Robust Equilibria in Continuous Games: From Strategic to Dynamic Robustness

Anonymous Author(s)
Affiliation
Address
email

Abstract

In this paper, we examine the robustness of Nash equilibria in continuous games, under both strategic and dynamic uncertainty. Starting with the former, we introduce the notion of a *robust equilibrium* as those equilibria that remain invariant to small—but otherwise arbitrary—perturbations to the game's payoff structure, and we provide a crisp geometric characterization thereof. Subsequently, we turn to the question of dynamic robustness, and we examine which equilibria are stable limit points of the dynamics of *"follow the regularized leader"* (FTRL) in the presence of randomness and uncertainty. Despite their very distinct origins, we establish a structural correspondence between these two notions of robustness: strategic robustness implies dynamic robustness, and, conversely, the requirement of strategic robustness cannot be relaxed if dynamic robustness is to be maintained. Finally, we examine the rate of convergence to robust equilibria as a function of the underlying regularizer, and we show that entropically regularized learning converges at a geometric rate in games with affinely constrained action spaces.

1 Introduction

A fundamental requirement in game theory—which predates even the cornerstone notion of a Nash equilibrium—concerns the robustness that should be inherent in any axiomatization of rational behavior. To quote a famous passage by von Neumann & Morgenstern [39, p. 32]: "In whatever way we formulate the guiding principles and the objective justification of rational behavior, provisos will have to be made for every possible conduct of "the others." If the superiority of rational behavior over any other kind is to be established, then its description must include rules of conduct for all conceivable situations—including those where "the others" behaved irrationally in the sense of the standards which the theory will set for them."

As a byproduct of this tenet, there has been a flurry of activity since the 1970s in proposing refinements of the Nash equilibrium concept trying to dismiss equilibria that are highly fragile or otherwise implausible (e.g., because they involve incredible threats).[1] This pursuit of robustness has recently gained increased momentum owing to the applications of game theory to machine learning and data science, two fields where the notion of robustness has been likewise elusive. Here, even though many game-theoretic solutions perform extremely well on specific tasks—such as a well-trained generative adversarial network (GAN) at equilibrium—the resulting models tend to have a narrow performance envelope, being brittle, and unable to adapt to situations that deviate from their initial configuration.

In game-theoretic terms, this highlights the fact that, even though a Nash equilibrium is resilient to unilateral deviations, it need not be robust to small perturbations in the *payoff data* of the game

[1] For a masterful introduction to the topic, see the textbook of van Damme [38].

(which, in a machine learning context, could represent distributional shifts, incomplete observations, and/or other sources of uncertainty). In view of this, it is natural to ask

Which equilibria are robust in the presence of uncertainty?

This question has been the lodestar of the equilibrium refinement literature, and it has led to a wide array of proposals aiming to get rid of "unreasonable" equilibria that may disappear even under the most minute perturbation to the players' payoffs—from Selten's notion of *trembling hand perfection* [32], to Myerson's concept of *properness* [26], and the various criteria of strategic stability introduced by Kohlberg & Mertens [20] (hyperstability, full stability, sequential stability, etc.).

Dually to the above theory of "strategic refinement", an important alternative approach has been based on *dynamic* considerations: that is, the players of a game start off-equilibrium, and in one sense or another learn (or fail to learn) to play an equilibrium over time. Here, the focus is on the players' learning protocol, the information available during play, and the presence (or absence) of players that may deviate from this protocol. By the so-called "folk theorem of evolutionary game theory" [18], it is well known that only *strict* equilibria are stable and attracting under the replicator dynamics, a result which was extended more recently to a broad class of "regularized learning" schemes, in both continuous [14] and discrete time [15, 16].

These two viewpoints are not always compatible: for instance, in 2×2 games with two pure equilibria and one mixed (such as the Chicken / Hawk-Dove game), the mixed equilibrium is ruled out by almost all game-theoretic learning algorithms and dynamics, even though it survives a broad range of strategic refinement attacks. A point of hope here is the equivalence between (setwise) strategic and dynamic stability proved by Ritzberger & Weibull [28], who showed that a span of pure strategies in the mixed extension of a finite game is strategically stable in the sense of Kohlberg & Mertens [20] if and only if it is asymptotically stable under the replicator dynamics—see also [10] for an extension to a wider class of discrete-time models for learning, with different information assumptions.

Notably, these considerations all concern finite games in normal (or extensive) form. By contrast, most applications of game theory to machine learning and data science involve *continuous games*, that is, games with a finite number of players and a continuum of actions per player—for example, GANs, multi-agent reinforcement learning, Kelly auctions, etc. In view of this, our paper seeks to answer the following questions in the context of continuous games:

Which equilibria remain robust to arbitrary, small perturbations of the underlying game?
Which equilibria remain robust to small perturbations in the initialization of the players' dynamics?

We refer to these two questions as *strategic* and *dynamic robustness*, respectively, and our paper seeks to quantify the interplay between the two.

Our contributions in the context of related work. Aiming for the strongest possible definition of robustness, we propose the following strategic refinement criterion: an equilibrium of a continuous game is *strategically robust* if it remains an equilibrium in any slightly perturbed, nearby game. This requirement is similar in spirit to—but considerably stronger than—the classical notion of *essentiality* of Wu & Jiang [41], which posits that any nearby game has a nearby, possibly different equilibrium. Importantly, our results apply to *local* Nash equilibria, which are especially relevant in machine learning applications where payoff landscapes are typically nonconcave. This distinction is crucial, as global Nash equilibria do not always exist in general continuous games, making local equilibrium guarantees both meaningful and necessary in practice.

In contrast to finite games—where the notion of "nearby" is fairly unambiguous—perturbations to a continuous game involve functional variations and, as such, the choice of distance metric plays a crucial role. Importantly, albeit natural, our proposed robustness requirement becomes vacuous if distances are measured with respect to the players' payoff functions: more precisely, it is always possible to find a payoff perturbation with arbitrarily small L^∞-norm that ends up upsetting *any* equilibrium. The underlying issue here is that a small payoff perturbation may exhibit very high local variability, which can disrupt the first-order stationarity conditions that characterize equilibria in continuous games, thereby eliminating them altogether. To circumvent this issue, we argue that deviations of continuous games should be measured by comparing their respective gradient fields, which encode all the strategic information in the game. This shift in perspective leads to a crisp geometric characterization of strategically robust equilibria: they are extreme points of the game's action space, and they are sharp in the sense that the game's individual payoff gradients form a strictly acute angle with any tangent direction (cf. Fig. 1 later in the paper).

From a dynamic standpoint, we focus throughout on the *"follow the regularized leader"* (FTRL) family of algorithms [23, 33–35], arguably one of the most—if not *the* most—popular class of policies for online learning due to its strong regret minimization and convergence guarantees, and containing as special cases gradient descent/ascent methods [3, 44], dual averaging [27, 42], the exponential / multiplicative weights algorithm [2, 5, 5, 24, 40], implicitly normalized forecasters [1, 4, 43], exponentiated gradient methods [7, 19, 37], and many iterative first-order learning schemes. In this general context, we examine which equilibria admit robust convergence guarantees as stable limit points of the dynamics of FTRL in the presence of randomness and uncertainty.

Our first main result in this setting is that *strategic robustness implies dynamic robustness*, i.e., any strategically robust equilibrium is stable and attracting with high probability under the dynamics of FTRL, for any choice of regularizer. Conversely, we show that the strategic robustness requirement cannot be lifted, and we provide an example of a game with an extreme, non-robust equilibrium which attracts *all* FTRL orbits under a certain choice of regularizer, and *none* under another.

To the best of our knowledge, this is the first result of its kind for continuous games. In the context of *finite* games, Flokas et al. [14] showed that a point is asymptotically stable under the continuous-time FTRL dynamics if and only if it is a strict Nash equilibrium, while [10] extended this equivalence to discrete-time models of regularized learning under uncertainty. Strict equilibria are prime examples of strategically robust equilibria , so this part of the analysis of [10] is subsumed in ours. In the context of concave games—that is, continuous games with individually concave payoff functions—Mertikopoulos & Zhou [25] showed that sharp global equilibria enjoy comparable convergence guarantees under FTRL with a vanishing step-size. While such step-size schedules are effective at suppressing noise in the long run, they do so at the cost of significantly slowing down the algorithm's convergence. By contrast, we focus on fast, *constant* step-size schedules, which are widely used in practice due to their simplicity and often superior empirical performance. In this regime, we show that entropically regularized learning with a constant step-size converges to robust equilibria at a geometric rate, compared to distinctly subgeometric rate in the case of vanishing step-size policies.

2 Preliminaries

We start by briefly reviewing some basics of game theory and regularized learning, introducing the necessary context for our results.

2.1. The game-theoretic framework. Throughout our paper, we focus on a class of *continuous games* consisting of a finite set of $i \in \mathcal{N} = \{1, \dots, N\}$, and defined by the following primitives:

1. Each player $i \in \mathcal{N}$ has access to a compact convex set $\mathcal{X}_i$ of a finite dimensional vector space $\mathcal{V}_i$, describing the set of *actions* available to said player. By $\mathcal{X} \coloneqq \prod_i \mathcal{X}_i$ we denote the space of all ensembles $x = (x_1, \dots, x_N)$ of actions $x_i \in \mathcal{X}_i$ that are independently chosen by each player $i \in \mathcal{N}$. We will also write $x = (x_i; x_{-i})$ to emphasize the action of player $i \in \mathcal{N}$ against the joint action profile $x_{-i} \equiv (x_j)_{j \neq i}$ of all other players.
2. The players' rewards are determined by their individual *payoff functions* $u_i \colon \mathcal{X} \to \mathbb{R}$, assumed to be continuously differentiable for all $i \in \mathcal{N}$. Denoting by $\mathcal{Y}_i \equiv \mathcal{V}_i^*$ the dual space of $\mathcal{V}_i$, we define the individual gradient vector $v_i \colon \mathcal{X} \to \mathcal{Y}_i$ of player $i \in \mathcal{N}$ by
$$v_i(x) = \nabla_{x_i} u_i(x_i; x_{-i}) \tag{1}$$
and the ensemble $v(x) = (v_1(x), \dots, v_N(x)) \in \mathcal{Y} \equiv \prod_{i \in \mathcal{N}} \mathcal{Y}_i$ thereof.

A *continuous game* is then defined as a tuple $\mathcal{G} \equiv \mathcal{G}(\mathcal{N}, \mathcal{X}, u)$ with players, actions and payoff functions as above.

Nash equilibrium. The best known solution concept in game theory is that of a *Nash equilibrium* (NE), which characterizes the actions $x^* \in \mathcal{X}$ from which no player has incentive to deviate unilaterally. Formally, $x^* \in \mathcal{X}$ is a Nash equilibrium if
$$u_i(x^*) \geq u_i(x_i; x^*_{-i}) \quad \text{for all } x_i \in \mathcal{X}_i,\, i \in \mathcal{N}. \tag{NE}$$

A game $\mathcal{G} \equiv \mathcal{G}(\mathcal{N}, \mathcal{X}, u)$ always admits a Nash equilibrium if $\mathcal{X}$ is compact and each player's payoff function u_i is *individually concave* in the sense that $u_i(x_i; x_{-i})$ is concave in x_i for all $x_{-i} \in \mathcal{X}_{-i}$ [13, 31]. In this case, basic arguments from convex analysis [29, 30] show that x^* is an equilibrium of $\mathcal{G}$ if and only if it satisfies the (Stampacchia) variational inequality

$$\langle v(x^*), x - x^* \rangle \leq 0 \quad \text{for all } x \in \mathcal{X}. \tag{VI}$$

If the players' functions are not individually concave, a game may not admit a Nash equilibrium. In that case, it is more meaningful to consider *local* Nash equilibria, i.e., profiles $x^* \in \mathcal{X}$ such that

$$u_i(x^*) \geq u_i(x_i; x^*_{-i}) \quad \text{for all } x \text{ in a neighborhood } \mathcal{U} \text{ of } x^* \text{ in } \mathcal{X}. \tag{LNE}$$

In stark contrast to games with individually concave payoff functions, (VI) no longer characterizes local Nash equilibria: specifically, by first-order stationarity, we have (LNE) $\implies$ (VI) but the converse need not hold; in fact, a solution x^* of (VI) may be a global payoff *maximizer* for all $i \in \mathcal{N}$.

Note. In the sequel, we will work with general continuous games that may not admit a global equilibrium—but admit *local* Nash equilibria. To streamline our presentation, we will use the term "equilibrium" without any further qualification to refer to *local* equilibria, and we will say explicitly "global equilibria" for profiles satisfying (NE).

2.2. Regularized learning in games. The most widely used framework for learning in games, is the so called *"follow the regularized leader"* (FTRL) template, primarily because it leads to no regret in a wide variety of settings [34, 35]. The corresponding update rule hinges on the notion of a *regularized best response*, and proceeds as

$$y_{n+1} = y_n + \gamma \hat{v}_n, \qquad x_n = Q(y_n) \qquad \text{for } n = 1, 2, \dots \tag{FTRL}$$

where (i) $x_n \in \mathcal{X}$ denotes the players' action profile at step n; (ii) $y_n = \left(y_{i,n}\right)_{i \in \mathcal{N}} \in \mathcal{Y}$ is an auxiliary process that aggregates historical feedback into a compact state representation, i.e., a proxy for the players' empirical performance up to time n; (iii) $\hat{v}_n = \left(\hat{v}_{i,n}\right)_{i \in \mathcal{N}} \in \mathcal{Y}$ denotes the current gradient-like payoff signal; (iv) $\gamma > 0$ is the learning rate, or step-size parameter of the process; and (v) $Q \colon \mathcal{Y} \to \mathcal{X}$ is a mapping between the auxiliary process on the dual space $\mathcal{Y}$, and the players' strategy space $\mathcal{X}$. In what follows, we analyze the key components of this framework.

The algorithm's step-size. Throughout this work, we adopt a *constant* step-size routine. This stands in contrast to the stochastic approximation literature [8, 11, 21], where (FTRL) is typically implemented with a *vanishing* step-size satisfying the classical summability conditions $\sum_n \gamma_n = \infty$, $\sum_n \gamma_n^2 < \infty$. Such schedules are known to promote convergence by gradually suppressing the effect of noise [25]. In contrast, constant step-size methods are widely favored in practice for their simplicity, reduced hyperparameter sensitivity, and improved robustness in real-world applications

The mirror map. A central ingredient of regularized learning is the mirror map $Q \equiv (Q_i)_{i \in \mathcal{N}}$, with each $Q_i \colon \mathcal{Y}_i \to \mathcal{X}_i$ induced by a strongly convex *regularizer* $h_i \colon \mathcal{X}_i \to \mathbb{R}$ that promotes stability during the learning process. To streamline our presentation and letting $h(x) = \sum_{i \in \mathcal{N}} h_i(x_i)$, the players' *mirror map* is defined as

$$Q(y) \coloneqq \arg\max\nolimits_{x \in \mathcal{X}} \{\langle y, x \rangle - h(x)\} \tag{2}$$

In the rest of our paper, we will write $\mathcal{X}_h = \operatorname{im} Q$ for the image of $\mathcal{Y}$ under Q—and, likewise, $\mathcal{X}_{h_i} = \operatorname{im} Q_i$ for each player $i \in \mathcal{N}$. In particular, if Q is interior-valued—that is, $\mathcal{X}_h = \operatorname{ri} \mathcal{X}$—we will say that h is *steep* because, in this case, the (sub)gradients of h explode to infinity as $x \to \operatorname{bd} \mathcal{X}$ (i.e., h becomes "infinitely steep"); instead, if $\operatorname{im} Q = \mathcal{X}$, we will say that h is *non-steep*. For a precise version of this discussion, see Appendix A.

Different choices of the regularizer h induce different projection-like operations, adapted to the geometry of the underlying space. We describe two mainstay examples below.

