
The Reflexivity Threshold: A Phase Transition for Multi-Agent Performative Prediction

Ahmed Mahrous
KAUST

Roberto Di Pietro
KAUST

Abstract

Many learning systems are performative: once deployed, their predictions change the data distribution on which future models are trained. In multi-agent settings, this feedback is routed through a network of interacting learners, so stability depends not only on the strength of performativity but also on who affects whom. We study repeated retraining in multi-agent performative prediction and identify a spectral threshold that separates stable learning from worst-case learning obstruction. Prior work often collapses cross-agent feedback into a scalar contraction constant, missing this network structure. We instead introduce the reflexivity matrix Γ , which records how strongly one agent’s deployment can move another agent’s best response, defined from per-agent gradient sensitivities and curvature parameters. The threshold occurs when the spectral radius of this matrix crosses one. Below the threshold, repeated retraining converges geometrically to a unique performatively stable equilibrium, with cumulative excess loss bounded uniformly in the time horizon. Above the threshold, we construct supercritical instances in which decentralized learning provably fails: every local-information learner incurs cumulative excess loss that grows linearly in the horizon. Together, these results characterize the threshold behavior through upper and lower bounds and recover the single-agent scalar threshold from prior work as a special case.

1 Introduction

Many learning systems, once deployed, change the environment that later generates their training data. A credit scoring model can change how applicants behave; an algorithmic trading strategy can move the price data on which future strategies are trained. This feedback effect is the focus of performative prediction. Existing theory largely studies the case of one learner facing the consequences of its own deployment. Many real systems, however, contain multiple learners whose deployments *jointly* shape the environment.

The multi-agent setting is not simply a larger version of the single-agent problem; it introduces different difficulties. Feedback can travel across agents: one agent’s deployment changes the data distribution, which changes another agent’s retraining problem, which may itself affect the data distribution or other agents. Stability therefore depends not only on the overall strength of performativity, but also on how feedback is routed through the agent network.

We ask whether there is a general threshold separating stable multi-agent performative learning from unavoidable learning failure. When does repeated retraining converge to a stable equilibrium?

We answer this question through a reflexivity matrix Γ with entries $\Gamma_{ij} := a_{ij}/\mu_i$, measuring how strongly agent j ’s deployment can move agent i ’s best response, after normalizing by the curvature of agent i ’s objective. Thus Γ records who affects whom, and how strongly, in the performative feedback network.

Our main result is a phase transition at $\rho(\Gamma) = 1$, where $\rho(\Gamma)$ is the spectral radius of Γ . This quantity captures the largest asymptotic amplification factor of the feedback network. If $\rho(\Gamma) < 1$,

cross-agent feedback dissipates: repeated retraining is a contraction in a Perron–Frobenius weighted norm, converges geometrically to a unique performatively stable equilibrium, and incurs bounded cumulative excess loss. Above the threshold, this positive guarantee cannot be extended. We construct supercritical systems in which every decentralized local-information learner suffers cumulative loss that grows linearly with time, regardless of the retraining rule. The spectral radius of the reflexivity matrix therefore separates stable adaptation from unavoidable amplification in multi-agent performative systems.

Contributions. First, we introduce the reflexivity matrix Γ , a nonnegative matrix whose entry Γ_{ij} records how strongly agent j ’s deployment can move agent i ’s best response, as determined by per-agent gradient sensitivities and curvature constants. We prove an entrywise best-response sensitivity bound that reduces the nonlinear retraining dynamics to a spectral comparison through Γ , making Perron–Frobenius theory available for the multi-agent performative setting. Second, we prove the subcritical side of the threshold: if $\rho(\Gamma) < 1$, repeated retraining is a contraction in a Perron–Frobenius weighted norm, yielding a unique performatively stable equilibrium, geometric convergence from any initialization, and cumulative dynamic excess loss bounded uniformly in the time horizon. Third, we prove the supercritical side: for a rank-one family with $\rho(\Gamma) = 1 + \varepsilon$, every decentralized local-information learner incurs cumulative excess loss $\Omega(\Delta^2 T)$ over horizons $T = O(1/\varepsilon)$. Together, these results identify $\rho(\Gamma) = 1$ as a boundary for stability guarantees and recover the single-agent scalar threshold from prior work when $k = 1$.

2 Related Work

This paper draws on three lines of research. The first is the single-agent theory of performative prediction, which introduces the framework we extend. The second is the multi-agent theory of performative prediction and decision-dependent games, which is where our contribution is most directly positioned. The third is the combined tradition of spectral methods for network games and statistical lower bound techniques, from which we import two standard technical tools.

Performative prediction. Perdomo et al. [2020] introduced performative prediction in a single-agent setting, formalizing performative stability and convergence of repeated retraining under sensitivity conditions. Subsequent work largely retained this single-agent viewpoint while extending the framework to stochastic, causal, nonconvex, and sequential settings [Mendler–Dünner et al., 2020, Miller et al., 2021, Brown et al., 2022, Hardt et al., 2022, Mofakhami et al., 2023, Mandal et al., 2023]. Closest to our threshold result, Anagnostides et al. [2026] established hardness and lower bounds at the single-agent scalar threshold. We depart from this line by studying decentralized multi-agent retraining, where performative feedback is cross-agent and the scalar threshold is replaced by the spectral condition $\rho(\Gamma) = 1$.

Multi-agent performativity. The multi-agent extension is the most direct predecessor of our work, and the literature divides into three strands.

The first establishes scalar contraction conditions for convergence. Narang et al. [2023] formalized n -player performative prediction with a scalar aggregate threshold $\rho < 1$, and analogous scalar conditions appear in cooperative consensus learning [Li et al., 2022], decentralized games with coupled constraints [Yan and Cao, 2024], and a Wasserstein-free variant [Zhong et al., 2025]. We depart from this strand on two axes. We replace the scalar aggregate with a matrix Γ that preserves per-agent-pair feedback, and we pair the upper bound with a worst-case information-theoretic lower bound above the threshold.

The second develops spectral characterizations in special models. Wang et al. [2023] derived a necessary-and-sufficient spectral condition for a linear-quadratic mean-squared-error model on multiplex networks. Their spectral object is a weighted adjacency matrix with exogenous structure, and the condition is specific to the linear-quadratic model. Our Γ is built from primitive per-agent quantities, applies to general strongly convex losses, and is paired with a lower bound.

