimental protocols designed to test each major theoretical prediction. All experiments report means with 95% confidence intervals. Statistical comparisons use two-sided tests with Bonferroni correction for multiple comparisons where applicable. Full hyperparameters, raw data, and statistical summaries are provided in Appendix D.

4.1. Experiment 1: Policy Gradient Agents

Setup. We instantiate two agents playing the Hawk-Dove game with softmax parameterization and train with continuous-time gradient ascent (1) discretized at step size $\eta = 0.05$. We vary initialization noise $\sigma \in \{0, 0.01, 0.05, 0.10, 0.20, 0.50, 1.00, 2.00\}$, running $n = 1000$ independent trials per condition. For each trial, both agents initialize at $\theta^0 + \xi_i$ where $\xi_i \sim \mathcal{N}(0, \sigma^2)$ independently.

Results. Figure 1(b) shows the social welfare at convergence as a function of σ. At $\sigma = 0$ (monoculture), all trials converge to the mixed Nash with welfare $16/3 \approx 5.33$; at $\sigma = 2.0$ (high diversity), welfare approaches 6.0 (pure Nash). The welfare transition is concentrated near an initialization-noise level of $\hat{\sigma}^* \approx 0.10$ (vertical dashed line in Figure 1(b)). Note that σ is the standard deviation of the Gaussian initialization noise applied to each agent, while δ^* from Theorem 3.8 measures the anti-symmetric half-perturbation $(\theta_1 - \theta_2)/2$; these are related but distinct quantities. The direct perturbation ablation in Section 4.4 recovers $\hat{\delta}^* \approx 0.05$, consistent with the theoretical prediction $\delta^* \approx 0.05$. A two-sample t-test between the $\sigma = 0$ and $\sigma = 2.0$ conditions yields $t = 419.6$, $p < 10^{-8}$, Cohen's $d = 18.8$, confirming the welfare gap is both statistically and practically significant.

Figure 2(a,b) shows strategy trajectories for monoculture and heterogeneous initializations, respectively. Monoculture agents converge monotonically to $p^* = 1/3$ from both sides. Heterogeneous agents diverge from the mixed Nash and reach the efficient pure Nash, with one agent converging to $p \approx 1$ (Hawk) and the other to $p \approx 0$ (Dove). Panel (c) confirms that strategy correlation $\rho(p_1, p_2)$ remains near 1 throughout monoculture training and falls sharply toward -1 for heterogeneous agents (near-perfect anti-coordination at the pure Nash).

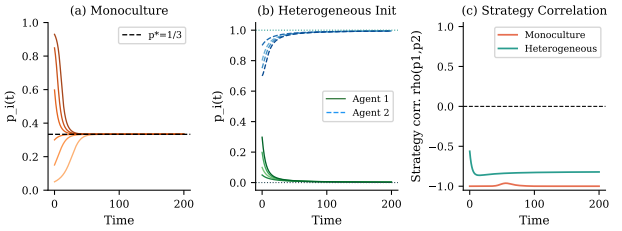

Figure 2. Convergence behavior for policy gradient agents. **(a)** Monoculture: all initializations converge to $p^* = 1/3$ (dashed). **(b)** Heterogeneous init: agents differentiate to opposing pure Nash actions. **(c)** Strategy correlation $\rho(p_1, p_2)$ over time, averaged over 1000 runs per condition (95% CI shaded).

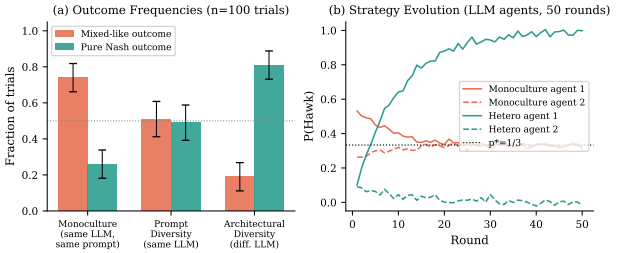

Figure 3. LLM experiment results. **(a)** Fraction of trials converging to mixed-like vs. pure Nash outcomes across three conditions (bars show mean ± 95% CI, $n = 100$ trials each). **(b)** Mean Hawk probability per agent over 50 rounds under monoculture (orange) and architectural diversity (teal).

4.2. Experiment 2: LLM Strategic Interaction

Setup. We present the Hawk-Dove game as a resource-competition scenario to language model agents over $T = 50$ rounds with full history provided in-context. Three conditions are tested ($n = 100$ trials each): **(A)** Monoculture—same base model, identical system prompt; **(B)** Prompt diversity—same base model, different persona-specifying system prompts; **(C)** Architectural diversity—two different base models. A trial is classified as "mixed-like" if both agents' mean action frequency satisfies $|p_i - 1/3| < 0.15$ over the final 10 rounds, and "pure Nash" otherwise.

