Look-Ahead Reasoning on Learning Platforms

Haiqing Zhu
Australian National University
haiqing.zhu@anu.edu.au

Tijana Zrnic
Stanford University
tijana.zrnic@stanford.edu

Celestine Mendler-Dünner
ELLIS Institute Tübingen
MPI for Intelligent Systems, Tübingen
Tübingen AI Center
celestine@tue.ellis.eu

Abstract

On many learning platforms, the optimization criteria guiding model training reflect the priorities of the designer rather than those of the individuals they affect. Consequently, users may act strategically to obtain more favorable outcomes. While past work has studied strategic user behavior on learning platforms, the focus has largely been on individual strategic responses to a deployed model, without considering the behavior of other users. In contrast, *look-ahead reasoning* takes into account that user actions are coupled, and—at scale—impact future predictions. Within this framework, we first formalize level-k thinking, a concept from behavioral economics, where users aim to outsmart their peers by looking one step ahead. We show that, while convergence to an equilibrium is accelerated, the equilibrium remains the same, providing no benefit of higher-level reasoning for individuals in the long run. Then, we focus on collective reasoning, where users take coordinated actions by optimizing through their joint impact on the model. By contrasting collective with selfish behavior, we characterize the benefits and limits of coordination; a new notion of alignment between the learner's and the users' utilities emerges as a key concept. Look-ahead reasoning can be seen as a generalization of algorithmic collective action; we thus offer the first results characterizing the utility trade-offs of coordination when contesting algorithmic systems.

1 Introduction

Digital platforms deploy learning algorithms that collect and analyze data about individuals to power services, personalize experiences, and allocate resources. As people come to understand how these systems make decisions, they often adapt strategically to improve their outcomes. Prior research has largely modeled such strategic behavior as *unilateral*: each agent responds to the platform's decision rule by optimizing their own outcome while treating that rule as fixed. For example, a job applicant might rephrase their resume to include keywords that align with an automated screening system's preferences. This perspective neglects the fact that many others may be doing the same—thereby collectively shifting the data distribution from which the platform learns in the future.

In reality, there is ample empirical evidence that users frequently reason about one another's behavior. They may act in solidarity [Tassinari and Maccarrone, 2020], anticipate other people's adaptations to gain an advantage [Kneeland, 2015], or coordinate to amplify their collective influence [Chen, 2018], in the latter case oftentimes facilitated by labor organizations [e.g., Rideshare Drivers United]. On learning platforms in particular, such reasoning involves not only anticipating the behavior of other platform participants but also how those behavioral changes will impact the

39th Conference on Neural Information Processing Systems (NeurIPS 2025).

 https://doi.org/10.52202/085713-2757

learning algorithm in the future. We call this *look-ahead reasoning*. In the resume-screening example, look-ahead reasoning might surface as choosing to emphasize distinct keywords that others have abandoned, anticipating that popular buzzwords will lose predictive value as they become widespread.

1.1 Our contributions

We characterize how look-ahead reasoning—user behavior that anticipates the actions of others in the population—reshapes learning dynamics and equilibria on learning platforms. We begin with selfish agents, who strategically account for the other agents' responses to the deployment of a predictive model and act independently. We then turn to coordinated behavior, where agents strategize together and account for their joint influence on the learner. Finally, contrasting the two settings allows us to characterize the benefits and limitations of coordination when contesting learning platforms.

To capture agents who selfishly aim to outsmart their peers, we formalize the concept of *level-k* thinking [Nagel, 1995] from behavioral economics in the context of learning systems. Level-k thinking captures different depths of strategic thought: a level-k thinker acts assuming they are "one step ahead" of everyone else, who is a level-$(k-1)$ thinker. A level-0 thinker is non-strategic. Higher levels k are defined recursively. In our setup, actions are determined by hallucinating the data distribution resulting from the assumed behavior of other agents and best-responding to the predictive model it induces. We study the dynamics of repeatedly retraining a model acting on a population of level-k thinkers. We show that "deeper" thinking achieved for larger k accelerates the learning dynamics, while resulting in the *same* equilibrium solution, no matter the depth of thinking.

Theorem 1 (Informal)**.** *For $k \geq 1$, let $\alpha_k \in (0,1)$ be the fraction of level k-thinkers in the population, $\sum_{k=1}^{\infty} \alpha_k = 1$. Assume the learner minimizes a loss function that is smooth and strongly convex, and suppose that the agent responses are sufficiently Lipschitz in the model parameters. Then, for some constant $c \in (0,1)$, repeated retraining converges to a unique stable point at rate*

$$O\left(\left[\sum_{k=1}^{\infty} c^k \alpha_k\right]^t\right).$$

Therefore, even if it relies on deeper reasoning, selfish behavior does not improve the agents' utility.

Next, we investigate whether agents can move past this obstacle if they coordinate. Look-ahead reasoning with coordination allows steering model updates by systematically taking into account that the learner learns from the data agents report. We show that the gap between coordination and lack thereof in terms of agent utility is governed by a notion of alignment between the objectives of the learning platform and the population. Below, we use $\ell(z,\theta)$ and $u(z,\theta)$ to denote the learner's loss and the agent utility for deploying model θ on instance z. Furthermore, $\langle a, b\rangle_M := a^\top M b$.

Theorem 2 (Informal)**.** *Let $\mathcal{D}^*$ and $\mathcal{D}^\sharp$ denote the population's data distributions at equilibrium under selfish reasoning and under collective reasoning, respectively. Then, assuming regularity conditions, the benefit of coordination* B*, defined as the difference in population utility at the two equilibria, is bounded as*

$$\mathrm{B} \leq \left(\left\langle \mathbb{E}_{z\sim\mathcal{D}^*}\left[\nabla_\theta u^*\right], \mathbb{E}_{z\sim\mathcal{D}^\sharp}\left[\nabla_\theta \ell^*\right]\right\rangle_{(\mathbf{H}^\star)^{-1}}\right)^2,$$

where $\mathbf{H}^\star = \mathbb{E}_{z\sim\mathcal{D}^*}\left[\nabla^2_\theta \ell(z,\theta^*)\right]$ and θ^* denotes the equilibrium model under selfish reasoning. We use the short-hand notation $\nabla_\theta u^* = \nabla_\theta u(z,\theta^*)$, $\nabla_\theta \ell^* = \nabla_\theta \ell(z,\theta^*)$.

If the average agent utility and the average loss of the learner are orthogonal, there is no benefit to coordination. Similarly, when $u = c \cdot \ell$ for either $c > 0$ or $c < 0$, the benefit of coordination is zero. However, when there is the right amount of overlap between the objectives that the collective can exploit, coordination can lead to more favorable outcomes than selfish reasoning. For example, think of instance z as a feature-label pair (x, y) and consider a learner and a collective who both care about predictions $f_\theta(x)$: the learner cares about making them accurate, and the collective cares about them having certain favorable values. Then, modifying the labels y can be an effective strategy for the collective to steer the predictive model θ, something that individuals reasoning selfishly cannot achieve, enabling a large benefit of coordination. We elaborate on this example later on.

In additional results, we study heterogeneous populations and collectives of varying size. The results explain why larger collectives, despite more leverage, do *not* always lead to a higher utility for participating agents. They also show how broader participation in the collective stabilizes learning dynamics.

https://doi.org/10.52202/085713-2757

1.2 Background and related work

Strategic classification [Hardt et al., 2016, Brückner and Scheffer, 2011] introduces a model to study strategic behavior in learning systems based on assumptions of individual rationality. It describes a population of agents best-responding to a decision rule by altering their features to achieve positive predictions, given a fixed decision rule. Several variations of this basic model have been studied [e.g., Dong et al., 2018, Chen et al., 2020, Bechavod et al., 2022, Ghalme et al., 2021, Jagadeesan et al., 2021, Levanon and Rosenfeld, 2022]; see Podimata [2025] for a recent survey of this literature. All these works focus on studying how agents strategize against a fixed decision rule. Our work introduces a new dimension of reasoning to strategic classification, taking into account how individual agents' actions are coupled and how this influences the predictive model the agents strategize against.

Performative prediction [Perdomo et al., 2020] introduces performative stability as an equilibrium notion that characterizes long-term outcomes in the interaction of a population with a learning system. Performative stability is a fixed point of repeated retraining by the learner in a dynamic environment. Prior work in performative prediction [e.g., Perdomo et al., 2020, Mendler-Dünner et al., 2020, Drusvyatskiy and Xiao, 2023, Brown et al., 2022, Narang et al., 2023] has studied the behavior of retraining and conditions that ensure its convergence to stability in different learning settings. We refer to Hardt and Mendler-Dünner [2025] for an overview of the performative prediction literature. A key concept in performative prediction is the "distribution map," which characterizes how different model deployments impact the population. In convergence analyses this map is typically treated as a fixed and unknown quantity. We study how different types of strategic reasoning impact the distribution map, thus also impacting the resulting convergence properties and equilibria.