The third identifies dynamical phase transitions. Piliouras and Yu [2023] proved a transition from stability to Li–Yorke chaos as update aggressiveness grows under scalar aggregate influence. Their threshold is dynamical and concerns the long-run behavior of specific update rules; ours is information-theoretic and applies to every decentralized local-information learner.

Spectral methods and lower bounds. Two technical traditions supply tools used in our analysis without serving as competing frameworks. Spectral conditions for equilibrium uniqueness and stability are well established in network games [Ballester et al., 2006, Bramoullé et al., 2014, Parise and Ozdaglar, 2019], where the spectral radius governs payoff coupling: agent j ’s action enters agent i ’s utility directly. In our setting, it governs distributional feedback through $D(w)$, and Γ is built from gradient sensitivities and curvature constants rather than payoff coefficients. Our lower bound uses the two-point method [Tsybakov, 2009]; recent extensions to interactive [Chen et al., 2024] and decision-dependent [Cutler et al., 2024] settings do not cover our performative learning setting. Our construction is tailored to multi-agent performative feedback over a finite horizon.

3 Model

This section specifies the multi-agent performative prediction model used throughout the paper. We describe the agents, their objectives, the dynamics of repeated retraining, and the assumptions that underlie the reduction of Section 4.

There are k agents. Each agent $i \in \{1, \dots, k\}$ chooses a deployment w_i from a feasible set $\mathcal{W}_i = \mathbb{R}^{d_i}$. The joint deployment profile is $w := (w_1, \dots, w_k) \in \mathcal{W} := \mathcal{W}_1 \times \dots \times \mathcal{W}_k$. The deployed profile induces a shared endogenous distribution $D(w)$ over observations $z \in \mathcal{Z}$. This distribution represents the data-generating process that all agents face after the profile w has been deployed. The dependence of D on w captures the performative nature of the system: the environment that agents observe is shaped by their own collective actions.

Each agent i has a loss function $\ell_i : \mathcal{W}_i \times \mathcal{Z} \rightarrow \mathbb{R}$ and aims to minimize its performative risk

$$F_i(u_i; w) := \mathbb{E}_{z \sim D(w)} [\ell_i(u_i; z)]. \quad (1)$$

The two arguments of F_i play distinct roles. The first, u_i , is the candidate action optimized by agent i . The second, w , is the deployed profile that determines the environment through $D(w)$. Retraining couples these roles by replacing the deployed profile with agents’ optimized actions.

For each agent i and each deployed profile w , the best response is

$$B_i(w) := \arg \min_{u_i \in \mathcal{W}_i} F_i(u_i; w). \quad (2)$$

Assumption 1 below ensures that this minimizer exists and is unique. The joint best-response map is $B(w) := (B_1(w), \dots, B_k(w))$. An equilibrium of the system is a profile that is consistent with every agent’s best response.

Definition 1 (Performatively stable equilibrium). *A profile $w^* \in \mathcal{W}$ is a performatively stable equilibrium if $w^* = B(w^*)$.*

At a performatively stable equilibrium, each agent’s deployed action is optimal for the distribution induced by the joint deployment. Existence of such an equilibrium is not assumed; it is derived in Section 5 as a consequence of the spectral condition on the reflexivity matrix.

We study the repeated retraining dynamics in which every agent simultaneously replaces its current deployment with its best response,

$$w_{t+1} = B(w_t), \quad t = 0, 1, 2, \dots, \quad (3)$$

with initialization $w_0 \in \mathcal{W}$. The central question of the paper is when the sequence $\{w_t\}$ converges to a performatively stable equilibrium, and when convergence fails.

Main assumptions. Three assumptions are used throughout the paper. The first two are local to individual agents and enter the reduction of Section 4. The third is a mild regularity condition on the feasible sets that allows application of Banach’s fixed-point theorem.

Assumption 1 (Strong convexity in own action). *For each agent i and every fixed deployed profile $w \in \mathcal{W}$, the map $u_i \mapsto F_i(u_i; w)$ is μ_i -strongly convex on $\mathcal{W}_i$, with $\mu_i > 0$.*

Strong convexity makes $B_i(w)$ unique and quantifies agent i ’s resistance to gradient perturbations. Larger μ_i means that the best response is less sensitive to changes in the data distribution.

Assumption 2 (Gradient sensitivity to the deployed profile). *For each i , F_i is continuously differentiable in u_i , and there exist nonnegative constants $a_{ij} \geq 0$ for all $i, j \in \{1, \dots, k\}$ such that for every agent i , every candidate action $u_i \in \mathcal{W}_i$, and every pair of deployed profiles $w, w' \in \mathcal{W}$,*

$$\|\nabla_{u_i} F_i(u_i; w) - \nabla_{u_i} F_i(u_i; w')\| \leq \sum_{j=1}^k a_{ij} \|w_j - w'_j\|.$$

Gradient sensitivity quantifies how a change in the deployed profile affects each agent's first-order condition. The coefficient a_{ij} measures how much agent j 's deployment disturbs agent i 's gradient. If agent j has no effect on agent i 's optimization problem, then $a_{ij} = 0$. Larger values of a_{ij} indicate stronger cross-agent coupling. The diagonal coefficient a_{ii} captures agent i 's own self-reflexive effect through the environment.

Assumptions 1 and 2 are the only two channels through which agents interact mathematically in this paper. Every downstream result is derived from these two assumptions and their combination in Section 4. The reflexivity matrix Γ of Section 4 is built from exactly the parameters $\{\mu_i\}$ and $\{a_{ij}\}$ introduced here.

Assumption 3 (Unconstrained domains). *Each feasible set is $\mathcal{W}_i = \mathbb{R}^{d_i}$.*

This makes $\mathcal{W}$ complete under the norms used below, enabling Banach's fixed-point theorem, and ensures that retraining stays in $\mathcal{W}$.

Throughout, x_i denotes agent i 's block of a vector $x \in \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_k}$, with Euclidean norm $\|\cdot\|$. The spectral radius of M is $\rho(M)$, and for $q \in \mathbb{R}_{++}^k$ we write $\|x\|_q := \max_i \|x_i\|/q_i$.

