SPADE: STOCHASTIC PAYOFF-BASED ALGORITHM FOR DECENTRALIZED EQUILIBRIA IN DECISION-DEPENDENT GAMES WITH COUPLED CONSTRAINTS

Anonymous authors
Paper under double-blind review

ABSTRACT

In learning problems within politics, finance, or operations, agents' decisions and predictions can influence the underlying data-generating process—a phenomenon known as the *performative effect* Perdomo et al. (2020). We study this effect in decentralized noncooperative games with coupled constraints, where data distributions shift endogenously from players' actions. Furthermore, in practical settings, agents may not know the functional form of their objectives and can only observe realized payoffs from played actions. Our main contribution is SPADE (**S**tochastic **P**ayoff-based **A**lgorithm for **D**ecentralized **E**quilibria), a low-communication cost distributed algorithm in which each player randomly perturbs its action and updates with the resulting payoff observations. We show that when data are sampled from players' unperturbed actions, SPADE converges to the performatively-stable equilibrium (PSE), and when both data and payoffs are induced by the deployed actions, SPADE-NE converges to the Nash equilibrium (NE) under additional smoothness. Finally, we propose SPADE-MP, a multipoint-feedback approach that retains near full-information guarantees and improves empirical convergence to NE. We support our results through numerical experiments on networked Cournot competition and show significant improvements in reducing communication overhead.

1 INTRODUCTION

Machine learning models are typically trained to generalize from observed data to new environments. However, *distribution shifts* in real-world environments, particularly *endogenous shifts*, where deployed models themselves influence the data they observe Dong et al. (2018), pose a major challenge, degrading predictive accuracy and decision quality Zhou (2022). For example, a pricing algorithm may raise prices in response to increased demand, which in turn can decrease observed demand and sales data. While exogenous shifts have been widely studied Chan et al. (2020); Wu et al. (2021); Lu et al. (2018), endogenous shifts have received comparatively less attention.

Endogenous distribution shifts can be modeled through the recent framework of *performative prediction* (PP) Perdomo et al. (2020) or *decision-dependent learning* Drusvyatskiy & Xiao (2023), where a deployed decision parameter $\boldsymbol{\theta}$ can influence the underlying data distribution through a decision-dependent mapping $\mathcal{D}(\boldsymbol{\theta})$. Following the seminal work of Perdomo et al. (2020), recent research has studied stability, optimality, and learning dynamics in performative systems Piliouras & Yu (2023); Miller et al. (2021), with applications to reinforcement learning Mandal et al. (2023), online learning Wood et al. (2021); Jagadeesan et al. (2022), and bilevel optimization Lu (2023).

Much of this work, however, focuses on centralized decision-making. Learning is often decentralized and involves multiple strategic agents whose actions jointly shape future outcomes. Examples include trading under liquidity and capital constraints, where market participants' actions affect prices and liquidity conditions Fattouh & Mahadeva (2014); competition in ride-sharing platforms, where matching decisions affect rider demand and driver supply Narang et al. (2023); and online advertising auctions, where bidding strategies can affect auction outcomes and the competitors faced Varian (2009). These applications naturally lead to the setting of this paper: a decentralized noncooperative game in which agents' decisions affect both system payoffs and data-generating distributions, where each player seeks to minimize a private cost function while also satisfying

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Work	Feedback	Info exchanged	Constraints	Target	Regret Rate
Narang et al. (2023)	payoff/stochastic gradient	all actions (public)	none	PSE/NE	— ($\star$)
Wang et al. (2023)	stochastic gradient	actions	none	PSE	— ($\star$)
Yan & Cao (2024)	stochastic gradient	actions, estimates, duals	coupled	PSE	$\mathcal{O}(T^{3/4})$
SPADE, SPADE-NE	**payoff only**	**duals only**	coupled	**PSE/NE**	$\boldsymbol{\mathcal{O}(T^{7/8})}$
SPADE-MP	**payoff(s) only**	**duals only**	coupled	**PSE/NE**	$\boldsymbol{\mathcal{O}(T^{3/4})}$

Table 1: Comparison with the closest multiplayer performative-learning settings. ($\star$) Narang et al. (2023) report equilibrium-error and sample-complexity guarantees. Wang et al. (2023) establish convergence in expected squared equilibrium error to the PSE.

shared system constraints. Two equilibrium notions, the canonical *Nash equilibrium* (NE) and the *performative-stable equilibrium* (PSE)–which instead holds fixed the data distribution induced by the equilibrium profile when evaluating unilateral deviations–characterize stable outcomes in this game.

Our work is closely related to the decentralized performative game framework introduced in Yan & Cao (2024). Yan & Cao propose a stochastic primal-dual algorithm that computes the PSE by exchanging actions, duals, and estimates of other players' actions across the communication graph. However, requiring actions of *all other agents* to compute gradients can invoke privacy concerns and lead to *severe communication overhead*, especially in large-scale networks Martin et al. (1997); Chow & Kwok (2002); Sancho et al. (2007). Moreover, this approach requires players to have *full information* of their cost functions to compute stochastic gradients, which may not be practical.

Our contributions. To address this limitation, we develop SPADE, a decentralized algorithm based on payoff-based learning frameworks Marden & Shamma (2012); Goto et al. (2012); Tatarenko & Kamgarpour (2018). Instead of requiring gradient information or explicit knowledge of other players' actions, each player constructs a *zeroth-order* gradient estimator using random perturbations and observed payoffs. This modification significantly reduces communication requirements in large networks, as players no longer need to pass action or model information through the network. Our main contributions are summarized as follows.

1. PSE seeking. In § 3.1, we establish SPADE's convergence to the PSE. Through a technical analysis handling payoff-based estimation error, performative drift from changing distributions, and consensus errors in dual variables to enforce shared constraints, we prove sublinear bounds on SPADE's *performative regret* and cumulative constraint violation.

2. NE seeking. In § 3.2, we introduce SPADE-NE for more challenging NE seeking, providing the first convergence guarantee to NE in performative games with coupled constraints under additional smoothness.

3. Matching full information. In § 3.3, we propose SPADE-MP, a multi-point improvement of SPADE-NE for when learners can query payoffs at two nearby deployments. Under a paired-feedback assumption, SPADE-MP converges to the NE with $\mathcal{O}(T^{3/4})$ regret, highlighting that appropriately designed two-point feedback can recover full information rates in Yan & Cao (2024).

4. Experimental validation. In § 4, we showcase the empirical effectiveness of SPADE, SPADE-NE, and SPADE-MP on networked Cournot (quantity choice) competition, demonstrating its effectiveness on PSE/NE convergence and constraint violation, along with a $6\times$ per-iteration communication reduction over the gradient-based baseline of Yan & Cao (2024).

Related work. Performativity and decision-dependent learning have been studied across several settings, including corrupted/open environments Shan et al. (2023), bilevel optimization Lu (2023), reinforcement learning Mandal et al. (2023), stochastic optimization Drusvyatskiy & Xiao (2023), and constrained performative prediction Yan & Cao (2023). Beyond stability, several works study optimality and the gap between performatively stable and performatively optimal decisions Miller et al. (2021); Izzo et al. (2021); Jagadeesan et al. (2022). At the multi-agent level, Narang et al. (2023) and Wang et al. (2023) study performative effects in game-theoretic environments, while Li et al. (2022); Piliouras & Yu (2023) consider decentralized or multi-agent performative learning. These works differ from ours in that they either focus on unconstrained games, cooperative settings,

centralized formulations, or restricted distribution families. In contrast, we study decentralized noncooperative games with coupled decision-dependent distributions and shared constraints. A comparison between the works closest to ours is provided in Table 1.

Our contributions also build on the literature in payoff-based learning Bravo et al. (2018); Tatarenko & Kamgarpour (2018); Drusvyatskiy et al. (2022), allowing agents to reduce communication by relying on realized payoff observations. Overall, our work bridges decentralized game-theoretic learning, performative prediction, and payoff-based learning for constrained multi-agent systems with endogenous distribution shifts. Related works are discussed in further detail in Appendix A.

2 PRELIMINARIES

Consider a game with n players, where each player $i \in [n]$ selects a strategy or action $\boldsymbol{\theta}_i \in \boldsymbol{\Omega}_i \subset \mathbb{R}^d$. Let $\boldsymbol{\theta}$ denote the vector of all players' actions in space $\boldsymbol{\Omega} = \times_i \boldsymbol{\Omega}_i$, and let $\boldsymbol{\theta}_{-i}$ be that for all players but i. Let $J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i})$ denote player i's cost function, where $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$ is a random variable drawn from a latent *decision-dependent* distribution mapping, $\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i})$. Thus, the payoffs of the players crucially depend on data whose distributions can depend on strategic interactions between players. Each player seeks to minimize their performative risk, $\mathrm{PR}_i(\boldsymbol{\theta}) := \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta})}[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta})]$, subject to a *coupled constraint*, solving

$$\min_{\boldsymbol{\theta}_i \in \boldsymbol{\Omega}_i} \quad \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i})} \left[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}) \right] \quad \text{s.t.} \quad \boldsymbol{g}_i(\boldsymbol{\theta}_i) + \sum_{j \neq i} \boldsymbol{g}_j(\boldsymbol{\theta}_j) \leq \mathbf{0}, \tag{1}$$

where $\boldsymbol{g}_i : \boldsymbol{\Omega}_i \to \mathbb{R}^m$. In this work, we assume players **do not know the functional form of** J_i. We aim to develop *payoff-based* (or "gradient-free") algorithms that learn stable equilibria of the game, where no player has an incentive to unilaterally deviate from their strategy. In particular, we define the generalized Nash equilibrium and performative-stable equilibrium below:

Definition 1 (Nash Equilibrium (NE)) *A joint strategy $\boldsymbol{\theta}^{\mathrm{NE}}$ is a NE if for all $i \in [n]$,*

$$\boldsymbol{\theta}_i^{\mathrm{NE}} \in \arg\min_{\boldsymbol{\theta}_i \in \boldsymbol{\Omega}_i} \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{\mathrm{NE}})} \left[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{\mathrm{NE}}) \right] \quad \textit{s.t.} \quad \boldsymbol{g}_i(\boldsymbol{\theta}_i) + \sum_{j \neq i} \boldsymbol{g}_j(\boldsymbol{\theta}_j^{\mathrm{NE}}) \leq \mathbf{0}.$$

Definition 2 (Performative-Stable Equilibrium (PSE)) *A joint strategy $\boldsymbol{\theta}^{\mathrm{PSE}}$ is a PSE if for all $i \in [n]$,*

$$\boldsymbol{\theta}_i^{\mathrm{PSE}} \in \arg\min_{\boldsymbol{\theta}_i \in \boldsymbol{\Omega}_i} \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta}^{\mathrm{PSE}})} \left[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{\mathrm{PSE}}) \right] \quad \textit{s.t.} \quad \boldsymbol{g}_i(\boldsymbol{\theta}_i) + \sum_{j \neq i} \boldsymbol{g}_j(\boldsymbol{\theta}_j^{\mathrm{PSE}}) \leq \mathbf{0}.$$

The key distinction between PSE and NE is whether a unilateral deviation *also changes the data distribution* $\mathcal{D}_i$. In particular, for any $\boldsymbol{\theta}, \boldsymbol{\delta} \in \boldsymbol{\Omega}$. define the *direct gradient mapping* for player i as

$$G_{\boldsymbol{\delta}}^{(i)}(\boldsymbol{\theta}) := \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\delta})} \left[\nabla_{\boldsymbol{\theta}_i} J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}) \right],$$

and define the *indirect distribution-shift mapping* for player i as

$$H_{\boldsymbol{\delta}}^{(i)}(\boldsymbol{\theta}) := \nabla_{\boldsymbol{u}_i} \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{u}_i, \boldsymbol{\delta}_{-i})} \left[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}) \right] \big|_{\boldsymbol{u}_i = \boldsymbol{\delta}_i}.$$

Since the coupled constraint also makes Game 1 a generalized NE problem, we target *variational* PSE and NE, characterized by a common multiplier, $\boldsymbol{\lambda}^{\mathrm{PSE}}$ or $\boldsymbol{\lambda}^{\mathrm{NE}}$, associated with the shared feasible set. Thus, using the above mappings, the PSE and NE satisfy the respective stationarity conditions

$$\begin{aligned} 0 &\in G_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top \boldsymbol{\lambda}^{\mathrm{PSE}} + N_{\boldsymbol{\Omega}_i}(\boldsymbol{\theta}_i^{\mathrm{PSE}}), \\ 0 &\in G_{\boldsymbol{\theta}^{\mathrm{NE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{NE}}) + H_{\boldsymbol{\theta}^{\mathrm{NE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{NE}}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top \boldsymbol{\lambda}^{\mathrm{NE}} + N_{\boldsymbol{\Omega}_i}(\boldsymbol{\theta}_i^{\mathrm{NE}}), \end{aligned} \tag{2}$$

for each $i \in [n]$, where $N_{\boldsymbol{\Omega}_i}$ denotes the normal cone to $\boldsymbol{\Omega}_i$.

The common multipliers $\boldsymbol{\lambda}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{NE}} \in \mathbb{R}_+^m$ also satisfy primal feasibility and complementary slackness, i.e. for $* \in \{\mathrm{PSE}, \mathrm{NE}\}$,

$$\sum_{i=1}^n \boldsymbol{g}_i(\boldsymbol{\theta}_i^*) \leq \mathbf{0}, \qquad \boldsymbol{\lambda}^* \odot \sum_{i=1}^n \boldsymbol{g}_i(\boldsymbol{\theta}_i^*) = \mathbf{0}.$$

Importantly, from these equations, we see that PSE's condition depends on the *direct* gradient $G^{(i)}$, whereas NE also includes the *indirect distribution-shift* term $H^{(i)}$. This is why PSE has traditionally been easier to compute, since gradient-based methods can estimate $G^{(i)}$ using $\nabla_{\boldsymbol{\theta}_i} J_i$ without explicit knowledge of $\mathcal{D}_i(\cdot)$. However, PSE may *still allow* profitable unilateral deviations once the induced distribution changes. NE rules out such deviations and is thus strategically more stable, but also harder to compute, motivating our algorithms (namely SPADE-NE and SPADE-MP).

The players are interconnected by a communication graph, denoted $\mathcal{G}(A)$, with doubly stochastic weight matrix $A = (a_{ij})_{n\times n}$. We assume $a_{ij} = a_{ji} > 0$ if players i and j communicate and $a_{ij} = 0$ otherwise. Let $\mathcal{N}_i$ denote player i's neighborhood, including i itself. Throughout, we assume $\mathcal{G}(A)$ is connected, and our algorithms aim to reduce the communication load over this network.

2.1 Standing assumptions

We first state here all assumptions used. Assumptions 1 to 6 review the necessary assumptions and results from Narang et al. (2023); Yan & Cao (2024) for the tractability of the problem, giving existence and uniqueness of PSE and NE in Game 1. Assumptions 7 and 8 are necessary for the algorithm guarantees.

Let $\boldsymbol{\xi} = (\boldsymbol{\xi}_i)_{i\in[n]}$ (living in space $\boldsymbol{\Xi} = \times_i \boldsymbol{\Xi}_i$) and $J(\boldsymbol{\xi};\boldsymbol{\theta}) = (J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta}))_{i\in[n]}$. We define a stochastic pseudogradient mapping of $J(\boldsymbol{\xi};\boldsymbol{\theta})$ as $\nabla J(\boldsymbol{\xi};\boldsymbol{\theta}) := \operatorname{col}\left(\nabla_{\boldsymbol{\theta}_1} J_1(\boldsymbol{\xi}_1;\boldsymbol{\theta}), \cdots, \nabla_{\boldsymbol{\theta}_n} J_n(\boldsymbol{\xi}_n;\boldsymbol{\theta})\right)$, and impose the standard game theory assumption:

Assumption 1 *There exists a constant $\mu > 0$ such that the stochastic gradient mapping $\nabla J(\boldsymbol{\xi};\boldsymbol{\theta})$ is μ-strongly monotone, i.e.,*

$$\langle \nabla J(\boldsymbol{\xi};\boldsymbol{\theta}) - \nabla J(\boldsymbol{\xi};\boldsymbol{\theta}'), \boldsymbol{\theta} - \boldsymbol{\theta}' \rangle \geq \mu \|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_2^2, \qquad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi},\ \boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Omega}.$$

Let $\mathcal{G}(\mathcal{P})$ be a graph where the weight $p_{ij} \geq 0$ represents the magnitude of how much player j's decision affects player i's data distribution. These weights are normalized as $\sum_{j=1}^n p_{ij} = 1$, $\forall i \in [n]$. Let $W_1(\mathcal{D}, \mathcal{D}')$ represent the *Wasserstein-1* distance between two probability measures $\mathcal{D}$ and $\mathcal{D}'$. Following Wang et al. (2023), we impose the following assumption on the distributions $\{\mathcal{D}_i\}_{i\in[n]}$, bounding the sensitivity of player i's distribution with respect to the players' decision differences.

Assumption 2 *For any $i \in [n]$, there exists a constant $\varepsilon_i \geq 0$ such that, for all $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \mathbb{R}^{nd}$, the distribution mapping $\mathcal{D}_i$ is constrained by $\mathcal{W}_1\big(\mathcal{D}_i(\boldsymbol{\theta}), \mathcal{D}_i(\boldsymbol{\theta}')\big) \leq \varepsilon_i \sqrt{\sum_{j=1}^n p_{ij} \|\boldsymbol{\theta}_j - \boldsymbol{\theta}'_j\|_2^2}$.*

We also require the following assumptions, standard across constrained optimization and performative prediction literature Bertsekas (2014); Yan & Cao (2024).

Assumption 3 *For each $i \in [n]$, the action set $\boldsymbol{\Omega}_i \subset \mathbb{R}^d$ is compact, convex, and has nonempty interior. Let $C < \infty$ be such that $\|\boldsymbol{\theta}_i\|_2 \leq C$ for all $\boldsymbol{\theta}_i \in \boldsymbol{\Omega}_i$.*

Assumption 4 *For any $i \in [n]$ and $\boldsymbol{\theta}_i \in \boldsymbol{\Omega}_i$, the cost function $J_i(\xi_i; \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i})$ is convex w.r.t. $\boldsymbol{\theta}_i$. Moreover, there exists a constant $L_i > 0$ such that $\nabla_{\boldsymbol{\theta}_i} J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta})$ is L_i-smooth, i.e.,*

$$\|\nabla_{\boldsymbol{\theta}_i} J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}_i} J_i(\boldsymbol{\xi}'_i;\boldsymbol{\theta}')\|_2 \leq L_i \left(\|\boldsymbol{\xi}_i - \boldsymbol{\xi}'_i\|_2 + \|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_2\right), \quad \boldsymbol{\xi}_i, \boldsymbol{\xi}'_i \in \boldsymbol{\Xi}_i, \boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Omega}.$$

Assumption 5 *$J_i(\boldsymbol{\xi}_i, \boldsymbol{\theta})$ is Lipschitz, i.e. for all i, there exists $G_i \geq 0$ such that $|J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta}) - J_i(\boldsymbol{\xi}'_i;\boldsymbol{\theta}')| \leq G_i \left(\|\boldsymbol{\xi}_i - \boldsymbol{\xi}'_i\|_2 + \|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_2\right), \forall\, \xi_i, \xi'_i \in \boldsymbol{\Xi}_i, \boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Omega}$.*

Assumption 6 *For any $i \in [n]$ and $\boldsymbol{\theta}_i \in \mathbb{R}^d$, the constraint function $g_i(\boldsymbol{\theta}_i)$ is differentiable and convex with respect to $\boldsymbol{\theta}_i$. Moreover, there exists a constant $G_g \geq 0$ such that g_i is G_g-Lipschitz, i.e.,*

$$\|g_i(\boldsymbol{\theta}_i) - g_i(\boldsymbol{\theta}'_i)\|_2 \leq G_g \|\boldsymbol{\theta}_i - \boldsymbol{\theta}'_i\|_2, \quad \forall\, \boldsymbol{\theta}_i, \boldsymbol{\theta}'_i \in \boldsymbol{\Omega}_i.$$

With these assumptions, we have the following additional property and equilibrium guarantees:

Lemma 1 (Yan & Cao (2024)) *The constraint function $g_i(\cdot)$ is bounded for any i, i.e. there exists a constant $B \geq 0$ such that for any $i \in [n], \boldsymbol{\theta}_i \in \boldsymbol{\Omega}_i$, $\|g_i(\boldsymbol{\theta}_i)\|_2 \leq B$.*

Proposition 1 (Theorems 3.3–3.4 of Yan & Cao (2024)) *Suppose Assumptions 1–6 hold.*

If $\widetilde{\mu} := \mu - \sum_{i=1}^{n} L_i \varepsilon_i \max_{j\in[n]} \sqrt{p_{ij}} > 0$, *then* (1) *admits a unique (variational) PSE.*

If $\widehat{\mu} := \widetilde{\mu} - \sqrt{\sum_{i=1}^{n} L_i^2 \varepsilon_i^2 \, p_{ii}} > 0$ *and the joint indirect distribution-shift mapping* $H_{\boldsymbol{\delta}}(\boldsymbol{\theta})$ *is monotone w.r.t.* $\boldsymbol{\delta}$, *then* $\nabla \mathrm{PR}$ *is* $\widehat{\mu}$*-strongly monotone and* (1) *admits a unique (variational) NE.*

We now state the remaining assumptions needed for the algorithm analysis, which are adaptations of those used in stochastic optimization and approximation for performative learning.

Assumption 7 (Finite payoff second moment) *For each* $i \in [n]$, *there exist a reference action profile* $\bar{\boldsymbol{\theta}} \in \boldsymbol{\Omega}$ *and a constant* $V_i < \infty$ *such that* $\sup_{\boldsymbol{\delta}\in\boldsymbol{\Omega}} \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\delta})} \left[J_i(\boldsymbol{\xi}_i; \bar{\boldsymbol{\theta}})^2 \right] \le V_i$.

Assumption 8 (Slater's condition and known dual set) *There exists a strictly feasible* $\widehat{\boldsymbol{\theta}} \in \boldsymbol{\Omega}$ *with* $\sum_{i=1}^{n} \boldsymbol{g}_i(\widehat{\boldsymbol{\theta}}_i) \prec \mathbf{0}$. *Then, for each* $* \in \{\mathrm{PSE}, \mathrm{NE}\}$, *players know* $\boldsymbol{\Lambda}_{\max}$ *satisfying* $\|\boldsymbol{\lambda}^*\|_2 \le \boldsymbol{\Lambda}_{\max}$. *We define the dual set* $\boldsymbol{\Lambda} := \{\boldsymbol{\lambda} \in \mathbb{R}^m_+ : \|\boldsymbol{\lambda}\|_2 \le \Lambda_{\max}\}$, *which is convex and compact.*

Assumption 7 bounds the second moment of the payoff under the performative distribution which, together with Lipschitz continuity, provides the moment control needed for the payoff-based estimation errors in our convergence analysis. Assumption 8 ensures the existence of finite equilibrium multipliers through Slater's condition and provides a common bound $\Lambda_{\max}$ and compact dual set containing these multipliers which keeps the decentralized dual iterates bounded.

3 Main algorithms and results

This section presents our three theoretical results. Section 3.1 introduces SPADE, a payoff-based algorithm for the canonical PSE solution. Section 3.2 introduces SPADE-NE, which targets the computationally-harder NE under deployed feedback and a Lipschitz-gradient assumption on the performative risk. Section 3.3 introduces SPADE-MP, a multipoint variant that uses paired payoff queries and exhibits improved convergence rates and empirical efficiency in computing equilibria.

As emphasized, all three algorithms rely on players estimating their own gradients using payoff observations from *randomly perturbed actions*. This idea and subsequent estimate construction is based on a payoff-based gradient identity (Lemma 2).

Let $\mathbb{S}^{d-1} := \{\boldsymbol{z} \in \mathbb{R}^d : \|\boldsymbol{z}\|_2 = 1\}$ and $\mathbb{B}^d := \{\boldsymbol{v} \in \mathbb{R}^d : \|\boldsymbol{v}\|_2 \le 1\}$. For any function $f : \mathbb{R}^{nd} \to \mathbb{R}$, smoothing parameter $\sigma > 0$, and center $\boldsymbol{\theta}$, define the player-i smoothed function

$$\widetilde{f}_i(\boldsymbol{\theta}; \sigma) := \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \left[f(\boldsymbol{\theta}_i + \sigma \boldsymbol{v}_i, (\boldsymbol{\theta}_j + \sigma \boldsymbol{z}_j)_{j\neq i}) \right], \tag{3}$$

where $\boldsymbol{v}_i, \boldsymbol{z}_j$ are drawn uniformly and all random variables mutually independent.