A more recent literature on algorithmic collective action [Hardt et al., 2023] studies coordinated agent efforts with the goal of steering learning systems; see [Baumann and Mendler-Dünner, 2024, Ben-Dov et al., 2024, Gauthier et al., 2025, Sigg et al., 2025] for recent developments in this area, as well as related discussions of data leverage [Vincent et al., 2021] and protective optimization technologies [Kulynych et al., 2020]. From the perspective of our work, collective action is a type of look-ahead reasoning: agents plan through model updates under the assumption that they coordinate with other agents. In this work we study the tradeoffs and implications of coordination on the utility of agents participating in the collective. Relatedly, Hardt et al. [2022] discuss how platforms can reduce risk by actively steering a population. Collective action reverses this perspective and investigates how the population can improve its utility by steering the learner. This perspective is related to [Zrnic et al., 2021], who also deviate from the classical model of strategic classification and instead model the population as the leader in the Stackelberg game against the learning platform. Our framework aims to bridge strategic classification and algorithmic collective action to illuminate the benefits and limits of coordination.

2 Setup

We consider a population of individuals interacting with a learning platform. We assume the platform trains a predictive model on the population's data, and individuals strategically alter their data to achieve favorable outcomes. We elaborate below.

Learning platform. Upon observing data about the population, the learner optimizes the parameters $\theta \in \Theta$ of their predictive model f_θ. We work with the following optimality assumption on the learning algorithm: given a loss function ℓ, the learner's response $\mathcal{A}(\mathcal{D})$ to a data distribution $\mathcal{D}$ is given by *risk minimization*, defined as

$$\mathcal{A}(\mathcal{D}) := \operatorname*{arg\,min}_{\theta \in \Theta} \ \mathbb{E}_{z \sim \mathcal{D}} \left[\ell(z, \theta)\right].$$

Strategic agents. We assume individuals are described by data points $z \in \mathcal{Z}$ sampled from a base distribution $\mathcal{D}_0$. Typically, $z = (x, y) \in \mathcal{X} \times \mathcal{Y}$ are feature–label pairs. Individuals implement a data modification strategy $h_\theta : \mathcal{Z} \to \mathcal{Z}$ that maps an individual's data point z to a modified data point $h_\theta(z)$; the strategy can depend on the learning platform's currently deployed model θ. We will sometimes omit the subscript θ if the strategy is independent of the current model, i.e., $h_\theta \equiv h_{\theta'}$ for all θ, θ'. We use $\mathcal{D}_{h_\theta}$ to denote the distribution of $h_\theta(z)$ for $z \sim \mathcal{D}_0$; in other words, this is the distribution of data points after applying strategy h_θ to all base data points. Following the terminology of Perdomo et al. [2020], we call $\mathcal{D}_{h_\theta}$ a *distribution map*. Note that different strategies h_θ correspond to different distribution maps. When the strategy is clear from the context, we will write $\mathcal{D}_{h_\theta} \equiv \mathcal{D}(\theta)$. The notation $\mathcal{D}_h$ equally applies to model-independent strategies h.

 https://doi.org/10.52202/085713-2757

Equilibria and learning dynamics. We study the long-term behavior of the learner repeatedly optimizing their model and the population strategically adapting to it. Formally, we study the learning dynamics of *repeated risk minimization*:

$$\theta_{t+1} = \mathcal{A}(\mathcal{D}_{h_{\theta_t}}).$$

The model θ is repeatedly updated based on observed data from the previous time step. The natural equilibrium of these dynamics is called *performative stability* [Perdomo et al., 2020]. We say a model θ^* is performatively stable with respect to a strategy h_θ if

$$\theta^* = \mathcal{A}(\mathcal{D}_{h_{\theta^*}}).$$

In words, there is no reason for the learner to deviate from the current model, given observations of the strategic response of the population.

Population utility. Different strategies h_θ lead to different equilibria. Rather than just focusing on the learner's loss, we evaluate equilibria in terms of the utility they imply for the population. Let θ^* denote the equilibrium under strategy h_θ, i.e., the performatively stable point. Then, $\mathcal{D}_{h_{\theta^*}}$ denotes the equilibrium distribution. We denote the population's utility after implementing strategy h_θ by

$$U(h_\theta) = \mathbb{E}_{z\sim\mathcal{D}_{h_{\theta^*}}}\left[u(z,\theta^*)\right],$$

where $u(z,\theta)$ is the utility of an individual with data point z when the deployed model is θ. Note that the stable point satisfies $\theta^* = \mathcal{A}(\mathcal{D}_{h_{\theta^*}})$ and thus the utility is evaluated against the data that determines the predictive model. Again, we will sometimes omit the subscript when denoting the strategy if it is independent of the deployed model.

3 Level-k reasoning

Strategic classification [Hardt et al., 2016] assumes that each individual selfishly best-responds to a deployed model θ. As explained earlier, this model does not account for the agents' awareness that they, as a whole, determine the deployed model. To account for this dimension of reasoning, we build on the cognitive hierarchy framework from behavioral economics [Nagel, 1995] and generalize strategic classification to allow individuals to reason through the other individuals' responses. In particular, we formalize *level-k* thinking, which categorizes players by the "depth" of their strategic thought. Intuitively, an individual reasoning at level k assumes a level of cognitive reasoning for the rest of the population and tries to "outsmart" them. In other words, they are always one step ahead: a level-k thinker best-responds to the model that would result from a population of level-$(k-1)$ thinkers. The basic level-k model starts with an explicit assumption about how individuals at level 0 behave. It then defines higher levels of thinking recursively.

Suppose that agents at level 0 are non-strategic and implement $h_\theta^{(0)}(z) = z$ in response to all θ. Then, for every higher level of thinking $k \geq 1$ we define the strategy for level-k thinkers recursively:

$$h_\theta^{(k)}(z) := \operatorname{argmax}_{z'} u\big(z', \mathcal{A}(\mathcal{D}_{k-1}(\theta))\big), \tag{1}$$

where $\mathcal{D}_{k-1}(\theta)$ is the distribution obtained by applying the model-dependent strategy $h_\theta^{(k-1)}$ to every $z \sim \mathcal{D}_0$. At level $k = 1$, we recover the standard microfoundation model of strategic classification [Hardt et al., 2016], where individuals best-respond to a fixed model. For larger k, the agents anticipate the actions of other agents and best-respond to the hypothetical model resulting from the shifted distribution. The hypothetical model being $\theta' = \mathcal{A}(\mathcal{D}_{k-1}(\theta))$.

Different individuals in the population might implement different levels of reasoning. To reflect this we deviate from a homogeneous population and let the population consists of level-k thinkers at different levels k. In particular, we assume that an α_k-fraction of the population has cognitive level k, for $k = 1, 2, \ldots$ and $\sum_{k=1}^{\infty} \alpha_k = 1$. If $\alpha_k = 1$ for some k, then all individuals in the population have the same level of reasoning. This model results in the distribution map:

$$\mathcal{D}(\theta) := \sum_{k=1}^{\infty} \alpha_k \mathcal{D}_k(\theta). \tag{2}$$

We characterize the learning dynamics for different levels of thinking. We use the following Lipschitzness assumption on the induced distribution at level $k = 1$:

$$\mathcal{W}(\mathcal{D}_1(\theta), \mathcal{D}_1(\theta')) \leq \epsilon \left\|\theta - \theta'\right\|_2, \quad \forall \theta, \theta' \in \Theta,$$

where $\mathcal{W}$ denotes the Wasserstein-1 distance. In performative prediction this condition is known as ϵ-sensitivity [Perdomo et al., 2020].

Theorem 3 (Retraining with level-k thinkers)**.** *Suppose the loss of the learner ℓ is γ-strongly convex in θ and β-smooth in z, and that the distribution map $\mathcal{D}_1(\theta)$ is ϵ-sensitive. Then, as long as $\epsilon < \frac{\gamma}{\beta}$, there is a unique stable point θ^* such that for any $(\alpha_k)_{k=1}^{\infty}$ retraining on the mixed population* (2) *converges as*

$$\|\theta_t - \theta^*\|_2 \leq \left(\sum_{k=1}^{\infty} \left(\frac{\epsilon\beta}{\gamma}\right)^k \alpha_k\right)^t \|\theta_0 - \theta^*\|_2 .$$

The core technical step in the proof is to derive how the sensitivity of the distribution map $\mathcal{D}_k(\theta)$ changes recursively with k. In particular, the distribution map $\mathcal{D}(\theta)$ in (2) has sensitivity $\sum_{k=1}^{\infty} \alpha_k \left(\epsilon\beta/\gamma\right)^{k-1} \epsilon$. We refer to Appendix A for the full poof.

For the case where $\alpha_1 = 1$ and thus all agents reason at level $k = 1$, we recover the retraining result of Perdomo et al. [2020]. There are two interesting implications of the generalization in Theorem 3. First, we observe that for populations with higher levels of thinking k, the rate of convergence increases (although the condition for convergence, $\epsilon < \gamma/\beta$, remains the same). This can be interpreted as saying that performative distribution shifts are mitigated when the population has a deeper level of strategic thought. The second implication is that, as long as the agents act selfishly, they cannot benefit from higher levels of reasoning at stability.