Example 2.1 (Euclidean projection)**.** The quadratic regularizer $h(x) = \|x\|_2^2/2$ gives rise to the Euclidean projection $Q(y) = \operatorname{proj}_{\mathcal{X}}(y) = \arg\min_{x \in \mathcal{X}} \|y - x\|_2$. In this case, h is non-steep. ❦

Example 2.2 (Multiplicative weights)**.** For $\mathcal{A}_i$ a *finite* set of actions per player $i \in \mathcal{N}$, and $\mathcal{X}_i \equiv \Delta(\mathcal{A}_i)$, the entropic regularizer $h_i(x_i) = \sum_{\alpha_i \in \mathcal{A}_i} x_{i\alpha_i} \log x_{i\alpha_i}$ gives rise to the logit map, defined via $Q_i(y_i) = \exp(y_i)/\|\exp(y_i)\|_1$, where $\exp(y_i)$ denotes the element-wise exponential of y_i. ❦

The feedback process. Throughout this work, we consider two distinct feedback models: (i) *stochastic gradients*; and (ii) *payoff-based* feedback. We describe both frameworks below.

Stochastic gradient feedback. At every time step n, each player $i \in \mathcal{N}$ has access to a *stochastic first-order oracle* (SFO) – that is, a noisy version of their individual gradient vector of the form:

$$\hat{v}_n = v(x_n) + U_n \quad \text{with} \quad \mathbb{E}[U_n \mid \mathcal{F}_n] = 0 \tag{SFO}$$

where U_n is zero-mean and conditionally sub-Gaussian given the information $\mathcal{F}_n$ generated up to time $n \in \mathbb{N}$. In other words, players observe unbiased estimates of their individual gradient vectors.

Payoff-based feedback. Unlike the (SFO) model where players have access to a black-box oracle that provides noisy gradient information, it is often more realistic to consider a payoff-based feedback paradigm where players observe *only* their realized payoffs – that is, a single scalar value – and have to reconstruct an estimate of their individual gradient vectors.

The most widely used method in this setting is the *single-point stochastic approximation* (SPSA) framework of [12, 36], which is based on finite differences along randomly sampled directions. Specifically, denoting the set of unit directions $\mathcal{E}_i \coloneqq \{\pm e_1, \dots, \pm e_{d_i}\}$ that span the affine hull of $\mathcal{X}_i$ of dimension d_i, each player $i \in \mathcal{N}$ draws a direction $w_{i,n} \in \mathcal{E}_i$ uniformly at random in every round $n \in \mathbb{N}$. Since the perturbed action $x_{i,n} + \varepsilon_n w_{i,n}$ may lie outside $\mathcal{X}_i$ for a perturbation radius $\varepsilon_n > 0$, we introduce a pivot element $p_i \in \operatorname{ri}(\mathcal{X}_i)$ and a radius $r_i > \varepsilon_n$ such that $p_i + r_i w_i \in \mathcal{X}_i$ for all $w_i \in \mathcal{E}_i$. Based on these, we define the feasibility-adjusted action $x^{\varepsilon}_{i,n} \coloneqq x_{i,n} + (\varepsilon_n/r_i)(p_i - x_{i,n}) \in \mathcal{X}_i$. Finally, each player queries the perturbed action $\hat{x}_{i,n} \equiv x^{\varepsilon}_{i,n} + \varepsilon_n w_{i,n}$ which is an element of $\mathcal{X}_i$, and observes the realized payoff value $u_i(\hat{x}_n)$.[2] The gradient vector is, then, estimated via the single-point stochastic approximation scheme:

$$\hat{v}_{i,n} \coloneqq (d_i/\varepsilon_n)\, u_i(\hat{x}_n)\, w_{i,n} \tag{SPSA}$$

Importantly, the feasibility adjustment ensures that the perturbed action $\hat{x}_n$ remains within the players' action set $\mathcal{X}$, while preserving the direction of the original perturbation w_n. As we show in Appendix A, (SPSA) enjoys the bounds

$$\|\mathbb{E}[\hat{v}_n \mid \mathcal{F}_n] - v(x_n)\|_* = \mathcal{O}(\varepsilon_n) \quad \text{and} \quad \|\hat{v}_n\|_* = \mathcal{O}(1/\varepsilon_n)\,. \tag{3}$$

With all these in hand, we are now ready to present the main contributions of this work.

3 Strategic robustness: Geometric and variational characterization

In this section, we address the strategic aspects of the equilibrium robustness question, and ask:

Which equilibria remain equilibria after a small—but otherwise arbitrary—perturbation of the game?

We take this desideratum as the starting point for our definition of *strategic robustness*, that is, action profiles that remain (local) equilibria under small disturbances in the underlying game. This leads to a delicate interplay between the variational and geometric aspects of the underlying game, which we detail later in this section.

3.1. A first approach and insights. A natural way to measure the distance between two concave games, $\mathcal{G} \equiv \mathcal{G}(\mathcal{N}, \mathcal{X}, u)$ and $\widetilde{\mathcal{G}} \equiv \mathcal{G}(\mathcal{N}, \mathcal{X}, \tilde{u})$, would be via the uniform distance

$$\rho(\mathcal{G}, \widetilde{\mathcal{G}}) \coloneqq \max\nolimits_{i \in \mathcal{N}} \sup\nolimits_{x \in \mathcal{X}} |u_i(x) - \tilde{u}_i(x)|\,. \tag{4}$$

Intuitively, if this quantity is small enough, the two games are nearly indistinguishable from a strategic perspective, since for every strategy profile $x \in \mathcal{X}$, the payoffs in $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ are almost the same. Thus, one might expect that at least some equilibria of $\mathcal{G}$ would persist under sufficiently small perturbations, especially since a Nash equilibrium appears to be a property of the payoff functions themselves.

[2] Since $r_i > \varepsilon_n$, we write $x^{\varepsilon}_{i,n} = x_{i,n}(1 - \varepsilon_n/r_i) + (\varepsilon_n/r_i)p_i$ which is a convex combination of points in $\mathcal{X}_i$. Regarding $\hat{x}_{i,n}$, note it can be written as $\hat{x}_{i,n} = x_{i,n}(1 - \varepsilon_n/r_i) + (\varepsilon_n/r_i)(p_i + r_i w_{i,n})$, which is also a convex combination of points in $\mathcal{X}_i$. Thus, both belong to $\mathcal{X}_i$.

Surprisingly, as the next examples illustrate, this definition of distance *cannot* provide a meaningful concept of equilibrium robustness. Namely, for *any* game $\mathcal{G}$ with equilibrium $x^* \in \mathcal{X}$, there exists a perturbed game $\widetilde{\mathcal{G}}$, *arbitrarily close* to $\mathcal{G}$ (in the ρ-metric) such that $x^* \in \mathcal{X}$ is no longer an equilibrium of $\widetilde{\mathcal{G}}$. To show this, we provide Examples 3.1 and 3.2 which together cover all possible types of equilibria in continuous games in the sense of (VI).

Example 3.1. Let $\mathcal{G}$ be a continuous game, and $x^* \in \mathcal{X}$ be an equilibrium point, such that $\langle v_i(x^*), p_i - x_i^* \rangle < 0$ for some player $i \in \mathcal{N}$ and $p_i \in \mathcal{X}_i$. For arbitrary $\varepsilon > 0$, define $\tilde{u}_i : \mathcal{X} \to \mathbb{R}$ as

$$\tilde{u}_i(x) \coloneqq u_i(x) - \varepsilon \exp\Big(2\,\varepsilon^{-1} \langle v_i(x^*), x_i - x_i^* \rangle\Big) \tag{5}$$

which is a continuously differentiable concave function in x_i, and let $\tilde{u}_j \equiv u_j$ for all $j \neq i$, $j \in \mathcal{N}$. Since $x^* \in \mathcal{X}$ is an equilibrium of $\mathcal{G}$, it holds $\langle v_i(x^*), x_i - x_i^* \rangle \le 0$ for all $x_i \in \mathcal{X}_i$, which implies that

$$\rho\big(\mathcal{G}, \widetilde{\mathcal{G}}\big) = \sup\nolimits_{x \in \mathcal{X}} |u_i(x) - \tilde{u}_i(x)| = \varepsilon \sup\nolimits_{x_i \in \mathcal{X}_i} \exp\big(2\,\varepsilon^{-1} \langle v_i(x^*), x_i - x_i^* \rangle\big) = \varepsilon. \tag{6}$$

Computing the individual gradient vector of player $i \in \mathcal{N}$, we obtain

$$\tilde{v}_i(x) = v_i(x) - 2\, v_i(x^*) \exp\Big(2\,\varepsilon^{-1} \langle v_i(x^*), x_i - x_i^* \rangle\Big) \tag{7}$$

and, evaluating it at $x^* \in \mathcal{X}$, we get, $\tilde{v}_i(x^*) = -v_i(x^*)$. Therefore, for $x = (p_i; x_{-i}^*) \in \mathcal{X}$, we have

$$\langle \tilde{v}(x^*), x - x^* \rangle = -\langle v_i(x^*), p_i - x_i^* \rangle > 0 \tag{8}$$

i.e., $x^* \in \mathcal{X}$ is not an equilibrium point of the nearby game $\widetilde{\mathcal{G}}$. ❦

Example 3.2. Let $\mathcal{G}$ be a continuous game, and $x^* \in \mathcal{X}$ an equilibrium point, such that $\langle v(x^*), x - x^* \rangle = 0$ for all $x \in \mathcal{X}$. Fix a player $i \in \mathcal{N}$ and $p_i \in \mathcal{X}_i$, and let $y_i \in \mathcal{V}_i^*$ with $\langle y_i, p_i - x_i^* \rangle > 0$. For arbitrary $\varepsilon > 0$, let $\tilde{u}_i : \mathcal{X} \to \mathbb{R}$ be defined as

$$\tilde{u}_i(x) \coloneqq u_i(x) + \varepsilon \operatorname{diam}(\mathcal{X}_i)^{-1} \|y_i\|_*^{-1} \langle y_i, x_i - x_i^* \rangle \tag{9}$$

which is a concave function in x_i, and $\tilde{u}_j \equiv u_j$ for all $j \neq i$, $j \in \mathcal{N}$. Then, we readily get that

$$\rho\big(\mathcal{G}, \widetilde{\mathcal{G}}\big) = \sup\nolimits_{x \in \mathcal{X}} |u_i(x) - \tilde{u}_i(x)| = \varepsilon \operatorname{diam}(\mathcal{X}_i)^{-1} \|y_i\|_*^{-1} \sup\nolimits_{x_i \in \mathcal{X}_i} |\langle y_i, x_i - x_i^* \rangle| \le \varepsilon. \tag{10}$$

Computing the individual gradient vector of player $i \in \mathcal{N}$, we obtain

$$\tilde{v}_i(x) = v_i(x) + \varepsilon \operatorname{diam}(\mathcal{X}_i)^{-1} \|y_i\|_*^{-1}\, y_i \tag{11}$$

Therefore, for $x = (p_i; x_{-i}^*) \in \mathcal{X}$, it holds by the example's assumptions that

$$\langle \tilde{v}(x^*), x - x^* \rangle = \langle v_i(x^*), p_i - x_i^* \rangle + \varepsilon \operatorname{diam}(\mathcal{X}_i)^{-1} \|y_i\|_*^{-1} \langle y_i, p_i - x_i^* \rangle > 0 \tag{12}$$

where we used that $\langle v_i(x^*), p_i - x_i^* \rangle = 0$ and $\langle y_i, p_i - x_i^* \rangle > 0$ by our assumptions. Thus, $x^* \in \mathcal{X}$ is not an equilibrium of the perturbed game $\widetilde{\mathcal{G}}$. ❦

Remark 1. In Examples 3.1 and 3.2, if $\mathcal{G}$ is concave, so is $\widetilde{\mathcal{G}}$, indicating that this notion of distance is not proper even within the class of concave games.

The preceding examples demonstrate that under the distance (4), even an arbitrarily small perturbation to the payoff function of a *single* player can destroy *any* equilibrium[3]. This phenomenon arises because, although an equilibrium is defined in terms of payoff functions, the first-order stationarity condition in (VI) shows that it fundamentally depends on the individual gradient vectors. Therefore, any meaningful notion of distance between two games must likewise be aware of the behavior of the individual gradient vectors.

3.2. Defining the notion of strategic robustness. As illustrated in Examples 3.1 and 3.2, small changes in the payoffs, though negligible in the uniform norm, can significantly alter the equilibrium landscape. To address this, we refine the notion of distance between games $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ as follows:

$$\operatorname{dist}\big(\mathcal{G}, \widetilde{\mathcal{G}}\big) \coloneqq \sup\nolimits_{x \in \mathcal{X}} \|v(x) - \tilde{v}(x)\|_* \tag{13}$$

With this definition in hand, we are now ready to state the concept of strategic robustness in the class of continuous games.

Definition 1. An equilibrium $x^* \in \mathcal{X}$ of a game $\mathcal{G}$ is called *strategically robust* if there exists $\varepsilon > 0$ such that for *any* game $\widetilde{\mathcal{G}}$ with $\operatorname{dist}\big(\mathcal{G}, \widetilde{\mathcal{G}}\big) < \varepsilon$, x^* is also an equilibrium of $\widetilde{\mathcal{G}}$.

As we explore next, this definition offers a meaningful notion of "closeness" for equilibrium stability, one that is grounded not in the payoff values themselves, but in the geometry they induce.

[3]Such variations are not possible in the class of finite games, so, in this much more restrictive class, the sup-norm of the payoff differences is a valid metric

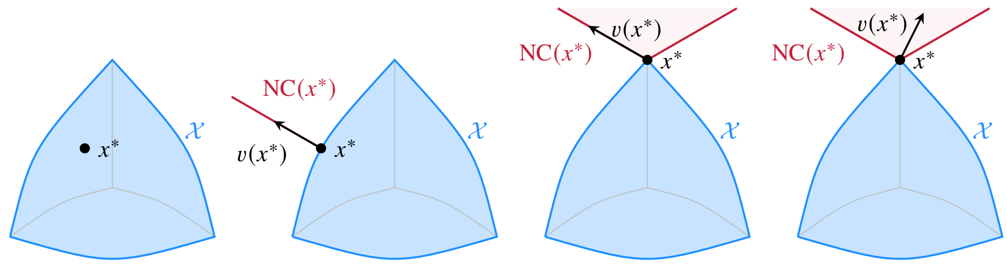

Figure 1: Different equilibrium configurations: an interior equilibrium ($v(x^*) = 0$); a boundary, non-extreme equilibrium (normal cone with empty topological interior); an extreme, non-robust equilibrium ($v(x^*)$ on the boundary of the normal cone); a robust equilibrium ($v(x^*)$ in the interior of the normal cone). Only the robust equilibrium remains invariant under strategic perturbations of the underlying game.

Geometric characterization. To provide a geometric characterization, we zoom in on the variational structure that governs Nash equilibria. Specifically, we show that strategically robust equilibria $x^* \in \mathcal{X}$ are precisely those solutions of (VI) for which the inequality is strict for all feasible deviations. Formally, we have the following characterization:

Theorem 1. *Let $x^* \in \mathcal{X}$ be a joint action profile in $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$. Then the following are equivalent:*

(i) *x^* is a strategically robust equilibrium.*

(ii) *$\langle v(x^*), z\rangle \leq -m\|z\|$ for some $m > 0$ and all $z \in \mathrm{TC}(x^*)$, where $\mathrm{TC}(x^*)$ is the closure of all rays emanating from x^* and intersecting $\mathcal{X}$ in at least one other point.*

(iii) *$v(x^*) \in \mathrm{int}(\mathrm{PC}(x^*))$, where $\mathrm{PC}(x^*) \coloneqq \{y \in \mathcal{Y} : \langle y, z\rangle \leq 0, \text{ for all } z \in \mathrm{TC}(x^*)\}$.*

Intuitively, Theorem 1 suggests that strategically robust equilibria are precisely the points $x^* \in \mathcal{X}$ whose gradient vector $v(x^*)$ lies in the topological interior of the polar cone $\mathrm{PC}(x^*)$, i.e., $\langle v(x^*), z\rangle < 0$ for all $z \in \mathrm{TC}(x^*)$. We, thus, conclude that strategic robustness can only occur at boundary points where the tangent cone is pointed; if the feasible set is locally flat at $x^* \in \mathcal{X}$, the corresponding polar cone has empty interior, and robustness is not possible. This phenomenon is illustrated in Fig. 1, and the full proof of Theorem 1 is provided in Appendix B.