Lemma 2 (Smoothing identity) *Let* $\boldsymbol{z}_i \sim \mathrm{Unif}(\mathbb{S}^{d-1})$ *independently for each* $i \in [n]$, *and write* $\boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n)$. *Then, for any* $\boldsymbol{\xi}_i \in \boldsymbol{\Xi}_i$,

$$\nabla_{\boldsymbol{\theta}_i} \widetilde{J}_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}, \sigma) = \frac{d}{\sigma} \mathbb{E}_{\boldsymbol{z}} \left[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta} + \sigma \boldsymbol{z}) \boldsymbol{z}_i \right]. \tag{4}$$

Lemma 2 then lets us construct the following payoff-based estimator for each $i \in [n]$ at iteration t,

$$\widehat{\nabla}_i J_i^t(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}^t) := \frac{d}{\sigma_t} J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t) \boldsymbol{z}_i^t, \tag{5}$$

where $\boldsymbol{\psi}_t$ is some joint action *perturbed* from $\boldsymbol{\theta}_t$ and $\{\sigma_t\}$ is a sequence of chosen smoothing parameters. Importantly, conditional on the current iterate, Lemma 2 shows that Estimator 5 can target the gradient of the (smoothed) objective evaluated at iterate $\boldsymbol{\theta}_t$.

However, note that perturbed actions may not lie in $\boldsymbol{\Omega}_i$. To resolve this, as admitted by Assumption 3, let $\boldsymbol{p}_i$ and R_i be the center and radius of the largest inscribed ball in $\boldsymbol{\Omega}_i$. For $0 < \sigma_t \le R_i$, define

$$\boldsymbol{\theta}_i^{t,\sigma} := \left(1 - \frac{\sigma_t}{R_i}\right) \boldsymbol{\theta}_i^t + \frac{\sigma_t}{R_i} \boldsymbol{p}_i \tag{6}$$

to be the *inward-shifted center* and thus sample $\boldsymbol{z}_i^t \sim \mathrm{Unif}(\mathbb{S}^{d-1})$ and $\boldsymbol{\psi}_i^t := \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{z}_i^t$. From convexity of $\boldsymbol{\Omega}_i$, we get $\boldsymbol{\psi}_i^t \in \boldsymbol{\Omega}_i$. Furthermore, since $\|\boldsymbol{\theta}^{t,\sigma} - \boldsymbol{\theta}^t\|_2 = O(\sigma_t)$, the difference between this smoothed gradient and the desired gradient at $\boldsymbol{\theta}^t$ is controlled by the smoothness assumptions.

Finally, we make the crucial remark that depending on how $\boldsymbol{\xi}_i^t$ is sampled, Estimator (5) can distinguish between PSE and NE targets. Specifically,

- With *intended-action sampling* $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t)$, the distribution is frozen at the *unperturbed* iterate. Lemma 2 then gives $\mathbb{E}[\widehat{\nabla}_i J_i^t \mid \boldsymbol{\theta}^t] = \widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma})$, the (smoothed) decoupled-risk gradient required by Definition 2 for PSE. This drives SPADE (Section 3.1).
- With *deployed-action sampling* $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\psi}^t)$, the payoff $J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)$ is an unbiased sample of $\mathrm{PR}_i(\boldsymbol{\psi}^t)$, and the same identity gives $\mathbb{E}[\widehat{\nabla}_i J_i \mid \boldsymbol{\theta}^t] = \nabla_{\boldsymbol{\theta}_i} \widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma})$, the (smoothed) performative-risk gradient aligned with Definition 1 for NE. This drives SPADE-NE (Section 3.2) (and SPADE-MP, Section 3.3).

3.1 SPADE: STOCHASTIC PAYOFF-BASED ALGORITHM FOR DECENTRALIZED EQUILIBRIA

We now present our first algorithm, SPADE, which targets the PSE. At each round t, every player i exchanges only the dual variable $\boldsymbol{\lambda}_i^t \in \mathbb{R}_+^m$ with its neighbors and performs the following updates:

$$\text{(sample/deploy)} \quad \boldsymbol{z}_i^t \sim \mathrm{Unif}(\mathbb{S}^{d-1}),\ \boldsymbol{\psi}_i^t = \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{z}_i^t,\ \boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t),\ \text{observe } J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t). \tag{7a}$$

$$\text{(primal/gradient step)} \quad \boldsymbol{\theta}_i^{t+1} = P_{\boldsymbol{\Omega}_i}\Big[\boldsymbol{\theta}_i^t - \gamma_t(\sigma_t)^2 \Big(\widehat{\nabla}_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t\Big)\Big], \tag{7b}$$

$$\text{(dual)} \quad \boldsymbol{\lambda}_i^{t+1} = P_{\boldsymbol{\Lambda}}\Big[\sum_{j \in \mathcal{N}_i} a_{ij} \boldsymbol{\lambda}_j^t + \gamma_t\, \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\Big]. \tag{7c}$$

We label this procedure Algorithm 1, with full pseudocode in Appendix B. SPADE follows a "gradient" update similar to Yan & Cao (2024) to solve the constrained optimization. However, unlike Yan & Cao (2024), the stochastic gradient $\nabla_{\boldsymbol{\theta}_i} J_i\big(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^t, \widehat{\boldsymbol{\theta}}_i^t\big)$–which needed costly communication of actions, estimates of others' actions ($\widehat{\boldsymbol{\theta}}_i^t$), and knowledge of J_i–is replaced by a payoff-based estimator. Moreover, we tune the step size $\gamma_t \sigma_t^2$ to help control the higher variance with payoff-based methods.

As a subtle remark, SPADE samples $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t)$ from the intended-action distribution, so Estimator 5 targets the PSE. This requires either a *simulator* that returns $J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)$ with $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t)$, or a *sticky* environment that responds to persistent or announced actions but not to transitory perturbations, as is common in economic models Bils & Klenow (2004); Rotemberg (1982); Klenow & Willis (2007). We study SPADE first because PSE is the standard performative-learning solution concept Yan & Cao (2024); Narang et al. (2023) and cleanly isolates the payoff-based machinery. Section 3.2 then considers the more strategically meaningful NE via SPADE-NE, which requires only standard deployed-payoff feedback from $\mathcal{D}(\boldsymbol{\psi}^t)$. We now present our main result.

Definition 3 (Performative Regret and Constraint Violation) *For each player $i \in [n]$, define the PSE-performative regret and expected constraint violation as*

$$\mathcal{R}_i^{\mathrm{PSE}}(T) := \sum_{t=1}^{T} \mathbb{E}\big[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}_i^t, \boldsymbol{\theta}_{-i}^{\mathrm{PSE}}) - J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}^{\mathrm{PSE}})\big], \qquad \mathcal{R}_g(T) := \mathbb{E}\left\| \left[\sum_{t=1}^{T} \sum_{i=1}^{n} \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right]_+ \right\|_2,$$

where $\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta}^{\mathrm{PSE}})$. Sublinear $\mathcal{R}_i^{\mathrm{PSE}}(T)$ and $\mathcal{R}_g(T)$ are consistent with convergence to $\boldsymbol{\theta}^{\mathrm{PSE}}$.

Theorem 1 (SPADE on PSE) *Suppose Assumptions 1–8 hold. Assume that $\widetilde{\mu} > 0$ and the stepsizes are nonincreasing, positive, and satisfy $\gamma_t, \sigma_t = \Theta(t^{-1/4})$, $\sigma_t \le \min\{1, R_{\min}\}$ and $\sup_{t \ge 1} \gamma_t \sigma_t^2 \le \frac{\widetilde{\mu}}{\nu}$, where $R_{\min} = \min_i R_i$ and $\nu := 6 \sum_{i=1}^n L_i^2 \left(1 + \varepsilon_i^2 \max_{j \in [n]} p_{ij}\right)$. Then, for every $i \in [n]$,*

$$\mathcal{R}_i^{\mathrm{PSE}}(T) \le \mathcal{O}(T^{7/8}) \quad \text{and} \quad \mathcal{R}_g(T) \le \mathcal{O}(T^{7/8}).$$

Indeed, the rate is slower than the $\mathcal{O}(T^{3/4})$ rate of Yan & Cao (2024) because the payoff-based estimator introduces the additional smoothing-parameter dependence $1/\sigma_T^2$ and the bias term $\sum_t \sigma_t$,

which constrain γ_t and σ_t to decay more slowly. Even with this asymptotic tradeoff for communication efficiency and gradient-free computation, SPADE is still sublinear in T and converges to PSE.

Proof idea. We give a high-level proof sketch of Theorem 1 here, deferring full details to Appendix D. Starting from (7b) and projection nonexpansiveness, we obtain the one-step recursion

$$\mathbb{E}\|\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 \le \mathbb{E}\|\boldsymbol{\theta}^{t} - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 + (\mathrm{I}) - (\mathrm{II}),$$

where $(\mathrm{I}) \sim \gamma_t^2\sigma_t^4\mathbb{E}\|\widehat{\phi} - \Phi^{\mathrm{PSE}}\|_2^2$ captures update errors and $(\mathrm{II}) \sim \gamma_t\sigma_t^2\mathbb{E}\langle\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}, \widehat{\phi} - \Phi^{\mathrm{PSE}}\rangle$ is the inner-product term producing descent ($\widehat{\phi}$ and Φ are update directions defined in the proof).

The core technical step is controlling the payoff-based estimator in (I). By the smoothing identity, its conditional mean is the gradient of a smoothed cost evaluated at the shifted center $\boldsymbol{\theta}^{t,\sigma}$. We therefore separate the error into a centered stochastic term (with second moment $\mathcal{O}(1/\sigma_t^2)$) and a smoothing/inward-shift bias (which is $\mathcal{O}(\sigma_t)$ by smoothness). The remaining terms are controlled using Assumptions 2-6 and the dual-consensus lemmas in Appendix C.

Term (II) contributes the strong-monotonicity descent $-2\widetilde{\mu}\gamma_t\sigma_t^2\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2$, where the Wasserstein-sensitivity error is absorbed into $\widetilde{\mu}$, together with an $\mathcal{O}(\gamma_t\sigma_t^3)$ smoothing-bias term and the dual tracking terms. Combining gives

$$\mathbb{E}\|\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 \le \left(1 - 2\widetilde{\mu}\gamma_t\sigma_t^2 + \nu\gamma_t^2\sigma_t^4\right)\mathbb{E}\|\boldsymbol{\theta}^{t} - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 + \pi\gamma_t^2\sigma_t^2 + \rho\gamma_t\sigma_t^3 + \text{dual terms}.$$

Under $\sup_t \gamma_t\sigma_t^2 \le \widetilde{\mu}/\nu$, telescoping the contraction and applying the dual-consensus bounds control the cumulative squared distance to the PSE. The regret and constraint-violation bounds then follow from Cauchy–Schwarz, Lipschitz continuity, and feasibility of $\boldsymbol{\theta}^{\mathrm{PSE}}$. □

3.2 SPADE-NE: A Payoff-Based Algorithm for Nash Equilibrium

We now turn to NE computation. Recall that the NE is the strategically meaningful solution concept, ruling out unilateral profitable deviations under each player's *own* induced distribution. NE has faced challenges in previous derivative-based performative learning literature because the optimality condition involves both the direct gradient $G^{(i)}$ and the indirect distribution-shift gradient $H^{(i)}$ of (2).

However, with payoff-based feedback using data drawn from the distribution *actually induced* by deployment (i.e. $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\psi}^t)$), the Estimator (5) targets $\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma})$ in expectation, sidestepping the explicit $G^{(i)}, H^{(i)}$ split — modulo the smoothing bias $\|\nabla\widetilde{\mathrm{PR}}_i - \nabla\mathrm{PR}_i\|_2$. To control this bias, we impose a smoothness condition on the performative-risk gradient itself Drusvyatskiy et al. (2022).

Assumption 9 (Lipschitz performative-risk gradient) *There exists a constant $L_{\mathrm{PR}} > 0$ such that, for all $i \in [n]$, $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Omega}$, $\|\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}')\|_2 \le L_{\mathrm{PR}}\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_2$.*

Assumption 9 is precisely what is necessary to translate the payoff-based smoothing identity into a controlled NE analysis by yielding a uniform, vanishing $\mathcal{O}(\sigma_t)$ bound on the smoothing bias. We emphasize that Assumptions 2 and 4 do *not* yield this property: the ε-sensitivity bound is a first-order regularity statement on the distribution map $\boldsymbol{\theta} \mapsto \mathcal{D}_i(\boldsymbol{\theta})$, whereas Lipschitzness of $\nabla\mathrm{PR}_i$ is a second-order property of the composite map $\boldsymbol{\theta} \mapsto \mathbb{E}_{\mathcal{D}_i(\boldsymbol{\theta})}[J_i(\cdot;\boldsymbol{\theta})]$. While this assumption is stronger than those used for PSE, it is standard in zeroth-order methods and holds in common performative models.

Example 1 (Linear performative effects) *Suppose $\boldsymbol{\xi}_i = \boldsymbol{A}_i\boldsymbol{\theta} + \boldsymbol{\eta}_i$ with $\boldsymbol{\eta}_i$ a $\boldsymbol{\theta}$-independent noise of finite second moment and $\|\boldsymbol{A}_i\|_2 \le L_A$ (e.g. our networked Cournot experimental setup). With L_i-smoothness (Assumption 4), this gives Assumption 9 with $L_{\mathrm{PR}} \le \max_i L_i(1 + L_A)^2$.*

Now we introduce SPADE-NE, which has the same per-round updates (7a)-(7c) as SPADE. The defining difference is the feedback model, where the sample $\boldsymbol{\xi}_i^t$ in (7a) is now drawn from $\mathcal{D}_i(\boldsymbol{\psi}^t)$, the distribution induced by the *deployed* action. We label the resulting procedure Algorithm 2, and the explicit per-step pseudocode is given in Appendix B.

The full theorem statement for SPADE-NE resembles that of Theorem 1, so we summarize it here and defer the details to Appendix E. As in Theorem 1, we measure convergence to the NE outcome via an **NE-performative regret**, defined under each iterate's *own* induced distribution:

$$\mathcal{R}_i^{\mathrm{NE}}(T) := \sum_{t=1}^{T}\mathbb{E}\left[\mathrm{PR}_i(\boldsymbol{\theta}_i^t, \boldsymbol{\theta}_{-i}^{\mathrm{NE}}) - \mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}})\right].$$

Theorem 2 (Convergence of SPADE-NE, informal) *Assume Assumptions 1–9 and $\widehat{\mu} > 0$. Letting stepsizes $\gamma_t, \sigma_t = \Theta(t^{-1/4})$ yields $\boldsymbol{\mathcal{R}_i^{\mathrm{NE}}(T)}$, $\boldsymbol{\mathcal{R}_g(T)} \leq \boldsymbol{\mathcal{O}(T^{7/8})}$, so SPADE-NE learns the NE.*

3.3 SPADE-MP: A MULTIPOINT-FEEDBACK APPROACH

We now ask if we can do better. In some settings, players can query their cost more than once before committing to an update. *Two-point* bandit methods cleverly exploit this additional data by deploying two *antithetic* perturbations near the current iterate. It turns out that using this feedback can admit even close-to full information rates for payoff-based methods Agarwal et al. (2010); Hikima (2023).

We use this approach for the performative game. At iteration t, sampling $\boldsymbol{z}^t$ and computing the inward-shifted center as before, the players synchronously realize the two joint perturbed actions

$$\boldsymbol{\psi}^{t,+} := \boldsymbol{\theta}^{t,\sigma} + \sigma_t \boldsymbol{z}^t, \qquad \boldsymbol{\psi}^{t,-} := \boldsymbol{\theta}^{t,\sigma} - \sigma_t \boldsymbol{z}^t, \tag{8}$$

while holding $\boldsymbol{\theta}^t$ and the dual variables fixed until both payoff observations have been collected. To target the NE[1], each payoff is generated under the distribution induced by the corresponding deployed action $\boldsymbol{\xi}_i^{t,+} \sim \mathcal{D}_i(\boldsymbol{\psi}^{t,+})$, $\boldsymbol{\xi}_i^{t,-} \sim \mathcal{D}_i(\boldsymbol{\psi}^{t,-})$, and player i observes $Y_i^{t,+} := J_i(\boldsymbol{\xi}_i^{t,+}; \boldsymbol{\psi}^{t,+})$, $Y_i^{t,-} := J_i(\boldsymbol{\xi}_i^{t,-}; \boldsymbol{\psi}^{t,-})$. To obtain the variance advantage of two-point feedback in the presence of stochastic performative outcomes, we require the two realizations to admit a sufficiently regular coupling.

Assumption 10 (Paired multipoint feedback) *At every iteration t, the two samples $(\boldsymbol{\xi}_i^{t,+}, \boldsymbol{\xi}_i^{t,-})$ have marginals $\mathcal{D}_i(\boldsymbol{\psi}^{t,+})$ and $\mathcal{D}_i(\boldsymbol{\psi}^{t,-})$, respectively, and are coupled so that, for some $\kappa_i \geq 0$,*

$$\mathbb{E}\left[\left\|\boldsymbol{\xi}_i^{t,+} - \boldsymbol{\xi}_i^{t,-}\right\|_2^2 \,\middle|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\right] \leq \kappa_i^2 \left\|\boldsymbol{\psi}^{t,+} - \boldsymbol{\psi}^{t,-}\right\|_2^2.$$

This assumption is analogous to the stochastic two-point oracle of Nesterov & Spokoiny (2017), extended to performative environments. It is naturally satisfied when the environment admits common-random-number queries, where the same exogenous shock is used for the two deployments while the action-dependent component of the environment still changes with $\boldsymbol{\psi}^{t,\pm}$ (e.g. $\boldsymbol{\eta}_i$ in Example 1).

Using the two observations, player i constructs the *two-point estimator*

$$\widehat{\nabla}_i^{\mathrm{MP}} J_i \;:=\; \frac{d\left(Y_i^{t,+} - Y_i^{t,-}\right)}{2\sigma_t} \boldsymbol{z}_i^t. \tag{9}$$

Since $\boldsymbol{\psi}_i^{t,\pm} = \boldsymbol{\theta}_i^{t,\sigma} \pm \sigma_t \boldsymbol{z}_i^t$, Lemma 3 shows that Estimator (9) can be viewed as the symmetric two-query counterpart of Estimator (5), with more desirable variance properties.

Lemma 3 *Under deployed-action sampling, $\mathbb{E}\left[\widehat{\nabla}_i^{\mathrm{MP}} J_i \,\middle|\, \boldsymbol{\theta}^t\right] = \nabla_{\boldsymbol{\theta}_i} \widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}; \sigma_t)$. Moreover, under Assumption 10, the second moment is bounded independently of σ_t:*

$$\mathbb{E}\left[\left\|\widehat{\nabla}_i^{\mathrm{MP}} J_i\right\|_2^2 \,\middle|\, \boldsymbol{\theta}^t\right] \leq 2nd^2 G_i^2 (1 + \kappa_i^2).$$

SPADE-MP replaces the single payoff observation in SPADE-NE with the paired evaluations (8) and estimator (9). Furthermore, since Lemma 3 removes the $1/\sigma_t^2$ growth in the estimator second moment, the primal step no longer requires the additional σ_t^2 attenuation used by the one-point SPADE and SPADE-NE. We thus keep the dual-consensus update (7c), while updating

$$\boldsymbol{\theta}_i^{t+1} = P_{\boldsymbol{\Omega}_i}\left[\boldsymbol{\theta}_i^t - \gamma_t \left(\widehat{\nabla}_i^{\mathrm{MP}} J_i + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t\right)\right], \tag{10}$$

Algorithm 3 in Appendix B displays the full pseudocode.

Theorem 3 (Convergence of SPADE-MP, informal) *Suppose Assumptions 1-6 and 8-10 hold. For sufficiently small, nonincreasing, positive stepsizes $\gamma_t, \sigma_t = \Theta(t^{-1/2})$, we get for all i,*

$$\mathcal{R}_i^{\mathrm{NE}}(T),\ \mathcal{R}_g(T) \leq \boldsymbol{\mathcal{O}}\left(\boldsymbol{T^{3/4}}\right).$$

[1]The same construction can target the PSE by replacing deployed-action sampling with intended-action sampling as in SPADE: draw a single $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t)$ and evaluate both $J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^{t,+})$ and $J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^{t,-})$ under this common frozen environment realization.

The improvement relative to the one-point SPADE-NE bound comes from the second-moment estimate in Lemma 3. This allows SPADE-MP to attain the same $\mathcal{O}(T^{3/4})$ rate (for NE) as the full-information method of Yan & Cao (2024) (for PSE). We defer the rest of the details to Appendix F.

4 Numerical Experiments

Here, we evaluate the empirical effectiveness of SPADE-MP on a foundational decentralized performative game model: oligopolistic networked Cournot competition Allaz & Vila (1993); Abolhassani et al. (2014). The experiment is intended to confirm Theorem 3 showing that SPADE-MP drives the time-averaged NE-performative regret and constraint violation toward zero. The interested reader can find full details and demonstrations of the one-point SPADE and SPADE-NE in Appendix G.

Specifically, we illustrate SPADE-MP's performance on the baseline five-firm, three-market model of Yan & Cao (2024), with sparse market participation, individual firm capacities, coupled market capacities, and a linear performative shift in stochastic demand. SPADE-MP uses an antithetic pair of payoff queries per iteration, coupled via a common exogenous market-shock draw. Full details of economic parameters, the communication graph, feasible query construction, randomization, and equilibrium references are in Appendix G.1. We plot each firm's time-averaged NE-performative regret $\mathcal{R}_i^{\mathrm{NE}}(t)/t$ and each market's signed time-averaged capacity residual $1/t \sum_{t'=1}^{t} \sum_i g_{ij}(\boldsymbol{\theta}_i^{t'})$.

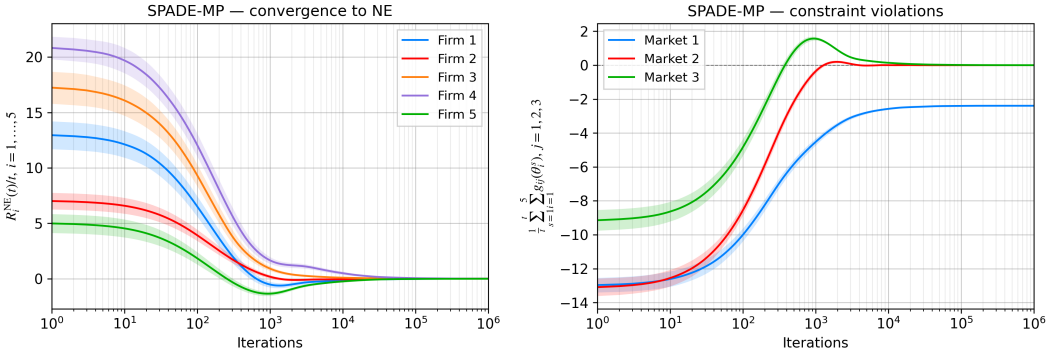


Figure 1: SPADE-MP time-averaged NE regret (left) and signed time-averaged market-capacity residuals (right). Curves show mean over 30 runs with $T = 10^6$ and bands are ± 1 sample SD. Note that NE regret need not be positive, since joint actions might not satisfy shared market constraints.

Results. Figure 1 shows that the time-averaged regret of every firm goes to 0, with mean final-iterate NE-distance at 0.07, while the cumulative constraint violation satisfies $\mathcal{R}_g(T)/T = 1.47 \times 10^{-3}$ on average, consistent with Theorem 3. The second and third market-capacity residuals approach zero, while the first approaches -2.39, agreeing with the independently computed NE's market 1 slackness. Visually, the smooth curves and comparatively narrow bands are consistent with the bounded-variance two-point estimator underlying Theorem 3, and the convergence shape closely resembles that of the $\mathcal{O}(T^{3/4})$ guarantee in the full-information setting of Yan & Cao (2024). While payoff-based approaches are still slower in absolute terms, all three SPADE algorithms crucially communicate only the local dual variable with each neighbor, incurring significantly less overhead. On our specified Cournot instance ($n = 5$, $d = 3$), this gives 42 directed scalar transmissions per round, compared with 252 ($\times 6$ burden increase) for the dense decentralized protocol of Yan & Cao (2024) (see Figure 6). In general, this performance improvement scales linearly with $n \cdot d$, i.e. the number of agents in the network multiplied with the dimensionality of the models. Full simulation figures of all three algorithms and details on the overhead calculations are given in Appendix G.3.

Conclusion. We studied decentralized noncooperative performative games with coupled constraints and developed payoff-based algorithms for computing PSE and NE under different feedback models. Our results show that communication-efficient learning is possible even when agents only observe payoffs. Our work is primarily theoretical, but may have societal impact in strategic learning domains. Future work includes improving rates and validating the approach in more realistic deployments.