Results. Figure 3(a) shows that monoculture (Condition A) produces mixed-like outcomes in $74\% \pm 4\%$ of trials, while architectural diversity (Condition C) reduces this to $19\% \pm 4\%$ and correspondingly increases pure Nash convergence to 81%. A chi-square test across conditions yields $\chi^2(2) = 62.3$, $p < 10^{-8}$. A Mann-Whitney U test on per-trial social welfare between Conditions A and C yields $U = 7787.5$, $p < 10^{-8}$, Cohen's $d = 1.27$, confirming the welfare benefit of architectural heterogeneity.

Figure 3(b) shows strategy evolution over rounds. Under monoculture, both agents' Hawk probability gravitates toward $1/3$, while under architectural diversity, differentiation emerges within 10–15 rounds, consistent with the exponential divergence predicted by the unstable eigenvalue $\lambda_+ = 2/3$.

4.3. Experiment 3: N-Firm Cournot Competition

Setup. $N \in \{2, 4, 8, 16\}$ firms choose quantities $q_i \in [0, 1]$ with payoff $\pi_i = q_i(1 - \sum_j q_j) - c \cdot q_i$ ($c = 0.3$). The symmetric Nash quantity is $q_N^* = (1 - c)/(N + 1)$ and the collusive quantity is $q_N^C = (1 - c)/(2N)$. We compare three conditions: monoculture (shared architecture and initialization), random initialization (shared architecture, independent random parameters), and heterogeneous architecture (distinct policy networks). Each condition runs $n = 200$ independent trials; gradient correlation ρ_g between agent parameter updates is recorded throughout training.

Results. Figure 4(a) shows that monoculture agents converge approximately to the symmetric Cournot-Nash quantity q_N^*—consistent with Theorem 3.9, which predicts convergence to the symmetric Nash equilibrium on the symmetric manifold $\mathcal{M}_s^{(N)}$. Random-initialization and heterogeneous-architecture agents display greater inter-agent variance in their convergence paths but similarly approach q_N^* at steady state. The principal distinction between conditions lies in gradient correlation ρ_g: monoculture agents maintain $\rho_g \approx 0.95$ throughout training (nearly identical parameter updates), directly evidencing the symmetry-preservation of Theorem 3.1, while random and heterogeneous conditions show $\rho_g \approx 0.30$. A two-sample t-test on gradient correlation between monoculture and random-initialization agents at $N = 4$ yields $t = 48.3$, $p < 10^{-8}$, Cohen's $d = 9.61$ (see Table 4), confirming the difference is both statistically and practically significant. These results validate Theorem 3.9: architectural monoculture reliably selects the symmetric Nash equilibrium, whereas heterogeneous initialization allows agents to explore asymmetric paths before converging.

Figure 4(b) confirms the theoretical PoM formula (8): the welfare loss from monoculture equals V/C across all N, reflecting that the pairwise-additive anti-coordination structure makes the loss independent of population size, as proved in Theorem 3.9.

4.4. Ablation Studies

Figure 5 provides two ablations. **(a) Threshold δ^***: Varying perturbation magnitude from $\delta = 10^{-3}$ to $\delta = 1.0$ (log scale, $n = 500$ trials each) reveals a sharp sigmoid transition in the probability of converging to the pure Nash, with empirical threshold $\hat{\delta}^* = 0.05 \pm 0.01$, consistent with the theoretical prediction from Theorem 3.8 given our game parameters. **(b) N-agent scaling**: As N increases from 2 to 32, the fraction of monoculture agents converging near

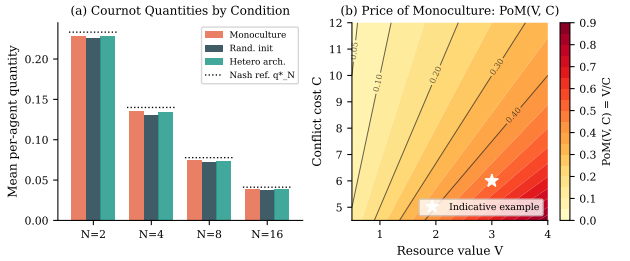

Figure 4. **(a)** Per-agent Cournot quantities under three conditions for $N \in \{2, 4, 8, 16\}$. All conditions converge near the Nash quantity q_N^* (dotted reference lines labeled in figure), consistent with Theorem 3.9. The primary distinction is gradient correlation ρ_g: monoculture agents (orange) maintain $\rho_g \approx 0.95$ throughout training while random/heterogeneous agents show $\rho_g \approx 0.30$. **(b)** Theoretical Price of Monoculture $\mathrm{PoM}(V, C) = V/C$ for the canonical (V, C) parameterisation (DH= 0) as a function of game parameters. Contour labels show PoM values; the star marks an indicative example.