Corollary 1. *Under the assumptions of Theorem 3, it holds that $U\big(h_\theta^{(1)}\big) = U\big(h_\theta^{(k)}\big),\ \forall k \geq 1$. Moreover, the utility at stability remains unaltered for any mixed population consisting of level-k thinkers regardless of $(\alpha_k)_{k=1}^{\infty}$.*

This corollary follows from the observation that the stable point θ^* is the same for any mixed population of level-k thinkers. Another consequence of this fact is that the equilibrium strategies $h_{\theta^*}^{(k)}$, and hence the induced distributions $\mathcal{D}_k(\theta^*)$, are identical for every k. We will denote this unique *optimal selfish strategy* by $h^* = h_{\theta^*}^{(k)}$ and the implied data distribution by $\mathcal{D}^*$.

4 Collective reasoning

Level-k thinkers anticipate model changes implied by the population's actions. We saw that higher levels of selfish reasoning do not improve the agents' utility at equilibrium. The fundamental reason is that individually they cannot steer the trajectory of the learning algorithm; they can merely anticipate it. In the following we show how individuals can achieve more favorable outcomes by joining forces and making decisions *collectively*; this gives them steering power.

We denote by $h^\sharp$ the *optimal collective strategy*:

$$h^\sharp = \arg\max_h U(h) = \arg\max_h \ \mathbb{E}_{z\sim\mathcal{D}_h}\left[u(z, \mathcal{A}(\mathcal{D}_h))\right].$$

Notice the difference compared to (1). In (1), the optimization variable z' does not enter the model training $\mathcal{A}$, while above h directly determines the subsequently deployed model. The optimal collective strategy is a Stackelberg equilibrium: the population acts as the Stackelberg leader.

To contrast the optimal collective and selfish strategies, we define the benefit of coordination.

Definition 1 (Benefit of coordination)**.** *Let h^* be the optimal selfish strategy and $h^\sharp$ the optimal collective strategy. We define the benefit of coordination as* $\mathrm{B} = U(h^\sharp) - U(h^*)$.

Since $h^\sharp$ is the globally optimal strategy for the population, it holds that $\mathrm{B} \geq 0$. How large B is depends on the goals pursued by the learner and the population, as characterized by ℓ and u, respectively. Through coordinated data modifications, the population can steer the model towards a common target. But to do so, they may have to deviate from the individually optimal strategy. Thus, what governs the benefit of coordination is the tradeoff between the return of steering the model to a better equilibrium and the cost of being evaluated against the modified data.

We start with a simple case where the benefit of coordination is zero.

Proposition 4. *Suppose $u = c \cdot \ell$ for some $c \neq 0$. Then, it holds that* $\mathrm{B} = 0$.

There are two different mechanisms at play, depending on the sign of c. In an adversarial setting where $c > 0$, the game between the learner and the collective is a zero-sum game. In this case the

 https://doi.org/10.52202/085713-2757

benefit of coordination is zero, as the cost of steering is equal to its return. When $c < 0$, the platform and the agents pursue the same goal and the game becomes a potential game between the two. In this case the benefit of coordination is zero as selfish actions are simultaneously optimal for the collective and there is no benefit to steering the model away from the selfish equilibrium. In general, however, the benefit of coordination can be arbitrarily large. Consider the following example.

Example 1 (Label modifications as an effective collective lever.)**.** *Consider a learner who aims to accurately predict labels from features and a collective that prefers these predictions to follow a target function g. We assume that agents can easily manipulate their labels, while their features cannot be changed. The population and the learner have the following utility and loss, respectively:*

$$u(z, \theta) = -(f_\theta(x) - g(x))^2, \quad \ell(z, \theta) = (f_\theta(x) - y)^2.$$

In this setting, the collective has a powerful lever to steer the model through label modifications. The optimal collective strategy is clearly $h^\sharp(z) = (x, g(x))$. In contrast, selfish agents have no leverage over the learner and would simply report $h^*(z) = (x, y)$. Assuming there exists θ such that $f_\theta(x) = \mathbb{E}[y|x]$, this gives a benefit of coordination equal to the suboptimality of the labeling function $\mathrm{B} = \mathbb{E}_{z\sim\mathcal{D}_0}(\mathbb{E}[y|x] - g(x))^2$, which can be arbitrarily large, depending on $\mathcal{D}_0$ and g.

To characterize when the benefit of coordination is large, we consider linear distribution maps. A generalized version of the result can be found in Appendix A.6.

Assumption 1 (Linearity)**.** *Let each strategy $h(\eta)$ be represented by a parameter vector $\eta \in \mathbb{R}^d$. We say that the induced distribution is linear with respect to the parameterization if, for all $\alpha \in [0, 1]$, $\mathcal{D}_{h(\alpha\eta+(1-\alpha)\eta')} = \alpha\mathcal{D}_{h(\eta)} + (1 - \alpha)\mathcal{D}_{h(\eta')}$.*

Intuitively, linearity means that the population's data distribution is the same whether agents linearly interpolate between two strategies η and η', or they split up in two subgroups and each implements one of the two strategies. Under this assumption, the following result provides a bound on the benefit of coordination. The proof can be found in Appendix A.

Theorem 5 (Benefit of coordination)**.** *Let Assumption 1 hold. Let $U(h(\eta))$ be γ-strongly concave in η and $U(\alpha h + (1 - \alpha)h')$ be differentiable with respect to α. Then, we have $\mathrm{B} \leq \frac{1}{2\gamma}\Phi^2$, where $\Phi := \left\langle \mathbb{E}_{z\sim\mathcal{D}_{h^*}}\left[\nabla_\theta u(z, \theta^*)\right], \mathbb{E}_{z\sim\mathcal{D}_{h^\sharp}}\left[\nabla_\theta \ell(z, \theta^*)\right]\right\rangle_{(\mathbf{H}^\star)^{-1}}$ and $\mathbf{H}^\star = \mathbb{E}_{z\in\mathcal{D}_{h^*}}\left[\nabla^2_{\theta,\theta}\ell(z, \theta^*)\right]$.*

This result shows how the benefit of coordination is governed by the alignment between the utility u of the population and the loss ℓ of the learner, quantified by the inner product of their gradients at the selfish equilibrium θ^*. We refer to Φ as the alignment term.

We have $\Phi = 0$ in the case where the gradients are orthogonal and the two functions u and ℓ are unrelated; in this case agents can optimize their utility independent of the model and there is no benefit to being able to steer the model. Further, in line with Proposition 4, when $u = c \cdot \ell$ for some constant c, we have $\Phi = 0$ because selfish actions are simultaneously collectively optimal and shifting the model away from θ^* would not benefit the collective. In particular, θ^* is a local optimum of the loss under $\mathcal{D}_{h^*}$ and thus the first term in the inner product becomes zero.

To see cases where Φ^2 can be large, we look at the individual terms in its definition. The first term in the inner product that defines Φ describes the suboptimality of θ^* for the agents under $\mathcal{D}_{h^*}$. It is non-zero if the collective's utility on $\mathcal{D}_{h^*}$ can be improved by moving the model away from θ^*. The second term in the inner product captures the response of the learner to the collective strategy. For this term to be non-zero the data modification $\mathcal{D}_{h^\sharp}$ must make θ^* suboptimal. Consequently, Φ is large if both terms in the inner product are non-zero, and there is overlap in the direction of improvement for the learner and the collective. Moreover, what really matters is not just the raw gradient alignment but one filtered through the local curvature of the loss landscape. The directions that the learner finds "flat" (small Hessian eigenvalues) allow for more influence on the model through small data modifications, and thus they offer more leverage for the collective. We provide an example satisfying the assumptions of Theorem 5 with a non-trivial closed-form expression for Φ in Appendix B.

Let us revisit Example 1. Although the example does not satisfy the assumptions in Theorem 5, the alignment term provides the right intuition for why B is large: since the learner aims to accurately predict labels, any systematic label modification induces a model change. The collective can fully control the direction of these changes, i.e., the gradient of the loss, and align it with their objectives. The label modification strategy leverages this to increase the agents' utility, providing non-zero return as long as θ^* is suboptimal for the collective under $\mathcal{D}_{h^*}$.

https://doi.org/10.52202/085713-2757

5 Heterogeneous populations

A perfectly coordinated population where every agent participates, or one that implements the *optimal* collective strategy, is unlikely to emerge in practice. In the following we consider some plausible deviations from the idealized collective studied in the previous section. Unless stated otherwise, we assume the collective implements any fixed strategy h (independent of θ), which could be a simpler alternative to a potentially hard-to-implement optimal strategy $h^\sharp$.

5.1 Scaling strategies in the presence of non-strategic agents

First, we consider a setting where only a fraction of the population participates in the collective and study conditions under which it is worth scaling up a strategy h, meaning a larger collective implies a higher utility for the collective. To study how the collective's utility changes with its size, we consider the following mixture model for the population as proposed in [Hardt et al., 2023]:

$$\mathcal{D}^\alpha = \alpha\mathcal{D}_h + (1-\alpha)\cdot\mathcal{D}_0 \tag{3}$$

where an α-fraction of the population implements the collective strategy h and the remaining $(1-\alpha)$-fraction is non-strategic. Here, the collective strategy h can be any fixed strategy, not necessarily the optimal one. We are interested in the average utility for agents participating in the collective, denoted as $U_\alpha := \mathbb{E}_{z\sim\mathcal{D}_h}\left[u(z;\theta^*_\alpha)\right]$, where θ^*_α is the equilibrium under the mixture model (3).

