State Dependent Performative Prediction with Stochastic Approximation

Qiang Li **Hoi-To Wai**
Dept. of SEEM, The Chinese University of Hong Kong, Hong Kong SAR of China.

Abstract

This paper studies the performative prediction problem which optimizes a stochastic loss function with data distribution that depends on the decision variable. We consider a setting where the agent(s) provides samples adapted to both the learner's and agent's previous states. The samples are then used by the learner to update his/her state to optimize a loss function. Such closed loop update dynamics is studied as a state dependent stochastic approximation (SA) algorithm, which is shown to find a fixed point known as the *performative stable* solution. Our setting captures the unforgetful nature and reliance on past experiences of agents. Our contributions are three-fold. First, we present a framework for modeling state dependent performative prediction with *biased* stochastic gradients driven by a controlled Markov chain whose transition probability depends on the learner's state. Second, we present a new finite-time performance analysis of the SA algorithm. We show that the expected squared distance to the performative stable solution decreases as $\mathcal{O}(1/k)$, where k is the iteration number. Third, numerical experiments verify our findings.

1 INTRODUCTION

Many supervised learning algorithms are built around the assumption that learners can obtain samples from a static distribution *independent* of the state of the learner and/or the agent who provides the sample. This assumption is reasonable for static tasks such as image classification. Oftentimes, it simplifies the design and analysis of algorithms.

On the other hand, in certain applications agents can be *performative* where the samples are drawn from a *decision-dependent* distribution. This is relevant to the framework of strategic classification [Hardt et al., 2016, Cai et al., 2015, Kleinberg and Raghavan, 2020]. For instance, when training a classifier for loan applications, given the classifier published by the *learner* (bank), the *agent(s)* (loan applicants) may manipulate their profile prior to the submission, e.g., by spending more with credit cards, making unnatural transactions, etc., in order to increase their chance of successful application. The latter may affect the convergence properties, or even the stability of learning algorithms.

Earlier works [Bartlett, 1992, Quiñonero-Candela et al., 2009] studied the effects of exogenous changes with shifts in data distribution. To capture the effects of data distribution shifts during learning, Perdomo et al. [2020] studied a scenario where the *learner* is interested in the following *performative prediction* problem:

$$\min_{\theta \in \mathbb{R}^d} \; V(\theta) = \mathbb{E}_{z \sim \mathcal{D}(\theta)} \big[\ell(\theta; z) \big], \tag{1}$$

where $\ell(\theta; z)$ is the loss function given the sample $z \in \mathsf{Z}$. The loss function is strongly-convex with respect to (w.r.t.) the parameter $\theta \in \mathbb{R}^d$, and the gradient map $\nabla_\theta \ell(\theta; z)$ is Lipschitz continuous w.r.t. z, θ. In addition, the distribution $\mathcal{D}(\theta)$ on Z is parameterized by the decision vector θ, which captures the distribution shift due to the learner's state. Problem (1) finds a parameter θ which minimizes the expected loss that takes care of the decision-dependent distribution.

Despite the strong convexity of $\ell(\theta; z)$, problem (1) is non-convex in general due to coupling with θ in the data distribution $\mathcal{D}(\theta)$. As a remedy, Perdomo et al. [2020] studied population-based algorithms that converge to a *performative stable* point, θ_{PS}, which is a fixed point to the system $\theta = \arg\min_{\theta' \in \mathbb{R}^d} \mathbb{E}_{z \sim \mathcal{D}(\theta)} \big[\ell(\theta'; z) \big]$. Along the same line, Mendler-Dünner et al. [2020] analyzed stochastic algorithms which deploy minibatches of i.i.d. samples from the shifted distribution at each

Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS) 2022, Valencia, Spain. PMLR: Volume 151.

iteration, Izzo et al. [2021], Miller et al. [2021] studied gradient estimation techniques and developed algorithms that converge to an optimal solution of (1) through introducing a gradient correction term (also see [Munro, 2020] which has considered a related setting), Drusvyatskiy and Xiao [2020] studied the stability of proximal stochastic gradient algorithms (and their variants), Brown et al. [2020] studied population-based algorithms where the state dependent distribution is updated iteratively, Zrnic et al. [2021] studied the impact of update frequencies of learners/agents on the regret to equilibrium.

This paper studies the convergence of stochastic algorithms for (1) where only one data sample (or a minibatch of samples) is required at each iteration. Specifically, we focus on a setting where agent(s) rely on their past experiences while adapting to the data distribution shifts due to the model deployed by the learner. Meanwhile, the learner follows a greedy deployment scheme similar to Mendler-Dünner et al. [2020] where s/he deploys the most recent model after each update round. In other words, the learner's and agent(s)' states are co-evolving dynamically. The closed-loop algorithm can be studied as a *state dependent stochastic approximation (SA)* algorithm. In contrast to Mendler-Dünner et al. [2020], Drusvyatskiy and Xiao [2020] where only the learner's state is incrementally updated and the agent draws i.i.d. samples from the distribution shifted by the learner's state, in our setting, the agent's state evolves according to a *controlled Markov chain (MC)* whose stationary distribution is the shifted distribution.

Our study is motivated by the *stateful (or unforgetful) nature* of the agents who depend on past experiences when adapting to a shifted target data distribution. For example, a loan applicant may take months to build up his/her credit history to adapt to changes in the published classifier. Several questions naturally arise from such dynamical performative prediction problems: *will the stochastic algorithm converge to a performative stable point similar to Mendler-Dünner et al. [2020], Drusvyatskiy and Xiao [2020]? what is the sample complexity?* This paper addresses these questions as we make the following contributions:

- We develop a *fully state dependent performative prediction* framework which extends the model and analysis in [Mendler-Dünner et al., 2020, Drusvyatskiy and Xiao, 2020]. The proposed extension relies on a state dependent stochastic approximation (SA) algorithm with noise originating from a controlled Markov chain [cf. Algorithm 1].
- Our main result consists of a finite-time convergence analysis of the state dependent SA algorithm under a setting which *does not assume* the iterates to be bounded a-priori. Previous works either assumed the latter condition a-priori (e.g., Benveniste et al. [2012]), or they require a compact constraint set (e.g., Atchadé et al. [2017]). Using a novel analysis, we show that the mean squared error between the SA iterates and the unique performative stable solution [cf. (2)] converges at a rate of $\mathcal{O}(1/k)$, where k is the iteration number. We also discuss the convergence to an approximate stationary point of (1) when the loss function $\ell(\theta; z)$ is possibly non-convex.
- We demonstrate the efficacy of the SA algorithm with numerical experiments. We show that it has a comparable performance with Mendler-Dünner et al. [2020] which assumes an ideal setting with i.i.d. samples taken from the shifted distribution.

The rest of this paper is organized as follows. §2 formally describes the performative prediction problem and a state dependent SA algorithm for tackling the problem, §3 presents the main theoretical results for the convergence of the state dependent SA algorithm, §4 gives an overview of the proof strategy, and §5 presents the numerical experiments.

Related Works Analysis for state dependent stochastic approximation (SA) algorithms with controlled MC, which extend over the classical SA [Robbins and Monro, 1951], has been considered in a number of works. Benveniste et al. [2012], Kushner and Yin [2003] studied the asymptotic convergence of such algorithms, also see [Tadić et al., 2017] which analyzed the case of biased SA. Recent works have analyzed the finite-time performance of state dependent SA algorithms that are related to ours. Atchadé et al. [2017] considered a proximal SA algorithm where the proximal function has a compact domain; Karimi et al. [2019] analyzed the plain SA algorithm without projection but assumed that the updates are bounded; Sun et al. [2018], Doan et al. [2020] studied SA algorithms with a static MC not suitable for performative prediction. Our analysis relaxes these restrictions and focuses on convergence to a performative stable solution unique to (1).

Lastly, our analysis technique is related to the recent endeavors on obtaining finite time bounds for reinforcement learning (RL) algorithms. Notice that (1) can be regarded as a special case of policy optimization [Sutton and Barto, 2018]. To this end, recent works [Wu et al., 2020, Xu et al., 2020, Zhang et al., 2020] studied the sample complexity of actor critic algorithms with controlled MCs in finding a stationary point of an average reward function. In comparison, we study the convergence to a unique performative stable solution.

Notations We denote Z as the state space of samples and $\mathcal{Z}$ is a σ-algebra on Z. A Markov transition kernel is a map given by $\mathrm{P} : \mathsf{Z} \times \mathcal{Z} \to \mathbb{R}_+$. At the state $z \in \mathsf{Z}$, the next state is drawn as $z' \sim \mathrm{P}(z, \cdot)$. It holds for any measurable function $f : \mathsf{Z} \to \mathbb{R}$ that $\mathbb{E}[f(z')|z] = \int_{\mathsf{Z}} f(z')\mathrm{P}(z, \mathrm{d}\, z') =: \mathrm{P}f(z)$. Unless otherwise specified, the operator ∇ takes the gradient of a function w.r.t. the first argument for a multivariate function, e.g., $\nabla \ell(\theta; z)$ denotes the gradient of $\ell(\theta; z)$ taken w.r.t. θ. For $x, y \in \mathbb{R}^d$, we denote the inner product as $\langle x \,|\, y \rangle = x^\top y$.

2 STATE DEPENDENT PERFORMATIVE PREDICTION

Since Problem (1) is non-convex in general, we are interested in algorithms that converge to the *performative stable* (PS) solution:

$$\theta_{PS} = \arg\min_{\theta \in \mathbb{R}^d} \; \mathbb{E}_{z \sim \mathcal{D}(\theta_{PS})}[\ell(\theta; z)]. \tag{2}$$

Note that the expectation is taken with respect to $z \sim \mathcal{D}(\theta_{PS})$. It is known that θ_{PS} is in general different from an optimal or stationary solution to (1); see [Perdomo et al., 2020, Example 2.2]. Instead, θ_{PS} is a fixed point solution to the procedure when the learner repeatedly update the parameter θ with the drifted data distribution provided by the agent.

We consider a *state dependent* stochastic approximation (SA) algorithm motivated by the stateful nature of agents. The latter is modeled using a controlled Markov chain. For any θ, we define a Markov transition kernel P_θ which induces a Markov chain with a unique stationary distribution $\pi_\theta(\cdot)$ such that $\mathbb{E}_{z \sim \pi_\theta(\cdot)}[\nabla \ell(\theta; z)] = \mathbb{E}_{z' \sim \mathcal{D}(\theta)}[\nabla \ell(\theta; z')]$. The *learner* and *agent* interact through the rules depicted in Algorithm 1 and the $\mathsf{Sample\text{-}n\text{-}Adapt}(\cdot)$ subroutine.

We observe that (3) is a standard SA recursion based on $\nabla \ell(\theta_k; z_{k+1})$, where the *learner* deploys the most recent model θ_k and takes the sample z_{k+1} directly from the agent. Specifically, we consider a *state dependent* setting where the sampling of z_{k+1} can be affected by both the *learner*'s and *agent*'s current states. Formally, the state dependency is captured by modeling the samples sequence $\{z_k\}_{k \geq 1}$ as a controlled MC in (4). Notice that the stationary distribution $\pi_\theta(\cdot)$ does not need to be the same as $\mathcal{D}(\theta)$ as long as the former yields an unbiased gradient estimate. However, for simplicity, we assume $\pi_\theta(\cdot) \equiv \mathcal{D}(\theta)$ for any $\theta \in \mathbb{R}^d$.

Unlike Algorithm 1, the greedy deployment scheme studied by Mendler-Dünner et al. [2020] assumed that the agent(s) draws z_{k+1} as independent samples directly from $\mathcal{D}(\theta_k)$. The latter implies that $\nabla \ell(\theta_k; z_{k+1})$ is an unbiased estimator of $\mathbb{E}_{z' \sim \mathcal{D}(\theta_k)}[\nabla \ell(\theta_k; z')]$. As a significant departure from Mendler-Dünner et al. [2020], in (3), the stochastic gradient $\nabla \ell(\theta_k; z_{k+1})$ is a *biased*

Algorithm 1: State Dependent SA

Input: initialization θ_0, step sizes $\{\gamma_k\}_{k \geq 0}$.

For $k = 0, 1, 2, \ldots$

Agent draws $z_{k+1} = \mathsf{Sample\text{-}n\text{-}Adapt}(\theta_k, z_k)$.

Learner updates

$$\theta_{k+1} = \theta_k - \gamma_{k+1} \nabla \ell(\theta_k; z_{k+1}), \tag{3}$$

and deploys θ_{k+1}.

$\mathsf{Sample\text{-}n\text{-}Adapt}(\theta, z)$:

Draw the next sample as

$$z' \sim \mathrm{P}_\theta(z, \cdot), \tag{4}$$

where $\mathrm{P}_\theta : \mathsf{Z} \times \mathcal{Z} \to \mathbb{R}_+$ is a Markov transition kernel that depends on the input θ.

Output: $z' \in \mathsf{Z}$.

estimator. We observe

$$\begin{aligned} \mathbb{E}[\nabla \ell(\theta; z_{k+1}) \,|\, z_k] &= \mathrm{P}_{\theta_k} \nabla \ell(\theta; z_k) \\ &= \textstyle\int_{\mathsf{Z}} \nabla \ell(\theta; z) \mathrm{P}_{\theta_k}(z_k, \mathrm{d}\, z), \end{aligned} \tag{5}$$

for any $\theta \in \mathbb{R}^d$. Since $\mathrm{P}_{\theta_k}(z_k, \cdot) \neq \pi_{\theta_k}(\cdot)$, we have

$$\mathbb{E}[\nabla \ell(\theta; z_{k+1}) \,|\, z_k] \neq \mathbb{E}_{z' \sim \mathcal{D}(\theta_k)}[\nabla \ell(\theta_k; z')].$$

To provide some insights, under (4) with restricted access to the shifted data distribution $\mathcal{D}(\theta_k)$, one possibility to obtain an *unbiased* gradient estimate is via holding θ_k as fixed and repeat the sampling process $z' = \mathsf{Sample\text{-}n\text{-}Adapt}(\theta_k, z)$ indefinitely. In this case, we have the unbiased estimate $\lim_{n \to \infty} \mathrm{P}^n_{\theta_k} \nabla \ell(\theta; z) = \mathbb{E}_{z' \sim \mathcal{D}(\theta_k)}[\nabla \ell(\theta; z')]$ for any initial state $z \in \mathsf{Z}$. The latter can be regarded as the case where the learner *waits a long time* for the agent(s) to adapt to the distribution $\mathcal{D}(\theta_k)$ prior to updating the model and deploying it.

In this paper, we are interested in a case where the learner *does not wait* for the agent(s) to adapt to the distribution $\mathcal{D}(\theta_k)$, as s/he takes the instantaneous sample z_{k+1} from agent(s). The state dependent SA algorithm uses samples that are *co-evolving* with the learner's iterate in (3). The model (4) captures the stateful and stochastic nature of the agent as the sample z_{k+1} depends on the previous one z_k. Below we provide a motivating application example for (3), (4):

Example 1 (Strategic Classification with Adapted Best Response)**.** We consider the problem of strategic classification [Hardt et al., 2016] involving some *agents* and a *learner*. In an ideal scenario, the agent provides the best-response (i.e., optimized) samples upon knowing the current learner's state θ_k. The sample $z_{k+1} \sim \mathcal{D}(\theta_k)$ is drawn as

$$z_{k+1} \in \arg\max_{z' \in \mathsf{Z}} \; U(z'; \tilde{z}_{k+1}, \theta_k), \quad \tilde{z}_{k+1} \sim \mathcal{D}_0, \tag{6}$$

where $\mathcal{D}_0$ is the base distribution and $U(z'; z, \theta)$ is strongly concave in z' for any (z, θ). Observe that the best response $\max_{z' \in \mathsf{Z}} U(z'; z, \theta)$ perturbs the base sample z in favor of the agent.

In practice, the exact maximization in (6) can not be obtained unless the agent(s) are given sufficient time to respond to the learner's state, e.g., due to the difficulty in solving (6). Instead, we consider a setting where the agent(s) improve their responses via a gradient ascent dynamics evolving simultaneously with the learner.

