Implicit Learning Dynamics in Stackelberg Games: Equilibria Characterization, Convergence Analysis, and Empirical Study

Tanner Fiez [1] **Benjamin Chasnov** [1] **Lillian Ratliff** [1]

Abstract

Contemporary work on learning in continuous games has commonly overlooked the hierarchical decision-making structure present in machine learning problems formulated as games, instead treating them as simultaneous play games and adopting the Nash equilibrium solution concept. We deviate from this paradigm and provide a comprehensive study of learning in Stackelberg games. This work provides insights into the optimization landscape of zero-sum games by establishing connections between Nash and Stackelberg equilibria along with the limit points of simultaneous gradient descent. We derive novel gradient-based learning dynamics emulating the natural structure of a Stackelberg game using the implicit function theorem and provide convergence analysis for deterministic and stochastic updates for zero-sum and general-sum games. Notably, in zero-sum games using deterministic updates, we show the only critical points the dynamics converge to are Stackelberg equilibria and provide a local convergence rate. Empirically, our learning dynamics mitigate rotational behavior and exhibit benefits for training generative adversarial networks compared to simultaneous gradient descent.

1. Introduction

The emerging coupling between game theory and machine learning can be credited to the formulation of learning problems as interactions between competing objectives and strategic agents. Indeed, generative adversarial networks (GANs) (Goodfellow et al., 2014), robust supervised learning (Madry et al., 2018), reinforcement and multi-agent reinforcement learning (Dai et al., 2018; Zhang et al., 2019), and hyperparameter optimization (Maclaurin et al., 2015) problems can be cast as zero-sum or general-sum continuous action games. To obtain solutions in a tractable manner, gradient-based algorithms have gained attention.

Given the motivating applications, much of the contemporary work on learning in games has focused on zero-sum games with non-convex, non-concave objective functions and seeking stable critical points or local equilibria. A number of techniques have been proposed including optimistic and extra-gradient algorithms (Daskalakis et al., 2018; Daskalakis & Panageas, 2018; Mertikopoulos et al., 2019), gradient adjustments (Balduzzi et al., 2018; Mescheder et al., 2017), and opponent modeling methods (Zhang & Lesser, 2010; Foerster et al., 2018; Letcher et al., 2019; Schäfer & Anandkumar, 2019). However, only a select number of algorithms can guarantee convergence to stable critical points satisfying sufficient conditions for a *local Nash equilibrium* (LNE) (Mazumdar et al., 2019; Adolphs et al., 2019).

The dominant perspective in machine learning applications of game theory has been focused on simultaneous play. However, there are many problems exhibiting a hierarchical order of play, and in a game theoretic context, such problems are known as Stackelberg games. The Stackelberg equilibrium (Von Stackelberg, 2010) solution concept generalizes the min-max solution to general-sum games. In the simplest formulation, one player acts as the leader who is endowed with the power to select an action knowing the other player (follower) plays a best-response. This viewpoint has long been researched from a control perspective on games (Basar & Olsder, 1998) and in the bilevel optimization community (Danskin, 1967; 1966; Zaslavski, 2012).

The work from a machine learning perspective on games with a hierarchical decision-making structure is sparse and exclusively focuses on zero-sum games. In the most relevant theoretical work, Jin et al. (2019) show that all stable critical points of simultaneous gradient descent with a timescale separation between players approaching infinity satisfy sufficient conditions for a *local Stackelberg equilibrium* (LSE). The closest empirical work we are aware of is on unrolled GANs (Metz et al., 2017), where the leader (generator) optimizes a surrogate cost function that depends on parameters

[1]Department of Electrical and Computer Engineering, University of Washington, Seattle, Washington. Correspondence to: Tanner Fiez <fiezt@uw.edu>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020.

of the follower (discriminator) that have been 'rolled out' until an approximate local optimum is reached. This behavior intuitively approximates a hierarchical order of play and consequently the success of the unrolling method as a training mechanism provides some evidence supporting the LSE solution concept. In this paper, we provide a step toward bridging the gap between theory and practice along this perspective by developing implementable learning dynamics with convergence guarantees to critical points satisfying sufficient conditions for a LSE.

Contributions. Motivated by the lack of algorithms focusing on games exhibiting an order of play, we provide a study of learning in Stackelberg games including equilibria characterization, novel learning dynamics and convergence analysis, and an illustrative empirical study. The primary benefits of this work to the community include an enlightened perspective on the consideration of equilibrium concepts reflecting the underlying optimization problems present in machine learning applications formulated as games and an algorithm that provably converges to critical points satisfying sufficient conditions for a LSE in zero-sum games.

