

Learnable Mixed Nash Equilibria are Collectively Rational

Geelon So

Yi-An Ma

University of California, San Diego

Abstract

We extend the study of learning in games to dynamics that exhibit *non-asymptotic stability*. We do so through the notion of *uniform stability*, which is concerned with equilibria of individually utility-seeking dynamics. Perhaps surprisingly, it turns out to be closely connected to economic properties of collective rationality. Under mild non-degeneracy conditions and up to strategic equivalence, if a mixed equilibrium is not uniformly stable, then it is not *weakly Pareto optimal*—there is a way for all players to improve by jointly deviating from the equilibrium. On the other hand, if it is locally uniformly stable, then the equilibrium must be weakly Pareto optimal. Moreover, we show that uniform stability determines the last-iterate convergence behavior for the family of *incremental smoothed best-response dynamics*, used to model individual and corporate behaviors in the markets. Unlike dynamics around strict equilibria, which can stabilize to socially-inefficient solutions, individually utility-seeking behaviors near mixed Nash equilibria lead to collective rationality.

Keywords: algorithmic game theory, evolutionary dynamics, non-asymptotic stability, last-iterate convergence, smoothed best-response

1 Introduction

The Nash (1951) equilibrium is a foundational solution concept in games, capturing when collective behavior or strategies may be stationary. These equilibria are meaningful to study, for once such strategies appear, they may persist for a long time. But, there is an important caveat: not all equilibria can be robustly reached by players in the game (Hart and Mas-Colell, 2003; Papadimitriou, 2007; Daskalakis et al., 2009, 2010; Milionis et al., 2023). Any equilibrium that cannot be found or sustained is unlikely to have practical relevance. This caveat leads to the question on *learnability*: which Nash equilibria can players eventually learn to play from repeated interactions?

To grasp individual and corporate behaviors, literature in classical economics considers a model of learning where players: (i) take an evolutionary approach and incrementally update their strategies; (ii) are utility-seeking, which means that they aim to improve their own payoffs; and (iii) are uncoupled, where players are unaware of the other players' utilities or methods to improve them (Alchian, 1950; Winter, 1971). The goal of the players in this model is not to compute any pre-determined Nash equilibrium, at least not explicitly. Rather, the equilibrium is to emerge out of their joint, but individually utility-seeking, behavior.

Within this model of learning, the impossibility result of Hart and Mas-Colell (2003) shows that no *uncoupled* and *asymptotically-stable* learning dynamics can converge to all Nash equilibria. Uncoupled means that players learn in uncoordinated and decentralized ways. Asymptotic stability means that the convergence is robust to small perturbations in the dynamics. Simply put, not all equilibria admit such a strong notion of learnability, viz. dynamical stability. The ones that do, however, have been characterized for certain classes of learning dynamics in standard normal-form games: an equilibrium is ‘asymptotically learnable’ in this way if and only if it is *strict*, where every player has a single, deterministic strategy that is clearly locally optimal (Samuelson and Zhang, 1992; Vlatakis-Gkaragkounis et al., 2020; Giannou et al., 2021).

A significant gap remains for the learnability of *mixed* Nash equilibria, where players may use randomized strategies. On the one hand, mixed equilibria are not strict, and as a result, they are not asymptotically stable under these learning dynamics. Observations corroborating this finding demonstrate that many dynamics are not able to generically learn mixed Nash equilibria. This has been a significant source of criticism on the

Correspondence to: geelon@ucsd.edu and yianma@ucsd.edu

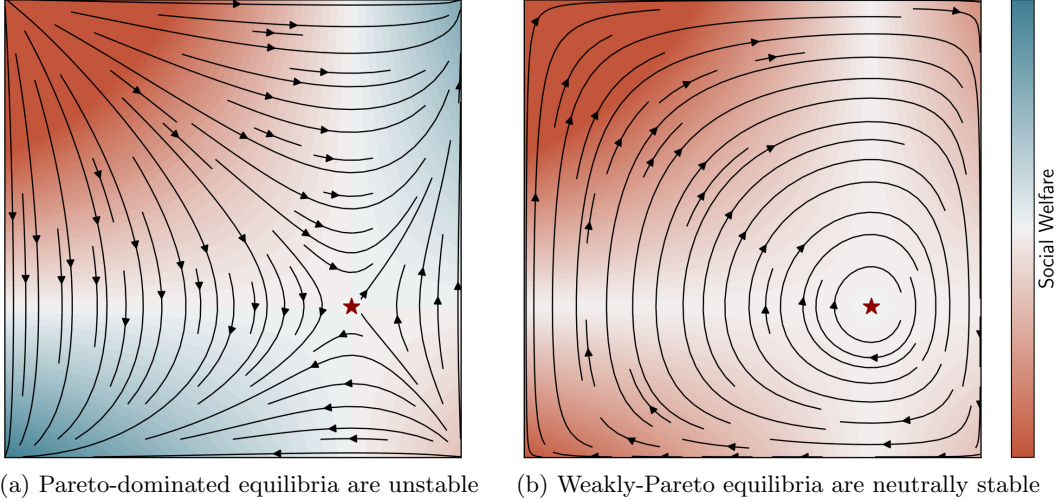


FIGURE 1: Uncoupled learning dynamics in two 2×2 normal-form games with purely-strategic utilities $\mathbf{f} = (f_1, f_2)$. The Nash equilibria are marked by stars. The streamlines visualize the trajectories of the learning dynamics. The heatmap plots the social welfare function $\min\{f_1, f_2\}$, measuring the utility of the player worst-off. The heatmap is white where the utilities are equal to the equilibrium; the darker the red, the worse the social welfare; the darker the blue, the better. (a) The mixed Nash equilibrium in this game is not weakly Pareto optimal, and the dynamics are unstable. (b) This equilibrium is weakly Pareto optimal, and the dynamics are non-asymptotically stable. The dynamics visualized here is continuous-time mirror ascent induced by the entropy mirror map.

viability of mixed equilibria as solutions in games (Shapley, 1963; Crawford, 1985; Stahl II, 1988; Jordan, 1993; Krishna and Sjöström, 1998; Ellison and Fudenberg, 2000; Kleinberg et al., 2011; Mertikopoulos et al., 2018; Bailey et al., 2021). On the other hand, there are cases in which mixed equilibria *are* approachable by simple, uncoupled dynamics, such as those in two-player zero-sum games (Robinson, 1951; Fudenberg and Kreps, 1993; Gjerstad, 1996; Hofbauer and Sigmund, 1998; Benam and Hirsch, 1999; Hofbauer and Sorin, 2006; Daskalakis and Panageas, 2019; Wei et al., 2021; Piliouras et al., 2022). To address this conundrum, we must go beyond the criterion of asymptotic stability, which is too stringent a requirement.

This work extends the study of learning in games to dynamics that exhibit *non-asymptotic* stability. It is a weaker notion of stability that includes neutral stability, which only requires that players starting off in a neighborhood of such an equilibrium will continue to remain close to it. The question we ask is:

Under the relaxed criterion of non-asymptotic stability, which Nash equilibria are learnable by uncoupled dynamics, and what are their economic properties?

There are two parts to this work. In the first part, we introduce and characterize a form of non-asymptotic stability that applies to broad classes of learning dynamics. We call it *uniform stability*. While it arises from dynamical considerations, we show that uniform stability is closely tied to the economic properties of the equilibrium, namely, *strategic Pareto optimality*. This is a game-variant of weak Pareto optimality defining a minimal form of collective rationality—it describes solutions where there is no way to strictly improve everyone’s utilities, up to strategic equivalence.¹ We show that an equilibrium that is not uniformly stable must not be strategically Pareto optimal. This formalizes and tightens an observation made by prior work: non-convergence to mixed equilibria can sometimes actually be a blessing in disguise, where players would be worse off had they managed to converge to the equilibrium (Kleinberg et al., 2011; Ostrovski and van Strien, 2013; Pangallo et al., 2022; Anagnostides et al., 2022). This may be surprising as it runs counter to the more prominent phenomenon where non-cooperative, individual behavior leads to worse social outcomes, or a high *price of anarchy*, such as in the Prisoner’s Dilemma or the Tragedy of the Commons (Luce and Raiffa, 1957; Hardin, 1968; Nachbar, 1990; Weibull, 1994; Koutsoupas and Papadimitriou, 1999).

In the second part, we focus on convergence and non-convergence of *incremental smoothed best-response dynamics*, a generalization of a canonical family of learning dynamics. For this class, we show that non-strict,

¹As standard learning dynamics generally factor through strategic equivalence classes of games (they do not depend on the ‘externalities’ or non-strategic components of the game), this reservation is necessary.

uniformly-stable equilibria can be learned, though with slower convergence than strict equilibria. Around a locally uniformly-stable Nash equilibrium, these dynamics can be *stabilized*: they lead to asymptotically-stable dynamics that approximate the equilibrium to arbitrary accuracy. While for non-uniformly stable mixed equilibria, there are dynamics that can never be stabilized to the equilibrium.

Thus we observe a dichotomy between dynamics around strict and mixed Nash equilibria under the context of ‘as if’ rationality (Friedman, 1953; Weibull, 1994). In the former, individually utility-seeking dynamics lead to the same equilibrium behaviors as if all the players have absolute, individual rationality. As in the prisoner’s dilemma, they can stabilize at collectively-irrational behaviors. Around mixed equilibria, uniform learnability disallows collective irrationality. As a result, individually utility-seeking dynamics robustly behave as if the players were collectively rational, reminiscent of the ‘invisible hand’ from classical economic theory (Smith, 1776; Debreu, 1959).

1.1 Main results

We consider uncoupled learning dynamics for standard N -player normal-form games, where players, through repeated interactions, evolve their individual strategies in order to maximize their own utilities.

For such dynamics, in Section 3, we introduce two non-asymptotic stability classes for mixed equilibria: those that have *pointwise uniform stability* and those that exhibit a stronger *local uniform stability*. These impose a second-order differential condition, either only at the equilibrium or in a neighborhood around it (note that the equilibrium condition itself is a first-order condition for mixed strategies). These conditions naturally arise from linear stability analysis, and they apply to broad classes of dynamics such as smoothed best-response, gradient ascent/adaptive dynamics, mirror ascent/positive-definite adaptive dynamics.

In Section 4, we show the chain of implications for mixed Nash equilibria under mild assumptions:

$$\text{local uniform stability} \implies \text{strategic Pareto optimality} \implies \text{pointwise uniform stability}.$$

Moreover, all of these solution concepts are equivalent for the class of *graphical* or *polymatrix games*, which are particularly succinct normal-form games that decompose across pairwise interactions between players.

In Section 5, we prove convergence results for the family of *incremental, smoothed best-response learning dynamics*. The dynamics are controlled by a smoothing parameter β and a learning rate η . The first governs the quality of any approximate Nash equilibria that the dynamics finds, while the second impacts the stability of its fixed points; whether the dynamics converge and at what rate. We prove that if a Nash equilibrium is locally uniformly stable, then for any choice of approximation β , the dynamics can always be stabilized to the equilibrium by taking a sufficiently fine learning rate η , whereby it converges with asymptotic stability. With T iterations, the overall convergence rate to the locally uniformly stable mixed Nash equilibria is of order $T^{-1/2}$. However, if a Nash equilibrium is not pointwise uniformly stable, then there are dynamics that can never be stabilized beyond a certain approximation error, no matter what learning rate is chosen.

Finally, as the first set of results are stated for fully-mixed equilibria, Section 6 extends these results to the case where the Nash equilibrium may be partially mixed. The following summarizes our main results:

Local uniform stability implies strategic Pareto optimality. If a Nash equilibrium falls within a region of uniformly-stable strategies, it must be collectively rational (Theorem 1). In this case, it is robustly approximable under incremental, smoothed best-response dynamics (Theorems 3 and 4).

Strategic Pareto optimality implies pointwise uniform stability. We show that strategic Pareto optimality implies *strategic Pareto stationarity*, which is a second-order condition for collective rationality (Proposition 1). Strategic stationarity is equivalent to pointwise uniform stability (Theorem 2), and it is also necessary for convergence with non-asymptotic stability (Proposition 2).

1.2 Related work

This work considers the convergence of learning dynamics in games. Two modes of convergence are often studied (i) time-averaged convergence, where players aim to achieve little or no regret in the long run, or the stronger (ii) day-to-day or last-iterate convergence, where players eventually learn to play an equilibrium of the game; see also Fudenberg and Kreps (1993) for discussion. This work falls in the latter category.

The learnability of strict Nash equilibria is fairly well-understood (Hart and Mas-Colell, 2003; Mertikopoulos and Sandholm, 2016; Vlatakis-Gkaragkounis et al., 2020; Giannou et al., 2021, and related works therein). In contrast, the learnability of mixed Nash equilibria is more complicated. Many works have pointed out that asymptotically-stable convergence to mixed Nash equilibria is generally ruled out because linearized dynamics about them have zero trace (Crawford, 1985; Jordan, 1993; Singh et al., 2000; Hart and Mas-Colell, 2003; Mertikopoulos et al., 2018). As mixed Nash equilibria are needed for the existence of Nash equilibria (Nash, 1951), their generic infeasibility is troubling for the viability of the solution concept, but has a resolution via Harsanyi’s purification theorem (Harsanyi, 1973).

Philosophically, unlike Harsanyi (1973), this work does not try to save all mixed equilibria, but proceeds from the fact that uncoupled dynamics can sometimes converge to mixed Nash equilibria, and aims to further refine the solution concept (Van Damme, 1992). Much of the existing work providing positive convergence results focuses on two-player games (Robinson, 1951; Fudenberg and Kreps, 1993; Gjerstad, 1996; Hofbauer and Sigmund, 1998; Benaim and Hirsch, 1999; Hofbauer and Sorin, 2006; Wei et al., 2021), and especially for zero-sum games (Gilpin et al., 2012; Daskalakis and Panageas, 2019; de Montbrun and Renault, 2022; Piliouras et al., 2022, and related works therein). Our results apply to general N -player normal-form games.

The results and techniques in this paper also relate to many areas of game theory. The connection between uniform stability and weak Pareto optimality generalizes existing characterizations of games with weakly Pareto optimal equilibria in two-player games (Moulin and Vial, 1978; Adler et al., 2009), and bridges to the classical notion of strong Nash equilibria (Aumann, 1959). Our notion of uniform stability builds on the game Jacobian (also sometimes called the game Hessian), which is also generally used to analyze smooth games (Isaacs, 1999; Candogan et al., 2011; Balduzzi et al., 2018; Letcher et al., 2019).

The algorithmic component of this paper focuses on smoothed best-response dynamics, which can be viewed as regularized learning/optimization (Rockafellar, 1997; Mertikopoulos and Sandholm, 2016). In normal-form games, the canonical regularizer is the logit map, yielding quantal response equilibria (McKelvey and Palfrey, 1995; Alós-Ferrer and Netzer, 2010; Goeree et al., 2016; Mertikopoulos and Sandholm, 2016). Our convergence result introduces the notion of stabilization to an equilibrium, which is similar to the tracing procedure discussed by McKelvey and Palfrey (1995).

2 Preliminaries

An N -player game $(\Omega, \mathbf{f})$ with players indexed by $n \in [N]$ consists of a *joint strategy space* $\Omega = \Omega_1 \times \dots \times \Omega_N$ and a set of *utilities* $\mathbf{f} = (f_1, \dots, f_N)$, where each player would like to maximize their own utility $f_n : \Omega \rightarrow \mathbb{R}$. In a round of the game, everyone independently chooses a strategy $x_n \in \Omega_n$. These choices constitute the joint decision vector $\mathbf{x} = (x_1, \dots, x_N)$, and players receive their respective payoffs $f_n(\mathbf{x}) \equiv f_n(x_n; \mathbf{x}_{-n})$. Here, we let $\mathbf{x}_{-n} = (x_1, \dots, x_{n-1}, x_{n+1}, \dots, x_N)$ denote the joint strategy without the n th component.

This work focuses on *normal-form games*, which we define here and expand upon in Section A.

Definition 1 (Normal-form game). An N -player *normal-form game* $(\Omega, \mathbf{f})$ is one where each player $n \in [N]$ has a set of k_n alternatives/pure strategies, the strategy space Ω_n is the $(k_n - 1)$ -dimensional probability simplex Δ^{k_n-1} , and each utility f_n is defined by a payoff tensor $T^n \in \mathbb{R}^{k_1 \times \dots \times k_N}$ as follows:

$$f_n(\mathbf{x}) = \sum_{i_1 \in [k_1]} \dots \sum_{i_N \in [k_N]} T^n_{i_1, \dots, i_N} x_{1, i_1} \dots x_{N, i_N} \quad \text{and} \quad x_n = (x_{n,1}, \dots, x_{n,k_n}) \in \Delta^{k_n-1}.$$

We say that the strategy $\mathbf{x} \in \Omega$ is *mixed* if it is in the interior of Ω , which we denote by $\text{int}(\Omega)$.

Normal-form games are a special case of *multilinear games*, which are games whose utilities are polynomials where the maximum degree in each variable is one. In the following, we first define multilinear polynomials of the form $p : \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_N} \rightarrow \mathbb{R}$, where we think of $d_n = k_n - 1$. The connection between normal-form games and multilinear games is formalized by Lemma A.1.

Definition 2 (Multilinear polynomial). Let $\mathbb{R}[\mathbf{x}]$ contain the polynomials of the form $p : \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_N} \rightarrow \mathbb{R}$, with variables $\mathbf{x} = (x_1, \dots, x_N)$ and $x_n = (x_{n,1}, \dots, x_{n,d_n})$ for each $n \in [N]$. We say that $p \in \mathbb{R}[\mathbf{x}]$ is *multilinear* over $\mathbf{x}$ if it is affine over each x_n . That is, for each $n \in [N]$, the polynomial is affine in x_n ,

$$p(\mathbf{x}) = A_n(\mathbf{x}_{-n})^\top x_n + b_n(\mathbf{x}_{-n}),$$

where $A_n \in \mathbb{R}[\mathbf{x}_{-n}]^{d_n}$ and $b_n \in \mathbb{R}[\mathbf{x}_{-n}]$ are polynomials, and $\mathbf{x}_{-n}$ includes all variables besides those in x_n .

Definition 3 (Multilinear game). An N -player *multilinear game* $(\Omega, \mathbf{f})$ is a game where for each $n \in [N]$, the strategy space Ω_n is an open, convex subset of $\mathbb{R}^{d_n}$, and the utility $f_n : \Omega \rightarrow \mathbb{R}$ is a multilinear polynomial. That is, each utility f_n extends to a multilinear polynomial on all of $\mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_N}$.