Remark 2. Both Examples 3.1 and 3.2 violate the condition in Definition 1, but for different reasons. In the former case, although the perturbed payoffs can be made arbitrarily close to the original, the perturbed gradient vector at the equilibrium can become arbitrarily large, making the distance $\mathrm{dist}\big(\mathcal{G}, \widetilde{\mathcal{G}}\big)$ exceed any $\varepsilon > 0$. In the latter case, the polar cone $\mathrm{PC}(x^*)$ at the equilibrium has empty interior, so strategic robustness cannot hold at x^*.

In the next section, we examine the dynamic implications of this result by studying the robustness of such equilibria under (FTRL).

4 From strategic to dynamic robustness: Convergence results

So far, we focused on strategic robustness, a static notion determined solely by the underlying structure of the game and the local geometry around the equilibria. In this section, we shift to the dynamic perspective of our central question and explore which equilibria admit robust convergence guarantees, namely, equilibria that can emerge as stable outcomes of regularized learning under feedback uncertainty, regardless of the specific choice of regularizer.

To this end, we first establish that non-equilibrium points cannot arise as limit points of the (FTRL) dynamics, even under perfect gradient feedback. Formally, we have the following proposition, whose proof is provided in Appendix C.

Proposition 1. *Let $x' \in \mathcal{X}$ be an non-equilibrium point of $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$. Then, x' cannot be a limit point of (FTRL) dynamics with $\hat{v}_n \equiv v(x_n)$ for all $n \in \mathbb{N}$.*

Having excluded non-equilibrium points as positive probability outcomes of a learning process, we now turn to identifying equilibria that are robust from a dynamic standpoint, and more precisely, under that of (FTRL). In this regard, strategically robust equilibria serve as natural candidates, as their stability with respect to game perturbations suggests they may also admit robust convergence guarantees. This is further supported by the finding that equilibrium points in the interior of the strategy space $\mathcal{X}$ cannot be limit points: in particular, we show below that, even under i.i.d. stochastic noise, the iterates of (FTRL) diverge from such equilibria almost surely.

Proposition 2. *Let $x^* \in \operatorname{ri}(\mathcal{X})$ be a Nash equilibrium of $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$, and $(x_n)_{n\in\mathbb{N}}$ be the sequence of play induced by* (FTRL) *with $\hat{v}_n = v(x_n) + U_n$, where U_n i.i.d. with $\mathbb{E}[U_n] = 0$ and $\operatorname{cov}(U_n) \succ 0$ for all $n \in \mathbb{N}$. Then:*

$$\mathbb{P}\Big(\lim_{n\to\infty} x_n = x^*\Big) = 0 \qquad \text{for any } x_1 \in \mathcal{X}_h. \tag{14}$$

Remark 3. The condition $\operatorname{cov}(U_n) \succ 0$ is not necessary. It suffices that $\operatorname{cov}(U_n)$ is non-degenerate in a direction $p - x^*$ for $p \in \mathcal{X}$; however, we state the stronger assumption for simplicity.

The key idea of the proof, which is deferred to Appendix C, is that, since $x^* \in \operatorname{ri}(\mathcal{X})$, we have $\langle v(x^*), x - x^*\rangle = 0$ for all $x \in \mathcal{X}$. At the same time, as $\operatorname{cov}(U_n) \succ 0$, the quantity $\langle U_n, x - x^*\rangle$ fluctuates and remains bounded away from zero infinitely often, thereby preventing convergence.

4.1. Learning with gradient-based feedback. In view of the impossibility result of Proposition 2, we shift our focus on the convergence of (FTRL) toward strategically robust equilibria. We first consider the gradient feedback model, where each player receives an unbiased estimate of their individual gradient vector via (SFO). Specifically, we analyze the behavior of (FTRL) and we establish local convergence guarantees toward strategically robust equilibria with high probability. This is encoded in the following theorem:

Theorem 2. *Let $x^* \in \mathcal{X}$ be a strategically robust equilibrium of $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$. Fix a confidence level $\delta > 0$, and let $(x_n)_{n\in\mathbb{N}}$ be the iterates of* (FTRL) *with feedback provided by* (SFO)*, and step-size $\gamma > 0$ sufficiently small. Then, there exists a neighborhood $\mathcal{U}$ of x^* in $\mathcal{X}_h$ such that:*

$$\mathbb{P}\Big(\lim_{n\to\infty} x_n = x^*\Big) \geq 1 - \delta \quad \text{if } x_1 \in \mathcal{U}. \tag{15}$$

Before proceeding, a few remarks are in order. Since continuous games may admit multiple Nash equilibria, global convergence guarantees are in general unattainable. As such, our analysis focuses on the local convergence landscape of (FTRL). As a sidenote, it is important to emphasize that the convergence result is robust to the choice of regularizer, relying solely on the general conditions outlined in Section 2 rather than any particular functional form.

While our main focus lies on the qualitative convergence behavior of (FTRL), stronger guarantees can be obtained under additional structural assumptions on the strategy space and the regularizer. In particular, suppose that $\mathcal{X}$ is a polyhedral domain of the form $\mathcal{X} \coloneqq \big\{x \in \mathbb{R}^d_+ \mid Ax = b\big\}$ for some $A \in \mathbb{R}^{m\times d}$ and $b \in \mathbb{R}^m$, and h is decomposable with kernel function θ, i.e., h can be written as $h(x) = \sum_{j=1}^d \theta(x_j)$ for some continuous function $\theta : \mathbb{R}_+ \to \mathbb{R}$ with locally Lipschitz θ'' and $\theta'' > 0$. Under these conditions, we obtain the explicit convergence rates for the (FTRL) dynamics, as follows.

Theorem 3. *If, in addition, $\mathcal{X}$ is a polyhedral domain and h is decomposable with kernel θ, on the event $E \coloneqq \{\lim_{n\to\infty} x_n = x^*\}$ it holds:*

$$\|x_n - x^*\| = \phi(-\Theta(n)) \tag{16}$$

where ϕ is the rate function defined via

$$\phi(z) \coloneqq \begin{cases} (\theta')^{-1}(z) & \text{if } z > \theta'(0^+) \\ 0 & \text{if } z \leq \theta'(0^+) \end{cases} \tag{17}$$

Remark 4. For the setting of Example 2.2, with $\mathcal{X} = \Delta(\mathcal{A})$, $\theta(z) = z \log z$ and x^* a *strict* Nash equilibrium, the convergence rate of (FTRL) as per Theorem 3, becomes $\|x_n - x^*\| = \exp(-\Theta(n))$.

Remark 5. For *finite games*, [16] showed that under a step-size schedule of the form $\gamma_n \propto 1/n^p$, the Robins-Monro summability conditions require $p \in (1/2, 1]$, leading to convergence rates from $\phi\big(-\Theta(n^{1-p})\big)$ to $\phi(-\Theta(\log n))$.

4.2. Learning with payoff-based feedback. We now turn to the payoff-based feedback model, where players observe only their realized payoffs and use them to estimate gradients indirectly via (SPSA). This feedback model introduces higher variance and structural bias due to the diminishing sampling radius and the feasibility corrections. Nevertheless, we show that strategically robust equilibria retain their dynamic robustness: they still locally attract the (FTRL) dynamics with high probability, as the following theorem suggests.

Theorem 4. *Let $x^* \in \mathcal{X}$ be a strategically robust equilibrium of $\mathcal{G}$. Fix a confidence level $\delta > 0$, and let $(x_n)_{n\in\mathbb{N}}$ be the iterates of (FTRL) run with (SPSA) with $\varepsilon_n \propto 1/n^p$ for some $p \in (0, 1/2)$ and step-size $\gamma > 0$ sufficiently small. Then, there exists a neighborhood $\mathcal{U}$ of x^* such that:*

$$\mathbb{P}\Big(\lim_{n\to\infty} x_n = x^*\Big) \geq 1 - \delta \qquad \text{if } x_1 \in \mathcal{U}. \tag{18}$$

If, in addition, $\mathcal{X}$ is affinely constrained and h is decomposable with kernel θ, then, whenever x_n converges to x^*, we have:

$$\|x_n - x^*\| = \phi(-\Theta(n))\,. \tag{19}$$

Despite the scarcity of information inherent in the payoff-based feedback model, strategically robust equilibria retain not only their convergence properties but also their convergence speed, under the additional structural assumptions on the regularizer and domain, matching that of the (SFO) feedback setting. This is further discussed along with the proof of the theorem in Appendix C.

4.3. Convergence landscape beyond strategic robustness. Having established the robust convergence properties of strategically robust equilibria, a natural question arises: *Can we expect robust convergence guarantees toward equilibria that lack this structural property?* As we show below, the answer is not encouraging: strategic robustness is essentially necessary for robust convergence.

To make this limitation precise, we move beyond the interior of the strategy space, where Proposition 2 rules out equilibria as potential limit points, and shift our focus to non-robust equilibria on the boundary. To illustrate the behavior of (FTRL) in this setting, we construct a game with a unique equilibrium that exhibits fundamentally different long-run behavior depending on the regularizer.

Proposition 3. *Consider the 1-player game $\mathcal{G}$ with $\mathcal{X} = [0, 1]$, $u(x) = -\frac{3}{4}x^{4/3}$ and $x^* = 0$. Let $(x_n)_{n\in\mathbb{N}}$ be the iterates of (FTRL) with $\gamma < 1$, and $\hat{v}_n = v(x_n) + U_n$, where U_n are i.i.d. standard normal random variables for all $n \in \mathbb{N}$. Then, for any initial condition $y_1 \in \mathbb{R}$, we have:*

(i) For $h(x) = x \log x$, it holds $\mathbb{P}(\lim_{n\to\infty} x_n = x^*) = 0$.

(ii) For $h(x) = -2\sqrt{x}$, it holds $\mathbb{P}(\lim_{n\to\infty} x_n = x^*) = 1$.

The core idea of the proof of Proposition 3 (which we present in detail in Appendix C) is to construct a process z_n that dominates y_n. Importantly, the process z_n can be then viewed as a random walk with a diminishing drift whose rate of decay depends on the choice of regularizer. Depending on the magnitude of this drift, the process exhibits two sharply contrasting long-term behaviors: if the drift decays sufficiently fast, the process behaves like a zero-mean random walk and returns infinitely often with probability 1 (recurrence); conversely, if the drift diminishes at a slower rate, the process behaves like a random walk with constant drift and escapes to infinity with probability 1 (transience).

In view of the above, we conclude that strategic robustness cannot be relaxed without compromising convergence guarantees, even when the equilibrium lies on the boundary.

5 Concluding remarks

Our aim in this paper was to examine the robustness of Nash equilibria in continuous games, under both strategic and dynamic uncertainty. From a strategic standpoint, we introduced the criterion of *strategic robustness* as those (local) equilibria which remain invariant under small perturbations of the underlying game, and we derived a tight geometric characterization thereof in terms of the variational geometry of the game. From a dynamic standpoint, we focused on the stability of regularized learning under uncertainty, and we established a deep structural connection between the two notions: strategic robustness guarantees dynamic robustness under (FTRL), and this implication is essentially tight: without strategic robustness, dynamic robustness cannot be ensured. To the best of our knowledge, this is the first result of its kind in the context of continuous game, and we find it particularly appealing for understanding the delicate interplay between the strategic and dynamic features of a game.

References

[1] Abernethy, J., Lee, C., and Tewari, A. Fighting bandits with a new kind of smoothness. In *NIPS '15: Proceedings of the 29th International Conference on Neural Information Processing Systems*, 2015.

[2] Arora, S., Hazan, E., and Kale, S. The multiplicative weights update method: A meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[3] Arrow, K. J., Hurwicz, L., and Uzawa, H. *Studies in linear and non-linear programming*. Stanford University Press, 1958.

[4] Audibert, J.-Y., Bubeck, S., and Lugosi, G. Minimax policies for combinatorial prediction games. In *COLT '11: Proceedings of the 24th Annual Conference on Learning Theory*, 2011.

[5] Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, 1995.

[6] Azizian, W., Iutzeler, F., Malick, J., and Mertikopoulos, P. On the rate of convergence of Bregman proximal methods in constrained variational inequalities. http://arxiv.org/abs/2211.08043, 2022.

[7] Beck, A. and Teboulle, M. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

[8] Benaïm, M. Dynamics of stochastic approximation algorithms. In Azéma, J., Émery, M., Ledoux, M., and Yor, M. (eds.), *Séminaire de Probabilités XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pp. 1–68. Springer Berlin Heidelberg, 1999.

[9] Billingsley, P. *Probability and measure*. A Wiley-Interscience publication. Wiley, 3. ed edition, 1995. ISBN 0471007102.

[10] Boone, V. and Mertikopoulos, P. The equivalence of dynamic and strategic stability under regularized learning in games. In *NeurIPS '23: Proceedings of the 37th International Conference on Neural Information Processing Systems*, 2023.

[11] Borkar, V. S. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press and Hindustan Book Agency, 2008.

[12] Bravo, M., Leslie, D. S., and Mertikopoulos, P. Bandit learning in concave N-person games. In *NeurIPS '18: Proceedings of the 32nd International Conference of Neural Information Processing Systems*, 2018.

[13] Debreu, G. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences of the USA*, October 1952.

[14] Flokas, L., Vlatakis-Gkaragkounis, E. V., Lianeas, T., Mertikopoulos, P., and Piliouras, G. No-regret learning and mixed Nash equilibria: They do not mix. In *NeurIPS '20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020.

[15] Giannou, A., Vlatakis-Gkaragkounis, E. V., and Mertikopoulos, P. Survival of the strictest: Stable and unstable equilibria under regularized learning with partial information. In *COLT '21: Proceedings of the 34th Annual Conference on Learning Theory*, 2021.

[16] Giannou, A., Vlatakis-Gkaragkounis, E. V., and Mertikopoulos, P. The convergence rate of regularized learning in games: From bandits and uncertainty to optimism and beyond. In *NeurIPS '21: Proceedings of the 35th International Conference on Neural Information Processing Systems*, 2021.

[17] Hall, P. and Heyde, C. C. *Martingale Limit Theory and Its Application*. Probability and Mathematical Statistics. Academic Press, New York, 1980.

[18] Hofbauer, J. and Sigmund, K. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4), July 2003.

[19] Kivinen, J. and Warmuth, M. K. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132(1):1–63, 1997.

[20] Kohlberg, E. and Mertens, J.-F. On the strategic stability of equilibria. *Econometrica*, 54(5):1003–1037, September 1986.

[21] Kushner, H. J. and Clark, D. S. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Springer, 1978.

[22] Lamperti, J. Criteria for the recurrence or transience of stochastic process. i. *Journal of Mathematical Analysis and Applications*, 1(3):314–330, 1960. ISSN 0022-247X. doi: https://doi.org/10.1016/0022-247X(60)90005-6. URL https://www.sciencedirect.com/science/article/pii/0022247X60900056.

[23] Lattimore, T. and Szepesvári, C. *Bandit Algorithms*. Cambridge University Press, Cambridge, UK, 2020.

[24] Littlestone, N. and Warmuth, M. K. The weighted majority algorithm. *Information and Computation*, 108 (2):212–261, 1994.

[25] Mertikopoulos, P. and Zhou, Z. Learning in games with continuous action sets and unknown payoff functions. *Mathematical Programming*, 173(1-2):465–507, January 2019.

[26] Myerson, R. B. Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, 7 (2):73–80, 1978.

[27] Nesterov, Y. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1): 221–259, 2009.

[28] Ritzberger, K. and Weibull, J. W. Evolutionary selection in normal-form games. *Econometrica*, 63(6): 1371–99, November 1995.

[29] Rockafellar, R. T. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[30] Rockafellar, R. T. and Wets, R. J. B. *Variational Analysis*, volume 317 of *A Series of Comprehensive Studies in Mathematics*. Springer-Verlag, Berlin, 1998.

[31] Rosen, J. B. Existence and uniqueness of equilibrium points for concave N-person games. *Econometrica*, 33(3):520–534, 1965.

[32] Selten, R. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.