Proposition 6 (Benefit of scaling up a strategy)**.** *Consider the mixture model in* (3)*, fix a strategy h, and denote the resulting equilibrium by θ^*_α. Then, the benefit of scaling up the strategy h for a collective of size α is positive, i.e., $\frac{\partial U_\alpha}{\partial\alpha} > 0$, if and only if $\Psi < 0$, where $\Psi = \langle \mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta u(z;\theta^*_\alpha)\right], \mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta^*_\alpha)\right]\rangle_{\mathbf{H}^{-1}}$ and $\mathbf{H} := \nabla^2_\theta \mathbb{E}_{z\sim\mathcal{D}^\alpha}\left[\ell(z;\theta^*_\alpha)\right]$.*

Whether a strategy is worth scaling up or not is again linked to a notion of alignment described by Ψ. To provide intuition for the result we consider some special cases. Suppose $\mathbf{H}$ is positive semidefinite, which happens when ℓ is convex; then, if $u = \ell$, $\Psi \geq 0$, and if $u = -\ell$, $\Psi \leq 0$. This means that, under convex losses, when considering the utility of participating agents, scaling up a fixed strategy is always harmful in a zero-sum game, and it is always beneficial when the learner and the population optimize the same target. Note that the result holds for any fixed strategy h. We provide additional intuition and empirical insights into how U_α and Ψ change with α in Section 6.

Optimal size-aware strategy. Next, we aim to understand what collectives can achieve if they are aware of partial participation and optimize their strategy accordingly. We define the optimal size-aware collective strategy for size α as: $h^\sharp_\alpha = \arg\max_h \mathbb{E}_{z\sim\mathcal{D}_h}\left[u(z;\mathcal{A}(\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0)\right]$.

The global Stackelberg solution $h^\sharp$ corresponds to the case where $\alpha = 1$ and the collective utility corresponds to the population utility. In the case where $\alpha < 1$, the collective optimizes the utility of participants, rather than the full population. Agents are informed of their collective size and will choose the best strategy $h^\sharp_\alpha$ accordingly. In the following proposition, we characterize the utility of agents participating in a collective that deploys a size-aware strategy. We use U^*_α to denote the utility of a population of size α implementing the optimal size-aware strategy $h^\sharp_\alpha$.

Proposition 7 (Benefit of larger collectives)**.** *Consider the mixture model* (3) *with a collective of size α implementing the optimal size-aware strategy $h = h^\sharp_\alpha$. Then, the average collective utility U^*_α achieved by implementing $h^\sharp_\alpha$ satisfies $\frac{\partial U^*_\alpha}{\partial\alpha} \geq 0$ if and only if $\frac{\partial U_\alpha}{\partial\alpha}|_{h=h^\sharp_\alpha} \geq 0$.*

Note that the derivative takes into account the dependence of the strategy on α. Thus, the result implies that reoptimizing the strategy as a function of collective size does not change whether scaling up is worth it or not. All that matters is the alignment of the pursued goals.

5.2 Learning dynamics in the presence of selfish agents

Finally, we study how partial participation impacts learning dynamics of repeated risk minimization. For $\alpha = 1$ and a fixed strategy h that is independent of θ the learning dynamics converge in a single step. This also holds for $\alpha < 1$ in the presence of non-strategic agents. However, this changes as soon as agents deviating from the collective strategy act selfishly. To reflect this scenario we consider the following alternative mixture model:

$$\mathcal{D}^\alpha(\theta) = \alpha\mathcal{D}_h + (1-\alpha)\cdot\mathcal{D}(\theta), \tag{4}$$

https://doi.org/10.52202/085713-2757

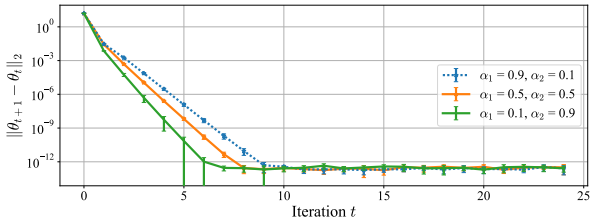


Figure 1: *Convergence of repeated risk minimization on a mixture population of level-k thinkers.* The curves show how the gap between iterates $\|\theta_{t+1} - \theta_t\|_2$ evolves across iterations t for different mixture weights. Error bars indicate one standard deviation over 10 runs.

where an α-fraction of the population implements the collective strategy and an $(1-\alpha)$ fraction deviates from it. We assume these latter act selfishly and their behavior can be characterized by $\mathcal{D}(\theta)$. In particular, the actions of these agents can depend on the deployed model, such as in level-k reasoning discussed in Section 3.

We characterize the rate of convergence of repeated risk minimization under this model and show that larger collectives have the advantage of stabilizing the learning dynamics.

Proposition 8. *Consider the heterogeneous population model* (4)*. Suppose ℓ is γ-strongly convex in θ and β-smooth in z, and that the distribution map $\mathcal{D}(\theta)$ is ϵ-sensitive. Then, as long as $\epsilon < \frac{\gamma}{\beta}$, repeated risk minimization is guaranteed to converge to a unique stable point θ^*_α at rate*

$$\|\theta_t - \theta^*_\alpha\|_2 \leq \left(\frac{\epsilon\beta(1-\alpha)}{\gamma}\right)^t \|\theta_0 - \theta^*_\alpha\|_2 .$$

This result explains how the sensitivity of the non-participating agents to changes in the deployed model, together with the fraction of these agents, determines the rate of convergence to stability. For $\alpha = 0$ we recover the result of Perdomo et al. [2020]. The smaller α the slower the convergence.

6 Simulations

We validate our theoretical findings empirically. We adapt the credit-scoring simulator from [Perdomo et al., 2020] that models how a lending institution classifies loan applicants by creditworthiness.[1] We first focus on the results related to level-k reasoning from Section 3 and then offer empirical insights into the trade-offs of collective reasoning from Section 4 and Section 5.

6.1 Retraining dynamics under level-k thinking

Assume the learner fits a logistic regression classifier θ using cross-entropy loss. The data has 10 features and we assume agents can manipulate the subset $S = \{$'remaining credit card balance','open credit lines', 'number of real estate loans'$\}$. Given some $\epsilon > 0$, the utility of the agents is given by $u_\epsilon((x,y),\theta) = -\langle\theta, x\rangle - \frac{1}{2\epsilon}\|x_0 - x\|_2^2$, where x_0 is their feature value under $\mathcal{D}_0$. The best response of the agents is given by $x^*_S = x_S - \epsilon\theta_S$, where S indexes the strategic features. Note that this corresponds to the strategy for agents reasoning at level-1; indeed, assuming other agents are non-strategic implies strategizing against a fixed model. It is not hard to see that the resulting distribution map $\mathcal{D}_1(\theta)$ is ϵ-sensitive. Under this model we simulate the repeated retraining dynamics for mixed populations of level-1 and level-2 thinkers of varying proportion.

In Figure 1 we report the speed of convergence by presenting the iterate gap $\|\theta_{t+1} - \theta_t\|_2$ against the number of iterations. We choose $\epsilon = 0.5$. First, we can see that under all three mixture weights the gap tends to zero and the dynamics converge. As the fraction of higher levels of thinking increases, the speed of convergence increases, which confirms our theoretical finding in Theorem 3. We also verified empirically that the dynamics converge to a unique equilibrium independent of α.

[1] For the implementation of the simulation, see https://github.com/haiqingzhu543/Look-Ahead-Reasoning-on-Learning-Platforms; for the dataset, see https://www.kaggle.com/c/GiveMeSomeCredit.

https://doi.org/10.52202/085713-2757

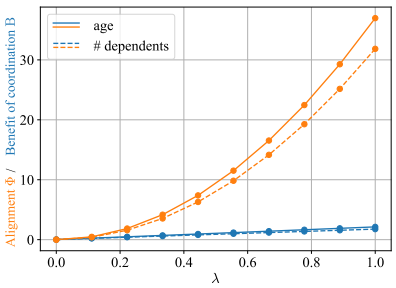


Figure 2: *Alignment serves as a good proxy for the benefit of coordination.* We consider the utility instantiation in (5) and evaluate alignment Φ and the benefit of coordination B for different values of λ. We show them for two strategies that modify the feature 'age', and '#dependents', respectively.

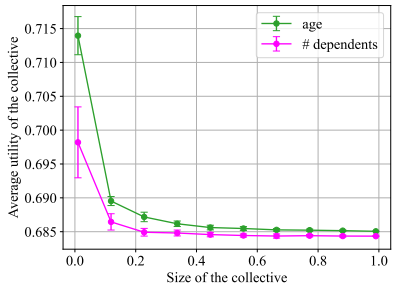


Figure 3: *Collective utility decreases with collective size in the zero-sum case*. The collective implements the optimal size-aware strategy for $\lambda = 0$ in a mixed population with non-strategic agents. Small collectives can realize large gains, but the response by the learner impedes gains at larger sizes α.