Concretely, consider a setting where the learning problem (1) utilizes data provided by m agents. Let $\mathcal{D}_0$ be the empirical distribution of m data points $\{\bar{d}_1, ..., \bar{d}_m\}$, where $\bar{d}_i \in \mathsf{Z}$ is the initial data held by agent i. At iteration k, a subset of agents $\mathcal{I}_k \subset \{1, ..., m\}$ (selected uniformly with $|\mathcal{I}_k| = pm$, $p \in (0, 1]$) becomes aware of the learner's state and they search for the best response through a gradient descent update. Then, the learner selects uniformly an agent $i_k \in \{1, ..., m\}$ and requests the data sample z_{k+1} from him/her. In summary, the inexact best response dynamics follows:

$$\textbf{Step 1: } d_i^{k+1} = \begin{cases} d_i^k + \alpha \nabla U(d_i^k; \bar{d}_i, \theta_k), & i \in \mathcal{I}_k, \\ d_i^k, & i \notin \mathcal{I}_k, \end{cases}$$
$$\textbf{Step 2: } z_{k+1} = d_{i_k}^{k+1}, \tag{7}$$

where the gradient is taken w.r.t. the first argument of $U(\cdot)$, $\alpha > 0$ is the agents' response rate. For the initialization, we set $d_i^0 = \bar{d}_i$ for all $i = 1, ..., m$. The above highlights the stateful nature of agents as they seek to improve their responses based on past experiences.

The best response dynamics executed by the agent(s) in (7) leads naturally to a controlled MC (4). Specifically, the MC's state is given by the tuple $\hat{z}_k = \{d_1^k, ..., d_m^k, z_k\}$ and the application of the Markov kernel P_{θ_k} to $\hat{z}_k$ yields the inexact best response dynamics (7). Furthermore, the latter admits a stationary distribution where $\lim_{n \to \infty} \mathrm{P}_{\theta_k}^n \nabla \ell(\theta; \hat{z}) = \mathbb{E}_{z \sim \mathcal{D}(\theta_k)}[\nabla \ell(\theta; z)]$. We provide detailed properties about the controlled MC in Appendix A.2; also see Algorithm 2 in Appendix D. □

We remark that from a stochastic algorithm design perspective, (3), (4) can also be motivated from the use of an MCMC sampler [Robert and Casella, 2013] when it is difficult to sample from $\mathcal{D}(\theta)$.

3 MAIN RESULTS

This section deals with the analysis of (3), (4) which entails unique challenges. First, the agent's states $\{z_k\}_{k \geq 1}$ form a *controlled MC* whose transition probabilities change according to the learner's states $\{\theta_k\}_{k \geq 0}$. Second, a plain SA update is used in (3) which does not require the projection to a compact constraint set in [Atchadé et al., 2017]. In fact, θ_k can be unbounded when the step size is not carefully selected.

To begin, let us define the following notations:

$$\begin{aligned} f(\theta_1; \theta_2) &= \mathbb{E}_{z \sim \mathcal{D}(\theta_2)}[\ell(\theta_1; z)], \\ \nabla f(\theta_1; \theta_2) &= \mathbb{E}_{z \sim \mathcal{D}(\theta_2)}[\nabla \ell(\theta_1; z)], \end{aligned} \tag{11}$$

where the first argument θ_1 controls the loss function value and the second argument θ_2 controls the distribution shift. Notice that $\nabla f(\theta_{PS}; \theta_{PS}) = 0$.

We consider the following assumptions. First, the learner's loss is strongly convex in θ, and its gradient map is Lipschitz continuous in (θ, z), i.e.,

Assumption 1. *For each $z \in \mathsf{Z}$, there exists $\mu > 0$ such that $\forall\ \theta, \theta' \in \mathbb{R}^d$,*

$$\ell(\theta; z) \geq \ell(\theta'; z) + \langle \nabla \ell(\theta'; z) \,|\, \theta - \theta' \rangle + \frac{\mu}{2} \|\theta' - \theta\|^2. \tag{12}$$

Assumption 2. *There exists $L \geq 0$ such that $\forall\ \theta, \theta' \in \mathbb{R}^d,\ z, z' \in \mathsf{Z}$,*

$$\|\nabla \ell(\theta; z) - \nabla \ell(\theta'; z')\| \leq L\big\{\|\theta - \theta'\| + \|z - z'\|\big\}. \tag{13}$$

Notice that as a consequence, the expected objective function $f(\theta_1, \theta_2)$ and gradient $\nabla f(\theta_1; \theta_2)$ are μ-strongly convex in θ_1, and L-Lipschitz in θ_1, respectively. These are standard assumptions in the optimization literature. As indicated in [Drusvyatskiy and Xiao, 2020], these conditions are closely related to the existence of a performative stable solution in (2).

Second, we have the following assumption on the oscillation of the stochastic gradient $\nabla \ell(\theta; z)$:

Assumption 3. *There exists $\sigma \geq 0$ such that $\forall\ \theta \in \mathbb{R}^d$,*

$$\sup_{z \in \mathsf{Z}} \|\nabla \ell(\theta; z) - \nabla f(\theta; \theta_{PS})\| \leq \sigma\big(1 + \|\theta - \theta_{PS}\|\big). \tag{14}$$

The above is slightly stronger than the assumptions on second order moments typically found in the stochastic gradient literature, e.g., Bottou et al. [2018], as we require a uniform bound on the gradient noise. This condition is common for the algorithms using Markovian samples [Sun et al., 2018, Srikant and Ying, 2019, Karimi et al., 2019], which requires that the oscillation of stochastic gradient is controlled. Moreover, similar to [Doan et al., 2020], this bound is adapted to the growth of $\|\theta - \theta_{PS}\|$ which is compatible with the strong convexity of the loss function $\ell(\theta; z)$. Lastly, for the Example 1 with strategic classification, this assumption is satisfied for the finite dataset setting in (7).

Our next set of assumptions pertain to the Markov kernels P_θ that generate $\{z_k\}_{k \geq 1}$:

Assumption 4. *There exists a solution $\widehat{\nabla} \ell : \mathbb{R}^d \times \mathsf{Z} \to \mathbb{R}^d$ to the Poisson equation: $\forall\ \theta, \theta' \in \mathbb{R}^d, z \in \mathsf{Z}$,*

$$\nabla \ell(\theta'; z) - \nabla f(\theta'; \theta) = \widehat{\nabla} \ell(\theta'; z) - \mathrm{P}_\theta \widehat{\nabla} \ell(\theta'; z). \tag{15}$$

Theorem 1. *Under Assumptions 1–6. Suppose that the MCs induced by P_θ are uniformly geometric ergodic for any $\theta \in \mathbb{R}^d$, the problem parameters satisfy $\epsilon < \frac{\mu}{L}$, the step sizes $\{\gamma_k\}_{k\geq 1}$ are non-increasing and for any $k \geq 1$,*

$$\frac{\gamma_{k-1}}{\gamma_k} \leq 1 + \frac{\gamma_k(\mu - L\epsilon)}{4}, \quad \gamma_k \leq \min\left\{ \frac{\mu - L\epsilon}{2L^2}, \frac{\mu - L\epsilon}{2\mathcal{C}_2}, \frac{\min\{(\mu - L\epsilon)/3, 3\widehat{L}\}}{\mathcal{C}_3 + 3\widehat{L}(\mu - L\epsilon)}, \frac{1}{6\widehat{L}} \right\}. \tag{8}$$

Then for any $k \geq 1$, the expected distance between θ_k and the performative stable solution θ_{PS} satisfies

$$\mathbb{E}[\|\theta_k - \theta_{PS}\|^2] \leq \prod_{i=1}^{k} \Big(1 - \gamma_i \frac{\mu - L\epsilon}{2}\Big) \|\theta_0 - \theta_{PS}\|^2 + \mathbb{C}\, \gamma_k, \tag{9}$$

where $\mathbb{E}[\cdot]$ is the expectation taken over all the randomness in (3), (4), and we have defined:

$$\mathbb{C} := 3\widehat{L}\overline{\Delta} + \frac{4\varsigma}{\mu - L\epsilon}\Big(2(2\sigma^2 + \mathcal{C}_1) + (\mu - L\epsilon)\widehat{L} + \Big(\mathcal{C}_3 + 3(\mu - L\epsilon)\widehat{L}\Big)\overline{\Delta}\Big), \tag{10}$$

with $\varsigma := 1 + \gamma_1(\mu - L\epsilon)/4$, and $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$, $\overline{\Delta}$ are constants defined in (23), (26), respectively.

Assumption 5. *Consider the solution to Poisson equation, $\widehat{\nabla}\ell(\cdot;\cdot)$, defined in (15). There exists $L_{PH} \geq 0$ such that $\forall\ \theta, \theta' \in \mathbb{R}^d$,*

$$\sup_{z \in \mathsf{Z}} \|\mathrm{P}_\theta \widehat{\nabla}\ell(\theta; z) - \mathrm{P}_{\theta'} \widehat{\nabla}\ell(\theta'; z)\| \leq L_{PH}\|\theta - \theta'\|. \tag{16}$$

For Assumption 4, the Poisson equation's solution $\widehat{\nabla}\ell$ in (15) exists under mild assumptions on the MCs. For instance, it holds when the MC is irreducible and aperiodic, and satisfying a Lyapunov drift condition, or in the simpler case, when the MC is uniform geometrically ergodic, see [Douc et al., 2018, Ch. 21.2]. For simplicity, we will focus on the case when the MC is uniform geometrically ergodic. For instance with Example 1 for the strategic classification problem, Assumption 4 holds when repeated applications of the iterative map (7) converges linearly to the best response.

Meanwhile, Assumption 5 is a smoothness condition on the kernel P_θ, which holds when perturbation to the Markov kernel is stable w.r.t. θ. In particular, a sufficient condition that implies the assumption when the controlled Markov chain is smooth with respect to changes in θ, i.e., $\sup_{z\in\mathsf{Z}} \|\mathrm{P}_\theta(z,\cdot) - \mathrm{P}_{\theta'}(z,\cdot)\|_{\mathrm{TV}} = \mathcal{O}(\|\theta - \theta'\|)$; see [Karimi et al., 2019, Appendix D]. In Appendix D, we also verify the condition under a special case of Example 1.

Moreover, Assumption 5 is linked to our next assumption which is central to the study of performative prediction. We require the distribution map $\mathcal{D}(\theta)$ to be ϵ-sensitive w.r.t. θ:

Assumption 6. *There exists $\epsilon \geq 0$ such that*

$$W_1(\mathcal{D}(\theta), \mathcal{D}(\theta')) \leq \epsilon\|\theta - \theta'\|,\ \forall\ \theta, \theta' \in \mathbb{R}^d, \tag{17}$$

where $W_1(\cdot,\cdot) = \inf_{J\in\mathcal{J}(\cdot,\cdot)} \mathbb{E}_{(z,z')\sim J}[\|z - z'\|_1]$. Notice that $W_1(\cdot)$ denotes the Wasserstein-1 distance and $\mathcal{J}(\mathcal{D}(\theta), \mathcal{D}(\theta'))$ is the set of all joint distributions on $\mathsf{Z} \times \mathsf{Z}$ with $\mathcal{D}(\theta), \mathcal{D}(\theta')$ as its marginal distribution.

The above is commonly assumed in performative prediction, e.g., [Perdomo et al., 2020]. Intuitively, it allows performative prediction algorithm to behave stably as the perturbation to $\mathcal{D}(\theta)$ is under control. In the subsequent analysis, we demonstrate that carefully controlling the step size in relation to μ, ϵ, L is crucial to the convergence of Algorithm 1.

Before presenting our main result, we notice that together with a uniform geometric ergodic assumption on the MCs induced by P_θ, Assumptions 2, 3, 4 imply that there exists constants $\overline{L}$, $\widehat{L} > 0$ such that

$$\max\left\{ \|\widehat{\nabla}\ell(\theta; z)\|, \left\|\mathrm{P}_\theta \widehat{\nabla}\ell(\theta; z)\right\| \right\} \leq \widehat{L}(1 + \|\theta - \theta_{PS}\|),$$
$$\|\nabla\ell(\theta; z)\| \leq \overline{L}(1 + \|\theta - \theta_{PS}\|), \tag{18}$$

for any $z \in \mathsf{Z}$, $\theta \in \mathbb{R}^d$. In other words, $\nabla\ell(\theta; z)$, $\widehat{\nabla}\ell(\theta; z)$, $\mathrm{P}_\theta\widehat{\nabla}\ell(\theta; z)$ are all locally bounded functions. Notice that $\overline{L}$ is proportional to σ in Assumption 3, while $\widehat{L}$ is proportional to the maximum mixing time of the Markov chain induced by the kernel P_θ over all $\theta \in \mathbb{R}^d$; see Appendix C.1 for their precise expressions.

Our main result for the state dependent SA algorithm is summarized in Theorem 1, which establishes the finite-time convergence of the state dependent SA algorithm (3), (4). To understand this result, we observe that the step size conditions in (8) can be satisfied by a variety of step size schedules. For instance, it can be satisfied by the constant step size $\gamma_k \equiv \gamma$; and the diminishing step size $\gamma_k = a_0/(a_1 + k) = \mathcal{O}(1/k)$ with appropriate $a_0, a_1 > 0$. Moreover, we require the SA algorithm to work in the regime when $\epsilon < \mu/L$. Similar condition is enforced in Perdomo et al. [2020] to ensure that the solution θ_{PS} is a stable fixed point to (2).

The main result of Theorem 1 is stated on the expected squared distance between θ_k, θ_{PS} in (9). The bound consists of a *transient* term and a *fluctuation* term. The *transient* term decays sub-exponentially as $\mathcal{O}(\exp(-\frac{\mu - L\epsilon}{2}\sum_{i=1}^{k}\gamma_i))$ and is scaled by the initial error $\|\theta_0 - \theta_{PS}\|^2$. The *fluctuation* term is in the order of $\mathcal{O}(\gamma_k)$ and is scaled by $\mathbb{C}$ which depends on the oscillation of stochastic gradient (via σ) and the mixing time of the controlled Markov chain (via $\widehat{L}$). With a diminishing step size schedule such as $\gamma_k = c_0/(c_1 + k)$, Theorem 1 shows that the state dependent SA algorithm finds the performative stable solution θ_{PS} at the rate of $\mathcal{O}(1/k)$ in expectation.

Non-strongly-convex Loss Function An obvious drawback with Theorem 1 is the requirement of strongly convex loss functions in Assumption 1. Below, we comment on the convergence of the state dependent SA algorithm (3), (4) when the loss function is possibly non-convex. In the absence of Assumption 1, the performative stable solution θ_{PS} may not be well defined. We resort to finding a stationary point to (1).

Our idea is to view (3), (4) as a biased SA algorithm with mean field $h(\theta) = \mathbb{E}_{z\sim\mathcal{D}(\theta)}[\nabla\ell(\theta; z)] = \nabla f(\theta; \theta)$. This mean field is correlated with the gradient for the performative loss in (1). Under additional assumptions on $\ell(\theta; z)$, $\mathcal{D}(\theta)$, in Appendix B we show

$$\langle h(\theta) \,|\, \nabla\mathbb{E}_{z\sim\mathcal{D}(\theta)}[\ell(\theta; z)]\rangle \geq \tfrac{1}{2}\|h(\theta)\|^2 - \mathrm{c}_0, \tag{19}$$

holds for all $\theta \in \mathbb{R}^d$, where c_0 is a bias term defined as:

$$\mathrm{c}_0 := \sup_{\theta\in\mathbb{R}^d} \frac{1}{2}\mathbb{E}_{z\sim\mathcal{D}(\theta)}\left[|\ell(\theta; z)|^2\|\nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\|^2\right],$$

where $p_{\mathcal{D}(\theta)} : \mathsf{Z} \to \mathbb{R}_+$ is the probability distribution function representing $\mathcal{D}(\theta)$. Observe that $\mathrm{c}_0 < \infty$ for compact Z and the constant is dependent on the sensitivity of $\mathcal{D}(\theta)$ to shifts in θ (cf. Assumption 6).