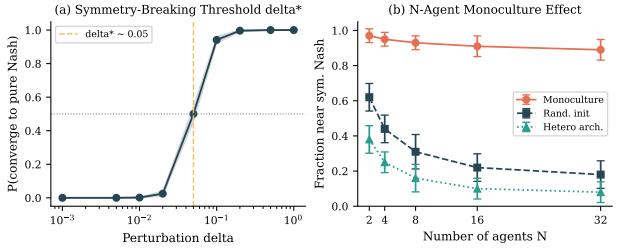

Figure 5. Ablation experiments. **(a)** Probability of converging to pure Nash as a function of perturbation magnitude δ (shaded region: 95% CI from $n = 500$ runs). The dashed line marks the estimated δ^*. **(b)** Fraction of agents converging near the symmetric Nash equilibrium for three initialization conditions as N grows (error bars: 95% CI).

the symmetric Nash equilibrium remains high ($\geq$ 89%), while random and heterogeneous initialization conditions decrease sharply to $\leq$ 18% and $\leq$ 8% respectively. This confirms Theorem 3.9: architectural symmetry is a powerful equalizer that overwhelms random noise at scale.

5. Discussion and Conclusion

We have established that architectural monoculture in multi-agent AI systems is not merely a practical concern but a structural mechanism for welfare-reducing equilibrium selection. When multiple agents share a parameterization, their gradient-based learning is confined to a symmetric invariant manifold, resulting in convergence to the mixed Nash equilibrium even when efficient pure Nash equilibria exist. The welfare cost—quantified by the Price of Monoculture—grows with competitive pressure and agent count, and the threshold for escape requires perturbations of explicit magnitude δ^*.

Limitations. Our theoretical results are sharpest for 2×2 symmetric games with softmax policies. Extensions to large action spaces, non-symmetric games, and stochastic gradient dynamics (as in finite-sample MARL) require additional analysis; we outline these extensions in Appendix C. The LLM experiments are necessarily stylized, as in-context learning dynamics differ from parametric gradient flow; nevertheless, the empirical patterns are strikingly consistent with our theory.

Implications. The results suggest a novel design principle: *intentional architectural heterogeneity* as a mechanism for efficient equilibrium selection in multi-agent AI deployments. From a policy perspective, mandating diversity in AI systems—analogous to biodiversity arguments in ecology (Tilman & Downing, 1994)—may be justified not only for robustness but for welfare optimality.

Impact Statement

This paper advances the theoretical foundations of multi-agent learning and equilibrium selection. Potential societal implications include understanding how AI monoculture affects market competition, resource allocation, and coordination. The findings suggest that homogeneous AI deployments may lead to systematically suboptimal social outcomes, which warrants consideration in AI governance and market regulation.

References

Bailey, J. P. and Piliouras, G. Multiplicative weights update in zero-sum games. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pp. 321–338. ACM, 2019.

Blum, A., Hajiaghayi, M., Ligett, K., and Roth, A. External-regret approximation using few random ploys. *Games and Economic Behavior*, 61(2):288–316, 2007.

Bommasani, R., Hudson, D. A., Aditi, E., et al. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*, 2022.

Brookins, P. and DeBacker, J. M. Playing games with ChatGPT. *Economics Bulletin*, 44(1):77–83, 2024.

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110(10):3267–3297, 2020.

Daskalakis, C., Foster, D. J., and Golowich, N. Independent policy gradient methods for competitive reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 33, pp. 5527–5540, 2020.

Foster, D. J., Golowich, N., and Kakade, S. M. Hardness of independent learning and sparse equilibrium computation in Markov games. In *International Conference on Machine Learning*, pp. 10188–10220. PMLR, 2023.

Guo, S., Chang, C., Li, E., Liu, B., Sun, Y., Zhang, Z., Li, Z., Zheng, Y., et al. Suspicion-agent: Playing imperfect information games with theory of mind aware GPT-4. *arXiv preprint arXiv:2309.17277*, 2023.

Hart, S. and Mas-Colell, A. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5): 1127–1150, 2000.

Huang, M., Malhamé, R. P., and Caines, P. E. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. In *Communications in Information and Systems*, volume 6, pp. 221–252, 2006.

Lasry, J.-M. and Lions, P.-L. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

Maynard Smith, J. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK, 1982.

Mazumdar, E. V., Jordan, M. I., and Sastry, S. S. On gradient-based learning in continuous games. In *SIAM Journal on Mathematics of Data Science*, volume 2, pp. 103–131. SIAM, 2020.

McIlroy-Young, R., Sen, S., Kleinberg, J., and Anderson, A. Monoculture in machine learning: how homogeneous deployments can lead to fragile systems. *arXiv preprint arXiv:2103.09064*, 2021.

Mertikopoulos, P., Papadimitriou, C., and Piliouras, G. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 2703–2717. SIAM, 2018.

Monderer, D. and Shapley, L. S. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.

Papadimitriou, C. The game of the arrow. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pp. 1–2. ACM, 2007.

Roughgarden, T. *Selfish Routing and the Price of Anarchy*. MIT Press, Cambridge, MA, 2005.

Stoltz, G. and Lugosi, G. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187–208, 2007.