[33] Shalev-Shwartz, S. *Online learning: Theory, algorithms, and applications*. PhD thesis, Hebrew University of Jerusalem, 2007.

[34] Shalev-Shwartz, S. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2), 2011.

[35] Shalev-Shwartz, S. and Singer, Y. Convex repeated games and Fenchel duality. In *NIPS' 06: Proceedings of the 19th Annual Conference on Neural Information Processing Systems*. MIT Press, 2006.

[36] Spall, J. C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Autom. Control*, 37(3):332–341, March 1992.

[37] Tsuda, K., Rätsch, G., and Warmuth, M. K. Matrix exponentiated gradient updates for on-line Bregman projection. *Journal of Machine Learning Research*, 6:995–1018, 2005.

[38] van Damme, E. *Stability and perfection of Nash equilibria*. Springer-Verlag, Berlin, 1987.

[39] von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[40] Vovk, V. G. Aggregating strategies. In *COLT '90: Proceedings of the 3rd Workshop on Computational Learning Theory*, pp. 371–383, 1990.

[41] Wu, W.-T. and Jiang, J.-H. Essential equilibrium points of N-person non-cooperative games. *Scientia Sinica*, 11(10):1307–1322, 1962.

[42] Xiao, L. Dual averaging methods for regularized stochastic learning and online optimization. *Journal of Machine Learning Research*, 11:2543–2596, October 2010.

[43] Zimmert, J. and Seldin, Y. Tsallis-INF: An optimal algorithm for stochastic and adversarial bandits. *Journal of Machine Learning Research*, 22(28):1–49, 2021.

[44] Zinkevich, M. Online convex programming and generalized infinitesimal gradient ascent. In *ICML '03: Proceedings of the 20th International Conference on Machine Learning*, pp. 928–936, 2003.

A Auxiliary results

As a preamble to our analysis, we provide some basic properties of the regularizers and the mirror maps, and present some auxiliary results from martingale theory and Markov processes that we will use throughout the sequel.

A.1. Mirror maps and results from convex analysis. In this section, we provide a more detailed discussion of key notions from convex analysis, including mirror maps and regularizers.

To begin, let $(\mathcal{V}, \|\cdot\|)$ be a finite-dimensional normed vector space. Its dual space is denoted by $(\mathcal{V}^*, \|\cdot\|_*)$, where the dual norm is defined as

$$\|y\|_* \equiv \max\{\langle y, x\rangle : \|x\| \leq 1\}, \tag{A.1}$$

and $\langle y, x\rangle$ denotes the canonical pairing between $y \in \mathcal{V}^*$ and $x \in \mathcal{V}$. To maintain consistency with the notation used throughout the paper, we will refer to $\mathcal{V}^*$ as $\mathcal{Y}$ from this point onward.

Given a closed convex set $\mathcal{X} \subseteq \mathcal{V}$ and a point $p \in \mathcal{X}$, we define the tangent cone $\mathrm{TC}(p)$ and the polar cone $\mathrm{PC}(p)$ as follows:

$$\mathrm{TC}(p) = \mathrm{cl}\{z \in \mathcal{V} : p + tz \in \mathcal{X} \text{ for some } t > 0\} \tag{A.2}$$

and

$$\mathrm{PC}(p) = \{y \in \mathcal{Y} : \langle y, z\rangle \leq 0, \text{ for all } z \in \mathrm{TC}(p)\} \tag{A.3}$$

For a strongly convex regularizer $h : \mathcal{X} \to \mathbb{R}$, the *subdifferential* of h at $x \in \mathcal{X}$ is defined as

$$\partial h(x) \coloneqq \{y \in \mathcal{Y} : h(x') \geq h(x) + \langle y, x' - x\rangle \text{ for all } x' \in \mathcal{X}\} \tag{A.4}$$

and we denote the *domain of subdifferentiability* of h as

$$\mathcal{X}_h = \{x \in \mathcal{X} : \partial h(x) \neq \varnothing\}\,. \tag{A.5}$$

In addition, the mirror map Q, defined via

$$Q(y) = \arg\max_{x \in \mathcal{X}}\{\langle y, x\rangle - h(x)\} \tag{A.6}$$

is single-valued on $\mathcal{Y}$, since the maximization problem admits a unique solution, as h is strongly convex. Finally, by the optimality conditions of (A.6), we get that

$$x = Q(y) \quad \text{if and only if} \quad y \in \partial h(x)\,. \tag{A.7}$$

since $0 \in y - \partial h(x)$. This readily implies that $\mathcal{X}_h = \operatorname{im} Q$. In general, we have

$$\mathrm{ri}(\mathcal{X}) \subseteq \mathcal{X}_h \subseteq \mathcal{X}\,, \tag{A.8}$$

where the first inclusion follows from standard results on the subdifferentiability of convex functions [29, Chap. 26], whereas the second is immediate from the definition of $\mathcal{X}_h$. This leads to two contrasting regimes: (i) $\mathcal{X}_h = \mathrm{ri}(\mathcal{X})$, in which case h is called *steep*; and (ii) $\mathcal{X}_h = \mathcal{X}$, in which case h is called *non-steep*.

Finally, we include here for future reference an elementary result concerning solid (convex) cones.

Lemma A.1. *Let $\mathcal{K}$ be a convex cone in $\mathbb{R}^d$ with nonempty topological interior, and let $z \in \mathrm{int}(\mathcal{K})$. Then there exists a finitely generated cone $\mathcal{K}'$ such that $z \in \operatorname{int}\mathcal{K}' \subseteq \operatorname{int}\mathcal{K}$.*

Remark. We stress here that, by $\operatorname{int}\mathcal{K}$ we mean the topological interior of $\mathcal{K}$ (which is nonempty by assumption), not the relative interior $\operatorname{ri}\mathcal{K}$ thereof (whis is always nonempty).

Proof. Since $\mathcal{K}$ is closed and $z \in \operatorname{int}\mathcal{K}$, there exists a closed ball $\mathcal{B}$ centered at z, which is entirely contained in $\operatorname{int}\mathcal{K}$ (an immediate consequence of the fact that z is well-separated from the boundary $\operatorname{bd}\mathcal{K}$ of $\mathcal{K}$). Since $\mathcal{B}$ is not contained in any lower-dimensional subspace of $\mathbb{R}^d$, it is possible to find inductively d linearly independent vectors $z_1, \ldots, z_d \in \mathcal{B}$ on the boundary $\operatorname{bd}\mathcal{B}$ of $\mathcal{B}$ such that z is contained in the convex hull $\Delta(z_1, \ldots, z_d)$ (and, in particular, in the relative interior thereof). Thus, letting $\mathcal{K}' \equiv \mathcal{K}(z_1, \ldots, z_d)$ be the polyhedral cone generated by $z_1, \ldots, z_d$, we have $\mathcal{K}' \subseteq \mathcal{K}$ and $z \in \operatorname{int}\mathcal{K}'$ by construction, and our proof is complete. ∎

A.2. Statistical bounds and results from probability theory. In this section, we provide some basic statistical bounds for (SPSA), and we present some results that we will use freely in the sequel. We start with our bounds for (SPSA), specifically:

Proposition A.1. *The estimator* (SPSA) *enjoys the following bounds:*

$$\|\mathbb{E}[\hat{v}_n \mid \mathcal{F}_n] - v(x_n)\|_* = \mathcal{O}(\varepsilon_n) \quad \text{and} \quad \|\hat{v}_n\|_* = \mathcal{O}(1/\varepsilon_n)\,. \tag{A.9}$$

Proof. Letting $\zeta_n \coloneqq \varepsilon_n(w_n + (p - x_n)/r)$, we write $\hat{x}_n = x_n + \zeta_n$, and we have for player $i \in \mathcal{N}$:

$$u_i(\hat{x}_n)w_{i,n} = u_i(x_n)w_{i,n} + \langle \nabla u_i(x_n), \zeta_n\rangle w_{i,n} + \int_0^1 \langle \nabla u_i(x_n + \tau\zeta_n) - \nabla u_i(x_n), \zeta_n\rangle d\tau w_{i,n} \tag{A.10}$$

Now, the middle term can be unfolded as

$$\langle \nabla u_i(x_n), \zeta_n\rangle w_{i,n} = \langle \nabla_i u_i(x_n), \zeta_{i,n}\rangle w_{i,n} + \sum_{j\neq i} \langle \nabla_j u_i(x_n), \zeta_{j,n}\rangle w_{i,n} \tag{A.11}$$

and, noting that $\mathbb{E}[w_{i,n} \mid \mathcal{F}_n] = 0$, we take conditional expectation, and we get:

$$\mathbb{E}[\langle \nabla u_i(x_n), \zeta_n\rangle w_{i,n} \mid \mathcal{F}_n] = \varepsilon_n\, \mathbb{E}[\langle \nabla_i u_i(x_n), w_{in}\rangle w_{i,n} \mid \mathcal{F}_n] = (\varepsilon_n/d_i) v_i(x_n) \tag{A.12}$$

and

$$\mathbb{E}[u_i(x_n) w_{i,n} \mid \mathcal{F}_n] = 0 \tag{A.13}$$

Therefore, we have:

$$\|\mathbb{E}[\hat{v}_{i,n} \mid \mathcal{F}_n] - v_i(x_n)\| = \left\|\mathbb{E}\left[\int_0^1 \langle \nabla u_i(x_n + \tau\zeta_n) - \nabla u_i(x_n), \zeta_n\rangle d\tau w_{i,n} \,\middle|\, \mathcal{F}_n\right]\right\| = \mathcal{O}(\varepsilon_n) \tag{A.14}$$

Now, for the second bound, since u_i is continuous on a compact domain, it is bounded, and we readily get that:

$$\|\hat{v}_{i,n}\|_* = \mathcal{O}(1/\varepsilon_n) \tag{A.15}$$

■

Moving forward, we provide some useful results from probability theory. The first two statements below are adapted from the classical textbook of Hall & Heyde [17], while the third one is a simplified version of [22, Theorem 3.2] on the recurrence of a nonnegative Markov process with diminishing drift. Namely, we have:

Theorem A.1. (Doob's maximal inequality, [17, Corollary 2.1]) *If S_n is a martingale, we have:*

$$\mathbb{P}\left(\sup_{k\leq n} |S_k| > t\right) \leq \frac{\mathbb{E}[|S_n|]}{t} \quad \text{for all } t > 0. \tag{A.16}$$

Theorem A.2. (Burkholder's inequality, [17, Theorem 2.10]) *Let $S_n \coloneqq \sum_{k=1}^n D_k$, where $(D_k)_{k\in\mathbb{N}}$ is a martingale difference sequence, and let $q \in (1, \infty)$. Then, there exists a constant C that depends only on q such that:*

$$\mathbb{E}[|S_n|^q] \leq C\, \mathbb{E}\left[\left|\sum_{k=1}^n D_k^2\right|^{q/2}\right] \tag{A.17}$$

Theorem A.3. (Lamperti [22, Theorem 3.2]) *Let the non-negative stochastic process $(x_n)_{n\in\mathbb{N}}$ be defined as*

$$x_{n+1} = (x_n + f(x_n) + \xi_n)^+ \tag{A.18}$$

for some $x \mapsto f(x)$ bounded measurable function, and ξ_n i.i.d. with $\mathbb{E}[\xi_n] = 0$, $\mathbb{V}(\xi_n) = \sigma^2 \neq 0$ and finite $2 + \varepsilon$ moment for some $\varepsilon > 0$. Then:

(i) if $f(x) \leq \sigma^2/2x$ for all x large enough, the process is recurrent in the sense there exists $c < \infty$ such that

$$\mathbb{P}\left(\liminf_{n\to\infty} x_n \leq c\right) = 1 \tag{A.19}$$

(ii) if $f(x) \geq \theta\sigma^2/2x$ for some $\theta > 1$ and all x large enough, the process is transient in the sense that

$$\mathbb{P}\left(\lim_{n\to\infty} x_n = \infty\right) = 1 \tag{A.20}$$

B Analysis and results for strategic robustness

Our aim in this appendix is to provide a detailed proof for Theorem 1, which we restate below for convenience.

Theorem 1. *Let $x^* \in \mathcal{X}$ be a joint action profile in $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$. Then the following are equivalent:*

(i) x^* is a strategically robust equilibrium.

(ii) $\langle v(x^*), z\rangle \leq -m\|z\|$ for some $m > 0$ and all $z \in \mathrm{TC}(x^*)$, where $\mathrm{TC}(x^*)$ is the closure of all rays emanating from x^* and intersecting $\mathcal{X}$ in at least one other point.

(iii) $v(x^*) \in \mathrm{int}(\mathrm{PC}(x^*))$, where $\mathrm{PC}(x^*) \coloneqq \{y \in \mathcal{Y} : \langle y, z\rangle \leq 0,$ for all $z \in \mathrm{TC}(x^*)\}$.

Proof. We will go full-circle by showing $(i) \implies (ii) \implies (iii) \implies (i)$.

(i) $\implies$ (ii). Suppose that $x^* \in \mathcal{X}$ is a strategically robust equilibrium, and let $\varepsilon > 0$ be such that x^* is an equilibrium of any $\widetilde{\mathcal{G}}$ with $\mathrm{dist}(\mathcal{G}, \widetilde{\mathcal{G}}) \leq \varepsilon$.

For the sake of contradiction, suppose that there exists $z \neq 0$, $z \in \mathrm{TC}(x^*)$ such that

$$\langle v(x^*), z\rangle = 0\,. \tag{B.1}$$

which readily implies that there exists player $i \in \mathcal{N}$ and $z_i \neq 0$, $z_i \in \mathrm{TC}_i(x_i^*)$, such that

$$\langle v_i(x^*), z_i\rangle = 0\,. \tag{B.2}$$

Fix some $y_i \in \mathcal{Y}_i$ such that $\langle y_i, z_i\rangle > 0$, and let $y \equiv (y_1, \dots, y_N) \in \mathcal{Y}$ with $y_j \equiv 0$ for $j \neq i$, $j \in \mathcal{N}$. Using (B.1) and the definition of y, we get that $\langle v(x^*) + \varepsilon\|y\|_*^{-1} y, z\rangle > 0$, and therefore, there exists $p \in \mathcal{X}$ such that:

$$\langle v(x^*) + \varepsilon\|y\|_*^{-1} y, p - x^*\rangle > 0 \tag{B.3}$$

Now, define the game $\widetilde{\mathcal{G}}$ with payoff functions

$$\tilde{u}_i(x) \coloneqq u_i(x) + \varepsilon\|y\|_*^{-1}\langle y_i, x_i - x_i^*\rangle \tag{B.4}$$

and $\tilde{u}_j \equiv u_j$ for all $j \in \mathcal{N}$, $j \neq i$. Then, the individual gradient vector of player $i \in \mathcal{N}$ is given by

$$\tilde{v}_i(x) = v_i(x) + \varepsilon\|y\|_*^{-1}\, y_i \tag{B.5}$$

and the distance between $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ is equal to

$$\mathrm{dist}(\mathcal{G}, \widetilde{\mathcal{G}}) = \sup_{x \in \mathcal{X}} \|v(x) - \tilde{v}(x)\|_* = \varepsilon\|y\|_*^{-1}\|y\|_* = \varepsilon. \tag{B.6}$$

Finally, we conclude that $x^* \in \mathcal{X}$ is not an equilibrium of $\widetilde{\mathcal{G}}$, since for $p \in \mathcal{X}$ as above, we have

$$\langle \tilde{v}(x^*), p - x^*\rangle = \langle v(x^*) + \varepsilon\|y\|_*^{-1} y, p - x^*\rangle > 0 \tag{B.7}$$

where the last inequality holds by (B.3). Thus, we arrive at a contradiction, i.e., $\langle v(x^*), z\rangle < 0$ for all $z \neq 0$, $z \in \mathrm{TC}(x^*)$. Finally, since $\{z \in \mathcal{V} : z \in \mathrm{TC}(x^*), \|z\| = 1\}$ is compact, we readily obtain

$$\sup\{\langle v(x^*), z\rangle : z \in \mathrm{TC}(x^*), \|z\| = 1\} \leq -m \tag{B.8}$$

for some $m > 0$. Therefore, for all $z \in \mathrm{TC}(x^*)$, we have:

$$\langle v(x^*), z\rangle \leq -m\|z\| \tag{B.9}$$

as was to be shown.