We provide a characterization of LSE via sufficient conditions on player objectives and term points satisfying the conditions *differential Stackelberg equilibria* (DSE). We show DSE are generic amongst LSE in zero-sum games. This means except on a set of measure zero in the class of zero-sum continuous games, DSE and LSE are equivalent. While the placement of *differential Nash equilibria* (DNE) amongst critical points in continuous games is reasonably well understood, an equivalent statement cannot be made regarding DSE. Accordingly, we draw connections between the solution concepts in the class of zero-sum games. We show that DNE are DSE, which indicates the solution concept in hierarchical play games is not as restrictive as the solution concept in simultaneous play games. Furthermore, we reveal that there exist stable critical points of simultaneous gradient descent dynamics that are DSE and not DNE. This insight gives meaning to a broad class of critical points previously thought to lack game-theoretic meaning and may give some explanation for the adequacy of solutions not satisfying sufficient conditions for LNE in GANs. To characterize this phenomenon, we provide necessary and sufficient conditions for when such points exist.

We derive novel gradient-based learning dynamics emulating the natural structure of a Stackelberg game from the sufficient conditions for a LSE and the implicit function theorem. The dynamics can be viewed as an analogue to simultaneous gradient descent incorporating the structure of hierarchical play games. In stark contrast to the simultaneous play counterpart, we show in zero-sum games the only stable critical points of the dynamics are DSE and such equilibria must be stable critical points of the dynamics. Using this fact and saddle avoidance results, we show the only critical points the discrete time algorithm converges to given deterministic gradients are DSE and provide a local convergence rate. In general-sum games, we cannot guarantee the only critical point attractors of the deterministic learning algorithms are DSE. However, we give a local convergence rate to critical points which are DSE. For stochastic gradient updates, we obtain analogous convergence guarantees asymptotically for each game class.

Empirically, we show that our dynamics result in stable learning compared to simultaneous gradient dynamics when training GANs. To gain insights into the placement of DNE and DSE in the optimization landscape, we analyze the eigenvalues of relevant game objects and observe convergence to neighborhoods of equilibria. Finally, we show that our dynamics can scale to computationally intensive problems.

2. Preliminaries

We now formalize the games we study, present equilibrium concepts accompanied by sufficient condition characterizations, and formulate Stackelberg learning dynamics.

2.1. Game Formalisms

Consider a non-cooperative game between two agents where player 1 is deemed the *leader* and player 2 the *follower*. The leader has cost $f_1 : X \to \mathbb{R}$ and the follower has cost $f_2 : X \to \mathbb{R}$, where $X = X_1 \times X_2 \in \mathbb{R}^m$ with $X_1 \in \mathbb{R}^{m_1}$ and $X_2 \in \mathbb{R}^{m_2}$ denoting the action spaces of the leader and follower, respectively.[1] We assume throughout that each f_i is sufficiently smooth: $f_i \in C^q(X, \mathbb{R})$ for some $q \geq 2$. For zero-sum games, the game is defined by costs $(f_1, f_2) = (f, -f)$. In words, we consider the class of two-player smooth games on continuous, unconstrained actions spaces. The designation of 'leader' and 'follower' indicates the order of play between the agents, meaning the leader plays first and the follower second.

In a *Stackelberg game*, the leader and follower aim to solve the following optimization problems, respectively:

$$\min_{x_1 \in X_1} \{f_1(x_1, x_2) \big| \, x_2 \in \arg\min_{y \in X_2} f_2(x_1, y)\}, \quad \text{(L)}$$

$$\min_{x_2 \in X_2} f_2(x_1, x_2). \quad \text{(F)}$$

This contrasts with a *simultaneous* play game in which each player i is faced with the optimization problem $\min_{x_i \in X_i} f_i(x_i, x_{-i})$. The learning algorithms we formulate are such that the agents follow myopic update rules which take steps in the direction of steepest descent for the respective optimizations problems.

[1] Our results hold more generally for action spaces that are precompact subsets of the Euclidean space since they are local.

2.2. Equilibria Concepts and Characterizations

Before formalizing learning rules, let us first discuss the equilibrium concept studied for simultaneous play games and contrast it with that which is studied in the hierarchical play counterpart. The typical equilibrium notion in continuous games is the pure strategy Nash equilibrium in simultaneous play games and the Stackelberg equilibrium in hierarchical play games. Each notion of equilibria can be characterized as the intersection points of the reaction curves of the players (Basar & Olsder, 1998). We focus our attention on local notions of the equilibrium concepts as is standard in learning in games since the objective functions we consider need not be convex or concave.

Definition 1 (Local Nash (LNE)). *The joint strategy $x^* \in X$ is a local Nash equilibrium on $U_1 \times U_2 \subset X_1 \times X_2$ if for each $i \in \{1, 2\}$, $f_i(x^*) \leq f_i(x_i, x^*_{-i})$, $\forall\, x_i \in U_i \subset X_i$.*

Definition 2 (Local Stackelberg (LSE)). *Consider $U_i \subset X_i$ for each $i \in \{1, 2\}$. The strategy $x_1^* \in U_1$ is a local Stackelberg solution for the leader if, $\forall x_1 \in U_1$,*

$$\sup_{x_2 \in R_{U_2}(x_1^*)} f_1(x_1^*, x_2) \leq \sup_{x_2 \in R_{U_2}(x_1)} f_1(x_1, x_2),$$

where $R_{U_2}(x_1) = \{y \in U_2 | f_2(x_1, y) \leq f_2(x_1, x_2), \forall x_2 \in U_2\}$. Moreover, (x_1^*, x_2^*) for any $x_2^* \in R_{U_2}(x_1^*)$ is a local Stackelberg equilibrium on $U_1 \times U_2$.