AI USE STATEMENT

In this work, we used generative AI tools to edit or shorten portions of the manuscript, improve the clarity and organization of our problem presentation and math, format LaTeX, and formulate revisions in response to reviewer feedback. We also used tools to assist with key ingredients of proofs, check and discuss intermediate mathematical derivations, provide feedback on the design and presentation of experiments, and assist with the implementation of algorithms and numerical simulations. All AI-assisted proofs, derivations, algorithmic implementations, experimental choices, and edited material were reviewed, rewritten if necessary, and verified to the authors' best knowledge. All literature sourcing was conducted manually by the authors, while AI was used to summarize and analyze sourced literature. LLM-generated code was tested and verified by multiple authors. We did not rely on generative AI as an independent source of research results or interpretations. We did not use AI to assist with translation, cleaning or reformatting datasets, generating synthetic datasets outside the exactly specified simulations, supporting qualitative or thematic data analysis, or formulating our main theoretical claims, model, and conceptual frameworks. The authors take full responsibility for the final content of this work, including all text, claims, proofs, experiments, code, and other artifacts produced with the help of AI.

ETHICS STATEMENT

This work is primarily theoretical and studies learning algorithms for decentralized noncooperative games with decision-dependent distributions. Our numerical experiments use only synthetic environments and do not involve human subjects, personal or sensitive data, or real-world deployment. Nonetheless, applications of performative learning may arise in consequential economic or decision-making settings, where strategic responses, fairness, transparency, or temporary constraint violations during learning may be important considerations. Our results should therefore not be interpreted as providing guarantees of fairness, privacy, safety, or suitability for deployment beyond the equilibrium and convergence properties established in the paper.

REPRODUCIBILITY STATEMENT

We provide the assumptions and algorithmic specifications needed to reproduce the theoretical results in the main text, with complete proofs in Appendices D, E, and F, and supporting lemmas in Appendix C. The numerical experiments are entirely synthetic and Appendix G specifies the networked Cournot environment, distributional mappings, coupled constraints, parameter choices, step-size schedules, randomization procedure, evaluation metrics, and computation of the reference equilibria, which are sufficient for reimplementation.

REFERENCES

Melika Abolhassani, Mohammad Hossein Bateni, Mohammad Taghi Hajiaghayi, Hamid Mahini, and Anshul Sawant. Network Cournot competition. In *International Conference on Web and Internet Economics*, pp. 15–29. Springer, 2014.

Alekh Agarwal, Ofer Dekel, and Lin Xiao. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *COLT*, pp. 28–40, 2010.

Blaise Allaz and Jean-Luc Vila. Cournot competition, forward markets and efficiency. *Journal of Economic Theory*, 59(1):1–16, 1993.

Yu Bai, Chi Jin, and Tiancheng Yu. Near-optimal reinforcement learning with self-play. *Advances in Neural Information Processing Systems*, 33:2159–2170, 2020.

Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

Mark Bils and Peter J Klenow. Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985, 2004.

Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8(6), 2007.

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

Mario Bravo, David Leslie, and Panayotis Mertikopoulos. Bandit learning in concave n-person games. *Advances in Neural Information Processing Systems*, 31, 2018.

Nicolo Cesa-Bianchi, Gábor Lugosi, et al. *Prediction, learning, and games*, volume 1. Cambridge university press Cambridge, 2006.

Alex Chan, Ahmed Alaa, Zhaozhi Qian, and Mihaela Van Der Schaar. Unlabelled data improves bayesian uncertainty calibration under covariate shift. In *International Conference on Machine Learning*, pp. 1392–1402. PMLR, 2020.

Ka-Po Chow and Yu-Kwong Kwok. On load balancing for distributed multiagent computing. *IEEE Transactions on Parallel and Distributed Systems*, 13(8):787–801, 2002.

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

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

Dmitriy Drusvyatskiy, Maryam Fazel, and Lillian J Ratliff. Improved rates for derivative free gradient play in strongly monotone games. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 3403–3408. IEEE, 2022.

Francisco Facchinei and Jong-Shi Pang. Chapter 12: Nash equilibria: The variational approach. *Convex Optimization in Signal Processing and Communications*, pp. 443, 2010.

Bassam Fattouh and Lavan Mahadeva. Causes and implications of shifts in financial participation in commodity markets. *Journal of Futures Markets*, 34(8):757–787, 2014.

Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

Tatsuhiko Goto, Takeshi Hatanaka, and Masayuki Fujita. Payoff-based inhomogeneous partially irrational play for potential game theoretic cooperative control: Convergence analysis. In *2012 American Control Conference (ACC)*, pp. 2380–2387. IEEE, 2012.

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

Yasunari Hikima. Multi-point feedback of bandit convex optimization with hard constraints. *arXiv preprint arXiv:2310.10946*, 2023.

Zachary Izzo, Lexing Ying, and James Zou. How to learn when data reacts to your model: performative gradient descent. In *International Conference on Machine Learning*, pp. 4641–4650. PMLR, 2021.

Meena Jagadeesan, Tijana Zrnic, and Celestine Mendler-Dünner. Regret minimization with performative feedback. In *International Conference on Machine Learning*, pp. 9760–9785. PMLR, 2022.

Chi Jin, Qinghua Liu, Yuanhao Wang, and Tiancheng Yu. V-learning–a simple, efficient, decentralized algorithm for multiagent RL. *arXiv preprint arXiv:2110.14555*, 2021.

Peter J Klenow and Jonathan L Willis. Sticky information and sticky prices. *Journal of Monetary Economics*, 54:79–99, 2007.

Qiang Li, Chung-Yiu Yau, and Hoi-To Wai. Multi-agent performative prediction with greedy deployment and consensus seeking agents. *Advances in Neural Information Processing Systems*, 35:38449–38460, 2022.

Jie Lu, Anjin Liu, Fan Dong, Feng Gu, Joao Gama, and Guangquan Zhang. Learning under concept drift: A review. *IEEE Transactions on Knowledge and Data Engineering*, 31(12):2346–2363, 2018.

Songtao Lu. Bilevel optimization with coupled decision-dependent distributions. In *International Conference on Machine Learning*, pp. 22758–22789. PMLR, 2023.

Debmalya Mandal, Stelios Triantafyllou, and Goran Radanovic. Performative reinforcement learning. In *International Conference on Machine Learning*, pp. 23642–23680. PMLR, 2023.

Jason R Marden and Jeff S Shamma. Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior*, 75(2):788–808, 2012.

Richard P Martin, Amin M Vahdat, David E Culler, and Thomas E Anderson. Effects of communication latency, overhead, and bandwidth in a cluster architecture. *ACM SIGARCH Computer Architecture News*, 25(2):85–97, 1997.

John P Miller, Juan C Perdomo, and Tijana Zrnic. Outside the echo chamber: Optimizing the performative risk. In *International Conference on Machine Learning*, pp. 7710–7720. PMLR, 2021.

Adhyyan Narang, Evan Faulkner, Dmitriy Drusvyatskiy, Maryam Fazel, and Lillian J Ratliff. Multiplayer performative prediction: Learning in decision-dependent games. *Journal of Machine Learning Research*, 24(202):1–56, 2023.

Yurii Nesterov and Vladimir Spokoiny. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017.

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

Georgios Piliouras and Fang-Yi Yu. Multi-agent performative prediction: From global stability and optimality to chaos. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pp. 1047–1074, 2023.

Julio J Rotemberg. Sticky prices in the United States. *Journal of Political Economy*, 90(6):1187–1211, 1982.

Tim Roughgarden. Algorithmic game theory. *Communications of the ACM*, 53(7):78–86, 2010.

Jose Carlos Sancho, Darren J Kerbyson, and Kevin J Barker. Efficient offloading of collective communications in large-scale systems. In *2007 IEEE International Conference on Cluster Computing*, pp. 169–178. IEEE, 2007.

Jia-Wei Shan, Peng Zhao, and Zhi-Hua Zhou. Beyond performative prediction: Open-environment learning with presence of corruptions. In *International Conference on Artificial Intelligence and Statistics*, pp. 7981–7998. PMLR, 2023.

Tatiana Tatarenko and Maryam Kamgarpour. Learning generalized nash equilibria in a class of convex games. *IEEE Transactions on Automatic Control*, 64(4):1426–1439, 2018.

Tatiana Tatarenko and Maryam Kamgarpour. Bandit online learning of nash equilibria in monotone games. *arXiv preprint arXiv:2009.04258*, 2021.

Hal R Varian. Online ad auctions. *American Economic Review*, 99(2):430–434, 2009.

Xiaolu Wang, Chung-Yiu Yau, and Hoi-To Wai. Network effects in performative prediction games. In *International Conference on Machine Learning*, pp. 36514–36540. PMLR, 2023.

Killian Wood, Gianluca Bianchin, and Emiliano Dall'Anese. Online projected gradient descent for stochastic optimization with decision-dependent distributions. *IEEE Control Systems Letters*, 6: 1646–1651, 2021.

Ruihan Wu, Chuan Guo, Yi Su, and Kilian Q Weinberger. Online adaptation to label distribution shift. *Advances in Neural Information Processing Systems*, 34:11340–11351, 2021.

Wenjing Yan and Xuanyu Cao. Zero-regret performative prediction under inequality constraints. *Advances in Neural Information Processing Systems*, 36:1298–1308, 2023.

Wenjing Yan and Xuanyu Cao. Decentralized noncooperative games with coupled decision-dependent distributions. *Advances in Neural Information Processing Systems*, 37:25592–25627, 2024.

Zhi-Hua Zhou. Open-environment machine learning. *National Science Review*, 9(8):nwac123, 2022.

APPENDIX CONTENTS

A Literature Review

B Algorithm Pseudocode

C Auxiliary Lemmas

D Convergence of SPADE (Algorithm 1): Payoff-Based Stochastic Primal-Dual Update for PSE

E Convergence of SPADE-NE (Algorithm 2): Payoff-Based Stochastic Primal-Dual Update for NE

F Convergence of SPADE-MP (Algorithm 3): Multipoint Payoff-Based Update for NE

G Experimental Details

A LITERATURE REVIEW

Here, we expand on the relevant literature and situate our results within their broader context.

Performative learning. As stated previously, performativity and decision-dependent learning has been studied in a variety of settings. Shan et al. (2023) studied decision-dependent distributions in open environments where data may be corrupted, proposing algorithms that decouple environmental noise from performative effects. For bilevel optimization, Lu (2023) analyzed the existence of performatively stable solutions and proposed stochastic algorithms for computing them. Performative effects have also been studied in reinforcement learning and stochastic optimization. Mandal et al. (2023) introduce *performative reinforcement learning*, where the deployed policy changes both rewards and transition dynamics, and establish convergence of repeated regularized policy optimization, projected gradient ascent, and finite-sample empirical procedures to performatively stable policies. Drusvyatskiy & Xiao (2023) study stochastic optimization under decision-dependent distributions, showing that distribution shift induces a gradient bias that decays with distance to equilibrium and using this to derive convergence guarantees for stochastic gradient, proximal-point, clipped-gradient, dual-averaging, and accelerated methods. Beyond stability analysis, several works have investigated optimality in performative prediction problems, including Miller et al. (2021); Izzo et al. (2021); Jagadeesan et al. (2022), which study the gap between performatively stable solutions and performatively optimal decisions.

At the multi-agent level, Narang et al. (2023) introduce the framework of *multiplayer performative prediction*, formalizing both PSE and NE. For NE, they derive sufficient conditions for strong monotonicity and study stochastic-gradient, adaptive, and derivative-free procedures. Wang et al. (2023) study performative games over multiplex networks, where agents cooperate over one network while their data distributions depend on neighboring models through a second population network. Related work also studies cooperative multi-agent performative learning. Li et al. (2022) formulate a decentralized consensus problem in which agents seek a common decision while each local data distribution depends on its own local decision. Piliouras & Yu (2023) instead analyze multi-agent performative dynamics under structured location-scale distribution maps, establishing regimes of global stability and optimality. Finally, Yan & Cao (2023) introduce performative prediction under inequality constraints in the single-agent setting, developing a robust primal-dual framework for performative optimality.

Our work differs from prior work by addressing *decentralized noncooperative games with coupled decision-dependent distributions and shared constraints*, a setting absent from this literature, and developing *communication-efficient* payoff-based dynamics that exchange only dual variables across the network rather than primal actions or gradients.

Derivative-free/payoff-based learning in games. Learning equilibria from payoff observations alone is comparatively underexplored relative to the extensive literature on first-order methods for games. The seminal contribution of Bravo et al. (2018) extends mirror descent to the bandit setting in concave N-person games, using a one-point perturbation estimator to construct a gradient surrogate from a single function evaluation. They establish almost-sure convergence to the Nash equilibrium in monotone games and an $O(1/T^{1/3})$ rate under strong monotonicity. Tatarenko & Kamgarpour

(2018) address the related problem of learning generalized Nash equilibria in convex games with *coupling constraints*, proposing a payoff-based primal-dual scheme in which players sample actions from Gaussian distributions whose means are updated using only realized cost values, and prove convergence under strict (or strong) monotonicity of the game mapping. Tatarenko & Kamgarpour (2021) subsequently relax this requirement, designing a doubly-regularized payoff-based procedure that converges to the least-norm Nash equilibrium in *merely* monotone games and quantifying its no-regret rate. Most recently, Drusvyatskiy et al. (2022) sharpen the analysis of Bravo et al. (2018) by reinterpreting the one-point method as stochastic gradient play on a perturbed strongly monotone game, improving the iteration complexity to $O(d^2/\varepsilon^2)$.

Payoff-based learning is also central to the *online-learning and multi-agent reinforcement-learning* literature Cesa-Bianchi et al. (2006); Roughgarden (2010); Bai et al. (2020). No-external-regret dynamics, including bandit variants of multiplicative-weights methods Freund & Schapire (1999), yield convergence of empirical play to coarse correlated equilibria, while swap-regret dynamics yield correlated equilibria Hart & Mas-Colell (2000); Blum & Mansour (2007). Related ideas underlie recent multi-agent reinforcement-learning algorithms, such as V-learning Jin et al. (2021), that learn equilibrium policies from trajectory-level reward feedback in Markov games. These works differ from our setting in both the solution concept and source of nonstationarity. We seek convergence to (unique) PSE or NE in continuous games with coupled constraints, where players' decisions endogenously alter the distributions generating their payoffs.

A complementary line of work shows that access to multiple payoff queries can substantially reduce the statistical cost of zeroth-order feedback. Agarwal et al. (2010) introduce the multi-point bandit model for online convex optimization and show that only two antithetic function evaluations per round suffice to recover near-full-information guarantees. Nesterov & Spokoiny (2017) develop random gradient-free methods based on finite-difference oracles and study a symmetric two-point estimator similar to our SPADE-MP estimate construction. Building on these ideas, Hikima (2023) studies two-point bandit convex optimization with constraints and obtain, for convex losses, $O(d^2\sqrt{T})$ regret together with $O(d^2T^{3/4})$ cumulative hard constraint violation.

However, these results treat the queried loss or stochastic environment as exogenous to the learner's current deployment. They therefore do not address the performative setting, where the perturbed actions themselves induce different distributions of the payoff-relevant variable. Our work fills a gap by developing payoff-based, derivative-free algorithms for decentralized noncooperative games with coupled decision-dependent distributions, and by carefully designing the gradient estimation procedure to target either PSE or NE efficiently depending on the form of payoff feedback available to the players.

B ALGORITHM PSEUDOCODE

For reference, we collect here the full pseudocode of the three algorithms presented in Section 3. The listings below add the explicit initialization, neighbor-exchange, and per-round loop structure.

Algorithm 1 SPADE: Stochastic Payoff-Based Algorithm for Decentralized Equilibria, for all $i \in [n]$

Inputs: stepsizes $\{\gamma_t\}$; smoothing $\{\sigma_t\}$; dual bound $\Lambda_{\max}$; ball center $\boldsymbol{p}_i$ and radius R_i.
Initialize $\boldsymbol{\theta}_i^1 \in \boldsymbol{\Omega}_i$ arbitrarily. Set $\boldsymbol{\lambda}_i^1 = 0$.
for $t = 1, \dots, T$ **do**
  Exchange $\boldsymbol{\lambda}_i^t$ with all neighbors.
  Compute the inward-shifted center $\boldsymbol{\theta}_i^{t,\sigma} = \left(1 - \frac{\sigma_t}{R_i}\right)\boldsymbol{\theta}_i^t + \frac{\sigma_t}{R_i}\boldsymbol{p}_i$.
  Sample $\boldsymbol{z}_i^t \sim \mathrm{Unif}(\mathbb{S}^{d-1})$ independently across players and set $\boldsymbol{\psi}_i^t = \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{z}_i^t$.
  Deploy $\boldsymbol{\psi}_i^t$ and sample $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i\big(\boldsymbol{\theta}_i^t, \boldsymbol{\theta}_{-i}^t\big)$.
  Observe the realized cost $J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)$ and local constraint violation $\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)$.
  Compute the zeroth-order gradient estimator

$$\widehat{\nabla}_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t) = \frac{d}{\sigma_t} J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t) \boldsymbol{z}_i^t.$$

  Update the primal variable:

$$\boldsymbol{\theta}_i^{t+1} = P_{\boldsymbol{\Omega}_i}\left[\boldsymbol{\theta}_i^t - \gamma_t(\sigma_t)^2 \left(\widehat{\nabla}_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t\right)\right]. \tag{11}$$

  Update the dual variable:

$$\boldsymbol{\lambda}_i^{t+1} = P_{\boldsymbol{\Lambda}}\left[\sum_{j \in \mathcal{N}_i} a_{ij}\, \boldsymbol{\lambda}_j^t + \gamma_t\, \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right].$$

end for

Algorithm 2 SPADE-NE: Payoff-Based Stochastic Primal–Dual Algorithm for NE, for all $i \in [n]$

Inputs: stepsizes $\{\gamma_t\}$; smoothing $\{\sigma_t\}$; dual bound $\Lambda_{\max}$; ball center $\boldsymbol{p}_i$ and radius R_i.
Initialize $\boldsymbol{\theta}_i^1 \in \boldsymbol{\Omega}_i$ arbitrarily. Set $\boldsymbol{\lambda}_i^1 = 0$.
for $t = 1, \dots, T$ **do**
  Exchange $\boldsymbol{\lambda}_i^t$ with all neighbors.
  Compute the inward-shifted center $\boldsymbol{\theta}_i^{t,\sigma} = \left(1 - \frac{\sigma_t}{R_i}\right)\boldsymbol{\theta}_i^t + \frac{\sigma_t}{R_i}\boldsymbol{p}_i$.
  Sample $\boldsymbol{z}_i^t \sim \mathrm{Unif}(\mathbb{S}^{d-1})$ independently across players and set $\boldsymbol{\psi}_i^t = \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{z}_i^t$.
  Deploy $\boldsymbol{\psi}_i^t$ and sample $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\psi}^t)$ *(distribution induced by deployed action)*.
  Observe the realized cost $J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)$ and local constraint violation $\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)$.
  Compute the zeroth-order gradient estimator

$$\widehat{\nabla}_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t) = \frac{d}{\sigma_t} J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t) \boldsymbol{z}_i^t.$$

  Update the primal variable:

$$\boldsymbol{\theta}_i^{t+1} = P_{\boldsymbol{\Omega}_i}\left[\boldsymbol{\theta}_i^t - \gamma_t(\sigma_t)^2 \left(\widehat{\nabla}_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t\right)\right]. \tag{12}$$

  Update the dual variable:

$$\boldsymbol{\lambda}_i^{t+1} = P_{\boldsymbol{\Lambda}}\left[\sum_{j \in \mathcal{N}_i} a_{ij}\, \boldsymbol{\lambda}_j^t + \gamma_t\, \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right].$$

end for

Algorithm 3 SPADE-MP: Multipoint Payoff-Based Algorithm for NE, for all $i \in [n]$

Inputs: stepsizes $\{\gamma_t\}$; smoothing parameters $\{\sigma_t\}$; dual bound $\Lambda_{\max}$; inscribed-ball center $\boldsymbol{p}_i$ and radius R_i.
Initialize $\boldsymbol{\theta}_i^1 \in \boldsymbol{\Omega}_i$ arbitrarily. Set $\boldsymbol{\lambda}_i^1 = 0$.
for $t = 1, \ldots, T$ **do**
  Exchange $\boldsymbol{\lambda}_i^t$ with all neighbors.
  Compute the inward-shifted center $\boldsymbol{\theta}_i^{t,\sigma} = \left(1 - \frac{\sigma_t}{R_i}\right) \boldsymbol{\theta}_i^t + \frac{\sigma_t}{R_i} \boldsymbol{p}_i$.
  Sample $\boldsymbol{z}_i^t \sim \mathrm{Unif}(\mathbb{S}^{d-1})$ independently across players.
  All players synchronously deploy the positive perturbation $\boldsymbol{\psi}_i^{t,+} = \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{z}_i^t$.
  Sample $\boldsymbol{\xi}_i^{t,+} \sim \mathcal{D}_i(\boldsymbol{\psi}^{t,+})$ and observe $Y_i^{t,+} = J_i(\boldsymbol{\xi}_i^{t,+}; \boldsymbol{\psi}^{t,+})$.
  Keeping $\boldsymbol{\theta}^t$, $\boldsymbol{\lambda}^t$, and $\boldsymbol{z}^t$ fixed, all players synchronously deploy the antithetic perturbation $\boldsymbol{\psi}_i^{t,-} = \boldsymbol{\theta}_i^{t,\sigma} - \sigma_t \boldsymbol{z}_i^t$.
  Sample $\boldsymbol{\xi}_i^{t,-} \sim \mathcal{D}_i(\boldsymbol{\psi}^{t,-})$, coupled with $\boldsymbol{\xi}_i^{t,+}$ according to Assumption 10, and observe $Y_i^{t,-} = J_i(\boldsymbol{\xi}_i^{t,-}; \boldsymbol{\psi}^{t,-})$.
  Compute the two-point payoff-based gradient estimator

$$\widehat{\nabla}_i^{\mathrm{MP}} J_i = \frac{d\left(Y_i^{t,+} - Y_i^{t,-}\right)}{2\sigma_t} \boldsymbol{z}_i^t.$$

  Observe the local constraint violation $\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)$.
  Update the primal variable:

$$\boldsymbol{\theta}_i^{t+1} = P_{\boldsymbol{\Omega}_i}\left[\boldsymbol{\theta}_i^t - \gamma_t \left(\widehat{\nabla}_i^{\mathrm{MP}} J_i + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t\right)\right]. \tag{13}$$

  Update the dual variable:

$$\boldsymbol{\lambda}_i^{t+1} = P_{\boldsymbol{\Lambda}}\left[\sum_{j \in \mathcal{N}_i} a_{ij}\, \boldsymbol{\lambda}_j^t + \gamma_t\, \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right].$$

end for

C AUXILIARY LEMMAS

In the following convergence proofs, we also utilize the following auxiliary lemmas on the dual variables. Define the average dual iterate $\bar{\boldsymbol{\lambda}}^t := \frac{1}{n}\sum_{i=1}^n \boldsymbol{\lambda}_i^t$ and neighbor-average dual iterate $\boldsymbol{w}_i^t := \sum_{j\in\mathcal{N}_i} a_{ij}\boldsymbol{\lambda}_j^t$.