It should be pointed out that, naturally, the algorithm (3), (4) may not provide a 'good' solution to the performative learning problem (1) as the learner generally ignores the performativity of agents. However, if the state dependent distribution is not sensitive to the change of state, the following corollary shows that (3) would still converge to a $\mathcal{O}(\mathrm{c}_0)$-neighborhood of a stationary solution. Before we discuss the main statement, we need two additional assumptions:

Assumption 2'. *The function $V(\theta) = \mathbb{E}_{z\sim\mathcal{D}(\theta)}[\ell(\theta; z)]$ is continuously differentiable and there exists $L_V \geq 0$ such that $\|\nabla V(\theta) - \nabla V(\theta')\| \leq L_V\|\theta - \theta'\|$, $\forall\, \theta, \theta' \in \mathbb{R}^d$.*

Assumption 3'. *There exists $\sigma \geq 0$ such that $\|\nabla\ell(\theta; z) - \nabla f(\theta; \theta)\| \leq \sigma$, $\forall\, \theta \in \mathbb{R}^d, z \in \mathsf{Z}$.*

The above are stronger conditions than Assumptions 2, 3, yet are reasonable settings for certain non-convex loss functions. For instance, Assumption 2' holds if $\ell(\theta; z), \nabla_\theta \log\left(p_{D(\theta)}(z)\right)$ are bounded, and $\nabla_\theta \log\left(p_{D(\theta)}(z)\right)$ to be Lipschitz w.r.t. θ, e.g., when

Corollary 1. *Under Assumptions 2', 3', 4, 5, and let (19) holds. With a step size sequence that decays as $\gamma_k = \mathcal{O}(1/\sqrt{k})$, it holds for any $K \geq 1$ that*

$$\begin{aligned}&\mathbb{E}[\|\nabla V(\theta_{\mathsf{K}})\|^2] = \mathcal{O}(\log K/\sqrt{K} + \mathrm{c}_0),\\ &\mathbb{P}(\mathsf{K} = k) = \gamma_k / \textstyle\sum_{j=1}^{K}\gamma_j,\ \forall\, k \in \{1, ..., K\},\end{aligned} \tag{20}$$

where $\mathsf{K} \in \{1, ..., K\}$ is a discrete r.v. independent of the randomness in the SA algorithm (defined in Appendix B) and $\mathbb{E}[\cdot]$ denotes the total expectation.

$\mathcal{D}(\theta)$ is 'smooth' such as a Gaussian or softmax distribution. Assumption 3' holds under similar condition as Assumption 3, e.g., Z is compact. We obtain Corollary 1 from [Karimi et al., 2019, Theorem 2].

Corollary 1 shows that even without strong convexity on $\ell(\cdot; z)$, the state dependent SA algorithm finds an $\mathcal{O}(\log K/\sqrt{K} + \mathrm{c}_0)$-stationary solution to (1) in at most K iterations. The proof can be found in Appendix B.

4 PROOF OUTLINE

We outline the main steps in proving Theorem 1. Our proof strategy consists in tracking the mean squared error $\Delta_k := \mathbb{E}[\|\theta_k - \theta_{PS}\|^2]$. To simplify notations, we define $\tilde{\mu} := \mu - L\epsilon$, the scalar product $G_{m:n} = \prod_{i=m}^{n}(1 - \gamma_i\tilde{\mu})$, for $n > m \geq 1$, $G_{m:n} = 1$ if $n \leq m$.

The following lemma describes the one-step progress of the SA algorithm.

Lemma 1. *Under Assumptions 1, 2, 3, 6. For any $k \geq 0$, it holds*

$$\begin{aligned}\|\theta_{k+1} - \theta_{PS}\|^2 \leq& \left(1 - 2\gamma_{k+1}\tilde{\mu} + 2L^2\gamma_{k+1}^2\right)\|\theta_k - \theta_{PS}\|^2\\ &+ 2\sigma^2\gamma_{k+1}^2 - 2\gamma_{k+1}\langle\theta_k - \theta_{PS} \,|\, \nabla\ell(\theta_k; z_{k+1}) - \nabla f(\theta_k; \theta_k)\rangle.\end{aligned} \tag{21}$$

The proof can be found in Appendix C.2, which involves a simple expansion of the squared error. The above lemma suggests that the sensitivity parameter shall satisfy $\epsilon < \mu/L$ to ensure $\tilde{\mu} > 0$. Furthermore, the step size condition $\sup_{k\geq 1}\gamma_k \leq \tilde{\mu}/(2L^2)$ in (8) leads to $1 - 2\gamma_{k+1}\tilde{\mu} + 2L^2\gamma_{k+1}^2 \leq 1 - \gamma_{k+1}\tilde{\mu}$ such that the first term in the r.h.s. of (21) is a contraction.

Under the above premises and suppose $z_{k+1} \overset{\text{i.i.d.}}{\sim} \mathcal{D}(\theta_k)$ as in [Mendler-Dünner et al., 2020, Drusvyatskiy and Xiao, 2020], the stochastic gradient in (21) is conditionally *unbiased*. Lemma 1 leads to $\Delta_{k+1} \leq (1 - \gamma_{k+1}\tilde{\mu})\Delta_k + 2\sigma^2\gamma_{k+1}^2$, implying $\Delta_k = \mathcal{O}(\gamma_k)$.

However, for the state dependent SA algorithm (3), (4), the stochastic gradient $\nabla\ell(\theta_k; z_{k+1})$ is conditionally *biased* and is driven by a controlled MC. Under the stepsize condition $\sup_{k\geq 1}\gamma_k \leq \tilde{\mu}/(2L^2)$, taking the total expectation and solving (21) yield

$$\Delta_k \le G_{1:k}\Delta_0 + 2\sigma^2 \sum_{s=0}^{k-1} G_{s+2:k}\gamma_{s+1}^2 + 2\sum_{s=0}^{k-1} G_{s+2:k} \times \gamma_{s+1}\mathbb{E}\big[\langle \theta_{PS} - \theta_s \,|\, \nabla \ell(\theta_s; z_{s+1}) - \nabla f(\theta_s;\theta_s)\rangle\big]. \quad (22)$$

It can be shown that the first two terms are bounded by $\mathcal{O}(\gamma_k)$. We are interested in the last term when the samples $\{z_k\}_{k\ge1}$ are drawn according to (4). Observe:

Lemma 2. *Under Assumptions 2–6 and the stepsize conditions in (8). For any $k \ge 1$, it holds*

$$\begin{aligned} &2\sum_{s=0}^{k-1} G_{s+2:k}\gamma_{s+1}\mathbb{E}\big[\langle \theta_{PS} - \theta_s \,|\, \nabla \ell(\theta_s; z_{s+1}) - \nabla f(\theta_s;\theta_s)\rangle\big] \\ &\le \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k}\big(\mathcal{C}_1 + \mathcal{C}_2\Delta_{s-1} + \mathcal{C}_3\Delta_{s-2}\big) \\ &\quad + \gamma_1 G_{2:k}\big\{\widehat{L}(1+3\Delta_0) + \gamma_1\mathcal{C}_1\big\} + \gamma_k\widehat{L}\big\{1+3\Delta_{k-1}\big\}, \end{aligned}$$

where we have defined the constants:

$$\begin{aligned} &\mathcal{C}_1 := \varsigma L_{PH}\overline{L} + 4\varsigma\overline{L}\widehat{L} + (1+\tilde{\mu})\varsigma\widehat{L}, \\ &\mathcal{C}_2 := 2\varsigma L_{PH}\overline{L}, \quad \mathcal{C}_3 := \mathcal{C}_1 + 2(1+\tilde{\mu})\varsigma\widehat{L}. \end{aligned} \quad (23)$$

The analysis is inspired by [Benveniste et al., 2012] and has been adopted in [Atchadé et al., 2017, Karimi et al., 2019]; see Appendix C.3. To handle the controlled MC, our technique involves applying Assumption 4 and decomposing the gradient error $\nabla \ell(\theta_s; z_{s+1}) - \nabla f(\theta_s;\theta_s)$ into Martingale and finite difference terms.

A key difference between Lemma 2 and analysis in the previous works such as [Benveniste et al., 2012] is that the latter assumed that the iterates or the stochastic gradients are bounded *a-priori* which can greatly simplify the proof. Even if such assumption holds, it may lead to unrealistic constants in the resultant bound. Our assumptions are significantly weaker as the stochastic gradients are allowed to grow as $\mathcal{O}(1 + \|\theta_k - \theta_{PS}\|)$ [cf. Assumption 3], compatible with the setting that the cost function $\ell(\cdot; z)$ is strongly convex. It demands a new proof technique as we present next.

Observe that substituting Lemma 2 into (22) yields:

$$\begin{aligned} \Delta_k &\le G_{1:k}\Delta_0 + \gamma_k\widehat{L}\big\{1+3\Delta_{k-1}\big\} \quad (24) \\ &\quad + \gamma_1 G_{2:k}\big\{\widehat{L}(1+3\Delta_0) + \gamma_1(2\sigma^2 + \mathcal{C}_1)\big\} \\ &\quad + \sum_{s=1}^{k-1}\gamma_{s+1}^2 G_{s+2:k}\big(2\sigma^2 + \mathcal{C}_1 + \mathcal{C}_2\Delta_s + \mathcal{C}_3\Delta_{s-1}\big). \end{aligned}$$

With the first step size condition in (8), we can apply the auxiliary result in Lemma 4 from the appendix, which simplifies the upper bound as

$$\begin{aligned} \Delta_k &\le G_{1:k}\Delta_0 + \Big(\tfrac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\gamma_k + 3\gamma_k\widehat{L}\Delta_{k-1} \\ &\quad + \gamma_1 G_{2:k}\big\{\widehat{L}(1+3\Delta_0) + \gamma_1(2\sigma^2 + \mathcal{C}_1)\big\} \\ &\quad + \textstyle\sum_{s=1}^{k-1}\gamma_{s+1}^2 G_{s+2:k}\big(\mathcal{C}_2\Delta_s + \mathcal{C}_3\Delta_{s-1}\big). \end{aligned} \quad (25)$$

Observe that the first row in (25) is already in a similar form to the bound presented in the theorem. The key issue lies with the last term $3\gamma_k\widehat{L}\Delta_{k-1}$ which may be unbounded. We show that our choice of step sizes in (8) ensures the convergence of Δ_k to $\mathcal{O}(\gamma_k)$:

Lemma 3. *Suppose that $\{\Delta_k\}_{k\ge0}$ satisfy (25) and the step sizes $\{\gamma_k\}_{k\ge1}$ satisfy (8). It holds* (i)

$$\sup_{k\ge0}\Delta_k \le \overline{\Delta} := 3\Delta_0 + \tfrac{\varsigma}{9\widehat{L}^2}\Big(2(2\sigma^2 + \mathcal{C}_1) + (\mu - L\epsilon)\widehat{L}\Big), \quad (26)$$

and (ii) *the following inequality holds for any $k \ge 1$:*

$$\begin{aligned} \Delta_k \le \prod_{i=1}^{k}(1 - \gamma_i\tfrac{\tilde{\mu}}{2})\Delta_0 + \Big\{&3\widehat{L}\overline{\Delta} + \frac{4\varsigma}{\tilde{\mu}}\Big(2(2\sigma^2 + \mathcal{C}_1) \\ &+ \tilde{\mu}\widehat{L} + \Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\overline{\Delta}\Big)\Big\}\gamma_k. \end{aligned} \quad (27)$$

Proving the above lemma requires one to establish the stability of the system (25), which demands a sufficiently small γ_k to control the remainder term $3\widehat{L}\gamma_k\Delta_{k-1}$. Our analysis leverages the special structure of this inequality system; see the proof details in Appendix C.4. The convergence bound (27) follows from the boundedness of Δ_k. Finally, we obtain Theorem 1 through applying Lemma 3.

5 NUMERICAL EXPERIMENTS

This section considers two performative prediction problems to validate our theories. All the experiments are performed with Python on a server using a single thread of an Intel Xeon 6138 CPU. Further details about the experiments below can be found in Appendix D.

Gaussian Mean Estimation The *first problem* is concerned with Gaussian mean estimation using synthetic data. Our aim is to validate Theorem 1 and illustrate how the controlled MC can lead to a variance reduced solution in this application. Here, (1) is specified as $\min_{\theta\in\mathbb{R}} \mathbb{E}_{z\sim\mathcal{D}(\theta)}[(z-\theta)^2/2]$ with $\mathcal{D}(\theta) \equiv \mathcal{N}(\bar{z} + \epsilon\theta; \sigma^2)$. For $0 < \epsilon < 1$, the performative stable solution has a closed form $\theta_{PS} = \frac{\bar{z}}{1-\epsilon}$. For the state dependent SA, the *agent* follows an autoregressive (AR) model $z_{k+1} = (1-\rho)z_k + \rho\tilde{z}_{k+1}$ with independent $\tilde{z}_{k+1} \sim \mathcal{N}(\bar{z} + \epsilon\theta_k; \sigma^2)$ and parameter $\rho \in (0,1)$. This AR recursion is a controlled MC with a stationary distribution that yields the unbiased gradient of (1) of reduced variance $\frac{\rho}{2-\rho}\sigma^2$. which implies that $\mathcal{D}(\theta) \ne \pi_\theta(\cdot)$ but it does not impact our final task of estimating the mean of Gaussian distribution. More details are in Appendix A.1.

We consider a setting with $\bar{z} = 10$, $\sigma = 50$, $\epsilon = 0.1$. The step size is $\gamma_k = \frac{c_0}{c_1+k}$, $c_0 = \frac{500}{\tilde{\mu}}$, $c_1 = \frac{800}{\tilde{\mu}^2}$. In Fig. 1 (left), we compare $|\theta_k - \theta_{PS}|^2$ against the iteration number k for the Gaussian estimation problem using

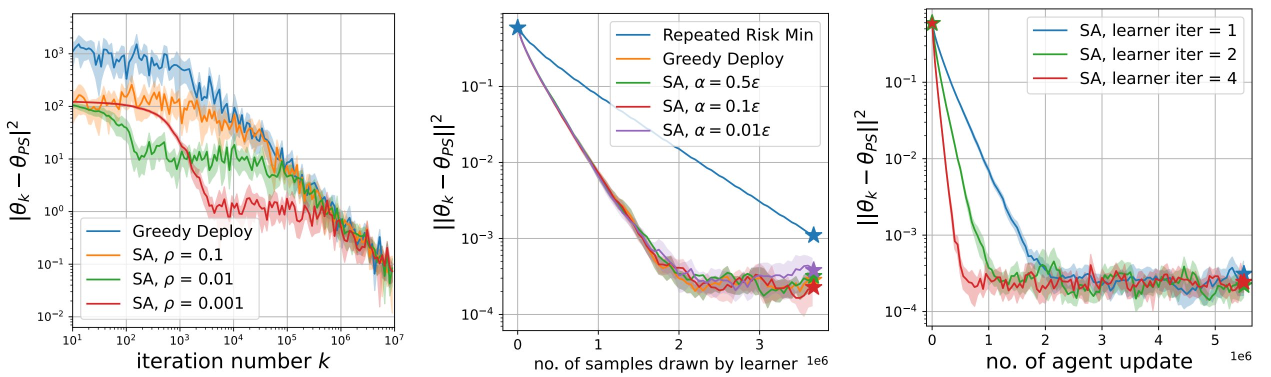

Figure 1: **Gaussian mean estimation** – *(Left)* Under different regression parameter ρ; **Strategic Classification** – *(Middle)* Under Linear BR $U_{\mathsf{q}}(\cdot)$ and different agent response rate α [cf. (7)]; *(Right)* Under Logistics BR $U_{\mathsf{lg}}(\cdot)$. The shaded region shows the 90% confidence interval over 20 trials.

state dependent SA and greedy deploy in [Mendler-Dünner et al., 2020]. As observed, both settings achieve an asymptotic convergence rate of $\mathcal{O}(1/k)$ towards θ_{PS} as predicted by Theorem 1. As $\rho \downarrow 0$, the state dependent SA delivers a smaller error as the AR model has a stationary distribution with lower variance.