While characterizing existence of equilibria is outside the scope of this work, we remark that Nash equilibria exist for convex costs on compact and convex strategy spaces and Stackelberg equilibria exist on compact strategy spaces (Basar & Olsder, 1998, Thm. 4.3, Thm. 4.8, & §4.9). This means the class of games on which Stackelberg equilibria exist is broader than on which Nash equilibria exist. Existence of local equilibria is guaranteed if the neighborhoods and cost functions restricted to those neighborhoods satisfy the assumptions of the cited results.

Predicated on existence, equilibria can be characterized in terms of sufficient conditions on player costs. We denote $D_i f_i$ as the derivative of f_i with respect to x_i, $D_{ij} f_i$ as the partial derivative of $D_i f_i$ with respect to x_j, and $D(\cdot)$ as the total derivative.[2] The following gives sufficient conditions for a LNE as given in Definition 1.

Definition 3 (Differential Nash (DNE) Ratliff et al. (2016)). *The joint strategy $x^* \in X$ is a differential Nash equilibrium if $D_i f_i(x^*) = 0$ and $D_i^2 f_i(x^*) > 0$ for each $i \in \{1, 2\}$.*

Analogous sufficient conditions can be stated to characterize a LSE from Definition 2.

Definition 4 (Differential Stackelberg (DSE)). *The joint strategy $x^* = (x_1^*, x_2^*) \in X$ is a differential Stackelberg equilibrium if $Df_1(x^*) = 0$, $D_2 f_2(x^*) = 0$, $D^2 f_1(x^*) > 0$, and $D_2^2 f_2(x^*) > 0$ where $x_2^* = r(x_1^*)$ and $r(\cdot)$ implicitly defined by $D_2 f_2(x^*) = 0$.*

Game Jacobians play a key role in determining stability of critical points. Let

$$\omega(x) = (D_1 f_1(x), D_2 f_2(x))$$

be the vector of individual gradients for the simultaneous play game and

$$\omega_{\mathcal{S}}(x) = (Df_1(x), D_2 f_2(x))$$

as the equivalent for the Stackelberg game. Observe that Df_1 is the total derivative of f_1 with respect to x_1 given x_2 is implicitly a function of x_1, capturing the fact that the leader operates under the assumption that the follower will play a (local) best response to x_1. We note that the reaction curve of the follower to x_1 may not be unique. However, under sufficient conditions on a local Stackelberg solution x, locally $D_2 f_2(x) = 0$ and $\det(D_2^2 f_2(x)) \neq 0$ so that Df_1 is well defined via the implicit mapping theorem (Lee, 2012).

The vector field $\omega(x)$ forms the basis of the well-studied simultaneous gradient learning dynamics and the Jacobian of the dynamics is given by

$$J(x) = \begin{bmatrix} D_1^2 f_1(x) & D_{12} f_1(x) \\ D_{21} f_2(x) & D_2^2 f_2(x) \end{bmatrix}.$$

Similarly, the vector field $\omega_{\mathcal{S}}(x)$ serves as the foundation of the learning dynamics we formulate in Section 2.4 and analyze throughout. The Jacobian of the Stackelberg vector field $\omega_{\mathcal{S}}(x)$ is given by

$$J_{\mathcal{S}}(x) = \begin{bmatrix} D_1(Df_1(x)) & D_2(Df_1(x)) \\ D_{21} f_2(x) & D_2^2 f_2(x) \end{bmatrix}. \tag{1}$$

A critical point is called non-degenerate if the determinant of the vector field Jacobian is non-zero. We denote by $\mathbb{C}_-^\circ$ and $\mathbb{C}_+^\circ$ the open left and right half complex planes. Moreover, a critical point x^* of $\dot{x} = -\omega(x)$ is stable if $\mathrm{spec}(-J(x^*)) \subset \mathbb{C}_-^\circ$ or equivalently $\mathrm{spec}(J(x^*)) \subset \mathbb{C}_+^\circ$. Similarly, a critical point x^* of $\dot{x} = -\omega_{\mathcal{S}}(x)$ is stable if $\mathrm{spec}(-J_{\mathcal{S}}(x^*)) \subset \mathbb{C}_-^\circ$ or equivalently $\mathrm{spec}(J_{\mathcal{S}}(x^*)) \subset \mathbb{C}_+^\circ$.