6.2 Trade-offs in collective reasoning

With the same credit-scoring data, we study the benefit of coordination and illustrate agent utility for different strategies and collective size. We again consider a learner that trains a logistic regression classifier using cross-entropy loss. For the population, we consider strategies that consist of misreporting a single feature. We choose this to be either 'age' or 'number of dependents'; 'age' is the most important feature for the classification problem, and 'number of dependents' is the least important one (see Figure 5 in the Appendix). Contrasting the two strategies allows us to vary the impact of a strategy on the learner in an isolated and systematic way.

Alignment as a proxy for the benefit of coordination. We investigate empirically how the alignment metric Φ in Theorem 5 relates to the benefit of coordination. We instantiate the agents' utility as

$$u\big((x,y);\theta\big) = \ell((x,y),\theta) - \lambda \cdot \|\theta\|^2, \tag{5}$$

where ℓ is the loss of the learner and the regularization term $\lambda \in [0,1]$ controls the alignment between the learner's and the agents' objectives. For $\lambda = 0$ we get a zero-sum game. The larger λ the more different the learner's and the agents' objectives are.

Note that, in this setting, the linearity and the strong concavity of our theory are not satisfied. We are still interested in investigating to what extent the alignment metric serves as a useful proxy for the benefit of coordination. In Figure 2 use λ to vary alignment and plot Φ against the benefit of coordination B. We find that $\mathrm{B} < \Phi$, which is in line with our theory. Furthermore, alignment correlates with the benefit of coordination and accurately predicts that one strategy is more effective than the other. Recall that we restrict the strategy space to the modification of a single feature; the solid and the dashed line compare two different settings.

Cost of steering and why larger collectives can be worse off. Next, we illustrate the challenges that come with large collectives in a misaligned setup. For this purpose we consider the zero-sum setting with $\lambda = 0$ and a mixed-population composed of a collective and non-strategic agents. The collective implements the optimal size-aware strategy $h^\sharp_\alpha$ for modifying each of the two features. We approximate this optimal strategy using gradient descent with learning rate 0.01 and 250 epochs.

In Figure 3 we visualize the collective utility as a function of the collective size α. We see that the collective utility is maximized as $\alpha \to 0$ and decreases with size. The gap at $\alpha = 0$ shows the benefit of strategic data reporting against a fixed model. Small α is an advantage in the zero-sum case as agents have diminishing influence on the learner and they can move almost independently of θ. For larger collectives it becomes increasingly hard to realize a benefit because the model responds to the agents' actions and an equivalent change to the feature has a larger effect on the learner's loss. This counter-force in the case of conflicting utilities is the reason why the collective utility may decrease with the collective size.

Scaling up a fixed strategy. Finally, we verify Proposition 6 empirically. For simplicity, we assume the learner performs binary classification to predict whether a person experienced 90 days past due

https://doi.org/10.52202/085713-2757

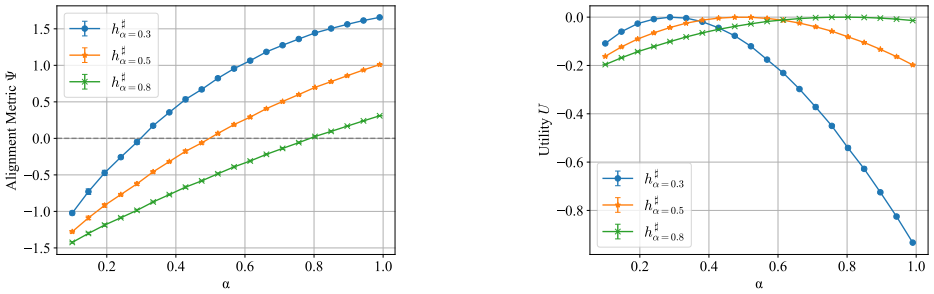


Figure 4: *Change in alignment metric (left) and utility (right) with collective size, for three fixed strategies.* The utility is non-monotonic in size, and the sign of the alignment metric Ψ accurately predicts whether it is worth scaling up a strategy or not. We consider the setting in (6) and evaluate three different strategies, corresponding to the optimal size-aware strategy $h^\sharp_\alpha$ at $\alpha \in \{0.3, 0.5, 0.8\}$.

delinquency or worse. The collective is uniformly sampled from the data with label 1 and aims to maximize their utility

$$u_i((x_i, y_i), \theta) = -\left\|\theta - \theta_{\text{target}}\right\|_2^2, \tag{6}$$

where θ_{target} a fixed target model that the collective would like to achieve (see Appendix C.2 for details). We again assume the collective can modify individual features and consider three strategies $h^\sharp_{\alpha=0.3}$, $h^\sharp_{\alpha=0.5}$, and $h^\sharp_{\alpha=0.8}$, where each strategy is optimal for a given collective size $\alpha \in \{0.3, 0.5, 0.8\}$. We refer to Appendix C.2 for details on how the strategies are computed.

In Figure 4 we illustrate the collective utility and the alignment metric Ψ as defined in Proposition 6 for different collective sizes α. We observe that, in accordance with our theory, when the alignment metric is positive, the average utilities of the collective decrease as the sizes of the collective increase. Analogously, when the alignment is negative, the average utilities of the collective increase with collective size. Notably, as α approaches the assumed sizes $0.3, 0.5$ and 0.8, the utility converges toward 0, indicating that the model closely matches the target θ_{target}. Beyond this point, however, the collective "over-pushes" the model, leading to a decrease in utility and a positive Ψ.

7 Conclusion

We study look-ahead reasoning as a new aspect of strategic reasoning on learning platforms. While traditional analyses of strategic classification treat users as reacting independently to a fixed model, look-ahead reasoning highlights that users' incentives and actions are inherently interdependent—each agent's actions influence future model deployments, and thus the utility of other agents in the population. Within this broad theme, we find that higher-order reasoning accelerates convergence toward equilibrium but does not improve individuals' long-run outcomes, suggesting that attempts to "outsmart" others may offer only transient advantages. In contrast, collective reasoning—where users coordinate their behavior through their shared impact on the model—allows the agents to steer the model towards a desirable state. Our results show that this can be a very effective lever for the collective when the loss of the learner and the utility of the population are appropriately aligned. However, for conflicting objectives, we find that the excessive steering power that comes with larger collectives can prevent large utility gains.

A central goal of our work was to provide a unifying framework to contrast selfish reasoning with collective reasoning when contesting machine learning predictions. There are several natural extensions of our work. One is investigating look-ahead reasoning under imperfect information. As common in economic models, we assumed agents determine their actions under perfect information. In our case, this concerns knowledge about the learner's loss function and the population's data distribution. However, in practice this information needs to be estimated from finite data which raises questions of statistical complexity. A related question would be to study how model misspecifications, estimation errors, or imperfect coordination impact outcomes in look-ahead reasoning.

More broadly, we hope our work can support recent discussions concerning data poisoning and multi-party adversarial attacks [e.g., Nestaas et al., 2025], and open up new pathways to transfer insights from strategic classification to algorithmic collective action, and vice versa.

https://doi.org/10.52202/085713-2757

Acknowledgements

Celestine Mendler-Dünner acknowledges the financial support of the Hector Foundation. The work was conducted during an internship of Haiqing Zhu at the Max Planck Institute for Intelligent Systems, Tübingen.

References

Joachim Baumann and Celestine Mendler-Dünner. Algorithmic collective action in recommender systems: promoting songs by reordering playlists. *Advances in Neural Information Processing Systems*, 37:119123–119149, 2024.

Yahav Bechavod, Chara Podimata, Steven Wu, and Juba Ziani. Information discrepancy in strategic learning. In *International Conference on Machine Learning*, pages 1691–1715, 2022.

Omri Ben-Dov, Jake Fawkes, Samira Samadi, and Amartya Sanyal. The role of learning algorithms in collective action. In *International Conference on Machine Learning*, pages 3443–3461, 2024.

Gavin Brown, Shlomi Hod, and Iden Kalemaj. Performative prediction in a stateful world. In *International Conference on Artificial Intelligence and Statistics*, volume 151, pages 6045–6061, 2022.

Michael Brückner and Tobias Scheffer. Stackelberg games for adversarial prediction problems. In *ACM SIGKDD*, pages 547–555, 2011.

Julie Yujie Chen. Thrown under the bus and outrunning it! the logic of didi and taxi drivers' labour and activism in the on-demand economy. *New Media & Society*, 20(8):2691–2711, 2018.

Yiling Chen, Yang Liu, and Chara Podimata. Learning strategy-aware linear classifiers. *Advances in Neural Information Processing Systems*, 33:15265–15276, 2020.

Jinshuo Dong, Aaron Roth, Zachary Schutzman, Bo Waggoner, and Zhiwei Steven Wu. Strategic classification from revealed preferences. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 55–70, 2018.