4 The Reduction: From Retraining Dynamics to Γ

The retraining dynamics $w_{t+1} = B(w_t)$ are nonlinear and couple all k agents through the shared endogenous distribution $D(w)$. Direct analysis of these dynamics is difficult without additional structure. This section introduces the reduction used throughout the paper: under the assumptions of Section 3, changes in the nonlinear map B are bounded entrywise by a nonnegative matrix Γ . The spectral properties of Γ can therefore be used to bound the dynamics of B .

The first step bounds the shift in any agent's best response induced by a shift in the deployed profile. The bound is linear and decomposes cleanly across agents.

Theorem 2 (Best-response sensitivity). *Under Assumptions 1–2, for every agent i and every pair of deployed profiles w, w' ,*

$$\|B_i(w) - B_i(w')\| \leq \sum_{j=1}^k \frac{a_{ij}}{\mu_i} \|w_j - w'_j\|.$$

Proof. Fix an agent i and let $u_i := B_i(w)$ and $u'_i := B_i(w')$. First-order optimality gives $\nabla_{u_i} F_i(u_i; w) = 0$ and $\nabla_{u_i} F_i(u'_i; w') = 0$. Strong convexity of $F_i(\cdot; w)$ (Assumption 1) yields $\mu_i \|u_i - u'_i\| \leq \|\nabla_{u_i} F_i(u_i; w) - \nabla_{u_i} F_i(u'_i; w)\|$. Substituting the first optimality condition and then the second gives $\mu_i \|u_i - u'_i\| \leq \|\nabla_{u_i} F_i(u'_i; w)\| = \|\nabla_{u_i} F_i(u'_i; w) - \nabla_{u_i} F_i(u'_i; w')\|$. The right-hand side is a gradient difference at a fixed action u'_i across two deployed profiles, which Assumption 2 bounds by $\sum_j a_{ij} \|w_j - w'_j\|$. Dividing by $\mu_i > 0$ yields the claim. $\square$

The proof uses each structural assumption once: gradient sensitivity bounds the performative change in agent i 's gradient, and strong convexity converts it into a best-response change. Their interaction appears only through a_{ij}/μ_i , the ratio of agent j 's disturbance to agent i 's curvature, which measures how much j can shift i 's best response.

The weights in Theorem 2 define the central object of the paper.

Definition 3 (Reflexivity matrix). *The reflexivity matrix is the $k \times k$ nonnegative matrix Γ with entries $\Gamma_{ij} := a_{ij}/\mu_i$, for $i, j = 1, \dots, k$.*

In terms of Γ , Theorem 2 states that for every i and every w, w' ,

$$\|B_i(w) - B_i(w')\| \leq \sum_{j=1}^k \Gamma_{ij} \|w_j - w'_j\|. \quad (4)$$

The vector of best-response shifts is therefore dominated entrywise by the product of Γ with the vector of action shifts.

The matrix Γ is the central reduction object. It is built from primitive per-agent quantities rather than imposed as an aggregate Lipschitz constant, and its nonnegativity makes Perron–Frobenius theory available for the weighted-norm contraction argument below. The rank-one construction in Section 6 saturates (4) along the Perron direction, so the same spectral radius governs both the upper and lower sides of the threshold.

5 Subcritical Regime: Contraction and Bounded Excess Loss

This section proves the positive side of the threshold. Under the assumptions of Section 3, if $\rho(\Gamma) < 1$, then the retraining map is a contraction in a Perron–Frobenius weighted norm. This yields a unique performatively stable equilibrium, geometric convergence from any initialization, and a cumulative excess-loss bound that does not grow with the time horizon.

Contraction via Perron–Frobenius. The best-response sensitivity inequality (4) bounds how differences between two deployed profiles translate into differences between their best-response profiles. When $\rho(\Gamma) < 1$, Perron–Frobenius theory provides positive weights under which this bound becomes a contraction in a weighted norm. A contraction in a weighted norm means that two deployment profiles become closer after one retraining step, once each agent’s deviation is rescaled by a positive weight. The weights account for heterogeneous feedback across agents.

Theorem 4 (Weighted contraction below threshold). *Suppose Assumptions 1–2 and 3 hold, and $\rho(\Gamma) < 1$. Then there exist a positive vector $q \in \mathbb{R}_{++}^k$ and a constant $\alpha \in (0, 1)$ such that*

$$\|B(w) - B(w')\|_q \leq \alpha \|w - w'\|_q \quad \text{for all } w, w' \in \mathcal{W},$$

where the weighted norm is defined by

$$\|x\|_q := \max_{1 \leq i \leq k} \frac{\|x_i\|}{q_i}.$$

Proof. Since Γ is nonnegative and $\rho(\Gamma) < 1$, Perron–Frobenius theory guarantees the existence of a vector $q \in \mathbb{R}_{++}^k$ and a constant $\alpha \in (0, 1)$ satisfying

$$\sum_{j=1}^k \Gamma_{ij} q_j \leq \alpha q_i \quad \text{for every } i. \quad (5)$$

If Γ is reducible, apply the argument to the perturbed matrix $\Gamma_\eta := \Gamma + \eta \mathbf{1}\mathbf{1}^\top$ for $\eta > 0$ chosen small enough that $\rho(\Gamma_\eta) < 1$; the strictly positive Perron eigenvector q of Γ_η then satisfies $\Gamma q \leq \Gamma_\eta q = \rho(\Gamma_\eta)q$, and we set $\alpha := \rho(\Gamma_\eta) \in (\rho(\Gamma), 1)$.

Fix $w, w' \in \mathcal{W}$ and an agent i . By the definition of the weighted norm, $\|w_j - w'_j\| \leq q_j \|w - w'\|_q$ for every j . Substituting this bound into the best-response sensitivity inequality (4) and then applying (5) yields

$$\|B_i(w) - B_i(w')\| \leq \sum_{j=1}^k \Gamma_{ij} q_j \|w - w'\|_q \leq \alpha q_i \|w - w'\|_q.$$

Dividing by $q_i > 0$ and taking the maximum over i gives the claim. $\square$

Equivalently, for any $\alpha \in (\rho(\Gamma), 1)$, one can choose positive weights q such that $\Gamma q \leq \alpha q$. The contraction factor can therefore be made arbitrarily close to $\rho(\Gamma)$. The proof uses only two facts: Γ is nonnegative and $\rho(\Gamma) < 1$.