Noting that the Schur complement of $J_{\mathcal{S}}(x)$ with respect to $D_2^2 f_2(x)$ is identically $D^2 f(x_1, r(x_1))$, we give alternative but equivalent sufficient conditions as those in Definition 4 in terms of $J_{\mathcal{S}}(x)$. Here, $\mathtt{S}_1(\cdot)$ denotes the Schur complement of $(\cdot)$ with respect to the bottom block matrix in $(\cdot)$. The proof of the following result is in Appendix B.

Proposition 1. *Consider a game (f_1, f_2) defined by $f_i \in C^q(X, \mathbb{R})$, $i = 1, 2$ with $q \geq 2$ and player 1 (without loss of generality) taken to be the leader. Let x^* satisfy $D_2 f_2(x^*) = 0$ and $D_2^2 f_2(x^*) > 0$. Then $Df_1(x^*) = 0$ and $\mathtt{S}_1(J_{\mathcal{S}}(x^*)) > 0$ if and only if x^* is a DSE. Moreover, in zero-sum games, $\mathtt{S}_1(J_{\mathcal{S}}(x)) = \mathtt{S}_1(J(x))$.*

[2] Example: given $f(x, r(x))$, $Df = D_1 f + D_2 f Dr$.

2.3. Genericity and Structural Stability

A natural question is how common is it for local equilibria to satisfy sufficient conditions, meaning in a formal mathematical sense, what is the gap between necessary and sufficient conditions in games. Towards addressing this, it has been shown that DNE are generic amongst LNE and structurally stable in the classes of zero-sum and general-sum continuous games, respectively (Ratliff et al., 2016; Mazumdar & Ratliff, 2019). The results say that except on a set of measure zero in each class of games, DNE = LNE and the equilibria persist under sufficiently smooth perturbations to the costs. We give analogous results for DSE in the class of zero-sum games in this section and provide proofs in Appendix C. The following result allows us to conclude that for a generic zero-sum game, DSE = LSE.

Theorem 1. *For the class of two-player, zero-sum continuous games $(f, -f)$ where $f \in C^q(\mathbb{R}^m, \mathbb{R})$ with $q \geq 2$,* DSE *are generic amongst* LSE*. That is, given a generic $f \in C^q(\mathbb{R}^m, \mathbb{R})$, all* LSE *of the game $(f, -f)$ are* DSE*.*

A critical point x^* of the vector field $\omega_{\mathcal{S}}(x)$ is hyperbolic if there are no eigenvalues of $J_{\mathcal{S}}(x^*)$ with zero real part. We now show that in generic zero-sum games, LSE are hyperbolic critical points of the vector field $\omega_{\mathcal{S}}(x)$, which is desirable property owing to the convergence implications.

Corollary 1. *For the class of two-player, zero-sum continuous games $(f, -f)$ where $f \in C^q(\mathbb{R}^m, \mathbb{R})$ with $q \geq 2$,* LSE *are generically non-degenerate, hyperbolic critical points of the vector field $\omega_{\mathcal{S}}(x)$.*

As a final result in this section, we show that DSE are structural stable in the class of zero-sum games. Structurally stability ensures that differential Stackelberg equilibria are robust and persist under smooth perturbations.

Theorem 2. *For the class of two-player, zero-sum continuous games $(f, -f)$ where $f \in C^q(\mathbb{R}^m, \mathbb{R})$ with $q \geq 2$,* DSE *are structurally stable: given $f \in C^q(\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}, \mathbb{R})$, $\zeta \in C^q(\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}, \mathbb{R})$, and a* DSE *$(x_1, x_2) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, there exists neighborhoods $U \subset \mathbb{R}$ of zero and $V \subset \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ such that $\forall\, t \in U$ there exists a unique* DSE *$(\tilde{x}_1, \tilde{x}_2) \in V$ for the zero-sum game $(f + t\zeta, -f - t\zeta)$.*

Before moving on, we remark that important classes of non-generic games certainly exist. In games where the cost function of the follower is bilinear, LSE can exist which do not satisfy the sufficient conditions outlined in Definition 4. As a simple example, $x^* = (0, 0)$ is a LSE for the zero-sum game defined by $f(x_1, x_2) = x_1 x_2$ and not a DSE since $D_2^2 f_2(x) = 0 \; \forall \; x \in X$. Since such games belong to a degenerate class in the context of the genericity result we provide, they naturally deserve special attention and algorithmic methods. While we do not focus our attention on this class of games, we do propose some remedies to allow our proposed learning algorithm to successfully seek out equilibria in them. In the experiments section, we discuss a regularized version of our dynamics that injects a small perturbation to cure degeneracy problems leveraging the fact that DSE are structurally stable. Further details can be found in Appendix H.1. Finally, for bimatrix games with finite actions it is common to reparameterize the problem using a softmax function to obtain mixed policies on the simplex (Fudenberg et al., 1998). We explore this viewpoint in Appendix H.4 on a parameterized bilinear game.