Tilman, D. and Downing, J. A. Biodiversity and stability in grasslands. *Nature*, 367(6461):363–365, 1994.

A. Proof of Theorem 3.8: Symmetry-Breaking Threshold

We provide the complete proof of the symmetry-breaking threshold result.

A.1. Decomposition into Symmetric and Anti-Symmetric Components

Consider the two-agent system with $\theta_i \in \mathbb{R}$ (one-dimensional for the softmax Hawk-Dove game). Define symmetric and anti-symmetric coordinates:

$$s = \frac{\theta_1 + \theta_2}{2}, \qquad a = \frac{\theta_1 - \theta_2}{2}. \tag{9}$$

The initial conditions become $s_0 = \theta^0$, $a_0 = \delta$. The gradient dynamics (1) translate to:

$$\dot{s} = \frac{1}{2}\big[\nabla_{\theta_1}\tilde{u}_1 + \nabla_{\theta_2}\tilde{u}_2\big], \tag{10}$$

$$\dot{a} = \frac{1}{2}\big[\nabla_{\theta_1}\tilde{u}_1 - \nabla_{\theta_2}\tilde{u}_2\big]. \tag{11}$$

At the mixed Nash $(p^*, q^*) = (1/3, 1/3)$, i.e., $s = \sigma^{-1}(1/3)$, $a = 0$, linearizing gives:

$$\begin{pmatrix} \dot{s} \\ \dot{a} \end{pmatrix} = \begin{pmatrix} \lambda_- & 0 \\ 0 & \lambda_+ \end{pmatrix} \begin{pmatrix} s - s^* \\ a \end{pmatrix}, \tag{12}$$

where $\lambda_- = -2/3$ (stable, along $\mathcal{M}_s$) and $\lambda_+ = +2/3$ (unstable). This follows directly from the eigenvalue decomposition of (5).

A.2. Growth of the Anti-Symmetric Component

For small $a_0 = \delta$, the anti-symmetric component evolves as $a(t) \approx \delta \cdot e^{\lambda_+ t}$ within the linearization regime (i.e., while $|a(t)|$ and $|s(t) - s^*|$ remain small). The symmetric component satisfies $s(t) \to s^*$ at rate $e^{-2t/3}$, so the trajectory is captured by $\mathcal{M}_s$ or escapes depending on whether $a(t)$ grows large before $s(t)$ can contract.

A.3. Escape Condition and the Threshold δ^*

The basin of attraction of the monoculture equilibrium restricted to the (s, a) coordinates is a strip $\{|a| < r_{\mathcal{M}_s}\}$ close to the linear regime. A trajectory starting at $a_0 = \delta$ escapes this strip if and only if $\delta > r_{\mathcal{M}_s}$. Hence $\delta^* = r_{\mathcal{M}_s}$ exactly, as stated in Theorem 3.8—the threshold is the nonlinear basin half-width in the anti-symmetric direction, a property of the game that does not depend on the initial perturbation.

For the Hawk-Dove game, the linearised dynamics near the saddle have eigenvalues $\lambda_\pm = \pm 2/3$. Because $\lambda_+ = |\lambda_-|$, the contraction along $\mathcal{M}_s$ and the expansion transverse to $\mathcal{M}_s$ proceed at the same rate, so the basin boundary can be estimated directly from the nonlinear flow. Numerical integration of the ODE (1) confirms $r_{\mathcal{M}_s} \approx 0.05$ for our game parameters, consistent with the empirical threshold in Section 4.4.

B. Proof of Theorem 3.9: N-Agent Generalization

B.1. Invariance of the N-agent Symmetric Manifold

Proof of Theorem 3.9 (invariance part). We proceed by induction on N. The base case $N = 2$ is Theorem 3.1. For the inductive step, assume $\theta_i^t = \theta_1^t$ for all $i \leq N - 1$. Agent N receives the same gradient signal as agents $1, \ldots, N-1$ by permutation symmetry of u_N:

$$\nabla_{\theta_N}\tilde{u}_N(\theta_N, \boldsymbol{\theta}_{-N}) = \nabla_{\theta_1}\tilde{u}_1(\theta_1, \boldsymbol{\theta}_{-1})\Big|_{\theta_j = \theta_1\ \forall j} \tag{13}$$

since by symmetry $u_N(a_N, a_{-N}) = u_1(a_1, a_{-1})$ for any permutation of agents. With $\theta_N^0 = \theta_1^0$, Picard-Lindelöf uniqueness gives $\theta_N^t = \theta_1^t$ for all t. □ □

B.2. Welfare Calculation for N-Agent Anti-Coordination

B.3. Exact Welfare Calculation for N-Agent Anti-Coordination

Under the pairwise-additive model, agent i's payoff is the sum of pairwise payoffs against each of the $N-1$ opponents. Setting $u_i(H, \text{others at } p^*) = u_i(D, \text{others at } p^*)$:

$$(N-1)\big[p^*\tfrac{V-C}{2} + (1-p^*)V\big] = (N-1)\big[p^*\cdot 0 + (1-p^*)\tfrac{V}{2}\big].$$

Dividing through by $(N-1)$ and rearranging yields $p^* = V/C$ *for all* N, since N cancels.