The contraction theorem gives the standard fixed-point consequences. Since B is a contraction on the complete space $(\mathcal{W}, \|\cdot\|_q)$, Banach’s fixed-point theorem applies.

Corollary 5 (Existence, uniqueness, and geometric convergence). *Under the conditions of Theorem 4:*

1. *The retraining map B admits a unique fixed point $w^* \in \mathcal{W}$, which is the unique performatively stable equilibrium of the system.*
2. *For every initialization $w_0 \in \mathcal{W}$, the iterates $w_{t+1} = B(w_t)$ satisfy*

$$\|w_t - w^*\|_q \leq \alpha^t \|w_0 - w^*\|_q \quad \text{for every } t \geq 0.$$

Proof. Assumption 3 ensures that $\mathcal{W}$ is complete under $\|\cdot\|_q$ and that B maps $\mathcal{W}$ into itself. Theorem 4 shows that B is an α -contraction with $\alpha \in (0, 1)$. Banach's fixed-point theorem yields the unique fixed point w^* . Iterating the contraction inequality with $w = w_t$ and $w' = w^*$ gives the geometric convergence bound. $\square$

Corollary 5 gives the basic learnability result in the subcritical regime. The system has a unique performatively stable equilibrium, and repeated retraining reaches it at a geometric rate from any initialization. The condition is stated through Γ , which is constructed from primitive per-agent quantities.

Contraction of the state dynamics alone does not give a performance bound. To control cumulative excess loss, we add smoothness in each agent's own action.

Assumption 4 (Smoothness in own action). *For each agent i and every fixed deployed profile w , the map $u_i \mapsto F_i(u_i; w)$ is L_i -smooth.*

The natural benchmark in a time-varying environment is the current best response. At round t , agent i deploys $w_{i,t}$, while the action minimizing its risk under the current profile w_t is $B_i(w_t)$. The gap between these two actions is the instantaneous excess loss, and summing over agents and rounds gives

$$\widetilde{\text{Reg}}_T := \sum_{t=0}^{T-1} \sum_{i=1}^k \left(F_i(w_{i,t}; w_t) - F_i(B_i(w_t); w_t) \right). \quad (6)$$

This benchmark compares each deployment to the best response in the environment it actually faces. It also avoids assumptions on how objective values vary with the deployed profile. Bounds against the fixed equilibrium comparator require such assumptions and are therefore treated separately.

Corollary 6 (Bounded cumulative dynamic excess loss). *Under Assumptions 1–2, 3, and 4, and assuming $\rho(\Gamma) < 1$, there exists a constant $C > 0$ depending only on Γ , q , and $\{L_i\}$ such that*

$$\widetilde{\text{Reg}}_T \leq \frac{C}{1 - \alpha^2} \|w_0 - w^*\|_q^2 \quad \text{for every } T \geq 1.$$

In particular, the cumulative excess loss is bounded uniformly in the horizon T .

Proof. Fix $t \geq 1$ and i . The best response $B_i(w_t)$ minimizes $u_i \mapsto F_i(u_i; w_t)$, which is L_i -smooth by Assumption 4. Standard smoothness gives

$$F_i(w_{i,t}; w_t) - F_i(B_i(w_t); w_t) \leq \frac{L_i}{2} \|w_{i,t} - B_i(w_t)\|^2. \quad (7)$$

For $t \geq 1$, the dynamics (3) give $w_{i,t} = B_i(w_{t-1})$, so

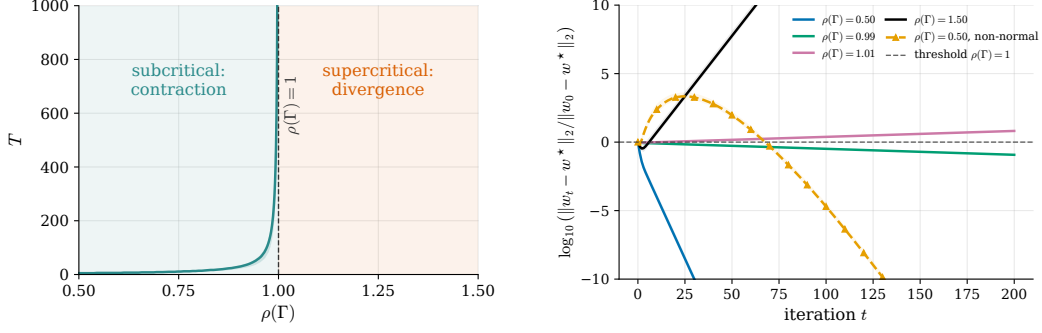
$$\|w_{i,t} - B_i(w_t)\| = \|B_i(w_{t-1}) - B_i(w_t)\| \leq \sum_{j=1}^k \Gamma_{ij} \|w_{j,t-1} - w_{j,t}\|,$$

where the inequality is (4). Using $\|w_{j,t} - w_{j,t-1}\| \leq q_j \|w_t - w_{t-1}\|_q$ gives

$$\|w_{i,t} - B_i(w_t)\| \leq \left(\sum_{j=1}^k \Gamma_{ij} q_j \right) \|w_t - w_{t-1}\|_q.$$

Substituting into (7) and summing over i gives

$$\sum_{i=1}^k \left(F_i(w_{i,t}; w_t) - F_i(B_i(w_t); w_t) \right) \leq C' \|w_t - w_{t-1}\|_q^2,$$



(a) Number of retraining steps needed to satisfy $\|w_t - w^*\|_q \leq \delta$. Convergence slows sharply as $\rho(\Gamma)$ approaches 1 from below. Supercritical ρ values are not plotted as iterates do not converge within the 1,000-iteration cap.

(b) Normalized distance to equilibrium (log scale) under repeated retraining. Subcritical instances contract geometrically, while the plotted supercritical instances move away from equilibrium. The non-normal subcritical curve shows transient growth before contraction.