Lemma 4 (Dual boundedness) *With $\boldsymbol{\lambda}_i^1 = \mathbf{0}$, we have $\boldsymbol{\lambda}_i^t \in \boldsymbol{\Lambda}$ and $\bar{\boldsymbol{\lambda}}^t \in \boldsymbol{\Lambda}$ for all $i \in [n]$ and $t \in [T]$. In particular $\|\boldsymbol{\lambda}_i^t\|_2 \le \Lambda_{\max}$, $\boldsymbol{\lambda}_i^t \ge \mathbf{0}$, and $\sum_{i=1}^n \|\boldsymbol{\lambda}_i^t\|_2^2 \le n\Lambda_{\max}^2$.*

Proof: $\mathbf{0} \in \boldsymbol{\Lambda}$, and the update returns a point of $\boldsymbol{\Lambda}$ by construction, so $\boldsymbol{\lambda}_i^t \in \boldsymbol{\Lambda}$ by induction. Since $\boldsymbol{\Lambda}$ is convex, the average $\bar{\boldsymbol{\lambda}}^t$ and the neighbor averages $\boldsymbol{w}_i^t$ also lie in $\boldsymbol{\Lambda}$. □

Lemma 5 (Dual consensus error) *Let $\sigma_2(\boldsymbol{A})$ denote the second largest singular value of $\boldsymbol{A}$, and assume $\sigma_2(\boldsymbol{A}) < 1$. Then*

$$\sum_{t=1}^{T}\sum_{i=1}^{n} \left\|\boldsymbol{\lambda}_i^t - \bar{\boldsymbol{\lambda}}^t\right\|_2 \;\le\; \frac{nB}{1-\sigma_2(\boldsymbol{A})}\sum_{t=1}^{T}\gamma_t \;=\; \mathcal{O}\!\left(\sum_{t=1}^{T}\gamma_t\right).$$

Proof: Since $\boldsymbol{w}_i^t \in \boldsymbol{\Lambda}$, write $\boldsymbol{\lambda}_i^{t+1} = \boldsymbol{w}_i^t + \boldsymbol{\omega}_i^t$ where $\boldsymbol{\omega}_i^t := P_{\boldsymbol{\Lambda}}[\boldsymbol{w}_i^t + \gamma_t \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)] - P_{\boldsymbol{\Lambda}}[\boldsymbol{w}_i^t]$. Nonexpansiveness of $P_{\boldsymbol{\Lambda}}$ gives

$$\|\boldsymbol{\omega}_i^t\|_2 = \left\|P_{\boldsymbol{\Lambda}}\left[\boldsymbol{w}_i^t + \gamma_t \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right] - P_{\boldsymbol{\Lambda}}[\boldsymbol{w}_i^t]\right\|_2 \le \gamma_t \|\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\|_2 \le \gamma_t B.$$

Stack the dual iterates row-wise into $\boldsymbol{\Lambda}^t \in \mathbb{R}^{n\times m}$ and similarly stack $\boldsymbol{\omega}_i^t$ into $\boldsymbol{\Omega}^t \in \mathbb{R}^{n\times m}$. Then the dual update across all players can be written as

$$\boldsymbol{\Lambda}^{t+1} = \boldsymbol{A}\boldsymbol{\Lambda}^t + \boldsymbol{\Omega}^t.$$

Let $\boldsymbol{J} := I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\top$ and $\boldsymbol{\Delta}^t := \boldsymbol{J}\boldsymbol{\Lambda}^t$ be the dual disagreement matrix at iteration t, or in other words, the i-th row of $\boldsymbol{\Delta}^t$ is $\boldsymbol{\lambda}_i^t - \bar{\boldsymbol{\lambda}}^t$. Since $\boldsymbol{A}$ is doubly stochastic, we have $\boldsymbol{J}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{J}$. Thus,

$$\begin{aligned}
\boldsymbol{\Delta}^{t+1} &= \boldsymbol{J}\boldsymbol{\Lambda}^{t+1} \\
&= \boldsymbol{J}\left(\boldsymbol{A}\boldsymbol{\Lambda}^t + \boldsymbol{\Omega}^t\right) \\
&= \boldsymbol{A}\boldsymbol{J}\boldsymbol{\Lambda}^t + \boldsymbol{J}\boldsymbol{\Omega}^t \\
&= \boldsymbol{A}\boldsymbol{\Delta}^t + \boldsymbol{J}\boldsymbol{\Omega}^t.
\end{aligned}$$

Since $\boldsymbol{\Delta}^t$ lies in the disagreement subspace $\{\boldsymbol{X} : \mathbf{1}^\top\boldsymbol{X} = 0\}$, multiplication by the doubly-stochastic $\boldsymbol{A}$ contracts this component by $\sigma_2(\boldsymbol{A})$. Thus, using $\|\boldsymbol{J}\|_2 \le 1$, we have that

$$\begin{aligned}
\|\boldsymbol{\Delta}^{t+1}\|_F &\le \|\boldsymbol{A}\boldsymbol{\Delta}^t\|_F + \|\boldsymbol{J}\boldsymbol{\Omega}^t\|_F \\
&\le \sigma_2(\boldsymbol{A})\|\boldsymbol{\Delta}^t\|_F + \|\boldsymbol{\Omega}^t\|_F \\
&\le \sigma_2(\boldsymbol{A})\|\boldsymbol{\Delta}^t\|_F + \sqrt{n}\,\gamma_t B,
\end{aligned} \tag{14}$$

where $\|\cdot\|_F$ denotes the Frobenius norm and the last inequality follows from $\|\boldsymbol{\Omega}^t\|_F^2 = \sum_{i=1}^n \|\boldsymbol{\omega}_i^t\|_2^2 \le n\gamma_t^2 B^2$.

Since $\boldsymbol{\lambda}_i^1 = \mathbf{0}$ for all i, we have $\boldsymbol{\Delta}^1 = \mathbf{0}$. Thus, iterating (14) gives

$$\begin{aligned}
\|\boldsymbol{\Delta}^t\|_F &\le \sqrt{n}B\sum_{s=1}^{t-1}\sigma_2(\boldsymbol{A})^{\,t-1-s}\gamma_s \\
&= \sqrt{n}B\sum_{k=0}^{t-2}\sigma_2(\boldsymbol{A})^k\gamma_{t-1-k}.
\end{aligned}$$

Moreover, since the ith row of $\boldsymbol{\Delta}^t$ is $\boldsymbol{\lambda}_i^t - \bar{\boldsymbol{\lambda}}^t$, Cauchy–Schwarz gives

$$\sum_{i=1}^{n}\|\boldsymbol{\lambda}_i^t - \bar{\boldsymbol{\lambda}}^t\|_2 \le \sqrt{n}\left(\sum_{i=1}^{n}\|\boldsymbol{\lambda}_i^t - \bar{\boldsymbol{\lambda}}^t\|_2^2\right)^{1/2} = \sqrt{n}\|\boldsymbol{\Delta}^t\|_F.$$

Thus, applying the bound on $\|\boldsymbol{\Delta}^t\|_F$,

$$\begin{aligned}\sum_{t=1}^{T}\sum_{i=1}^{n}\|\boldsymbol{\lambda}_i^t-\bar{\boldsymbol{\lambda}}^t\|_2 &\le nB\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sigma_2(\boldsymbol{A})^{\,t-1-s}\gamma_s\\ &= nB\sum_{s=1}^{T-1}\gamma_s\sum_{t=s+1}^{T}\sigma_2(\boldsymbol{A})^{\,t-1-s}\\ &\le nB\sum_{s=1}^{T-1}\gamma_s\sum_{k=0}^{\infty}\sigma_2(\boldsymbol{A})^k\\ &\le \frac{nB}{1-\sigma_2(\boldsymbol{A})}\sum_{t=1}^{T}\gamma_t,\end{aligned}$$

which proves the stated bound. □

Lemma 6 (Dual tracking) *Let $\{\gamma_t\}$ be positive and nonincreasing. Then, for every $\boldsymbol{\lambda}\in\boldsymbol{\Lambda}$,*

$$\sum_{t=1}^{T}\sum_{i=1}^{n}\left\langle\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^t,\ \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle \ \le\ \frac{2n\Lambda_{\max}^2}{\gamma_T}+\frac{nB^2}{2}\sum_{t=1}^{T}\gamma_t+2B\sum_{t=1}^{T}\sum_{i=1}^{n}\left\|\boldsymbol{\lambda}_i^t-\bar{\boldsymbol{\lambda}}^t\right\|_2.$$

Together with Lemma 5 this is $\mathcal{O}\Big(1/\gamma_T+\sum_{t=1}^{T}\gamma_t\Big)$.

Proof: Fix $\boldsymbol{\lambda}\in\boldsymbol{\Lambda}$ and let $D_t:=\sum_{i=1}^{n}\|\boldsymbol{\lambda}_i^t-\boldsymbol{\lambda}\|_2^2$. Since $P_{\boldsymbol{\Lambda}}$ is nonexpansive and $P_{\boldsymbol{\Lambda}}[\boldsymbol{\lambda}]=\boldsymbol{\lambda}$,

$$\begin{aligned}D_{t+1}&=\sum_{i=1}^{n}\|\boldsymbol{\lambda}_i^{t+1}-\boldsymbol{\lambda}\|_2^2\\ &=\sum_{i=1}^{n}\left\|P_{\boldsymbol{\Lambda}}\left[\boldsymbol{w}_i^t+\gamma_t\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right]-P_{\boldsymbol{\Lambda}}[\boldsymbol{\lambda}]\right\|_2^2\\ &\le\sum_{i=1}^{n}\left\|\boldsymbol{w}_i^t+\gamma_t\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)-\boldsymbol{\lambda}\right\|_2^2\\ &=\sum_{i=1}^{n}\|\boldsymbol{w}_i^t-\boldsymbol{\lambda}\|_2^2+2\gamma_t\sum_{i=1}^{n}\left\langle\boldsymbol{w}_i^t-\boldsymbol{\lambda},\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle+\gamma_t^2\sum_{i=1}^{n}\|\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\|_2^2.\end{aligned}$$

By convexity of $\|\cdot\|_2^2$ and double stochasticity of $\boldsymbol{A}$, we have $\sum_i\|\boldsymbol{w}_i^t-\boldsymbol{\lambda}\|_2^2\le\sum_i\|\boldsymbol{\lambda}_i^t-\boldsymbol{\lambda}\|_2^2=D_t$. Together with $\|\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\|_2\le B$ from Lemma 1, we rearrange the above to get

$$2\gamma_t\sum_{i=1}^{n}\left\langle\boldsymbol{\lambda}-\boldsymbol{w}_i^t,\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle\le D_t-D_{t+1}+\gamma_t^2nB^2.$$

To replace $\boldsymbol{w}_i^t$ by $\boldsymbol{\lambda}_i^t$, we first note by Cauchy–Schwarz and $\|\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\|_2\le B$,

$$\left\langle\boldsymbol{w}_i^t-\boldsymbol{\lambda}_i^t,\ \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle\le\|\boldsymbol{w}_i^t-\boldsymbol{\lambda}_i^t\|_2\,\|\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\|_2\le B\|\boldsymbol{w}_i^t-\boldsymbol{\lambda}_i^t\|_2.$$

Then, noting that $\sum_i\|\boldsymbol{w}_i^t-\boldsymbol{\lambda}_i^t\|_2\le\sum_i\sum_j a_{ij}\|\boldsymbol{\lambda}_j^t-\boldsymbol{\lambda}_i^t\|_2\le 2\sum_i\|\boldsymbol{\lambda}_i^t-\bar{\boldsymbol{\lambda}}^t\|_2$, we have

$$\sum_{i=1}^{n}\left\langle\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^t,\ \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle\le\sum_{i=1}^{n}\left\langle\boldsymbol{\lambda}-\boldsymbol{w}_i^t,\ \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle+2B\sum_{i=1}^{n}\|\boldsymbol{\lambda}_i^t-\bar{\boldsymbol{\lambda}}^t\|_2.$$

Thus, after dividing by $2\gamma_t$ and plugging in,

$$\sum_{i=1}^{n}\left\langle\boldsymbol{\lambda}-\boldsymbol{\lambda}_i^t,\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle\le\frac{D_t-D_{t+1}}{2\gamma_t}+\frac{\gamma_tnB^2}{2}+2B\sum_{i=1}^{n}\left\|\boldsymbol{\lambda}_i^t-\bar{\boldsymbol{\lambda}}^t\right\|_2.$$

Summing over t, using Abel summation on the first term, $D_t\le 4n\Lambda_{\max}^2$ (Lemma 4), and the monotonicity of $1/\gamma_t$, we get

$$\sum_{t=1}^{T}\frac{D_t-D_{t+1}}{2\gamma_t}\le\frac{D_1}{2\gamma_1}+\sum_{t=2}^{T}D_t\left(\frac{1}{2\gamma_t}-\frac{1}{2\gamma_{t-1}}\right)\le\frac{4n\Lambda_{\max}^2}{2\gamma_T}=\frac{2n\Lambda_{\max}^2}{\gamma_T},$$

which proves the claim. □

D CONVERGENCE OF SPADE (ALGORITHM 1): PAYOFF-BASED STOCHASTIC PRIMAL-DUAL UPDATE FOR PSE

We present the proof of Theorem 1.

Proof: For each player $i \in [n]$, define the Lagrangian as

$$\mathcal{L}_{\boldsymbol{\delta}}^{(i)}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}, \boldsymbol{\lambda}) \;=\; \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\delta})}[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i})] + \left\langle \boldsymbol{\lambda},\, \boldsymbol{g}_i(\boldsymbol{\theta}_i) + \sum_{j \neq i} \boldsymbol{g}_j(\boldsymbol{\theta}_j) \right\rangle,$$

where $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ is the dual variable. Denote by $\nabla \boldsymbol{g}_i(\cdot)$ the Jacobian matrix of $\boldsymbol{g}_i(\cdot)$. From primal-dual theory Boyd & Vandenberghe (2004); Facchinei & Pang (2010), let $\boldsymbol{\lambda}^{\mathrm{PSE}}$ denote the common Lagrange multiplier of the system (2), which exists and satisfies $\boldsymbol{\lambda}^{\mathrm{PSE}} \in \boldsymbol{\Lambda}$ under Assumption 8. Then, for any $\gamma > 0$,

$$\begin{aligned}
\boldsymbol{\theta}_i^{\mathrm{PSE}} &= P_{\boldsymbol{\Omega}_i}\Big[\boldsymbol{\theta}_i^{\mathrm{PSE}} - \gamma\Big(G_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top \boldsymbol{\lambda}^{\mathrm{PSE}}\Big)\Big], \quad \forall i \in [n], \\
\boldsymbol{\lambda}^{\mathrm{PSE}} &= \left[\boldsymbol{\lambda}^{\mathrm{PSE}} + \gamma\left(\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}}) + \sum_{j \neq i} \boldsymbol{g}_j(\boldsymbol{\theta}_j^{\mathrm{PSE}})\right)\right]_+,
\end{aligned}$$

where γ is a control parameter. Thus, given $\boldsymbol{\theta}_{-i}^{\mathrm{PSE}}$ and under $\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta}^{\mathrm{PSE}})$, $(\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}})$ is a saddle point of $\mathcal{L}_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{\mathrm{PSE}}, \boldsymbol{\lambda})$ for every $i \in [n]$. The joint saddle point $(\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}})$ achieves the PSE of the game (1) under strong duality Boyd & Vandenberghe (2004).

Thus, for every $t \in [T]$ and $i \in [n]$, define

$$\widehat{\phi}_i^t\big(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^t, \boldsymbol{\psi}^t, \boldsymbol{\lambda}\big) := \widehat{\nabla}_i J_i\big(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t\big) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}, \tag{15}$$

where $\widehat{\nabla}_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t)$ is estimator (5). Let $\widehat{\boldsymbol{\phi}}^t(\cdot) := \big[\widehat{\phi}_1^t(\cdot), \ldots, \widehat{\phi}_n^t(\cdot)\big]^\top$.

We also define

$$\Phi_{\boldsymbol{\delta}}^{i,t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}) := G_{\boldsymbol{\delta}}^{(i)}(\boldsymbol{\theta}^t) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}, \qquad G_{\boldsymbol{\delta}}^{(i)}(\boldsymbol{\theta}^t) := \mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\delta})}\big[\nabla_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}^t)\big], \tag{16}$$

$$\widehat{\Phi}_{\boldsymbol{\delta}}^{i,t}(\boldsymbol{\theta}_i^t, \boldsymbol{\psi}^t, \boldsymbol{\lambda}) := \widehat{G}_{\boldsymbol{\delta}}^{(i)}(\boldsymbol{\theta}_i^t, \boldsymbol{\psi}^t) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}, \qquad \widehat{G}_{\boldsymbol{\delta}}^{(i)}(\boldsymbol{\theta}_i^t, \boldsymbol{\psi}^t) := \mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\delta})}\big[\widehat{\nabla}_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t, \boldsymbol{\theta}_i^t)\big]. \tag{17}$$

Moreover, define $\Phi_{\boldsymbol{\delta}}^t(\boldsymbol{\theta}, \boldsymbol{\lambda}) := \big[\Phi_{\boldsymbol{\delta}}^{1,t}(\boldsymbol{\theta}, \boldsymbol{\lambda}), \ldots, \Phi_{\boldsymbol{\delta}}^{n,t}(\boldsymbol{\theta}, \boldsymbol{\lambda})\big]^\top$, and define $\widehat{\Phi}_{\boldsymbol{\delta}}^t(\boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{\lambda})$ similarly.

We start from the primal update and use the nonexpansiveness of Euclidean projection. Together with the saddle point $\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}}$, we write

$$\begin{aligned}
\mathbb{E}\|\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 = \sum_{i=1}^n \mathbb{E}\Bigg\| & P_{\Omega_i}\Big[\boldsymbol{\theta}_i^t - \gamma_t(\sigma_t)^2 \widehat{\phi}_i^t(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^t, \boldsymbol{\psi}^t, \boldsymbol{\lambda}_i^t)\Big] \\
& - P_{\Omega_i}\Big[\boldsymbol{\theta}_i^{\mathrm{PSE}} - \gamma_t(\sigma_t)^2 \Phi_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{i,t}(\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}})\Big]\Bigg\|_2^2 \\
\leq\; & \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 + \underbrace{\gamma_t^2(\sigma_t)^4 \sum_{i=1}^n \mathbb{E}\Big\|\widehat{\phi}_i^t(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^t, \boldsymbol{\psi}^t, \boldsymbol{\lambda}_i^t) - \Phi_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{i,t}(\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}})\Big\|_2^2}_{:=(\mathrm{I})} \\
& \underbrace{- 2\gamma_t(\sigma_t)^2 \sum_{i=1}^n \mathbb{E}\Big\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{\phi}_i^t(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^t, \boldsymbol{\psi}^t, \boldsymbol{\lambda}_i^t) - \Phi_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{i,t}(\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}})\Big\rangle}_{:=(\mathrm{II})}.
\end{aligned} \tag{18}$$

We now bound (I) and (II) separately.

Decomposition of (I)**.** We first add and subtract $\Phi^{i,t}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}^t_i)$ inside the norm. This gives

$$\begin{aligned}(\mathrm{I}) &= \gamma_t^2(\sigma_t)^4 \sum_{i=1}^n \mathbb{E}\left\| \widehat{\phi}^t_i(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}^t_i, \boldsymbol{\psi}^t, \boldsymbol{\lambda}^t_i) - \Phi^{i,t}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}}) \right\|_2^2 \qquad (19)\\ &= \gamma_t^2(\sigma_t)^4 \sum_{i=1}^n \mathbb{E}\Big\| \widehat{\phi}^t_i(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}^t_i, \boldsymbol{\psi}^t, \boldsymbol{\lambda}^t_i) - \Phi^{i,t}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}^t_i) \\ &\qquad + \Phi^{i,t}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}^t_i) - \Phi^{i,t}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}}, \boldsymbol{\lambda}^{\mathrm{PSE}}) \Big\|_2^2. \qquad (20)\end{aligned}$$

Using the Cauchy-Schwarz inequality, we can further bound

$$\begin{aligned}(\mathrm{I}) \le\ & \underbrace{3\gamma_t^2(\sigma_t)^4 \sum_{i=1}^n \mathbb{E}\left\| \widehat{\nabla}_i J_i(\boldsymbol{\xi}^t_i; \boldsymbol{\psi}^t, \boldsymbol{\theta}^t_i) - G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t) \right\|_2^2}_{:=(\mathrm{a})} \\ & + \underbrace{3\gamma_t^2(\sigma_t)^4 \sum_{i=1}^n \mathbb{E}\left\| G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t) - G^{(i)}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}}) \right\|_2^2}_{:=(\mathrm{b})} \\ & + \underbrace{3\gamma_t^2(\sigma_t)^4 \sum_{i=1}^n \mathbb{E}\left\| \nabla \boldsymbol{g}_i(\boldsymbol{\theta}^t_i)^\top \boldsymbol{\lambda}^t_i - \nabla \boldsymbol{g}_i(\boldsymbol{\theta}^{\mathrm{PSE}}_i)^\top \boldsymbol{\lambda}^{\mathrm{PSE}} \right\|_2^2}_{:=(\mathrm{c})}. \qquad (21)\end{aligned}$$

We first bound (a) using techniques for payoff-based learning with perturbations. For each $i \in [n]$, let the smoothed *cost* be

$$\widetilde{J}_i(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}) := \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \left[J_i\big(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}_i + \sigma_t \boldsymbol{v}_i, (\boldsymbol{\theta}_j + \sigma_t \boldsymbol{z}_j)_{j \neq i}\big) \right]. \qquad (22)$$

We first verify the smoothing identity, Lemma 2. Holding $\boldsymbol{z}_{-i}$ fixed and using the divergence theorem,

$$\begin{aligned}\nabla_i \widetilde{J}_i(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}) &= \frac{1}{\mathrm{vol}(\mathbb{B}^d)} \mathbb{E}_{\boldsymbol{z}_{-i}} \left[\int_{\mathbb{B}^d} \nabla_i J_i\big(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}_i + \sigma_t \boldsymbol{v}_i, \boldsymbol{\theta}_{-i} + \sigma_t \boldsymbol{z}_{-i}\big)\, d\boldsymbol{v}_i \right] \\ &= \frac{1}{\sigma_t\, \mathrm{vol}(\mathbb{B}^d)} \mathbb{E}_{\boldsymbol{z}_{-i}} \left[\int_{\mathbb{S}^{d-1}} J_i\big(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}_i + \sigma_t \boldsymbol{z}_i, \boldsymbol{\theta}_{-i} + \sigma_t \boldsymbol{z}_{-i}\big)\, \boldsymbol{z}_i\, dS(\boldsymbol{z}_i) \right] \\ &= \frac{d}{\sigma_t} \mathbb{E}_{\boldsymbol{z}} \left[J_i(\boldsymbol{\xi}^t_i; \boldsymbol{\theta} + \sigma_t \boldsymbol{z}) \boldsymbol{z}_i \right],\end{aligned}$$

where the last equality uses $\mathrm{area}(\mathbb{S}^{d-1}) = d\ \mathrm{vol}(\mathbb{B}^d)$.

For SPADE, $\boldsymbol{\xi}^t_i \sim \mathcal{D}_i(\boldsymbol{\theta}^t)$ is sampled independently of the perturbation $\boldsymbol{z}^t$. Define

$$\widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma}) := \mathbb{E}_{\boldsymbol{\xi}^t_i \sim \mathcal{D}_i(\boldsymbol{\theta}^t)} \left[\nabla_i \widetilde{J}_i(\boldsymbol{\xi}^t_i; \boldsymbol{\theta}^{t,\sigma}) \right].$$

Then, by the tower rule and the smoothing identity,

$$\mathbb{E}\left[\widehat{\nabla}_i J^t_i \,\middle|\, \boldsymbol{\theta}^t \right] = \mathbb{E}_{\boldsymbol{\xi}^t_i, \boldsymbol{z}^t} \left[\frac{d}{\sigma_t} J_i(\boldsymbol{\xi}^t_i; \boldsymbol{\psi}^t) \boldsymbol{z}^t_i \,\middle|\, \boldsymbol{\theta}^t \right] = \widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma}). \qquad (23)$$

Thus, write

$$\begin{aligned}3\gamma_t^2\sigma_t^4 \sum_{i=1}^n \mathbb{E}\left\| \widehat{\nabla}_i J^t_i - G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t) \right\|_2^2 \le\ & 6\gamma_t^2\sigma_t^4 \sum_{i=1}^n \mathbb{E}\left\| \widehat{\nabla}_i J^t_i - \widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma}) \right\|_2^2 \\ & + 6\gamma_t^2\sigma_t^4 \sum_{i=1}^n \mathbb{E}\left\| \widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma}) - G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t) \right\|_2^2. \qquad (24)\end{aligned}$$

Let $R_i := \widehat{\nabla}_i J_i^t - \widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma})$ and $Q_i := \widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma}) - G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t)$.