(ii) $\implies$ (iii). First, note that the $\|\cdot\|_*-$ ball of radius $\varepsilon > 0$ centered at $v(x^*)$ can be written as:

$$\mathbb{B}_\varepsilon(v(x^*)) = v(x^*) + \varepsilon\,\mathbb{B}_1(0) \tag{B.10}$$

where $\mathbb{B}_\varepsilon(y) \coloneqq \{y' \in \mathcal{Y} : \|y' - y\|_* \leq r\}$ for $y \in \mathcal{Y}$. Now, take any $y \in \mathbb{B}_1(0)$ and $z \in \mathrm{TC}(x^*)$. Then, for $\varepsilon > 0$ we have

$$\begin{aligned}\langle v(x^*) + \varepsilon y, z\rangle &= \langle v(x^*), z\rangle + \varepsilon\langle y, z\rangle \\ &\leq -m\|z\| + \varepsilon\|y\|_*\|z\| \\ &\leq -(m - \varepsilon)\|z\|\end{aligned} \tag{B.11}$$

Setting $\varepsilon = m/2$, we have for all $z \in \mathrm{TC}(x^*)$

$$\langle v(x^*) + (m/2)y, z\rangle < -(m/2)\|z\| \tag{B.12}$$

which implies that $v(x^*) + (m/2)y \in \mathrm{PC}(x^*)$. Thus, we readily get that $\mathbb{B}_{m/2}(v(x^*)) \subseteq \mathrm{PC}(x^*)$, i.e., $v(x^*) \in \mathrm{int}(\mathrm{PC}(x^*))$.

(iii) $\implies$ ***(i)*.** Suppose that $v(x^*) \in \operatorname{int}(\mathrm{PC}(x^*))$. First, it directly implies that $\langle v(x^*), z\rangle \le 0$, for all $z \in \mathrm{TC}(x^*)$, i.e., x^* is an equilibrium of $\mathcal{G}$. In addition, there exists $\varepsilon > 0$ such that $\mathbb{B}_\varepsilon(v(x^*)) \subseteq \mathrm{PC}(x^*)$. Therefore, for any game $\widetilde{\mathcal{G}}$ with $\operatorname{dist}(\mathcal{G}, \widetilde{\mathcal{G}}) < \varepsilon$, we immediately get that $\tilde{v}(x^*) \in \mathrm{PC}(x^*)$, which implies that $x^* \in \mathcal{X}$ is an equilibrium of $\widetilde{\mathcal{G}}$. Thus, x^* is strategically robust, and our proof is complete. ∎

C Analysis and results for dynamic robustness

In this appendix, we provide detailed proofs of the statements presented in Section 4, along with several intermediate results that will serve as key building blocks.

C.1. Intermediate results. We begin this section with two results establishing sufficient conditions for convergence, followed by a high-probability deviation bound for martingales. We conclude with a variant of Farkas' Lemma, which will be instrumental in deriving convergence rates.

Proposition C.1. *Let $x^* \in \mathcal{X}$ and $\mathcal{Z} :=\{z_1, \dots, z_m\} \subseteq \mathcal{V}$ be a set of unit vectors, such that any $z \in \mathrm{TC}(x^*)$ can be written as $z = \sum_{j=1}^m \lambda_j z_j$ for some $\lambda_j \ge 0$. If $\lim_{n\to\infty}\langle y_n, z_j\rangle = -\infty$ for all $z_j \in \mathcal{Z}$, then $\lim_{n\to\infty} Q(y_n) = x^*$.*

Proof. Denote $Q(y_n)$ by x_n, and suppose that $\limsup_{n\to\infty}\|x_n - x^*\| > 0$. Then, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $\|x_{n_k} - x^*\|$ stays bounded away from zero, i.e., $\|x_{n_k} - x^*\| \ge c$ for some $c > 0$ and all $k \in \mathbb{N}$. Since $y_{n_k} \in \partial h(x_{n_k})$, we readily get for $z_{n_k} = (x_{n_k} - x^*)/\|x_{n_k} - x^*\|$:

$$\begin{aligned} h(x^*) &\ge h(x_{n_k}) + \langle y_{n_k}, x^* - x_{n_k}\rangle \\ &= h(x_{n_k}) - \langle y_{n_k}, z_{n_k}\rangle \|x_{n_k} - x^*\| \\ &\ge \min h - \langle y_{n_k}, z_{n_k}\rangle\|x_{n_k} - x^*\|\,. \end{aligned} \tag{C.1}$$

Now, we have $z_{n_k} \in \mathrm{TC}(x^*)$, and by assumption, $z_{n_k} = \sum_{j=1}^m \lambda_{j,k} z_j$ for some coefficients $\lambda_{j,k} \ge 0$. Therefore, the above inequality can be written as:

$$h(x^*) \ge \min h - \|x_{n_k} - x^*\| \sum_{j=1}^m \lambda_{j,k}\langle y_{n_k}, z_j\rangle \tag{C.2}$$

$$\ge \min h - \left(\max_{j'}\langle y_{n_k}, z_{j'}\rangle\right)\|x_{n_k} - x^*\| \sum_{j=1}^m \lambda_{j,k} \tag{C.3}$$

Now, note that by the definition of z_{n_k}, we have $\|z_{n_k}\| = 1$, and, thus:

$$1 = \|z_{n_k}\| = \left\|\sum_{j=1}^m \lambda_{j,k} z_j\right\| \le \sum_{j=1}^m \lambda_{j,k}\|z_j\| = \sum_{j=1}^m \lambda_{j,k} \tag{C.4}$$

where we used that $\|z_j\| = 1$ for all j. Now, since $\lim_{n\to\infty}\langle y_n, z_j\rangle = -\infty$ for all $z_j \in \mathcal{Z}$, it readily implies that

$$\lim_{n\to\infty}\max_{j'}\langle y_n, z_{j'}\rangle = -\infty \tag{C.5}$$

Therefore, for all k large enough, we have $-\max_{j'}\langle y_{n_k}, z_{j'}\rangle > 0$, and, using that $\sum_{j=1}^m \lambda_{j,k} \ge 1$ and $\|x_{n_k} - x^*\| \ge c$ for all $k \in \mathbb{N}$, we obtain:

$$h(x^*) \ge \min h + \left(-\max_{j'}\langle y_{n_k}, z_{j'}\rangle\right)\|x_{n_k} - x^*\| \sum_{j=1}^m \lambda_{j,k} \tag{C.6}$$

$$\ge \min h - c\max_{j'}\langle y_{n_k}, z_{j'}\rangle \tag{C.7}$$

Finally, letting $k \to \infty$, we get that $h(x^*) \ge \infty$, which is a contradiction. Thus, the result follows. ∎

Finally, using the above proposition, we establish the following corollary.

Corollary C.1. *Let $\mathcal{W}(M) :=\{y \in \mathcal{Y} : \max_{z\in\mathcal{Z}}\langle y, z\rangle < -M\}$ for $M > 0$. Then, for any $\varepsilon > 0$, there exists $M_\varepsilon > 0$ such that for all $y \in \mathcal{W}(M_\varepsilon)$ it holds $\|x^* - Q(y)\| < \varepsilon$.*

Proof. Suppose it does not hold. Then, there exists $\varepsilon > 0$ such that for any $n \in \mathbb{N}$, one can find $y_n \in \mathcal{Y}$ such that $\max_{z\in\mathcal{Z}}\langle y_n, z\rangle < -n$ and $\|x^* - Q(y_n)\| \geq \varepsilon$. Taking $n \to \infty$ leads to a contradiction with Proposition C.1. ∎

Lemma C.1. *Let $S_n := \gamma \sum_{k=1}^n \xi_k$ be a martingale with respect to a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ such that $\mathbb{E}[|\xi_n|^q] \leq \sigma_n^q$ for all $n \in \mathbb{N}$ and some $q \geq 2$. Then, for any $\mu \in (0,1)$ and $c > 0$, it holds:*

$$\mathbb{P}\left(\sup_{\ell\leq n}|S_\ell| > c(\gamma n)^\mu\right) \leq C_q \frac{\gamma^{q(1-\mu)}\sum_{k=1}^n \sigma_k^q}{n^{1+q(\mu-1/2)}} \tag{C.8}$$

where C_q is a constant that depends only on c and q.

Proof. To bound the maximum absolute deviation of S_n, we apply Doob's maximal inequality (see Theorem A.1), and obtain:

$$\mathbb{P}\left(\sup_{\ell\leq n}|S_\ell| > c(\gamma n)^\mu\right) \leq \frac{\mathbb{E}[|S_n|^q]}{c^q(\gamma n)^{q\mu}} \tag{C.9}$$

Now, we invoke Burkholder's inequality (see Theorem A.2), from which we get:

$$\mathbb{E}[|S_n|^q] \leq C'_q\, \mathbb{E}\left[\left(\sum_{k=1}^n \gamma^2|\xi_k|^2\right)^{q/2}\right] \leq C'_q\gamma^q\, \mathbb{E}\left[\left(\sum_{k=1}^n|\xi_k|^2\right)^{q/2}\right] \tag{C.10}$$

where C_q is a constant that depends only on q. Since $q \geq 2$, applying Jensen's inequality, we obtain:

$$\left(\frac{1}{n}\sum_{k=1}^n|\xi_k|^2\right)^{q/2} \leq \frac{1}{n}\left(\sum_{k=1}^n|\xi_k|^q\right) \tag{C.11}$$

and, therefore,

$$\mathbb{E}\left[\left(\sum_{k=1}^n|\xi_k|^2\right)^{q/2}\right] \leq n^{q/2-1}\,\mathbb{E}\left[\sum_{k=1}^n|\xi_k|^q\right] \tag{C.12}$$

$$\leq n^{q/2-1}\sum_{k=1}^n \sigma_k^q \tag{C.13}$$

Thus, combining the above with (C.9) and (C.10), we obtain:

$$\mathbb{P}\left(\sup_{\ell\leq n}|S_\ell| > c(\gamma n)^\mu\right) \leq \frac{C'_q\gamma^q n^{q/2-1}\sum_{k=1}^n\sigma_k^q}{c^q(\gamma n)^{q\mu}} \tag{C.14}$$

$$\leq C_q\frac{\gamma^{q(1-\mu)}\sum_{k=1}^n\sigma_k^q}{n^{1+q(\mu-1/2)}} \tag{C.15}$$

for $C_q \equiv C'_q/c^q$, and the proof is complete. ∎

We finally provide a separation result in the spirit of Farkas' lemma, that we will need for establishing the convergence rates.

Lemma C.2. *Let $\mathcal{X} = \{x \in \mathcal{V} : Ax = b, x \geq 0\}$ for $A \in \mathbb{R}^{m\times d}$, $b \in \mathbb{R}^d$. Then, for all $x^* \in \mathcal{X}$ with $act(x^*) := \{\beta \in \{1,\dots,d\} : x^*_\beta = 0\}$, there exists $P \equiv P(x^*) \geq 1$ such that for all $\mathcal{I} \subseteq act(x^*)$ at least one of the following is true:*

(i) $\mathcal{I} \neq \emptyset$ and there exists $\beta \in act(x^*) \setminus \mathcal{I}$ such that $x_\beta \leq P\max\{x_\alpha : \alpha \in \mathcal{I}\}$ for all $x \in \mathcal{X}$.

(ii) There exists $z \in \ker(A)$ such that $\|z\| \leq P$, $z_\beta = 0$ for $\beta \in \mathcal{I}$ and $1 \leq z_\beta \leq P$ for $\beta \in act(x^*)\setminus\mathcal{I}$. Then, there exists

Proof. For the proof, see Azizian et al. [6, Lemma 6]. ∎

C.2. Main results of Section 4. With the necessary tools in place, we proceed to prove the main results stated in Section 4. We start with the first result, establishing that a non-equilibrium point cannot arise as a limit point of the sequence of play induced by (FTRL).

Proposition 1. *Let $x' \in \mathcal{X}$ be an non-equilibrium point of $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$. Then, x' cannot be a limit point of (FTRL) dynamics with $\hat{v}_n \equiv v(x_n)$ for all $n \in \mathbb{N}$.*

Proof. Since x' is not an equilibrium, there exists $p \in \mathcal{X}$ with $\langle v(x'), p - x'\rangle > 0$. Therefore, by continuity of the function $x \mapsto \langle v(x), p - x\rangle$, there exists a neighborhood $\mathcal{U}$ of x' and $c > 0$ such that $\langle v(x), p - x\rangle \geq c$ for all $x \in \mathcal{U}$.

Moreover, since $\mathrm{cl}(\mathcal{U})$ compact, we have $\sup_{x\in\mathrm{cl}(\mathcal{U})} \|v(x)\|_* = B < \infty$. For the sake of contradiction, suppose that $x_n \to x'$. Then, $x_n \in \mathcal{U} \cap \mathbb{B}_{c/4B}(x')$ eventually, i.e., there exists n_0 such that $x_n \in \mathcal{U}$ and $\|x_n - x'\| < c/4B$ for all $n \geq n_0$.

Finally, since $y_n \in \partial h(x_n)$, we have for $n > n_0$:

$$\begin{aligned}
h(p) &\geq h(x_n) + \langle y_n, p - x_n\rangle \\
&\geq h(x_n) + \langle y_{n_0}, p - x_n\rangle + \gamma \sum_{k=n_0}^{n-1} \langle v(x_k), p - x_n\rangle \\
&\geq h(x_n) + \langle y_{n_0}, p - x_n\rangle + \gamma \sum_{k=n_0}^{n-1} \langle v(x_k), p - x_k + x_k - x_n\rangle \\
&\geq h(x_n) + \langle y_{n_0}, p - x_n\rangle + \gamma \sum_{k=n_0}^{n-1} (\langle v(x_k), p - x_k\rangle + \langle v(x_k), x_k - x_n\rangle) \\
&\geq h(x_n) + \langle y_{n_0}, p - x_n\rangle + \gamma \sum_{k=n_0}^{n-1} (\langle v(x_k), p - x_k\rangle - \|v(x_k)\|_* \|x_k - x_n\|) \\
&\geq h(x_n) + \langle y_{n_0}, p - x_n\rangle + \gamma \sum_{k=n_0}^{n-1} (c - B\|x_k - x_n\|) \\
&\geq h(x_n) - \|y_{n_0}\|_* \|p - x_n\| + \gamma \sum_{k=n_0}^{n-1} (c - c/2) \\
&\geq \min h - \|y_{n_0}\|_* \operatorname{diam}(\mathcal{X}) + \gamma c(n - n_0)/2
\end{aligned} \tag{C.16}$$

Taking $n \to \infty$, we obtain $h(p) \geq \infty$, which is a contradiction. Therefore, the result follows. ■

Moving forward, we show that equilibrium points in the relative interior cannot be limit points of (FTRL), either. Formally, we have:

Proposition 2. *Let $x^* \in \mathrm{ri}(\mathcal{X})$ be a Nash equilibrium of $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$, and $(x_n)_{n\in\mathbb{N}}$ be the sequence of play induced by (FTRL) with $\hat{v}_n = v(x_n) + U_n$, where U_n i.i.d. with $\mathbb{E}[U_n] = 0$ and $\mathrm{cov}(U_n) \succ 0$ for all $n \in \mathbb{N}$. Then:*

$$\mathbb{P}\Big(\lim_{n\to\infty} x_n = x^*\Big) = 0 \qquad \textit{for any } x_1 \in \mathcal{X}_h. \tag{14}$$

Proof. Since x^* an equilibrium point in $\mathrm{ri}(\mathcal{X})$, we readily get that $\langle v(x^*), x - x^*\rangle = 0$ for all $x \in \mathcal{X}$, and $x^* \in \mathcal{X}_h$. In view of this, there exists $y^* \in \mathcal{Y}$ such that $y^* \in \partial h(x^*)$, i.e., $x^* = Q(y^*)$. Our goal is to show that the auxiliary process y_n does not converge to $\partial h(x^*)$. However, there are infinitely many points in $\mathcal{Y}$ that belong to $\partial h(x^*)$, so this attempt is insufficient, in the sense that, showing that y^* is not a limit point of the y_n dynamics, does not preclude that some other $y \in \partial h(x^*)$ is not.