Algorithm 1 Deterministic Stackelberg Learning Dynamics

Input: $x_0 \in X$, learning rates $\gamma_2 > \gamma_1 > 0$
for $k = 0, 1, \ldots$ **do**
  $\omega_{\mathcal{S},1} \leftarrow D_1 f_1(x_k) - D_{21} f_2(x_k)^\top (D_2^2 f_2(x_k))^{-1} D_2 f_1(x_k)$
  $\omega_{\mathcal{S},2} \leftarrow D_2 f_2(x_k)$
  $x_{1,k+1} \leftarrow x_{1,k} - \gamma_1 \omega_{\mathcal{S},1}$
  $x_{2,k+1} \leftarrow x_{2,k} - \gamma_2 \omega_{\mathcal{S},2}$
end for

2.4. Stackelberg Learning Dynamics

Recall that $\omega_{\mathcal{S}}(x) = (Df_1(x), D_2 f_2(x))$ is the vector field for Stackelberg games and it, along with its Jacobian $J_{\mathcal{S}}(x)$, characterize sufficient conditions for a DSE. Letting $\omega_{\mathcal{S},i}$ be the i–th component of $\omega_{\mathcal{S}}$, the leader total derivative is $\omega_{\mathcal{S},1}(x) = D_1 f_1(x) - D_{21} f_2(x)^\top (D_2^2 f_2(x))^{-1} D_2 f_1(x)$ where $Dr \equiv -(D_2^2 f_2(x))^{-1} \circ D_{21} f_2$ with r defined by the implicit function theorem (Lee, 2012) in a neighborhood of a differential Stackelberg, meaning where $\omega_{\mathcal{S},2}(x) = 0$ with $\det(D_2^2 f_2(x)) \neq 0$ (which is also holds generically at critical points by Lemma C.3). The Stackelberg learning rule we study for each player in discrete time is given by

$$x_{i,k+1} = x_{i,k} - \gamma_{i,k} h_{\mathcal{S},i}(x_k). \tag{2}$$

In deterministic learning players have oracle gradient access so that $h_{\mathcal{S},i}(x) = \omega_{\mathcal{S},i}(x)$. We study convergence for deterministic learning in Section 4.1 and Algorithm 1 provides example pseudocode. In stochastic learning players have unbiased gradient estimates and $h_{\mathcal{S},i}(x_k) = \omega_{\mathcal{S},i}(x_k) + w_{k+1,i}$ where $\{w_{i,k}\}$ is player i's noise process. We provide convergence analysis for stochastic learning in Section 4.2.

3. Implications for Zero-Sum Settings

Before presenting convergence analysis of the update in (2), we draw connections between Nash and Stackelberg equilibria in zero-sum games and discuss the relevance to applications such as adversarial learning. To do so, we evaluate the limiting behavior of the dynamics from a continuous time viewpoint since the discrete time system closely approximates this behavior for suitably selected learning rates. While we provide the intuition behind the results here, the formal proofs of the results are in Appendix D.

Let us first show that for zero-sum games, all stable critical points of $\dot{x} = -\omega_{\mathcal{S}}(x)$ are DSE and vice versa.

Proposition 2. *In zero-sum games $(f, -f)$ with $f \in C^q(X, \mathbb{R})$ for $q \geq 2$, a joint strategy $x \in X$ is a stable critical point of $\dot{x} = -\omega_{\mathcal{S}}(x)$ if and only if x is a* DSE. *Moreover, if f is generic, a point x is a stable critical point of $\dot{x} = -\omega_{\mathcal{S}}(x)$ if and only if it is a* LSE.

The result follows from the structure of the Jacobian of $\omega_{\mathcal{S}}(x)$, which is lower block triangular with player 1 and 2 as the leader and follower, respectively. Proposition 2 implies that with appropriate stepsizes the update rule in (2) will only converge to Stackelberg equilibria and thus, unlike simultaneous gradient descent, will not converge to spurious locally asymptotically stable points that lack game-theoretic meaning (see, e.g., Mazumdar et al. (2020)).

This previous result begs the question of which stable critical points of the dynamics $\dot{x} = -\omega(x)$ are DSE? The following gives a partial answer to the question and also indicates that recent works seeking DNE are also seeking DSE.

Proposition 3. *In zero-sum games $(f, -f)$ with $f \in C^q(X, \mathbb{R})$ for $q \geq 2$,* DNE *are* DSE. *Moreover, if f is generic,* LNE *are* LSE.

This result follows from the facts that the conditions of a DNE imply $\mathtt{S}_1(J(x)) > 0$ and that non-degenerate DNE are generic amongst LNE within the class of zero-sum games (Mazumdar & Ratliff, 2019). In the zero-sum setting, the fact that Nash equilibria are a subset of Stackelberg equilibria for finite games is well-known (Basar & Olsder, 1998). We extend this result locally to continuous action space games. Similar to our work and concurrently, Jin et al. (2019) show that LNE are local min-max solutions.