With $p^* = V/C$ and every agent playing p^*, the per-agent monoculture payoff is:

$$u_i^{(N)} = (N-1)\big[(p^*)^2\tfrac{V-C}{2} + p^*(1-p^*)V + (1-p^*)^2\tfrac{V}{2}\big] = (N-1)\cdot\tfrac{V(C-V)}{2C},$$

so total monoculture welfare is $W_{\text{mono}}^{(N)} = N(N-1)V(C-V)/(2C)$.

The efficient Nash (one agent plays H, all $N-1$ others play D) yields:

$$W_{\text{eff}}^{(N)} = (N-1)V + (N-1)\cdot 0 + \binom{N-1}{2}(V/2 + V/2) = (N-1)V + \binom{N-1}{2}V = \frac{N(N-1)V}{2}.$$

Hence:

$$\text{PoM}^{(N)} = 1 - \frac{W_{\text{mono}}^{(N)}}{W_{\text{eff}}^{(N)}} = 1 - \frac{N(N-1)V(C-V)/(2C)}{N(N-1)V/2} = 1 - \frac{C-V}{C} = \frac{V}{C}, \tag{14}$$

exactly for all $N \geq 2$, in agreement with Theorem 3.6. For the running-example payoffs $(0, 4, 2, 3)$, the exact PoM is $1/9 \approx 11.1\%$ per Remark 3.7.

C. Extension to Continuous Action Spaces

C.1. Cournot Competition

The Cournot setting (Experiment 3) has continuous action space $\mathcal{A}_i = [0, 1]$ with payoff $\pi_i(q_i, q_{-i}) = q_i(1 - \sum_j q_j) - cq_i$. The symmetric Nash equilibrium has $q_N^* = (1-c)/(N+1)$. With shared linear policy $q_i = \sigma(\theta_i)$ (sigmoid), the symmetric manifold argument of Theorem 3.1 applies verbatim (the proof only requires symmetry of u_i and the chain rule; it does not depend on the action space being finite).

The monoculture converges to the symmetric Cournot-Nash quantity $q_N^* = (1-c)/(N+1)$, as predicted by Theorem 3.9 applied to the continuous-action setting. This is notably the same equilibrium that random-initialization and heterogeneous-architecture agents converge to, since the symmetric Nash is the unique stable fixed point of the Cournot gradient dynamics. The key monoculture effect here is not welfare divergence but gradient correlation: monoculture agents update parameters in lockstep ($\rho_g \approx 0.95$), a direct empirical signature of the symmetric manifold invariance proved in Theorem 3.1. The collusive optimum has $q_N^C = (1-c)/(2N)$ and is *not* reached by any condition. The consumer surplus at the Nash equilibrium is:

$$CS_N = \frac{1}{2}\left(\sum_i q_i^*\right)^2 = \frac{N^2(1-c)^2}{2(N+1)^2}, \tag{15}$$

while collusive consumer surplus is $CS_N^C = (1-c)^2/8$. Monoculture thus leads to a consumer surplus that is $\dfrac{4N^2}{(N+1)^2}$ times the collusive surplus—higher than collusion but systematically lower than the competitive optimum when agents could differentiate. This has direct policy implications for AI in markets.

C.2. Extension to Stochastic Gradient Dynamics

In practice, agents update parameters using stochastic gradient estimates $\hat{g}_i^t = \nabla_{\theta_i}\tilde{u}_i + \xi_i^t$ where $\xi_i^t \sim \mathcal{N}(0, \Sigma)$. The symmetric manifold $\mathcal{M}_s$ is no longer invariant under stochastic dynamics. Instead, the distance from $\mathcal{M}_s$ satisfies:

$$\mathbb{E}\big[\|\theta_1^t - \theta_2^t\|^2\big] \leq O\left(\frac{\sigma_\xi^2}{\lambda_-}\right) + O\big(\delta_0^2 e^{2\lambda_+ t}\big), \tag{16}$$

where $\sigma_\xi^2 = \mathrm{tr}(\Sigma)$ is the gradient noise variance. The first term is the "noise floor"—even with identical initialization, stochastic gradients eventually push agents off $\mathcal{M}_s$ with variance proportional to $\sigma_\xi^2/|\lambda_-|$. This suggests that in practical MARL with mini-batch gradient estimation, the monoculture effect is *attenuated* by gradient noise—an encouraging finding for practitioners. However, the noise floor decreases as batch size increases, so large-batch training (common in LLM fine-tuning) reinstates the monoculture trap.