Figure 1: The reflexivity threshold. The simulations show that convergence slows as $\rho(\Gamma)$ approaches 1 from below. Below the threshold, repeated retraining moves geometrically toward equilibrium; above the threshold, the plotted trajectories move away from equilibrium. Setup details are in Appendix B.

for a constant C' absorbing the L_i and the row-sum factors $(\sum_j \Gamma_{ij} q_j)^2$. By Corollary 5 and the triangle inequality,

$$\|w_t - w_{t-1}\|_q \leq \|w_t - w^*\|_q + \|w_{t-1} - w^*\|_q \leq (1 + \alpha) \alpha^{t-1} \|w_0 - w^*\|_q,$$

so the per-round excess loss decays as $\alpha^{2(t-1)}$. Summing the geometric series from $t = 1$ to $T - 1$ gives a contribution of order $\|w_0 - w^*\|_q^2 / (1 - \alpha^2)$. For $t = 0$, $\|w_0 - B(w_0)\|_q \leq \|w_0 - w^*\|_q + \|B(w^*) - B(w_0)\|_q \leq (1 + \alpha) \|w_0 - w^*\|_q$, which fits the same geometric series. Combining yields the claim. $\square$

The bound depends on the initialization and the contraction factor but not on the time horizon, with the $1/(1 - \alpha^2)$ factor arising from the geometric decay of the one-step gap together with smoothness. Below the threshold, cumulative excess loss is thus $O(1)$; Section 6 constructs supercritical instances with $\Omega(T)$ cumulative excess loss over a critical horizon.

As $\rho(\Gamma) \rightarrow 1^-$, the iteration complexity to a fixed accuracy and the excess-loss bound both scale like $1/(1 - \rho(\Gamma))$, matching the supercritical lower bound's horizon of order $1/(\rho(\Gamma) - 1)$. Figure 1 illustrates this threshold behavior.

6 Supercritical Regime

Section 5 established that $\rho(\Gamma) < 1$ is sufficient for geometric convergence and bounded cumulative excess loss. Whether the supercritical regime produces an intrinsic learning obstruction is delicate: failure of contraction does not by itself imply information-theoretic impossibility, since with enough observations a learner may recover the correct equilibrium even when repeated retraining diverges. We show that the threshold has a sharp negative side in the worst case. For every $\varepsilon > 0$, we construct a supercritical family with $\rho(\Gamma) = 1 + \varepsilon$ in which any decentralized local-information learner—one that at each round uses only its own history of observations and actions—incurs linear cumulative excess loss for $T = O(1/\varepsilon)$ rounds.

All agents know the model parameters $(\Gamma_\varepsilon, \Delta, \sigma_\varepsilon^2, \nu)$; only the hidden parameter ν is unknown.

Definition 7 (Decentralized local-information learner). A decentralized local-information learner is a policy profile $A = (A_1, \dots, A_k)$ in which agent i 's action $w_{i,t}$ at round t is measurable with respect to the private history $\mathcal{H}_{i,t}^{\text{pre}} := \sigma(w_{i,0:t-1}, Y_{i,0:t-1}, U_{i,0:t})$, where $Y_{i,s}$ is agent i 's private observation at round s and $U_{i,s}$ is agent i 's internal randomization.

No agent observes the full system state: each sees only its own signal, without access to other agents' actions or observations. The lower bound applies to any learner satisfying this constraint, regardless of the algorithm used within each agent.

Hard family. The construction places all endogenous amplification along a single direction, which simplifies the analysis and makes the separation between the two environments explicit.

Fix $\varepsilon \in (0, 1]$ and a unit vector $v \in \mathbb{R}^k$ with nonnegative entries. Define the reflexivity matrix

$$\Gamma_\varepsilon := (1 + \varepsilon) vv^\top. \quad (8)$$

The matrix Γ_ε is nonnegative, symmetric, and rank one, with single nonzero eigenvalue $1 + \varepsilon$, so $\rho(\Gamma_\varepsilon) = 1 + \varepsilon$. The system is therefore supercritical by construction, with ε measuring the distance above the threshold. Each agent has $d_i = 1$ and $\mathcal{W}_i = \mathbb{R}$, and the quadratic loss gives $\mu_i = 1$, so the abstract reflexivity matrix of Section 4 coincides with Γ_ε in this construction.

The environment contains a hidden binary parameter $\nu \in \{-1, +1\}$ that determines the true state of the world. Fix an amplitude $\Delta > 0$ and let the noise variance be $\sigma_\varepsilon^2 := 8\Delta^2/\varepsilon$. The observation model is

$$Y_t = \nu\Delta v + \Gamma_\varepsilon w_t + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \sigma_\varepsilon^2 I_k), \quad (9)$$

where each agent i observes only the coordinate $Y_{i,t}$.

Each agent uses quadratic loss $\ell_i(a; y) := \frac{1}{2}(a - y)^2$, which gives the performative risk

$$F_i(a; w, \nu) = \frac{1}{2} \left(a - \nu\Delta v_i - (\Gamma_\varepsilon w)_i \right)^2 + \frac{\sigma_\varepsilon^2}{2}. \quad (10)$$

The additive constant $\sigma_\varepsilon^2/2$ cancels in differences and is dropped below. Minimizing over a gives the best-response map $B(w; \nu) = \nu\Delta v + \Gamma_\varepsilon w$. The fixed-point equation $w^{*,\nu} = B(w^{*,\nu}; \nu)$ has the explicit solution

$$w^{*,\nu} = (I - \Gamma_\varepsilon)^{-1} \nu\Delta v = -\frac{\nu\Delta}{\varepsilon} v. \quad (11)$$

The last equality uses the identity $(I - \Gamma_\varepsilon)v = -\varepsilon v$, which follows from the definition (8).

The two equilibria $w^{*,+}$ and $w^{*, -}$ are therefore located at distance $2\Delta/\varepsilon$ along v . As ε decreases, these equilibria move apart, and the geometric penalty for choosing the wrong sign grows. This divergence is the mechanism that converts slow information acquisition into large cumulative excess loss.