In Proposition D.1 of Appendix D, we show the previous results imply all DNE are stable critical points of both $\dot{x} = -\omega(x)$ and $\dot{x} = -\omega_{\mathcal{S}}(x)$. This leaves the question of the meaning of stable points of $\dot{x} = -\omega(x)$ which are not DNE.

Finding Meaning in Spurious Stable Critical Points. We focus on the question of when stable fixed points of $\dot{x} = -\omega(x)$ are DSE and not DNE. It was shown by Jin et al. (2019) that not all stable points of $\dot{x} = -\omega(x)$ are local min-max or local max-min equilibria since one can construct a function such that $D_1^2 f(x)$ and $-D_2^2 f(x)$ are *both* not positive definite but the real parts of the eigenvalues of $J(x)$ are positive. It appears to be much harder to characterize when a stable critical point of $\dot{x} = -\omega(x)$ is not a DNE but is a DSE since it requires the follower's individual Hessian to be positive definite. Indeed, it reduces to a fundamental problem in linear algebra in which the relationship between the eigenvalues of the sum of two matrices is largely unknown without assumptions on the structure of the matrices (Knutson & Tao, 2001).

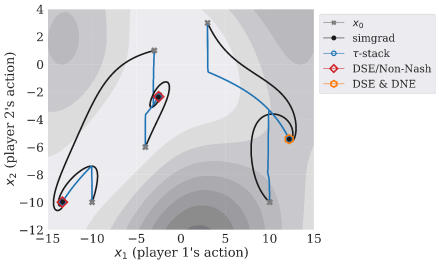


Figure 1. Example demonstrating existence of DSE and DSE that are not DNE: $\mathcal{G} = (f, -f)$ where f is defined in (3) with $a = 0.15, b = 0.25$. There are two stable points of simultaneous gradient descent which are DSE, but not DNE.

In Appendix E, we provide necessary and sufficient conditions for attractors at which the follower's Hessian is positive definite to be DSE. Taking intuition from the expression $\mathtt{S}_1(J(x)) = D_1^2 f(x) - D_{21} f(x)^\top (D_2^2 f(x))^{-1} D_{21} f(x)$, the conditions are derived from relating $\mathrm{spec}(D_1^2 f)$ to $\mathrm{spec}(D_2^2 f)$ via $D_{12} f$. To illustrate this fact, consider the following example in which stable points are DSE and not DNE—meaning points $x \in X$ at which $D_1^2 f(x) \not\succ 0$, $-D_2^2 f(x) > 0$ and $\mathrm{spec}(-J(x^*)) \subset \mathbb{C}_-^\circ$ and $\mathtt{S}_1(J(x)) > 0$.

Example: Non-Nash Attractors are Stackelberg. Consider the zero-sum game defined by

$$f(x) = -e^{-0.01(x_1^2 + x_2^2)}((ax_1^2 + x_2)^2 + (bx_2^2 + x_1)^2). \quad (3)$$

Let player 1 be the leader who aims to minimize f with respect to x_1 taking into consideration that player 2 (follower) aims to minimize $-f$ with respect to x_2. In Fig. 1, we show the trajectories for different initializations for this game; it can be seen that simultaneous gradient descent can lead to stable critical points which are DSE and not DNE. In fact, it is the case that all stable critical points with $-D_2^2 f(x) > 0$ are DSE in games on $\mathbb{R}^2$ (see Corollary E.1, Appendix E).

This example, along with Propositions E.1 and E.2 in Appendix E, implies some stable fixed points of $\dot{x} = -\omega(x)$ which are not DNE are in fact DSE. This is a meaningful result since recent works have proposed schemes to avoid stable critical points which are not DNE as they have been thought to lack game-theoretic meaning (Adolphs et al., 2019; Mazumdar et al., 2019). Moreover, some recent empirical studies show a number of successful approaches to training GANs do not converge to DNE, but rather to stable fixed points of the dynamics at which the follower is at a local optimum (Berard et al., 2020). This may suggest reaching DSE is desirable in GANs.

The 'realizable' assumption in the GAN literature says the discriminator network is zero near an equilibrium parameter configuration (Nagarajan & Kolter, 2017). The assumption implies the Jacobian of $\dot{x} = -\omega(x)$ is such that $D_1^2 f(x) = 0$. Under this assumption, we show stable critical points

which are not DNE are DSE given $-D_2^2 f(x) > 0$.

Proposition 4. *Consider a zero-sum GAN satisfying the realizable assumption. Any stable critical point of $\dot{x} = -\omega(x)$ at which $-D_2^2 f(x) > 0$ is a* DSE *and a stable critical point of $\dot{x} = -\omega_{\mathcal{S}}(x)$.*

4. Convergence Analysis

In this section, we provide convergence guarantees for both the deterministic and stochastic settings. In the former, players have oracle access to their gradients at each step while in the latter, players are assumed to have an unbiased estimator of the gradient appearing in their update rule. Proofs of the deterministic results can be found in Appendix F and the stochastic results in Appendix G.