2.1 Individual and collective rationality: two solution concepts

A fundamental aspect of non-cooperative games is that, even though players cannot directly control how others behave, the payoff that each receives nevertheless depends on their joint decision. This leads to the Nash equilibrium as a solution concept for games—joint strategies where it is not possible to improve one’s own utility without coordinating with other players. It is a notion of *individual rationality*:

Definition 4 (Nash equilibrium). A joint decision vector $\mathbf{x}^* \in \Omega$ is a *Nash equilibrium* if:

$$\forall n \in [N], \quad x_n^* \in \arg \max_{x_n \in \Omega_n} f_n(x_n; \mathbf{x}_{-n}^*).$$

In contrast, if as in the multi-objective optimization, players are cooperative and can coordinate, then a minimal solution concept is weak Pareto optimality, which describes joint decisions where there is no way to strictly improve everyone’s utilities. It is a notion of *collective rationality*:

Definition 5 (Weak Pareto optimality). A joint decision vector $\mathbf{x}^* \in \Omega$ is *weakly Pareto optimal* if there does not exist any $\mathbf{x} \in \Omega$ such that $f_n(\mathbf{x}) > f_n(\mathbf{x}^*)$ for all $n \in [N]$.

These solution concepts are generally incomparable and independent of each other. Famous examples, including the *prisoner’s dilemma* or the *tragedy of the commons*, demonstrate that Nash equilibria and Pareto optimal strategies can even be disjoint—in the absence of mechanisms to enable or enforce cooperation, socially-optimal outcomes can be unstable under individually-rational or selfish behavior.

This work refines the relationship between Nash stability and Pareto optimality from the perspective of *learning in games*, which is a dynamical perspective in game theory motivated by a simple observation: a solution concept has no bearing on practice if players of a game cannot find it. Eventually, we will show that the interior Nash equilibria that can be robustly found must, in fact, be strategically Pareto optimal, a notion related to Pareto optimality that is well-suited for games. We will motivate and introduce this in the next section. To do so, we now turn to how players come to discover and play a Nash equilibrium.

2.2 Uncoupled learning dynamics and strategic equivalence

Consider a model of learning where players learn to play the game via trial-and-error: they simply play many rounds of the game. As players may observe the joint strategies $\mathbf{x}(t)$ chosen over time t , they can also continually adjust their own strategies. A *learning rule/dynamics* specifies how a player $n \in [N]$ chooses to act each round $x_n(t)$ as a function of past experiences, namely $\mathbf{x}(s)$ where $s < t$, as well as any additional information they may have about the game—for example, we assume players know their own utilities.

A basic informational constraint we impose is that the players are *uncoupled*, which means that a player’s learning rule cannot depend on the utilities or learning rules of other players (Hart and Mas-Colell, 2003). Intuitively, this means that players do not know what others want, nor how others learn from past experiences. Even if they were so inclined, players do not know how to directly be altruistic or malicious. This constraint in a sense enforces selfish behavior, or individual rationality.

To discuss the connection between the economic and the dynamic properties of games, we need to focus on the ‘active ingredients’ in the utilities that play a role in the dynamics. In existing literature, this motivates the decomposition of utilities into their *strategic* and *non-strategic* components (Candogan et al., 2011). In multilinear games, the strategic component precisely consists of the terms in the polynomial f_n that depends on the n th set of variables x_n , while the non-strategic terms are those that vary independently of x_n . Games with the same strategic components are called *strategically equivalent* (Monderer and Shapley, 1996).

Definition 6 (Strategic and non-strategic components). Let $(\Omega, \mathbf{f})$ be a multilinear game. For each $n \in [N]$, the utility f_n is a polynomial that is affine in x_n , with the following decomposition given in Definition 2,

$$f_n(\mathbf{x}) = \underbrace{A_n(\mathbf{x}_{-n})^\top x_n}_{\text{strategic}} + \underbrace{b_n(\mathbf{x}_{-n})}_{\text{non-strategic}}.$$

The *strategic component* of f_n is the term linear in x_n . The *non-strategic component* of f_n is the offset term. Two utilities are *strategically equivalent* if they have the same strategic component, denoted by $f_n \simeq_{\text{SE}} f'_n$.

The target of this work is to illuminate the game-theoretic properties of equilibria found or not found by simple, uncoupled learning dynamics. As the dynamics we study, Equation (1) below, indeed turn out to be invariant to changes in the non-strategic component of the game (Lemma C.1), we expect results only up to strategic equivalence. This motivates the following notion of strategic Pareto optimality.

Definition 7 (Strategic Pareto optimality). Let $(\Omega, \mathbf{f})$ be a multilinear game. A joint decision vector $\mathbf{x}^* \in \Omega$ is *strategically Pareto optimal* if it is weakly Pareto optimal for the strategic component of the game.

Akin to the classic story about Nash equilibria and weak Pareto optimality, a simple example shows that neither the condition of being a Nash equilibrium nor being strategically Pareto optimal imply each other (Example A.1). But surprisingly, despite uncoupledness, players can be guaranteed to converge to a mixed Nash equilibrium that is strategically Pareto optimal, while they are not guaranteed to converge to a mixed Nash equilibrium without strategic Pareto optimality under mild non-degeneracy assumptions. In the next section, we formally introduce the learning dynamics considered in this work.

2.3 A simple class of learning rules

Let $(\Omega, \mathbf{f})$ be an N -player game where each Ω_n is a convex domain. Let $\mathbf{h} = (h_1, \dots, h_N)$ be a collection of strictly convex regularizers $h_n : \Omega_n \rightarrow \mathbb{R}$. The dynamics we study are based on smoothed best-responses, where players maximize their own utilities by solving a regularized optimization problem:

Definition 8 (β -smoothed best-response). The β -smoothed best-response map $\Phi^\beta = (\Phi_1^\beta, \dots, \Phi_N^\beta) : \Omega \rightarrow \Omega$ with respect to a set of strictly convex regularizers $\mathbf{h}$ for any $\beta > 0$ is given by:

$$\forall n \in [N], \quad \Phi_n^\beta(\mathbf{x}) = \arg \max_{x'_n \in \Omega_n} f_n(x'_n; \mathbf{x}_{-n}) - \beta h_n(x'_n).$$

Throughout this work, we assume that the regularizers are *steep*, which is a standard assumption from convex optimization that guarantees that Φ^β map into the interior of Ω (Rockafellar, 1997, Theorem 26.1).

Definition 9 (Steep regularizer). Let Ω be a compact and convex domain, and $h : \Omega \rightarrow \mathbb{R}$ be strictly convex and smooth on $\text{int}(\Omega)$. It is steep if for any sequence $x_k \in \Omega$ converging to a boundary point of Ω ,

$$\lim_{k \rightarrow \infty} \|\nabla h(x_k)\| = \infty.$$

In the case of normal-form games, a canonical choice is entropy regularization, $h_n(x_n) = \sum_i x_{n,i} \log x_{n,i}$, which is a steep regularizer on the probability simplex. The smoothed best-response map is also called *quantal response*, and the smoothed-best responses are given by the softmax function:

$$\Phi_n^\beta(\mathbf{x}) \propto \exp\left(\frac{1}{\beta} \nabla_n f_n(\mathbf{x})\right).$$

The specific class of discrete-time dynamics that we study in this work is the following:

Incremental smoothed best-response dynamics. Let $\eta \in (0, 1)$ be a learning rate and Φ^β be a β -smoothed best-response map induced by steep regularizers. The (Φ^β, η) -averaging dynamics is the dynamics $\mathbf{x}(t)$ given by the transition map:

$$\mathbf{x}(t) = (1 - \eta) \mathbf{x}(t-1) + \eta \Phi^\beta(\mathbf{x}(t-1)), \quad (1)$$

where the joint strategy $\mathbf{x}(t)$ played in the round t is an weighted average of the strategy $\mathbf{x}(t-1)$ played in the previous round $t-1$ with its smoothed best response $\Phi^\beta(\mathbf{x}(t-1))$.

The fixed points of (1) are precisely the fixed points of the smoothed best-response map Φ^β , which are called *smoothed equilibria*. Importantly, smoothed equilibria are *approximate Nash equilibria* (Lemma B.1), and they converge to Nash equilibria as the smoothness parameter β goes to zero (Lemma B.2). When the dynamics converge, they achieve meaningful solutions in a game-theoretic sense.

Definition 10 (β -smoothed equilibrium). Let Φ^β be a β -smoothed best-response map. A joint decision vector $\mathbf{x}^\beta \in \Omega$ is a β -smoothed equilibrium if it is a fixed point $\mathbf{x}^\beta = \Phi^\beta(\mathbf{x}^\beta)$.

Definition 11 (ε -approximate Nash equilibrium). Let $\varepsilon \geq 0$. We say that a joint decision vector $\mathbf{x}^\varepsilon \in \Omega$ is an ε -approximate Nash equilibrium if $f_n(x_n, \mathbf{x}_{-n}^\varepsilon) \leq f_n(\mathbf{x}^\varepsilon) + \varepsilon$ for all $\mathbf{x} \in \Omega$ and $n \in [N]$.

Economic interpretation of game dynamics Smoothed best-response is a common model of *bounded rationality*, where the solution that a player finds is only an approximate best-reply to the strategies $\mathbf{x}_{-n}$ chosen by the others, perhaps due to noise, or constraints on information or computation. The player approaches ‘perfect rationality’ as the parameter β goes to zero.

Typically, smoothed best-response and its limiting best-response are rational from the perspective of players that are *myopic*, which means that they do not try to steer the long-term behavior of the dynamics. Myopia is often justified in the evolutionary game theory setting, where each ‘player’ corresponds to a large population of individuals. As each individual has an infinitesimal impact on the long-term behavior of the dynamics, the most individually-rational behavior is to optimize for immediate outcomes. Myopia is also justified if the individual’s lifespan is much shorter than the timescale over which the dynamics evolve.

Under this context, the incremental or averaging aspect of the dynamics (1) also arise naturally; Fudenberg and Levine (1998) also call these “partial” dynamics. It captures situations where only a random fraction of individual within the populations play the game at each round. As only a fraction of individuals update their strategies, the strategy profile only partially evolves toward the smoothed best-response direction. In short, the dynamics (1) can be considered as a model for strategic interactions across large, imperfectly-rational populations, such as at a traffic stop where individuals with bounded rationality are intermittently matched up with others to play a dangerous game of chicken.

3 Uniform stability of Nash equilibria

The local behavior of a smooth, non-degenerate dynamical system is determined by linearizing the dynamics. In learning in games literature, this motivates the *game Jacobian* (e.g. Letcher et al. (2019)). It arises naturally from the analysis of gradient ascent dynamics, where players update their strategies along the direction of steepest ascent $\nabla_n f_n(\mathbf{x})$, and the game Jacobian is the Jacobian of the gradient ascent dynamics:

Definition 12 (Jacobian of the game). Let $(\Omega, \mathbf{f})$ be an N -player multilinear game and let $\mathbf{x} \in \Omega$. The *game Jacobian* $\mathbf{J}(\mathbf{x})$ is the Jacobian of the map $\mathbf{x} \mapsto (\nabla_1 f_1(\mathbf{x}), \dots, \nabla_N f_N(\mathbf{x}))$, so that:

$$\forall n, m \in [N], \quad J_{nm}(\mathbf{x}) = \nabla_{nm}^2 f_n(\mathbf{x}).$$

In particular, the block diagonals $J_{nn}(\mathbf{x}) = 0$ are zero matrices.

We use the game Jacobian to define a notion of stability of a Nash equilibrium $\mathbf{x}^*$. It depends on the eigenvalues of the matrices $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x}^*)$, where $\mathbf{H}$ ranges over the family of positive-definite, block-diagonal matrices. If such an eigenvalue has positive real part, the dynamics around $\mathbf{x}^*$ become unstable for some regularizers (Proposition 2). Due to that fact that the block diagonals of the game Jacobian are zero matrices, sum of all eigenvalues is zero: $\text{tr}(\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})) = 0$. As a result, instability also arises if there is an eigenvalue with negative real part—some other eigenvalue must have positive real part. And so, the only possibility for stability is when all eigenvalues of matrices $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x}^*)$ are purely imaginary. We call this uniform stability. Formally, this definition uses the following, purely linear-algebraic matrix condition (Definition 13). Notice that uniform stability is a bilinear (that is, a second-order) condition depending only on the strategic component of the game.

Definition 13 (Uniformly-stable block matrix). Let $\mathbf{J} \in \mathbb{R}^{d \times d}$ be a block-matrix, where d is the sum of its block dimensions $(d_1, \dots, d_N)$. We say that $\mathbf{J}$ is *uniformly stable* if the eigenvalues of $\mathbf{H}^{-1}\mathbf{J}$ are purely imaginary for all positive-definite, block-diagonal $\mathbf{H} \in \mathbb{R}^{d \times d}$,

$$\text{spec}(\mathbf{H}^{-1}\mathbf{J}) \subseteq i\mathbb{R},$$

where $\mathbf{H} = \text{diag}(H_1, \dots, H_N)$ and $H_n \in \mathbb{R}^{d_n \times d_n}$ is positive definite.

Definition 14 (Uniformly stable equilibria). Let $(\Omega, \mathbf{f})$ be an N -player multilinear game. We say that a Nash equilibrium $\mathbf{x}^* \in \Omega$ is (pointwise) *uniformly stable* if the Jacobian $\mathbf{J}(\mathbf{x}^*)$ is uniformly stable. It is *locally uniformly stable* if it is contained in an open set $U \subset \Omega$ such that $\mathbf{J}(\mathbf{x})$ is uniformly stable for all $\mathbf{x} \in U$.

Section 4 develops the economic meaning of uniform stability. Section 5 shows the consequences of this spectral condition for dynamic stability. These two sections are independent of each other and can be read in either order. Before doing so, we provide some intuitive and formal motivation for uniform stability.

3.1 Motivation for uniform stability

The concept of uniform stability arises when we relax the assumption that players are ‘perfectly rational’. Unlike gradient-ascent players, smoothed best-response players do not necessarily update their strategies in the direction of steepest ascent (Lemma 2). With the existence of the regularizers, players are no longer maximally efficient—due to either bounded rationality or a need for better stability in their strategies, players only aim to improve modestly.

At a current joint strategy $\mathbf{x}$, let’s say that the n th player makes an update along the v_n direction. This leads to a local improvement whenever $\nabla_n f_n(\mathbf{x})^\top v_n > 0$. And this occurs if and only if there is a positive-definite matrix $H_n \succ 0$ so that the update direction v_n and the gradient $\nabla_n f_n(\mathbf{x})$ are related by:

$$v_n = H_n^{-1} \nabla_n f_n(\mathbf{x}).$$

This is a consequence of a simple linear-algebraic result, proved in Section D:

Lemma 1. *Let $u, v \in \mathbb{R}^m \setminus \{0\}$. Then, $u^\top v > 0$ if and only if there is a positive-definite matrix H so that:*

$$u = Hv.$$

In optimization, this result justifies methods such as preconditioned gradient ascent and mirror ascent. In learning in games, it also directly justifies smoothed best-response dynamics as well as *positive-definite adaptive dynamics* (Hopkins, 1999). In the current paper, this result leads to the notion of uniform stability by pointing us toward the preconditioned Jacobian maps $\mathbf{x} \mapsto (H_1^{-1} \nabla_1 f_1(\mathbf{x}), \dots, H_N^{-1} \nabla_N f_N(\mathbf{x}))$. As uncoupled players do not exactly know how others are learning, stability is defined uniformly over all positive-definite matrices. This is a fairly strong condition; at a high level, this is the price that comes with considering a broad class of learning dynamics, where players may have bounded rationality.

A formal connection between uniform stability and smoothed best-response dynamics is obtained by computing the Jacobian of the smoothed best-response map Φ^β . From Lemma 2 below, it becomes clear that the eigenvalues of $\nabla \Phi^\beta(\mathbf{x}^*)$ are related to the spectra of matrices $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x}^*)$. This result allows us to achieve in Section 5 the convergence guarantees assuming local uniform stability, as well as non-convergence without pointwise uniform stability. The proof of Lemma 2 is given in Section E.

Lemma 2 (Gradient of smoothed best-response). *Let $(\Omega, \mathbf{f})$ be a multilinear game, where $\Omega_n \subset \mathbb{R}^{d_n}$ for each $n \in [N]$, and let $\mathbf{h}$ be a set of smooth and strictly convex regularizers, $h_n : \Omega_n \rightarrow \mathbb{R}$. Then, the associated β -smoothed best-response map Φ^β is well-defined and smooth, where the gradient of Φ^β at $\mathbf{x}$ is given by:*

$$\nabla \Phi^\beta(\mathbf{x}) = \frac{1}{\beta} \mathbf{H}(\mathbf{x})^{-1} \mathbf{J}(\mathbf{x}), \quad \text{where } \mathbf{H}(\mathbf{x})_{nm} = \begin{cases} \nabla_n^2 h_n(x_n) & n = m \\ 0_{d_n \times d_m} & n \neq m, \end{cases}$$

and $\mathbf{H}(\mathbf{x}) \in \mathbb{R}^{d \times d}$ is a positive-definite, block-diagonal matrix with $d = d_1 + \dots + d_N$.

Conversely, Lemma L.2 shows that for any choice $\mathbf{x}$ and positive-definite, block diagonal matrix $\mathbf{H}^{-1}$, there exists a collection of steep regularizers such that $\nabla \Phi^\beta(\mathbf{x})$ is precisely given by $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})$, up to scaling.

4 The economic meaning of uniform stability

We will introduce assumptions in Section 4.1, define and motivate strategic stationarity in Section 4.2, before finally providing the economic characterizations of uniform stability in Section 4.3.

4.1 Structural assumptions on the game

The following assumptions will help us exclude trivial ways in which a strategy is strategically Pareto optimal. For example, if one player has constant utility, then every joint decision is strategically Pareto optimal, since it is impossible to strictly improve that player's utility. Specifically, we assume players have *bi-directional interactions*, ruling out one-sided strategic interactions between any two players. For example, when one player's utility directly depends on another, but not vice versa.

Definition 15 (Bi-directional interactions). Let $(\Omega, \mathbf{f})$ be an N -player game with smooth utilities and let $\mathbf{x} \in \text{int}(\Omega)$. We say that two players $n, m \in [N]$ have *bi-directional interactions* at $\mathbf{x}$ if for all $n, m \in [N]$:

$$\ker(J_{nm}(\mathbf{x})) = \ker(J_{mn}(\mathbf{x})^\top).$$

We will also assume that the interaction graph is *connected*, where this is a graph that captures when a player is strategically responsive to the action of another: an edge from player n to player m denotes that the block in the game Jacobian $J_{nm}(\mathbf{x})$ is non-zero.