Dmitriy Drusvyatskiy and Lin Xiao. Stochastic optimization with decision-dependent distributions. *Mathematics of Operations Research*, 48(2):954–998, 2023.

Etienne Gauthier, Francis Bach, and Michael I. Jordan. Statistical collusion by collectives on learning platforms. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pages 18897–18919, 2025.

Ganesh Ghalme, Vineet Nair, Itay Eilat, Inbal Talgam-Cohen, and Nir Rosenfeld. Strategic classification in the dark. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pages 3672–3681, 2021.

Moritz Hardt and Celestine Mendler-Dünner. Performative Prediction: Past and Future. *Statistical Science*, 40(3):417 – 436, 2025.

Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou, and Mary Wootters. Strategic classification. In *ACM Conference on Innovations in Theoretical Computer Science*, page 111–122, 2016.

Moritz Hardt, Meena Jagadeesan, and Celestine Mendler-Dünner. Performative power. In *Advances in Neural Information Processing Systems*, volume 35, pages 22969–22981, 2022.

Moritz Hardt, Eric Mazumdar, Celestine Mendler-Dünner, and Tijana Zrnic. Algorithmic collective action in machine learning. In *International Conference on Machine Learning*, 2023.

Meena Jagadeesan, Celestine Mendler-Dünner, and Moritz Hardt. Alternative microfoundations for strategic classification. In *International Conference on Machine Learning*, volume 139, pages 4687–4697, 2021.

Terri Kneeland. Identifying higher-order rationality. *Econometrica*, 83(5):2065–2079, 2015.

 https://doi.org/10.52202/085713-2757

Bogdan Kulynych, Rebekah Overdorf, Carmela Troncoso, and Seda Gürses. Pots: protective optimization technologies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, page 177–188. Association for Computing Machinery, 2020. ISBN 9781450369367.

Sagi Levanon and Nir Rosenfeld. Generalized strategic classification and the case of aligned incentives. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162, pages 12593–12618. PMLR, 2022.

Celestine Mendler-Dünner, Juan Perdomo, Tijana Zrnic, and Moritz Hardt. Stochastic optimization for performative prediction. In *Advances in Neural Information Processing Systems*, volume 33, pages 4929–4939, 2020.

Rosemarie Nagel. Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, 85(5):1313–1326, 1995.

Adhyyan 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.

Fredrik Nestaas, Edoardo Debenedetti, and Florian Tramèr. Adversarial search engine optimization for large language models. In *The Thirteenth International Conference on Learning Representations*, 2025.

Juan Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *International Conference on Machine Learning*, pages 7599–7609, 2020.

Chara Podimata. Incentive-aware machine learning; robustness, fairness, improvement & causality. *arXiv preprint arXiv:2505.05211*, 2025.

Rideshare Drivers United. `https://www.drivers-united.org/`.

Dorothee Sigg, Moritz Hardt, and Celestine Mendler-Dünner. Decline now: A combinatorial model for algorithmic collective action. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2025.

Arianna Tassinari and Vincenzo Maccarrone. Riders on the storm: Workplace solidarity among gig economy couriers in italy and the uk. *Work, Employment and Society*, 34(1):35–54, 2020.

Nicholas Vincent, Hanlin Li, Nicole Tilly, Stevie Chancellor, and Brent Hecht. Data leverage: A framework for empowering the public in its relationship with technology companies. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, page 215–227, 2021.

Tijana Zrnic, Eric Mazumdar, Shankar Sastry, and Michael Jordan. Who leads and who follows in strategic classification? *Advances in Neural Information Processing Systems*, 34:15257–15269, 2021.

A Proofs

A.1 Auxiliary results

Lemma 9. *Let $\mathcal{W}$ denote the Wasserstein-1 distance, and $\sum_{i=1}^n \alpha_i \geq 0$ with $\alpha_i \geq 0$, then*

$$\mathcal{W}\left(\sum_{i=1}^n \alpha_i \mathcal{D}_i, \sum_{i=1}^n \alpha_i \mathcal{D}'_i\right) \leq \sum_{i=1}^n \alpha_i \mathcal{W}(\mathcal{D}_i, \mathcal{D}'_i).$$

Proof. By definition, the Wasserstein-1 distance can be written as

$$\min_{\mu(X,Y)} \mathbb{E}_{\mu(X,Y)} \left[\|X - Y\|\right],$$

where μ is the joint distribution of X, Y and $X \sim \sum_{i=1}^n \alpha_i \mathcal{D}_i$, $Y \sim \sum_{i=1}^n \alpha_i \mathcal{D}'_i$. Consider the measures μ_i defined as

$$\mu_i = \min_{\mu(X_i,Y_i)} \mathbb{E}_{\mu(X_i,Y_i)} \left[\|X_i - Y_i\|\right],$$

where $X_i \sim \mathcal{D}_i, Y_i \sim \mathcal{D}'_i$. Then, with $\hat{\mu} = \sum_i \alpha_i \mu_i$, we can notice that

$$\mathcal{W}\left(\sum_{i=1}^n \alpha_i \mathcal{D}_i, \sum_{i=1}^n \alpha_i \mathcal{D}'_i\right) \leq \mathbb{E}_{\hat{\mu}} \left[\|X - Y\|\right] = \sum_{i=1}^n \alpha_i \mathbb{E}_{\mu_i} \left[\|X - Y\|\right] = \sum_{i=1}^n \alpha_i \mathcal{W}(\mathcal{D}_i, \mathcal{D}'_i),$$

where the first inequality follows from the minimization property of the Wasserstein-1 distance. □

Below we state a bound on the benefit of coordination that does not assume linearity (Assumption 1).

Theorem 5'. *Let each strategy $h(\eta)$ be represented by a parameter vector $\eta \in \mathbb{R}^d$. Suppose $U(h(\eta))$ is γ-strongly concave on η. Then, we have*

$$0 \leq \mathrm{B} \leq \frac{1}{2\gamma} \cdot \left\|\nabla_\eta \nabla_\theta \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\ell(z, \theta^*)\right] (\mathbf{H}^\star)^{-1} \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta u(z, \theta^*)\right]\right\|_2^2.$$

with $\mathbf{H}^\star := \mathbb{E}_{z \in \mathcal{D}_{h^*}} \left[\nabla^2_{\theta,\theta} \ell(z, \theta^*)\right]$ and θ^* denoting the stable point corresponding to the selfish strategy h^*.

A.2 Proof of Theorem 3

The key step is to prove Lemma 10 below. The claim in the theorem follows by combining Lemma 10 with Theorem 3.5 in Perdomo et al. [2020].

Lemma 10. *Let α_k be the portion of the population with cognitive level k. Suppose $\mathcal{D}_1(\theta)$ is ϵ-sensitive and the loss ℓ is γ-strongly convex and β-smooth in z, θ. Then the distribution map $\overline{\mathcal{D}}(\theta) := \sum_{k=1}^\infty \alpha_k \mathcal{D}_k(\theta)$ is $\bar{\epsilon}$-sensitive with*

$$\bar{\epsilon} = \sum_{k=1}^\infty \alpha_k \left(\frac{\epsilon\beta}{\gamma}\right)^{k-1} \epsilon.$$

Proof. Denote $\theta^k := \mathcal{A}(\mathcal{D}_1(\theta^{k-1}))$ and $\theta^0 = \theta$. Similarly, $\phi^k := \mathcal{A}(\mathcal{D}_1(\phi^{k-1}))$ and $\phi^0 = \phi$. From Equation (1), we can see that $\mathcal{D}_k(\theta^0) = \mathcal{D}_1(\theta^{k-1})$. Consider the map $\mathcal{D}_k$; then, we have the recursion:

$$\mathcal{W}(\mathcal{D}_k(\theta), \mathcal{D}_k(\phi)) = \mathcal{W}(\mathcal{D}_1(\theta^{k-1}), \mathcal{D}_1(\phi^{k-1})) \leq \epsilon \left\|\theta^{k-1} - \phi^{k-1}\right\|_2.$$

By Perdomo et al. [2020], Theorem 3.5, we also have

$$\left\|\theta^{k-1} - \phi^{k-1}\right\|_2 \leq \left(\frac{\epsilon\beta}{\gamma}\right)^{k-1} \|\theta - \phi\|_2.$$

Finally, we notice that

$$\mathcal{W}\left(\sum_{k=1}^\infty \alpha_k \mathcal{D}_k(\theta), \sum_{k=1}^\infty \alpha_k \mathcal{D}_k(\phi)\right) \leq \sum_{k=1}^\infty \alpha_k \mathcal{W}(\mathcal{D}_k(\theta), \mathcal{D}_k(\phi)) \leq \sum_{k=1}^\infty \alpha_k \left(\frac{\epsilon\beta}{\gamma}\right)^{k-1} \epsilon \|\theta - \phi\|_2,$$

where the first inequality follows from Lemma 9. □

https://doi.org/10.52202/085713-2757

A.3 Proof of Corollary 1

We start from Theorem 3. Define the contraction factor

$$\rho_\alpha = \left(\sum_{k=1}^{\infty} \left(\frac{\epsilon\beta}{\gamma} \right)^k \alpha_k \right).$$

It can be seen that if ρ_α is positive for some α, it holds that the dynamics will converge to the equilibrium for any α with $\sum_{k=1}^{\infty} \alpha_k = 1$. Similarly, if it is zero, this holds so for any α. Thus, a simple contraction argument shows that the trajectory converges to the same stable point independent of α. The same holds for the special case $\alpha_k = 1$. At this point, no agent is moving and thus the equilibrium strategies $h_{\theta^*}^{(k)}$ are identical, and so are the utilities:

$$h_{\theta^*}^{(k)} = h_{\theta^*}^{(k')} \quad \Rightarrow \quad U(h_{\theta^*}^{(k)}) = U(h_{\theta^*}^{(k')}).$$

A.4 Proof of Proposition 4

Since (h^*, θ^*) is the performatively stable point, by definition, it is clear that

$$\mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta \ell(z, \theta^*) \right] = 0.$$

Therefore, since $u = c \cdot \ell$, we can conclude that $\mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta u(z, \theta^*) \right] = c \cdot \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta \ell(z, \theta^*) \right] = 0$. Then, by Theorem 5', we have $\mathrm{B} = 0$.