We begin with Q_i. By definition,

$$\widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma}) = \mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t)} \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \left[\nabla_i J_i\big(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{v}_i, (\boldsymbol{\theta}_j^{t,\sigma} + \sigma_t \boldsymbol{z}_j)_{j \neq i}\big) \right].$$

Thus, since SPADE samples $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t)$ and by the Lipschitz gradients (Assumption 4),

$$\begin{aligned}
\|Q_i\|_2 &= \left\| \mathbb{E}_{\boldsymbol{\xi}_i^t} \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \left[\nabla_i J_i\big(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{v}_i, (\boldsymbol{\theta}_j^{t,\sigma} + \sigma_t \boldsymbol{z}_j)_{j \neq i}\big) - \nabla_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}^t) \right] \right\|_2 \\
&\leq \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \left[\left\| \nabla_i J_i\big(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{v}_i, (\boldsymbol{\theta}_j^{t,\sigma} + \sigma_t \boldsymbol{z}_j)_{j \neq i}\big) - \nabla_i J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\theta}^t) \right\|_2 \right] \\
&\leq L_i \, \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \left[\left\| \boldsymbol{\theta}^{t,\sigma} + \sigma_t (\boldsymbol{v}_i, \boldsymbol{z}_{-i}) - \boldsymbol{\theta}^t \right\|_2 \right] \\
&\leq L_i \left(\|\boldsymbol{\theta}^{t,\sigma} - \boldsymbol{\theta}^t\|_2 + \sigma_t \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \|(\boldsymbol{v}_i, \boldsymbol{z}_{-i})\|_2 \right).
\end{aligned}$$

Since $\|\boldsymbol{v}_i\|_2 \leq 1$ and $\|\boldsymbol{z}_j\|_2 = 1$ for every $j \neq i$,

$$\|(\boldsymbol{v}_i, \boldsymbol{z}_{-i})\|_2^2 = \|\boldsymbol{v}_i\|_2^2 + \sum_{j \neq i} \|\boldsymbol{z}_j\|_2^2 \leq n,$$

and so $\mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}} \|(\boldsymbol{v}_i, \boldsymbol{z}_{-i})\|_2 \leq \sqrt{n}$. Thus, we can write

$$\begin{aligned}
\|\boldsymbol{\theta}^{t,\sigma} - \boldsymbol{\theta}^t\|_2^2 &= \sum_{j=1}^{n} \left(\frac{\sigma_t}{R_j} \right)^2 \|\boldsymbol{p}_j - \boldsymbol{\theta}_j^t\|_2^2 \\
&\leq \frac{4nC^2\sigma_t^2}{R_{\min}^2},
\end{aligned}$$

and so $\|\boldsymbol{\theta}^{t,\sigma} - \boldsymbol{\theta}^t\|_2 \leq \frac{2C\sqrt{n}}{R_{\min}} \sigma_t$. Combining the above bounds gives

$$\|Q_i\|_2 \leq L_i \sigma_t \sqrt{n} \left(1 + \frac{2C}{R_{\min}} \right).$$

We next bound R_i. Since $\widetilde{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{t,\sigma})$ is the conditional mean of the payoff estimator, $\mathbb{E}\|X - \mathbb{E}X\|_2^2 \leq \mathbb{E}\|X\|_2^2$ gives

$$\mathbb{E}\left[\|R_i\|_2^2 \,\middle|\, \boldsymbol{\theta}^t = \boldsymbol{\theta} \right] \leq \mathbb{E}\left[\left\| \widehat{\nabla}_i J_i^t \right\|_2^2 \,\middle|\, \boldsymbol{\theta}^t = \boldsymbol{\theta} \right] = \frac{d^2}{\sigma_t^2} \mathbb{E}\left[J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)^2 \,\middle|\, \boldsymbol{\theta}^t = \boldsymbol{\theta} \right], \tag{25}$$

where we used $\|\boldsymbol{z}_i^t\|_2 = 1$.

Since both $\boldsymbol{\psi}^t$ and the reference profile $\bar{\boldsymbol{\theta}}$ belong to $\boldsymbol{\Omega}$, Lipschitz continuity of the cost on $\boldsymbol{\Omega}$ (Assumption 5) and Assumption 7 give

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta})} \left[J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)^2 \right] &\leq 2\mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta})} \left[J_i(\boldsymbol{\xi}_i^t; \bar{\boldsymbol{\theta}})^2 \right] + 2G_i^2 \|\boldsymbol{\psi}^t - \bar{\boldsymbol{\theta}}\|_2^2 \\
&\leq 2V_i + 2G_i^2 \|\boldsymbol{\psi}^t - \bar{\boldsymbol{\theta}}\|_2^2.
\end{aligned} \tag{26}$$

By Assumption 3, $\|\boldsymbol{\psi}^t\|_2, \|\bar{\boldsymbol{\theta}}\|_2 \leq \sqrt{n}C$, and so $\|\boldsymbol{\psi}^t - \bar{\boldsymbol{\theta}}\|_2^2 \leq 4nC^2$. Thus, estimator variance is bounded by

$$\mathbb{E}\left[\|R_i\|_2^2 \,\middle|\, \boldsymbol{\theta}^t = \boldsymbol{\theta} \right] \leq \frac{d^2}{\sigma_t^2} \left(2V_i + 8nG_i^2 C^2 \right). \tag{27}$$

Plugging the bounds on R_i and Q_i into (24) then gives

$$\begin{aligned}
\text{(a)} &\leq 6\gamma_t^2 \sigma_t^4 \sum_{i=1}^{n} \left[\frac{d^2}{\sigma_t^2} \left(2V_i + 8nG_i^2 C^2 \right) + L_i^2 \sigma_t^2 n \left(1 + \frac{2C}{R_{\min}} \right)^2 \right] \\
&\leq \left(12nd^2 V_{\max} + 48n^2 d^2 G_{\max}^2 C^2 \right) \gamma_t^2 \sigma_t^2 + 6n^2 L_{\max}^2 \left(1 + \frac{2C}{R_{\min}} \right)^2 \gamma_t^2 \sigma_t^6.
\end{aligned} \tag{28}$$

Thus, for fixed problem dimensions, the leading contribution is

$$\mathcal{O}\big(\big(nd^2V_{\max} + n^2d^2G_{\max}^2C^2\big)\,\gamma_t^2\sigma_t^2\big)\,,$$

while the smoothing and inward-shift contribution is of higher order $\mathcal{O}(\gamma_t^2\sigma_t^6)$.

For (b), we add and subtract $G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{\mathrm{PSE}})$:

$$\begin{aligned}
(\mathrm{b}) &= 3\gamma_t^2(\sigma_t)^4\sum_{i=1}^n \mathbb{E}\left\|G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t) - G^{(i)}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\|_2^2 && (29)\\
&= 3\gamma_t^2(\sigma_t)^4\sum_{i=1}^n \mathbb{E}\left\|G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t) - G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{\mathrm{PSE}}) + G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{\mathrm{PSE}}) - G^{(i)}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\|_2^2 && (30)\\
&\le 6\gamma_t^2(\sigma_t)^4\sum_{i=1}^n \mathbb{E}\left\|G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t) - G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\|_2^2\\
&\quad + 6\gamma_t^2(\sigma_t)^4\sum_{i=1}^n \mathbb{E}\left\|G^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^{\mathrm{PSE}}) - G^{(i)}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\|_2^2. && (31)
\end{aligned}$$

By Assumptions 4 and 2,

$$(\mathrm{b}) \le 6\gamma_t^2(\sigma_t)^4\sum_{i=1}^n L_i^2\left(1+\varepsilon_i^2\max_{j\in[n]}p_{ij}\right)\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2. \tag{32}$$

For (c), because each $\boldsymbol{g}_i(\cdot)$ is G_g-Lipschitz,

$$\begin{aligned}
(\mathrm{c}) &\le 6\gamma_t^2(\sigma_t)^4\sum_{i=1}^n \mathbb{E}\left\|\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top\boldsymbol{\lambda}_i^t\right\|_2^2 + 6\gamma_t^2(\sigma_t)^4\sum_{i=1}^n\left\|\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top\boldsymbol{\lambda}^{\mathrm{PSE}}\right\|_2^2 && (33)\\
&\le 6\gamma_t^2(\sigma_t)^4G_g^2\sum_{i=1}^n\mathbb{E}\|\boldsymbol{\lambda}_i^t\|_2^2 + 6\gamma_t^2(\sigma_t)^4nG_g^2\|\boldsymbol{\lambda}^{\mathrm{PSE}}\|_2^2 && (34)\\
&\le 12\gamma_t^2(\sigma_t)^4nG_g^2\Lambda_{\max}^2, && (35)
\end{aligned}$$

where the last step uses Lemma 4 and $\boldsymbol{\lambda}^{\mathrm{PSE}} \in \boldsymbol{\Lambda}$.

Substituting the bounds on (a), (b), and (c) into (21), and using $0 < \sigma_t \le 1$, we obtain

$$\begin{aligned}
(\mathrm{I}) &\le 6\gamma_t^2\sigma_t^4\sum_{i=1}^n L_i^2\left(1+\varepsilon_i^2\max_{j\in[n]}p_{ij}\right)\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2\\
&\quad + \left[12nd^2V_{\max} + 48n^2d^2G_{\max}^2C^2 + 12nG_g^2\Lambda_{\max}^2 + 6n^2L_{\max}^2\left(1+\frac{2C}{R_{\min}}\right)^2\right]\gamma_t^2\sigma_t^2. && (36)
\end{aligned}$$

Decomposition of (II)**.** We now turn to (II). By definition,

$$\begin{aligned}
-(\mathrm{II}) &= -2\gamma_t(\sigma_t)^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{\phi}_i^t(\boldsymbol{\xi}_i^t;\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t,\boldsymbol{\lambda}_i^t) - \Phi^{i,t}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}},\boldsymbol{\lambda}^{\mathrm{PSE}})\right\rangle && (37)\\
&= \underbrace{-2\gamma_t(\sigma_t)^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{G}^{(i)}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t) - G^{(i)}_{\boldsymbol{\theta}^{\mathrm{PSE}}}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle}_{:=(\mathrm{i})}\\
&\quad \underbrace{-2\gamma_t(\sigma_t)^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top\boldsymbol{\lambda}_i^t - \nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top\boldsymbol{\lambda}^{\mathrm{PSE}}\right\rangle}_{:=(\mathrm{ii})}, && (38)
\end{aligned}$$

since by the tower rule and $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\theta}^t)$,

$$\mathbb{E}\left[\widehat{\phi}_i^t(\boldsymbol{\xi}_i^t;\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t,\boldsymbol{\lambda}_i^t) - \Phi_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{i,t}(\boldsymbol{\theta}^{\mathrm{PSE}},\boldsymbol{\lambda}^{\mathrm{PSE}})\right]$$
$$= \mathbb{E}\left[\mathbb{E}_{\boldsymbol{\xi}_i^t}\left[\widehat{\phi}_i^t(\boldsymbol{\xi}_i^t;\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t,\boldsymbol{\lambda}_i^t) - \Phi_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{i,t}(\boldsymbol{\theta}^{\mathrm{PSE}},\boldsymbol{\lambda}^{\mathrm{PSE}}) \,\middle|\, \boldsymbol{\theta}^t,\boldsymbol{\psi}^t,\boldsymbol{\lambda}^t\right]\right] \tag{39}$$
$$= \mathbb{E}\left[\widehat{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t) - G_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right] + \mathbb{E}\left[\nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t - \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top \boldsymbol{\lambda}^{\mathrm{PSE}}\right]. \tag{40}$$

We now bound (i). Add and subtract $G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t)$ and $G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})$:

$$\text{(i)} = \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t) - G_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle \tag{41}$$
$$= \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t)\right\rangle$$
$$+ \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle$$
$$+ \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}}) - G_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle. \tag{42}$$

Therefore,

$$-2\gamma_t\sigma_t^2\text{(i)} = -2\gamma_t\sigma_t^2 \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t)\right\rangle$$
$$- 2\gamma_t\sigma_t^2 \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle$$
$$- 2\gamma_t\sigma_t^2 \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}}) - G_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle. \tag{43}$$

We bound the first term on the right-hand side by Cauchy–Schwarz, applied only to the bias of the estimator. By the tower property,

$$\mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t)\right\rangle = \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widetilde{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{t,\sigma}) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t)\right\rangle,$$

since $\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}$ is $\boldsymbol{\theta}^t$-measurable. Using bound on the Q_i-term in (a) in Term I, we see

$$\left\|\widetilde{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{t,\sigma}) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t)\right\|_2 \le L_i\sigma_t\sqrt{n}\left(1 + \frac{2C}{R_{\min}}\right).$$

Using $\|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}\|_2 \le 2C$ then gives

$$-2\gamma_t\sigma_t^2 \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, \widehat{G}_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}_i^t,\boldsymbol{\psi}^t) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t)\right\rangle \le 4C\gamma_t\sigma_t^3\sqrt{n}\left(1 + \frac{2C}{R_{\min}}\right)\sum_{i=1}^n L_i$$
$$\le 4Cn^{3/2}L_{\max}\left(1 + \frac{2C}{R_{\min}}\right)\gamma_t\sigma_t^3. \tag{44}$$

For the middle term in (43), strong monotonicity gives

$$-2\gamma_t\sigma_t^2 \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^t) - G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle \le -2\mu\gamma_t\sigma_t^2\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2. \tag{45}$$

For the last term in (43), Assumption 2 gives

$$- 2\gamma_t\sigma_t^2 \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}, G_{\boldsymbol{\theta}^t}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}}) - G_{\boldsymbol{\theta}^{\mathrm{PSE}}}^{(i)}(\boldsymbol{\theta}^{\mathrm{PSE}})\right\rangle$$
$$\le 2\gamma_t\sigma_t^2 \sum_{i=1}^n L_i\varepsilon_i \max_{j\in[n]}\sqrt{p_{ij}}\,\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2. \tag{46}$$

Combining the above bounds yields

$$-2\gamma_t\sigma_t^2(\mathrm{i}) \le 4Cn^{3/2}L_{\max}\left(1+\frac{2C}{R_{\min}}\right)\gamma_t\sigma_t^3 + 2\gamma_t\sigma_t^2\left(-\mu+\sum_{i=1}^n L_i\varepsilon_i\max_{j\in[n]}\sqrt{p_{ij}}\right)\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2. \tag{47}$$

We next bound (ii). We have

$$-2\gamma_t\sigma_t^2(\mathrm{ii}) = -2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{PSE}},\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top\boldsymbol{\lambda}_i^t-\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top\boldsymbol{\lambda}^{\mathrm{PSE}}\right\rangle \tag{48}$$

$$\le 2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^{\mathrm{PSE}}-\boldsymbol{\theta}_i^t,\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top\boldsymbol{\lambda}_i^t\right\rangle+2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{PSE}},\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top\boldsymbol{\lambda}^{\mathrm{PSE}}\right\rangle \tag{49}$$

$$\le 2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left[\left(\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})-\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right)^\top\boldsymbol{\lambda}_i^t\right]+2\gamma_t\sigma_t^2\,\mathbb{E}\left\langle\boldsymbol{\lambda}^{\mathrm{PSE}},\sum_{i=1}^n\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle \tag{50}$$

$$= 2\gamma_t\sigma_t^2\mathbb{E}\left[\sum_{i=1}^n\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})^\top(\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t)+\boldsymbol{g}(\boldsymbol{\theta}^{\mathrm{PSE}})^\top\overline{\boldsymbol{\lambda}}^t+\Delta_t\right], \tag{51}$$

where $\overline{\boldsymbol{\lambda}}^t := \frac{1}{n}\sum_{i=1}^n\boldsymbol{\lambda}_i^t$ and $\Delta_t := \sum_{i=1}^n\left\langle\boldsymbol{\lambda}^{\mathrm{PSE}}-\boldsymbol{\lambda}_i^t,\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle$. The second inequality applies convexity of each component of $\boldsymbol{g}_i$ twice, at $\boldsymbol{\theta}_i^t$ and at $\boldsymbol{\theta}_i^{\mathrm{PSE}}$, together with $\boldsymbol{\lambda}_i^t\ge\mathbf{0}$ (Lemma 4) and $\boldsymbol{\lambda}^{\mathrm{PSE}}\ge\mathbf{0}$, and then uses complementary slackness $\left\langle\boldsymbol{\lambda}^{\mathrm{PSE}},\sum_i\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})\right\rangle=0$.

Furthermore, since $\boldsymbol{g}(\boldsymbol{\theta}^{\mathrm{PSE}})^\top\overline{\boldsymbol{\lambda}}^t\le 0$, we obtain

$$-2\gamma_t\sigma_t^2(\mathrm{ii}) \le 2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left[\|\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}})\|_2\|\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t\|_2\right]+2\gamma_t\sigma_t^2\mathbb{E}[\Delta_t]. \tag{52}$$

Now define

$$\widetilde{\mu} := \mu-\sum_{i=1}^n L_i\varepsilon_i\max_{j\in[n]}\sqrt{p_{ij}}, \tag{53}$$

$$\nu := 6\sum_{i=1}^n L_i^2\left(1+\varepsilon_i^2\max_{j\in[n]}p_{ij}\right), \tag{54}$$

$$\pi := 12nd^2V_{\max}+48n^2d^2G_{\max}^2C^2+12nG_g^2\Lambda_{\max}^2+6n^2L_{\max}^2\left(1+\frac{2C}{R_{\min}}\right)^2, \tag{55}$$

$$\rho := 4Cn^{3/2}L_{\max}\left(1+\frac{2C}{R_{\min}}\right). \tag{56}$$

Substituting (36), (47), and (52) into (18), we obtain

$$\mathbb{E}\|\boldsymbol{\theta}^{t+1}-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 \le \left(1-2\gamma_t\sigma_t^2\widetilde{\mu}+\gamma_t^2\sigma_t^4\nu\right)\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 + 2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left[B\|\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t\|_2\right]+2\gamma_t\sigma_t^2\mathbb{E}[\Delta_t] + \pi\gamma_t^2\sigma_t^2+\rho\gamma_t\sigma_t^3. \tag{57}$$

By assumption, we have $\sup_{t\geq 1}\gamma_t\sigma_t^2 \leq \frac{\widetilde{\mu}}{\nu}$. Then $1-2\widetilde{\mu}\gamma_t\sigma_t^2+\nu\gamma_t^2\sigma_t^4 \leq 1-\widetilde{\mu}\gamma_t\sigma_t^2$. Thus, (57) implies

$$\begin{aligned}\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 \leq \frac{1}{\widetilde{\mu}\gamma_t\sigma_t^2}\left(\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2-\mathbb{E}\|\boldsymbol{\theta}^{t+1}-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2\right)\\ +\frac{2B}{\widetilde{\mu}}\sum_{i=1}^n\mathbb{E}\left[\|\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t\|_2\right]+\frac{2}{\widetilde{\mu}}\mathbb{E}[\Delta_t]\\ +\frac{\pi\gamma_t}{\widetilde{\mu}}+\frac{\rho}{\widetilde{\mu}}\sigma_t. \qquad (58)\end{aligned}$$

Since $\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 \leq 4nC^2$, we can telescope the first term. Indeed,

$$\begin{aligned}&\sum_{t=1}^T\frac{1}{\gamma_t\sigma_t^2}\left(\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2-\mathbb{E}\|\boldsymbol{\theta}^{t+1}-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2\right)\\ &=\frac{1}{\gamma_1\sigma_1^2}\mathbb{E}\|\boldsymbol{\theta}^1-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2-\frac{1}{\gamma_T\sigma_T^2}\mathbb{E}\|\boldsymbol{\theta}^{T+1}-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2\\ &\quad+\sum_{t=2}^T\left(\frac{1}{\gamma_t\sigma_t^2}-\frac{1}{\gamma_{t-1}\sigma_{t-1}^2}\right)\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 \qquad (59)\\ &\leq\frac{4nC^2}{\gamma_T\sigma_T^2}, \qquad (60)\end{aligned}$$

where the last inequality uses that $\gamma_t\sigma_t^2$ is nonincreasing.

Summing over $t=1,\ldots,T$, applying (60), and bounding the two dual terms by Lemma 6 with the admissible comparator $\boldsymbol{\lambda}=\boldsymbol{\lambda}^{\mathrm{PSE}}\in\boldsymbol{\Lambda}$ and by Lemma 5, namely

$$\sum_{t=1}^T\mathbb{E}[\Delta_t] \leq \frac{2n\Lambda_{\max}^2}{\gamma_T}+\left(\frac{nB^2}{2}+\frac{2nB^2}{1-\sigma_2(\boldsymbol{A})}\right)\sum_{t=1}^T\gamma_t,$$

and

$$\sum_{t=1}^T\sum_{i=1}^n\mathbb{E}\|\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t\|_2 \leq \frac{nB}{1-\sigma_2(\boldsymbol{A})}\sum_{t=1}^T\gamma_t,$$

and using $\sigma_T\leq 1$ so that $1/\gamma_T\leq 1/(\gamma_T\sigma_T^2)$, we obtain

$$\sum_{t=1}^T\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2 \leq \mathcal{O}\left(\frac{1}{\widetilde{\mu}\gamma_T\sigma_T^2}+\frac{1}{\widetilde{\mu}}\sum_{t=1}^T\gamma_t+\frac{1}{\widetilde{\mu}}\sum_{t=1}^T\sigma_t\right). \qquad (61)$$

We now derive the regret bound. By Assumption 5,

$$\mathbb{E}_{\boldsymbol{\xi}_i\sim\mathcal{D}_i(\boldsymbol{\theta}^{\mathrm{PSE}})}\left|J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta}_i^t,\boldsymbol{\theta}_{-i}^{\mathrm{PSE}})-J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta}^{\mathrm{PSE}})\right| \leq G_i\|\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{PSE}}\|_2.$$

Taking expectation over the algorithmic randomness and using the definition of $\mathcal{R}_i^{\mathrm{PSE}}(T)$, we obtain

$$\begin{aligned}\mathcal{R}_i^{\mathrm{PSE}}(T)&=\sum_{t=1}^T\mathbb{E}\left[J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta}_i^t,\boldsymbol{\theta}_{-i}^{\mathrm{PSE}})-J_i(\boldsymbol{\xi}_i;\boldsymbol{\theta}^{\mathrm{PSE}})\right] \qquad (62)\\ &\leq G_i\sum_{t=1}^T\mathbb{E}\|\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{PSE}}\|_2 \qquad (63)\\ &\leq G_i\sqrt{T\sum_{t=1}^T\mathbb{E}\|\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{PSE}}\|_2^2} \qquad (64)\\ &\leq \mathcal{O}\left(\sqrt{\frac{T}{\widetilde{\mu}}\left(\frac{1}{\gamma_T\sigma_T^2}+\sum_{t=1}^T\gamma_t+\sum_{t=1}^T\sigma_t\right)}\right), \qquad \forall i\in[n]. \qquad (65)\end{aligned}$$

Finally, we bound the cumulative constraint violation. Since $\sum_{i=1}^{n} \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}}) \leq \mathbf{0}$, we have pathwise

$$\left\| \left[\sum_{t=1}^{T} \sum_{i=1}^{n} \boldsymbol{g}_i(\boldsymbol{\theta}_i^t) \right]_+ \right\|_2 \leq \left\| \sum_{t=1}^{T} \sum_{i=1}^{n} \left(\boldsymbol{g}_i(\boldsymbol{\theta}_i^t) - \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{PSE}}) \right) \right\|_2$$

$$\leq G_g \sum_{t=1}^{T} \sum_{i=1}^{n} \|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{PSE}}\|_2 \tag{66}$$

$$\leq G_g \sqrt{n} \sum_{t=1}^{T} \|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2, \tag{67}$$

using the G_g-Lipschitzness of $\boldsymbol{g}_i$ (Assumption 6). Taking expectations, applying Cauchy–Schwarz and (61),

$$\mathcal{R}_g(T) \leq G_g \sqrt{nT \sum_{t=1}^{T} \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2^2} \leq \mathcal{O}\left(\sqrt{\frac{T}{\widetilde{\mu}} \left(\frac{1}{\gamma_T \sigma_T^2} + \sum_{t=1}^{T} \gamma_t + \sum_{t=1}^{T} \sigma_t \right)} \right). \tag{68}$$

From here, if we optimize $\gamma_t = \tau_1(\tau_2 t + \tau_1)^{-1/4}$ and $\sigma_t = \tau_3(\tau_4 t + \tau_3)^{-1/4}$ for some $\tau_1, \tau_2, \tau_3, \tau_4 > 0$ that satisfy the stepsize conditions, standard summation arguments give

$$\sum_{t=1}^{T} \gamma_t = \mathcal{O}(T^{3/4}), \quad \sum_{t=1}^{T} \sigma_t = \mathcal{O}(T^{3/4}), \quad \frac{1}{\gamma_T \sigma_T^2} = \mathcal{O}(T^{3/4})$$

and bounds $\mathcal{R}_i^{\mathrm{PSE}}(T) \leq \mathcal{O}(T^{7/8})$, $\forall i \in [n]$ and $\mathcal{R}_g(T) \leq \mathcal{O}(T^{7/8})$. This completes the proof. □

E CONVERGENCE OF SPADE-NE (ALGORITHM 2): PAYOFF-BASED STOCHASTIC PRIMAL-DUAL UPDATE FOR NE

We first present the full convergence guarantee statement of SPADE-NE.