Definition 16 (Interaction graph). Let $(\Omega, \mathbf{f})$ be an N -player game and let $\mathbf{x} \in \text{int}(\Omega)$. The *interaction graph* $G(\mathbf{x}) = (V, E)$ at $\mathbf{x}$ is the directed graph whose vertices are the players $V = [N]$, and the edges are:

$$(n, m) \in E \iff \nabla_{nm}^2 f_n(\mathbf{x}) \neq 0.$$

If the game is bi-directional at $\mathbf{x}$, then $G(\mathbf{x}) = (V, E)$ is undirected, since $(n, m) \in E$ if and only if $(m, n) \in E$.

Mathematically speaking, this assumption is mild and generic in multilinear games. It ensures that the lower-order, bilinear terms around an equilibrium dominate each player's utility.

4.2 Necessity and sufficiency for strategic Pareto optimality

We define a second-order (bilinear) relaxation of strategic Pareto optimality, which we call *strategic Pareto stationarity*, or *strategic stationarity*, for short. It will make use of the following matrix concept:

Definition 17 (λ -skew matrix). Let $\mathbf{J} \in \mathbb{R}^{d \times d}$ be a block-matrix, where d is the sum of its block dimensions $(d_1, \dots, d_N)$. Let $\lambda = (\lambda_1, \dots, \lambda_N)$ be a tuple of positive scalars. We say that $\mathbf{J}$ is λ -*skew-symmetric* if:

$$\forall n, m \in [N], \quad \lambda_n J_{nm} = -\lambda_m J_{mn}^\top.$$

In other words, the matrix $\mathbf{\Lambda J}$ is skew-symmetric, where $\mathbf{\Lambda}$ is block-diagonal with $\Lambda_{nn} = \lambda_n I_{d_n \times d_n}$.

Definition 18 (Strategic Pareto stationarity). Let $(\Omega, \mathbf{f})$ be an N -player multilinear game. A Nash equilibrium $\mathbf{x}^*$ is *strategically Pareto stationary* if there is a set of positive scalars λ so that $\mathbf{J}(\mathbf{x}^*)$ is λ -skew.

To understand the meaning of strategic Pareto stationarity, it is useful to first focus on *polymatrix games*, which are multilinear games whose utilities are polynomials of degree at most two—they have up to bilinear terms. These games are also called *graphical games* (Kearns, 2008; Cai et al., 2016). Formally:

Definition 19 (Polymatrix game). A multilinear game is *polymatrix* if its utilities are bilinear polynomials.

The payoffs in these games only depend on the pairwise interactions between players (higher-order k -way interactions are captured by the degree- k terms in the polynomial utilities). Furthermore, the Jacobian of a polymatrix game is a constant $\mathbf{J}(\mathbf{x}) \equiv \mathbf{J}$, so if a polymatrix game has a Nash equilibrium at $\mathbf{x}^*$, then a particularly nice and strategically-equivalent form of the utility f_n is given by:

$$f_n(\mathbf{x}) \simeq_{\text{SE}} \sum_{m \neq n} (x_n - x_n^*)^\top J_{nm} (x_m - x_m^*).$$

From this form of the utilities, we can see that when the Nash equilibrium of a polymatrix game is strategically Pareto stationary, every pair of player must be in *strict competition*. Classically, strict competition describes two-player games where any strategy that improves one player’s utility must worsen the other’s; that is, every strategy is Pareto optimal (Adler et al., 2009). Indeed, we have here that $J_{nm} = -J_{mn}^\top$, up to a change of units. Thus, strategic Pareto stationarity in the bilinear setting means that if two players n and m deviate away from the Nash equilibrium, any resulting change in the strategic components of their utilities must be opposite in sign. Generalizing this, we obtain a sufficient condition for polymatrix games:

Lemma 3 (Sufficient condition for strategic Pareto optimality). *Let $(\Omega, \mathbf{f})$ be an N -player polymatrix game with a Nash equilibrium $\mathbf{x}^*$. If $\mathbf{x}^*$ is strategically Pareto stationary, then it is strategically Pareto optimal.*

This is proved in Section G. Under the assumptions of connected interaction graph and bi-directional interactions at the equilibrium—defined in Section 4.1—the converse is also true, where strategic Pareto optimality implies strategic Pareto stationarity. In fact, this holds generally for multilinear games. With these assumptions, we show that in multilinear games, strategic Pareto stationarity is necessary for optimality. The proof of this result is given in Section H.

Proposition 1 (Strategic Pareto stationarity is necessary for strategic Pareto optimality). *Let $(\Omega, \mathbf{f})$ be an N -player multilinear game with a Nash equilibrium $\mathbf{x}^*$. Let the interaction graph $G(\mathbf{x}^*)$ be connected and the game be bi-directional at $\mathbf{x}^*$. If the equilibrium $\mathbf{x}^*$ is strategically Pareto optimal, then it is strategically Pareto stationary.*

4.3 The economic characterization of uniform stability

The formal results concerning the economic meaning of uniform stability follow.

Theorem 1 (Local uniform stability implies strategic Pareto optimality). *Let $\mathbf{x}^*$ be a Nash equilibrium of an N -player multilinear game, at which it is locally bi-directional. Suppose the interaction graph is connected at $\mathbf{x}^*$. If the game is locally uniformly stable at $\mathbf{x}^*$, then the equilibrium is strategically Pareto optimal.*

Theorem 2 (Pointwise uniform stability is equivalent to strategic Pareto stationarity). *Let $\mathbf{x}^*$ be a Nash equilibrium of an N -player multilinear game, at which it is bi-directional. Suppose that the interaction graph is connected at $\mathbf{x}^*$. Then, $\mathbf{x}^*$ is uniformly stable if and only if it is strategically Pareto stationary.*

Theorem 2 and Proposition 1 prove that strategic Pareto optimality implies pointwise uniform stability. In a way, it is not surprising that a bilinear condition like strategic stationarity ends up being the adequate solution concept to characterizing uniformly-stable Nash equilibria. Observe that at an equilibrium $\mathbf{x}^*$, the first-order or linear strategic terms $\nabla_n f_n(\mathbf{x}^*)$ vanish across every player’s utility, and that uniform stability at $\mathbf{x}^*$ can only reveal information only about the bilinear terms $\nabla_{nm}^2 f_n(\mathbf{x}^*)$ within the strategic component of the utilities; the condition depends purely on the Jacobian of the game $\mathbf{J}(\mathbf{x}^*)$ at that point (see Definition 14).

5 Convergence under uniform stability

The fundamental result by Hart and Mas-Colell (2003) shows that, in general, uncoupled dynamics cannot be guaranteed to asymptotically converge to mixed Nash equilibria. In this section, we refine this result for smoothed best-response learning dynamics (1), providing conditions for when the dynamics are unstable and fail to converge, and when they are stable and, in fact, globally converge. We consider two stability classes for mixed equilibria: those that are pointwise uniformly stable, and those that are locally uniformly stable.

For simplicity, we consider a Nash equilibrium $\mathbf{x}^*$ that is approximated by the sequence of β -smoothed equilibria $\mathbf{x}^\beta$ as β goes to zero (see Lemma B.2). The first result shows that when $\mathbf{x}^*$ is *not* uniformly stable, then the (Φ^β, η) -averaging dynamics can become unstable around $\mathbf{x}^\beta$ once β becomes sufficiently small. Intuitively, this means that $\mathbf{x}^*$ is ‘inapproximable’ by generic smoothed best-response dynamics. In contrast, our second result shows that if $\mathbf{x}^*$ is contained in a region that is uniformly stable, then it can be approximated to arbitrary precision by generic, smoothed best-response dynamics. It turns out that the price of higher precision, and in turn smaller β , is a slower rate of convergence due to the requirement of a smaller learning rate η . Informally:

Non-convergence result. If a Nash equilibrium $\mathbf{x}^*$ is not uniformly stable, then there are smoothed best-responses that cannot be stabilized: there are regularizers $\mathbf{h}$ so that when $\beta > 0$ is sufficiently small, all (Φ^β, η) -averaging dynamics become unstable about the smoothed equilibrium $\mathbf{x}^\beta$ (Proposition 2).

Convergence result. If a Nash equilibrium $\mathbf{x}^*$ is contained in a region that is uniformly stable, then every smoothed best-responses can be stabilized: for all Φ^β and sufficiently small $\eta > 0$, the (Φ^β, η) -averaging dynamics globally converges to $\mathbf{x}^\beta$ (Theorem 3).

5.1 Stability concepts in dynamical systems

Before we state the main results, recall the standard notion of stability from dynamical systems:

Definition 20 (Stability of fixed points). Let $\Omega \subset \mathbb{R}^d$ be a domain and let $\Phi : \Omega \rightarrow \Omega$ be a transition map defining the discrete-time dynamical system $\mathbf{x}(t+1) = \Phi(\mathbf{x}(t))$. Let $\mathbf{x}^* \in \Omega$ be a fixed point of Φ .

- The fixed point $\mathbf{x}^*$ is *stable* if for every $\varepsilon > 0$, there exists some $\delta > 0$ such that:

$$\|\mathbf{x}(0) - \mathbf{x}^*\| < \delta \implies \forall t \geq 0, \quad \|\mathbf{x}(t) - \mathbf{x}^*\| < \varepsilon.$$

- Moreover, it is *asymptotically stable* if there exists some $\delta > 0$ such that:

$$\|\mathbf{x}(0) - \mathbf{x}^*\| < \delta \implies \lim_{t \rightarrow \infty} \|\mathbf{x}(t) - \mathbf{x}^*\| = 0.$$

We say that $\mathbf{x}^*$ is *unstable* if it is not stable.

We will consider families of β -smoothed best-response dynamics, whose fixed points $\mathbf{x}^\beta$ converge to a Nash equilibrium $\mathbf{x}^*$ as β goes to zero. This motivates the following notion of *stabilization* to Nash equilibria, where each $\mathbf{x}^\beta$ is an asymptotically-stable fixed point of sufficiently fine (Φ^β, η) -averaging dynamics.

Definition 21 (Stabilizing smoothed best-response dynamics). Let $(\Omega, \mathbf{f})$ be an N -player multilinear game with a Nash equilibrium $\mathbf{x}^*$. Let $\mathbf{h}$ be a set of smooth and strictly convex regularizers. We say that smoothed best-response dynamics with respect to $\mathbf{h}$ *stabilizes* to $\mathbf{x}^*$ if the following hold:

- The β -smoothed equilibria $\mathbf{x}^\beta$ converge to $\mathbf{x}^*$ as β goes to zero.
- For each $\beta > 0$, the β -smoothed equilibrium $\mathbf{x}^\beta$ is an asymptotically-stable fixed point of the (Φ^β, η) -averaging dynamics (1) whenever $\eta \in (0, 1)$ is sufficiently small.

5.2 Convergence results for smoothed best-response dynamics

The first result shows that if a Nash equilibrium is not uniformly stable, then not only is it not possible to stabilize generic smooth best-response dynamics to it (the dynamics to $\mathbf{x}^\beta$ are not asymptotically stable for arbitrarily small β), but in fact, the dynamics can become unstable. The following is proved in Section J.

Proposition 2 (Inapproximability of unstable equilibria). *Let $(\Omega, \mathbf{f})$ be an N -player multilinear game with an interior Nash equilibrium $\mathbf{x}^*$. Suppose that $\mathbf{x}^*$ is not uniformly stable. Then, there is some collection of smooth and strictly convex regularizers $\mathbf{h}$ such that smoothed best-response dynamics does not stabilize to $\mathbf{x}^*$. In fact, even if the sequence of β -smoothed equilibria $\mathbf{x}^\beta$ converges to $\mathbf{x}^*$, each $\mathbf{x}^\beta$ is an unstable fixed point of the (Φ^β, η) -averaging dynamics (1) for all $\eta \in (0, 1)$ and sufficiently small $\beta > 0$.*

Now, we present the main convergence result concerning Nash equilibria with local uniform stability. It turns out that due to additional structure in multilinear games, convergence is global: no matter where the dynamics is initialized, it will converge to the smoothed equilibrium. The following is proved in Section K.

Theorem 3 (Convergence to equilibria in uniformly-stable regions). *Let $(\Omega, \mathbf{f})$ be an N -player multilinear game with a Nash equilibrium $\mathbf{x}^*$ that is locally uniformly stable. Over arbitrary choice of smooth and strictly convex regularizer $\mathbf{h}$, all smoothed best-response dynamics stabilize to $\mathbf{x}^*$. In particular, there is a constant $C_f > 0$ depending only on the game so that when $\eta \leq C_f \beta^2$, the (Φ^β, η) -averaging dynamics globally converge:*

$$\|\mathbf{x}(t) - \mathbf{x}^\beta\| \leq \exp\left(-\frac{\eta t + \ln N}{2}\right).$$

Using the fact that β -smoothed equilibria are $O(\beta)$ -approximate Nash equilibria, this implies that smoothed best-response dynamics with averaging achieve $T^{-1/2}$ -convergence rate to Nash equilibria.

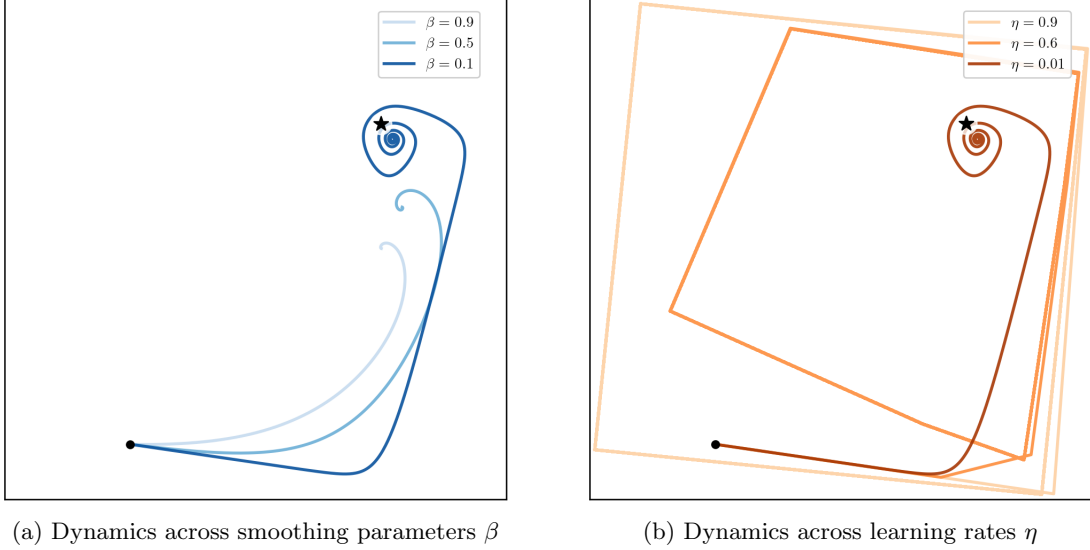


FIGURE 2: The trajectories of β -smoothed best-response dynamics with η -learning rate toward a uniformly-stable, mixed Nash equilibrium (star) in a two-player normal-form game. (a) The β -smoothed equilibria become better approximations of the Nash equilibrium as β shrinks, but the dynamics become less stable and exhibit more cycling. The figure on the left plots the trajectories initialized at the black dot for varying smoothing β and fixed averaging η . (b) The learning dynamics converge to β -smoothed equilibria once η becomes sufficiently small, and it does so with greater stability with smaller η . However, stability comes at the expense of slower rate of convergence. The figure on the right plots the trajectories for fixed β and varying η .

6 Uniform stability for partially-mixed equilibria

The main convergence result (Theorem 3) can be extended to equilibria that are not fully mixed, and are not interior equilibria. A Nash equilibrium $\mathbf{x}^*$ is on the boundary of the probability simplex if there are players who never play certain actions. This can happen if these actions are strictly dominated, in which case a perfectly rational player will never play those strategies, while a player with bounded rationality will play them with decreasing frequency. Intuitively, as long as the players are not overly irrational, the stability of smoothed best-response dynamics around these equilibria should not hinge upon these suboptimal actions.

Before we make this intuition formal, recall some notations for normal-form game:

- The joint strategy space $\Omega = \Omega_1 \times \cdots \times \Omega_N$ is a product of probability simplices, $\Omega_n = \Delta^{k_n-1}$, where the n th player may randomize over k_n pure strategies.
- The vertices of the simplex Ω_n correspond to pure strategies of player n . For each alternative $i \in [k_n]$, we let $e_{n,i} \in \Omega_n$ denote the pure strategy where the n th player deterministically chooses the i th action.

The *support* of a strategy is the set of pure strategies with positive probability of being played:

Definition 22 (Support of a strategy). For any player n , the *support* of a strategy $x_n \in \Omega_n$ is the set:

$$\text{supp}(x_n) = \{i \in [k_n] : x_{n,i} \neq 0\}.$$

Given $\mathbf{x} \in \Omega$, define the reduced strategy sets:

$$\Omega_n(x_n) := \{x'_n \in \Omega_n : \text{supp}(x'_n) \subset \text{supp}(x_n)\} \quad \text{and} \quad \Omega(\mathbf{x}) := \prod_{i \in [N]} \Omega_i(x_i). \quad (2)$$

Note that $\Omega_n(x_n)$ is an $(k'_n - 1)$ -dimensional simplex where $k'_n = |\text{supp}(x_n)|$.

For the convergence result, we consider *quasi-strict equilibria*, which are Nash equilibria where each player fully mixes on the set of best responses (Harsanyi and Selten, 1988; Van Damme, 1992).² This captures the property that if an alternative is never played in an equilibrium, then it must be strictly dominated.

Definition 23 (Quasi-strict equilibrium). We say that a Nash equilibrium $\mathbf{x}^*$ of an N -player normal form game is *fully-mixed on the best-response set* or *quasi-strict* if for each player $n \in [N]$:

$$i \in \text{supp}(x_n^*) \iff i \in \arg \max_{i \in [k_n]} f_n(e_{n,i}; \mathbf{x}_{-n}^*)$$

Since actions that are not supported by a quasi-strong equilibrium are dominated, we expect that they are strategically irrelevant. This motivates the following refinement of the game (Harsanyi and Selten, 1988), where the irrelevant actions are removed from the game.

Definition 24 (Reduced game). Let $(\Omega, \mathbf{f})$ be a normal-form game with a Nash equilibrium $\mathbf{x}^*$. The *reduced game* around $\mathbf{x}^*$ is the game on the domain $\Omega(\mathbf{x}^*)$ with the utilities $\mathbf{f}$ restricted to $\Omega(\mathbf{x}^*)$.