Strategic Classification The *second problem* is a strategic classification (SC) problem similar to Perdomo et al. [2020] for a credit scoring classifier with `GiveMeSomeCredit` dataset[1]. Our aim is to demonstrate the effects on convergence rate when stateful agents adapt slowly to the shifted distribution. Recall that our theory implies while the algorithm will still converge to θ_{PS}, a slower convergence will be observed under the state dependent setting. To specify (1), let $z \equiv (x, y)$ where $x \in \mathbb{R}^d$ is feature vector, $y \in \{0, 1\}$ is label. The *learner* finds $\theta \in \mathbb{R}^d$ that minimizes $\mathbb{E}_{z\sim\mathcal{D}(\theta)}[\ell(\theta; z)]$, where:

$$\ell(\theta; z) = \tfrac{\beta}{2}\|\theta\|^2 + \log(1 + \exp(\langle\theta\,|\,x\rangle)) - y\langle\theta\,|\,x\rangle \quad (28)$$

is the logistic loss. With $\beta > 0$, $\ell(\theta; z)$ is a β-strongly convex function w.r.t. θ satisfying Assumption 1. For any $\theta \in \mathbb{R}^d$, the shifted data distribution $\mathcal{D}(\theta)$ is obtained through evaluating the best response (BR) in (6) of *Example 1*. We consider two types of strongly concave utility functions adopted by the *agents*:

$$U_{\mathsf{q}}(z'; z, \theta) = \langle\theta\,|\,x'\rangle - \tfrac{\|x'-x\|^2}{2\epsilon},$$
$$U_{\mathsf{lg}}(z'; z, \theta) = y\langle\theta\,|\,x'\rangle - \log\left(1 + \exp(\langle\theta\,|\,x'\rangle)\right) - \tfrac{\|x'-x\|^2}{2\epsilon},$$

where $z \equiv (x, y)$ is the original (unshifted) data. The label $y \in \{0, 1\}$ is unchanged in the BR. Notice that $U_{\mathsf{q}}(\cdot)$, $U_{\mathsf{lg}}(\cdot)$ have respectively linear and logistics costs. Both utility functions include a quadratic regularizer where $\epsilon > 0$ controls the sensitivity of distribution shift.

[1]Available: https://kaggle.com/c/GiveMeSomeCredit.

With a published θ_k, the agents maximize the utility function prior to giving data to the learner for the next round. For both $U_{\mathsf{q}}(\cdot)$ and $U_{\mathsf{lg}}(\cdot)$, the BR obtained steers the classifier in favor of the agent(s). Furthermore, $U_{\mathsf{lg}}(\cdot)$ is motivated by logistic regression which favors towards samples with label '1'.

In the experiments, we set $\beta = 1000/m$ in (28), $\epsilon = 0.01$ in the utility functions, and in (7), we set number of selected agents as $|\mathcal{I}_k| = 5$, agents' response rate as $\alpha = 0.5\epsilon$ unless otherwise specified. The step size for (3) is $\gamma_k = c_0/(c_1 + k)$, $c_0 = 100/\tilde{\mu}, c_1 = 8L^2/\tilde{\mu}^2$.

We first consider when θ_{PS} is computed with $\mathcal{D}(\theta)$ defined by the linear BR function $U_{\mathsf{q}}(\cdot)$ and compare state dependent SA (3), (4) with the greedy deploy scheme in [Mendler-Dünner et al., 2020] and repeated risk minimization [Perdomo et al., 2020]. As shown in Fig. 1 (middle), all algorithms converge to θ_{PS}. As $\alpha \downarrow 0$, the state dependent SA converges at slightly slower rates as the agents adapt to the distribution shift with increased mixing time of the MC[2] The result corroborates with Theorem 1 which established the $\mathcal{O}(1/k)$ convergence rate with state dependent SA.

We next consider the same SC problem as before, but under a different setting where θ_{PS} is computed with $\mathcal{D}(\theta)$ defined by the logistics BR function $U_{\mathsf{lg}}(\cdot)$. The agents follow a more complicated dynamics since the BR does not admit a closed form solution. Again, we aim to validate Theorem 1 on the convergence of Algorithm 1 under the stateful agent setting. We compare the distance $\|\theta_k - \theta_{PS}\|^2$ as the algorithm proceeds. In addition to showing the convergence to θ_{PS} as pre-

[2]Due to the ill conditioning of `GiveMeSomeCredit` dataset, the effect of mixing time are not obvious in the middle figure. An additional simulation on synthetic data with logistic loss can be found in Appendix D which shows a more significant effect of the slower mixing time on the convergence of SA.

dicted in Theorem 1, we demonstrate the effects when the learner adopts a lazy deployment scheme by varying the number of θ-update by the *learner* per the agents' update (cf. Mendler-Dünner et al. [2020]).

Fig. 1 (right) shows the error $\|\theta_k - \theta_{PS}\|^2$ against the number of adaptation steps performed at the agents via (7) as we illustrate the convergence rate from the perspectives of the agents. We observe that the error decreases at a faster rate when the number of learner's iteration increases. Further details about our experiments and additional results are in Appendix D.

Conclusion We consider a state dependent SA algorithm for performative prediction. We showed a convergence rate of $\mathcal{O}(1/k)$ in mean-squared error towards the performative stable solution when the agents provide data drawn from a controlled MC. Our study paved the first step towards understanding and applying performative prediction in a dynamical setting.

Acknowledgements

This work has been supported by the CUHK Direct Grant #4055113.

References

Y. F. Atchadé, G. Fort, and E. Moulines. On perturbed proximal gradient algorithms. *The Journal of Machine Learning Research*, 18(1):310–342, 2017.

P. L. Bartlett. Learning with a slowly changing distribution. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 243–252, 1992.

A. Benveniste, M. Métivier, and P. Priouret. *Adaptive algorithms and stochastic approximations*, volume 22. Springer Science & Business Media, 2012.

L. Bottou, F. E. Curtis, and J. Nocedal. Optimization methods for large-scale machine learning. *Siam Review*, 60(2):223–311, 2018.

G. Brown, S. Hod, and I. Kalemaj. Performative prediction in a stateful world. *arXiv preprint arXiv:2011.03885*, 2020.

Y. Cai, C. Daskalakis, and C. Papadimitriou. Optimum statistical estimation with strategic data sources. In *Conference on Learning Theory*, pages 280–296. PMLR, 2015.

T. T. Doan, L. M. Nguyen, N. H. Pham, and J. Romberg. Finite-time analysis of stochastic gradient descent under markov randomness. *arXiv preprint arXiv:2003.10973*, 2020.

R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov chains*. Springer, 2018.

D. Drusvyatskiy and L. Xiao. Stochastic optimization with decision-dependent distributions. *arXiv preprint arXiv:2011.11173*, 2020.

M. Hardt, N. Megiddo, C. Papadimitriou, and M. Wootters. Strategic classification. In *Proceedings of the 2016 ACM conference on innovations in theoretical computer science*, pages 111–122, 2016.

Z. Izzo, L. Ying, and J. Zou. How to learn when data reacts to your model: Performative gradient descent. *arXiv preprint arXiv:2102.07698*, 2021.

B. Karimi, B. Miasojedow, E. Moulines, and H.-T. Wai. Non-asymptotic analysis of biased stochastic approximation scheme. In *Conference on Learning Theory*, pages 1944–1974. PMLR, 2019.

J. Kleinberg and M. Raghavan. How do classifiers induce agents to invest effort strategically? *ACM Transactions on Economics and Computation (TEAC)*, 8(4):1–23, 2020.

H. Kushner and G. G. Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35. Springer Science & Business Media, 2003.

C. Mendler-Dünner, J. Perdomo, T. Zrnic, and M. Hardt. Stochastic optimization for performative prediction. *Advances in Neural Information Processing Systems*, 33, 2020.

J. Miller, J. C. Perdomo, and T. Zrnic. Outside the echo chamber: Optimizing the performative risk. *arXiv preprint arXiv:2102.08570*, 2021.

E. Munro. Learning to personalize treatments when agents are strategic. *arXiv preprint arXiv:2011.06528*, 2020.

A. Patrascu and I. Necoara. Efficient random coordinate descent algorithms for large-scale structured nonconvex optimization. *Journal of Global Optimization*, 61(1):19–46, 2015.

J. C. Perdomo, C. Zrnic, Tijana aCnd Mendler-Dünner, and M. Hardt. Performative prediction. In *ICML*, 2020.

J. Quiñonero-Candela, M. Sugiyama, N. D. Lawrence, and A. Schwaighofer. *Dataset shift in machine learning*. Mit Press, 2009.

P. Richtárik and M. Takáč. Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function. *Mathematical Programming*, 144(1):1–38, 2014.

H. Robbins and S. Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.

R. Srikant and L. Ying. Finite-time error bounds for linear stochastic approximation and TD learning. In *Conference on Learning Theory*, pages 2803–2830. PMLR, 2019.

T. Sun, Y. Sun, and W. Yin. On markov chain gradient descent. In *NeurIPS*, 2018.

R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

V. B. Tadić, A. Doucet, et al. Asymptotic bias of stochastic gradient search. *Annals of Applied Probability*, 27(6):3255–3304, 2017.

Y. Wu, W. Zhang, P. Xu, and Q. Gu. A finite time analysis of two time-scale actor critic methods. In *NeurIPS*, 2020.

T. Xu, Z. Wang, and Y. Liang. Non-asymptotic convergence analysis of two time-scale (natural) actor-critic algorithms. *arXiv preprint arXiv:2005.03557*, 2020.

S. Zhang, B. Liu, H. Yao, and S. Whiteson. Provably convergent two-timescale off-policy actor-critic with function approximation. In *International Conference on Machine Learning*, pages 11204–11213. PMLR, 2020.

T. Zrnic, E. Mazumdar, S. S. Sastry, and M. I. Jordan. Who leads and who follows in strategic classification? *arXiv preprint arXiv:2106.12529*, 2021.

Supplementary Material: State Dependent Performative Prediction with Stochastic Approximation

---

A SUPPLEMENTARY INFORMATION FOR SECTION 2

A.1 Example on Gaussian Estimation (A Case where $\pi_\theta(\cdot) \neq \mathcal{D}(\theta)$)

Consider the following instance of (1) with:

$$\min_{\theta \in \mathbb{R}} \; \mathbb{E}_{z \sim \mathcal{D}(\theta)}[(z-\theta)^2/2] \quad \text{where} \quad \mathcal{D}(\theta) \equiv \mathcal{N}(\bar{z} + \epsilon\theta; \sigma^2). \tag{29}$$

Following (3), the state dependent SA algorithm reads

$$\theta_{k+1} = \theta_k - \gamma_{k+1} \nabla \ell(\theta_k; z_{k+1}) = \theta_k - \gamma_{k+1}(\theta_k - z_{k+1}). \tag{30}$$

where the sequence $\{z_k\}_{k \geq 1}$ is generated by an autoregressive (AR) model, with $\rho \in (0, 1]$,

$$z_{k+1} = (1-\rho) z_k + \rho \tilde{z}_{k+1} \quad \text{where} \quad \tilde{z}_{k+1} \sim \mathcal{D}(\theta_k) = \mathcal{N}(\bar{z} + \epsilon\theta_k; \sigma^2), \tag{31}$$

such that the draw of $\tilde{z}_{k+1}$ are independent. We show that the algorithm (30), (31) can be analyzed as a state dependent SA (3), (4) considered in our framework. Precisely, we show that the controlled MC in (31) admits a stationary distribution $\pi_{\theta_k}(\cdot)$ such that $\mathbb{E}_{z \sim \pi_{\theta_k}(\cdot)}[\nabla \ell(\theta; z)] = \mathbb{E}_{z' \sim \mathcal{D}(\theta_k)}[\nabla \ell(\theta; z')]$.

Observe that (31) defines a controlled MC with a transition kernel denoted by $\mathrm{P}_{\theta_k} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ at the kth iteration, as in (4). For every $\theta \in \mathbb{R}$, $z \in \mathbb{R}$, the kernel P_θ has a unique stationary distribution given by

$$\lim_{n \to \infty} \mathrm{P}_\theta^n(z, \cdot) = \pi_\theta(\cdot) \equiv \mathcal{N}\big(\bar{z} + \epsilon\theta \,;\, \frac{\rho}{2-\rho}\sigma^2\big). \tag{32}$$

Notice that the above is different from the distribution $\mathcal{D}(\theta)$ desired in (29) unless $\rho = 1$. In the latter case, the AR model (31) reduces to drawing i.i.d. samples from $\mathcal{D}(\theta)$. For general $\rho < 1$, it still satisfies the asymptotically unbiasedness of the stochastic gradient estimate. In particular,

$$\mathbb{E}_{z \sim \pi_\theta(\cdot)}[\nabla \ell(\theta; z)] = \mathbb{E}_{z \sim \pi_\theta(\cdot)}[\theta - z] = \mathbb{E}_{z \sim \mathcal{D}(\theta)}[\nabla \ell(\theta; z)]. \tag{33}$$

The key observation is that for this particular performative prediction problem (29), the gradient of the loss function is *linear* in the sample z. As such, with (32) yielding a stationary distribution that has the same mean as $\mathcal{D}(\theta)$, the asymptotic unbiasedness property is unaffected. In fact, the stationary distribution in (32) has a reduced variance compared to $\mathcal{D}(\theta)$. Therefore, we expect the estimation error of θ_{PS} to be more stable using (30), (31) than [Mendler-Dünner et al., 2020].

A.2 Details on Example 1 for Adapted Best Response

We continue the discussions in the paper with the procedure (7). When $\theta \in \mathbb{R}^d$ is fixed, the procedure in (7) is modelled as a Markov Chain (MC) with unique stationary distribution that corresponds to the best response distribution $\mathcal{D}(\theta)$ described in (6).

To this end, we model the state of the MC by the tuple $\hat{z} \equiv (d_1, ..., d_m, z)$. Consider the state space given by Z^{m+1} and denote $\mathrm{P}_\theta : \mathsf{Z}^{m+1} \times \mathcal{Z}^{m+1} \to \mathbb{R}_+$ as the Markov transition kernel. We remark that there is a slight abuse of notation here as the stochastic gradient $\nabla \ell(\theta; \hat{z})$ used by the learner depends only on the last term, z, in the agents' state variable $\hat{z}$. We have decided to use the current notation in the main paper to avoid introducing complicated notation for the implementation focused readers. Nevertheless, the SC example fits our proposed model.

Turning back on the MC. Observe when the current state is $\hat{z}$, under the action of kernel P_θ, by following the description in (7), we obtain the next state $\hat{z}' = (d_1', ..., d_m', z')$ as

$$d_j' = d_j + \alpha \mathbb{1}_{j \in \mathcal{I}} \nabla U(d_j; \bar{d}_j, \theta),\ j = 1, ..., m, \quad z' = d_i', \tag{34}$$

with probability

$$\mathbb{P}(\mathcal{I}_k = \mathcal{I}, i_k = i) = \frac{1}{\binom{m}{pm}} \times \frac{1}{m},$$

for any $\mathcal{I} \subseteq \{1, ..., m\}$, $|\mathcal{I}| = pm$ and $i \in \{1, ..., m\}$.

At each transition, the data points $\{d_1, ..., d_m\}$ are updated by the first equation in (34). The latter can be treated as one iteration of the *random block coordinate gradient descent (RBCD)* algorithm to the *separable* problem:

$$\max_{d_i, i=1,...,m} \sum_{i=1}^m U(d_i; \bar{d}_i, \theta). \tag{35}$$

Note that the optimal solution to the above, $\{d_1^\star, ..., d_m^\star\}$, is a set of data points that forms the empirical distribution $\mathcal{D}(\theta)$. Furthermore, it is known that the RBCD algorithm converges linearly with high probability and almost surely to the optimal solution for strongly concave maximization; see Richtárik and Takáč [2014], Patrascu and Necoara [2015].