To tackle this issue, we will show that the space $\mathcal{Y}$ can be decomposed as $\mathcal{Y} = \widehat{\mathcal{Y}} \oplus \overline{\mathcal{Y}}$ where all the "essential" deviations of the problem is in $\overline{\mathcal{Y}}$. For this, we define the set

$$\widehat{\mathcal{Y}} = \{y \in \mathcal{Y} : \langle y, p - x\rangle = 0, \text{ for all } x, p \in \mathcal{X}\}. \tag{C.17}$$

which is a subspace of $\mathcal{Y}$, and as the following lemma suggests, is equal to the polar cone at any point in the relative interior.

Lemma C.3. *Let* $x_0 \in \mathrm{ri}(\mathcal{X})$ *and* $\widehat{\mathcal{Y}} = \{y \in \mathcal{Y} : \langle y, p - x\rangle = 0,$ *for all* $x, p \in \mathcal{X}\}$*. Then* $\widehat{\mathcal{Y}} = \mathrm{PC}(x_0)$.

To preserve the clarity of the argument, we defer the proof of Lemma C.3 until the end of this proposition. Letting $\overline{\mathcal{Y}}$ be the orthocomplement of $\widehat{\mathcal{Y}}$, we readily get that $\mathcal{Y} = \widehat{\mathcal{Y}} \oplus \overline{\mathcal{Y}}$, and any point y in $\mathcal{Y}$ can be uniquely written as $y = \hat{y} + \bar{y}$ with $\hat{y} \in \widehat{\mathcal{Y}}$ and $\bar{y} \in \overline{\mathcal{Y}}$. Defining the linear map $\Pi : \mathcal{Y} \to \mathcal{Y}$ as $\Pi y = \bar{y}$, and more importantly, under all points in $\partial h(x^*)$ under Π are essentially unique.

This is formalized in the following lemma, whose proof is relegated after this proposition.

Lemma C.4. *Let* $x_0 \in \mathrm{ri}(\mathcal{X})$ *and* $y, y' \in \partial h(x_0)$*. Then* $\Pi y \in \partial h(x_0)$*, and* $\Pi y = \Pi y'$.

In view of the above, we are now ready to prove the result. Namely, fix some $p \in \mathcal{X}$, $p \neq x^*$ and let $\xi_n := \langle \Pi U_n, p - x^*\rangle$.

Then, setting $\sigma^2 \equiv (p - x^*)^\top \Sigma (p - x^*) > 0$, we have $\xi_n \sim (0, \sigma^2)$ i.i.d., and, so, there exists $\varepsilon, \delta > 0$ such that $\mathbb{P}(\xi_n > \varepsilon) = \delta$ for all $n \in \mathbb{N}$. Therefore, by the second Borel-Cantelli lemma [9], we get $\mathbb{P}(A) = 1$ for $A \equiv \{\xi_n > \varepsilon$ infinitely often$\}$. For the sake of contradiction, suppose that $\mathbb{P}(B) > 0$ for $B \equiv \{\lim_{n\to\infty} x_n = x^*\}$. Fix some $\omega \in B$. Then, for all n large enough, we readily get that $x_n(\omega) \in \mathrm{ri}(\mathcal{X})$, and, denoting $z_n := \Pi y_n$ and $z^* := \Pi y^*$, we readily get that

$$\lim_{n\to\infty} z_n(\omega) = z^* \tag{C.18}$$

Thus, setting $\alpha_n \equiv \langle z_n - z^*, p - x^*\rangle$ we conclude by the above equality that $\lim_{n\to\infty} \alpha_n = 0$, and therefore it holds

$$\begin{aligned} 0 &= \lim_{n\to\infty} (\alpha_n - \alpha_{n-1}) \\ &= \lim_{n\to\infty} \langle \Pi \hat{v}_n, p - x^*\rangle \\ &= \lim_{n\to\infty} \langle \Pi v(x_n), p - x^*\rangle + \langle \Pi U_n, p - x^*\rangle \\ &= \langle \Pi v(x^*), p - x^*\rangle + \lim_{n\to\infty} \xi_n \\ &= \lim_{n\to\infty} \xi_n \end{aligned} \tag{C.19}$$

Therefore, $\omega \notin A$, which implies that $B \subseteq A^c$, with $\mathbb{P}(A^c) = 0$. Thus, $\mathbb{P}(B) = 0$, which is a contradiction, and the result follows.

■

We now prove the two auxiliary lemmas presented in the proof of Proposition 2.

Lemma C.3. *Let* $x_0 \in \mathrm{ri}(\mathcal{X})$ *and* $\widehat{\mathcal{Y}} = \{y \in \mathcal{Y} : \langle y, p - x\rangle = 0,$ *for all* $x, p \in \mathcal{X}\}$*. Then* $\widehat{\mathcal{Y}} = \mathrm{PC}(x_0)$.

Proof. First, we will show that

$$\mathrm{PC}(x_0) = \{y \in \mathcal{Y} : \langle y, p - x_0\rangle = 0 \text{ for all } p \in \mathcal{X}\} \tag{C.20}$$

For this, suppose that there exist $y \in \mathrm{PC}(x_0)$ and $p' \in \mathcal{X}$ such $\langle y, p' - x_0\rangle < 0$. Then, since $x_0 \in \mathrm{ri}(\mathcal{X})$, there exists $\alpha > 0$ such that $x_0 - \alpha(p' - x_0) \in \mathcal{X}$. By the definition of the polar cone, $\langle y, x_0 - \alpha(p' - x_0) - x_0\rangle \le 0$, or equivalently, $\langle y, p' - x_0\rangle \ge 0$, which is a contradiction. Therefore, (C.20) holds, which implies that $\widehat{\mathcal{Y}} \subseteq \mathrm{PC}(x)$.

Now, for the inverse inclusion, let $y \in \mathrm{PC}(x_0)$ and $p, x \in \mathcal{X}$. Then, we have:

$$\begin{aligned} \langle y, p - x\rangle &= \langle y, p - x_0 + x_0 - x\rangle \\ &= \langle y, p - x_0\rangle + \langle y, x_0 - x\rangle \\ &= 0 \end{aligned} \tag{C.21}$$

where the last equality follows by (C.20). Thus, $y \in \widehat{\mathcal{Y}}$, and we conclude the result. ■

Lemma C.4. *Let* $x_0 \in \mathrm{ri}(\mathcal{X})$ *and* $y, y' \in \partial h(x_0)$*. Then* $\Pi y \in \partial h(x_0)$*, and* $\Pi y = \Pi y'$.

Proof. For the first part, note that

$$\begin{aligned}\langle y, p - x_0\rangle = \langle \hat{y} + \bar{y}, p - x_0\rangle &= \langle \hat{y}, p - x_0\rangle + \langle \bar{y}, p - x_0\rangle \\ &= \langle \bar{y}, p - x_0\rangle \\ &= \langle \Pi y, p - x_0\rangle \end{aligned} \tag{C.22}$$

which directly implies that $\Pi y \in \partial h(x_0)$. For the second part, since $x_0 \in \mathrm{ri}(\mathcal{X})$, and $y, y' \in \partial h(x_0)$, we have that

$$\langle y - y', p - x_0\rangle = 0 \quad \text{for all } p \in \mathcal{X} \tag{C.23}$$

Thus $y - y' \in \mathrm{PC}(x_0)$, and invoking Lemma C.3 we obtain that $y - y' \in \widehat{\mathcal{Y}}$. Therefore, applying the linear projection operator Π, we readily get that $\Pi(y - y') = 0$, and, using linearity, the result follows. ∎

We now turn to our main convergence theorems, showing that the iterates of (FTRL) converge with high probability under both gradient-based and payoff-based feedback

Theorem 2. *Let $x^* \in \mathcal{X}$ be a strategically robust equilibrium of $\mathcal{G}(\mathcal{N}, \mathcal{X}, u)$. Fix a confidence level $\delta > 0$, and let $(x_n)_{n\in\mathbb{N}}$ be the iterates of (FTRL) with feedback provided by (SFO), and step-size $\gamma > 0$ sufficiently small. Then, there exists a neighborhood $\mathcal{U}$ of x^* in $\mathcal{X}_h$ such that:*

$$\mathbb{P}\left(\lim_{n\to\infty} x_n = x^*\right) \geq 1 - \delta \quad \text{if } x_1 \in \mathcal{U}. \tag{15}$$

Proof. Since x^* strategically robust, $v(x^*)$ lies in the interior of the $\mathrm{PC}(x^*)$. By Lemma A.1, this implies in turn that there exists a polyhedral cone $\mathcal{K}$ generated by $\mathcal{Z} \equiv \{z_1, \dots, z_r\}$ for $r \in \mathbb{N}$, such that $\mathrm{TC}(x^*) \subseteq \mathcal{K}$ and $\langle v(x^*), z\rangle < 0$ for all $z \in \mathcal{Z}$.[4] Therefore, for all $z \in \mathcal{Z}$, we have $\langle v(x^*), z\rangle \leq -m$, and by continuity of the vector field v, there exists a neighborhood $\mathcal{U}$ of x^* and $c > 0$ such that $\langle v(x), z\rangle \leq -c$ for all $z \in \mathcal{Z}$ and $x \in \mathcal{U}$.

Fixing some $z \in \mathcal{Z}$, we obtain:

$$\begin{aligned}\langle y_{n+1}, z\rangle &= \langle y_n, z\rangle + \gamma\langle \hat{v}_n, z\rangle \\ &= \langle y_n, z\rangle + \gamma\langle v(x_n), z\rangle + \gamma\langle U_n, z\rangle \\ &= \langle y_1, z\rangle + \gamma\sum_{k=1}^{n}\langle v(x_k), z\rangle + \gamma\sum_{k=1}^{n}\langle U_k, z\rangle \end{aligned} \tag{C.24}$$

Now, we define the stochastic process $(S_n)_{n\in\mathbb{N}}$ via $S_n \coloneqq \gamma\sum_{k=1}^{n}\langle U_k, z\rangle$, which is a martingale, since $\mathbb{E}[\langle U_k, z\rangle \,|\, \mathcal{F}_k] = 0$.

Therefore, by Lemma C.1 for $\sigma_n \equiv \sigma$, $q > 2$ and $\mu \in (0, 1)$, whose value is determined later, we get:

$$\delta_n \coloneqq \mathbb{P}\left(\sup_{\ell\leq n} |S_\ell| > c(\gamma n)^{\mu}\right) \leq C_q \frac{\gamma^{q(1-\mu)}\sigma^q}{n^{q(\mu-1/2)}} \tag{C.25}$$

where C_q is a constant that depends only on c and q. Thus, we readily have that:

$$\begin{aligned}\mathbb{P}\left(\bigcap_{n\geq 1}\left\{\sup_{\ell\leq n}|S_\ell| \leq c(\gamma n)^{\mu}\right\}\right) &= 1 - \mathbb{P}\left(\bigcup_{n\geq 1}\left\{\sup_{\ell\leq n}|S_\ell| > c(\gamma n)^{\mu}\right\}\right) \\ &\geq 1 - \sum_{n=1}^{\infty}\mathbb{P}\left(\sup_{\ell\leq n}|S_\ell| > c(\gamma n)^{\mu}\right) \\ &\geq 1 - \sum_{n=1}^{\infty}\delta_n \end{aligned} \tag{C.26}$$

where the second inequality comes from the union bound. Now, we need to ensure that $\sum_{n=1}^{\infty}\delta_n \leq \delta/r$. For this, we need the sequence to be summable, which, using (C.25), is guaranteed for $q(\mu - 1/2) > 1$, or equivalently, $\mu \in (1/2 + 1/q, 1)$. Therefore, for $\gamma > 0$ small enough, we obtain that $\sum_{n=1}^{\infty}\delta_n \leq \delta/r$.

[4] To resolve any ambiguities, the cone in question here is the polar of the cone provided by Lemma A.1.

Therefore, with probability at least $1 - \delta/r$, the template inequality becomes:

$$\langle y_{n+1}, z\rangle \leq \langle y_1, z\rangle + \gamma \sum_{k=1}^{n} \langle v(x_k), z\rangle + c(\gamma n)^{\mu} \tag{C.27}$$

If we initialize y_1 such that $\langle y_1, z'\rangle < -M - c$ for all $z' \in \mathcal{Z}$, we get that $\langle y_n, z\rangle < -M$ for all $n \in \mathbb{N}$ with probability at least $1 - \delta/r$. To see this, suppose that $\langle y_k, z\rangle < -M$ for all $k = 1, \dots, n$. Then

$$\begin{aligned}\langle y_{n+1}, z\rangle &= \langle y_1, z\rangle + \gamma \sum_{k=1}^{n} \langle v(x_k), z\rangle + \gamma \sum_{k=1}^{n} \langle U_k, z\rangle \\ &\leq -M - c - c\gamma n + c(\gamma n)^{\mu}\end{aligned} \tag{C.28}$$

For $n \in \mathbb{N}$ with $\gamma n < 1$, we have $-c + c(\gamma n)^{\mu} < 0$, while for $\gamma n \geq 1$, it holds $-c\gamma n + c(\gamma n)^{\mu} < 0$. In both cases, we conclude that $\langle y_{n+1}, z\rangle < -M$, and by induction, we get the inequality.

Therefore, with probability at least $1 - \delta/r$, we have:

$$\langle y_{n+1}, z\rangle \leq -M - c - c\gamma n + c(\gamma n)^{\mu} \tag{C.29}$$

and sending $n \to \infty$, we get $\langle y_n, z\rangle \to -\infty$.

Finally, repeating the same argument for all $z \in \mathcal{Z}$ and applying a union bound, we readily get that $\langle y_n, z\rangle \to -\infty$ with probability at least $1 - \delta$, and invoking Proposition C.1, the result follows. ∎

Having established the local convergence to x^* with high probability, we proceed to the convergence rate in the case of affinely constrained $\mathcal{X}$ and decomposable regularizer h.

Theorem 3. *If, in addition, $\mathcal{X}$ is a polyhedral domain and h is decomposable with kernel θ, on the event $E := \{\lim_{n\to\infty} x_n = x^*\}$ it holds:*

$$\|x_n - x^*\| = \phi(-\Theta(n)) \tag{16}$$

where ϕ is the rate function defined via

$$\phi(z) := \begin{cases} (\theta')^{-1}(z) & \text{if } z > \theta'(0^+) \\ 0 & \text{if } z \leq \theta'(0^+) \end{cases} \tag{17}$$

Proof. By the definition of the iterates of (FTRL), we have:

$$Q(y_n) = \arg\min_{x\in\mathcal{X}} \{h(x) - \langle y_n, x\rangle : Ax = b, x \geq 0\} \tag{C.30}$$

Introducing the Lagrangian

$$\mathcal{L}(x, \lambda, \mu) = h(x) - \langle y_n, x\rangle + \sum_{i=1}^{m} \lambda_i (a_i^\top x - b_i) - \sum_{j=1}^{d} \mu_j x_j \tag{C.31}$$

with $\lambda_i \in \mathbb{R}$ and $\mu_j \geq 0$, by the KKT conditions, we readily obtain:

$$y_n = \nabla h(x_n) + \sum_{\ell=1}^{m} \lambda_i a_i - \mu \tag{C.32}$$

where $\nabla h(x) = \sum_{\beta=1}^{d} \theta'(x_{\beta,n}) e_\beta$, since θ is continuously differentiable.

For the sequel, we define the set of active constraints at x^* as $\operatorname{act}(x^*) := \{\beta \in \{1, \dots, d\} : x^*_\beta = 0\}$. Note that on the event of $\{\lim_{n\to\infty} x_n = x^*\}$, the iterates x_n lie in a neighborhood of x^*, as shown in Theorem 2. Thus, all non-active indices $\alpha \notin \operatorname{act}(x^*)$ stay bounded away from zero, and so $|\theta(x_{\alpha,n})|$ remains bounded for all n.

We treat the two cases separately: (i) the steep case, where h is steep – equivalently $\theta(0^+) = -\infty$, and (ii) the non-steep case, where is h not steep, i.e., $\theta'(0^+) > -\infty$.