4.1. Deterministic Setting

Consider the deterministic Stackelberg update

$$x_{k+1} = x_k - \gamma_2 \omega_{\mathcal{S}_\tau}(x_k)$$

where $\omega_{\mathcal{S}_\tau}(x_k)$ is the m-dimensional vector with entries $\tau^{-1}(D_1 f_1(x_k) - D_{21}^\top f_2(x_k)(D_2^2 f_2(x_k))^{-1} D_2 f_1(x_k)) \in \mathbb{R}^{m_1}$ and $D_2 f_2(x_k) \in \mathbb{R}^{m_2}$, and $\tau = \gamma_2/\gamma_1$ is the "timescale" separation with $\gamma_2 > \gamma_1$. We refer to these dynamics as the τ-Stackelberg update.

To get convergence guarantees, we apply well known results from discrete time dynamical systems. For a dynamical system $x_{k+1} = F(x_k)$, when the spectral radius $\rho(DF(x^*))$ of the Jacobian at fixed point is less than one, F is a contraction at x^* so that x^* is locally asymptotically stable (see Prop. F.1, Appendix F). In particular, $\rho(DF(x^*)) \leq c < 1$ implies that $\|DF\| \leq c + \varepsilon < 1$ for $\varepsilon > 0$ on a neighborhood of x^* (Ortega & Rheinboldt, 1970, 2.2.8). Hence, Prop. F.1 implies that if $\rho(DF(x^*)) = 1 - \alpha < 1$ for some α, then there exists a ball $B_p(x^*)$ of radius $p > 0$ such that for any $x_0 \in B_p(x^*)$, and some constant $K > 0$, $\|x_k - x^*\|_2 \leq K(1 - \alpha/2)^k \|x_0 - x^*\|_2$ using $\varepsilon = \alpha/4$.

For a zero-sum setting defined by cost function $f \in C^q(X, \mathbb{R})$ with $q \geq 2$, recall that $\mathtt{S}_1(J(x)) = D_1^2 f(x) - D_{21} f(x)^\top (D_2^2 f(x))^{-1} D_{21} f(x)$ is the first Schur complement of the Jacobian $J(x)$. Let $H(x)$ be the Hessian of the function $f : X \to \mathbb{R}$.

Theorem 3 (Zero-Sum Rate of Convergence.)**.** *Consider a zero-sum game defined by $f \in C^q(X, \mathbb{R})$ with $q \geq 2$. For a* DSE *x^* with $\alpha = \min\{\lambda_{\min}(-D_2^2 f(x^*)), \lambda_{\min}(\frac{1}{\tau}\mathtt{S}_1(J(x^*)))\}$ and $\beta = \rho(H(x^*))$, the τ–Stackelberg update converges locally with a rate of $O((1 - \alpha^2/(2\beta^2))^k)$.*

Corollary 2 (Zero-Sum Finite Time Guarantee)**.** *Given $\varepsilon > 0$, under the assumptions of Theorem 3, τ-Stackelberg learning obtains an ε-*DSE *in $\lceil \frac{2\beta^2}{\alpha^2} \log(\|x_0 - x^*\|^2/\varepsilon) \rceil$ iterations for any $x_0 \in B_\delta(x^*)$ with $\delta = \sqrt{\alpha^2/(2L\beta^2)}$ where L is the local Lipschitz constant of $I - \gamma_2 J_{\mathcal{S}_\tau}(x^*)$.*

The proofs leverage the structure of the Jacobian $J_{\mathcal{S}_\tau}$, which is lower block diagonal, along with the above noted result from dynamical systems theory. The key insight is that at a given x, the spectrum of $J_{\mathcal{S}_\tau}(x)$ is the union of the spectrum of $\tau^{-1}\mathtt{S}_1(J(x))$ and $-D_2^2 f(x)$ for zero-sum settings.

We now show a discrete-time analogue to Proposition 2.

Proposition 5. *Consider a zero-sum game defined by $f \in C^q(X, \mathbb{R})$, $q \geq 2$. Suppose that $\gamma_2 < 1/L$ where $\max\{\mathrm{spec}(\frac{1}{\tau}\mathtt{S}_1(J(x))) \cup \mathrm{spec}(-D_2^2 f(x))\} \leq L$. Then, x is a stable critical point of τ–Stackelberg update if and only if x is a* DSE*.*

The next result shows that τ-Stackelberg avoids saddle points almost surely in general-sum games. We remark that DSE are never saddle points in zero-sum games.