In the reduced game, $\mathbf{x}^*$ becomes an interior or fully-mixed Nash equilibrium (Corollary L.1), allowing the results in Section 5.2 to be applied. The remaining problem is about how the players learn to converge towards the support of the equilibrium and avoid playing the suboptimal strategies that have been removed from the reduced game. In order to control the rate at which these suboptimal strategies appear in the β -smoothed equilibria $\mathbf{x}^\beta$, we introduce the following *quantitative* version of steepness (c.f. Definition 9):

Definition 25 (Linear steepness). Let $h : \Delta^{k-1} \rightarrow \mathbb{R}$ be a steep regularizer. For all $v \in \mathbb{R}^k$, define the map:

$$x^\beta(v) = \arg \max_{x \in \Delta^{k-1}} v^\top x - \beta h(x).$$

For each $i \in [i]$ and $\varepsilon > 0$, let $V_i(\varepsilon) = \{v \in \mathbb{R}^k : v_i < \max_{j \in [k]} v_j - \varepsilon\}$ be the set of vectors where the i th component is ε -suboptimal. We say that h is *linearly steep* if the following holds for all $i \in [k]$ and $\varepsilon > 0$:

$$\lim_{\beta \rightarrow 0} \sup_{v \in V_i(\varepsilon)} \frac{x^\beta(v)_i}{\beta} = 0.$$

In words, this states that the probability mass $x^\beta(v)_i$ placed on any alternative $i \in [k]$ must shrink sub-linearly with β when the value v_i is ε -suboptimal. In the case that h is the entropy map, Example L.1 shows that, in fact, this distance shrinks as $\exp(-\Omega(1/\beta))$ at a much faster, exponential rate. As a consequence of linear steepness, the following shows that the rate at which suboptimal alternatives are played in β -smoothed equilibria $\mathbf{x}^\beta$ also vanish at a sublinear rate $o(\beta)$:

Lemma 4 (Dominated strategies occur at β -sublinear rates in β -smoothed equilibria). *Let $(\Omega, \mathbf{f})$ be an N -player normal-form game and $\mathbf{h}$ be a collection of linearly steep regularizers. Suppose that as β goes to zero, the sequence of β -smoothed equilibria $\mathbf{x}^\beta$ converges to a quasi-strict equilibrium $\mathbf{x}^*$. Let $\Pi : \Omega \rightarrow \Omega(\mathbf{x}^*)$ be the orthogonal projection onto the support of $\mathbf{x}^*$. Then:*

$$\lim_{\beta \rightarrow 0} \frac{\|\mathbf{x}^\beta - \Pi \mathbf{x}^\beta\|}{\beta} = 0.$$

Before we can give the main convergence result for all quasi-strict equilibria, we need a technical condition to ensure that pseudoinverse $\nabla^2 h(x)^+$ extends smoothly to the boundary of the simplex:

Definition 26 (Proper regularizer). A map $h : \Delta^{k-1} \rightarrow \mathbb{R}$ on the simplex is a *proper regularizer* if:

- The restriction of h onto each face of the simplex is steep regularizer.
- The Moore-Penrose pseudoinverse of the Hessian $\nabla^2 h(x)^+$ is smooth on the simplex.

²Classically, these are called *quasi-strong* equilibria. We prefer the term *quasi-strict* since *strong Nash equilibria* is another, unrelated refinement of the Nash equilibrium concept, which we also make use of. For discussion, see Van Damme (1992).

See Section L.1 for formal definition of the *faces* of the simplex and what it means to be *smooth* on it. The following convergence result shows that if $\mathbf{x}^*$ is a quasi-strict equilibrium, it suffices to analyze its uniform stability within the reduced game, neglecting all suboptimal alternatives. This is because whenever the regularizers are linearly steep, any potentially de-stabilizing effect of these additional strategies are overwhelmed by the stabilizing effects of the averaging dynamics.

Theorem 4 (Convergence to non-interior equilibria). *Let $(\Omega, \mathbf{f})$ be an N -player, normal-form game and let $\mathbf{h}$ be a collection of linearly steep, proper regularizers. Suppose that $\mathbf{x}^*$ is a quasi-strict Nash equilibrium and that it is the limit of the β -smoothed equilibria $\mathbf{x}^\beta$ as β goes to zero. If the reduced game around $\mathbf{x}^*$ is locally uniformly stable, then smoothed best-response dynamics stabilize to $\mathbf{x}^*$. In particular, there exists a constant $C_{\mathbf{f}} > 0$ such that for all sufficiently small $\beta > 0$ and $\eta \leq C_{\mathbf{f}}\beta^2$, the (Φ^β, η) -averaging dynamics is locally asymptotically stable at $\mathbf{x}^\beta$, where the operator norm of its Jacobian at $\mathbf{x}^\beta$ is bounded by:*

$$\|(1 - \eta)\mathbf{I} + \eta \nabla \Phi^\beta(\mathbf{x}^\beta)\|_2 \leq \exp\left(-\frac{\eta}{2}\right).$$

The results of this section are proved in Section L.

7 Discussion

In this work, we have introduced a theory of non-asymptotic stability for learning in games. We showed a fairly tight connection between uniform stability and collective rationality. And through uniform stability, we studied the last-iterate convergence behavior for the class of incremental, smoothed best-response dynamics. We close with a few open questions that we believe would be important to pursue.

Strengthening the necessity of uniform stability for convergence to mixed Nash equilibria Our main non-convergence result (Proposition 2) shows that if an equilibrium $\mathbf{x}^*$ is not uniformly stable, then there exist smoothed best-responses that cannot be stabilized. We believe this can be strengthened, where ‘there exist’ is replaced with ‘almost all’ (in terms of the Hessian of the regularizer at $\mathbf{x}^*$).

A second strengthening has to do with the economic properties of the non-converging behaviors. We have shown that players can avoid stabilizing to Pareto inefficient equilibria, which can be seen as a form of collective rationality. However, we have not analyzed if players escape inefficient equilibria in a collectively-rational way:

Open Question 1. *Are there uncoupled learning dynamics that not only do not stabilize to mixed equilibria that are not uniformly stable, but moreover locally escape them in collectively-rational ways. That is, whenever $\|\mathbf{x}(t) - \mathbf{x}^*\| < \varepsilon$ is sufficiently close, there exists some $\delta > 0$ such that for some time, all players improve:*

$$\forall s \in (t, t + \delta) \quad \text{and} \quad n \in [N], \quad f_n(\mathbf{x}(s)) > f_n(\mathbf{x}(t)).$$

Non-asymptotic analyses beyond smoothed best-response While incremental, smoothed best-response dynamics is a natural family of learning dynamics. It would be of interest to consider other classes of updates that can lead to mixed Nash equilibria.

Relaxation of bi-directionality assumption It would be interesting to generalize beyond the assumption of bi-directional interactions. In particular, the bi-directional interactions are represented by undirected graphs. One can also consider interaction structures that are represented by directed graphs, which may lead to longer-range interactions (e.g., players form a directed cycle as in the Cyclic Matching Pennies game of Kleinberg et al. (2011)).

Acknowledgments

We express our gratitude toward Mike Jordan, for his inspiration and suggestions.

References

- Ilan Adler, Constantinos Daskalakis, and Christos H Papadimitriou. A note on strictly competitive games. In *International Workshop on Internet and Network Economics*, pages 471–474. Springer, 2009.
- Armen A Alchian. Uncertainty, evolution, and economic theory. *Journal of political economy*, 58(3):211–221, 1950.
- Carlos Alós-Ferrer and Nick Netzer. The logit-response dynamics. *Games and Economic Behavior*, 68(2): 413–427, 2010.
- Ioannis Anagnostides, Ioannis Panageas, Gabriele Farina, and Tuomas Sandholm. On last-iterate convergence beyond zero-sum games. In *International Conference on Machine Learning*, pages 536–581. PMLR, 2022.
- Robert J Aumann. Acceptable points in general cooperative n -person games. *Contributions to the Theory of Games*, 4:287, 1959.
- James P Bailey, Sai Ganesh Nagarajan, and Georgios Piliouras. Stochastic multiplicative weights updates in zero-sum games. *arXiv preprint arXiv:2110.02134*, 2021.
- David Balduzzi, Sebastien Racaniere, James Martens, Jakob Foerster, Karl Tuyls, and Thore Graepel. The mechanics of n -player differentiable games. In *International Conference on Machine Learning*, pages 354–363. PMLR, 2018.
- Michel Benaïm and Morris W Hirsch. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behavior*, 29(1-2):36–72, 1999.
- Nicoletta Bof, Ruggero Carli, and Luca Schenato. Lyapunov theory for discrete time systems. *arXiv preprint arXiv:1809.05289*, 2018.
- Yang Cai, Ozan Candogan, Constantinos Daskalakis, and Christos Papadimitriou. Zero-sum polymatrix games: A generalization of minmax. *Mathematics of Operations Research*, 41(2):648–655, 2016.
- Ozan Candogan, Ishai Menache, Asuman Ozdaglar, and Pablo A Parrilo. Flows and decompositions of games: Harmonic and potential games. *Mathematics of Operations Research*, 36(3):474–503, 2011.
- Vincent P Crawford. Learning behavior and mixed-strategy Nash equilibria. *Journal of Economic Behavior & Organization*, 6(1):69–78, 1985.
- Constantinos Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. In *10th Innovations in Theoretical Computer Science (ITCS) conference, ITCS 2019*, 2019.
- Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a Nash equilibrium. *Communications of the ACM*, 52(2):89–97, 2009.
- Constantinos Daskalakis, Rafael Frongillo, Christos H Papadimitriou, George Pierrakos, and Gregory Valiant. On learning algorithms for Nash equilibria. In *International Symposium on Algorithmic Game Theory*, pages 114–125. Springer, 2010.
- Étienne de Montbrun and Jérôme Renault. Optimistic gradient descent ascent in zero-sum and general-sum bilinear games. *arXiv preprint arXiv:2208.03085*, 2022.
- Gerard Debreu. *Theory of value: An axiomatic analysis of economic equilibrium*. Wiley, New York, 1959.
- Glenn Ellison and Drew Fudenberg. Learning purified mixed equilibria. *Journal of Economic Theory*, 90(1): 84–115, 2000.
- Ludwig Elsner. An optimal bound for the spectral variation of two matrices. *Linear algebra and its applications*, 71:77–80, 1985.

- Milton Friedman. The methodology of positive economics. *The Essays in Positive Economics*, 1953.
- Drew Fudenberg and David M Kreps. Learning mixed equilibria. *Games and economic behavior*, 5(3): 320–367, 1993.
- Drew Fudenberg and David K Levine. *The theory of learning in games*, volume 2. MIT press, 1998.
- Angeliki Giannou, Emmanouil Vasileios Vlatakis-Gkaragkounis, and Panayotis Mertikopoulos. Survival of the strictest: Stable and unstable equilibria under regularized learning with partial information. In *Conference on Learning Theory*, pages 2147–2148. PMLR, 2021.
- Andrew Gilpin, Javier Pena, and Tuomas Sandholm. First-order algorithm with convergence for-equilibrium in two-person zero-sum games. *Mathematical programming*, 133(1):279–298, 2012.
- Steven Gjerstad. The rate of convergence of continuous fictitious play. *Economic Theory*, 7:161–177, 1996.
- Jacob K Goeree, Charles A Holt, and Thomas R Palfrey. Quantal response equilibrium: A stochastic theory of games. In *Quantal response equilibrium*. Princeton University Press, 2016.
- Garrett Hardin. The tragedy of the commons. *Science*, 162(3859):1243–1248, 1968.
- John C Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International journal of game theory*, 2(1):1–23, 1973.
- John C Harsanyi and Reinhard Selten. A general theory of equilibrium selection in games. *MIT Press Books*, 1, 1988.
- Sergiu Hart and Andreu Mas-Colell. Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review*, 93(5):1830–1836, 2003.
- Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.
- Josef Hofbauer and Sylvain Sorin. Best response dynamics for continuous zero-sum games. *Discrete and Continuous Dynamical Systems Series B*, 6(1):215, 2006.
- Ed Hopkins. A note on best response dynamics. *Games and Economic Behavior*, 29(1-2):138–150, 1999.
- Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.
- Rufus Isaacs. *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*. Courier Corporation, 1999.
- James S Jordan. Three problems in learning mixed-strategy Nash equilibria. *Games and Economic Behavior*, 5(3):368–386, 1993.
- Michael Kearns. Graphical games. In *The New Palgrave Dictionary of Economics*, pages 1–4. Springer, 2008.
- Robert D Kleinberg, Katrina Ligett, Georgios Piliouras, and Éva Tardos. Beyond the Nash equilibrium barrier. In *ICS*, volume 20, pages 125–140, 2011.
- Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *Annual symposium on theoretical aspects of computer science*, pages 404–413. Springer, 1999.
- Vijay Krishna and Tomas Sjöström. On the convergence of fictitious play. *Mathematics of Operations Research*, 23(2):479–511, 1998.
- Peter D Lax. *Linear algebra and its applications*. John Wiley & Sons, 2007.
- Alistair Letcher, David Balduzzi, Sébastien Racaniere, James Martens, Jakob Foerster, Karl Tuyls, and Thore Graepel. Differentiable game mechanics. *Journal of Machine Learning Research*, 20(84):1–40, 2019.

- R Duncan Luce and Howard Raiffa. *Games and decisions: Introduction and critical survey*. John Wiley & Sons, Inc., 1957.
- Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.
- Panayotis Mertikopoulos and William H Sandholm. Learning in games via reinforcement and regularization. *Mathematics of Operations Research*, 41(4):1297–1324, 2016.
- Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the twenty-ninth annual ACM-SIAM symposium on discrete algorithms*, pages 2703–2717. SIAM, 2018.
- Jason Milionis, Christos Papadimitriou, Georgios Piliouras, and Kelly Spendlove. An impossibility theorem in game dynamics. *Proceedings of the National Academy of Sciences*, 120(41):e2305349120, 2023.
- Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.
- Hervé Moulin and J-P Vial. Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory*, 7(3-4):201–221, 1978.
- John H Nachbar. “Evolutionary” selection dynamics in games: Convergence and limit properties. *International Journal of Game Theory*, 19(1):59–89, 1990.
- John Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.
- Georg Ostrovski and Sebastian van Strien. Payoff performance of fictitious play. *arXiv preprint arXiv:1308.4049*, 2013.
- Marco Pangallo, James BT Sanders, Tobias Galla, and J Doyne Farmer. Towards a taxonomy of learning dynamics in 2×2 games. *Games and Economic Behavior*, 132:1–21, 2022.
- Christos H Papadimitriou. The complexity of finding Nash equilibria. *Algorithmic game theory*, 2:30, 2007.
- Georgios Piliouras, Lillian Ratliff, Ryann Sim, and Stratis Skoulakis. Fast convergence of optimistic gradient ascent in network zero-sum extensive form games. In *International Symposium on Algorithmic Game Theory*, pages 383–399. Springer, 2022.
- Julia Robinson. An iterative method of solving a game. *Annals of mathematics*, 54(2):296–301, 1951.
- R Tyrrell Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1997.
- Larry Samuelson and Jianbo Zhang. Evolutionary stability in asymmetric games. *Journal of economic theory*, 57(2):363–391, 1992.
- Lloyd S Shapley. Some topics in two-person games. *Advances in game theory*, 1963.
- Satinder Singh, Michael J Kearns, and Yishay Mansour. Nash convergence of gradient dynamics in general-sum games. In *UAI*, pages 541–548, 2000.
- Adam Smith. *An inquiry into the nature and causes of the wealth of nations*. London: printed for W. Strahan; and T. Cadell, 1776.
- Dale O Stahl II. On the instability of mixed-strategy Nash equilibria. *Journal of Economic Behavior & Organization*, 9(1):59–69, 1988.
- Eric Van Damme. Refinements of Nash equilibrium. In *Advances in Economic Theory: Sixth World Congress*, volume 1, pages 32–75. Cambridge University Press New York, 1992.
- Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, Thanasis Lianas, Panayotis Mertikopoulos, and Georgios Piliouras. No-regret learning and mixed Nash equilibria: They do not mix. *Advances in Neural Information Processing Systems*, 33:1380–1391, 2020.

- Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Linear last-iterate convergence in constrained saddle-point optimization. *The Ninth International Conference on Learning Representations*, 2021.
- Jörgen W Weibull. The ‘as if’ approach to game theory: Three positive results and four obstacles. *European Economic Review*, 38(3-4):868–881, 1994.
- Sidney G Winter. Satisficing, selection, and the innovating remnant. *The Quarterly Journal of Economics*, 85(2):237–261, 1971.

A Normal-form games and multilinear games

A.1 Background for normal-form games

In a *normal-form game*, each player $n \in [N]$ chooses an action out of k_n alternatives or pure strategies, indexed over $[k_n]$. To allow for randomized strategies, let Δ^{k_n-1} denote the *probability simplex* over $[k_n]$,

$$\Delta^{k_n-1} = \left\{ x_n \in \mathbb{R}^{k_n} : \sum_{i \in [k_n]} x_{n,i} = 1 \text{ and } x_{n,i} \geq 0 \right\}.$$

The n th player's strategy space is $\Omega_n = \Delta^{k_n-1}$, where we say that $x_n \in \Omega_n$ is a *mixed strategy profile*, with:

$$\forall i \in [k_n], \quad x_{n,i} = \text{probability that player } n \text{ plays action } i.$$

Assuming players randomize their actions independently, each joint mixed strategy $\mathbf{x} = (x_1, \dots, x_N) \in \Omega$ defines the product distribution $x_1 \otimes \dots \otimes x_N$ over the joint space of alternatives $[k_1] \times \dots \times [k_N]$.