The steep case. We define the set of "good" indices $\mathcal{I}$ at step n as: $\beta \in \mathcal{I}$ if $\theta'(x_{\beta,n}) \leq -\Theta(n)$. Our goal is to show that all indices $\operatorname{act}(x^*)$ of x_n are "good". Fix some $n \in \mathbb{N}$.

Suppose that $\operatorname{act}(x^*) \setminus \mathcal{I} \neq \emptyset$, and let $P \geq 1$, as per Lemma C.2. Then,

- If condition *(i)* of *Lemma C.2* holds, there exists β' such that $x_{\beta',n} \leq P \max\{x_{\alpha,n} : \alpha \in \mathcal{I}\}$, and thus, $\mathcal{I} \leftarrow \mathcal{I} \cup \{\beta'\}$.

- If condition *(ii)* of *Lemma C.2* holds, there exists $z' \in \ker(A)$ such that $\|z'\| \leq P$, $z'_\beta = 0$ if $\beta \in \mathcal{I}$ and $1 \leq z'_\beta \leq P$ if $\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}$. By (C.32), and noting that $z' \in \ker(A)$ and $\mu = 0$, since all constraints are non-active due to steepness of h, we have:

$$\langle \nabla h(x_n), z' \rangle = \langle y_n, z' \rangle \tag{C.33}$$

  Moreover, it holds:

$$\begin{aligned} \langle \nabla h(x_n), z' \rangle = \sum_{\beta=1}^{d} \theta'(x_{\beta,n}) z'_\beta &= \sum_{\beta \in \mathcal{I}} \theta'(x_{\beta,n}) z'_\beta + \sum_{\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}} \theta'(x_{\beta,n}) z'_\beta + \sum_{\beta \notin \operatorname{act}(x^*)} \theta'(x_{\beta,n}) z'_\beta \\ &= \sum_{\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}} \theta'(x_{\beta,n}) z'_\beta + C \end{aligned} \tag{C.34}$$

  for a constant C, since all non-active indices remain bounded away from zero, as explained in the beginning. Now, note that $z' \in \operatorname{TC}(x^*)$, and thus, by Lemma A.1, we can write z' as $z' = \sum_{i=1}^{r} \ell_i z_i$ with $\ell_i \geq 0$, such that $\langle y_n, z_i \rangle \leq -\Theta(n)$ for all $i = 1, \dots, r$ as in the proof of Theorem 2. So, combining it with (C.39), (C.34), we obtain:

$$\sum_{\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}} \theta'(x_{\beta,n}) z'_\beta \leq -\Theta(n) \tag{C.35}$$

  and therefore, there exists at least one $\beta' \in \operatorname{act}(x^*) \setminus \mathcal{I}$ such that

$$\theta'(x_{\beta',n}) z'_{\beta'} \leq -\Theta(n) \tag{C.36}$$

  Thus, $\mathcal{I} \leftarrow \mathcal{I} \cup \{\beta'\}$.

Therefore, as $\operatorname{act}(x^*)$ is finite, we conclude inductively that $\theta'(x_{\beta,n}) \leq -\Theta(n)$ for all $\beta \in \operatorname{act}(x^*)$. Finally, we have that $\mathbb{R}^d = \operatorname{row}(A) + \operatorname{span}\{e_\beta : \beta \in \operatorname{act}(x^*)\}$, and thus, for all i, we can write the standard basis vector e_i as $e_i = \sum_{\beta \in \operatorname{act}(x^*)} \lambda_{i,\beta} e_\beta + a_i$ for some $a_i \in \operatorname{row}(A)$

$$\begin{aligned} x_{i,n} - x_i^* = \langle x_n - x^*, e_i \rangle &= \left\langle x_n - x^*, \sum_{\beta \in \operatorname{act}(x^*)} \lambda_{i,\beta} e_\beta + a_i \right\rangle \\ &= \left\langle x_n - x^*, \sum_{\beta \in \operatorname{act}(x^*)} \lambda_{i,\beta} e_\beta \right\rangle \\ &= \sum_{\beta \in \operatorname{act}(x^*)} \lambda_{i,\beta} x_{\beta,n} \end{aligned} \tag{C.37}$$

where we used that $\langle x_n - x^*, a_i \rangle = 0$. Thus, since $\theta'(x_{\beta,n}) \leq -\Theta(n)$ for all $\beta \in \operatorname{act}(x^*)$, by the equivalence of norms and the above, we conclude that

$$\|x_n - x^*\| = (\theta')^{-1}(-\Theta(n)) \tag{C.38}$$

The non-steep case. For the non-steep case, we follow a similar approach, but with some modifications since the iterates of (FTRL) are not always in the interior of $\mathcal{X}$.

Specifically, let the set of "good" indices $\mathcal{I}$ be defined as: $\beta \in \mathcal{I}$ if $x_{\beta,n} = 0$ or $\theta'(x_{\beta,n}) \leq -\Theta(n)$. Our goal is to show that all indices $\operatorname{act}(x^*)$ of x_n are "good". We construct $\mathcal{I}$ sequentially, as before.

Suppose that $\operatorname{act}(x^*) \setminus \mathcal{I} \neq \emptyset$, and let $P \geq 1$, as per Lemma C.2. Then,

- If condition *(i)* of *Lemma C.2* holds, there exists β' such that $x_{\beta',n} \leq P \max\{x_{\alpha,n} : \alpha \in \mathcal{I}\}$, and thus, $\mathcal{I} \leftarrow \mathcal{I} \cup \{\beta'\}$.

- If condition *(ii)* of *Lemma C.2* holds, there exists $z' \in \ker(A)$ such that $\|z'\| \le P$, $z'_\beta = 0$ if $\beta \in \mathcal{I}$ and $1 \le z'_\beta \le P$ if $\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}$. Therefore, we have

$$\begin{aligned}\langle \nabla h(x_n), z' \rangle = \langle y_n, z' \rangle + \langle \mu, z' \rangle &= \langle y_n, z' \rangle + \sum_{\beta \in \mathcal{I}} \mu_\beta z'_\beta + \sum_{\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}} \mu_\beta z'_\beta + \sum_{\beta \notin \operatorname{act}(x^*)} \mu_\beta z'_\beta \\ &= \langle y_n, z' \rangle \end{aligned} \tag{C.39}$$

  where, in this case, we used that (i) $z'_\beta = 0$ for $\beta \in \mathcal{I}$, (ii) $\mu_\beta = 0$ by complementary slackness for $\beta \notin \operatorname{act}(x^*)$ since these constraints remain non-active for the whole process, and (iii) $\mu_\beta = 0$, again by complementary slackness for $\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}$ since if they were active, we would have $\beta \in \mathcal{I}$. This, with the same argument as before, we conclude that

$$\sum_{\beta \in \operatorname{act}(x^*) \setminus \mathcal{I}} \theta'(x_{\beta,n}) z'_\beta \le -\Theta(n) \tag{C.40}$$

  and therefore, there exists at least one $\beta' \in \operatorname{act}(x^*) \setminus \mathcal{I}$ such that

$$\theta'(x_{\beta',n}) z'_{\beta'} \le -\Theta(n) \tag{C.41}$$

  This holds until β' vanishes, which can lead to $\mu_{\beta'} > 0$. In either case, we have $\mathcal{I} \leftarrow \mathcal{I} \cup \{\beta'\}$.

Finally, since $\operatorname{act}(x^*)$ is finite, we conclude inductively that all for all $\beta \in \operatorname{act}(x^*)$, we have either $\theta'(x_{\beta,n}) \le -\Theta(n)$ or $x_{\beta,n} = 0$. As in the steep case, we conclude

$$\|x_n - x^*\| = \phi(-\Theta(n)) \tag{C.42}$$

■

We now shift to the payoff-based setting. The relevant result is restated below.

Theorem 4. *Let $x^* \in \mathcal{X}$ be a strategically robust equilibrium of $\mathcal{G}$. Fix a confidence level $\delta > 0$, and let $(x_n)_{n \in \mathbb{N}}$ be the iterates of* (FTRL) *run with* (SPSA) *with $\varepsilon_n \propto 1/n^p$ for some $p \in (0, 1/2)$ and step-size $\gamma > 0$ sufficiently small. Then, there exists a neighborhood $\mathcal{U}$ of x^* such that:*

$$\mathbb{P}\Big(\lim_{n\to\infty} x_n = x^*\Big) \ge 1 - \delta \qquad \text{if } x_1 \in \mathcal{U}. \tag{18}$$

If, in addition, $\mathcal{X}$ is affinely constrained and h is decomposable with kernel θ, then, whenever x_n converges to x^*, we have:

$$\|x_n - x^*\| = \phi(-\Theta(n))\,. \tag{19}$$

Proof. First of all, we write $\hat{v}_n$ in the following convenient form:

$$\hat{v}_n = v(x_n) + U_n + b_n \tag{C.43}$$

with

$$U_n = \hat{v}_n - \mathbb{E}[\hat{v}_n \mid \mathcal{F}_n] \qquad \text{and} \qquad b_n = \mathbb{E}[\hat{v}_n \mid \mathcal{F}_n] - v(x_n) \tag{C.44}$$

which, by Proposition A.1, satisfy the bounds $\|U_n\|_* = \mathcal{O}(1/\varepsilon_n)$ and $\|b_n\|_* = \mathcal{O}(\varepsilon_n)$. Now, as in the proof of Theorem 2, $v(x^*)$ lies in the interior of the $\operatorname{PC}(x^*)$. By Lemma A.1 this in turn implies that there exists a polyhedral cone $\mathcal{K}$ generated by $\mathcal{Z} \equiv \{z_1, \dots, z_r\}$ for $r \in \mathbb{N}$, such that $\operatorname{TC}(x^*) \subseteq \mathcal{K}$ and $\langle v(x^*), z \rangle < 0$ for all $z \in \mathcal{Z}$. [5] Therefore, for all $z \in \mathcal{Z}$, we have $\langle v(x^*), z \rangle \le -m$, and by continuity of the vector field v, there exists a neighborhood $\mathcal{U}$ of x^* and $c > 0$ such that $\langle v(x), z \rangle \le -c$ for all $z \in \mathcal{Z}$ and $x \in \mathcal{U}$. Fix some $z \in \mathcal{Z}$. Then, unfolding the evolution of y_n, we have:

$$\begin{aligned}\langle y_{n+1}, z \rangle &= \langle y_n, z \rangle + \gamma \langle \hat{v}_n, z \rangle \\ &= \langle y_n, z \rangle + \gamma \langle v(x_n), z \rangle + \gamma \langle U_n, z \rangle + \gamma \langle b_n, z \rangle \\ &= \langle y_1, z \rangle + \gamma \sum_{k=1}^n \langle v(x_k), z \rangle + \gamma \sum_{k=1}^n \langle U_k, z \rangle + \gamma \sum_{k=1}^n \langle b_k, z \rangle \end{aligned} \tag{C.45}$$

Now, we define the stochastic process $(S_n)_{n \in \mathbb{N}}$ via $S_n \coloneqq \gamma \sum_{k=1}^n \langle U_k, z \rangle$, which is a martingale, since $\mathbb{E}[\langle U_k, z \rangle \mid \mathcal{F}_k] = 0$, and $\mathbb{E}[|\langle U_k, z \rangle|^q \mid \mathcal{F}_k] \le \mathbb{E}[\|U_k\|_*^q \mid \mathcal{F}_k] = \mathcal{O}((1/\varepsilon_n)^q)$.

[5] As before, to resolve any ambiguities, the cone in question here is the polar of the cone provided by Lemma A.1.

Therefore, by Lemma C.1 for $\sigma_n = \Theta(1/\varepsilon_n)$, $q > 2$ and $\mu \in (0,1)$, whose value is determined later, we get:

$$\delta_n \coloneqq \mathbb{P}\left(\sup_{\ell \le n} |S_\ell| > (c/2)(\gamma n)^\mu\right) \le C_q \frac{\gamma^{q(1-\mu)} \sum_{k=1}^n \sigma_k^q}{n^{1+q(\mu-1/2)}} \tag{C.46}$$

where C_q is a constant that depends only on c and q. Thus, for $\varepsilon_n = \varepsilon/n^p$, there exist $B > 0$ such that:

$$\sum_{k=1}^n \sigma_k^q \le B\varepsilon^{-q} \sum_{k=1}^n k^{pq} \le B'\varepsilon^{-q} n^{1+pq} \tag{C.47}$$

where we used that $\sum_{k=1}^n k^{pq} = \Theta(n^{1+pq})$. So, using the above bound, (C.48) becomes:

$$\delta_n \le C_q' \frac{\gamma^{q(1-\mu)} \varepsilon^{-q} n^{1+pq}}{n^{1+q(\mu-1/2)}} \le C_q' \frac{\gamma^{q(1-\mu)} \varepsilon^{-q}}{n^{q(\mu-1/2-p)}} \tag{C.48}$$

Thus, we readily have that:

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{n\ge 1} \left\{\sup_{\ell \le n} |S_\ell| \le c(\gamma n)^\mu\right\}\right) &= 1 - \mathbb{P}\left(\bigcup_{n\ge 1} \left\{\sup_{\ell \le n} |S_\ell| > c(\gamma n)^\mu\right\}\right) \\
&\ge 1 - \sum_{n=1}^\infty \mathbb{P}\left(\sup_{\ell \le n} |S_\ell| > c(\gamma n)^\mu\right) \\
&\ge 1 - \sum_{n=1}^\infty \delta_n
\end{aligned} \tag{C.49}$$

Now, we need to show that there exists $\mu \in (0,1)$ and $q > 2$ such that $\sum_{n=1}^\infty \delta_n \le \delta/r$. In order for the sum to be finite, we need $q(\mu - 1/2 - p) > 1$ which readily implies that $p < \mu - 1/2 - 1/q$.

For the bias term, since $\|b_k\|_* = \Theta(\varepsilon_k)$, we have:

$$\sum_{k=1}^n \langle b_k, z\rangle \le \sum_{k=1}^n \|b_k\|_* \|z\| \le \sum_{k=1}^n \|b_k\|_* \le D \sum_{k=1}^n \varepsilon_k \le D \sum_{k=1}^n \varepsilon/k^p \le D'\varepsilon n^{1-p} \tag{C.50}$$

for some $D' > 0$, where in the last inequality we used that $\sum_{k=1}^n 1/k^p = \Theta(n^{1-p})$.

Therefore, for $1 - \mu < p$, and $\gamma < 1$, we readily get that

$$\gamma \sum_{k=1}^n \langle b_k, z\rangle \le D'\gamma\varepsilon n^\mu \le D'\varepsilon(\gamma n)^\mu \tag{C.51}$$

Therefore, we need to satisfy

$$1 - \mu < p < \mu - 1/2 - 1/q \tag{C.52}$$

from which we obtain $\mu \in (3/4, 1)$. Thus, for $p \in (0, 1/2)$, there exist $\mu \in (3/4, 1)$ and $q > 2$ that satisfy (C.51). So, for ε, γ sufficiently small we can guarantee that

$$\gamma \sum_{k=1}^n \langle b_k, z\rangle \le (c/2)(\gamma n)^\mu \qquad \text{and} \qquad \sum_{n=1}^\infty \delta_n \le \delta/r \tag{C.53}$$

Therefore, with probability at least $1 - \delta/r$, the template inequality becomes:

$$\begin{aligned}
\langle y_{n+1}, z\rangle &\le \langle y_1, z\rangle + \gamma \sum_{k=1}^n \langle v(x_k), z\rangle + (c/2)(\gamma n)^\mu + (c/2)(\gamma n)^\mu \\
&\le \langle y_1, z\rangle + \gamma \sum_{k=1}^n \langle v(x_k), z\rangle + c(\gamma n)^\mu
\end{aligned} \tag{C.54}$$

Initializing y_1 such that $\langle y_1, z'\rangle < -M - c$ for all $z' \in \mathcal{Z}$, we have $\langle y_n, z\rangle < -M$ for all $n \in \mathbb{N}$ with probability at least $1 - \delta/r$. To see this, we proceed by induction, and suppose that $\langle y_k, z\rangle < -M$ for all $k = 1, \dots, n$. Then

$$\langle y_{n+1}, z\rangle = \langle y_1, z\rangle + \gamma \sum_{k=1}^n \langle v(x_k), z\rangle + \gamma \sum_{k=1}^n \langle U_k, z\rangle + \gamma \sum_{k=1}^n \langle b_k, z\rangle$$

$$\leq -M - c - c\gamma n + c(\gamma n)^{\mu} \tag{C.55}$$

For $n \in \mathbb{N}$ with $\gamma n < 1$, we have $-c + c(\gamma n)^{\mu} < 0$, while for $\gamma n \geq 1$, it holds $-c\gamma n + c(\gamma n)^{\mu} < 0$. In both cases, we conclude that $\langle y_{n+1}, z\rangle < -M$, and by induction, we get the inequality.