With the above observations, the MC induced by P_θ has a stationary distribution $\pi_\theta(\cdot)$ where for any measurable function $f : \mathsf{Z}^{m+1} \to \mathbb{R}^n$, it holds

$$\lim_{k \to \infty} \mathrm{P}_\theta^k f(d_1, ..., d_m, z) = \frac{1}{m} \sum_{i=1}^m f(d_1^\star, ..., d_m^\star, d_i^\star) = \mathbb{E}_{z' \sim \mathcal{D}(\theta)}[f(d_1^\star, ..., d_m^\star, z')], \tag{36}$$

for any initial state $\hat{z} = (d_1, ..., d_m, z)$. The above identity can be derived from the fact that the RBCD algorithm converges almost surely to the optimal solution to (35) and the random variable z^k is uniformly drawn from $\{d_1^k, ..., d_m^k\}$. Furthermore, for any L-Lipschitz continuous f, it holds

$$\begin{aligned}
&\left\| \mathrm{P}_\theta^k f(d_1, ..., d_m, z) - \mathbb{E}_{z' \sim \mathcal{D}(\theta)}[f(d_1^\star, ..., d_m^\star, z')] \right\| \\
&\overset{(a)}{\leq} \frac{1}{m} \sum_{i=1}^m \|\mathbb{E}[f(d_1^k, ..., d_m^k, d_i^k)] - f(d_1^\star, ..., d_m^\star, d_i^\star)\| \\
&\overset{(b)}{\leq} L\big(1 + \frac{1}{m}\big) \sum_{i=1}^m \mathbb{E}[\|d_i^k - d_i^\star\|] \leq \overline{\mathrm{C}} \rho^k,
\end{aligned} \tag{37}$$

where (a) uses $\mathrm{P}_\theta^k f(d_1, ..., d_m, z) = \sum_{i=1}^m \mathbb{E}[f(d_1^k, ..., d_m^k, d_i^k)]/m$ and the expectation is taken with respect to the random subset selection of $\mathcal{I}_k$ in (7). In the expression that follows (b), the constants $\overline{\mathrm{C}}$, $\rho \in [0, 1)$ depend on the initial value $\hat{z}$ and the strong concavity property of $U(\cdot)$. The above property is important for establishing the existence of the solution to Poisson equation in Assumption 4.

B CONVERGENCE ANALYSIS WITH NON-CONVEX LOSS FUNCTION

We first verify the inequality (19) by observing the following expression for the gradient of performative loss:

$$\nabla V(\theta) = \nabla \int_{\mathsf{Z}} \ell(\theta; z) p_{\mathcal{D}(\theta)}(z) \mathrm{d}\, z = \mathbb{E}_{z \sim \mathcal{D}(\theta)}[\nabla \ell(\theta; z)] + \mathbb{E}_{z \sim \mathcal{D}(\theta)}\big[\ell(\theta; z) \nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\big], \tag{38}$$

where we have denoted $p_{\mathcal{D}(\theta)}(z)$ as the probability distribution function for $\mathcal{D}(\theta)$. The above identity is derived using chain rule and the property $\nabla_\theta \log p_{\mathcal{D}(\theta)}(z) = \frac{\nabla_\theta p_{\mathcal{D}(\theta)}(z)}{p_{\mathcal{D}(\theta)}(z)}$ similar to the policy gradient theorem; see [Sutton and Barto, 2018, Ch. 13].

Observe that

$$\langle \nabla V(\theta) \,|\, h(\theta) \rangle = \|h(\theta)\|^2 + \langle \mathbb{E}_{z \sim \mathcal{D}(\theta)}\big[\ell(\theta; z) \nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\big] \,|\, h(\theta) \rangle \tag{39}$$

We note

$$|\langle \mathbb{E}_{z\sim\mathcal{D}(\theta)}\big[\ell(\theta;z)\nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\big] \,|\, h(\theta)\rangle| \leq \frac{1}{2}\|h(\theta)\|^2 + \frac{1}{2}\|\mathbb{E}_{z\sim\mathcal{D}(\theta)}\big[\ell(\theta;z)\nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\big]\|^2$$
$$\leq \frac{1}{2}\|h(\theta)\|^2 + \frac{1}{2}\mathbb{E}_{z\sim\mathcal{D}(\theta)}\big[|\ell(\theta;z)|^2\|\nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\|^2\big] \leq \frac{1}{2}\|h(\theta)\|^2 + \mathrm{c}_0$$

where we have used the Jensen's inequality and set

$$\mathrm{c}_0 := \sup_{\theta\in\mathbb{R}^d} \frac{1}{2}\mathbb{E}_{z\sim\mathcal{D}(\theta)}\big[|\ell(\theta;z)|^2\|\nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\|^2\big]. \tag{40}$$

The above can be shown to be bounded when the loss function is bounded (e.g., a sigmoid loss), and the state dependent distribution has bounded gradient w.r.t. θ (e.g., a soft-max distribution). Together, we obtain the desired inequality:

$$\langle \nabla V(\theta) \,|\, h(\theta)\rangle \geq \frac{1}{2}\|h(\theta)\|^2 - \mathrm{c}_0,\ \forall\, \theta \in \mathbb{R}^d. \tag{41}$$

Notice that (38) also implies

$$\begin{aligned}\|\nabla V(\theta)\| &\leq \|h(\theta)\| + \|\mathbb{E}_{z\sim\mathcal{D}(\theta)}\big[\ell(\theta;z)\nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\big]\| \\ &\leq \|h(\theta)\| + \mathbb{E}_{z\sim\mathcal{D}(\theta)}\big[|\ell(\theta;z)|\|\nabla_\theta \log(p_{\mathcal{D}(\theta)}(z))\|\big] \leq \|h(\theta)\| + \sqrt{2\mathrm{c}_0}.\end{aligned} \tag{42}$$

Proof of Corollary 1 Notice that (41), (42) imply A1, A2 of [Karimi et al., 2019], respectively. Moreover, the stated assumptions in the corollary imply A3, A5-A7 of [Karimi et al., 2019]. Applying Theorem 2 from [Karimi et al., 2019] shows that

$$\mathbb{E}[\|\nabla V(\theta_{\mathsf{K}})\|^2] \lesssim \mathbb{E}[\|h(\theta_{\mathsf{K}})\|^2] + \mathrm{c}_0 \lesssim \frac{1+\sum_{k=1}^K \gamma_k^2}{\sum_{k=1}^K \gamma_k} + \mathrm{c}_0, \tag{43}$$

where we have omitted the constants from [Karimi et al., 2019]. Note that $\mathsf{K} \in \{1, ..., K\}$ is a discrete r.v. selected independently with the probability $\mathbb{P}(\mathsf{K} = k) = \gamma_k / \sum_{j=1}^K \gamma_j$. Setting the step sizes as $\gamma_k = \mathcal{O}(1/\sqrt{k})$ shows the desired bound in the corollary.

C MISSING PROOFS IN SECTION 3 & 4

Below, we present the detailed proof for the lemmas presented in §4.

C.1 Constants $\widehat{L}, \overline{L}$ in (18)

Below we derive the expressions for the two constants in (18) under Assumptions 2, 3, 4. Furthermore, we assume that the MC induced by P_θ are uniform geometric ergodic such that there exists $\rho \in [0, 1)$, $K \geq 0$, such that for any $n \in \mathbb{N}$,

$$\sup_{\theta\in\mathbb{R}^d, z\in\mathsf{Z}} \|\mathrm{P}_\theta^n(z,\cdot) - \pi_\theta(\cdot)\|_{\mathrm{TV}} \leq \rho^n K, \tag{44}$$

where $\|\cdot\|_{\mathrm{TV}}$ is total variation norm.

Firstly, we observe that under Assumption 4 and by [Douc et al., 2018, Proposition 21.2.3], the solution to the Poisson equation (15) is given as:

$$\widehat{\nabla}\ell(\theta;z) = \sum_{n=0}^{\infty} \mathrm{P}_\theta^n\big(\nabla\ell(\theta;z) - \nabla f(\theta;\theta)\big) = \sum_{n=0}^{\infty}\big\{\mathrm{P}_\theta^n - \pi_\theta\big\}\big(\nabla\ell(\theta;z) - \nabla f(\theta;\theta)\big). \tag{45}$$

for any $\theta \in \mathbb{R}^d$, $z \in \mathsf{Z}$, since we noted that $\pi_\theta(\nabla\ell(\theta;z)) = \pi_\theta(\nabla f(\theta;\theta)) = \nabla f(\theta;\theta)$. As such, for any $\theta \in \mathbb{R}^d$, $z \in \mathsf{Z}$,

$$\begin{aligned}\|\widehat{\nabla}\ell(\theta;z)\| &\leq \sum_{n=0}^{\infty}\|\mathrm{P}_\theta^n(z,\cdot) - \pi_\theta(\cdot)\|_{\mathrm{TV}} \sup_{z\in\mathsf{Z}}\|\nabla\ell(\theta;z) - \nabla f(\theta;\theta)\| \\ &\leq \sum_{n=0}^{\infty}\rho^n K\sigma(1 + \|\theta - \theta_{PS}\|) = \frac{K\sigma}{1-\rho}(1 + \|\theta - \theta_{PS}\|),\end{aligned} \tag{46}$$

where the last inequality is due to Assumption 3 and (44). Similarly, it can be shown that $\|\mathrm{P}_\theta \widehat{\nabla}\ell(\theta; z)\| \leq \frac{\rho K \sigma}{1-\rho}(1 + \|\theta - \theta_{PS}\|)$. Secondly, we observe for any $\theta \in \mathbb{R}^d$, $z \in \mathsf{Z}$, it holds

$$\begin{aligned} \|\nabla \ell(\theta; z)\| &\leq \|\nabla \ell(\theta_{PS}; z) - \nabla \ell(\theta; z)\| + \|\nabla \ell(\theta_{PS}; z)\| \\ &\overset{(a)}{\leq} L\|\theta - \theta_{PS}\| + \|\nabla \ell(\theta_{PS}; z) - \nabla f(\theta_{PS}; \theta_{PS})\| \overset{(b)}{\leq} L\|\theta - \theta_{PS}\| + \sigma, \end{aligned} \tag{47}$$

where (a) is due to Assumption 2 and the fact $\nabla f(\theta_{PS}; \theta_{PS}) = 0$ and (b) is due to Assumption 3. From the above observations, we yield the following constants for (18):

$$\widehat{L} = \frac{K\sigma}{1-\rho}, \quad \overline{L} = \max\{L, \sigma\}. \tag{48}$$

C.2 Proof of Lemma 1

We begin our analysis by observing that as $\nabla f(\theta_{PS}; \theta_{PS}) = 0$, we have:

$$\begin{aligned} &\|\theta_{k+1} - \theta_{PS}\|^2 = \|\theta_k - \gamma_{k+1}\nabla \ell(\theta_k; z_{k+1}) - \theta_{PS}\|^2 \\ &= \underbrace{\|\theta_k - \theta_{PS}\|^2}_{=:B_1} - 2\gamma_{k+1} \underbrace{\langle \theta_k - \theta_{PS} \,|\, \nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_{PS}; \theta_{PS}) \rangle}_{=:B_2} \\ &\quad + \gamma_{k+1}^2 \underbrace{\|\nabla f(\theta_{PS}; \theta_{PS}) - \nabla \ell(\theta_k; z_{k+1})\|^2}_{=:B_3} \end{aligned}$$

The inner product can be lower bounded as

$$\begin{aligned} B_2 &= \langle \theta_k - \theta_{PS} \,|\, \nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_{PS}; \theta_{PS}) \rangle \\ &= \langle \theta_k - \theta_{PS} \,|\, \nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_k; \theta_k) \rangle + \langle \theta_k - \theta_{PS} \,|\, \nabla f(\theta_k; \theta_k) - \nabla f(\theta_k; \theta_{PS}) \rangle \\ &\quad + \langle \theta_k - \theta_{PS} \,|\, \nabla f(\theta_k; \theta_{PS}) - \nabla f(\theta_{PS}; \theta_{PS}) \rangle \\ &\overset{(a)}{\geq} \langle \theta_k - \theta_{PS} \,|\, \nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_k; \theta_k) \rangle \\ &\quad - \|\theta_k - \theta_{PS}\| \, \|\nabla f(\theta_k; \theta_k) - \nabla f(\theta_k; \theta_{PS})\| + \mu \|\theta_k - \theta_{PS}\|^2 \\ &\overset{(b)}{\geq} \langle \theta_k - \theta_{PS} | \nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_k; \theta_k) \rangle + (\mu - L\varepsilon) \|\theta_k - \theta_{PS}\|^2 \end{aligned} \tag{49}$$

where (a) is due to the Cauchy-schwarz inequality and the μ-strong convexity of $\nabla f(\cdot;\cdot)$; (b) is due to the L-smoothness of f and the ϵ-sensitivity of the distribution [c.f Assumption 6]; also see Perdomo et al. [2020]. Furthermore,

$$\begin{aligned} B_3 &= \|\nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_{PS}; \theta_{PS}) + \nabla \ell(\theta_{PS}; z_{k+1}) - \nabla \ell(\theta_{PS}; z_{k+1})\|^2 \\ &\leq 2\left(\|\nabla \ell(\theta_{PS}; z_{k+1}) - \nabla \ell(\theta_k; z_{k+1})\|^2 + \|\nabla f(\theta_{PS}; \theta_{PS}) - \nabla \ell(\theta_{PS}; z_{k+1})\|^2 \right) \\ &\leq 2L^2 \|\theta_k - \theta_{PS}\|^2 + 2\sigma^2 \end{aligned} \tag{50}$$

where the third inequality is due to Assumptions 2, 3. Combing the bounds for B_1, B_2 and B_3, we can get the desired inequality.

$$\begin{aligned} &\|\theta_{k+1} - \theta_{PS}\|^2 \\ &\leq \|\theta_k - \theta_{PS}\|^2 + 2\gamma_{k+1}^2 \cdot \left(\sigma^2 + L^2 \|\theta_k - \theta_{PS}\|^2 \right) \\ &\quad - 2\gamma_{k+1} \left(\langle \theta_k - \theta_{PS} | \nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_k; \theta_k) \rangle + (\mu - L\varepsilon) \|\theta_k - \theta_{PS}\|^2 \right) \\ &= \left(1 - 2\gamma_{k+1}(\mu - L\epsilon) + 2\gamma_{k+1}^2 L^2\right) \|\theta_k - \theta_{PS}\|^2 \\ &\quad + 2\gamma_{k+1}^2 \sigma^2 - 2\gamma_{k+1} \langle \theta_k - \theta_{PS} \,|\, \nabla \ell(\theta_k; z_{k+1}) - \nabla f(\theta_k; \theta_k) \rangle. \end{aligned} \tag{51}$$

It is noted that if we consider a case when the SA scheme (4) is non-state-dependent, e.g., z_{k+1} is drawn from $\mathcal{D}(\theta_k)$ independently, then proving Lemma 1 suffices to show our desired Theorem 1 since the last term in equation (21) is zero mean when conditioned on the previous iterates [cf. (22)].

C.3 Proof of Lemma 2

Applying Assumption 4 shows that the sum of inner product can be evaluated as

$$\begin{aligned}&\sum_{s=1}^{k} \gamma_s G_{s+1:k} \mathbb{E}\langle \theta_{PS} - \theta_{s-1} \,|\, \nabla \ell(\theta_{s-1}; z_s) - \nabla f(\theta_{s-1}; \theta_{s-1}) \rangle \\ &= \sum_{s=1}^{k} \gamma_s G_{s+1:k} \mathbb{E}\langle \theta_{PS} - \theta_{s-1} \,|\, \widehat{\nabla \ell}(\theta_{s-1}; z_s) - \mathrm{P}_{\theta_{s-1}} \widehat{\nabla \ell}(\theta_{s-1}; z_s) \rangle \equiv \mathbb{E}\left(A_1 + A_2 + A_3 + A_4 + A_5\right),\end{aligned}$$

where we decomposed the sum of inner product into five sub-terms A_1, A_2, A_3, A_4, A_5 such that

$$\begin{aligned}A_1 &:= -\sum_{s=2}^{k} \gamma_s G_{s+1:k} \langle \theta_{s-1} - \theta_{PS} \,|\, \widehat{\nabla \ell}(\theta_{s-1}; z_s) - \mathrm{P}_{\theta_{s-1}} \widehat{\nabla \ell}(\theta_{s-1}; z_{s-1}) \rangle \\ A_2 &:= -\sum_{s=2}^{k} \gamma_s G_{s+1:k} \langle \theta_{s-1} - \theta_{PS} \,|\, \mathrm{P}_{\theta_{s-1}} \widehat{\nabla \ell}(\theta_{s-1}; z_{s-1}) - \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \rangle \\ A_3 &:= -\sum_{s=2}^{k} \gamma_s G_{s+1:k} \langle \theta_{s-1} - \theta_{s-2} \,|\, \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \rangle \\ A_4 &:= -\sum_{s=2}^{k} (\gamma_s G_{s+1:k} - \gamma_{s-1} G_{s:k}) \langle \theta_{s-2} - \theta_{PS} \,|\, \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \rangle \\ A_5 &:= -\gamma_1 G_{2:k} \langle \theta_0 - \theta_{PS} \,|\, \widehat{\nabla \ell}(\theta_0; z_1) \rangle + \gamma_k \langle \theta_{k-1} - \theta_{PS} \,|\, \mathrm{P}_{\theta_{k-1}} \widehat{\nabla \ell}(\theta_{k-1}; z_k) \rangle.\end{aligned}$$

We remark that a similar decomposition can be found in [Benveniste et al., 2012]. However, Benveniste et al. [2012] proceeded with the analysis by assuming that θ_k stays in the compact set for all $k \geq 0$. We do not make such assumption in this work.