The utilities in a normal-form game are defined by a collection of *payoff tensors* $T^1, \dots, T^N \in \mathbb{R}^{k_1 \times \dots \times k_N}$,

$$T_{i_1, \dots, i_N}^n = \text{payoff for player } n \text{ when the } m\text{th player chooses action } i_m \text{ for } m \in [N].$$

The utilities $\mathbf{f}$ of a normal-form game extend these payoffs to mixed strategies in all of Ω , and they compute the *expected payoffs* when each player randomize independently:

$$f_n(\mathbf{x}) = \sum_{i_1 \in [k_1]} \dots \sum_{i_N \in [k_N]} T_{i_1, \dots, i_N}^n x_{1,i_1} \dots x_{N,i_N}.$$

Lemma A.1 (Normal-form games are multilinear games). *Let $(\Omega, \mathbf{f})$ be an N -player normal-form game, where the n th player has k_n alternatives. There is a multilinear game $(\tilde{\Omega}, \tilde{\mathbf{f}})$ with $\tilde{\mathbf{f}} : \mathbb{R}^{k_1-1} \times \dots \times \mathbb{R}^{k_N-1} \rightarrow \mathbb{R}^N$ and an affine bijection $\phi : \text{int}(\Omega) \rightarrow \tilde{\Omega}$ such that $\mathbf{f}(\mathbf{x}) = (\tilde{\mathbf{f}} \circ \phi)(\mathbf{x})$ for all $\mathbf{x} \in \text{int}(\Omega)$.*

Proof of Lemma A.1. This follows from the form of the utility of a normal form game. $\square$

A.2 Individual and collective rationality are generally incomparable

Example A.1 (Strategic Pareto optimality does not imply Nash stability). *Consider the following 2-player normal-form game, which is purely strategic, since all rows and columns sum to zero:*

f_A	B_1	B_2	B_3	f_B	B_1	B_2	B_3
A_1	4	-1	-3	A_1	4	-10	6
A_2	-10	2	8	A_2	-1	2	-1
A_3	6	-1	-5	A_3	-3	8	-5

Here, we show the payoff matrices for the A and B players, and bold the preferred action conditioned on the other player. In particular, the joint strategy (A_2, B_2) is a Nash equilibrium, but not a strategic Pareto optimum, since it can be improved to (A_1, B_1) . And while (A_1, B_1) is Pareto optimal, it is not a Nash equilibrium.

B Smoothed equilibria and dynamics

Lemma B.1 (β -smoothed equilibria are $O(\beta)$ -approximate Nash equilibria). *Let $\mathbf{h}$ be a set of strictly convex and bounded regularizers, where $\max h_n - \min h_n \leq C$ for $n \in [N]$, inducing the β -smoothed best-response map Φ^β . If a joint strategy is a β -smoothed equilibrium of Φ^β , then it is a $C\beta$ -approximate Nash equilibrium.*

Proof of Lemma B.1. Let $\mathbf{x}$ be a β -smoothed equilibrium with respect to $\mathbf{h}$. By the definition of Φ^β , we have that the following inequality holds for any alternate strategy $z_n \in \Omega_n$:

$$f_n(z_n; \mathbf{x}_{-n}) - \beta h_n(z_n) \leq f_n(\mathbf{x}) - \beta h_n(x_n).$$

By rearranging and applying the boundedness of the regularizer, we obtain:

$$\begin{aligned} f_n(z_n; \mathbf{x}_{-n}) &\leq f_n(\mathbf{x}) + \beta(h_n(z_n) - h_n(x_n)) \\ &\leq f_n(\mathbf{x}) + C\beta. \end{aligned}$$

Thus, $\mathbf{x}$ is a $C\beta$ -approximate Nash equilibrium. $\square$

Lemma B.2 (β -smoothed equilibria converge to Nash equilibria). *Let Ω be a convex, joint strategy space with a metric ρ . Let $(\Omega, \mathbf{f})$ be an N -player game with continuous utilities, and $\mathbf{h}$ be a set of strictly convex and bounded regularizers. Suppose that the game has a unique Nash equilibrium $\mathbf{x}^*$. Then:*

$$\lim_{\beta \rightarrow 0} \rho(\mathbf{x}^\beta, \mathbf{x}^*) = 0.$$

Proof of Lemma B.2. This follows from Lemma B.1 and Lemma B.3. $\square$

Lemma B.3 (Convergence to Nash equilibria). *Let Ω be a joint strategy space with a metric ρ , and suppose that $(\Omega, \mathbf{f})$ is an N -player game with continuous utilities. Let $(\mathbf{x}^{\varepsilon_k})_{k \in \mathbb{N}}$ be any sequence of ε_k -Nash equilibria where $\varepsilon_k \rightarrow 0$. If $\mathbf{x}^*$ is a limit point of the sequence, then $\mathbf{x}^*$ is a Nash equilibrium.*

Proof. The result follows since:

1. Let $\mathbf{X}^\varepsilon$ denote the set of ε -Nash equilibria. It is a closed set for all $\varepsilon > 0$.
2. The ε -Nash equilibria satisfies $\mathbf{X}^\varepsilon \subset \mathbf{X}^{\varepsilon'}$ whenever $\varepsilon \leq \varepsilon'$. Also, the set of Nash equilibria is given by:

$$\mathbf{X}^* = \bigcap_{\varepsilon > 0} \mathbf{X}^\varepsilon.$$

3. For each ε , the sequence $\mathbf{x}^{\varepsilon_k}$ eventually remains in $\mathbf{X}^\varepsilon$, and so $\mathbf{x}^* \in \mathbf{X}^\varepsilon$ for all ε .

This implies that $\mathbf{x}^*$ is Nash. $\square$

C Strategic equivalence and games in canonical form

For many results of this work, we may often restrict our consideration to games in *canonical form*:

Definition C.1 (Canonical form of a multilinear game). Let $(\Omega, \mathbf{f})$ be an N -player multilinear game, where the joint strategy space includes the origin $\mathbf{0} \in \Omega$. We say that the game is in *canonical form* if the utilities have no non-strategic component and if the origin is a Nash equilibrium, with $\mathbf{x}^* = \mathbf{0}$.

We can do this without loss of generality because the solution concepts we are concerned with all respect strategic equivalence, namely strategic Pareto optimality (Definition 7), uniform stability (Definition 14), and strategic Pareto stationarity (Definition 18). Therefore, it is enough to prove them for games that are purely strategic (i.e. they have no non-strategic component; see Definition 6). Furthermore, in multilinear games, we can assume that the Nash equilibrium is set at the origin $\mathbf{x}^* = \mathbf{0}$. We may do this without loss of generality since the translation of a multilinear polynomial remains a multilinear polynomial.

Lemma C.1. *The dynamics (1) are uncoupled and are preserved across strategic equivalence classes.*

Proof. The dynamics depend only on the strategic components of the game. $\square$

D Proof of Lemma 1

Lemma 1. *Let $u, v \in \mathbb{R}^m \setminus \{0\}$. Then, $u^\top v > 0$ if and only if there is a positive-definite matrix H so that:*

$$u = Hv.$$

Proof. ($\Rightarrow$). Suppose that $u^\top v > 0$. There are two cases:

1. If u and v are linearly dependent, then we can let $H = \lambda I$ where $\lambda = \|u\|/\|v\|$, so that $Hv = u$.
2. Otherwise, u and v span two dimensions. We can choose an orthonormal basis B such that:

$$B^\top u = \alpha_1 e_1 + \alpha_2 e_2 \quad \text{and} \quad B^\top v = \beta_1 e_1 + \beta_2 e_2, \quad \alpha_1, \alpha_2, \beta_1, \beta_2 > 0.$$

This is possible because $u^\top v > 0$. Now, let $\Lambda = \text{diag}(\alpha_1/\beta_1, \alpha_2/\beta_2, 1, \dots, 1)$ and set:

$$H = B\Lambda B^\top,$$

so that by construction, $u = Hv$.

($\Leftarrow$). Suppose that $u = Hv$ for some positive-definite matrix H . Then $u^\top v = v^\top H v > 0$. $\square$

E Proof of Lemma 2

Lemma 2 (Gradient of smoothed best-response). *Let $(\Omega, \mathbf{f})$ be a multilinear game, where $\Omega_n \subset \mathbb{R}^{d_n}$ for each $n \in [N]$, and let $\mathbf{h}$ be a set of smooth and strictly convex regularizers, $h_n : \Omega_n \rightarrow \mathbb{R}$. Then, the associated β -smoothed best-response map Φ^β is well-defined and smooth, where the gradient of Φ^β at $\mathbf{x}$ is given by:*

$$\nabla \Phi^\beta(\mathbf{x}) = \frac{1}{\beta} \mathbf{H}(\mathbf{x})^{-1} \mathbf{J}(\mathbf{x}), \quad \text{where } \mathbf{H}(\mathbf{x})_{nm} = \begin{cases} \nabla_n^2 h_n(x_n) & n = m \\ 0_{d_n \times d_m} & n \neq m, \end{cases}$$

and $\mathbf{H}(\mathbf{x}) \in \mathbb{R}^{d \times d}$ is a positive-definite, block-diagonal matrix with $d = d_1 + \dots + d_N$.

Proof. The map $x'_n \mapsto f_n(x'_n, \mathbf{x}_{-n}) - \beta h_n(x'_n)$ is a strictly concave function over a convex set Ω_n . And so, it has a unique maximizer, implying that $\Phi^\beta(\mathbf{x})_n$ is well-defined.

Given any $\mathbf{x}$ that maps to a fully mixed strategy $\Phi^\beta(\mathbf{x}) \in \text{int}(\Omega)$, we can apply the Lagrange multiplier theorem to express $\Phi^\beta(\mathbf{x})_n$ for each $n \in [N]$ as the stationary point to the Lagrangian:

$$\mathcal{L}_i(x'_n, \lambda; \mathbf{x}) = f_n(x'_n, \mathbf{x}_{-n}) - \beta h_n(x'_n) - \lambda \mathbb{1}^\top x'_n,$$

where $\mathbb{1} \in \mathbb{R}^{d_n}$ is the all-ones vector. This implies that $\Phi^\beta(\mathbf{x})_n$ is the unique stationary point satisfying:

$$0 = \nabla_{x'_n} \mathcal{L}_n(x'_n, \lambda; \mathbf{x}) = \nabla_{x'_n} (f_n(x'_n, \mathbf{x}_{-n}) - \beta h_n(x'_n)) - \lambda \mathbb{1}.$$

Let $\Pi_n = \mathbf{I} - \frac{1}{d_n} \mathbb{1} \mathbb{1}^\top$ be the linear projection and $\Psi_n : \Omega_n \times \Omega \rightarrow T\Omega_n$ be the map:

$$\Psi_n(x'_n, \mathbf{x}) = \nabla_n (f_n(x'_n, \mathbf{x}_{-n}) - \beta h_n(x'_n)).$$

Then, $x'_n = \Phi^\beta(\mathbf{x})_n$ is the unique solution to $\Psi_n(x'_n, \mathbf{x}) = 0$. By the implicit function theorem, we obtain the gradient of Φ^β . The blocks on the diagonals are zero because f_n is multilinear, while $\nabla_n^2 h_n$ is positive definite on Ω_n since h_n is strictly convex. $\square$

F Proof of Theorem 1

Theorem 1 (Local uniform stability implies strategic Pareto optimality). *Let $\mathbf{x}^*$ be a Nash equilibrium of an N -player multilinear game, at which it is locally bi-directional. Suppose the interaction graph is connected at $\mathbf{x}^*$. If the game is locally uniformly stable at $\mathbf{x}^*$, then the equilibrium is strategically Pareto optimal.*

Proof. Under these assumptions, we claim that if a Nash equilibrium $\mathbf{x}^*$ has local uniform stability, then the game must be bilinear. Theorem 2 then shows that $\mathbf{x}^*$ must be strategically Pareto stationary. As the game is bilinear, Lemma 3 shows that $\mathbf{x}^*$ is strategically Pareto optimal.

To finish the proof, we show that the strategic component of the game is bilinear: we prove that for all triples of players $n, m, k \in [N]$, the higher-order derivative $\nabla_{nmk}^3 f_n \equiv 0$ is identically zero everywhere.

First, note that as the game is locally uniformly stable at $\mathbf{x}^*$, the Jacobian map $\mathbf{J}$ is uniformly $\boldsymbol{\lambda}$ -skew, where $\boldsymbol{\lambda}$ is a set of positive scalars independent of $\mathbf{x}$ (Lemma F.1). By a positive rescaling of the utilities, we may assume that $\mathbf{J}(\mathbf{x})$ is skew-symmetric everywhere, and that for any triple of players $n, m, k \in [N]$,

$$\forall \mathbf{x} \in \Omega, \quad J_{nm}(\mathbf{x}) = -J_{mn}(\mathbf{x})^\top, \quad J_{mk}(\mathbf{x}) = -J_{km}(\mathbf{x})^\top, \quad J_{kn}(\mathbf{x}) = -J_{nk}(\mathbf{x})^\top.$$

Taking derivatives and using the fact that the partial derivatives commute over smooth functions, we have:

$$\begin{aligned} \nabla_{nmk}^3 f_n + \nabla_{nmk}^3 f_m &\equiv 0 \\ \nabla_{nmk}^3 f_m + \nabla_{nmk}^3 f_k &\equiv 0 \\ \nabla_{nmk}^3 f_k + \nabla_{nmk}^3 f_n &\equiv 0 \end{aligned}$$

This can only happen if the triples $\nabla_{nmk}^3 f_n = \nabla_{nmk}^3 f_m = \nabla_{nmk}^3 f_k = 0$ are identically zero everywhere. $\square$

Lemma F.1. *Let $(\Omega, \mathbf{f})$ be an N -player multilinear game in canonical form. Let $\mathbf{x}^*$ be a Nash equilibrium that is locally bi-directional and is locally uniformly stable. There exists a set of positive scalars $\boldsymbol{\lambda}^*$ such that $\mathbf{J}(\mathbf{x})$ is $\boldsymbol{\lambda}^*$ -skew for all $\mathbf{x} \in \Omega$.*

Proof. Let the Nash equilibrium be locally bi-directional and locally uniformly stable, in particular, on a neighborhood U of $\mathbf{x}^*$. Then, Theorem 2 shows that $\mathbf{J}(\mathbf{x})$ is $\boldsymbol{\lambda}(\mathbf{x})$ -skew, for some local map $\mathbf{x} \mapsto \boldsymbol{\lambda}(\mathbf{x})$ on U to the positive orthant of $\mathbb{R}^N$. In particular, if we let $\lambda_{nm}(\mathbf{x}) = \lambda_m(\mathbf{x})/\lambda_n(\mathbf{x})$, we obtain:

$$J_{nm}(\mathbf{x}) = -\lambda_{nm}(\mathbf{x})J_{mn}(\mathbf{x})^\top. \quad (3)$$

We claim that $\lambda_{nm}(\mathbf{x})$ is in fact constant in $\mathbf{x}$. From this, it follows that $\mathbf{J}(\mathbf{x})$ is $\boldsymbol{\lambda}^*$ -skew where $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}(\mathbf{x}^*)$.

To prove the claim, fix some $\mathbf{z} \in \Omega$ and consider the ℓ th player. For this proof, define the vector $\mathbf{z}(\varepsilon)$ by:

$$\mathbf{z}(\varepsilon) = (z_\ell + \varepsilon, \mathbf{z} - \ell),$$

which is a perturbation only in the ℓ th player's strategy. By multilinearity, we have that for all $n, m \in [N]$:

$$\nabla_\ell J_{nm}(\mathbf{z}(\varepsilon)) = \nabla_\ell J_{nm}(\mathbf{z}),$$

showing that $\nabla_\ell J_{nm}(\mathbf{z}(\varepsilon))$ is constant in ε . In particular, this follows from $J_{nm}(\mathbf{z}(\varepsilon)) = J_{nm}(\mathbf{z}) + \nabla_\ell J_{nm}(\mathbf{z})\varepsilon$. To compress notation, define the following tensors:

$$T_{rst} = [\nabla_\ell J_{nm}(\mathbf{z})_{rs}]_t \quad \text{and} \quad T'_{rst} = [\nabla_\ell J_{mn}(\mathbf{z})_{sr}]_t.$$

Let $\lambda \equiv \lambda_{nm}(\mathbf{z}(\varepsilon))$. Then, applying the product rule to Equation (3) at $\mathbf{x} = \mathbf{z}(\varepsilon)$, we obtain that:

$$\begin{aligned} T_{rst} &= -\partial_t \lambda \cdot T'_{rsu} \varepsilon_u - \lambda T'_{rst} \\ &= -T'_{rsu} \cdot [\varepsilon_u \cdot \partial_t \lambda + \lambda \cdot \delta_{ut}]. \end{aligned}$$

The product rule applies because λ_{nm} is smooth; it is the quotient of two non-zero polynomials. We obtain:

$$T'_{rsu} \varepsilon_u \partial_t \lambda + (T_{rsu} + \lambda T'_{rsu}) \cdot \delta_{ut} = 0.$$

For each fixed pair (r, s) , the following (u, t) -slice is either the zero matrix or a full-rank (diagonal) matrix:

$$(T_{rsu} + \lambda T'_{rsu}) \cdot \delta_{ut}.$$

On the other hand, the following (u, t) -slice is either the zero matrix or a rank-1 matrix:

$$T'_{rsu} \varepsilon_u \partial_t \lambda.$$

As the sum of these two matrices is zero, the first must have been the zero matrix, whereby we obtain:

$$T_{rsu} + \lambda T'_{rsu} = 0.$$

This shows that $\lambda \equiv \lambda_{nm}(\mathbf{z}(\varepsilon))$ is constant in ε . This holds for all choices of joint strategy $\mathbf{z} \in U$ and player ℓ , so this implies that λ_{nm} is constant everywhere on the open set U . Finally, due to multilinearity, this implies that it is constant on the whole space Ω . $\square$

Notation 1 (Translation operator). For any $c \in \mathbb{R}^n$, let $T_c : \mathbb{M} \rightarrow \mathbb{M}$ be the *translation operator*:

$$(T_c f)(x) = f(x - c).$$

It is a linear map, and under the standard polynomial basis, it has matrix representation:

$$[T_c]_{S, S'} = \begin{cases} (-1)^{|S' \setminus S|} \prod_{i \in S' \setminus S} c_i & S' \supseteq S \\ 0 & S' \not\supseteq S. \end{cases} \quad (4)$$

Lemma F.2. Let $f, g \in \mathbb{M}$ and $S \subset [n]$. Suppose that for all $c \in \mathbb{R}^n$, we have that:

$$(T_c f)_S = (T_c g)_S.$$

Then, for all $S' \supseteq S$, the coefficients are equal $f_{S'} = g_{S'}$.

Proof. We proceed by induction. Let $P(k)$ denote the proposition: $f_{S'} = g_{S'}$ for all $S' \supseteq S$ where $|S' \setminus S| \leq k$.

- Base case: If $S' = S$, then by assumption $f_S = (T_0 f)_S = (T_0 g)_S = g_S$. Thus, $P(0)$ is true.
- Inductive step: Suppose that $P(k)$ holds. Let $S' \supseteq S$ and $|S' \setminus S| = k + 1$. Define $c = (c_1, \dots, c_n)$:

$$c_i = \begin{cases} 1 & i \in S' \\ 0 & i \notin S', \end{cases}$$

and let T_c be the translation with offset c . By (4), the S -coefficients of $T_c f$ and $T_c g$ are given by:

$$(T_c f)_S = \sum_{I \subseteq [n]} \alpha_I f_I \quad \text{and} \quad (T_c g)_S = \sum_{I \subseteq [n]} \alpha_I g_I,$$

where $\alpha_I = (-1)^{|I \setminus S|} \cdot \mathbb{1}\{S \subseteq I \subseteq S'\}$. By $P(k)$, all coefficients are equal $f_I = g_I$ when $S \subseteq I \subsetneq S'$. By the equality $(T_c f)_S = (T_c g)_S$, we obtain that:

$$\begin{aligned} 0 &= (T_c f)_S - (T_c g)_S \\ &= \sum_{S \subseteq I \subsetneq S'} \alpha_I \cdot (f_I - g_I) + \alpha_{S'} \cdot (f_{S'} - g_{S'}) \\ &= \alpha_{S'} \cdot (f_{S'} - g_{S'}), \end{aligned}$$

which implies that $f_{S'} = g_{S'}$ since $\alpha_{S'} \neq 0$. This proves $P(k + 1)$ holds.