Therefore, with probability at least $1 - \delta/r$, we have:

$$\langle y_{n+1}, z\rangle \leq -M - c - c\gamma n + c(\gamma n)^{\mu} \tag{C.56}$$

and sending $n \to \infty$, we get $\langle y_n, z\rangle \to -\infty$.

As a final step, using the same argument for all $z \in \mathcal{Z}$ and applying a union bound, we have that $\langle y_n, z\rangle \to -\infty$ with probability at least $1 - \delta$, and by Proposition C.1, we get that

$$\mathbb{P}\Big(\lim_{n\to\infty} x_n = x^*\Big) \geq 1 - \delta\,. \tag{C.57}$$

If, in addition, $\mathcal{X}$ is affinely constrained and h is decomposable with kernel θ, then the argument in the proof of Theorem 3 applies verbatim, yielding:

$$\|x_n - x^*\| = \phi(-\Theta(n))\,. \tag{C.58}$$

whenever x_n converges to x^*. ■

We conclude this appendix with Proposition 3, which illustrates that an extreme, non-strategically robust equilibrium may exhibit fundamentally different behavior depending on the choice of regularizer.

Proposition 3. *Consider the 1-player game $\mathcal{G}$ with $\mathcal{X} = [0,1]$, $u(x) = -\frac{3}{4}x^{4/3}$ and $x^* = 0$. Let $(x_n)_{n\in\mathbb{N}}$ be the iterates of (FTRL) with $\gamma < 1$, and $\hat{v}_n = v(x_n) + U_n$, where U_n are i.i.d. standard normal random variables for all $n \in \mathbb{N}$. Then, for any initial condition $y_1 \in \mathbb{R}$, we have:*

(i) For $h(x) = x\log x$, it holds $\mathbb{P}(\lim_{n\to\infty} x_n = x^*) = 0$.

(ii) For $h(x) = -2\sqrt{x}$, it holds $\mathbb{P}(\lim_{n\to\infty} x_n = x^*) = 1$.

Proof. We show each case separately.

(i) Writing down the (FTRL) dynamics, we have

$$\begin{aligned} y_{n+1} &= y_n + \gamma(-x_n^{1/3} + U_n) \\ x_n &= \sup_{x\in[0,1]} (y_n x - x\log x) \end{aligned} \tag{C.59}$$

Solving the maximization problem in the definition of x_n, we obtain:

$$x_n = \begin{cases} \exp(y_n - 1), & \text{if } y_n \leq 1 \\ 1, & \text{if } y_n > 1 \end{cases}$$

or, equivalently, $x_n = \mathbb{1}(y_n > 1) + \mathbb{1}(y_n \leq 1)\exp(y_n - 1)$ and the dual process can be written as:

$$y_{n+1} = y_n - \gamma\,\mathbb{1}(y_n > 1) - \gamma\,\mathbb{1}(y_n \leq 1)\exp\big((y_n - 1)/3\big) + \gamma U_n \tag{C.60}$$

It is clear that $x_n \to 0$ if and only if $y_n \to -\infty$ as n goes to infinity. For notational convenience, set $z_n \equiv -y_n$. Then, the evolution of the dual process becomes:

$$z_{n+1} = z_n + \gamma\,\mathbb{1}(z_n < -1) + \gamma\,\mathbb{1}(z_n \geq -1)\exp\big((-z_n - 1)/3\big) - \gamma U_n \tag{C.61}$$

Now, define the process

$$z'_{n+1} \equiv \big(z'_n + \gamma\,\mathbb{1}(z'_n < -1) + \gamma\,\mathbb{1}(z'_n \geq -1)\exp\big((-z'_n - 1)/3\big) - \gamma U_n\big)^+, \quad z'_1 = z_1 \tag{C.62}$$

where U_n is the same random variable as in (C.61).

The rest of our proof relies on a series of claims, which we state and prove one-by-one.

Claim 1. *The process $(z'_n)_{n\in\mathbb{N}}$ dominates $(z_n)_{n\in\mathbb{N}}$, i.e., $z'_n \geq z_n$ for all $n \in \mathbb{N}$.*

The proof of Claim 1 lies at the end. Now, invoking Theorem A.3 with

$$f(z) \equiv \gamma\,\mathbb{1}(z_n' < -1) + \gamma\,\mathbb{1}(z_n' \geq -1)\exp\big((-z_n' - 1)/3\big) \tag{C.63}$$

bounded and $\sigma^2 = \gamma^2$, it holds that

$$f(z) \leq \frac{\sigma^2}{2z} \quad \text{for all } z \text{ large enough} \tag{C.64}$$

Thus, $\big(z_n'\big)_{n\in\mathbb{N}}$ is recurrent, which implies that

$$\mathbb{P}\Big(\lim_{n\to\infty} z_n' = \infty\Big) = 0 \tag{C.65}$$

Finally, since $\big(z_n'\big)_{n\in\mathbb{N}}$ dominates $(z_n)_{n\in\mathbb{N}}$ by Claim 1, we obtain

$$\Big\{\lim_{n\to\infty} z_n = \infty\Big\} \subseteq \Big\{\lim_{n\to\infty} z_n' = \infty\Big\} \tag{C.66}$$

which implies that

$$\mathbb{P}\Big(\lim_{n\to\infty} x_n = x^*\Big) = \mathbb{P}\Big(\lim_{n\to\infty} z_n = \infty\Big) \leq \mathbb{P}\Big(\lim_{n\to\infty} z_n' = \infty\Big) = 0 \tag{C.67}$$

and the result follows.

Proof of Claim 1. Consider the function

$$g(z) \coloneqq z + \gamma\,\mathbb{1}(z_n < -1) + \gamma\,\mathbb{1}(z_n \geq -1)\exp\big((-z_n - 1)/3\big) \tag{C.68}$$

Then:

- For $z < -1$: $g'(z) = 1$
- For $z > -1$: $g'(z) = 1 - \gamma/3\,\exp\big((-z-1)/3\big) > 1 - \gamma/3 > 0$

Thus, g is strictly increasing for all $z \in \mathbb{R}$. Now, for the sake of contradiction, suppose that there exists $\omega \in \Omega$ and a first time $k+1 \in \mathbb{N}$ where the dominance does not hold, i.e.,

$$z_k(\omega) \leq z_k'(\omega) \quad \text{and} \quad z_{k+1}'(\omega) < z_{k+1}(\omega) \tag{C.69}$$

By the monotonicity property of g, we get that

$$g(z_k(\omega)) \leq g(z_k'(\omega)) \tag{C.70}$$

and, therefore, adding $-\gamma U_k(\omega)$ in both sides

$$\begin{aligned} z_{k+1}(\omega) &\leq z_k' + \gamma\,\mathbb{1}(z_n' < -1) + \gamma\,\mathbb{1}(z_n' \geq -1)\exp\big((-z_n' - 1)/3\big) - \gamma U_k \\ &\leq \big(z_k' + \gamma\,\mathbb{1}(z_n' < -1) + \gamma\,\mathbb{1}(z_n' \geq -1)\exp\big((-z_n' - 1)/3\big) - \gamma U_k\big)^+ \\ &\leq z_{k+1}'(\omega) \end{aligned} \tag{C.71}$$

which is a contradiction. Thus, the proof of Claim 1 is complete.

(ii) In this setup, the (FTRL) dynamics are described by the system

$$\begin{aligned} y_{n+1} &= y_n + \gamma(-x_n^{1/3} + U_n) \\ x_n &= \sup_{x\in[0,1]}\big(y_n\,x + 2\sqrt{x}\big) \end{aligned} \tag{C.72}$$

Solving the maximization problem in the definition of x_n, we obtain:

$$x_n = \begin{cases} (-y_n)^{-2}, & \text{if } y_n \leq -1 \\ 1, & \text{if } y_n > -1 \end{cases} \tag{C.73}$$

or, equivalently, $x_n = \mathbb{1}(y_n > -1) + \mathbb{1}(y_n \leq -1)(-y_n)^{-2}$ and the dual process can be written as:

$$y_{n+1} = y_n - \gamma\,\mathbb{1}(y_n > -1) - \gamma\,\mathbb{1}(y_n \leq -1)(-y_n)^{-2/3} + \gamma U_n \tag{C.74}$$

For notational convenience, set $z_n \equiv -y_n$. Then, the evolution of the dual process becomes:

$$z_{n+1} = z_n + \gamma\,\mathbb{1}(z_n < 1) + \gamma\,\mathbb{1}(z_n \geq 1) z_n^{-2/3} - \gamma U_n \tag{C.75}$$

It is clear that $x_n \to 0$ if and only if $z_n \to \infty$ as n goes to infinity. Now, define the process

$$z_{n+1}' = \Big(z_n' + \gamma\,\mathbb{1}(z_n' < 1) + \gamma\,\mathbb{1}(z_n' \geq 1) {z_n'}^{-2/3} - \gamma U_n\Big)^+, \quad z_1' = z_1 \tag{C.76}$$

where U_n is the same randomness as in (C.75).

Claim 2. *The process* $(z'_n)_{n\in\mathbb{N}}$ *dominates* $(z_n)_{n\in\mathbb{N}}$*, i.e.,* $z'_n \geq z_n$ *for all* $n \in \mathbb{N}$.

The proof of Claim 2 lies at the end. Now, invoking Theorem A.3 with

$$f(z) \equiv \gamma\,\mathbb{1}(z<1) + \gamma\,\mathbb{1}(z \geq 1) z^{-2/3} \tag{C.77}$$

bounded, $\sigma^2 = \gamma^2$, and $\theta > 1$, we have

$$f(z) \geq \frac{\sigma^2\theta}{2z} \quad \text{for all } z \text{ large enough} \tag{C.78}$$

Thus, $\left(z'_n\right)_{n\in\mathbb{N}}$ is transient, which implies that $\mathbb{P}(A) = 1$ for $A = \{\omega \in \Omega : \lim_{n\to\infty} z'_n(\omega) = \infty\}$.

Now, fix some $\omega \in A$. Since $\lim_{n\to\infty} z'_n(\omega) = \infty$, there exists $n_\omega \in \mathbb{N}$ such that $z'_n > 1$ for all $n \geq n_\omega$, and therefore

$$\begin{aligned} z'_{n+1} &= z'_n + \gamma {z'_n}^{-2/3} - \gamma U_n \\ &= z'_{n_\omega} + \gamma \sum_{k=n_\omega+1}^{n} \left({z'_k}^{-2/3} - U_k\right) \end{aligned} \tag{C.79}$$

from which we conclude that

$$\sum_{k=n_\omega+1}^{n} \left({z'_k}^{-2/3} - U_k\right) \to \infty \quad \text{as } n \to \infty \tag{C.80}$$

Finally, we have that

$$z_{n+1} = z_{n_\omega} + \gamma \sum_{k=n_\omega+1}^{n} \left(\mathbb{1}(z_k<1) + \mathbb{1}(z_k \geq 1) z_k^{-2/3} - U_k\right) \tag{C.81}$$

and, since $(z'_n)_{n\in\mathbb{N}}$ dominates $(z_n)_{n\in\mathbb{N}}$, and $z'_n > 1$ for all $n \geq n_\omega$, we readily get that

$$\sum_{k=n_\omega+1}^{n} \left(\mathbb{1}(z_k<1) + \mathbb{1}(z_k \geq 1) z_k^{-2/3} - U_k\right) \geq \sum_{k=n_\omega+1}^{n} \left({z'_k}^{-2/3} - U_k\right) \tag{C.82}$$

Thus, by (C.80), we conclude that

$$\sum_{k=n_\omega+1}^{n} \left(\mathbb{1}(z_k<1) + \mathbb{1}(z_k \geq 1) z_k^{-2/3} - U_k\right) \to \infty \quad \text{as } n \to \infty \tag{C.83}$$

which implies that $\lim_{n\to\infty} z_n(\omega) = \infty$. Therefore, we obtain that $\lim_{n\to\infty} z_n(\omega) = \infty$ for all $\omega \in A$, and since $\mathbb{P}(A) = 1$, it follows that

$$\mathbb{P}\left(\lim_{n\to\infty} x_n = x^*\right) = \mathbb{P}\left(\lim_{n\to\infty} z_n = \infty\right) = 1 \tag{C.84}$$

and the proof is complete.

Proof of Claim 2. Consider the function

$$g(z) := z + \gamma\,\mathbb{1}(z<1) + \gamma\,\mathbb{1}(z \geq 1) z^{-2/3} \tag{C.85}$$

Then:

- For $z < 1$: $g'(z) = 1 + \gamma > 0$
- For $z > 1$: $g'(z) = 1 - 2\gamma z^{-5/3}/3 > 1 - 2\gamma/3 > 0$

Thus, g is strictly increasing for all $z \in \mathbb{R}$. Now, for the sake of contradiction, suppose that there exists $\omega \in \Omega$ and a first time $k+1 \in \mathbb{N}$ where the dominance does not hold, i.e.,

$$z_k(\omega) \leq z'_k(\omega) \quad \text{and} \quad z'_{k+1}(\omega) < z_{k+1}(\omega) \tag{C.86}$$

By the monotonicity property of g, we get that

$$g(z_k(\omega)) \leq g(z'_k(\omega)) \tag{C.87}$$

and therefore, adding $-\gamma U_k(\omega)$ in both sides

$$\begin{aligned} z_{k+1}(\omega) &\leq z'_k + \gamma\,\mathbb{1}(z'_k<1) + \gamma\,\mathbb{1}(z'_k \geq 1) {z'_k}^{-2/3} - \gamma U_k \\ &\leq \left(z'_k + \gamma\,\mathbb{1}(z'_k<1) + \gamma\,\mathbb{1}(z'_k \geq 1) {z'_k}^{-2/3} - \gamma U_k\right)^+ \\ &\leq z'_{k+1}(\omega) \end{aligned} \tag{C.88}$$

which is a contradiction. Thus, the proof of Claim 2 is complete. ∎

NeurIPS Paper Checklist

1. **Claims**

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

   Answer: [Yes]

   Justification: It can be found in Section 3, Section 4 and the appendix.

   Guidelines:

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

2. **Limitations**

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

   Answer: [Yes]

   Justification: It can be found in Section 1.

   Guidelines:

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

3. **Theory assumptions and proofs**

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

   Answer: [Yes]

   Justification: It can be found in Section 3, Section 4 and the appendix.

Guidelines:

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

4. **Experimental result reproducibility**

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

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

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

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

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

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

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

6. **Experimental setting/details**

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

   Answer: [NA]

   Justification: The paper does not include experiments.

   Guidelines:

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

7. **Experiment statistical significance**

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

   Answer: [NA]

   Justification: No statistical significance applicable.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).

- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

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

   Answer: [NA]

   Justification: The paper does not include experiments.

   Guidelines:

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

9. **Code of ethics**

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

   Answer: [Yes]

   Justification: The paper conforms with the NeurIPS Code of Ethics.

   Guidelines:

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

10. **Broader impacts**

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

    Answer: [NA]

    Justification: There is no societal impact.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.

- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

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

    Answer: [NA]

    Justification: There are no such risks.

    Guidelines:

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

12. **Licenses for existing assets**

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

    Answer: [NA]

    Justification: The paper does not use existing assets.

    Guidelines:

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

13. **New assets**

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

    Answer: [NA]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.

- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

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

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

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

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

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

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

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

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

16. **Declaration of LLM usage**

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

    Answer: [NA]

    Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

    Guidelines:

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