For A_1, we note that $\widehat{\nabla \ell}(\theta_{s-1}; z_s) - \mathrm{P}_{\theta_{s-1}} \widehat{\nabla \ell}(\theta_{s-1}; z_{s-1})$ is a martingale difference sequence and therefore we have $\mathbb{E}\left[A_1\right] = 0$ by taking the total expectation.

For A_2, as $\theta_{k+1} = \theta_k - \gamma_{k+1} \nabla \ell\left(\theta_k; z_{k+1}\right)$, we get $\theta_{s-1} - \theta_{s-2} = -\gamma_{s-1} \nabla \ell(\theta_{s-2}; z_{s-1})$. Applying the smoothness condition Assumption 5 shows that

$$\begin{aligned}A_2 &= -\sum_{s=2}^{k} \gamma_s G_{s+1:k} \langle \theta_{s-1} - \theta_{PS} \,|\, \mathrm{P}_{\theta_{s-1}} \widehat{\nabla \ell}(\theta_{s-1}; z_{s-1}) - \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \rangle \\ &\leq L_{PH} \sum_{s=2}^{k} \gamma_s G_{s+1:k} \left\|\theta_{s-1} - \theta_{PS}\right\| \left\|\theta_{s-1} - \theta_{s-2}\right\| \\ &\leq L_{PH} \sum_{s=2}^{k} \gamma_{s-1} \gamma_s G_{s+1:k} \left\|\theta_{s-1} - \theta_{PS}\right\| \left\|\nabla \ell(\theta_{s-2}; z_{s-1})\right\|.\end{aligned} \tag{52}$$

Combining with the implied bound (18) from the assumptions as well as (8) yield

$$\begin{aligned}A_2 &\leq \varsigma L_{PH} \overline{L} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \|\theta_{s-1} - \theta_{PS}\| \left(1 + \|\theta_{s-2} - \theta_{PS}\|\right) \\ &\leq \varsigma L_{PH} \overline{L} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \left\{ \frac{1}{2} + \frac{1}{2} \|\theta_{s-2} - \theta_{PS}\|^2 + \|\theta_{s-1} - \theta_{PS}\|^2 \right\} \\ &\leq \varsigma L_{PH} \overline{L} \left\{ \frac{1}{2} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} + \frac{1}{2} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \|\theta_{s-2} - \theta_{PS}\|^2 + \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \|\theta_{s-1} - \theta_{PS}\|^2 \right\},\end{aligned}$$

where the second inequality applies $a(1+c) \leq \frac{1}{2} + \frac{1}{2}c^2 + a^2$ for any $a, c \in \mathbb{R}$.

For A_3, again using (18), we observe that

$$
\begin{aligned}
A_3 &= -\sum_{s=2}^{k} \gamma_s G_{s+1:k} \langle \theta_{s-1} - \theta_{s-2} \,|\, \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \rangle \\
&\leq \sum_{s=2}^{k} \gamma_s G_{s+1:k} \, \|\theta_{s-1} - \theta_{s-2}\| \cdot \left\| \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \right\| \\
&\leq \sum_{s=2}^{k} \gamma_s \gamma_{s-1} G_{s+1:k} \, \|\nabla \ell(\theta_{s-2}; z_{s-1})\| \cdot \widehat{L} \, (1 + \|\theta_{s-2} - \theta_{PS}\|) \\
&\leq \varsigma \overline{L} \widehat{L} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} (1 + \|\theta_{s-2} - \theta_{PS}\|)^2 \\
&\leq 2\varsigma \overline{L} \widehat{L} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \{1 + \|\theta_{s-2} - \theta_{PS}\|^2\}.
\end{aligned} \tag{53}
$$

For A_4, we notice that

$$
\begin{aligned}
A_4 &= -\sum_{s=2}^{k} (\gamma_s G_{s+1:k} - \gamma_{s-1} G_{s:k}) \langle \theta_{s-2} - \theta_{PS} \,|\, \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \rangle \\
&\leq \sum_{s=2}^{k} |\gamma_s G_{s+1:k} - \gamma_{s-1} G_{s:k}| \, \|\theta_{s-2} - \theta_{PS}\| \cdot \left\| \mathrm{P}_{\theta_{s-2}} \widehat{\nabla \ell}(\theta_{s-2}; z_{s-1}) \right\|.
\end{aligned} \tag{54}
$$

It can be shown that $|\gamma_s G_{s+1:k} - \gamma_{s-1} G_{s:k}| \leq (1 + \tilde{\mu}) \varsigma \gamma_s^2 G_{s+1:k}$, therefore

$$
\begin{aligned}
A_4 &\leq (1 + \tilde{\mu}) \varsigma \widehat{L} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \, \|\theta_{s-2} - \theta_{PS}\| \, (1 + \|\theta_{s-2} - \theta_{PS}\|) \\
&\leq (1 + \tilde{\mu}) \varsigma \widehat{L} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \Big\{ \frac{1}{2} + \frac{3}{2} \|\theta_{s-2} - \theta_{PS}\|^2 \Big\} \\
&\leq (1 + \tilde{\mu}) \varsigma \widehat{L} \Big\{ \frac{1}{2} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} + \frac{3}{2} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \|\theta_{s-2} - \theta_{PS}\|^2 \Big\}.
\end{aligned} \tag{55}
$$

Finally, for A_5, we have

$$
\begin{aligned}
A_5 &= -\gamma_1 G_{2:k} \langle \theta_0 - \theta_{PS} \,|\, \widehat{\nabla \ell}(\theta_0; z_1) \rangle + \gamma_k \langle \theta_{k-1} - \theta_{PS} \,|\, \mathrm{P}_{\theta_{k-1}} \widehat{\nabla \ell}(\theta_{k-1}; z_k) \rangle \\
&\leq \gamma_1 G_{2:k} \, \|\theta_0 - \theta_{PS}\| \left\| \widehat{\nabla \ell}(\theta_0; z_1) \right\| + \gamma_k \, \|\theta_{k-1} - \theta_{PS}\| \left\| \mathrm{P}_{\theta_{k-1}} \widehat{\nabla \ell}(\theta_{k-1}; z_k) \right\| \\
&\leq \gamma_1 \widehat{L} \, G_{2:k} \|\theta_0 - \theta_{PS}\| \big(1 + \|\theta_0 - \theta_{PS}\|\big) + \gamma_k \widehat{L} \, \|\theta_{k-1} - \theta_{PS}\| \big(1 + \|\theta_{k-1} - \theta_{PS}\|\big) \\
&\leq \frac{\gamma_1 \widehat{L} \, G_{2:k}}{2} + \frac{\gamma_k \widehat{L}}{2} + \frac{3\gamma_1 \widehat{L}}{2} \, G_{2:k} \|\theta_0 - \theta_{PS}\|^2 + \frac{3\gamma_k \widehat{L}}{2} \, \|\theta_{k-1} - \theta_{PS}\|^2
\end{aligned}
$$

Summing up A_1 to A_5 and taking the full expectation yield:

$$
\begin{aligned}
& 2 \big| \mathbb{E} \left[A_1 + A_2 + A_3 + A_4 + A_5 \right] \big| \\
& \leq \varsigma L_{PH} \overline{L} \left\{ \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} + \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \Delta_{s-2} + 2 \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \Delta_{s-1} \right\} \\
& \quad + 4\varsigma \overline{L} \widehat{L} \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \big\{1 + \Delta_{s-2}\big\} + (1 + \tilde{\mu}) \varsigma \widehat{L} \Big\{ \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} + 3 \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k} \Delta_{s-2} \Big\} \\
& \quad + \gamma_1 \widehat{L} \, G_{2:k} + \gamma_k \widehat{L} + 3\gamma_1 \widehat{L} \, G_{2:k} \Delta_0 + 3\gamma_k \widehat{L} \, \Delta_{k-1}.
\end{aligned}
$$

Recall the following constants:

$$\mathcal{C}_1 := \varsigma L_{PH}\overline{L} + 4\varsigma\overline{L}\widehat{L} + (1+\tilde{\mu})\varsigma\widehat{L}, \quad \mathcal{C}_2 := 2\varsigma L_{PH}\overline{L}, \quad \mathcal{C}_3 := \varsigma L_{PH}\overline{L} + 4\varsigma\overline{L}\widehat{L} + 3(1+\tilde{\mu})\varsigma\widehat{L}. \tag{56}$$

We obtain the desirable bound for the lemma:

$$\begin{aligned}
&2\big|\mathbb{E}\left[A_1 + A_2 + A_3 + A_4 + A_5\right]\big| \\
&\leq \sum_{s=2}^{k} \gamma_s^2 G_{s+1:k}\big(\mathcal{C}_1 + \mathcal{C}_2\Delta_{s-1} + \mathcal{C}_3\Delta_{s-2}\big) + \widehat{L}\gamma_k\big\{1 + 3\Delta_{k-1}\big\} + \gamma_1 G_{2:k}\big(\widehat{L}(1+3\Delta_0) + \gamma_1\mathcal{C}_1\big).
\end{aligned}$$

This concludes the proof.

C.4 Proof of Lemma 3

Consider the inequality in (25). We consider a non-negative upper bound sequence $\{\mathrm{U}_k\}_{k\geq 0}$ defined by the recursion:

$$\begin{aligned}
\mathrm{U}_k &= G_{1:k}\mathrm{U}_0 + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\gamma_k + \sum_{s=1}^{k-1}\gamma_{s+1}^2 G_{s+2:k}\big(\mathcal{C}_2\mathrm{U}_s + \mathcal{C}_3\mathrm{U}_{s-1}\big) \\
&\quad + \gamma_1 G_{2:k}\big\{\widehat{L}(1+3\mathrm{U}_0) + \gamma_1(2\sigma^2 + \mathcal{C}_1)\big\} + 3\gamma_k\widehat{L}\mathrm{U}_{k-1},
\end{aligned} \tag{57}$$

for any $k \geq 1$, and we have defined $\mathrm{U}_0 = \Delta_0$. Notice that by construction, we have $\Delta_k \leq \mathrm{U}_k$ for any $k \geq 0$.

Using the convention that $\mathrm{U}_{-1} = 0$, we observe that for any $k \geq 1$,

$$\begin{aligned}
\mathrm{U}_k &= (1-\gamma_k\tilde{\mu})\mathrm{U}_{k-1} + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\Big(\gamma_k - (1-\gamma_k\tilde{\mu})\gamma_{k-1}\Big) \\
&\quad + \gamma_k^2(\mathcal{C}_2\mathrm{U}_{k-1} + \mathcal{C}_3\mathrm{U}_{k-2}) + 3\widehat{L}\Big(\gamma_k\mathrm{U}_{k-1} - (1-\gamma_k\tilde{\mu})\gamma_{k-1}\mathrm{U}_{k-2}\Big) \\
&\leq \big(1 - \gamma_k\tilde{\mu} + \gamma_k^2\mathcal{C}_2\big)\mathrm{U}_{k-1} + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\tilde{\mu}\varsigma\gamma_k^2 + \Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\gamma_k\gamma_{k-1}\mathrm{U}_{k-2} \\
&\quad + 3\widehat{L}\big(\gamma_k\mathrm{U}_{k-1} - \gamma_{k-1}\mathrm{U}_{k-2}\big) \\
&\leq \big(1 - \gamma_k\tilde{\mu}/2\big)\mathrm{U}_{k-1} + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\tilde{\mu}\varsigma\gamma_k^2 + \Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\gamma_k\gamma_{k-1}\mathrm{U}_{k-2} \\
&\quad + 3\widehat{L}\big(\gamma_k\mathrm{U}_{k-1} - \gamma_{k-1}\mathrm{U}_{k-2}\big),
\end{aligned} \tag{58}$$

where the last inequality is due to $\gamma_k \leq \tilde{\mu}/2\mathcal{C}_2$.

We prove part (i) of the lemma. From (58), we consider an upper bound sequence $\{\overline{\mathrm{U}}_k\}_{k\geq -1}$ defined by the recursion:

$$\begin{aligned}
\overline{\mathrm{U}}_k &= \big(1 - \gamma_k\tilde{\mu}/2\big)\overline{\mathrm{U}}_{k-1} + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\tilde{\mu}\varsigma\gamma_k^2 + \Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\gamma_k\gamma_{k-1}\overline{\mathrm{U}}_{k-2} \\
&\quad + 3\widehat{L}\big(\gamma_k\overline{\mathrm{U}}_{k-1} - \gamma_{k-1}\overline{\mathrm{U}}_{k-2}\big), \quad \forall k \geq 1.
\end{aligned} \tag{59}$$

We have also defined $\overline{\mathrm{U}}_0 = \mathrm{U}_0$, $\overline{\mathrm{U}}_{-1} = 0$. For any $t \geq 1$, summing up the equation (59) from $k = 1$ to $k = t$ yields

$$\begin{aligned}
\sum_{k=1}^{t}\overline{\mathrm{U}}_k &= \sum_{k=1}^{t}\Big\{\big(1 - \gamma_k\tilde{\mu}/2\big)\overline{\mathrm{U}}_{k-1} + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\tilde{\mu}\varsigma\gamma_k^2 + \Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\gamma_k\gamma_{k-1}\overline{\mathrm{U}}_{k-2}\Big\} \\
&\quad + 3\widehat{L}\sum_{k=1}^{t}\big(\gamma_k\overline{\mathrm{U}}_{k-1} - \gamma_{k-1}\overline{\mathrm{U}}_{k-2}\big),
\end{aligned}$$

Rearranging terms leads to

$$\overline{\mathrm{U}}_t = \overline{\mathrm{U}}_0 + \sum_{k=1}^{t}\Big\{\Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\gamma_k\gamma_{k-1}\overline{\mathrm{U}}_{k-2} + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L}\Big)\tilde{\mu}\varsigma\gamma_k^2 - \frac{\tilde{\mu}}{2}\gamma_k\overline{\mathrm{U}}_{k-1}\Big\} + 3\widehat{L}\gamma_t\overline{\mathrm{U}}_{t-1}$$

Using the step size conditions $\gamma_k \leq \gamma_{k-1}$, $\gamma_k \leq \big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\big)^{-1} \min\{\tilde{\mu}/2, 3\widehat{L}\}$, $\gamma_k \leq (6\widehat{L})^{-1}$,

$$
\begin{aligned}
\overline{\mathrm{U}}_t &\leq \overline{\mathrm{U}}_0 + 3\widehat{L}\gamma_t \overline{\mathrm{U}}_{t-1} \\
&\quad + \sum_{k=1}^{t} \Big\{ \Big[\Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\gamma_k^2 - \frac{\tilde{\mu}}{2}\gamma_k \Big] \overline{\mathrm{U}}_{k-1} + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L} \Big) \tilde{\mu}\varsigma\gamma_k^2 \Big\} \\
&\leq 3\widehat{L}\gamma_t \overline{\mathrm{U}}_{t-1} + \overline{\mathrm{U}}_0 + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L} \Big) \tilde{\mu}\varsigma \sum_{k=1}^{t} \gamma_k^2 \\
&\leq \frac{1}{2}\overline{\mathrm{U}}_{t-1} + \Big\{ \overline{\mathrm{U}}_0 + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L} \Big) \tilde{\mu}\varsigma \sum_{k=1}^{t} \gamma_k^2 \Big\},
\end{aligned}
\tag{60}
$$

where we obtain the first inequality after shifting the summation's index and it is noted that $\overline{\mathrm{U}}_{-1} = 0$. Rearranging terms and solving the recursion lead to

$$
\begin{aligned}
\overline{\mathrm{U}}_t &\leq \big(\tfrac{1}{2}\big)^t \overline{\mathrm{U}}_0 + \sum_{s=1}^{t} \big(\tfrac{1}{2}\big)^{t-s} \Big\{ \overline{\mathrm{U}}_0 + \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L} \Big) \tilde{\mu}\varsigma \sum_{k=1}^{s} \gamma_k^2 \Big\} \\
&\leq 3\overline{\mathrm{U}}_0 + 2\tilde{\mu}\varsigma \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L} \Big) \sum_{\ell=1}^{t} \gamma_\ell^2 \big(\tfrac{1}{2}\big)^{t-\ell} \\
&\leq 3\overline{\mathrm{U}}_0 + \frac{\tilde{\mu}\varsigma}{9\widehat{L}^2} \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L} \Big)
\end{aligned}
\tag{61}
$$

Recall that $\overline{\Delta} := 3\overline{\mathrm{U}}_0 + \frac{\varsigma}{9\widehat{L}^2}\Big(2(2\sigma^2 + \mathcal{C}_1) + \tilde{\mu}\widehat{L}\Big)$, the above shows $\Delta_t \leq \mathrm{U}_t \leq \overline{\mathrm{U}}_t \leq \overline{\Delta}$ for any $t \geq 1$, thus establishing part (i).