The result follows by induction. □

G Proof of Lemma 3

Lemma 3 (Sufficient condition for strategic Pareto optimality). Let $(\Omega, \mathbf{f})$ be an N -player polymatrix game with a Nash equilibrium $\mathbf{x}^*$. If $\mathbf{x}^*$ is strategically Pareto stationary, then it is strategically Pareto optimal.

Proof. Suppose that $\mathbf{x}^*$ is strategically Pareto stationary. Then, there exist a set of positive scalars $\lambda_n > 0$ such that the weighted strategic components of the utilities sum to zero for all $\mathbf{x} \in \Omega$:

$$\sum_{n \in [N]} \lambda_n f_n(\mathbf{x}) \simeq_{\text{SE}} \sum_{n \in [N]} \sum_{m \neq n} \lambda_n (x_n - x_n^*)^\top J_{nm} (x_m - x_m^*) = 0.$$

It follows that there does not exist any $\mathbf{x} \in \Omega$ such that the following holds simultaneously for all $n \in [N]$,

$$f_n(\mathbf{x}) - f_n(\mathbf{x}^*) \simeq_{\text{SE}} \sum_{m \neq n} (x_n - x_n^*)^\top J_{nm} (x_m - x_m^*) > 0.$$

In particular, the Nash equilibrium is strategically Pareto optimal. □

H Proof of Proposition 1

In this section, we show that strategic Pareto optimality implies stationarity. We break down the proof of Proposition 1 into a sequence of results of increasing generality. The core of each proof is by contradiction: if a Nash equilibrium is not strategically stationary, then it is not Pareto optimal.

1. In Section H.1, we prove the result in the two-player case. The main innovation here is Lemma H.2; it is a linear-algebraic result showing that bilinear forms $(x, y) \mapsto x^\top My$ can be recovered from the related map $(x, y) \mapsto \text{sign}(x^\top My)$, up to a positive scale.
2. In Section H.2, we prove the general result for N players. This result makes use of Lemma H.3, which shows under some assumptions, that a Nash equilibrium is weakly Pareto optimal if and only if it is a strong Nash equilibrium. This will give us a set of easier conditions to check, which we make use of in the proof of Lemma H.4, which is precisely Proposition 1 in the case of games in canonical form.

Proposition 1 (Strategic Pareto stationarity is necessary for strategic Pareto optimality). *Let $(\Omega, \mathbf{f})$ be an N -player multilinear game with a Nash equilibrium $\mathbf{x}^*$. Let the interaction graph $G(\mathbf{x}^*)$ be connected and the game be bi-directional at $\mathbf{x}^*$. If the equilibrium $\mathbf{x}^*$ is strategically Pareto optimal, then it is strategically Pareto stationary.*

Proof. This follows from Lemma H.4. □

H.1 Weak Pareto optimality implies stationarity: the two-player case

In this section, we let $(\Omega, \mathbf{f})$ be a two-player polymatrix game in canonical form, with individual strategy spaces $\Omega_n = \mathbb{R}^{d_n}$ for $n \in \{1, 2\}$, and a Nash equilibrium $\mathbf{x}^*$ at the origin. As the game is bilinear, the game Jacobian is constant, which we denote by $\mathbf{J}$. In particular, we may assume that the utilities are given by:

$$f_1(\mathbf{x}) = x_1^\top J_{12} x_2 \quad \text{and} \quad f_2(\mathbf{x}) = x_2^\top J_{21} x_1,$$

so that the payoffs for both players at the Nash equilibrium are $f_1(\mathbf{0}) = f_2(\mathbf{0}) = 0$.

Recall that a Nash equilibrium $\mathbf{x}^*$ is strategically Pareto stationary if $\mathbf{J}$ is block skew-symmetrizable. In this case, this means that there exists $\Sigma \in \mathbb{R}^{d_1 \times d_2}$ and $\lambda > 0$ such that:

$$\mathbf{J} = \begin{bmatrix} & J_{12} \\ J_{21} & \end{bmatrix} = \begin{bmatrix} & \\ \Sigma^\top & -\lambda \Sigma \end{bmatrix}.$$

Lemma H.1 (Strategic Pareto optimality in two-player games). *Let $(\Omega, \mathbf{f})$ be a bi-directional, two-player polymatrix game in canonical form with a Nash equilibrium $\mathbf{x}^*$ and Jacobian $\mathbf{J}$. Suppose $\mathbf{x}^*$ is strategically Pareto optimal. Then, it is strategically Pareto stationary. That is, there exists some $\lambda > 0$ such that:*

$$J_{12} = -\lambda J_{21}^\top.$$

Proof. Let $\mathbf{x}^*$ be strategically Pareto optimal. Note that in this case, as the utilities are purely strategic, this simply means that $\mathbf{x}^*$ is weakly Pareto optimal. We claim that for all $\mathbf{x} \in \Omega$:

$$\text{sign}(f_1(\mathbf{x})) = \text{sign}(-f_2(\mathbf{x})), \tag{5}$$

where $\text{sign} : \mathbb{R} \rightarrow \{-, 0, +\}$ is the sign function. Assuming the claim for now, we obtain the result immediately from Lemma H.2, which shows that the sign of a bilinear map characterizes it up to a positive scalar. That is, the two matrices J_{12} and $-J_{21}^\top$ are related by some positive scaling factor $\lambda > 0$.

Proof of claim. To show the sign condition (5), we only need to rule out the two cases: (1) one utility is strictly negative $f_n(\mathbf{x}) < 0$ while the other is non-positive $f_{-n}(\mathbf{x}) \leq 0$, and the opposite case, (2) one utility is strictly positive $f_n(\mathbf{x}) > 0$ while the other is non-negative $f_{-n}(\mathbf{x}) \geq 0$.

We will only show how to rule out case (1) by contradiction, as the other case is completely analogous. Without loss of generality, suppose that:

$$f_1(\mathbf{x}) = x_1^\top J_{12} x_2 < 0 \quad \text{and} \quad f_2(\mathbf{x}) = x_1^\top J_{21}^\top x_2 \leq 0.$$

Since $f_1(\mathbf{x}) \neq 0$, we have that $x_2 \notin \ker(J_{12})$. By bi-directionality, we also have that $x_2 \notin \ker(J_{21}^\top)$. In particular, the two vectors $J_{12}x_2$ and $J_{21}^\top x_2$ are non-zero and lie on in same halfspace $\{w \in \mathbb{R}^{d_1} : x_1^\top w \leq 0\}$, where at least one is contained in the interior of the halfspace. And so, there exists some w_1 such that:

$$w_1^\top J_{12}x_2 > 0 \quad \text{and} \quad w_1^\top J_{21}^\top x_2 > 0.$$

For example, let:

$$w_1 = \frac{1}{2} \left(\frac{J_{12}x_2}{\|J_{12}x_2\|} + \frac{J_{21}^\top x_2}{\|J_{21}^\top x_2\|} \right).$$

When $\mathbf{x}' = (w_1, x_2)$, we obtain that both $f_1(\mathbf{x}') > 0$ and $f_2(\mathbf{x}') > 0$. Recall that $f_n(\mathbf{x}^*) = 0$. Thus, the joint vector $\mathbf{x}'$ strictly improves both utilities over $\mathbf{x}^*$, violating the weak Pareto optimality of $\mathbf{x}^*$. $\square$

Lemma H.2 (Bilinear maps are characterized by their signs up to scaling). *Let $A, B \in \mathbb{R}^{d_1 \times d_2}$ be matrices such that for all $x \in \mathbb{R}^{d_1}$ and $y \in \mathbb{R}^{d_2}$,*

$$\text{sign}(x^\top A y) = \text{sign}(x^\top B y),$$

where $\text{sign} : \mathbb{R} \rightarrow \{-, 0, +\}$ is the sign function,

$$\text{sign}(z) = \begin{cases} + & z > 0 \\ 0 & z = 0 \\ - & z < 0. \end{cases}$$

Then, there is some $\lambda > 0$ such that $A = \lambda B$.

Proof. Consider the singular value decomposition of A :

$$A = U \Sigma^A V^\top,$$

where $U \in \mathbb{R}^{d_1 \times d_1}$ and $V \in \mathbb{R}^{d_2 \times d_2}$ are orthogonal and the singular values $\sigma_j^A = \Sigma_{jj}^A$ are non-negative, where $j \in \{1, \dots, d_1 \wedge d_2\}$. Now define $\Sigma^B = U^\top B V$. We now show that in fact this yields a simultaneous singular value decomposition:

$$B = U \Sigma^B V^\top,$$

where all singular values σ_j^B are non-negative, and $\sigma_j^A = 0$ if and only if $\sigma_j^B = 0$. This is true because the signs of Σ^A and Σ^B match coordinate-wise. To see this, let e_j^1 and e_k^2 be from the standard basis vectors of $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, respectively. Then:

$$\Sigma_{jk}^A = (U e_j^1)^\top A (V e_k^2) \quad \text{and} \quad \Sigma_{jk}^B = (U e_j^1)^\top B (V e_k^2).$$

By assumption, the entries Σ_{jk}^B and Σ_{jk}^A share same signs.

Now, we show that $\Sigma^A = \alpha \Sigma^B$ for some $\lambda > 0$. Let u_j and v_k denote the left and right singular vectors—the columns of U and V . Without loss of generality, we may assume that $\sigma_1^A, \sigma_1^B > 0$, and we let:

$$\lambda = \frac{u_1^\top A v_1}{u_1^\top B v_1}.$$

For each $i > 1$, let $\alpha_i = \sqrt{\sigma_i^A / \sigma_1^A}$, so that:

$$(u_i + \alpha_i u_1)^\top A (v_i - \alpha_i v_1) = u_i^\top A v_i - \alpha_i^2 (u_1^\top A v_1) = 0,$$

where cross-terms such as $u_1^\top A v_i = 0$ because u_1 and v_i are unpaired singular vectors. By assumption that the signs match, this also implies that the same holds for B , where:

$$\begin{aligned} 0 &= (u_i + \alpha_i u_1)^\top B (v_i - \alpha_i v_1) \\ &= u_i^\top B v_i - \alpha_i^2 (u_1^\top B v_1) \\ &= \sigma_i^B - \frac{\sigma_i^A}{\sigma_1^A} \frac{u_1^\top A v_1}{\alpha}. \end{aligned}$$

And since $u_1^\top A v_1 / \sigma_1^A = 1$, we obtain that $\lambda \sigma_i^B = \sigma_i^A$. It follows that $A = \lambda B$. $\square$

H.2 Strategic Pareto optimality implies stationarity: the N -player case

We now generalize the result from Section H.1 to N -player multilinear games. We prove this result by way of the classical notion of *strong Nash equilibria* (Aumann, 1959). At the usual Nash equilibrium, no single player can improve their own utility without coordinating with others. The strong Nash equilibrium extends this form of stability to groups or coalitions of players. For any subset of players $S \subset [N]$, let $\mathbf{x}_S$ and $\mathbf{x}_{-S}$ denote the tuples $(x_n : n \in S)$ and $(x_n : n \notin S)$, respectively. Then:

Definition H.1 (Strong Nash equilibrium). Let $S \subset [N]$ be any subset or *coalition* of players. A joint decision vector $\mathbf{x}^* \in \Omega$ is *weakly S -Pareto optimal* if there does not exist some $\mathbf{x} \in \Omega$ such that:

$$\forall n \in S, \quad f_n(\mathbf{x}_S; \mathbf{x}_{-S}^*) > f_n(\mathbf{x}^*).$$

We say that $\mathbf{x}^*$ is a *strong Nash equilibrium* if it is weakly S -Pareto optimal for all $S \subset [N]$.

The following lemma establishes an equivalence between strong Nash stability and weak Pareto optimality for games in canonical form under mild structural assumptions.

Lemma H.3 (Strongly Nash equilibria are weakly Pareto). *Let $(\Omega, \mathbf{f})$ be a bi-directional, N -player multilinear game in canonical form with a Nash equilibrium $\mathbf{x}^*$. Suppose the interaction graph is connected at $\mathbf{x}^*$. Then, the equilibrium $\mathbf{x}^*$ is a strong Nash equilibrium if and only if $\mathbf{x}^*$ is weakly Pareto optimal.*

The forward direction is immediate by definition: when $\mathbf{x}^*$ is a strong Nash equilibrium, then applying the condition to $S = [N]$ shows that it also weakly Pareto optimal. For the reverse direction, suppose that $\mathbf{x}^*$ is not a strong Nash equilibrium. Thus, there is a coalition of players $S \subset [N]$ that can coordinate together to strictly improve their own utilities. We show that this group can be grown to include all N players, which then implies that $\mathbf{x}^*$ is not weakly Pareto optimal. In particular, the coalition can be extended to include any player that is adjacent to S in the interaction graph.

Proof of Lemma H.3. ($\Rightarrow$). If $\mathbf{x}^*$ is a strong Nash equilibrium, then by definition, it is weakly S -Pareto optimal where $S = [N]$. Thus, it is weakly Pareto optimal.

($\Leftarrow$). Suppose that $\mathbf{x}^*$ is not a strong Nash equilibrium. Let $S \subseteq [N]$ be a maximal coalition such that $\mathbf{x}^*$ is not weakly S -Pareto optimal: there is a joint decision vector of the form $\mathbf{x} = (\mathbf{x}_S, \mathbf{x}_{-S}^*)$ where all players in S strictly improve. If $S = [N]$, then this shows that $\mathbf{x}^*$ is not weakly Pareto optimal. To finish the proof, we now rule out the possibility that S is not all of $[N]$.

Let $G(\mathbf{x}^*) = (V, E)$ be the interaction graph at $\mathbf{x}^*$. Choose any player $n \notin S$ that is adjacent to some player $m \in S$ in $G(\mathbf{x}^*)$; this player exists because the graph is connected, and let $S' = S \cup \{n\}$. We will construct a joint strategy that is not S' -Pareto optimal, contradicting the maximality of S as desired.

Consider joint decision vectors $\mathbf{x}'$ of the form:

$$x'_\ell = \begin{cases} x_\ell + \varepsilon_\ell & \ell \in S' \\ x_\ell^* & \ell \notin S'. \end{cases}$$

Players in the coalition S play a perturbation of the joint strategy $\mathbf{x}_S$, the player n plays a perturbed Nash strategy, and the remainder play the Nash equilibrium strategy $\mathbf{x}_{-S'}^*$. By the continuity of the utilities, there exists $\delta > 0$ so that if $\|\mathbf{x} - \mathbf{x}'\| < \delta$, then all coalition members $m \in S$ maintain a strict improvement:

$$\forall m \in S, \quad f_m(\mathbf{x}') - f_m(\mathbf{x}^*) > 0.$$

Because the game is in canonical form, the n th player's utility at $\mathbf{x}'$ can be expressed in the form:

$$f_n(\mathbf{x}') - f_n(\mathbf{x}^*) = \varepsilon_n^\top A_n(\varepsilon_S),$$

where A_n is a polynomial in ε_S . In particular, it has a non-zero linear term:

$$A_n(\varepsilon_S) = \sum_{m \in S} J_{nm}(\mathbf{x}^*) \varepsilon_m + o(\|\varepsilon_S\|),$$

since the player n is connected to at least one player $m \in S$. As the polynomial A_n is not identically zero, there exist choices of $\mathbf{x}'$ satisfying $f_n(\mathbf{x}') - f_n(\mathbf{x}^*) > 0$ with $\|\varepsilon_S\| < \delta$. By this choice, we have $\|\mathbf{x}' - \mathbf{x}^*\| < \delta$, so that all players in S' achieve positive utilities. Thus, the equilibrium $\mathbf{x}^*$ is not weakly S' -Pareto optimal. $\square$

Lemma H.4 (Strategic Pareto optimality in N -player games). *Let $(\Omega, \mathbf{f})$ be a bi-directional, N -player multilinear game in canonical form with a Nash equilibrium $\mathbf{x}^*$. Suppose that $\mathbf{x}^*$ is a strategically Pareto optimal. Then, it is strategically Pareto stationary.*

Proof. In the following, let us denote $\mathbf{J} \equiv \mathbf{J}(\mathbf{x}^*)$ for short.

Let $\mathbf{x}^*$ be strategically Pareto optimal. Lemma H.3 shows that it is a strong Nash equilibrium. Thus, it is weakly S -Pareto optimal for all pairs $S = \{n, m\}$. The game restricted to these two players is bilinear; Lemma H.1 implies that there is some $a_{mn} > 0$ so that:

$$J_{nm} = -a_{mn} J_{mn}^\top.$$

We prove that $\mathbf{x}^*$ is strategically Pareto stationary by showing that $\mathbf{J}$ is block skew-symmetrizable, where there is a set of positive scalars $(\lambda_n)_{n \in [N]}$ such that:

$$\forall n, m \in [N], \quad a_{mn} = \frac{\lambda_m}{\lambda_n}. \quad (6)$$

We prove the consistency condition (6) by induction over the subgraphs of the interaction graph G of the game. Without loss of generality, we may assume that G is connected, since players in different connected components are completely independent of each other. In the following, if $G' = (V', E')$ is a subgraph of G , we let $\mathbf{J}[G']$ denote the matrix:

$$J[G']_{nm} = \begin{cases} J_{nm} & (n, m) \in E' \\ 0 & (n, m) \notin E', \end{cases}$$

where the block $J[T]_{nm} = 0$ is zeroed out when $(n, m) \notin E'$.

Base case. Let $T = (V_T, E_T)$ be a spanning tree of G . Endow T with a rooted tree structure, so that there is a root vertex $n_0 \in V_T$ and each other vertex $n \in V_T$ has a unique parent vertex $\text{pa}(n) \in V_T$. Inductively define $\lambda_n > 0$ for $n \in V_T$ so that:

$$\lambda_{n_0} = 1 \quad \text{and} \quad \lambda_n = a_{mn} \lambda_m \text{ where } m = \text{pa}(n).$$

By construction, we have that $\mathbf{J}[T]$ is block skew-symmetrizable by the set of positive scalars $(\lambda_n)_{n \in [N]}$.