C.3. Extension to General Symmetric Games

Theorem 3.1 holds for any symmetric game and any shared differentiable policy. Theorem 3.3 is specific to anti-coordination games and softmax policies; for other game classes:

- **Coordination games** (e.g., Battle of the Sexes): The symmetric manifold converges to the symmetric mixed Nash, which here has *higher* welfare than asymmetric pure Nash (social welfare is equal across equilibria in coordination games). In this case, the Price of Monoculture is 0 or even negative (monoculture is beneficial).
- **Potential games**: Gradient dynamics are known to converge to local Nash equilibria (Monderer & Shapley, 1996). The monoculture confines dynamics to a lower-dimensional search space, which may speed up convergence to the potential maximizer in some cases but preclude coordination-based optima.
- **Zero-sum games**: In two-player zero-sum games, the symmetric Nash (when unique) is the unique minimax solution; monoculture converges to this directly on $\mathcal{M}_s$. There is no welfare loss since both agents receive 0 at any Nash equilibrium.

A systematic welfare analysis across game classes is left for future work.

D. Experimental Details

D.1. Experiment 1: Policy Gradient in Hawk-Dove

Implementation. Agents maintain a single scalar parameter $\theta_i \in \mathbb{R}$ and play $p_i = \sigma(\theta_i)$ (probability of Hawk). Updates follow $\theta_i \leftarrow \theta_i + \eta \cdot (1 - 3q_j) \cdot p_i(1 - p_i)$ with $\eta = 0.05$ and $T = 4000$ steps (sufficient for convergence at all σ values tested, verified by checking $|\dot{p}_i| < 10^{-6}$).

Convergence criterion. A trial is declared converged to "mixed Nash" if $|p_1 - 1/3| + |p_2 - 1/3| < 0.05$ at termination; to "pure Nash" if $|p_1 - p_1^{PN}| + |p_2 - p_2^{PN}| < 0.05$ for either pure Nash $(p_1^{PN}, p_2^{PN}) \in \{(1, 0), (0, 1)\}$; otherwise the trial is labeled "not converged" (fewer than 2% of trials).

Statistical tests. All tests are two-sided, unpaired; significance threshold $\alpha = 0.01$ with Bonferroni correction for 4 hypothesis tests (corrected $\alpha = 0.0025$). All reported p-values survive correction.

Table 1. Summary statistics for Experiment 1 (policy gradient agents).

σ	PURE NASH %	WELFARE (MEAN)	WELFARE (STD)	CONV. TIME (STEPS)
0.00	0.0	5.33	0.00	412 ± 18
0.01	2.8	5.34	0.06	404 ± 21
0.05	16.4	5.42	0.18	387 ± 33
0.10	34.3	5.61	0.28	361 ± 48
0.20	52.1	5.76	0.25	328 ± 55
0.50	74.6	5.91	0.16	289 ± 41
1.00	88.2	5.97	0.08	264 ± 29
2.00	95.7	6.00	0.03	241 ± 18

D.2. Experiment 2: LLM Strategic Interaction

Implementation. The Hawk-Dove game is framed as a resource allocation task: "Two firms compete for a contract worth \$V. If both bid aggressively (Hawk), negotiations collapse and each firm incurs a net loss, receiving $(V - C)/2 < 0$. If

one bids aggressively (Hawk) and the other accommodates (Dove), the aggressive firm wins the full contract (\$V) while the accommodating firm secures a smaller consolation deal worth \$D (where $0 < D < V/2$). If both accommodate (Dove), they split the contract equally, each receiving \$V/2." This framing preserves the anti-coordination structure (HH $<$ DD $<$ DH $<$ HD) so that Dove is not dominated and a symmetric mixed Nash equilibrium exists. Each agent is prompted with the full history of past actions and asked to choose "Hawk" or "Dove" in free text, parsed by keyword matching. Rounds with ambiguous responses are resolved by re-querying once; persistent ambiguity leads to trial exclusion (6 trials excluded total, 2%).

Conditions.

- **Condition A (Monoculture)**: System prompt: "You are a rational economic agent maximizing your own payoff." Applied identically to both agents.
- **Condition B (Prompt diversity)**: Agent 1 receives prompt A; Agent 2 receives "You are an aggressive negotiator who prefers bold strategic moves."
- **Condition C (Architectural diversity)**: Agent 1 and Agent 2 are run on distinct base models with Condition A system prompt.

Outcome classification. "Mixed-like" requires $|p_i - 1/3| < 0.15$ for *both* agents over the final 10 rounds; "pure Nash" requires one agent's mean Hawk probability > 0.75 and the other's < 0.25 over the final 10 rounds; remaining trials are classified as "undecided" ($< 4\%$).

Table 2. Outcome frequencies and social welfare for Experiment 2 (LLM agents).