We now proceed to proving part (ii) of the lemma. We define $\overline{G}_{m:n} = \prod_{\ell=m}^{n} (1 - \gamma_\ell \tilde{\mu}/2)$ and observe from (59) that

$$
\begin{aligned}
\mathrm{U}_k &\leq \overline{G}_{1:k}\mathrm{U}_0 + \sum_{s=1}^{k} \overline{G}_{s+1:k} \Big\{ \Big(\frac{2}{\tilde{\mu}}(2\sigma^2 + \mathcal{C}_1) + \widehat{L} \Big) \tilde{\mu}\varsigma\gamma_s^2 + \Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu}\Big)\gamma_s\gamma_{s-1}\mathrm{U}_{s-2} \Big\} \\
&\quad + 3\widehat{L} \sum_{s=1}^{k} \overline{G}_{s+1:k} \Big\{ \Big(\gamma_s \mathrm{U}_{s-1} - \gamma_{s-1}\mathrm{U}_{s-2} \Big) \Big\}.
\end{aligned}
\tag{62}
$$

Notice that

$$
\begin{aligned}
&\sum_{s=1}^{k} \overline{G}_{s+1:k} \big(\gamma_s \mathrm{U}_{s-1} - \gamma_{s-1}\mathrm{U}_{s-2} \big) \\
&= \sum_{s=1}^{k} \overline{G}_{s+1:k} \big(\gamma_s \mathrm{U}_{s-1} + (1 - \gamma_s\tilde{\mu}/2)(\gamma_{s-1}\mathrm{U}_{s-2} - \gamma_{s-1}\mathrm{U}_{s-2}) - \gamma_{s-1}\mathrm{U}_{s-2} \big) \\
&= \sum_{s=1}^{k} \Big\{ \Big(\overline{G}_{s+1:k}\gamma_s\mathrm{U}_{s-1} - \overline{G}_{s:k}\gamma_{s-1}\mathrm{U}_{s-2} \Big) - \gamma_s\gamma_{s-1}\tilde{\mu}\mathrm{U}_{s-2}/2 \Big\} \leq \gamma_k \mathrm{U}_{k-1} \leq \gamma_k \overline{\Delta}.
\end{aligned}
\tag{63}
$$

By Lemma 4, we have $\sum_{s=1}^{k} \overline{G}_{s+1:k}\gamma_s^2 \leq 4\gamma_k/\tilde{\mu}$ and the following is obtained

$$
\mathrm{U}_k \leq \overline{G}_{1:k}\mathrm{U}_0 + \Big\{ \frac{4\varsigma}{\tilde{\mu}} \Big(2(2\sigma^2 + \mathcal{C}_1) + \tilde{\mu}\widehat{L} \Big) + 3\widehat{L}\overline{\Delta} + \frac{4\varsigma}{\tilde{\mu}} \Big(\mathcal{C}_3 + 3\widehat{L}\tilde{\mu} \Big) \overline{\Delta} \Big\} \gamma_k.
\tag{64}
$$

The proof is completed.

C.5 Auxiliary Lemmas

Lemma 4. *Let $a > 0$ and $(\gamma_k)_{k \geq 1}$ be a non-increasing sequence such that $\gamma_1 < 2/a$. If $\gamma_{k-1}/\gamma_k \leq 1 + (a/2)\gamma_k$ for any $k \geq 1$, then for any $k \geq 2$,*

$$
\sum_{j=1}^{k} \gamma_j^2 \prod_{\ell=j+1}^{k} (1 - \gamma_\ell a) \leq \frac{2}{a}\gamma_k.
\tag{65}
$$

Proof. The proof is elementary. Observe that:

$$\begin{aligned}
\sum_{j=1}^{k} \gamma_j^2 \prod_{\ell=j+1}^{k} (1-\gamma_\ell a) &= \gamma_k \sum_{j=1}^{k} \gamma_j \prod_{\ell=j+1}^{k} \frac{\gamma_{\ell-1}}{\gamma_\ell} (1-\gamma_\ell a) \\
&\leq \gamma_k \sum_{j=1}^{k} \gamma_j \prod_{\ell=j+1}^{k} (1+(a/2)\gamma_\ell)(1-\gamma_\ell a) \\
&\leq \gamma_k \sum_{j=1}^{k} \gamma_j \prod_{\ell=j+1}^{k} (1-\gamma_\ell (a/2)) \\
&= \frac{2\gamma_k}{a} \sum_{j=1}^{k} \left(\prod_{\ell=j+1}^{k} (1-\gamma_\ell a/2) - \prod_{\ell'=j}^{k} (1-\gamma_{\ell'} a/2) \right) \\
&= \frac{2\gamma_k}{a} \left(1 - \prod_{\ell'=1}^{k} (1-\gamma_{\ell'} a/2) \right) \leq \frac{2\gamma_k}{a}.
\end{aligned} \tag{66}$$

The proof is concluded. □

D DETAILS OF THE NUMERICAL EXPERIMENTS

This section provides details about the numerical experiments on the second problem of strategic classification (SC) in §5. Moreover, we provide additional experiment results to better illustrate the performance of the state dependent SA algorithm for this problem.

The experiments conducted in this section are based on the Credit simulator provided at `https://github.com/zykls/performative-prediction`. Our experiments are conducted on a server with Intel Xeon Gold 6138 CPU. The Python codes are executed in a single-thread environment.

There are two roles in the SC problem – *learner* and *agents*. The learner utilizes agents' information to obtain a classifier f_θ. Meanwhile, individual agents hope to be assigned to a favorable class. To do so, they modify their features and thereby shifting the data distribution towards the target $\mathcal{D}(\theta)$. Specifically, our experiments are done on the `GiveMeSomeCredit` dataset with $m = 18357$ samples as we select $d = 3$ features to build the classifier. Each (original) data sample is given by $\bar{z}_i = (\bar{x}_i, \bar{y}_i)$ with the label $\bar{y}_i \in \{0, 1\}$ and selected feature $\bar{x}_i \in \mathbb{R}^3$. We associate each data sample to an agent. The task for the learner (bank) is to design a classifier that distinguishes whether the application of an individual (agent) who want to default a loan should be granted or not.

We simulate the *adapted best response* presented in **Example 1** of the main paper. In this setting, the agents rely on their past experience to present data to the learner that is favorable to to agents. The latter is achieved by a gradient descent step that depends on the current learner's state (θ_k), past agent's state (z_k) and the original data ($\mathcal{D}_0$). As the dynamics is coupled between the agents' and learner's update, we present the overall algorithm based on (3), (4) as follows:

Algorithm 2: State Dependent SA with Adapted Best Response.

Input: initial iterate $\theta_0 \in \mathbb{R}^d$, agents' state $x_i^0 = \bar{x}_i$, $i \in \{1, ..., m\}$ such that $\bar{x}_i$ is the ith original feature vector, step sizes $\{\gamma_k\}_{k \geq 0}$, agents' response rate $\alpha > 0$, update parameter b.

For $k = 0, 1, 2, \ldots$

1. A subset of *agents*, $\mathcal{I}_k$ with $|\mathcal{I}_k| = \mathsf{b}$, is selected uniformly from $\{1, ..., m\}$. They adapt their feature vectors based on past experience and θ_k as:

$$x_i^{k+1} = x_i^k + \alpha \nabla U(x_i^k; \bar{z}_i, \theta_k),\ \forall\, i \in \mathcal{I}_k, \quad x_i^{k+1} = x_i^k,\ \forall\, i \notin \mathcal{I}_k. \tag{67}$$

2. An agent $i_k \in \{1, ..., m\}$ is drawn uniformly to present data. Set $z_{k+1} = (x_{i_k}^{k+1}, y_{i_k})$.
3. The *learner* computes the $k+1$th iterate by:

$$\theta_{k+1} = \theta_k - \gamma_{k+1} \nabla \ell(\theta_k; z_{k+1}).$$

The most recent iterate θ_{k+1} is deployed and made available to the agent(s).

Steps 1 & 2 in Algorithm 2 resemble the adaptive best response update in (7). We emphasize that these two steps are *agnostic* to the learner as the latter only sees z_{k+1} at iteration k, similarly, the last step is not known to the agents as the latter only sees the classifier given as θ_{k+1}.

Furthermore, we recall that the following two types of utility functions are considered as $U(\cdot)$:

$$\begin{aligned} U_{\mathsf{q}}(x'; z, \theta) &= \langle \theta \,|\, x' \rangle - \frac{\|x' - x\|^2}{2\epsilon}, \\ U_{\mathsf{lg}}(x'; z, \theta) &= y \langle \theta \,|\, x' \rangle - \log\left(1 + \exp(\langle \theta \,|\, x' \rangle)\right) - \frac{\|x' - x\|^2}{2\epsilon}. \end{aligned} \tag{68}$$

In step 1, the agents' response rate α and parameter b control the speed of adaptation among the group of m agents. These parameters will affect the mixing time of the MC which determines the bounds in Theorem 1. Overall, we observe that the agents' states and learner's iterates are evolving simultaneously, highlighting the coupled nature in the analysis of the state dependent SA algorithm.

In cases such as $U_{\mathsf{lg}}(\cdot)$ where the ideal best response $\arg\max_{x'} U(x'; z, \theta)$ must be obtained via an iterative algorithm. From an algorithmic standpoint, the stateful nature for the agent is necessary for the performative prediction algorithm to converge to θ_{PS}.

Verification of Assumption 5 We demonstrate that Assumption 5 holds for a special case of our numerical experiments on strategic classification. Here, we consider for simplicity that $\mathsf{b} = m$ such that all the m agents adapt their feature vectors at iteration k with a gradient ascent step[3]. Furthermore, we concentrate on the case with quadratic best response function $U_{\mathsf{q}}(\cdot)$.

Recall that $\ell(\theta; z) = \frac{\beta}{2} \|\theta\|^2 + \log(1 + \exp(\langle \theta \,|\, x \rangle)) - y \langle \theta \,|\, x \rangle$ and our objective function is $\mathbb{E}_{z \sim \mathcal{D}(\theta)} \ell(z; \theta)$. By [Douc et al., 2018, Proposition 21.2.3], the Possion equation (15) admits the following solution:

$$\widehat{\nabla} \ell(\theta; z) = \sum_{k=0}^{\infty} \left(\mathsf{P}_\theta^k - \pi_\theta\right) \{\nabla \ell(\theta; z) - \nabla f(\theta; \theta)\} = \sum_{k=0}^{\infty} \left(\mathsf{P}_\theta^k - \pi_\theta\right) \left[\left(\frac{\exp(\langle \theta \,|\, x \rangle)}{1 + \exp(\langle \theta \,|\, x \rangle)} - y\right) x\right]$$

Notice that $\widetilde{\nabla} \ell(\theta; z) := \left(\frac{\exp(\langle \theta \,|\, x \rangle)}{1 + \exp(\langle \theta \,|\, x \rangle)} - y\right) x$ is an L-Lipschitz continuous map with respect to θ for any $z \in \mathsf{Z}$ and it is bounded by $\sup_{z \in \mathsf{Z}} 2 \|x\|$. For any $\theta, \theta' \in \mathbb{R}^d$, we obtain

$$\begin{aligned} \left\|\mathsf{P}_\theta \widehat{\nabla} \ell(\theta, z) - \mathsf{P}_{\theta'} \widehat{\nabla} \ell(\theta', z)\right\| &\leq \left\|\sum_{k=0}^{\infty} \left\{\mathrm{P}_\theta^k - \pi_\theta - (\mathrm{P}_{\theta'}^k - \pi_{\theta'})\right\} \widetilde{\nabla} \ell(\theta; z)\right\| \\ &\quad + \left\|\sum_{k=0}^{\infty} \left(\mathrm{P}_{\theta'}^k - \pi_{\theta'}\right) \left[\widetilde{\nabla} \ell(\theta; z) - \widetilde{\nabla} \ell(\theta'; z)\right]\right\| \end{aligned} \tag{69}$$

Note that $\left\|\sum_{k=0}^{\infty} \left(\mathrm{P}_{\theta'}^k - \pi_{\theta'}\right) \left[\widetilde{\nabla} \ell(\theta; z) - \widetilde{\nabla} \ell(\theta'; z)\right]\right\| \leq \frac{\overline{\mathrm{C}} L}{1 - \rho} \|\theta - \theta'\|$ as a consequence of the uniform geometric ergodicity of P_θ and the L-Lipschitz continuity of $\widetilde{\nabla} \ell$.

To handle the first term of (69), we define

$$x_i^k(\theta) = \rho^k x_i^0 + (\bar{x}_i + \epsilon \theta)\{1 - \rho^{k+1}\} \quad \text{and} \quad \bar{x}_i(\theta) = \bar{x}_i + \epsilon \theta$$

where $\rho = 1 - \alpha/\epsilon$. For any $k \geq 0$, we note that

$$\left\{\mathrm{P}_{\theta'}^k - \pi_{\theta'}\right\} \widetilde{\nabla} \ell(\theta; z) = \frac{1}{m} \sum_{i=1}^{m} \left\{\left(\frac{\exp(\langle \theta \,|\, x_i^k(\theta') \rangle)}{1 + \exp(\langle \theta \,|\, x_i^k(\theta') \rangle)} - y_i\right) x_i^k(\theta') - \left(\frac{\exp(\langle \theta \,|\, \bar{x}_i(\theta') \rangle)}{1 + \exp(\langle \theta \,|\, \bar{x}_i(\theta') \rangle)} - y_i\right) \bar{x}_i(\theta')\right\}$$

[3] The general case with random BCD involve tedious calculations and is skipped for brevity, yet the general proof idea can be applied.