Inductive step. Assume that there is a subgraph $G' \subset G$ such that $\mathbf{J}[G']$ is block skew-symmetrizable by the positive scalars $(\lambda_n)_{n \in [N]}$. Let $e \in E \setminus E'$ be an edge not in G' . Moreover, choose e so that there is a minimal cycle graph $C = (V_C, E_C) \subset G$ containing the edge e such that all other edges in C are in G' :

$$e \in E_C \quad \text{and} \quad E_C \setminus \{e\} \subset E'.$$

In particular, there are no edges in G between any two vertices of C that are not already in C .

Suppose there are r players in the cycle. Let us rename these players so that $V_C = [r]$. Actually, for convenience, let $V_C = \mathbb{Z}/r\mathbb{Z}$, where player r can also be called player 0. Naturally, order the players so that:

$$(n, m) \in E_C \quad \Longleftrightarrow \quad n - m \pmod{r} = \pm 1.$$

Let us also choose the labels so that $e = (0, 1)$. We now prove that the consistency condition (6) holds; that is, $a_{r1} = \lambda_r / \lambda_1$. We proceed by contradiction and construct $\mathbf{x}_C$ such that:

$$\forall n \in V_C, \quad f_n(\mathbf{x}_C; \mathbf{x}_{-C}^*) - f_n(\mathbf{x}^*) > 0,$$

This would contradict the assumption that $\mathbf{x}^*$ is weakly Pareto optimality over the coalition of players in C .

Assume without loss of generality that $a_{1r} < \lambda_1 / \lambda_r$ (otherwise, we can use the relation $a_{r1} = a_{1r}^{-1}$ and reverse the order of the players). Choose any $0 < \alpha < 1 - a_{1r} \cdot \frac{\lambda_r}{\lambda_1}$ and define the sequence:

$$\forall n \in [r], \quad \alpha_n = 1 - \alpha \cdot \frac{n-1}{r-1}.$$

Now, for each $n \in [r]$, choose any sequence $w_n \notin \ker(J_{n-1,n})$. We define $\mathbf{x}_C = (x_1, \dots, x_r)$ as follows:

$$x_n = \begin{cases} w_1 & n = 1 \\ \frac{\alpha_n w_n}{\lambda_n w_n^\top J_{n,n-1} x_{n-1}} & 1 < n \leq r. \end{cases}$$

Since the game is in canonical form, we have that $\mathbf{x}^* = \mathbf{0}$ is centered at the origin, and that when all player outside of V_C plays $\mathbf{x}_{-C}^*$, the change in utility is:

$$f_n(\mathbf{x}_C; \mathbf{x}_{-C}^*) - f_n(\mathbf{x}^*) = x_n^\top J_{n,n-1} x_{n-1} + x_n^\top J_{n,n+1} x_{n+1} + o(\|\mathbf{x}_C\|^2),$$

where $o(\|\mathbf{x}_C\|^2)$ captures terms in the polynomial that have degree at least three. To obtain the contradiction, it suffices to show that there exists $\mathbf{x}_C$ such that the bilinear terms are positive for all players in C ; by replacing $\mathbf{x}_C$ with $\varepsilon \mathbf{x}_C$ with sufficiently small $\varepsilon > 0$, the higher-order terms can be made to have vanishing contribution to the utility relative to the bilinear term.

By the inductive hypothesis, $\mathbf{J}[G']$ is block skew-symmetrizable by the set of scalars $(\lambda_n)_{n \in [N]}$, so that:

$$\begin{aligned} \forall n \in [r-1], \quad f_n(\mathbf{x}_C; \mathbf{x}_{-C}^*) - f_n(\mathbf{x}^*) &= x_n^\top J_{n,n-1} x_{n-1} - \frac{\lambda_{n+1}}{\lambda_n} \cdot x_{n+1}^\top J_{n+1,n} x_n + o(\|\mathbf{x}_C\|^2) \\ &= \lambda_n^{-1} \cdot (\alpha_n - \alpha_{n+1}) + o(\|\mathbf{x}_C\|^2) \\ &= \lambda_n^{-1} \cdot \alpha / (r-1) + o(\|\mathbf{x}_C\|^2). \end{aligned}$$

We also have that the change of the utility of the r th player is:

$$\begin{aligned} f_r(\mathbf{x}_C; \mathbf{x}_{-C}^*) - f_r(\mathbf{x}^*) &= \lambda_r^{-1} \cdot \alpha_r - a_{1r} \cdot x_1 J_{1r} x_r + o(\|\mathbf{x}_C\|^2) \\ &= \lambda_r^{-1} \left(1 - \alpha - a_{1r} \cdot \frac{\lambda_r}{\lambda_1} \right) + o(\|\mathbf{x}_C\|^2), \end{aligned}$$

where in the first step, we used the equality $J_{r1} = -a_{1r} J_{1r}^\top$. By construction, the lowest-order, bilinear term of the utility is positive, of each player in the coalition C using this choice of $\mathbf{x}_C$, obtaining the desired contradiction. And so, we must have that $a_{1r} = \lambda_1 / \lambda_r$, implying the consistency condition (6). $\square$

I Proof of Theorem 2

Theorem 2 (Pointwise uniform stability is equivalent to strategic Pareto stationarity). *Let $\mathbf{x}^*$ be a Nash equilibrium of an N -player multilinear game, at which it is bi-directional. Suppose that the interaction graph is connected at $\mathbf{x}^*$. Then, $\mathbf{x}^*$ is uniformly stable if and only if it is strategically Pareto stationary.*

Proof of Theorem 2. Uniform stability and strategic Pareto stationarity both depend only on the bilinear terms of the strategic component of games. Thus, we may assume without loss of generality that the game is a purely-strategic, N -player polymatrix game with the Jacobian $\mathbf{J} \equiv \mathbf{J}(\mathbf{x}^*)$. In fact, we may assume it is in canonical form (Definition C.1), where the Nash equilibrium is centered at the origin $\mathbf{x}^* = \mathbf{0}$, with utilities:

$$f_n(\mathbf{x}) = \sum_{m \neq n} x_n J_{nm} x_m.$$

($\Rightarrow$). For Nash equilibria in polymatrix games in canonical form, Proposition 1 shows that the strategic Pareto stationarity is equivalent to weak Pareto optimality. And so, to prove the forward direction, we may show the contrapositive: if $\mathbf{x}^*$ is not weakly Pareto optimal, then it is not uniformly stable.

When $\mathbf{x}^*$ is not weakly Pareto optimal, there is a joint strategy $\mathbf{x} \in \Omega$ such that all players improve:

$$f_n(\mathbf{x}) - f_n(\mathbf{x}^*) = x_n^\top \sum_{m \neq n} J_{nm} x_m > 0.$$

Lemma 1 shows that for each player $n \in [N]$, there exists $H_n > 0$ such that:

$$H_n \sum_{m \neq n} J_{nm} x_m = x_n.$$

Let $\mathbf{H} = \text{diag}(H_1, \dots, H_N)$. We obtain that $\mathbf{H}\mathbf{J}\mathbf{x} = \mathbf{x}$, so that $\mathbf{H}\mathbf{J}$ has a positive eigenvalue $\lambda = 1$. Thus, the game is not uniformly stable at $\mathbf{x}^*$.

($\Leftarrow$). When $\mathbf{x}^*$ is strategically Pareto stationary, then by definition, there is a block-diagonal matrix $\mathbf{\Lambda}$ such that $\mathbf{\Lambda}\mathbf{J} = \mathbf{\Sigma}$ is skew-symmetric, where each diagonal block is the diagonal matrix $\Lambda_{nn} = \lambda_n I_{d_n \times d_n}$.

Let $\mathbf{H}$ be any positive-definite block-diagonal matrix. Then, $\mathbf{H}^{-1}\mathbf{J}$ is similar to a skew-symmetric matrix:

$$\mathbf{H}^{-1}\mathbf{J} \sim (\mathbf{H}^{-1/2}\mathbf{\Lambda}^{1/2})\mathbf{\Sigma}(\mathbf{\Lambda}^{1/2}\mathbf{H}^{-1/2}),$$

where we used the fact that $\mathbf{\Lambda}\mathbf{H}^{-1} = \mathbf{H}^{-1}\mathbf{\Lambda}$. By Lemma I.1, $\mathbf{H}^{-1}\mathbf{J}$ has purely imaginary eigenvalues. Since this holds for all possible choice of $\mathbf{H}$, the game is uniformly stable at $\mathbf{x}^*$. $\square$

Lemma I.1 (Skew-symmetric matrices have imaginary eigenvalues, Theorem 7 in Chapter 8 of Lax (2007)). *If $M \in \mathbb{R}^{m \times m}$ is skew-symmetric, then $\text{spec}(M) \subset i\mathbb{R}$.*

Theorem I.1 (Eigenvalues of matrix powers, Theorem 1.1.6 of Horn and Johnson (2012)). *Let $M \in \mathbb{C}^{m \times m}$ be a square matrix. Let $p(t)$ be a given polynomial of degree k . If (λ, v) is an eigenvalue–eigenvector pair of M , then $(p(\lambda), v)$ is an eigenvalue–eigenvector pair of $p(M)$. Conversely, if $k \geq 1$ and if μ is an eigenvalue of $p(M)$, then there is some eigenvalue λ of M such that $\mu = p(\lambda)$.*

J Proof of Proposition 2

Proposition 2 (Inapproximability of unstable equilibria). *Let $(\Omega, \mathbf{f})$ be an N -player multilinear game with an interior Nash equilibrium $\mathbf{x}^*$. Suppose that $\mathbf{x}^*$ is not uniformly stable. Then, there is some collection of smooth and strictly convex regularizers $\mathbf{h}$ such that smoothed best-response dynamics does not stabilize to $\mathbf{x}^*$. In fact, even if the sequence of β -smoothed equilibria $\mathbf{x}^\beta$ converges to $\mathbf{x}^*$, each $\mathbf{x}^\beta$ is an unstable fixed point of the (Φ^β, η) -averaging dynamics (1) for all $\eta \in (0, 1)$ and sufficiently small $\beta > 0$.*

Proof. In the following, let $\mathbf{H}(\mathbf{x})$ denote the block-diagonal matrix with $H_n(\mathbf{x}) = \nabla_n^2 h_n(x_n)$ given a set of smooth and strictly convex regularizers $\mathbf{h}$.

As $\mathbf{x}^*$ is not uniformly stable, we can choose $\mathbf{h}$ such that the eigenvalues of $\mathbf{H}^{-1}(\mathbf{x}^*)\mathbf{J}(\mathbf{x}^*)$ are not all purely imaginary (such a choice exists by Lemma L.2). In fact, it must have an eigenvalue λ^* with positive real part $\Re(\lambda^*) > c > 0$. This is because the matrix $\mathbf{H}^{-1}(\mathbf{x}^*)\mathbf{J}(\mathbf{x}^*)$ has zero trace—the diagonal blocks of the matrix $H_n(\mathbf{x}^*)J_{nn}(\mathbf{x}^*)$ in multilinear games is zero—and so, the sum of all eigenvalues must also be zero.

The existence of eigenvalues with positive real parts, bounded away from zero, extends to an open region around $\mathbf{x}^*$ by continuity. In particular, the matrix-valued function $\mathbf{x} \mapsto \mathbf{H}^{-1}(\mathbf{x})\mathbf{J}(\mathbf{x})$ is continuous, so by the continuity of eigenvalues (Theorem J.1), there is an open set U around $\mathbf{x}^*$ so that for all $\mathbf{x} \in U$, each matrix $\mathbf{H}^{-1}(\mathbf{x})\mathbf{J}(\mathbf{x})$ also has an eigenvalue $\lambda(\mathbf{x})$, where its real part is bounded away from zero:

$$\Re(\lambda(\mathbf{x})) > c/2 > 0.$$

We now show that the dynamics do not stabilize to $\mathbf{x}^*$. That is, either the sequence of β -smoothed equilibria $\mathbf{x}^\beta$ does not converge to $\mathbf{x}^*$ as β goes to zero, or the dynamics become unstable around $\mathbf{x}^\beta$. If the sequence does not converge, then there is nothing to do. So, without loss of generality, we may assume that there exists some sufficiently small $\beta_0 < 2/c$ so that all β -smoothed equilibria remain in U when $\beta < \beta_0$:

$$\forall \beta < \beta_0, \quad \mathbf{x}^\beta \in U.$$

We can now analyze the linear stability of $\mathbf{x}^\beta$ under the (Φ^β, η) -averaging dynamics (1). Its Jacobian is:

$$\nabla \left((1 - \eta) \mathbf{x} + \eta \Phi^\beta(\mathbf{x}) \right) = (1 - \eta) \mathbf{I} + \frac{\eta \mathbf{H}^{-1}(\mathbf{x})\mathbf{J}(\mathbf{x})}{\beta},$$

where $\mathbf{I}$ is the identity matrix, making use of the gradient computation Lemma 2. The eigenvalues are:

$$(1 - \eta) + \frac{\eta\lambda}{\beta}, \quad \text{where } \lambda \in \text{spec}(\mathbf{H}^{-1}(\mathbf{x}^\beta)\mathbf{J}(\mathbf{x}^\beta)).$$

Whenever $\beta < \beta_0$, we have that $\mathbf{x}^\beta \in U$, so that there is at least one λ whose real part is at least $c/2$. Moreover, as $\beta_0 < 2/c$, it follows that the modulus is bounded below for all $\eta \in (0, 1)$:

$$\left| (1 - \eta) + \frac{\eta\lambda}{\beta} \right| > 1 - \eta + \frac{\eta c}{2\beta} > 1.$$

The fixed point $\mathbf{x}^\beta$ is unstable when $\beta < \beta_0$ as the Jacobian of the dynamics has an eigenvalue whose modulus is greater than 1. This follows from Lyapunov's indirect method (Bof et al., 2018, Theorem 3.3). $\square$

Theorem J.1 (Continuity of eigenvalues, Theorem 1 of Elsner (1985)). *Let $A, B \in \mathbb{C}^{m \times m}$ have spectra $\{\lambda_1, \dots, \lambda_m\}$ and $\{\mu_1, \dots, \mu_m\}$, respectively. Then:*

$$\max_{j \in [m]} \min_{i \in [m]} |\lambda_i - \mu_j| \leq (\|A\|_F + \|B\|_F)^{(m-1)/m} \cdot \|A - B\|_F^{1/m}.$$

K Proof of Theorem 3

Theorem 3 (Convergence to equilibria in uniformly-stable regions). *Let $(\Omega, \mathbf{f})$ be an N -player multilinear game with a Nash equilibrium $\mathbf{x}^*$ that is locally uniformly stable. Over arbitrary choice of smooth and strictly convex regularizer $\mathbf{h}$, all smoothed best-response dynamics stabilize to $\mathbf{x}^*$. In particular, there is a constant $C_{\mathbf{f}} > 0$ depending only on the game so that when $\eta \leq C_{\mathbf{f}}\beta^2$, the (Φ^β, η) -averaging dynamics globally converge:*

$$\|\mathbf{x}(t) - \mathbf{x}^\beta\| \leq \exp\left(-\frac{\eta t + \ln N}{2}\right).$$

Proof. Let Φ^β be any choice of smoothed best-response map induced by a set of smooth and strictly convex regularizer $\mathbf{h}$ and $\beta > 0$. Recall from Lemma 2 that its Jacobian is:

$$\nabla \Phi^\beta(\mathbf{x}) = \frac{1}{\beta} \mathbf{H}(\mathbf{x})^{-1} \mathbf{J}(\mathbf{x}),$$

where $\mathbf{H}(\mathbf{x})$ is the block-diagonal matrix with the n th block being $\nabla_n^2 h_n(x_n)$ and $\mathbf{J}(\mathbf{x})$ is the game Jacobian. Let $U \subset \Omega$ be any open ball containing $\mathbf{x}^*$ such that $\|\mathbf{H}(\mathbf{x})^{-1} \mathbf{J}(\mathbf{x})\| \leq L$ is bounded by a constant $L > 0$. Such a ball exists by the continuity of $\nabla \Phi^\beta$. We now show that for sufficiently small η , the (Φ^β, η) -averaging dynamics contracts toward $\mathbf{x}^\beta$ when initialized within U , implying that $\mathbf{x}^\beta$ is asymptotically stable. In particular, we show that at each iteration, the distance to $\mathbf{x}^\beta$ contracts by a factor of $\exp(-\Omega(\eta))$ whenever:

$$\eta \leq \frac{\beta^2}{1 + L^2}.$$

Fix any $\mathbf{x} \in U$. By the mean-value theorem, there is some $\bar{\mathbf{x}} \in U$, a convex combination of $\mathbf{x}$ and $\mathbf{x}^\beta$, such that $\Phi^\beta(\mathbf{x}) - \Phi^\beta(\mathbf{x}^\beta) = \nabla \Phi^\beta(\bar{\mathbf{x}})(\mathbf{x} - \mathbf{x}^\beta)$. We use $\bar{\mathbf{x}}$ to define $\mathbf{M}_\eta(\mathbf{x})$ as the matrix:

$$\mathbf{M}_\eta(\mathbf{x}) = (1 - \eta)\mathbf{I} + \eta \nabla \Phi^\beta(\bar{\mathbf{x}}),$$

where $\mathbf{I} \in \mathbb{R}^{d \times d}$ is the identity. The (Φ^β, η) -averaging dynamics can be re-centered at $\mathbf{x}^\beta$, which becomes:

$$\begin{aligned} \mathbf{x}(t+1) - \mathbf{x}^\beta &= (1 - \eta)(\mathbf{x}(t) - \mathbf{x}^\beta) + \eta(\Phi^\beta(\mathbf{x}(t)) - \Phi^\beta(\mathbf{x}^\beta)) \\ &= \mathbf{M}_\eta(\mathbf{x}(t))(\mathbf{x}(t) - \mathbf{x}^\beta). \end{aligned}$$

The dynamics contract to $\mathbf{x}^\beta$ exponentially quickly if the spectral norm $\|\mathbf{M}_\eta(\mathbf{x})\|_2$ is bounded by $1 - \Omega(1)$,

$$\|\mathbf{x}(t+1) - \mathbf{x}^\beta\| \leq \|\mathbf{M}_\eta(\mathbf{x}(t))\|_2 \|\mathbf{x}(t) - \mathbf{x}^\beta\|.$$

Indeed, we have:

$$\begin{aligned}
\|\mathbf{M}_\eta(\mathbf{x})\|_2 &\leq \left| 1 - \eta + i\eta \frac{L}{\beta} \right| \\
&= \sqrt{1 - 2\eta + \eta^2 \left(1 + \frac{L^2}{\beta^2} \right)} \\
&\leq \exp \left(-\eta + \frac{1}{2}\eta^2 \left(1 + \frac{L^2}{\beta^2} \right) \right) \\
&\leq \exp \left(-\frac{\eta}{2} \right),
\end{aligned}$$

where the first inequality follows from $\|\nabla\Phi\| \leq L$, the second from the inequality $\sqrt{1-z} \leq e^{-z/2}$, and the last inequality holds whenever $\eta \leq \beta^2/(1+L^2)$.