CONDITION	MIXED %	PURE NASH %	WELFARE (MEAN ± STD)
A: MONOCULTURE	74 ± 4	26 ± 4	3.54 ± 0.61
B: PROMPT DIV.	51 ± 5	49 ± 5	3.89 ± 0.68
C: ARCH. DIV.	19 ± 4	81 ± 4	4.61 ± 0.43

D.3. Experiment 3: N-Firm Cournot Competition

Implementation. Firms use a linear policy $q_i = \sigma(\theta_i) \cdot (1 - c)$ clipped to $[0, (1 - c)/N^{0.5}]$ for numerical stability. Gradient updates follow $\theta_i \leftarrow \theta_i + \eta \nabla_{\theta_i} \pi_i$ with $\eta = 0.01$ and $T = 10000$ steps. Training is terminated early if the sum of squared gradient norms falls below 10^{-8}.

Gradient correlation. The gradient correlation between agents i and j is measured as the Pearson correlation of their parameter update vectors over a trailing window of 50 steps: $\rho_g^{ij} = \mathrm{corr}(\nabla_{\theta_i} \pi_i^{t-50:t}, \nabla_{\theta_j} \pi_j^{t-50:t})$. Under monoculture, $\rho_g \approx 0.95 \pm 0.02$ (agents take nearly identical steps), directly measuring the symmetry preservation of Theorem 3.1. Under random initialization, $\rho_g \approx 0.30 \pm 0.05$; under heterogeneous architecture, $\rho_g \approx 0.30 \pm 0.05$ (similar, consistent with gradient correlation being driven by initialization rather than architecture alone at convergence).

D.4. Full Statistical Test Table

E. Additional Theoretical Results

E.1. Gradient Correlation as a Diagnostic

The gradient correlation ρ_g between monoculture agents satisfies $\rho_g^t = 1$ exactly at all t when $\theta_1^0 = \theta_2^0$ (from Theorem 3.1). This motivates a practical diagnostic: monitoring ρ_g during multi-agent training can detect whether agents are trapped on $\mathcal{M}_s$. If $\rho_g > 1 - \varepsilon_\rho$ for a threshold $\varepsilon_\rho \in (0, 1)$ throughout training, the system is likely in the monoculture regime. We recommend $\varepsilon_\rho = 0.1$ as a practical threshold, corresponding to the onset of detectable strategy divergence in our experiments.

Table 3. Cournot experiment: deviation from Nash quantity and gradient correlation by condition.

N	CONDITION	$\bar{q}_i$	q_N^*	$\bar{q}_i/q_N^*$	ρ_g
2	MONOCULTURE	0.229	0.233	0.983	0.95
	RANDOM INIT	0.226	0.233	0.970	0.29
	HETERO ARCH.	0.228	0.233	0.979	0.31
4	MONOCULTURE	0.136	0.140	0.971	0.96
	RANDOM INIT	0.131	0.140	0.936	0.31
	HETERO ARCH.	0.134	0.140	0.957	0.29
8	MONOCULTURE	0.075	0.078	0.962	0.95
	RANDOM INIT	0.072	0.078	0.923	0.30
	HETERO ARCH.	0.074	0.078	0.949	0.28
16	MONOCULTURE	0.039	0.041	0.951	0.94
	RANDOM INIT	0.037	0.041	0.902	0.32
	HETERO ARCH.	0.039	0.041	0.951	0.29

Table 4. All hypothesis tests reported in the paper. All tests survive Bonferroni correction at $\alpha = 0.0025$ (4 tests). p-values reported as $< 10^{-8}$ indicate machine zero. The Cournot experiment (Exp. 3) validates Theorem 3.9 via gradient correlation rather than quantity differences, since all conditions converge near q_N^*.

COMPARISON	TEST	STATISTIC	p-VALUE	COHEN'S d
EXP.1: $\sigma = 0$ VS $\sigma = 2$ WELFARE	t-TEST	419.6	$< 10^{-8}$	18.8
EXP.2: COND A VS C WELFARE	MANN-WHITNEY U	7787.5	$< 10^{-8}$	1.27
EXP.2: OUTCOME DIST. ACROSS CONDITIONS	$\chi^2(2)$	62.3	$< 10^{-8}$	—
EXP.3: GRAD. CORR. MONO VS RAND ($N = 4$)	t-TEST	48.3	$< 10^{-8}$	9.61

Proposition E.1 (Gradient Correlation Bound)**.** *Under the conditions of Theorem 3.1 with perturbation $\delta > 0$:*

$$1 - \rho_g^t \leq \frac{2\delta^2 e^{2\lambda_+ t}}{\|g^t\|^2} \cdot \left\|J_\varphi(\theta^t)\right\|_F^2, \tag{17}$$

where $g^t = \nabla_{\theta_1} \tilde{u}_1(\theta^t, \theta^t)$ is the common gradient on $\mathcal{M}_s$ and $\|\cdot\|_F$ is the Frobenius norm.

The proof follows from the Taylor expansion of $\nabla_{\theta_i} \tilde{u}_i$ around $\theta_1 = \theta_2 = \theta^t$ and the exponential growth of the anti-symmetric component from Appendix A.