Taking the Jacobian of the above expression with respect to θ' yields

$$\mathrm{D}_{\theta'}\left(\left\{\mathrm{P}_{\theta'}^k-\pi_{\theta'}\right\}\widetilde{\nabla}\ell(\theta;z)\right)=\frac{\epsilon}{m}\sum_{i=1}^m\left\{y_i\rho^{k+1}+\frac{1-\rho^{k+1}}{1+\exp(-\langle\theta\,|\,x_i^k(\theta')\rangle)}-\frac{1}{1+\exp(-\langle\theta\,|\,\bar{x}_i(\theta')\rangle)}\right\}\mathrm{I}_d$$
$$+\frac{\epsilon}{m}\sum_{i=1}^m\left\{\frac{1-\rho^{k+1}}{(1+\exp(-\langle\theta\,|\,x_i^k(\theta')\rangle))(1+\exp(\langle\theta\,|\,x_i^k(\theta')\rangle))}x_i^k(\theta')\right.$$
$$\left.-\frac{1}{(1+\exp(-\langle\theta\,|\,\bar{x}_i(\theta')\rangle))(1+\exp(\langle\theta\,|\,\bar{x}_i(\theta')\rangle))}\bar{x}_i(\theta')\right\}\theta^\top$$

Using the fact that $\left\|x_i^k(\theta')-\bar{x}_i(\theta')\right\|\leq\rho^k(\left\|x_i^0\right\|+\|\bar{x}_i+\epsilon\theta'\|)$ and the 1-Lipschitz continuity of $\frac{1}{1+e^x}$, $\frac{x}{(1+e^x)(1+e^{-x})}$, it can be shown that the above Jacobian matrix can be bounded for any k,θ',θ as

$$\|\mathrm{D}_{\theta'}\left(\left\{\mathrm{P}_{\theta'}^k-\pi_{\theta'}\right\}\widetilde{\nabla}\ell(\theta;z)\right)\|_{\max}\leq\overline{\mathrm{D}}\,\rho^k, \tag{70}$$

where $\overline{\mathrm{D}}<\infty$ under some regularity conditions[4]. Collecting terms implies that

$$\left\|\left\{\mathrm{P}_{\theta}^k-\pi_\theta-(\mathrm{P}_{\theta'}^k-\pi_{\theta'})\right\}\widetilde{\nabla}\ell(\theta;z)\right\|\leq\overline{\mathrm{D}}\,\rho^k\,\|\theta-\theta'\|\,,$$

and thus

$$\left\|\mathsf{P}_\theta\widehat{\nabla}\ell(\theta,z)-\mathsf{P}_{\theta'}\widehat{\nabla}\ell(\theta',z)\right\|\leq\frac{\overline{\mathrm{C}}L+\overline{\mathrm{D}}}{1-\rho}\,\|\theta-\theta'\|\,.$$

We note that $L_{PH}:=\frac{\overline{\mathrm{C}}L+\overline{\mathrm{D}}}{1-\rho}$ in the last inequality where it depends on the mixing time. This verifies Assumption 5 for this example.

Lastly, we remark that [Karimi et al., 2019, Lemma 7] may be used for verifying Assumption 5. Notice that the assumption A13 therein is implied by our Assumption 2, and the assumption A14 therein is implied by the geometric convergence established in (37). It remains to verify the assumption A12 in [Karimi et al., 2019] to establish Assumption 5. In addition, if the following smoothness condition on the Markov kernel

$$\sup_{z\in\mathsf{Z}}\|\mathrm{P}_\theta(z,\cdot)-\mathrm{P}_{\theta'}(z,\cdot)\|_{\mathsf{TV}}\leq L_P\,\|\theta-\theta'\|\,,\;\forall\,\theta,\theta'\in\mathbb{R}^d \tag{71}$$

holds, then [Karimi et al., 2019, Lemma 7] can be applied. The condition (71) may be applicable when the agents' updates follow a softmax like policy.

Additional Experiments Next, we provide additional experiments to illustrate the performance of the state dependent SA algorithm from a few perspectives that are skipped in the interest of space for the main paper. Unless otherwise specified, we adopt the same parameters set in the experiments presented in the main paper. In particular, we set $\beta=1000/m$ in (28), $\epsilon=0.01$ in the utility functions, and in (7), we set number of selected agents $|\mathcal{I}_k|=5$, agents' response rate $\alpha=0.5\epsilon$. The step size for (3) is $\gamma_k=c_0/(c_1+k)$, $c_0=100/\tilde{\mu}$, $c_1=8L^2/\tilde{\mu}^2$, where $L,\tilde{\mu}$ are estimated as $\sqrt{2\beta m+\|X\|_F^2/2}$, $(1-\epsilon)\beta-\epsilon\|X\|_F^2/4m$, respectively. By default, the SA algorithm is executed as presented in Algorithm 2 with a batch size of $\mathsf{batch}=1$ and the agents perform only $\mathsf{BR}=1$ best response update per SA update in step 3 of Algorithm 2.

Besides, we compare the convergence rates of the algorithms from the perspective of the *agents* – measured by the number of BR updates performed by the agents. This is the setting used in the plot of Fig. 1 (right) and is denoted with the x-axis label of 'no. of agent update'. We also compare the convergence from the perspective of the *learner* – measured by the number of samples requested from the agents by the learner. This setting is denoted with the x-axis label of 'no. of samples drawn by learner'.

Effects of Stateful Updates at Agents Notice that the comparison has been made in Fig. 1 (right). Here, in Fig. 2 we again plot the convergence of the SA algorithm to illustrate the convergence rates from the learner's perspective as well. We observe that the SA algorithms with stateful update converges as k increases. We vary the 'learner's iteration' parameter to observe the effects on convergence when the learner is adapting at faster rate

[4] We note that (70) can be implied by either $\|\theta\|<\infty$ or a milder condition such as $\bar{x}_i(\theta)\neq 0$ for any θ,i.

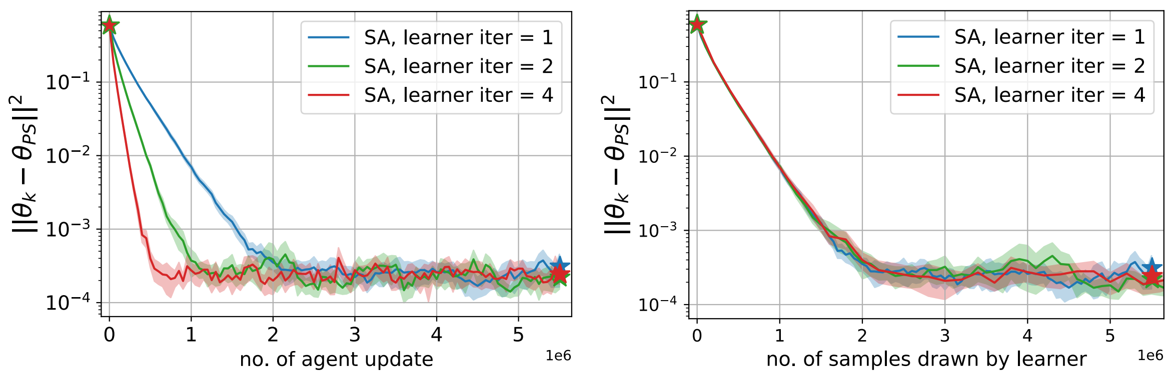

Figure 2: Convergence of SA algorithm with varying number of learner's updates per iteration.

than the agents. This is achieved by repeating steps 2 and 3 in Algorithm 2 for multiple times. Notice that this setting is similar to the lazy deploy scheme in [Mendler-Dünner et al., 2020]. From the figure, we observe that doing so improves the convergence from the agents' perspective, while the sample efficiency (from the learner's perspective) is unaffected.

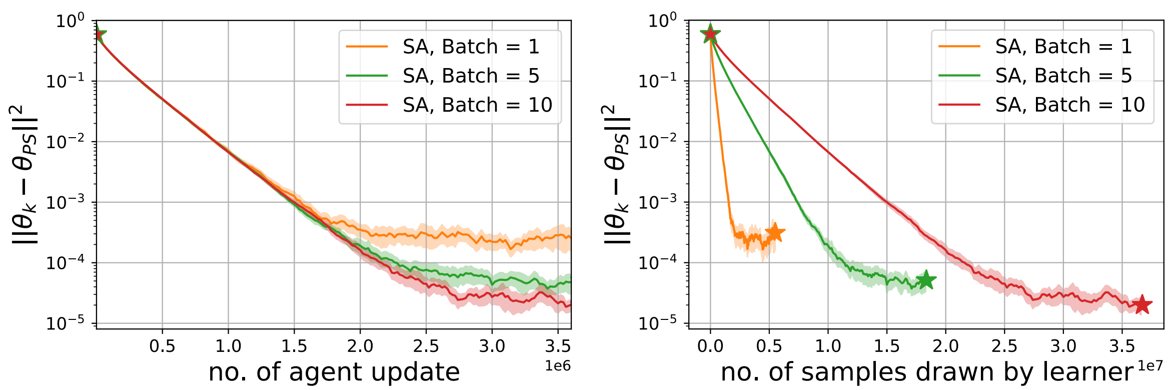

Figure 3: Convergence of SA algorithm with varying number of minibatch size.

Effects of Minibatch Size In this experiment, we consider the setting of $U_{\sf lg}(\cdot)$ and we draw different batch size of samples ($\hat{\sf b} \in \{1, 5, 10\}$) per iteration. To implement this, at step 2 of Algorithm 2, the learner draws $\hat{\sf b}$ agents uniformly as $\hat{\mathcal{I}}_k$, and at step 3, we update the iterate through:

$$\theta_{k+1} = \theta_k - \gamma_{k+1} \frac{1}{\hat{\sf b}} \sum_{j \in \hat{\mathcal{I}}_k} \nabla \ell(\theta_k; z_{k+1,j}).$$

In Fig. 3, we compare the error $\|\theta_k - \theta_{PS}\|^2$ in terms of the number of agents' best response update[5] and the number of samples drawn by the *learner*. We find that increasing the minibatch reduces the variance of the gradient estimate, yet it can be less sample efficient from the learner's perspective.

Effects of Number of Adaptive Best Responses In this experiment, we consider the setting of $U_{\sf lg}(\cdot)$ and at each iteration, the agents execute multiple rounds of adapted best response ($\mathsf{BR} \in \{1, 2, 4\}$) to simulate the scenario when the agents are allowed with more time to respond to the published classifier θ_k. To implement this, we repeat the update in (67) of step 1 in Algorithm 2 for BR times. Notice that this is reverse of Fig. 1 (right) where the learner performs multiple iterations per agents' best response update.

In Fig. 4, we compare the error $\|\theta_k - \theta_{PS}\|^2$ in terms of the number of agent update and the number of samples drawn by the *learner*. We observe that increasing the number of best responses improves the performance slightly. However, as a drawback, it requires more computations/updates at the agents to reach the same performance.

[5]Since the agents only perform one best response update per iteration, the x-axis is equivalent to the iteration number k.

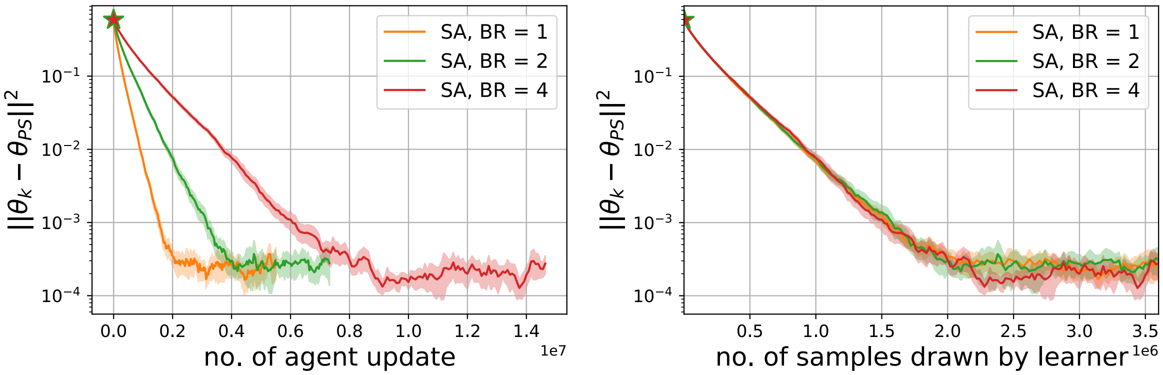

Figure 4: Convergence of SA algorithm with varying number of best responses.

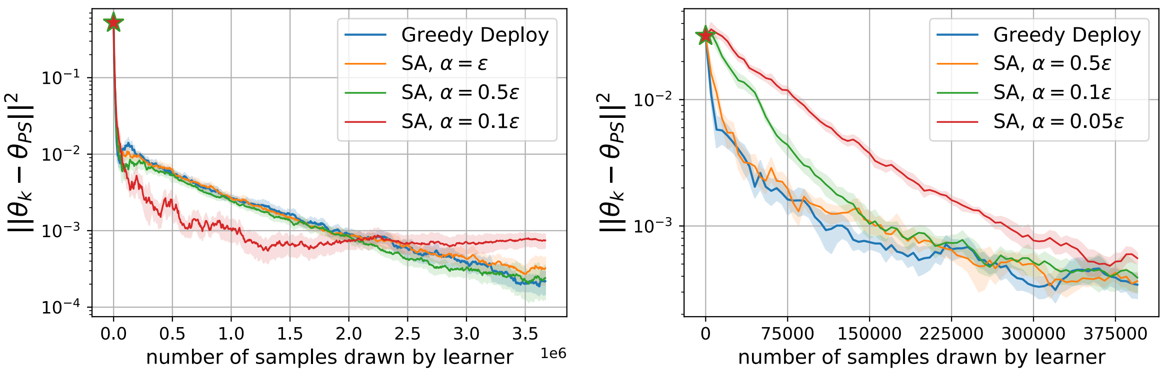

Figure 5: Convergence of SA algorithm with different values of agents' response rate α. (Left) In Setting (A) with large ϵ on the `GiveMeSomeCredit` dataset. (Right) In Setting (B) with synthetic data.

Large ϵ and Synthetic Data Lastly, we consider two additional settings with $U_{\mathsf{q}}(\cdot)$ on the SC problem as we further investigate the effects of agent response rate α on the convergence of state dependent SA:

(A) The first setting has sensitivity parameter ϵ that exceeds the theoretical bound μ/L. Specifically, we set $\epsilon = 100$ and evaluate the effect of the agents' response rate (α) on the convergence rate of θ_k to θ_{PS}. This is similar to the experiment considered in Mendler-Dünner et al. [2020]. Notice that as $\epsilon \gg \mu/L$, the SA algorithm and the greedy deployment scheme in Mendler-Dünner et al. [2020] are not guaranteed to converge, yet a larger ϵ shall amplify the effects of the data distribution shift. Note we have considered a different set of step sizes for (3) as $\gamma_k = c_0/(c_1 + k)$ with $c_0 = L, c_1 = L^2/\mu$, and the inexact best response dynamics (7) takes $|\mathcal{I}_k| = 100$.

(B) The second setting considers the same logistic regression problem (28), but with *synthetic* data. Here, we first generate a ground truth classifier $\theta_{true} \in \mathcal{U}[-1, 1]^{11}$, and generate $m = 5000$ samples of $\{d_i\}_{i=1}^m$ with $d_i = (x_i, y_i)$, $x_i \sim \mathcal{U}[-1, 1]^{11}$, and $y_i = \text{sign}(\langle x_i \,|\, \theta_{true} \rangle)$. We also set $\beta = \frac{1000}{m}$ in (28) and $\epsilon = 0.05$ in (68), where the latter satisfies $\epsilon < \frac{\mu}{L}$ for the problem instance considered. Furthermore, we consider the step sizes for (3) as $\gamma_k = c_0/(c_1 + k)$ with $c_0 = \frac{3}{\mu - L\epsilon}, c_1 = \frac{80L^2}{(\mu - L\epsilon)^2}$, and the inexact best response dynamics (7) takes $|\mathcal{I}_k| = 1$.

In Fig. 5, we compare the error $\|\theta_k - \theta_{PS}\|^2$ against the number of samples drawn by the *learner* as we examine different agents' response rate: with $\alpha \in \{0.1\epsilon, 0.5\epsilon, \epsilon\}$ in setting (A), or with $\alpha \in \{0.05\epsilon, 0.1\epsilon, 0.5\epsilon\}$ in setting (B). Note that Fig. 5 (Left) is similar to the comparison made in Fig. 1 (Middle), yet the increased ϵ leads to more significant data distribution shift. In both settings (A) and (B), we find that when the agents' response rate α decreases, the state dependent SA algorithm slows down. Moreover, as α increases towards ϵ, the convergence rate of state dependent SA algorithm approaches that of the greedy deployment scheme in [Mendler-Dünner et al., 2020]. This corroborates with the conclusion from Theorem 1 as decreasing α leads to a slower Markov chain, therefore increasing the constants such as $\widehat{L}$ in the theorem.