A.5 Proof of Theorem 5

We use the parameterization in Assumption 1. The optimal strategy is defined as $h^\sharp := h(\eta^\sharp)$ where $\eta^\sharp := \arg\max_\eta U(h(\eta))$. The performatively stable point $(h^*, \theta^*) = (h(\eta^*), \theta^*)$ is the point such that

$$\eta^* = \arg\max_\eta \mathbb{E}_{z \sim \mathcal{D}_{h(\eta)}} \left[u(z, \theta^*) \right],$$
$$\theta^* = \arg\min_\theta \mathbb{E}_{z \sim \mathcal{D}_{h(\eta^\star)}} \left[\ell(z, \theta) \right].$$

We consider a hypothetical mixture population which is represented as $\alpha \mathcal{D}_{h^\sharp} + (1-\alpha)\mathcal{D}_{h^*} = \mathcal{D}_{h(\alpha\eta^\sharp + (1-\alpha)\eta^*)}$. For this mixture population, $\alpha = 0$ indicates the equilibrium of selfish actions. This means that, if we fix the learner's action to be θ^*, the action $h_\alpha = h(\alpha\eta^\sharp + (1-\alpha)\eta^*)$ will maximize the population's expected utility only when $\alpha = 0$. Formally, consider the function $\iota(\alpha) = \mathbb{E}_{z \sim \mathcal{D}_{\alpha h^\sharp + (1-\alpha)h^*}} \left[u(z, \theta^*) \right] = \mathbb{E}_{z \sim \mathcal{D}_{h(\alpha\eta^\sharp + (1-\alpha)\eta^*)}} \left[u(z, \theta^*) \right]$. By Assumption 1, we must have

$$\left. \frac{\partial \iota}{\partial \alpha} \right|_{\alpha=0} = \mathbb{E}_{z \sim \mathcal{D}_{h^\sharp}} \left[u(z, \theta^*) \right] - \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[u(z, \theta^*) \right] = 0.$$

At the stable point (h^*, θ^*), we also have that $\mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta \ell(z, \theta^*) \right] = 0$. Consider the function $e(\alpha) = U\left(\alpha h^\sharp + (1-\alpha)h^* \right) = U\left(h(\alpha\eta^\sharp + (1-\alpha)\eta^*) \right)$; we have

$$\begin{aligned}
\left. \frac{\partial e}{\partial \alpha} \right|_{\alpha=0} &= \mathbb{E}_{z \sim \mathcal{D}_{h^\sharp}} \left[u(z, \theta^*) \right] - \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[u(z, \theta^*) \right] \\
&\quad - \left\langle \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta u(z, \theta^*) \right], \mathbb{E}_{z \sim \mathcal{D}_{h^\sharp}} \left[\nabla_\theta \ell(z, \theta^*) \right] - \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta \ell(z, \theta^*) \right] \right\rangle_{(\mathbf{H}^\star)^{-1}} \\
&= - \left\langle \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta u(z, \theta^*) \right], \mathbb{E}_{z \sim \mathcal{D}_{h^\sharp}} \left[\nabla_\theta \ell(z, \theta^*) \right] - \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta \ell(z, \theta^*) \right] \right\rangle_{(\mathbf{H}^\star)^{-1}} \\
&= - \left\langle \mathbb{E}_{z \sim \mathcal{D}_{h^*}} \left[\nabla_\theta u(z, \theta^*) \right], \mathbb{E}_{z \sim \mathcal{D}_{h^\sharp}} \left[\nabla_\theta \ell(z, \theta^*) \right] \right\rangle_{(\mathbf{H}^\star)^{-1}}.
\end{aligned}$$

Finally, the result follows from the PL-inequality.

A.6 Proof of Theorem 5'

For notational simplicity, we set $f(\eta, \theta) := \mathbb{E}_{z \sim \mathcal{D}_{h(\eta)}} \left[u(z; \theta) \right]$ and $g(\eta, \theta) := \mathbb{E}_{z \sim \mathcal{D}_{h(\eta)}} \left[\ell(z; \theta) \right]$. Recall that

$$U(h(\eta)) = f(\eta, \mathcal{A}(\mathcal{D}_{h(\eta)})) = \mathbb{E}_{z \sim \mathcal{D}_{h(\eta)}} \left[u(z, \mathcal{A}(\mathcal{D}_{h(\eta)})) \right],$$

where the first argument in f only applies in distribution that is taken against and the second argument is corresponding to the second argument of u. Then, by the implicit function theorem we have

$$\nabla_\eta U(h(\eta)) = \nabla_\eta f(\eta, \mathcal{A}(\mathcal{D}_{h(\eta)})) = \nabla_1 f(\eta, \mathcal{A}(\mathcal{D}_{h(\eta)})) - \nabla^2_{1,2} g(\eta, \mathcal{A}(\mathcal{D}_{h(\eta)})) \left[\nabla^2_{2,2} g(\eta, \mathcal{A}(\mathcal{D}_{h(\eta)}))\right]^{-1} \nabla_2 f(\eta, \mathcal{A}(\mathcal{D}_{h(\eta)})),$$

where ∇_1 and ∇_2 denote the gradient operator on the first/second argument of the function. For the equilibrium of the selfish strategy $(h(\eta^*), \theta^*)$, we must have $\nabla_1 f(\eta, \mathcal{A}(\mathcal{D}_h)) = 0$. Let $\eta^\sharp := \arg\max_\eta U(h(\eta))$, by the PL-inequality, we obtain that

$$U\left(h(\eta^\sharp)\right) - U\left(h(\eta^*)\right) \leq \frac{1}{2\gamma} \cdot \left\| \nabla^2_{1,2} g(\eta^*, \mathcal{A}(\mathcal{D}_{h(\eta^*)})) \left[\nabla^2_{2,2} g(\eta^*, \mathcal{A}(\mathcal{D}_{h(\eta^*)}))\right]^{-1} \nabla_2 f(\eta^*, \mathcal{A}(\mathcal{D}_{h(\eta^*)})) \right\|_2^2.$$

A.7 Proof of Proposition 6

Consider the derivative of U_α with respect to variable α,

$$\frac{\partial U_\alpha}{\partial \alpha} = \frac{\partial \mathbb{E}_{z\sim\mathcal{D}_h}\left[u(z;\theta^*_\alpha)\right]}{\partial \alpha} = \frac{\partial \mathbb{E}_{z\sim\mathcal{D}_h}\left[u(z;\theta^*_\alpha)\right]}{\partial \theta} \cdot \frac{\partial \theta^*_\alpha}{\partial \alpha}.$$

Next, we notice that

$$\nabla_\theta \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right] = 0,$$

where we can further write the left-hand side as

$$\nabla_\theta \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right] = \alpha \cdot \mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right] + (1-\alpha) \cdot \mathbb{E}_{z\sim\mathcal{D}_0}\left[\nabla_\theta \ell(z;\theta)\right] = 0.$$

Therefore, $\mathbb{E}_{z\sim\mathcal{D}_0}\left[\nabla_\theta \ell(z;\theta)\right] = -\frac{\alpha}{1-\alpha} \cdot \mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right]$. Then, we consider the term $\frac{\partial \theta^*_\alpha}{\partial \alpha}$. By the implicit function theorem, with $\alpha > 0$, we have

$$\begin{aligned}
\frac{\partial \theta^*_\alpha}{\partial \alpha} &= -\left(\nabla^2_{\theta,\theta} \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right]\right)^{-1} \left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right] - \mathbb{E}_{z\sim\mathcal{D}_0}\left[\nabla_\theta \ell(z;\theta)\right]\right) \\
&= -\frac{1}{1-\alpha}\left(\nabla^2_{\theta,\theta} \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right]\right)^{-1} \left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right]\right).
\end{aligned}$$

Therefore, we can finally write

$$\frac{\partial U_\alpha}{\partial \alpha} = -\frac{1}{1-\alpha}\left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta u(z;\theta^*_\alpha)\right]\right)^T \left(\nabla^2_{\theta,\theta} \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right]\right)^{-1} \left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right]\right).$$