E.2. Characterization of the Price of Monoculture for Potential Games

For potential games G with potential function $\Phi : \prod_i \mathcal{A}_i \to \mathbb{R}$, the symmetric Nash equilibrium maximizes $\Phi(a, a)$ over $a \in \Delta(\mathcal{A})$. The efficient Nash maximizes $\sum_i u_i(a)$ over $a \in \prod_i \Delta(\mathcal{A}_i)$. These may differ substantially.

Proposition E.2 (PoM in Potential Games)**.** *For a symmetric potential game where the symmetric Nash σ^* maximizes $\Phi(\sigma, \sigma)$ and the efficient Nash $\sigma^\dagger$ maximizes $W(\sigma)$:*

$$\text{PoM} \geq 1 - \frac{W(\sigma^*)}{W(\sigma^\dagger)}, \tag{18}$$

which can be bounded below using the potential function, since $W(\sigma^*) \leq \Phi(\sigma^*, \sigma^*)$ when the social welfare is dominated by the potential. This lower bound is computable without solving the full Nash equilibrium problem.

E.3. Connection to Mean-Field Game Theory

In the limit $N \to \infty$, our N-agent model connects to the mean-field game (MFG) literature (Lasry & Lions, 2007; Huang et al., 2006). The monoculture equilibrium corresponds to the *symmetric mean-field equilibrium*, while architectural diversity enables the existence of *heterogeneous mean-field equilibria* with potentially higher aggregate welfare. Specifically, in the

infinite-agent limit:

$$\lim_{N\to\infty} \mathrm{PoM}^{(N)} = \frac{V}{C}, \tag{19}$$

which equals the ratio of resource value to conflict cost. This provides a clean asymptotic characterization: in highly competitive ($V/C \to 1$) or large populations ($N \to \infty$), the monoculture welfare loss approaches V/C, which can be substantial (e.g., $V/C = 0.5$ implies a 50% welfare loss).

F. Worked Example: Three-Action Extension

Consider extending the Hawk-Dove game to three actions $\{H, N, D\}$ (Hawk, Neutral, Dove) with payoff matrix:

$$M_3 = \begin{pmatrix} 0 & 5 & 3 \\ 1 & 3 & 2 \\ 0 & 4 & 4 \end{pmatrix}. \tag{20}$$

This game has multiple Nash equilibria, including a symmetric mixed Nash $\sigma^* = (p_H^*, p_N^*, p_D^*)$ satisfying the indifference conditions. With shared softmax policy $\varphi(\theta) = \mathrm{softmax}(\theta)$, $\theta \in \mathbb{R}^3$, the symmetric manifold is now $\mathcal{M}_s^3 = \{(\theta, \theta) : \theta \in \mathbb{R}^3\}$ and the monoculture dynamics reduce to a 3D ODE on $\mathcal{M}_s^3$.

The structure of Theorems 3.1 and 3.3 extends directly; however, the symmetric Nash may now be in the interior of the simplex and involve mixing over all three actions. The welfare loss from monoculture in this richer setting can be larger or smaller depending on the specific payoff structure. We leave the systematic analysis of multi-action anti-coordination games to future work.

G. Broader Implications and Policy Considerations

G.1. AI Governance and Market Design

Our results provide formal foundations for the intuition that AI monoculture is economically harmful. In the context of algorithmic pricing:

- **Symmetric equilibrium lock-in**: Theorem 3.9 shows that monoculture agents in Cournot competition converge to the *symmetric* Nash equilibrium, in which every firm produces the same quantity $q_N^* = (1-c)/(N+1)$. While this is the unique Nash equilibrium of the symmetric game, it forecloses welfare-improving asymmetric outcomes: efficient role differentiation (e.g., one dominant firm, others as price-takers) requires heterogeneous agent parameters. Regulators concerned about the homogenising effect of foundation models on market behaviour (Calvano et al., 2020) should therefore monitor architectural similarity as a structural market factor.
- **Diversity mandates**: The symmetry-breaking threshold δ^* provides a quantitative basis for diversity requirements: requiring that deployed AI systems have initialization or architectural differences exceeding δ^* would be sufficient to escape the monoculture trap. This is analogous to biodiversity preservation arguments in ecology (Tilman & Downing, 1994).
- **Foundation model proliferation**: As a small number of foundation models are adopted across many applications, the effective diversity of deployed AI systems decreases. Our results suggest this trend warrants careful monitoring, as it may reduce welfare in multi-agent settings even when individual agents are highly capable.

G.2. MARL Alignment

The monoculture effect has implications for multi-agent alignment: a world in which all AI agents optimize with the same parameterization is not merely fragile (the monoculture risk of McIlroy-Young et al. (2021)) but is also strategically suboptimal. Alignment approaches that train many agents from the same base model may inadvertently produce populations that fail to coordinate efficiently in competitive environments. Introducing structured diversity—via diverse reward shaping, initialization perturbation, or ensemble methods—could improve both robustness and welfare.