The rate then follows from Theorem 2, along with the bound on the diameter of the joint simplex space:

$$\sup_{\mathbf{x}, \mathbf{x}' \in \Omega} \|\mathbf{x} - \mathbf{x}'\| \leq \sqrt{N}.$$

□

L Beyond interior Nash equilibria

The convergence results Proposition 1 and Theorem 3 considered only interior Nash equilibria. This section extends these result to Nash equilibria when they are on the boundary of the simplex.

The following lemma shows that if $\mathbf{x}^*$ is an isolated, non-interior equilibrium, then the game can be locally *reduced* by removing actions that are not in the support of $\mathbf{x}^*$. These actions turn out to be strictly dominated, and so the game-theoretic equilibrium $\mathbf{x}^*$ is preserved after removing dominated strategies.

Lemma L.1. *Let $(\Omega, \mathbf{f})$ be an N -player normal-form game with a quasi-strong equilibrium $\mathbf{x}^* \in \Omega$. There exists a neighborhood U containing $\mathbf{x}^*$ such that actions not supported by $\mathbf{x}^*$ are strictly dominated. That is, there exists some $\varepsilon > 0$ such that for each player $n \in [N]$ and for all $\mathbf{x} \in U$ and $i \notin \text{supp}(x_n^*)$,*

$$f_n(e_{n,i}; \mathbf{x}_{-n}) < \max_{j \in [k_n]} f_n(e_{n,j}; \mathbf{x}_{-n}) - \varepsilon.$$

Proof. As $\mathbf{x}^*$ is quasi-strong, for each player $n \in [N]$, there is some $\varepsilon_n > 0$ such that whenever a pure strategy $i \in [k_n]$ is not supported, meaning that $i \notin \text{supp}(\mathbf{x}^*)$, then it is strictly dominated:

$$f_n(e_{n,i}; \mathbf{x}_{-n}^*) < \max_{j \in [k_n]} f_n(e_{n,j}; \mathbf{x}_{-n}^*) - \varepsilon_n.$$

This is because the set of pure strategies that are best-responses to $\mathbf{x}^*$ are precisely those in $\text{supp}(x_n^*)$. As this is a finite set of alternatives, there must be some positive gap between those that maximize utility and those that do not. And as f_n is continuous, for a sufficiently small ball B_n around $\mathbf{x}^*$, the same domination condition remains true, although perhaps with a smaller gap. That is, when $i \notin \text{supp}(\mathbf{x}^*)$,

$$\forall \mathbf{x} \in B_n, \quad f_n(e_{n,i}; \mathbf{x}_{-n}) < \max_{j \in [k_n]} f_n(e_{n,j}; \mathbf{x}_{-n}) - \varepsilon_n/2.$$

The result follows by letting U be the intersection $B_1 \cap \dots \cap B_N$ and $\varepsilon = \min_{n \in [N]} \varepsilon_n/2$. □

Corollary L.1 (Reduction to interior equilibria). *Let $\mathbf{x}^*$ be a quasi-strong equilibrium of a normal-form game. Then, it is an interior Nash equilibrium of the reduced game at $\mathbf{x}^*$.*

Proof. For each player $n \in [N]$ and alternative $i \in \text{supp}(x_n^*)$, the probability mass on i is bounded away from zero since it is positive $x_{n,i}^* > 0$. Then, either it is the unique action that is supported, in which case we say that x_n^* is trivially contained in the interior of the reduced simplex $\Omega_n(x_n^*)$. Otherwise, it must be bounded away from one $x_{n,i}^* < 1$, since there is at least one other supported action. Either way:

$$x_n^* \in \text{int}(\Omega_n(x_n^*)).$$

Thus, $\mathbf{x}^*$ is an interior equilibrium of the reduced game. □

L.1 Calculus on the probability simplex

This section formalizes smoothness on the boundary of the simplex. Let $\Delta \equiv \Delta^{k-1}$ be the $(k-1)$ -dimensional probability simplex. We view it as the union of open manifolds of dimensions ranging from 0 through $k-1$. In particular, we decompose the simplex as follows:

$$\Delta = \bigcup_{S \subset [k]} \Delta_S, \quad \text{where } \Delta_S := \{x \in \Delta^{k-1} : \text{supp}(x) = S\},$$

so that Δ_S is the interior of a probability simplex with $|S| - 1$ dimensions. For each $S \subset [k]$, we say that the closure of Δ_S , denoted by $\overline{\Delta}_S$, is a *face* of the simplex. To take the gradient and Hessian of map $h : \Delta \rightarrow \mathbb{R}$ at a point $x \in \Delta$ with $\text{supp}(x) = S$, we simply take them with respect to the simplex Δ_S .

Definition L.1 (Calculus on the simplex). For each $S \subset [k]$, let $T\Delta_S$ denote the tangent space of Δ_S , which we may naturally identify with the following subspace of $\mathbb{R}^k$:

$$T\Delta_S := \{v \in \mathbb{R}^k : \text{supp}(v) = S \text{ and } \mathbf{1}^\top v = 0\}.$$

When a map $h_S : \Delta_S \rightarrow \mathbb{R}$ is smooth, let $\nabla_S h_S : \Delta_S \rightarrow T\Delta_S$ denote its gradient. Let $h : \Delta \rightarrow \mathbb{R}$ be a map whose restriction onto Δ_S is smooth for all $S \subset [k]$. Its *gradient* and *Hessian* at x are defined as:

$$\nabla h(x) := \nabla_S h(x) \quad \text{and} \quad \nabla^2 h(x) := \nabla_S^2 h(x),$$

where $S = \text{supp}(x)$. Thus, $\nabla h(x) \in \mathbb{R}^k$ and $\nabla^2 h(x) \in \mathbb{R}^{k \times k}$.

Definition L.2 (Smoothness on the simplex). A map $g : \Delta \rightarrow \mathbb{R}$ is *smooth* on the simplex if it can be extended to a smooth map on an open set $U \subset \mathbb{R}^k$ containing Δ .

Any utility $f : \Omega \rightarrow \mathbb{R}$ in a normal-form game is smooth on the simplex, since it immediately extends to a multilinear map $p : \mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_N} \rightarrow \mathbb{R}$, where p is formally the same polynomial as f . While their function values coincide on Ω , they are not the same function because their domains are distinct. In particular, their derivatives are not generally equal, as they map to distinct tangent spaces. Nevertheless, we will use the representation inherited from the ambient space for each $\Delta^{k_n-1} \subset \mathbb{R}^{k_n}$ specified in Definition L.1.

L.2 The rate of suboptimal plays in smoothed-best responses

Example L.1 (Entropic regularizer is linearly steep). Let $h(x) = \sum x_i \log x_i$ on $\Omega = \Delta^{k-1}$. For any $v \in T\Omega$, the solution $x^\beta(v) := \nabla h^{-1}(v/\beta)$ is given by:

$$x^\beta(v)_i = \frac{\exp(v_i/\beta)}{\sum_{j=1}^k \exp(v_j/\beta)} \leq \exp\left(-\frac{1}{\beta} \cdot \left(\max_{j \in [k]} v_j - v_i\right)\right).$$

It follows that the entropic map is linearly steep since $\exp(-\varepsilon/\beta)$ decays to zero much faster than β .

Lemma 4 (Dominated strategies occur at β -sublinear rates in β -smoothed equilibria). Let $(\Omega, \mathbf{f})$ be an N -player normal-form game and $\mathbf{h}$ be a collection of linearly steep regularizers. Suppose that as β goes to zero, the sequence of β -smoothed equilibria $\mathbf{x}^\beta$ converges to a quasi-strict equilibrium $\mathbf{x}^*$. Let $\Pi : \Omega \rightarrow \Omega(\mathbf{x}^*)$ be the orthogonal projection onto the support of $\mathbf{x}^*$. Then:

$$\lim_{\beta \rightarrow 0} \frac{\|\mathbf{x}^\beta - \Pi \mathbf{x}^\beta\|}{\beta} = 0.$$

Proof of Lemma 4. By Lemma L.1, there exists an open region U around $\mathbf{x}^*$ and some $\varepsilon > 0$ such that for each player $n \in [N]$ and for all $\mathbf{x} \in U$ and $i \notin \text{supp}(\mathbf{x}^*)$, the action i is ε -suboptimal. In particular, $\nabla_n f_n(\mathbf{x}) \in T\Omega_n$ denote the gradient of f_n on the manifold Ω_n . We obtain that:

$$\nabla_n f_n(\mathbf{x})_i < \max_{j \in [k_n]} \nabla_n f_n(\mathbf{x})_j - \varepsilon.$$

Since $(\mathbf{x}^\beta)_\beta$ converges to $\mathbf{x}^*$, for sufficiently small β , the β -smoothed equilibrium $\mathbf{x}^\beta$ eventually remains in U as β goes to zero. In particular, $\nabla f_n(\mathbf{x}^\beta)$ is contained in the set $V_i(\varepsilon)$ where the i th alternative is ε -suboptimal, defined in Definition 25. As the regularizer h_n is linearly steep, the β -smoothed equilibrium places sublinearly little mass on any action $i \notin \text{supp}(x_n^*)$:

$$\lim_{\beta \rightarrow 0} \frac{x_{n,i}^\beta}{\beta} = 0.$$

Thus, the projection of the smoothed equilibrium $\mathbf{x}^\beta$ onto $\Omega(\mathbf{x}^*)$ satisfies:

$$\frac{\|\mathbf{x}^\beta - \Pi \mathbf{x}^\beta\|}{\beta} \leq \sum_{n \in [N]} \sum_{i \notin \text{supp}(x_n^*)} \frac{2x_{n,i}^\beta}{\beta},$$

by the triangle inequality. Taking the limit as β goes to zero yields the result. $\square$

Lemma L.2 (Hessians of linearly steep regularizers). *Let $\Delta = \Delta^{k-1}$ be the $(k-1)$ -dimensional simplex. Let $H \in T\Delta \otimes T\Delta$ be positive definite and $x \in \text{int}(\Delta)$. There is a linearly steep regularizer $h : \Delta \rightarrow \mathbb{R}$ where:*

$$\nabla h(x) = 0 \quad \text{and} \quad \nabla^2 h(x) = H.$$

Proof. Define $\text{Sym}^+(T\Delta)$ to be the set of positive-definite matrices in $T\Delta \otimes T\Delta$. Fix $x \in \text{int}(\Omega_i)$ and consider the family of steep regularizers:

$$h(x; \lambda, A, w) = \lambda \sum_{i=1}^k x_i \log x_i + \frac{1}{2} \|A(x - w)\|_2^2,$$

where $\lambda > 0$, $A \in \mathbb{R}^{k \times k}$ is invertible, and $w \in \mathbb{R}^{m_i}$. We show that h is steep and that for this fixed x ,

$$\text{Sym}^+(T\Delta) = \left\{ \nabla^2 h(x; \lambda, A) : \lambda > 0 \text{ and } A \text{ invertible} \right\}.$$

Under the coordinate representation given by Definition L.1, the gradient on the simplex Δ is given by:

$$\nabla h(x; \lambda, A) = \left(I - \frac{1}{k} \mathbb{1} \mathbb{1}^\top \right) \left[\lambda \log x + A^\top A(x - w) \right],$$

where $\log x$ is applied component-wise. Thus, h is steep, since $\log x_i \rightarrow -\infty$ as $x_i \rightarrow 0$. There exists $w \in \mathbb{R}^k$ such that $\nabla h(x_i) = 0$, by selecting:

$$w = x_i - \lambda(A^\top A)^{-1} \log x_i.$$

The Hessian on the simplex is given by:

$$\nabla^2 h(x; \lambda, A) = \left(I - \frac{1}{k} \mathbb{1} \mathbb{1}^\top \right) \left[\lambda \cdot \text{diag}(1/x) + A^\top A \right] \left(I - \frac{1}{k} \mathbb{1} \mathbb{1}^\top \right),$$

where $\text{diag}(1/x)$ is the diagonal matrix with diagonal entries $1/x_i$. For every $M \in \text{Sym}^+(T\Delta)$, there is a sufficiently small $\lambda > 0$ such that the following matrix is positive-definite, so there is an invertible matrix A :

$$A^\top A = M + \mathbb{1} \mathbb{1}^\top - \lambda \cdot \text{diag}(1/x).$$

It follows that $\nabla^2 h(x; \lambda, A) = M$, showing that $\text{Sym}^+(T\Delta) \subset \{\nabla_{\Omega_i}^2 h_i(z; \lambda, A) : \lambda > 0 \text{ and } A \text{ invertible}\}$. The reverse inclusion holds as h is strongly convex, so its Hessian on the simplex is contained in $\text{Sym}^+(T\Delta)$. $\square$

L.3 Convergence to non-interior Nash equilibria

Theorem 4 (Convergence to non-interior equilibria). *Let $(\Omega, \mathbf{f})$ be an N -player, normal-form game and let $\mathbf{h}$ be a collection of linearly steep, proper regularizers. Suppose that $\mathbf{x}^*$ is a quasi-strict Nash equilibrium and that it is the limit of the β -smoothed equilibria $\mathbf{x}^\beta$ as β goes to zero. If the reduced game around $\mathbf{x}^*$ is locally uniformly stable, then smoothed best-response dynamics stabilize to $\mathbf{x}^*$. In particular, there exists a constant $C_{\mathbf{f}} > 0$ such that for all sufficiently small $\beta > 0$ and $\eta \leq C_{\mathbf{f}}\beta^2$, the (Φ^β, η) -averaging dynamics is locally asymptotically stable at $\mathbf{x}^\beta$, where the operator norm of its Jacobian at $\mathbf{x}^\beta$ is bounded by:*

$$\|(1 - \eta)\mathbf{I} + \eta\nabla\Phi^\beta(\mathbf{x}^\beta)\|_2 \leq \exp\left(-\frac{\eta}{2}\right).$$

Proof. Let Φ^β be any choice of smoothed best-response map induced by a set of linearly steep and proper regularizers $\mathbf{h}$ and $\beta > 0$. Its Jacobian under the coordinate representation specified by Definition L.1 is:

$$\nabla\Phi^\beta(\mathbf{x}) = \frac{1}{\beta}\mathbf{H}(\mathbf{x})^+\mathbf{J}(\mathbf{x}),$$

where $\mathbf{H}(\mathbf{x})$ is the block-diagonal matrix with the n th block being $\nabla_n^2 h_n(x_n)^+$ and $\mathbf{J}(\mathbf{x})$ is the game Jacobian, as computed by Lemma 2. Define $\mathbf{G}(\mathbf{x}) := \mathbf{H}(\mathbf{x})^+\mathbf{J}(\mathbf{x})$ and let $\Pi : \Omega \rightarrow \Omega(\mathbf{x}^*)$ be the orthogonal projection to the support of $\mathbf{x}^*$. Let $U \subset \Omega$ be any open ball containing $\mathbf{x}^*$ such that $\|\mathbf{H}(\mathbf{x})^+\mathbf{J}(\mathbf{x})\| \leq L$ is bounded by a constant $L > 0$. Such a ball exists because the regularizers are proper, so that $\nabla\Phi^\beta$ is smooth. Moreover, the reduced game around $\mathbf{x}^*$ is locally uniformly stable, so this ball can be chosen such that:

$$\text{spec}(\mathbf{G}(\Pi\mathbf{x}^\beta)) \subset i\mathbb{R}.$$

To prove that the (Φ^β, η) -averaging dynamics is a local contraction at $\mathbf{x}^\beta$ whenever η is sufficiently small, we show that the Jacobian of these dynamics have operator norm bounded by 1. The Jacobian of the (Φ^β, η) -averaging dynamics can be decomposed as follows:

$$\begin{aligned} (1 - \eta)\mathbf{I} + \eta\nabla\Phi^\beta(\mathbf{x}^\beta) &= (1 - \eta)\mathbf{I} + \frac{\eta}{\beta}\mathbf{G}(\mathbf{x}^\beta) \\ &= (1 - \eta)\mathbf{I} + \frac{\eta}{\beta}\mathbf{G}(\Pi\mathbf{x}^\beta) + \frac{\eta}{\beta}\left(\mathbf{G}(\mathbf{x}^\beta) - \mathbf{G}(\Pi\mathbf{x}^\beta)\right) \\ &= (1 - \eta)\mathbf{I} + \frac{\eta}{\beta}\mathbf{G}(\Pi\mathbf{x}^\beta) + \eta\left(\frac{\nabla\mathbf{G}(\bar{\mathbf{x}}^\beta)(\mathbf{x}^\beta - \Pi\mathbf{x}^\beta)}{\beta}\right), \end{aligned}$$

where in the last step applies the mean-value theorem to find some $\bar{\mathbf{x}}^\beta$ that is a convex combination of $\mathbf{x}^\beta$ and $\Pi\mathbf{x}^\beta$, once again making use of the smoothness of the pseudoinverse of the Hessian of the regularizers.

When β is sufficiently small, then $\Pi\mathbf{x}^\beta$ is contained in U , so that the eigenvalue of $\mathbf{G}(\Pi\mathbf{x}^\beta)$ is $i\lambda$ for some real-valued scalar λ with $|\lambda| < L$. By Lemma 4, the norm of the last term can be made arbitrarily small:

$$\left\|\frac{\nabla\mathbf{G}(\bar{\mathbf{x}}^\beta)(\mathbf{x}^\beta - \Pi\mathbf{x}^\beta)}{\beta}\right\|_2 \leq \frac{C\|\mathbf{x}^\beta - \Pi\mathbf{x}^\beta\|}{\beta} \leq o(\beta),$$

where we may uniformly bound the norm of $\nabla\mathbf{G}$ over Ω since it is a smooth map on a compact set. Let us choose β small enough so that this term is no more than $1/2$. Thus, the operator norm of the Jacobian of the dynamics is bounded by:

$$\|(1 - \eta)\mathbf{I} + \eta\nabla\Phi^\beta(\mathbf{x}^\beta)\|_2 \leq \left|1 - \frac{\eta}{2} + i\eta\frac{L}{\beta}\right| \leq \exp\left(-\frac{\eta}{2}\right),$$

where the last inequality holds when $\eta \leq \beta^2/(1 + 4L^2)$; see the proof of Theorem 3 for this computation. $\square$