The instantaneous excess performative loss at round t is

$$r_t^\nu := \sum_{i=1}^k \left(F_i(w_{i,t}; w_t, \nu) - F_i(w_i^{*,\nu}; w^{*,\nu}, \nu) \right), \quad (12)$$

and the cumulative excess loss over T rounds is $\text{Reg}_T^\nu(A) := \sum_{t=0}^{T-1} r_t^\nu$.

The subcritical result uses a dynamic comparator (the current best response $B_i(w_t)$), whereas (12) uses the fixed equilibrium comparator $w^{*,\nu}$: the dynamic comparator measures one-step adaptation to the current caused environment, while the equilibrium comparator measures whether the learner identifies the correct stable equilibrium. In the supercritical construction these goals decouple, so the equilibrium comparator is the appropriate benchmark for the two-environment testing argument. For the quadratic family in (10) the two comparators in fact yield identical instantaneous excess loss, so the $\Omega(\Delta^2 T)$ bound holds under either benchmark.

Theorem 8 (Linear lower bound in a supercritical family). *There exist absolute constants $c_0, c_1 > 0$ such that the following holds. For every $\varepsilon \in (0, 1]$, in the supercritical family defined by (8)–(9) with $\sigma_\varepsilon^2 = 8\Delta^2/\varepsilon$, every decentralized local-information learner A satisfies*

$$\frac{1}{2} \mathbb{E}_+[\text{Reg}_T^+(A)] + \frac{1}{2} \mathbb{E}_-[\text{Reg}_T^-(A)] \geq c_1 \Delta^2 T \quad \text{for every } T \leq c_0/\varepsilon.$$

In particular,

$$\sup_{\nu \in \{-1, +1\}} \mathbb{E}_\nu[\text{Reg}_T^\nu(A)] \geq c_1 \Delta^2 T \quad \text{for every } T \leq c_0/\varepsilon.$$

The proof proceeds in three steps. Step 1 shows that the two environments $\nu = +1$ and $\nu = -1$ remain statistically indistinguishable over the critical horizon. Step 2 shows that misidentifying the environment forces the learner into a region of large excess loss. Step 3 combines these two facts.

Proof sketch; full proof in Appendix A. The argument has two parts.

Indistinguishability. Conditional on the realized pre-observation history $\mathcal{F}_t$, the action w_t is fixed by the learner’s policy. Under the two hypotheses, the next observation is Gaussian with common covariance $\sigma_\varepsilon^2 I_k$ and conditional means differing by $2\Delta v$. The per-round Kullback–Leibler divergence is therefore $\|2\Delta v\|^2/(2\sigma_\varepsilon^2) = 2\varepsilon/C$ by the noise scaling. Chain rule and Pinsker’s inequality give $\text{TV}(P_+^t, P_-^t) < 1/2$ for $t \leq c_0/\varepsilon$ with $c_0 = C/8$, so the minimax testing error is bounded below by an absolute constant $p_0 > 0$.

Geometric cost of misidentification. The two equilibria (11) lie at $\pm(\Delta/\varepsilon)v$, so on the event that the midpoint decoder picks the wrong sign, $|\langle w_t - w^{*,\nu}, v \rangle| \geq \Delta/\varepsilon$. The closed form $r_t^\nu = \frac{1}{2}\|(I - \Gamma_\varepsilon)(w_t - w^{*,\nu})\|^2$ together with $(I - \Gamma_\varepsilon)v = -\varepsilon v$ then gives $r_t^\nu \geq \Delta^2/2$ on that event. The factor ε from the contraction of $(I - \Gamma_\varepsilon)$ along v cancels exactly with the factor $1/\varepsilon$ from the equilibrium separation, which is why the obstruction survives the limit $\varepsilon \rightarrow 0$ and produces a horizon of order $1/\varepsilon$.

Combining the two parts: averaging over ν and using minimax testing error at least p_0 gives $\frac{1}{2}\mathbb{E}_+[r_t^+] + \frac{1}{2}\mathbb{E}_-[r_t^-] \geq p_0\Delta^2/2$ for every $t \leq c_0/\varepsilon$, and summing yields the claim with $c_1 = p_0/2$. $\square$

7 Discussion

Reflexivity matrix and interpretation. The framework replaces a scalar Lipschitz constant with the reflexivity matrix Γ , whose entries a_{ij}/μ_i retain the per-pair structure of performative feedback. This matters because two systems with the same scalar bound can have different stability behavior if their feedback is routed through different agent networks. The Perron weights q in Theorem 4 encode this heterogeneity by rescaling agents according to their role in the collective retraining dynamics. When $k = 1$, Γ reduces to the scalar ratio a_{11}/μ_1 , recovering the single-agent contraction threshold of Perdomo et al. [2020]. More generally, $\rho(\Gamma) = 1$ separates dissipative feedback, where repeated retraining contracts, from supercritical feedback, where uniform stability guarantees fail. We use *reflexivity* for this deployment-to-data channel, following the financial-market terminology of Soros [1987]; we do not formalize the broader economic concept.

Beyond the spectral radius. The threshold $\rho(\Gamma) = 1$ governs asymptotic behavior, but finite-horizon dynamics can depend on finer matrix properties and initial conditions. One important case is non-normality: $\|\Gamma^t\|$ may grow transiently even if $\rho(\Gamma) < 1$, as illustrated by the non-normal subcritical trajectory in Panel 1b. Thus the spectral radius identifies the asymptotic threshold, while quantities such as the numerical range or pseudospectra may better capture finite-horizon amplification. We leave this refinement to future work.

Empirical measurement. The framework identifies a stability quantity, $\rho(\Gamma)$, that is measurable in principle but difficult to identify empirically. Each entry a_{ij} is counterfactual: it describes how agent i ’s gradient would respond to alternative deployments by agent j , while μ_i is also not directly observed. Estimation is therefore a causal identification problem, since common system shocks may affect both agent j ’s deployment and agent i ’s gradient. Testing whether estimated values of $\rho(\Gamma)$ predict observed instability is a concrete direction for future work.