Theorem 4 (Convergence of SPADE-NE, full version of Theorem 2) *Suppose Assumptions 1–9 hold. Assume that $\widehat{\mu} > 0$ as defined in Proposition 1 and the stepsizes are nonincreasing, positive, and satisfy $0 < \sigma_t \le \min\{1, R_{\min}\}$, $\sup_{t\ge 1} \gamma_t \sigma_t^2 \le \frac{\widehat{\mu}}{\widehat{\nu}}$, and $\sigma_t \to 0$, where $\widehat{\nu} := 6L_{\mathrm{PR}}^2 n$. Then for every $i \in [n]$,*

$$\mathcal{R}_i^{\mathrm{NE}}(T) \;\le\; \mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}}\left(\frac{1}{\gamma_T \sigma_T^2} + \sum_{t=1}^{T}\gamma_t + \sum_{t=1}^{T}\sigma_t\right)}\right),$$

$$R_g(T) \;\le\; \mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}}\left(\frac{1}{\gamma_T \sigma_T^2} + \sum_{t=1}^{T}\gamma_t + \sum_{t=1}^{T}\sigma_t\right)}\right).$$

Theorem 4 matches the rate of Theorem 1 up to the replacement $\widetilde{\mu} \mapsto \widehat{\mu}$. Choosing the same stepsizes $\gamma_t = \tau_1(\tau_2 t + \tau_1)^{-1/4}$ and $\sigma_t = \tau_3(\tau_4 t + \tau_3)^{-1/4}$ give final bounds $\mathcal{R}_i^{\mathrm{NE}}(T)$, $R_g(T) \le \mathcal{O}(T^{7/8})$.

We present the proof of Theorem 2. The argument closely follows the proof structure of Theorem 1, and we focus on the three places where the analysis distinctly differs from the PSE case:

- the gradient estimator 5 targets the gradient of the smoothed performative risk,
- the performative drift term vanishes, and
- the descent inequality inherits strong monotonicity from $\nabla\mathrm{PR}$ with modulus $\widehat{\mu}$ rather than from ∇J.

Proof: For each player $i \in [n]$, define the Lagrangian against the performative risk PR_i as

$$\mathcal{L}_{\mathrm{PR}}^{(i)}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}, \boldsymbol{\lambda}) = \mathrm{PR}_i(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}) + \left\langle \boldsymbol{\lambda}, \boldsymbol{g}_i(\boldsymbol{\theta}_i) + \sum_{j\ne i} \boldsymbol{g}_j(\boldsymbol{\theta}_j) \right\rangle,$$

where $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ is the dual variable. In contrast to the PSE Lagrangian $\mathcal{L}_{\boldsymbol{\delta}}^{(i)}$ in Appendix D, the primal objective here is the full performative risk $\mathrm{PR}_i(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\xi}_i \sim \mathcal{D}_i(\boldsymbol{\theta})}[J_i(\boldsymbol{\xi}_i; \boldsymbol{\theta})]$ and absorbs the distributional dependence on $\boldsymbol{\theta}$. Denote by $\nabla \boldsymbol{g}_i(\cdot)$ the Jacobian matrix of $\boldsymbol{g}_i(\cdot)$. From primal-dual theory Boyd & Vandenberghe (2004); Facchinei & Pang (2010), let $\boldsymbol{\lambda}^{\mathrm{NE}}$ denote the common Lagrange multiplier of the system (2), which exists and satisfies $\boldsymbol{\lambda}^{\mathrm{NE}} \in \boldsymbol{\Lambda}$ under Assumption 8. Then, for any $\gamma > 0$,

$$\boldsymbol{\theta}_i^{\mathrm{NE}} = P_{\boldsymbol{\Omega}_i}\left[\boldsymbol{\theta}_i^{\mathrm{NE}} - \gamma\left(\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top \boldsymbol{\lambda}^{\mathrm{NE}}\right)\right], \quad \forall i \in [n],$$

$$\boldsymbol{\lambda}^{\mathrm{NE}} = \left[\boldsymbol{\lambda}^{\mathrm{NE}} + \gamma\left(\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}}) + \sum_{j\ne i} \boldsymbol{g}_j(\boldsymbol{\theta}_j^{\mathrm{NE}})\right)\right]_+.$$

Thus, given $\boldsymbol{\theta}_{-i}^{\mathrm{NE}}$, $(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})$ is a saddle point of $\mathcal{L}_{\mathrm{PR}}^{(i)}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}^{\mathrm{NE}}, \boldsymbol{\lambda})$ for every $i \in [n]$. Existence of the saddle point relies on the strong monotonicity of $\nabla\mathrm{PR}$ (Proposition 1) and the Lipschitzness of $\nabla\mathrm{PR}_i$ (Assumption 9), which together with strong duality Boyd & Vandenberghe (2004) yield that the joint saddle point $(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})$ achieves the NE of the game.

Now, for every $t \in [T]$ and $i \in [n]$, define the PR-based target mapping

$$\Psi^{i,t}(\boldsymbol{\theta}, \boldsymbol{\lambda}) := \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i)^\top \boldsymbol{\lambda}.$$

By nonexpansiveness of Euclidean projection applied to the primal update (12), and using that $(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})$ is a saddle point, we obtain

$$\mathbb{E}\|\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 \le \mathbb{E}\|\boldsymbol{\theta}^{t} - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + \underbrace{\gamma_t^2(\sigma_t)^4 \sum_{i=1}^{n} \mathbb{E}\left\|\widehat{\phi}_i^t - \Psi^{i,t}(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})\right\|_2^2}_{:=(\mathrm{I}')}$$
$$\underbrace{- 2\gamma_t(\sigma_t)^2 \sum_{i=1}^{n} \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \widehat{\phi}_i^t - \Psi^{i,t}(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})\right\rangle}_{:=(\mathrm{II}')}, \tag{69}$$

where $\widehat{\phi}_i^t := \widehat{\nabla}_i J_i^t + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t$ is as defined in the PSE proof.

Bounding (I')**.** We add and subtract $\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t$ and $\Psi^{i,t}_{\boldsymbol{\theta}^t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}_i^t)$ inside the norm of (I'), then apply Cauchy–Schwarz to obtain

$$(\mathrm{I}') \le 3\gamma_t^2(\sigma_t)^4 \sum_{i=1}^{n} \underbrace{\mathbb{E}\left\|\widehat{\nabla}_i J_i^t - \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma})\right\|_2^2}_{(\mathrm{a}')\text{ smoothed-gradient variance}}$$
$$+ 3\gamma_t^2(\sigma_t)^4 \sum_{i=1}^{n} \underbrace{\mathbb{E}\left\|\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t)\right\|_2^2}_{(\mathrm{b}')\text{ smoothing bias}}$$
$$+ 3\gamma_t^2(\sigma_t)^4 \sum_{i=1}^{n} \underbrace{\mathbb{E}\left\|\Psi^{i,t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}_i^t) - \Psi^{i,t}(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})\right\|_2^2}_{(\mathrm{c}')\text{ primal–dual drift}}. \tag{70}$$

For (a'), we first verify the smoothing identity. Let the smoothed performative risk be,

$$\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}) := \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}}\left[\mathrm{PR}_i(\boldsymbol{\theta}_i + \sigma_t \boldsymbol{v}_i, (\boldsymbol{\theta}_j + \sigma_t \boldsymbol{z}_j)_{j\neq i})\right].$$

Thus, the key identity from Lemma 2 follows from noting that,

$$\mathbb{E}\left[\widehat{\nabla}_i J_i^t \,\middle|\, \boldsymbol{\theta}^t\right] = \frac{d}{\sigma_t}\mathbb{E}_{\boldsymbol{z}^t}\left[\mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\psi}^t)}\left[J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)\right] \boldsymbol{z}_i^t\right]$$
$$= \frac{d}{\sigma_t}\mathbb{E}_{\boldsymbol{z}^t}\left[\mathrm{PR}_i(\boldsymbol{\psi}^t)\boldsymbol{z}_i^t\right] = \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}). \tag{71}$$

Thus, deployed-action sampling $\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\psi}^t)$ makes the payoff estimator an unbiased estimator of the smoothed performative-risk gradient evaluated at the feasible inward-shifted center.

Thus, again using $\mathbb{E}\|X - \mathbb{E}X\|_2^2 \le \mathbb{E}\|X\|_2^2$ gives

$$(\mathrm{a}') \le \mathbb{E}\left[\left\|\widehat{\nabla}_i J_i^t\right\|_2^2 \,\middle|\, \boldsymbol{\theta}^t = \boldsymbol{\theta}\right] = \frac{d^2}{\sigma_t^2}\mathbb{E}\left[J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)^2 \,\middle|\, \boldsymbol{\theta}^t = \boldsymbol{\theta}\right],$$

where we used $\|\boldsymbol{z}_i^t\|_2 = 1$. Lipschitz continuity of the cost on $\boldsymbol{\Omega}$ and Assumption 7 give

$$\mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\psi}^t)}\left[J_i(\boldsymbol{\xi}_i^t; \boldsymbol{\psi}^t)^2\right] \le 2\mathbb{E}_{\boldsymbol{\xi}_i^t \sim \mathcal{D}_i(\boldsymbol{\psi}^t)}\left[J_i(\boldsymbol{\xi}_i^t; \bar{\boldsymbol{\theta}})^2\right] + 2G_i^2\|\boldsymbol{\psi}^t - \bar{\boldsymbol{\theta}}\|_2^2$$
$$\le 2V_i + 2G_i^2\|\boldsymbol{\psi}^t - \bar{\boldsymbol{\theta}}\|_2^2.$$

By Assumption 3, $\|\boldsymbol{\psi}^t\|_2, \|\bar{\boldsymbol{\theta}}\|_2 \le \sqrt{n}C$, so $\|\boldsymbol{\psi}^t - \bar{\boldsymbol{\theta}}\|_2^2 \le 4nC^2$. Thus,

$$(\mathrm{a}') \le \frac{d^2}{\sigma_t^2}\left(2V_i + 8nG_i^2C^2\right). \tag{72}$$

(b') is the smoothing bias. Differentiating the smoothing under the expectation gives

$$\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) = \mathbb{E}_{\boldsymbol{v}_i, \boldsymbol{z}_{-i}}\left[\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i\big(\boldsymbol{\theta}_i^{t,\sigma} + \sigma_t \boldsymbol{v}_i, (\boldsymbol{\theta}_j^{t,\sigma} + \sigma_t \boldsymbol{z}_j)_{j\neq i}\big)\right].$$

Thus, by Jensen's inequality and Assumption 9,

$$\begin{aligned}&\left\|\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t)\right\|_2 \\ &\le L_{\mathrm{PR}}\mathbb{E}_{\boldsymbol{v}_i,\boldsymbol{z}_{-i}}\left[\left\|\boldsymbol{\theta}^{t,\sigma} + \sigma_t(\boldsymbol{v}_i,\boldsymbol{z}_{-i}) - \boldsymbol{\theta}^t\right\|_2\right] \\ &\le L_{\mathrm{PR}}\left(\|\boldsymbol{\theta}^{t,\sigma} - \boldsymbol{\theta}^t\|_2 + \sigma_t\mathbb{E}_{\boldsymbol{v}_i,\boldsymbol{z}_{-i}}\|(\boldsymbol{v}_i,\boldsymbol{z}_{-i})\|_2\right).\end{aligned}$$

Then, noting that $\|\boldsymbol{\theta}^{t,\sigma} - \boldsymbol{\theta}^t\|_2 \le \frac{2C\sqrt{n}}{R_{\min}}\sigma_t$ and $\mathbb{E}\|(\boldsymbol{v}_i,\boldsymbol{z}_{-i})\|_2 \le \sqrt{n}$, we have that

$$\left\|\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t)\right\|_2 \le L_{\mathrm{PR}}\sigma_t\sqrt{n}\left(1 + \frac{2C}{R_{\min}}\right), \tag{73}$$

and so,

$$(\mathrm{b}') \le L_{\mathrm{PR}}^2\sigma_t^2 n\left(1 + \frac{2C}{R_{\min}}\right)^2. \tag{74}$$

Finally, (c') is the primal-dual drift, which we expand as,

$$\begin{aligned}(\mathrm{c}') \le\ & 2\,\mathbb{E}\left\|\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}})\right\|_2^2 \\ &+ 2\,\mathbb{E}\left\|\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top\boldsymbol{\lambda}_i^t - \nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top\boldsymbol{\lambda}^{\mathrm{NE}}\right\|_2^2.\end{aligned}$$

Applying Assumption 9 to the first term gives $\le 2L_{\mathrm{PR}}^2\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2$. The second term is treated as in the PSE proof: $\boldsymbol{g}_i$ is G_g-Lipschitz (Assumption 6) and the dual iterates $\boldsymbol{\lambda}_i^t$ together with $\boldsymbol{\lambda}^{\mathrm{NE}}$ are uniformly bounded by $\Lambda_{\max}$ via Lemma 4 and Assumption 8. Combining,

$$\begin{aligned}(\mathrm{c}') &\le 2L_{\mathrm{PR}}^2\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + 2G_g^2\mathbb{E}\|\boldsymbol{\lambda}_i^t\|_2^2 + 2G_g^2\|\boldsymbol{\lambda}^{\mathrm{NE}}\|_2^2 \\ &\le 2L_{\mathrm{PR}}^2\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + 4G_g^2\Lambda_{\max}^2.\end{aligned} \tag{75}$$

Substituting (72)–(75) into (70), we obtain

$$\begin{aligned}(\mathrm{I}') \le\ & 6L_{\mathrm{PR}}^2 n\gamma_t^2\sigma_t^4\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 \\ &+ \left(6nd^2V_{\max} + 24n^2d^2G_{\max}^2C^2\right)\gamma_t^2\sigma_t^2 \\ &+ 12nG_g^2\Lambda_{\max}^2\gamma_t^2\sigma_t^4 + 3n^2L_{\mathrm{PR}}^2\left(1 + \frac{2C}{R_{\min}}\right)^2\gamma_t^2\sigma_t^6.\end{aligned}$$

Finally, since $0 < \sigma_t \le 1$, suppressing the lower-order terms into t-independent constant $\widehat{\pi}$, we have

$$(\mathrm{I}') \le 6L_{\mathrm{PR}}^2 n\gamma_t^2\sigma_t^4\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + \widehat{\pi}\gamma_t^2\sigma_t^2. \tag{76}$$

Bounding (II'). Starting from the definition of (II'), and using (71) inside the conditional expectation together with the tower property, we can write

$$\begin{aligned}-(\mathrm{II}') = &-2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}})\right\rangle \\ &- 2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top\boldsymbol{\lambda}_i^t - \nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top\boldsymbol{\lambda}^{\mathrm{NE}}\right\rangle.\end{aligned}$$

For the first term, add and subtract $\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t)$. Using the smoothing bias bound (73) and the fact that $\|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}\|_2 \le 2C$ first gives

$$-2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t)\right\rangle \le 4Cn^{3/2}L_{\mathrm{PR}}\left(1 + \frac{2C}{R_{\min}}\right)\gamma_t\sigma_t^3.$$

Then, using $\widehat{\mu} > 0$ and strong monotonicity of $\nabla\mathrm{PR}$ from Proposition 1, the other piece is bounded as

$$-2\gamma_t\sigma_t^2\sum_{i=1}^n\mathbb{E}\left\langle\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}})\right\rangle \le -2\widehat{\mu}\gamma_t\sigma_t^2\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2.$$

For the constraint second term, this is handled as in the PSE proof. Convexity of each component of $\boldsymbol{g}_i$ applied at $\boldsymbol{\theta}_i^t$ and at $\boldsymbol{\theta}_i^{\mathrm{NE}}$, nonnegativity of $\boldsymbol{\lambda}_i^t$ and $\boldsymbol{\lambda}^{\mathrm{NE}}$, complementary slackness, and feasibility of $\boldsymbol{\theta}^{\mathrm{NE}}$ yield

$$-2\gamma_t\sigma_t^2\sum_{i=1}^{n}\mathbb{E}\left\langle\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{NE}},\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top\boldsymbol{\lambda}_i^t-\nabla\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top\boldsymbol{\lambda}^{\mathrm{NE}}\right\rangle\le 2\gamma_t\sigma_t^2\sum_{i=1}^{n}\mathbb{E}\left[B\|\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t\|_2\right]+2\gamma_t\sigma_t^2\mathbb{E}[\Delta_t],$$

where $\Delta_t:=\sum_{i=1}^{n}\left\langle\boldsymbol{\lambda}^{\mathrm{NE}}-\boldsymbol{\lambda}_i^t,\boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\right\rangle$.

Combining yields the bound,

$$\begin{aligned}-(\mathrm{II}')\le{}&-2\widehat{\mu}\gamma_t\sigma_t^2\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{NE}}\|_2^2+4Cn^{3/2}L_{\mathrm{PR}}\left(1+\frac{2C}{R_{\min}}\right)\gamma_t\sigma_t^3\\&+2\gamma_t\sigma_t^2\sum_{i=1}^{n}\mathbb{E}\left[B\|\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t\|_2\right]+2\gamma_t\sigma_t^2\mathbb{E}[\Delta_t].\end{aligned}\tag{77}$$

Now define the constants:

- $\widehat{\nu}:=6L_{\mathrm{PR}}^2n$,
- $\widehat{\rho}:=4Cn^{3/2}L_{\mathrm{PR}}\left(1+\frac{2C}{R_{\min}}\right)$.

Plugging (76) and (77) into (69) gives the descent recursion

$$\begin{aligned}\mathbb{E}\|\boldsymbol{\theta}^{t+1}-\boldsymbol{\theta}^{\mathrm{NE}}\|_2^2\le{}&\left(1-2\widehat{\mu}\gamma_t\sigma_t^2+\widehat{\nu}\gamma_t^2\sigma_t^4\right)\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{NE}}\|_2^2\\&+2\gamma_t\sigma_t^2\sum_{i=1}^{n}\mathbb{E}\left[B\|\boldsymbol{\lambda}_i^t-\overline{\boldsymbol{\lambda}}^t\|_2\right]+2\gamma_t\sigma_t^2\mathbb{E}[\Delta_t]\\&+\widehat{\pi}\gamma_t^2\sigma_t^2+\widehat{\rho}\gamma_t\sigma_t^3.\end{aligned}\tag{78}$$

Under the stepsize assumption $\sup_t\gamma_t\sigma_t^2\le\widehat{\mu}/\widehat{\nu}$, we have $1-2\widehat{\mu}\gamma_t\sigma_t^2+\widehat{\nu}\gamma_t^2\sigma_t^4\le 1-\widehat{\mu}\gamma_t\sigma_t^2$. Following the steps of the PSE proof, i.e. summing (78) from $t=1$ to T, telescoping as in (60), and applying Lemma 6 with the admissible comparator $\boldsymbol{\lambda}=\boldsymbol{\lambda}^{\mathrm{NE}}\in\boldsymbol{\Lambda}$ together with Lemma 5, yields

$$\sum_{t=1}^{T}\mathbb{E}\|\boldsymbol{\theta}^t-\boldsymbol{\theta}^{\mathrm{NE}}\|_2^2\le\mathcal{O}\left(\frac{1}{\widehat{\mu}\gamma_T\sigma_T^2}+\frac{1}{\widehat{\mu}}\sum_{t=1}^{T}\gamma_t+\frac{1}{\widehat{\mu}}\sum_{t=1}^{T}\sigma_t\right).\tag{79}$$

We then note that under Assumptions 3 and 9, the function PR_i is Lipschitz on $\boldsymbol{\Omega}$. Indeed, $\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i$ is continuous on the compact set $\boldsymbol{\Omega}$ (since it is L_{PR}-Lipschitz by Assumption 9), so let

$$G_i^{\mathrm{PR}}:=\sup_{\boldsymbol{\theta}\in\boldsymbol{\Omega}}\|\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta})\|_2<\infty.$$

Then, the mean-value inequality applied along the segment $\{(\tau\boldsymbol{\theta}_i^t+(1-\tau)\boldsymbol{\theta}_i^{\mathrm{NE}},\boldsymbol{\theta}_{-i}^{\mathrm{NE}}):\tau\in[0,1]\}\subset\boldsymbol{\Omega}$ (which is contained in $\boldsymbol{\Omega}$ by convexity, Assumption 3) gives

$$\left|\mathrm{PR}_i(\boldsymbol{\theta}_i^t,\boldsymbol{\theta}_{-i}^{\mathrm{NE}})-\mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}})\right|\le G_i^{\mathrm{PR}}\|\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{NE}}\|_2.$$

Therefore, by the definition of NE-performative regret and applying Cauchy–Schwarz and (79),

$$\begin{aligned}\mathcal{R}_i^{\mathrm{NE}}(T)&\le G_i^{\mathrm{PR}}\sum_{t=1}^{T}\mathbb{E}\|\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{NE}}\|_2\\&\le G_i^{\mathrm{PR}}\sqrt{T\sum_{t=1}^{T}\mathbb{E}\|\boldsymbol{\theta}_i^t-\boldsymbol{\theta}_i^{\mathrm{NE}}\|_2^2}\\&\le\mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}}\left(\frac{1}{\gamma_T\sigma_T^2}+\sum_{t=1}^{T}\gamma_t+\sum_{t=1}^{T}\sigma_t\right)}\right),\qquad\forall i\in[n].\end{aligned}$$

The constraint-violation bound is obtained by the same argument as in the PSE proof. Since $\sum_{i=1}^{n} \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}}) \leq \mathbf{0}$ and $\boldsymbol{g}_i$ is G_g-Lipschitz, we have pathwise

$$\left\| \left[\sum_{t=1}^{T} \sum_{i=1}^{n} \boldsymbol{g}_i(\boldsymbol{\theta}_i^t) \right]_+ \right\|_2 \leq \left\| \sum_{t=1}^{T} \sum_{i=1}^{n} \left(\boldsymbol{g}_i(\boldsymbol{\theta}_i^t) - \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}}) \right) \right\|_2 \leq G_g \sqrt{n} \sum_{t=1}^{T} \|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2.$$

Taking expectations, applying Cauchy–Schwarz and (79),

$$R_g(T) \leq \mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}} \left(\frac{1}{\gamma_T \sigma_T^2} + \sum_{t=1}^{T} \gamma_t + \sum_{t=1}^{T} \sigma_t \right)} \right),$$

which matches the statement of Theorem 4. This completes the proof. □

F CONVERGENCE OF SPADE-MP (ALGORITHM 3): MULTIPOINT PAYOFF-BASED UPDATE FOR NE

F.1 PROOF OF LEMMA 3

Proof: Conditioning first on $\boldsymbol{z}^t$ and using the respective marginals in Assumption 10, we have that

$$\mathbb{E}\big[Y_i^{t,\pm} \,\big|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\big] = \mathrm{PR}_i(\boldsymbol{\psi}^{t,\pm}),$$

where $\boldsymbol{\psi}_i^{t,\pm} = \boldsymbol{\theta}_i^{t,\sigma} \pm \sigma_t \boldsymbol{z}_i^t$. Thus, by the definition of the multipoint Estimator (9),

$$\mathbb{E}\Big[\widehat{\nabla}_i^{\mathrm{MP}} J_i \,\Big|\, \boldsymbol{\theta}^t\Big] = \frac{d}{2\sigma_t}\mathbb{E}_{\boldsymbol{z}^t}\left[\big(\mathrm{PR}_i(\boldsymbol{\psi}^{t,+}) - \mathrm{PR}_i(\boldsymbol{\psi}^{t,-})\big)\,\boldsymbol{z}_i^t\right].$$

Since $\boldsymbol{z}^t \overset{d}{=} -\boldsymbol{z}^t$, we have

$$\mathbb{E}_{\boldsymbol{z}^t}\left[\mathrm{PR}_i(\boldsymbol{\psi}^{t,-})\boldsymbol{z}_i^t\right] = -\mathbb{E}_{\boldsymbol{z}^t}\left[\mathrm{PR}_i(\boldsymbol{\psi}^{t,+})\boldsymbol{z}_i^t\right],$$

and so

$$\mathbb{E}\Big[\widehat{\nabla}_i^{\mathrm{MP}} J_i \,\Big|\, \boldsymbol{\theta}^t\Big] = \frac{d}{\sigma_t}\mathbb{E}_{\boldsymbol{z}^t}\left[\mathrm{PR}_i(\boldsymbol{\psi}^{t,+})\boldsymbol{z}_i^t\right] = \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}),$$

where the final equality is the smoothing identity of Lemma 2.