A.8 Proof of Proposition 7

The proof is the same as the proof of Proposition 6 up to a use of the envelope theorem. For completeness, we restate the proof here and highlight the use of the envelope theorem. For notational simplicity, we abbreviate $h^\sharp_\alpha$ as h. Consider the derivative of U^*_α with respect to α,

$$\frac{\partial U^*_\alpha}{\partial \alpha} = \frac{\partial \mathbb{E}_{z\sim\mathcal{D}_h}\left[u(z;\theta^*_\alpha)\right]}{\partial \alpha} = \frac{\partial \mathbb{E}_{z\sim\mathcal{D}_h}\left[u(z;\theta^*_\alpha)\right]}{\partial \theta} \cdot \frac{\partial \theta^*_\alpha}{\partial \alpha},$$

where the second equality follows from the implicit function theorem and the envelope theorem. Next, we notice that

$$\nabla_\theta \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right] = 0,$$

where we can further write the left-hand side as

$$\nabla_\theta \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right] = \alpha \cdot \mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right] + (1-\alpha) \cdot \mathbb{E}_{z\sim\mathcal{D}_0}\left[\nabla_\theta \ell(z;\theta)\right] = 0.$$

Therefore, $\mathbb{E}_{z\sim\mathcal{D}_0}\left[\nabla_\theta \ell(z;\theta)\right] = -\frac{\alpha}{1-\alpha} \cdot \mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right]$. Consider the term $\frac{\partial \theta^*_\alpha}{\partial \alpha}$. By the implicit function theorem, with $\alpha > 0$, we have

$$\begin{aligned}
\frac{\partial \theta^*_\alpha}{\partial \alpha} &= -\left(\nabla^2_{\theta,\theta} \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right]\right)^{-1} \left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right] - \mathbb{E}_{z\sim\mathcal{D}_0}\left[\nabla_\theta \ell(z;\theta)\right]\right) \\
&= -\frac{1}{1-\alpha}\left(\nabla^2_{\theta,\theta} \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right]\right)^{-1} \left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right]\right).
\end{aligned}$$

We finally write

$$\frac{\partial U^*_\alpha}{\partial \alpha} = -\frac{1}{1-\alpha}\left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta u(z;\theta^*_\alpha)\right]\right)^T \left(\nabla^2_{\theta,\theta} \mathbb{E}_{z\sim\alpha\mathcal{D}_h + (1-\alpha)\mathcal{D}_0}\left[\ell(z;\theta)\right]\right)^{-1} \left(\mathbb{E}_{z\sim\mathcal{D}_h}\left[\nabla_\theta \ell(z;\theta)\right]\right).$$

https://doi.org/10.52202/085713-2757

A.9 Proof of Proposition 8

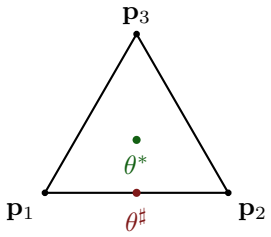
By Lemma 10, the sensitivity of the mixture distribution can be computed as

$$\sum_{k=1}^{\infty} \alpha_k \left(\frac{\epsilon\beta}{\gamma}\right)^{k-1} \epsilon = (1-\alpha)\epsilon,$$

where $\alpha_1 = 1 - \alpha$ and $\alpha_0 = \alpha$. Combining this with Theorem 3.5 in Perdomo et al. [2020] yields the result.

B Example to illustrate Theorem 5

Consider a toy setting to derive a closed-form expression for the benefit of coordination. Assume the learner estimates the centroid θ of a distribution supported on three anchor points $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$—the corners of an equilateral triangle. The collective applies a strategy h that moves their data point z to one of the three anchors with probability (w_1, w_2, w_3), respectively, $\sum_i w_i = 1$ and $w_i \geq 0$. This implies a distribution $\mathcal{D}_h = \sum_i w_i \delta_{\mathbf{p}_i}$ where the strategy determines $\mathbf{w}$. The learner minimizes the squared loss; the solution is $\mathcal{A}(\mathcal{D}_h) = \sum_i w_i x_i$. The collective, on the other hand, prefers the centroid to lie between $\mathbf{p}_1$ and $\mathbf{p}_2$, and maximizes



$$u(z, \theta) = -\|\mathbf{p}_1 - \theta\|_2^2 - \|\mathbf{p}_2 - \theta\|_2^2 - \|z - \theta\|_2^2.$$

When the collective distributes mass uniformly, i.e., $h^* = (1/3, 1/3, 1/3)$, the centroid $\theta^* = \frac{1}{3}\sum_{i=1}^{3} \mathbf{p}_i$ is a performatively stable point. There is no incentive for either party to change their strategy since they are both best-responding to the current state. However, a look-ahead collective would prefer $h^\sharp = (1/2, 1/2, 0)$, as they are aware that they could collectively shift the model θ away from this stable state further down in the triangle. This leads to a look-ahead optimal point $\theta^\sharp = \frac{1}{2}(\mathbf{p}_1 + \mathbf{p}_2)$, which deviates from θ^*.

In this example, Assumption 1 is satisfied. Moreover, $\Phi^2 > 0$ since $\Phi = -2r^2$ where $r := \|\mathbf{p}_3\|_2$. The benefit of coordination $\mathrm{B} = U(h^\sharp) - U(h^*) = \frac{3}{4}r^2$ is strictly positive as long as the anchor points are appropriately spaced. One can check that U is $2r^2$-strongly concave, and Theorem 5 can be verified since $\mathrm{B} \leq \frac{\Phi^2}{2\gamma} = r^2$ which is tight up to a factor 1/4 coming from the slack in the strong concavity assumption.

C Simulations

C.1 Feature importance

In Figure 5 we show the importance of different features in the credit-scoring simulator. We simulate modifications to feature i by replacing z_i with a value z_i', which we sample independently from a standard normal distribution. Subsequently we train a logistic regression classifier on the modified data. The values of the bars indicate the *drop* of test accuracy compared to the baseline classifier trained without misreporting. The error bars indicate one standard deviations over 10 different train-test splits.

C.2 Binary prediction setup

Consider the case in which the collective wants to implement a fixed strategy, which modifies their strategic features "age" and "number of dependents" as

$$h_\eta((x_i, y_i)) = (x_{i,S} + \eta \cdot \mu_{0,S}, y_i),$$

where $\mu_{0,S}$ is the mean of the strategic features when the label equals 0. Since the features are all centered across the whole dataset, this transformation can be interpreted geometrically as translating the samples with label 1 along the direction of μ_0, moving them toward the center of the distribution with label 0 in feature space.

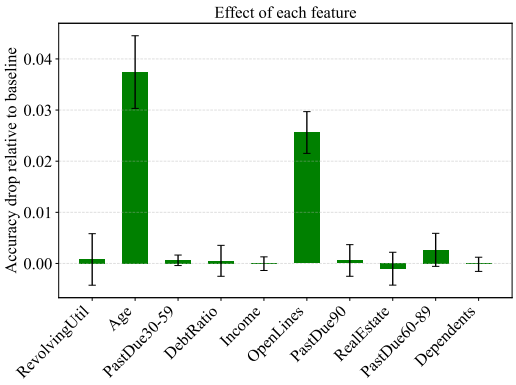


Figure 5: Accuracy drop against modifying individual features.

We run simulations with three strategies $h^{\sharp}_{\alpha=0.3}$, $h^{\sharp}_{\alpha=0.5}$, and $h^{\sharp}_{\alpha=0.8}$, where each strategy optimally chooses η for a given collective size $\alpha \in \{0.3, 0.5, 0.8\}$. The target model θ_{target} we used in the simulation was fixed to be the model produced by the strategy with parameters $(\eta = 0.5, \alpha = 0.3)$. For each $\alpha \in \{0.3, 0.5, 0.8\}$, we can solve for an optimal η_α such that the resulting model coincides with θ_{target}. In other words, for all three collective sizes, there exists a manipulation strength η_α that enables the collective to attain maximum utility (equal to 0). This ensures that comparisons across different values of α are well-defined.

https://doi.org/10.52202/085713-2757

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction accurately reflect the scope and contributions.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Limitations are discussed where necessary.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: The proofs are complete. They are either in the main text or the supplementary material.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The source code of the experiments are included in the supplementary material.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

https://doi.org/10.52202/085713-2757

Answer: [Yes]

Justification: Data is available publicly/randomly generated (in the source code). The code is in the supplementary material.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: The setup is discussed in the simulation section. The implementation is in the supplementary material.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Error bars and other setup info are discussed in the appendix.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).

- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: All simulations in this paper are executable with a standard personal laptop.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: [NA]

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Our paper mainly focuses on the theoretical contributions. The societal implications/explanations of the results are discussed in the paper.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

https://doi.org/10.52202/085713-2757

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: [NA]

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The data used in the experimental section is open for public and research use.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: [NA]

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: No experiments involving human participants are conducted.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: No experiments involving human participants are conducted.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

 https://doi.org/10.52202/085713-2757

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: LLM is only used for wording and grammar checking.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.