Limitations and extensions. Several modeling choices delimit our results. We assume strong convexity in each agent’s own action and exact best-response retraining; extending the analysis to weaker curvature assumptions, approximate best responses, gradient updates, or stochastic retraining is a natural direction for future work. We also treat the reflexivity matrix as fixed across rounds. If the interaction structure evolves, a natural extension is to study time-varying matrices Γ_t , where tools such as joint spectral radii may be needed to characterize stability. The lower bound establishes that no universal positive guarantee can extend above the threshold, since at least one supercritical family forces linear cumulative excess loss; it does not assert that every supercritical Γ is unlearnable. Characterizing which supercritical systems remain learnable is open for future work. Coordination among agents may shift the threshold, while damped retraining is a natural design lever for stabilizing otherwise unstable systems.

References

- Ioannis Anagnostides, Rohan Chauhan, Ioannis Panageas, Tuomas Sandholm, and Jingming Yan. On the computational complexity of performative prediction. *arXiv preprint arXiv:2601.20180*, 2026.
- Coralio Ballester, Antoni Calvó-Armengol, and Yves Zenou. Who’s who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417, 2006.
- Yann Bramoullé, Rachel Kranton, and Martin D’Amours. Strategic interaction and networks. *American Economic Review*, 104(3):898–930, 2014.
- Gavin Brown, Shlomi Hod, and Iden Kalemaj. Performative prediction in a stateful world. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pages 6045–6061. PMLR, 2022.
- Fan Chen, Dylan J. Foster, Yanzun Han, Jian Qian, Alexander Rakhlin, and Yunbei Xu. Assouad, Fano, and Le Cam with interaction: A unifying lower bound framework and characterization for bandit learnability. In *Advances in Neural Information Processing Systems*, volume 37, 2024.
- Joshua Cutler, Mateo Díaz, and Dmitriy Drusvyatskiy. Stochastic approximation with decision-dependent distributions: Asymptotic normality and optimality. *Journal of Machine Learning Research*, 25(90):1–49, 2024.
- Moritz Hardt, Meena Jagadeesan, and Celestine Mendler-Dünnér. Performative power. In *Advances in Neural Information Processing Systems*, volume 35, 2022.
- Qiang Li, Chung-Yiu Yau, and Hoi-To Wai. Multi-agent performative prediction with greedy deployment and consensus seeking agents. In *Advances in Neural Information Processing Systems*, volume 35, 2022.
- Debmalya Mandal, Stelios Triantafyllou, and Goran Radanovic. Performative reinforcement learning. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 23642–23680. PMLR, 2023.
- Celestine Mendler-Dünnér, Juan C. Perdomo, Tijana Zrnic, and Moritz Hardt. Stochastic optimization for performative prediction. In *Advances in Neural Information Processing Systems*, volume 33, pages 4929–4939, 2020.
- John P. Miller, Juan C. Perdomo, and Tijana Zrnic. Outside the echo chamber: Optimizing the performative risk. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7710–7720. PMLR, 2021.
- Mehrnaz Mofakhami, Ioannis Mitliagkas, and Gauthier Gidel. Performative prediction with neural networks. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 11079–11093. PMLR, 2023.
- Adhyayan Narang, Evan Faulkner, Dmitriy Drusvyatskiy, Maryam Fazel, and Lillian J. Ratliff. Multiplayer performative prediction: Learning in decision-dependent games. *Journal of Machine Learning Research*, 24(202):1–56, 2023.
- Francesca Parise and Asuman Ozdaglar. A variational inequality framework for network games: Existence, uniqueness, convergence and sensitivity analysis. *Games and Economic Behavior*, 114: 47–82, 2019.
- Juan C. Perdomo, Tijana Zrnic, Celestine Mendler-Dünnér, and Moritz Hardt. Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 7599–7609. PMLR, 2020.
- Georgios Piliouras and Fang-Yi Yu. Multi-agent performative prediction: From global stability and optimality to chaos. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 1047–1074, 2023.
- George Soros. *The Alchemy of Finance: Reading the Mind of the Market*. Simon and Schuster, New York, 1987.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer, 2009.

Xiaolu Wang, Chung-Yiu Yau, and Hoi-To Wai. Network effects in performative prediction games. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 36514–36540. PMLR, 2023.

Wenjing Yan and Xuanyu Cao. Decentralized noncooperative games with coupled decision-dependent distributions. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Guangzheng Zhong, Yang Liu, and Jiming Liu. Towards performatively stable equilibria in decision-dependent games for arbitrary data distribution maps. *arXiv preprint*, 2025.

A Full Proof of Theorem 8

This appendix gives the full proof of Theorem 8, expanding the sketch in Section 6. We retain the same notation: P_ν denotes the law of the full system trajectory under learner A in environment $\nu \in \{-1, +1\}$, and $\mathcal{F}_t$ denotes the history up to and including the action w_t but excluding the observation Y_t .

Proof of Theorem 8. Step 1: Bounding the Kullback–Leibler divergence. Conditional on the realized pre-observation history $\mathcal{F}_t$, the action w_t is fixed by the learner’s policy. The observation Y_t is then drawn from a Gaussian whose mean differs between the two environments only through the term $\nu\Delta v$. Specifically,

$$P_\nu(Y_t | \mathcal{F}_t) = \mathcal{N}(\nu\Delta v + \Gamma_\varepsilon w_t, \sigma_\varepsilon^2 I_k).$$

The means differ by $2\Delta v$, and the covariance matrices are identical. The Kullback–Leibler divergence between two Gaussians with common covariance $\sigma^2 I$ and mean difference μ equals $\|\mu\|^2/(2\sigma^2)$. Applying this formula and using $\sigma_\varepsilon^2 = 8\Delta^2/\varepsilon$,

$$\text{KL}(P_+(Y_t | \mathcal{F}_t) \| P_-(Y_t | \mathcal{F}_t)) = \frac{\|2\Delta v\|^2}{2\sigma_\varepsilon^2} = \frac{2\Delta^2}{\sigma_\varepsilon^2} = \frac{\varepsilon}{4}.$$

By the chain rule for Kullback–Leibler divergence, the total divergence after t rounds is

$$\text{KL}(P_+^t \| P_-^t) = \frac{t\varepsilon}{4}. \quad (13)$$

Choose $c_0 := 1/2$. For every $t \leq c_0/\varepsilon$, Pinsker’s inequality gives

$$\text{TV}(P_+^t, P_-^t) \leq \sqrt{\frac{1}{2} \text{KL}(P_+^t \| P_-^t)} \leq \sqrt{\frac{t\varepsilon}{8}} \leq \sqrt{1/16} = \frac{1}{4}.$$

The minimax testing error for deciding ν from the observed history is therefore bounded below by an absolute constant $p_0 > 0$, independent of ε and t .