For the second-moment bound, we first control payoff difference. Since $\boldsymbol{\psi}^{t,+}, \boldsymbol{\psi}^{t,-} \in \boldsymbol{\Omega}$, we have

$$|Y_i^{t,+} - Y_i^{t,-}| = \big|J_i(\boldsymbol{\xi}_i^{t,+}; \boldsymbol{\psi}^{t,+}) - J_i(\boldsymbol{\xi}_i^{t,-}; \boldsymbol{\psi}^{t,-})\big| \le G_i\left(\|\boldsymbol{\xi}_i^{t,+} - \boldsymbol{\xi}_i^{t,-}\|_2 + \|\boldsymbol{\psi}^{t,+} - \boldsymbol{\psi}^{t,-}\|_2\right).$$

Squaring and applying Cauchy-Schwarz gives

$$\mathbb{E}\big[|Y_i^{t,+} - Y_i^{t,-}|^2 \,\big|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\big] \le 2G_i^2\mathbb{E}\big[\|\boldsymbol{\xi}_i^{t,+} - \boldsymbol{\xi}_i^{t,-}\|_2^2 \,\big|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\big] + 2G_i^2\|\boldsymbol{\psi}^{t,+} - \boldsymbol{\psi}^{t,-}\|_2^2. \tag{80}$$

For the first term, Assumption 10 yields

$$\mathbb{E}\big[\|\boldsymbol{\xi}_i^{t,+} - \boldsymbol{\xi}_i^{t,-}\|_2^2 \,\big|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\big] \le \kappa_i^2\|\boldsymbol{\psi}^{t,+} - \boldsymbol{\psi}^{t,-}\|_2^2 = 4\kappa_i^2\sigma_t^2\|\boldsymbol{z}^t\|_2^2. \tag{81}$$

For the second term, $\boldsymbol{\psi}^{t,+} - \boldsymbol{\psi}^{t,-} = 2\sigma_t\boldsymbol{z}^t$, so $\|\boldsymbol{\psi}^{t,+} - \boldsymbol{\psi}^{t,-}\|_2^2 = 4\sigma_t^2\|\boldsymbol{z}^t\|_2^2$. Substituting into (80) yields

$$\mathbb{E}\big[|Y_i^{t,+} - Y_i^{t,-}|^2 \,\big|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\big] \le 8G_i^2(1+\kappa_i^2)\sigma_t^2\|\boldsymbol{z}^t\|_2^2.$$

Plugging this into the second moment of the estimator, we get

$$\begin{aligned}\mathbb{E}\left[\left\|\widehat{\nabla}_i^{\mathrm{MP}} J_i\right\|_2^2 \,\middle|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\right] &= \frac{d^2\|\boldsymbol{z}_i^t\|_2^2}{4\sigma_t^2}\mathbb{E}\big[|Y_i^{t,+} - Y_i^{t,-}|^2 \,\big|\, \boldsymbol{\theta}^t, \boldsymbol{z}^t\big] \\ &\le 2d^2G_i^2(1+\kappa_i^2)\|\boldsymbol{z}^t\|_2^2\|\boldsymbol{z}_i^t\|_2^2.\end{aligned}$$

Since $\|\boldsymbol{z}_j^t\|_2 = 1$ for every $j \in [n]$, we have $\|\boldsymbol{z}^t\|_2^2 = n$. Taking the expectation over $\boldsymbol{z}^t$ by the tower rule therefore yields

$$\mathbb{E}\left[\left\|\widehat{\nabla}_i^{\mathrm{MP}} J_i\right\|_2^2 \,\middle|\, \boldsymbol{\theta}^t\right] \le 2nd^2G_i^2(1+\kappa_i^2),$$

which proves the claim. □

F.2 PROOF OF THEOREM 3

We now prove convergence of SPADE-MP to the NE. Two features distinguish this proof from those of SPADE and SPADE-NE. Notably, Estimator (9) has a second moment bounded independently of the smoothing radius σ_t, so the primal update uses an outer step γ_t rather than $\gamma_t\sigma_t^2$ and the squared-update term is bounded by a constant.

Theorem 5 (Convergence of SPADE-MP, full version of Theorem 3) *Suppose Assumptions 1–6 and 8–10 hold. Assume that $\widehat{\mu} > 0$ and the stepsizes are positive, nonincreasing, and satisfy*

$0 < \sigma_t \le \min\{1, R_{\min}\}$, $\sup_{t\ge 1}\gamma_t \le \frac{\widehat{\mu}}{\widehat{\nu}_{\mathrm{MP}}}$ *where* $\widehat{\nu}_{\mathrm{MP}} := 6nL_{\mathrm{PR}}^2$, *and* $\sigma_t \to 0$. *Then, for every* $i \in [n]$,

$$\mathcal{R}_i^{\mathrm{NE}}(T) \le \mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}}\left(\frac{1}{\gamma_T} + \sum_{t=1}^{T}\gamma_t + \sum_{t=1}^{T}\sigma_t\right)}\right), \tag{82}$$

and

$$\mathcal{R}_g(T) \le \mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}}\left(\frac{1}{\gamma_T} + \sum_{t=1}^{T}\gamma_t + \sum_{t=1}^{T}\sigma_t\right)}\right). \tag{83}$$

In particular, choosing $\gamma_t = \tau_1(\tau_2 t + \tau_1)^{-1/2}$ *and* $\sigma_t = \tau_3(\tau_4 t + \tau_3)^{-1/2}$ *for positive constants* $\tau_1, \tau_2, \tau_3, \tau_4$ *gives* $\mathcal{R}_i^{\mathrm{NE}}(T) = \mathcal{O}(T^{3/4})$ *and* $\mathcal{R}_g(T) = \mathcal{O}(T^{3/4})$.

Proof: For each player $i \in [n]$, define the Lagrangian against the performative risk as in Appendix E,

$$\mathcal{L}_{\mathrm{PR}}^{(i)}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}, \boldsymbol{\lambda}) = \mathrm{PR}_i(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}) + \left\langle \boldsymbol{\lambda},\, \boldsymbol{g}_i(\boldsymbol{\theta}_i) + \sum_{j\ne i}\boldsymbol{g}_j(\boldsymbol{\theta}_j)\right\rangle.$$

From primal–dual theory Boyd & Vandenberghe (2004); Facchinei & Pang (2010), with $\boldsymbol{\lambda}^{\mathrm{NE}} \in \boldsymbol{\Lambda}$ the common Lagrange multiplier of the stationarity system (2) as in Appendix E, the saddle point $(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})$ satisfies the projected primal–dual fixed-point conditions for every control parameter $\gamma > 0$. Hence, at every iteration t, we may take $\gamma = \gamma_t$ and write

$$\boldsymbol{\theta}_i^{\mathrm{NE}} = P_{\boldsymbol{\Omega}_i}\left[\boldsymbol{\theta}_i^{\mathrm{NE}} - \gamma_t\left(\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top \boldsymbol{\lambda}^{\mathrm{NE}}\right)\right], \tag{84}$$

$$\boldsymbol{\lambda}^{\mathrm{NE}} = \left[\boldsymbol{\lambda}^{\mathrm{NE}} + \gamma_t\left(\boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}}) + \sum_{j\ne i}\boldsymbol{g}_j(\boldsymbol{\theta}_j^{\mathrm{NE}})\right)\right]_+. \tag{85}$$

Recall from Appendix E the PR-based target mapping $\Psi^{i,t}(\boldsymbol{\theta}, \boldsymbol{\lambda}) := \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i)^\top \boldsymbol{\lambda}$, and define the multipoint counterpart of $\widehat{\phi}_i^t$,

$$\widehat{\phi}_i^{\mathrm{MP},t}(\boldsymbol{\theta}_i^t, \boldsymbol{\lambda}) := \widehat{\nabla}_i^{\mathrm{MP}} J_i + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda},$$

so that the primal update (13) is $\boldsymbol{\theta}_i^{t+1} = P_{\boldsymbol{\Omega}_i}[\boldsymbol{\theta}_i^t - \gamma_t \widehat{\phi}_i^{\mathrm{MP},t}(\boldsymbol{\theta}_i^t, \boldsymbol{\lambda}_i^t)]$.

By nonexpansiveness of Euclidean projection applied to the primal update (13), and comparing with the fixed-point identity (84), we obtain

$$\mathbb{E}\|\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 \le \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 \;+\; \underbrace{\gamma_t^2 \sum_{i=1}^{n}\mathbb{E}\left\|\widehat{\phi}_i^{\mathrm{MP},t}(\boldsymbol{\theta}_i^t, \boldsymbol{\lambda}_i^t) - \Psi^{i,t}(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})\right\|_2^2}_{:=(\mathrm{I}^{\mathrm{MP}})}$$
$$\underbrace{- 2\gamma_t \sum_{i=1}^{n}\mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}},\, \widehat{\phi}_i^{\mathrm{MP},t}(\boldsymbol{\theta}_i^t, \boldsymbol{\lambda}_i^t) - \Psi^{i,t}(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})\right\rangle}_{:=(\mathrm{II}^{\mathrm{MP}})}. \tag{86}$$

This is the recursion (69) of the SPADE-NE proof with the outer step $\gamma_t\sigma_t^2$ replaced by γ_t and the one-point estimator replaced by the two-point estimator.

We first bound $(\mathrm{I}_{\mathrm{MP}})$. Add and subtract $\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t$ and $\Psi^{i,t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}_i^t)$ inside the norm of $(\mathrm{I}_{\mathrm{MP}})$, then apply Cauchy–Schwarz to obtain

$$
\begin{aligned}
(\mathrm{I}^{\mathrm{MP}}) \le & \underbrace{3\gamma_t^2 \sum_{i=1}^n \mathbb{E}\left\| \widehat{\phi}_i^{\mathrm{MP},t} - \left(\nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t\right)\right\|_2^2}_{(a)} \\
& + \underbrace{3\gamma_t^2 \sum_{i=1}^n \mathbb{E}\left\| \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) + \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t - \Psi^{i,t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}_i^t)\right\|_2^2}_{(b)} \\
& + \underbrace{3\gamma_t^2 \sum_{i=1}^n \mathbb{E}\left\| \Psi^{i,t}(\boldsymbol{\theta}^t, \boldsymbol{\lambda}_i^t) - \Psi^{i,t}(\boldsymbol{\theta}^{\mathrm{NE}}, \boldsymbol{\lambda}^{\mathrm{NE}})\right\|_2^2}_{(c)}.
\end{aligned} \tag{87}
$$

For the term (a), the constraint terms cancel, which leaves the variance of the estimator around its conditional mean. By $\mathbb{E}\|\boldsymbol{X} - \mathbb{E}\boldsymbol{X}\|_2^2 \le \mathbb{E}\|\boldsymbol{X}\|_2^2$ with $\boldsymbol{X} = \widehat{\nabla}_i^{\mathrm{MP}} J_i$, together with Lemma 3,

$$
(a) \le 3\gamma_t^2 \sum_{i=1}^n \mathbb{E}\left\| \widehat{\nabla}_i^{\mathrm{MP}} J_i \right\|_2^2 \le 6nd^2 \sum_{i=1}^n \left(G_i^2(1+\kappa_i^2)\right)\gamma_t^2. \tag{88}
$$

Note the departure from Appendix E: the one-point counterpart (72) carries a factor $1/\sigma_t^2$, whereas (88) is a constant multiple of γ_t^2.

For (b), this is the same smoothing and inward-shift bias as in Appendix E. By (74),

$$
(b) \le 3L_{\mathrm{PR}}^2 n^2 \left(1 + \frac{2C}{R_{\min}}\right)^2 \gamma_t^2 \sigma_t^2. \tag{89}
$$

Finally, for (c), similar to Appendix E, by $\|\boldsymbol{a}+\boldsymbol{b}\|_2^2 \le 2\|\boldsymbol{a}\|_2^2 + 2\|\boldsymbol{b}\|_2^2$, Assumption 9 for the $\nabla\mathrm{PR}$-piece, and the G_g-Lipschitzness of $\boldsymbol{g}_i$ (Assumption 6) together with $\|\boldsymbol{\lambda}_i^t\|_2 \le \Lambda_{\max}$ (Lemma 4) and $\|\boldsymbol{\lambda}^{\mathrm{NE}}\|_2 \le \Lambda_{\max}$ (Assumption 8) for the constraint piece,

$$
(c) \le 6nL_{\mathrm{PR}}^2 \gamma_t^2 \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + 24nG_g^2\Lambda_{\max}^2\gamma_t^2. \tag{90}
$$

Substituting (88)–(90) into (87) and using $\sigma_t \le 1$, we obtain

$$
(\mathrm{I}_{\mathrm{MP}}) \le 6nL_{\mathrm{PR}}^2\gamma_t^2\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + \widehat{\pi}_{\mathrm{MP}}\gamma_t^2, \tag{91}
$$

where the lower order σ_t terms have been suppressed into the constant

$$
\widehat{\pi}_{\mathrm{MP}} := 6nd^2 \sum_{i=1}^n G_i^2(1+\kappa_i^2) + 24nG_g^2\Lambda_{\max}^2 + 3n^2L_{\mathrm{PR}}^2\left(1 + \frac{2C}{R_{\min}}\right)^2.
$$

Now we bound $(\mathrm{II}^{\mathrm{MP}})$. Since $\boldsymbol{\theta}^t$ and $\boldsymbol{\lambda}^t$ are measurable with respect to the past, the tower property and Lemma 3 give

$$
\begin{aligned}
-(\mathrm{II}_{\mathrm{MP}}) = & -2\gamma_t \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}})\right\rangle \\
& - 2\gamma_t \sum_{i=1}^n \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t - \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top \boldsymbol{\lambda}^{\mathrm{NE}}\right\rangle.
\end{aligned} \tag{92}
$$

Add and subtract $\nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t)$ inside the first inner product. Using the smoothing and inward-shift bias bound

$$
\left\| \nabla_{\boldsymbol{\theta}_i}\widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i}\mathrm{PR}_i(\boldsymbol{\theta}^t)\right\|_2 \le L_{\mathrm{PR}}\sigma_t\sqrt{n}\left(1 + \frac{2C}{R_{\min}}\right)
$$

and $\|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}\|_2 \le 2C$ then gives

$$-2\gamma_t \sum_{i=1}^{n} \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla_{\boldsymbol{\theta}_i} \widetilde{\mathrm{PR}}_i(\boldsymbol{\theta}^{t,\sigma}) - \nabla_{\boldsymbol{\theta}_i} \mathrm{PR}_i(\boldsymbol{\theta}^t) \right\rangle \le 4Cn^{3/2} L_{\mathrm{PR}} \left(1 + \frac{2C}{R_{\min}}\right) \gamma_t \sigma_t. \tag{93}$$

while the remaining piece is

$$-2\gamma_t \sum_{i=1}^{n} \mathbb{E}\big\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla_{\boldsymbol{\theta}_i} \mathrm{PR}_i(\boldsymbol{\theta}^t) - \nabla_{\boldsymbol{\theta}_i} \mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}}) \big\rangle \le -2\widehat{\mu}\gamma_t \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2,$$

by the strong monotonicity of $\nabla \mathrm{PR}$ in Proposition 1, which holds because $\widehat{\mu} > 0$ by hypothesis.

The constraint-coupling inner product is handled as in Step 3 of Appendices D and E yielding

$$-2\gamma_t \sum_{i=1}^{n} \mathbb{E}\left\langle \boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}, \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)^\top \boldsymbol{\lambda}_i^t - \nabla \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}})^\top \boldsymbol{\lambda}^{\mathrm{NE}} \right\rangle \le 2\gamma_t \sum_{i=1}^{n} \mathbb{E}\left[B\|\boldsymbol{\lambda}_i^t - \overline{\boldsymbol{\lambda}}^t\|_2\right] + 2\gamma_t \mathbb{E}[\Delta_t],$$

where, as before, $\Delta_t := \sum_{i=1}^n \langle \boldsymbol{\lambda}^{\mathrm{NE}} - \boldsymbol{\lambda}_i^t, \boldsymbol{g}_i(\boldsymbol{\theta}_i^t)\rangle$. Combining,

$$-(\mathrm{II}^{\mathrm{MP}}) \le -2\widehat{\mu}\gamma_t \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + 4Cn^{3/2} L_{\mathrm{PR}} \left(1 + \frac{2C}{R_{\min}}\right) \gamma_t \sigma_t + 2\gamma_t \sum_{i=1}^{n} \mathbb{E}\left[B\|\boldsymbol{\lambda}_i^t - \overline{\boldsymbol{\lambda}}^t\|_2\right] + 2\gamma_t \mathbb{E}[\Delta_t]. \tag{94}$$

Thus, define the constants:

- $\widehat{\nu}_{\mathrm{MP}} := 6nL_{\mathrm{PR}}^2$,
- $\widehat{\rho}_{\mathrm{MP}} := 4Cn^{3/2} L_{\mathrm{PR}} \left(1 + \frac{2C}{R_{\min}}\right)$.

Plugging (91) and (94) into (86) gives the descent recursion

$$\mathbb{E}\|\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 \le \big(1 - 2\widehat{\mu}\gamma_t + \widehat{\nu}_{\mathrm{MP}}\gamma_t^2\big)\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 + 2\gamma_t \sum_{i=1}^{n} \mathbb{E}\left[B\|\boldsymbol{\lambda}_i^t - \overline{\boldsymbol{\lambda}}^t\|_2\right] + 2\gamma_t \mathbb{E}[\Delta_t] + \widehat{\pi}_{\mathrm{MP}}\gamma_t^2 + \widehat{\rho}_{\mathrm{MP}}\gamma_t\sigma_t. \tag{95}$$

Under the stepsize assumption $\sup_t \gamma_t \le \widehat{\mu}/\widehat{\nu}_{\mathrm{MP}}$, $1 - 2\widehat{\mu}\gamma_t + \widehat{\nu}_{\mathrm{MP}}\gamma_t^2 \le 1 - \widehat{\mu}\gamma_t$. Dividing by $\widehat{\mu}\gamma_t$ and summing over $t \in [T]$, telescoping as before in (60), and using that γ_t is nonincreasing and $\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 \le 4nC^2$ yields

$$\sum_{t=1}^{T} \frac{1}{\gamma_t} \left(\mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 - \mathbb{E}\|\boldsymbol{\theta}^{t+1} - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2\right) \;\le\; \frac{4nC^2}{\gamma_T}. \tag{96}$$

Bounding the two dual terms by Lemma 6 with the admissible comparator $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\mathrm{NE}} \in \boldsymbol{\Lambda}$ and by Lemma 5, namely

$$\sum_{t=1}^{T} \mathbb{E}[\Delta_t] \le \frac{2n\Lambda_{\max}^2}{\gamma_T} + \left(\frac{nB^2}{2} + \frac{2nB^2}{1 - \sigma_2(\boldsymbol{A})}\right) \sum_{t=1}^{T} \gamma_t, \qquad \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbb{E}\|\boldsymbol{\lambda}_i^t - \overline{\boldsymbol{\lambda}}^t\|_2 \le \frac{nB}{1 - \sigma_2(\boldsymbol{A})} \sum_{t=1}^{T} \gamma_t,$$

we obtain

$$\sum_{t=1}^{T} \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2 \;\le\; \mathcal{O}\left(\frac{1}{\widehat{\mu}\gamma_T} + \frac{1}{\widehat{\mu}} \sum_{t=1}^{T} \gamma_t + \frac{1}{\widehat{\mu}} \sum_{t=1}^{T} \sigma_t\right). \tag{97}$$

Finally, to obtain the regret bounds, as in Appendix E, compactness of $\boldsymbol{\Omega}$ implies the existence of

$$G_i^{\mathrm{PR}} := \sup_{\boldsymbol{\theta} \in \boldsymbol{\Omega}} \|\nabla_{\boldsymbol{\theta}_i} \mathrm{PR}_i(\boldsymbol{\theta})\|_2 < \infty.$$

The mean-value inequality therefore gives

$$\mathrm{PR}_i(\boldsymbol{\theta}_i^t, \boldsymbol{\theta}_{-i}^{\mathrm{NE}}) - \mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}}) \le G_i^{\mathrm{PR}} \|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}\|_2.$$

Applying Cauchy–Schwarz and (97),

$$\begin{aligned}
\mathcal{R}_i^{\mathrm{NE}}(T) &\le G_i^{\mathrm{PR}} \sum_{t=1}^{T} \mathbb{E}\|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}\|_2 \\
&\le G_i^{\mathrm{PR}} \sqrt{T \sum_{t=1}^{T} \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2} \\
&\le \mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}} \left(\frac{1}{\gamma_T} + \sum_{t=1}^{T} \gamma_t + \sum_{t=1}^{T} \sigma_t \right)} \right),
\end{aligned} \tag{98}$$

which proves (82).

For the cumulative constraint violation, we have pathwise

$$\begin{aligned}
\left\| \left[\sum_{t=1}^{T} \sum_{i=1}^{n} \boldsymbol{g}_i(\boldsymbol{\theta}_i^t) \right]_+ \right\|_2 &\le \left\| \sum_{t=1}^{T} \sum_{i=1}^{n} \left(\boldsymbol{g}_i(\boldsymbol{\theta}_i^t) - \boldsymbol{g}_i(\boldsymbol{\theta}_i^{\mathrm{NE}}) \right) \right\|_2 \\
&\le G_g \sum_{t=1}^{T} \sum_{i=1}^{n} \|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{\mathrm{NE}}\|_2 \\
&\le G_g \sqrt{n} \sum_{t=1}^{T} \|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2.
\end{aligned} \tag{99}$$

Taking expectations, applying Cauchy–Schwarz, and using (97),

$$\mathcal{R}_g(T) \le G_g \sqrt{nT} \sqrt{\sum_{t=1}^{T} \mathbb{E}\|\boldsymbol{\theta}^t - \boldsymbol{\theta}^{\mathrm{NE}}\|_2^2} \;\le\; \mathcal{O}\left(\sqrt{\frac{T}{\widehat{\mu}} \left(\frac{1}{\gamma_T} + \sum_{t=1}^{T} \gamma_t + \sum_{t=1}^{T} \sigma_t \right)} \right), \tag{100}$$

which proves (83).

This completes the proof. □

G EXPERIMENTAL DETAILS

Here, we document the instantiation of our networked Cournot competition model appearing in §4. It parallels the construction of (Yan & Cao, 2024, Appendix F.1), allowing for a side-by-side comparison of the two approaches. This section specifies the economic model, feasible action sets, stochastic feedback mechanisms, computation of the reference equilibria, updates, randomization, and reported metrics used to produce Figures 1-3. Unless otherwise stated, all algorithms use the same fixed game parameters, initialization rules, random seed set, and time horizon.

G.1 COURNOT MODEL DETAILS

There are $n = 5$ firms and $m = 3$ markets. Firm i chooses to produce quantity vector $\boldsymbol{\theta}_i = (\theta_{i1}, \theta_{i2}, \theta_{i3}) \in \mathbb{R}^3_+$, however, firms may be restricted from participating in some markets. We instantiate the participation matrix as

$$\boldsymbol{S} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \tag{101}$$

where $S_{ij} = 1$ means that firm i participates in market j. Thus, the numbers of participants in the three markets are $(2, 3, 3)$ and the relevant action dimensions of the firms are $(2, 1, 2, 2, 1)$, with other excluded market production quantity choices set to 0.

Firm i has production capacity Q_i, so its feasible set is

$$\boldsymbol{\Omega}_i = \left\{ \boldsymbol{\theta}_i \in \mathbb{R}^3_+ : \theta_{ij} = 0 \text{ if } S_{ij} = 0, \sum_{j=1}^{3} \theta_{ij} \le Q_i \right\}.$$

For market j, let $L_j(\boldsymbol{\theta}) = \sum_i \theta_{ij}$ denote aggregate supply and B_j denote market capacity. Thus, the *market-wise* coupled constraints are

$$g_j(\boldsymbol{\theta}) = L_j(\boldsymbol{\theta}) - B_j \le 0, \qquad j \in [3].$$

We split each market constraint into individual firm constraint components across their participating markets according to

$$g_{ij}(\boldsymbol{\theta}_i) = S_{ij} \left(\theta_{ij} - \frac{B_j}{\sum_{k=1}^{5} S_{kj}} \right),$$

and thus we can rewrite market j's coupled constraint as $\sum_{i=1}^{5} g_{ij}(\boldsymbol{\theta}_i) = L_j(\boldsymbol{\theta}) - B_j \le 0$.

At an evaluated production profile $\boldsymbol{\theta}$, we specify the deterministic inverse-demand component in market j as

$$q_j(L_j(\boldsymbol{\theta})) = 10 \left(10 + \frac{L_j(\boldsymbol{\theta})}{5} \right)^{-1/2}.$$

Meanwhile, if $\boldsymbol{\delta}$ is the action profile that induces the environment (i.e. is the argument for $\mathcal{D}$), the stochastic component is

$$\xi_j(\boldsymbol{\delta}; Z_j) = 2.5 + \min\{\max\{Z_j, 0\}, 5\} + \beta_j L_j(\boldsymbol{\delta}), \qquad Z_j \sim \mathcal{N}(0, 1). \tag{102}$$

Thus, this distribution shift is linear in the inducing action $\boldsymbol{\delta}$, so Assumption 9 is satisfied. As described in the paper, the inducing profile in (102) is the unperturbed intended action, $\boldsymbol{\delta} = \boldsymbol{\theta}^t$ for SPADE and the deployed perturbed profile(s) for SPADE-NE and SPADE-MP. Additionally, the exogenous term $2.5 + \min\{\max\{Z_j, 0\}, 5\}$ can be written as a $\mathcal{N}(2.5, 1)$ random variable clipped to $[2.5, 7.5]$. The three exogenous Z_j shocks are independent across all markets and iterations, and common to all firms. The resulting induced price in market j is

$$p_j(\boldsymbol{\theta}, \boldsymbol{\delta}; Z_j) = \xi_j(\boldsymbol{\delta}; Z_j) + q_j(L_j(\boldsymbol{\theta})).$$

Firm i cost function is negative profit across the three markets, so they seek to minimize

$$J_i(\boldsymbol{Z};\boldsymbol{\theta},\boldsymbol{\delta}) = \sum_{j=1}^{3} \left[d_{ij} - p_j(\boldsymbol{\theta},\boldsymbol{\delta};Z_j)\right]\theta_{ij}.$$

where d_{ij} is firm i's constant marginal cost of supplying one unit to market j.

We use a common performativity coefficient $\varepsilon = 0.6$ for all firms, which scales the distribution shift with respect to the aggregate market supply. Specifically, let β_j in Equation (102) be given by $\boldsymbol{\beta} = \varepsilon\boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is a (negative) weight vector with $\sum_j |\alpha_j| = 1$ and $\alpha_j < 0$.

Now we detail the parameters of our simulated model instance. We randomly draw

- $B_j = 10 + 5U_j$,
- $Q_i = 10 + 2V_i$,
- $d_{ij} = 1 + 0.5W_{ij}$,
- $\alpha_j = -R_j \Big/ \sum_{k=1}^{3} R_k$,

where U, V, W, R are uniform on $[0, 1)$. All draws are made once before the repeated-run loop and are fixed across algorithms and seeds. In particular, our resulting capacities are

$$\boldsymbol{B} = (14.074,\ 14.529,\ 10.635),$$
$$\boldsymbol{Q} = (11.827,\ 11.265,\ 10.195,\ 10.557,\ 11.094),$$

the marginal cost matrix is

$$\boldsymbol{D} = \begin{bmatrix} 1.479 & 1.243 & 1.396 \\ 1.482 & 1.400 & 1.480 \\ 1.079 & 1.071 & 1.328 \\ 1.485 & 1.211 & 1.018 \\ 1.479 & 1.458 & 1.425 \end{bmatrix},$$

and the performative effect parameters are $\boldsymbol{\alpha} = (-0.394, -0.286, -0.320)$, giving distribution shift factors $\boldsymbol{\beta} = (-0.236, -0.172, -0.192)$.

The communication graph is represented by the following doubly stochastic "Metropolis" matrix

$$\boldsymbol{A} = \begin{bmatrix} .25 & .25 & .25 & 0 & .25 \\ .25 & .25 & .25 & 0 & .25 \\ .25 & .25 & .25 & .25 & 0 \\ 0 & 0 & .25 & .50 & .25 \\ .25 & .25 & 0 & .25 & .25 \end{bmatrix}. \tag{103}$$

Finally, the PSE and NE references are computed independently of the stochastic trajectories by solving their deterministic KKT complementarity systems. The resulting equilibrium profiles are

$$\boldsymbol{\theta}^{\mathrm{PSE}} = \begin{bmatrix} 7.365 & 0 & 2.445 \\ 0 & 7.431 & 0 \\ 6.708 & 3.487 & 0 \\ 0 & 3.611 & 6.946 \\ 0 & 0 & 1.243 \end{bmatrix}, \tag{104}$$

$$\boldsymbol{\theta}^{\mathrm{NE}} = \begin{bmatrix} 5.808 & 0 & 3.004 \\ 0 & 4.612 & 0 \\ 5.868 & 4.327 & 0 \\ 0 & 5.590 & 4.759 \\ 0 & 0 & 2.872 \end{bmatrix}. \tag{105}$$

The common market multipliers are respectively $\boldsymbol{\lambda}^{\mathrm{PSE}} = (0.726, 1.626, 2.277)$ and $\boldsymbol{\lambda}^{\mathrm{NE}} = (0, 0.895, 1.687)$.

G.2 Algorithm Implementation Details

Note that because quantity choices in inactive markets are fixed at zero, we use perturbations within each firm's active-coordinate affine hull. If $d_i = \sum_j S_{ij}$ is the "relevant" dimension of firm i, the center and radius of the inscribed ball used by the algorithm are

$$R_i = \frac{Q_i}{d_i + \sqrt{d_i}}, \qquad p_{ij} = S_{ij} R_i. \tag{106}$$

Numerically, we have

$$\boldsymbol{R} = (3.4639753771,\ 5.6323592462,\ 2.9860700345,\ 3.0920726677,\ 5.5468815192).$$

For an independent direction $\boldsymbol{z}_i^t \sim \mathrm{Unif}(\mathbb{S}^{d_i-1})$ with zeros on inactive coordinates, define the usual

$$\boldsymbol{\theta}_i^{t,\sigma} = \left(1 - \frac{\sigma_t}{R_i}\right) \boldsymbol{\theta}_i^t + \frac{\sigma_t}{R_i} \boldsymbol{p}_i, \qquad \boldsymbol{\psi}_i^{t,\pm} = \boldsymbol{\theta}_i^{t,\sigma} \pm \sigma_t \boldsymbol{z}_i^t. \tag{107}$$

Thus, perturbed queries belong to the feasible set $\boldsymbol{\Omega}_i$ whenever $0 < \sigma_t \le R_i$. Equivalently, to be pedantic about Assumption 3, one could reparameterize each firm's feasible set in its d_i-dimensional active-coordinate space, where all perturbations and estimator dimensions below are taken in this reduced space.

As defined in the paper, the one-point estimator (5) used by SPADE and SPADE-NE is

$$\widehat{\nabla}_i J_i^t = \frac{d_i}{\sigma_t} J_i(\boldsymbol{Z}^t; \boldsymbol{\psi}^{t,+}, \boldsymbol{\delta}^t) \boldsymbol{z}_i^t. \tag{108}$$

SPADE sets $\boldsymbol{\delta}^t = \boldsymbol{\theta}^t$ (intended-action sampling), whereas SPADE-NE sets $\boldsymbol{\delta}^t = \boldsymbol{\psi}^{t,+}$ (deployed-action sampling). SPADE-MP uses the symmetric two-point estimator (9)

$$\widehat{\nabla}_i^{\mathrm{MP}} J_i^t = \frac{d_i}{2\sigma_t} \left[J_i(\boldsymbol{Z}^t; \boldsymbol{\psi}^{t,+}, \boldsymbol{\psi}^{t,+}) - J_i(\boldsymbol{Z}^t; \boldsymbol{\psi}^{t,-}, \boldsymbol{\psi}^{t,-}) \right] \boldsymbol{z}_i^t. \tag{109}$$

Note that the plus and minus deployments both use the same exogenous market-shock vector $\boldsymbol{Z}^t$ and their respective decision-dependent distribution shifts. This coupling thus satisfies Assumption 10. The primal and dual updates are carried out as specified in Appendix B.

Note that 0 production is a strict Slater point for the coupled market capacities. Then in line with Assumption 8, the players are supplied with the common compact set

$$\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^3_+ : \|\boldsymbol{\lambda}\|_2 \le 10\}. \tag{110}$$

The Euclidean norms of the computed common market multipliers are 2.89051 for the PSE and 1.90993 for the NE, so both lie strictly inside Λ.

For stepsizes and smoothing schedules, Figure 1 (SPADE-MP in the main body) and Figures 2 (SPADE) and 3 (SPADE-NE) use

Algorithm	γ_t	σ_t	primal multiplier a_t
SPADE	$0.3(t+100)^{-1/4}$	$1.5(t+100)^{-1/4}$	$\gamma_t \sigma_t^2$
SPADE-NE	$0.1(t+100)^{-1/4}$	$3.0(t+100)^{-1/4}$	$\gamma_t \sigma_t^2$
SPADE-MP	$0.1(t+100)^{-1/2}$	$3.0(t+100)^{-1/2}$	γ_t

Thus, the one-point schedules have $\gamma_t, \sigma_t = \Theta(t^{-1/4})$ and SPADE-MP has $\gamma_t, \sigma_t = \Theta(t^{-1/2})$, as prescribed in the theorems.

G.3 Experimental Results

All final figures use $T = 10^6$ horizon and average over 30 seeds, which are $\{0, 1, \ldots, 29\}$. For each seed, every active entry of $\boldsymbol{\theta}_i^1$ is independently drawn from $\mathrm{Unif}[0, 1)$ and projected onto $\boldsymbol{\Omega}_i$. The dual variables are initialized at $\boldsymbol{\lambda}_i^1 = \boldsymbol{0}$. The Cournot instance, graph, and equilibrium references are fixed across seeds, while initial actions, perturbation directions, and exogenous shocks vary

across seeds. Matching seed numbers use the same initialization and direction/shock streams across algorithms, while the algorithms then generate different trajectories through their respective updates.

We report (signed) time-averaged PSE/NE-performative regret and (signed) time-averaged market constraint slack. For $\star \in \{\mathrm{PSE}, \mathrm{NE}\}$, define the unilateral hybrid profile $\boldsymbol{\theta}^{t',(i;\star)} = \left(\boldsymbol{\theta}_i^{t'}, \boldsymbol{\theta}_{-i}^{\star}\right)$. In the left panel of each algorithm figure, the reported PSE regret for SPADE is

$$\frac{\mathcal{R}_i^{\mathrm{PSE}}(t)}{t} = \frac{1}{t}\sum_{t'=1}^{t}\left[\mathbb{E}_{\boldsymbol{Z}} J_i\Big(\boldsymbol{Z}; \boldsymbol{\theta}^{t',(i;\mathrm{PSE})}, \boldsymbol{\theta}^{\mathrm{PSE}}\Big) - \mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{PSE}})\right], \tag{111}$$

and for SPADE-NE and SPADE-MP, it is

$$\frac{\mathcal{R}_i^{\mathrm{NE}}(t)}{t} = \frac{1}{t}\sum_{t'=1}^{t}\left[\mathbb{E}_{\boldsymbol{Z}} J_i\Big(\boldsymbol{Z}; \boldsymbol{\theta}^{t',(i;\mathrm{NE})}, \boldsymbol{\theta}^{t',(i;\mathrm{NE})}\Big) - \mathrm{PR}_i(\boldsymbol{\theta}^{\mathrm{NE}})\right]. \tag{112}$$

As noted in the main body, regret remains signed. A unilateral action is required to lie in Ω_i, however, the associated hybrid profile need not satisfy the shared market capacities. Since the equilibrium profiles satisfy the market constraints, negative gaps are possible.

The right panel of each algorithm figure reports the signed per-market quantity

$$\overline{g}_j(t) = \frac{1}{t}\sum_{t'=1}^{t}\sum_{i=1}^{5} g_{ij}(\boldsymbol{\theta}_i^{t'}) = \frac{1}{t}\sum_{t'=1}^{t}\Big[L_j(\boldsymbol{\theta}^{t'}) - B_j\Big]. \tag{113}$$

The solid curves are means and the shaded bands are 1 sample standard deviation across the 30 runs.

SPADE-MP Figure 1 in the main text yields the cleanest convergence result. At $T = 10^6$, the five firms' mean SPADE-MP NE-regrets are $(-0.00184, 0.00012, 0.00347, 0.00473, -0.00271)$, and the three markets' signed average constraints are $(-2.39171, -0.00003, 0.00143)$. The first market is slack at the NE by 2.39740, while the second and third market capacities are active. The mean cumulative violation ratio is $\mathcal{R}_g(T)/T = 0.001469$ and the mean instantaneous distance from $\boldsymbol{\theta}^T$ to the NE is 0.06793, which are consistent with the regret and violation bounds of Theorem 3.

SPADE & SPADE-NE Figures 2 and 3 report SPADE and SPADE-NE, respectively, under the theory-compatible schedules. The mean endpoint distances $\|\boldsymbol{\theta}^T - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2$ and $\|\boldsymbol{\theta}^T - \boldsymbol{\theta}^{\mathrm{NE}}\|_2$ are 5.58 (for PSE) and 1.81 (for NE) after 10^9 iterations. Visually, both one-point algorithms exhibit steady, but very slow convergence toward their respective equilibria, with minimal constraint violation. SPADE-NE shows relatively steady convergence toward the NE, while SPADE shows steady convergence toward the PSE for Firms 1, 2, and 3, with slower convergence for Firms 4 and 5, which experienced the greatest deviations from their PSE actions. This highlights that although one-point payoff-based methods can learn equilibria in decentralized performative games with coupled constraints, their finite-horizon performance is limited by the high variance of the one-point estimator, whose second moment scales as $O(1/\sigma_t^2)$, together with the conservative step-size attenuation required to control this variance. Correspondingly, the theoretical regret guarantees for SPADE and SPADE-NE scale as $O(T^{7/8})$. This further motivates our SPADE-MP construction, which preserves the key properties of SPADE–namely, payoff-based gradient estimation, no knowledge of the functional form of the cost, and no costly exchange of action information–while achieving substantially faster convergence and an $O(T^{3/4})$ regret rate, matching the order obtained by the full-information stochastic-gradient baseline.

Since the theory-compatible one-point schedules exhibit substantial finite-horizon payoff-estimator variance and slow convergence–which, as detailed in the main body, is alleviated by the two-point approach in SPADE-MP–we also report a more aggressive schedule comparison to illustrate the convergence behavior of the algorithms, departing slightly from the theorem-prescribed schedules. These simulations use

Algorithm	γ_t	σ_t
SPADE	$0.03u_t^{-3/4}$	$0.8u_t^{-1/20}$
SPADE-NE	$0.10u_t^{-1}$	$0.8u_t^{-1/20}$

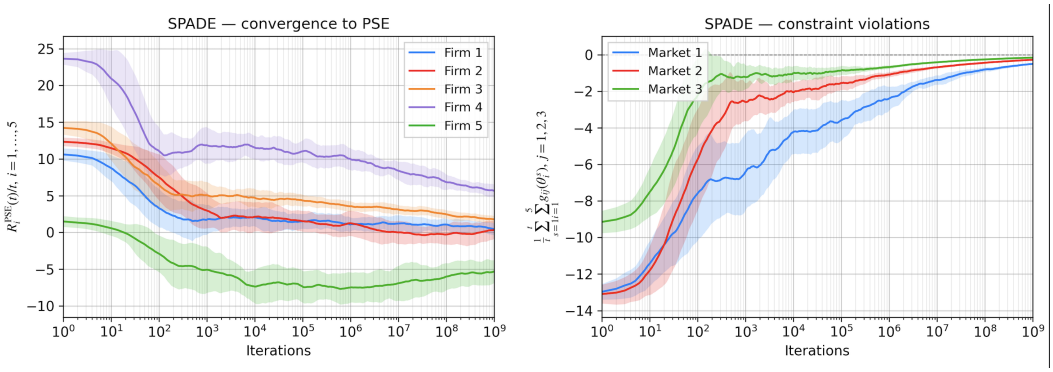


Figure 2: SPADE, showing time-averaged PSE regret (left) and signed time-averaged market-capacity residuals (right). Curves show the mean across 30 runs, with bands corresponding to ± 1 sample SD.

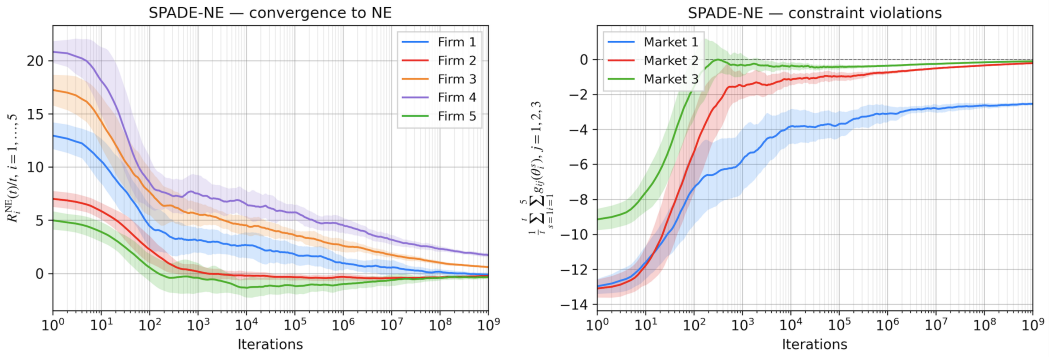


Figure 3: SPADE-NE, showing time-averaged NE regret (left) and signed time-averaged market-capacity residuals (right). Curves show the mean across 30 runs, with bands corresponding to ± 1 sample SD.

where $u_t := \frac{t+1000}{1001}$. These schedules preserve all other algorithmic features, including the payoff-based estimator, feasible-query construction, primal gain $\gamma_t \sigma_t^2$, dual update, and projections, but change the decay exponents and are therefore not explicitly covered by the $t^{-1/4}$ theorem schedules. They were selected by screening 12 pre-specified candidates for each one-point method on four seeds disjoint from the final seed set, using the normalized last-quarter equilibrium distance plus a positive-constraint penalty. The selected schedules were then evaluated on 30 matched final runs.

Figures 4 and 5 report the selected aggressive schedules for SPADE on PSE and SPADE-NE on NE, respectively. These more aggressive schedules improve finite-horizon convergence behavior substantially. The mean endpoint distances $\|\boldsymbol{\theta}^T - \boldsymbol{\theta}^{\mathrm{PSE}}\|_2$ and $\|\boldsymbol{\theta}^T - \boldsymbol{\theta}^{\mathrm{NE}}\|_2$ are 2.51644 and 1.14287, respectively, highlighting convergence toward their respective equilibria after just 10^6 iterations. For SPADE, Figure 4 also more clearly illustrates the convergence of Firms 4 and 5 toward their PSE profiles, revealing that, at least for this game instance, these firms begin converging more steadily on average only after Firms 1, 2, and 3 are sufficiently close to their equilibrium profiles.

The selected step-size schedules are included to more clearly demonstrate the converging behavior of SPADE and SPADE-NE over a finite horizon. In the runs for Figures 4 and 5, the larger and more slowly decaying smoothing radius helps balance signal against noise in the one-point estimator. In particular, the expected gradient contribution in the primal update is scaled by $\gamma_t \sigma_t^2$, whereas the payoff-level noise is approximately scaled by $\gamma_t \sigma_t$. At $t = 10^6$, the selected SPADE schedule has approximately 2.535 times the effective gradient multiplier of the theory-compatible schedule and approximately 4.7 times less payoff-level noise scaling. For SPADE-NE, the corresponding factors are approximately 1.127 and 5.3, respectively. These comparisons are intended only as finite-horizon empirical explanations and do not constitute additional asymptotic guarantees.

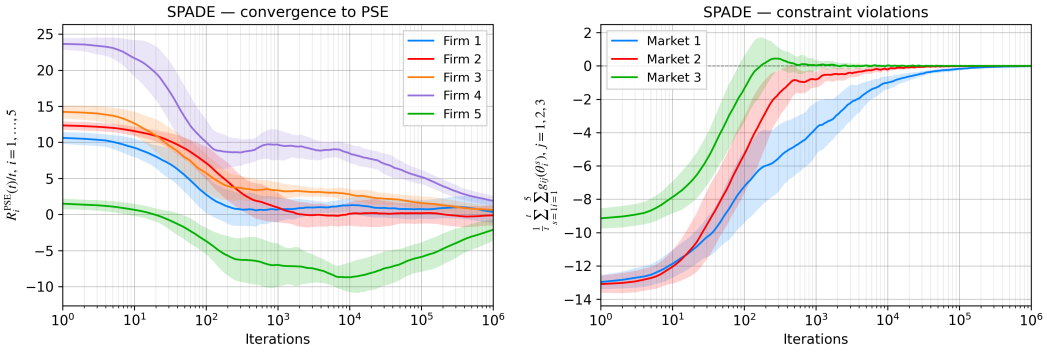


Figure 4: SPADE with the selected aggressive schedule, $\gamma_t = 0.03u_t^{-3/4}$ and $\sigma_t = 0.8u_t^{-1/20}$.

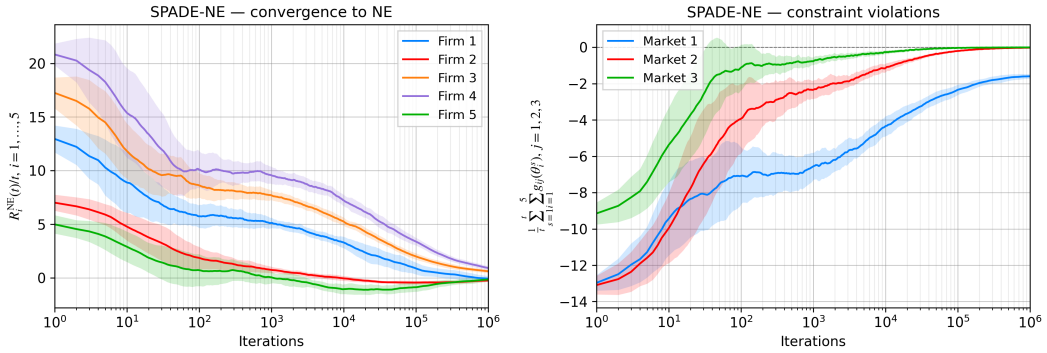


Figure 5: SPADE-NE with the selected aggressive schedule, $\gamma_t = 0.10u_t^{-1}$ and $\sigma_t = 0.8u_t^{-1/20}$.

Communication Costs As a final simple, yet significant analysis, we compare the communication cost of SPADE compared with that of the decentralized primal-dual algorithm of Yan & Cao (2024). All three SPADE variants exchange only the local dual vector $\boldsymbol{\lambda}_i^t \in \mathbb{R}^3$ with each neighbor, and the graph in (103) has 14 directed neighbor transmissions per round (excluding self weights). Thus, every SPADE variant sends $14 \cdot 3 = 42$ scalar values per round and 42 million over 10^6 rounds.

In addition to the three-market dual vector, Algorithm 1 of Yan & Cao (2024) communicates, per directed link, a three-market local action and four three-market estimates of the other players. Under the dense-vector convention stated in that paper, this is $3 + 4 \cdot 3 + 3 = 18$ scalars per link, or 252 per round and 252 million by $T = 10^6$ iterations[2]. Figure 6 plots these cumulative counts. Thus, SPADE significantly reduces the communication cost (as measured by scalars exchanged) by a factor of 6 in this particular Cournot instance, and this comparative advantage grows as the number of players and the dimensionality of the action space increase.

[2]If known inactive action coordinates are removed from every baseline message, this count decreases to 154 scalars per round. The 252 count is used in the figure because it corresponds to the paper's general $nd + m$ message dimension.

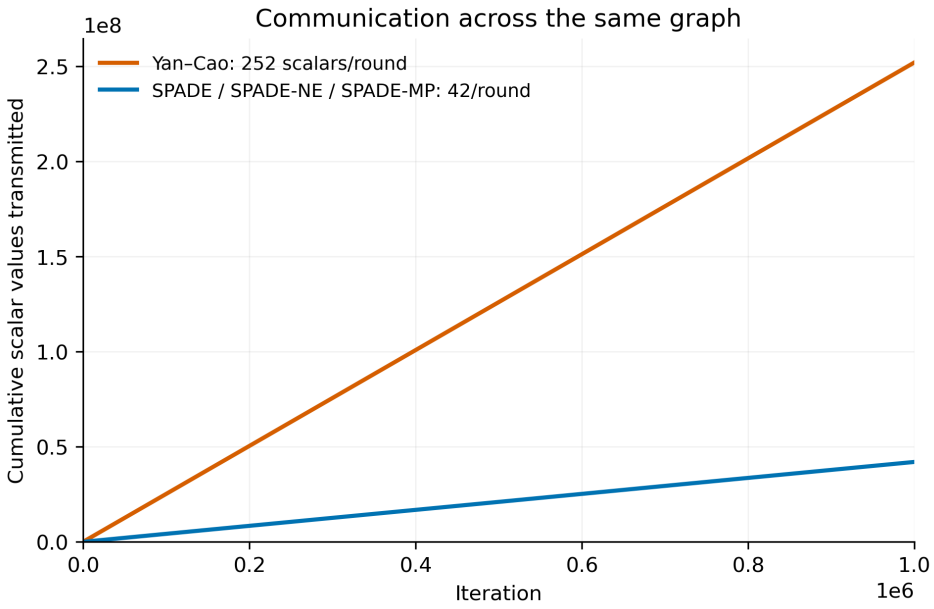


Figure 6: Cumulative scalar transmissions across the communication graph $\mathcal{G}(\boldsymbol{A})$ in our specified 5 firm, 3 market Cournot model.