Step 2: Geometric cost of misidentification. Define the midpoint decoder

$$\hat{\nu}_t := \arg \min_{s \in \{-1, +1\}} \|w_t - w^{*,s}\|.$$

On the event $\{\hat{\nu}_t \neq \nu\}$, the action w_t lies strictly closer to the incorrect equilibrium than to the correct one. Using (11), the two equilibria lie on opposite sides of the origin along v at distance Δ/ε each. The midpoint between them is the origin, so $\hat{\nu}_t \neq \nu$ implies that the projection of w_t onto v has the same sign as ν . Consequently,

$$|\langle w_t - w^{*,\nu}, v \rangle| \geq \frac{\Delta}{\varepsilon} \quad \text{on } \{\hat{\nu}_t \neq \nu\}. \quad (14)$$

For the quadratic family, the equilibrium identity $w^{*,\nu} = \nu\Delta v + \Gamma_\varepsilon w^{*,\nu}$ implies $F_i(w_i^{*,\nu}; w^{*,\nu}, \nu) = \sigma_\varepsilon^2/2$. Substituting into (12) and summing over i gives

$$r_t^\nu = \frac{1}{2} \|w_t - \nu\Delta v - \Gamma_\varepsilon w_t\|^2 = \frac{1}{2} \|(I - \Gamma_\varepsilon)(w_t - w^{*,\nu})\|^2, \quad (15)$$

where the second equality uses $\nu\Delta v = (I - \Gamma_\varepsilon)w^{*,\nu}$.

The identity $(I - \Gamma_\varepsilon)v = -\varepsilon v$ implies that $(I - \Gamma_\varepsilon)$ contracts the direction v by the factor ε . Combining (14) with the lower bound

$$\|(I - \Gamma_\varepsilon)(w_t - w^{*,\nu})\| \geq |\langle (I - \Gamma_\varepsilon)(w_t - w^{*,\nu}), v \rangle| = \varepsilon |\langle w_t - w^{*,\nu}, v \rangle|,$$

and substituting (14),

$$\|(I - \Gamma_\varepsilon)(w_t - w^{*,\nu})\| \geq \varepsilon \cdot \frac{\Delta}{\varepsilon} = \Delta \quad \text{on } \{\hat{\nu}_t \neq \nu\}.$$

Substituting into (15) yields

$$r_t^\nu \geq \frac{\Delta^2}{2} \quad \text{on } \{\hat{\nu}_t \neq \nu\}. \quad (16)$$

The factor ε from the contraction of $(I - \Gamma_\varepsilon)$ and the factor $1/\varepsilon$ from the equilibrium separation cancel exactly. This cancellation is why the lower bound survives the limit $\varepsilon \rightarrow 0$ and produces a horizon of order $1/\varepsilon$ rather than a faster-vanishing horizon.

Step 3: Combining the bounds. Taking expectation of (16) under environment ν gives

$$\mathbb{E}_\nu[r_t^\nu] \geq \frac{\Delta^2}{2} P_\nu(\hat{\nu}_t \neq \nu).$$

Averaging over $\nu \in \{-1, +1\}$,

$$\frac{1}{2}\mathbb{E}_+[r_t^+] + \frac{1}{2}\mathbb{E}_-[r_t^-] \geq \frac{\Delta^2}{2} \cdot \frac{1}{2} (P_+(\hat{\nu}_t \neq +1) + P_-(\hat{\nu}_t \neq -1)).$$

The bracketed quantity is twice the minimax testing error, which is at least p_0 by Step 1 for every $t \leq c_0/\varepsilon$. Therefore,

$$\frac{1}{2}\mathbb{E}_+[r_t^+] + \frac{1}{2}\mathbb{E}_-[r_t^-] \geq \frac{p_0\Delta^2}{2} \quad \text{for every } t \leq c_0/\varepsilon.$$

Summing over $t = 0, \dots, T-1$ gives the first claim with $c_1 := p_0/2$. The supremum bound follows because the maximum over ν dominates the average. $\square$

B Simulation Details for Figure 1

Both panels use Γ matrices generated as sparse nonnegative matrices with i.i.d. uniform entries on a randomly chosen 15% of positions, rescaled so that $\rho(\Gamma)$ equals the target spectral radius. Trajectories are generated by the linear iteration $w_{t+1} = \Gamma w_t$ with equilibrium $w^* = 0$.

Panel 1a (phase diagram). Dimension $k = 100$ (100 agents). We average over 20 base matrices, each paired with 5 initial states drawn uniformly from the unit ball of the Perron-weighted norm $\|\cdot\|_q$ associated with the base. For each ρ on a grid of 200 values in $[0.5, 1.5]$ (densely sampled near 1), we record the first iteration at which $\|w_t\|_q \leq \delta = 10^{-3}$, with cap $T_{\max} = 1,000$. The plotted curve is the median across the 100 trajectories with an interquartile band. Supercritical ρ values are masked from the plot because they do not converge. The Perron-weighted norm is used because the contraction in Theorem 4 holds in this norm; for our rescaled construction $\|\cdot\|_q$ and $\|\cdot\|_2$ differ only by a ρ -independent constant, so the curve shape is unaffected by the choice.

Panel 1b (trajectories). Each curve plots $\log_{10}(\|w_t - w^*\|_2 / \|w_0 - w^*\|_2)$ averaged over 50 random initializations on the unit sphere, with an interquartile band. The four baseline instances use the same sparse construction as Panel 1a at the labeled spectral radii. The non-normal instance uses a dense upper-triangular Γ with $\rho(\Gamma) = 0.5$ and a transient-growth ratio $\sup_t \|\Gamma^t\|_2 / \rho(\Gamma)^t \approx 10$, achieved by setting upper-triangular entries to amplify a target direction; initializations are drawn from the same distribution as the other curves rather than aligned to the worst-case direction.

Since the experiments are synthetic simulations, they can be reproduced with standard CPU resources.