Theorem 4 (Almost Sure Avoidance of Saddles)**.** *Consider a general sum game defined by $f_i \in C^q(X, \mathbb{R})$, $q \geq 2$ for $i = 1, 2$ and where, without loss of generality, player 1 is the leader. Suppose that $\omega_{\mathcal{S}_\tau}$ is L-Lipschitz with $\tau > 1$ and that $\gamma_2 < 1/L$. The τ–Stackelberg learning dynamics converge to saddle points of $\dot{x} = -\omega_{\mathcal{S}_\tau}(x)$ on a set of measure zero.*

In the zero-sum setting, $\omega_{\mathcal{S}_\tau}$ being Lipschitz is equivalent to $\max\{\mathrm{spec}(\frac{1}{\tau}\mathtt{S}_1(J(x))) \cup \mathrm{spec}(-D_2^2 f(x))\} \leq L$. In this case, using the structure of the Jacobian $J_{\mathcal{S}_\tau}$, we know that the eigenvalues are real, and hence the only admissible types of critical points are stable, unstable, or saddle points. Consequently, the previous pair of results imply that the only critical points τ-Stackelberg learning converges to in zero-sum games are DSE almost surely.

We now provide a convergence guarantee for deterministic general-sum games. However, the convergence guarantee is no longer a global guarantee to the set of attractors of which critical points are DSE since there is potentially stable critical points which are not DSE. This can be seen by examining the Jacobian which is no longer lower block triangular.

Given a critical point x^*, let $\alpha = \lambda_{\min}^2(\frac{1}{2}(J_{\mathcal{S}_\tau}^\top(x^*) + J_{\mathcal{S}_\tau}(x^*)))$ and $\beta = \lambda_{\max}(J_{\mathcal{S}_\tau}(x^*)^\top J_{\mathcal{S}_\tau}(x^*))$.

Theorem 5 (General Sum Rate of Convergence)**.** *Consider a general sum game (f_1, f_2) with $f_i \in C^q(X, \mathbb{R})$, $q \geq 2$ for $i = 1, 2$ and where, without loss of generality, player 1 is the leader. For a* DSE *x^* such that $J_{\mathcal{S}_\tau}^\top(x^*) + J_{\mathcal{S}_\tau}(x^*) > 0$, the τ–Stackelberg update converges locally with a rate of $O((1 - \frac{\alpha}{2\beta})^{k/2})$.*

Corollary 3 (General Sum Finite Time Guarantee)**.** *Given $\varepsilon > 0$, under the assumptions of Theorem 5, τ–Stackelberg learning obtains an ε–*DSE *in $\lceil \frac{2\beta}{\alpha} \log(\|x_0 - x^*\|/\varepsilon) \rceil$ iterations for any $x_0 \in B_\delta(x^*)$ with $\delta = \alpha/(2L\beta)$ where L is the local Lipschitz constant of $I - \gamma_2 J_{\mathcal{S}_\tau}(x)$.*

4.2. Stochastic Setting

In the stochastic setting, players use updates of the form

$$x_{i,k+1} = x_{i,k} - \gamma_{i,k}(\omega_{\mathcal{S},i}(x_k) + w_{i,k+1}) \tag{4}$$

where $\gamma_{1,k} = o(\gamma_{2,k})$ and $\{w_{i,k+1}\}$ is a stochastic process for each $i = 1, 2$. The results in this section assume the following. The maps $Df_1 : \mathbb{R}^m \to \mathbb{R}^{m_1}$, $D_2 f_2 : \mathbb{R}^m \to \mathbb{R}^{m_2}$ are Lipschitz, and $\|Df_1\| < \infty$. For each $i \in \{1, 2\}$, the learning rates satisfy $\sum_k \gamma_{i,k} = \infty$, $\sum_k \gamma_{i,k}^2 < \infty$. The noise processes $\{w_{i,k}\}$ are zero mean, martingale difference sequences. That is, given the filtration $\mathcal{F}_k = \sigma(x_s, w_{1,s}, w_{2,s},\ s \le k)$, $\{w_{i,k}\}_{i\in\mathcal{I}}$ are conditionally independent, $\mathbb{E}[w_{i,k+1}|\ \mathcal{F}_k] = 0$ a.s., and $\mathbb{E}[\|w_{i,k+1}\|\ \mathcal{F}_k] \le c_i(1 + \|x_k\|)$ a.s. for some constants $c_i \ge 0$, $i \in \mathcal{I}$.

The primary technical machinery we use in this section is stochastic approximation theory (Borkar, 2008) and tools from dynamical systems. The convergence guarantees in this section are analogous to that for deterministic learning but asymptotic in nature. We first provide a non-convergence guarantee: the dynamics avoid saddle points in the stochastic learning regime.

Theorem 6 (Almost Sure Avoidance of Saddles.)**.** *Consider a game (f_1, f_2) with $f_i \in C^q(\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}, \mathbb{R})$, $q \ge 2$ for $i = 1, 2$ and where without loss of generality, player 1 is the leader. Suppose that for each $i = 1, 2$, there exists a constant $b_i > 0$ such that $\mathbb{E}[(w_{i,t} \cdot v)^+ | \mathcal{F}_{i,t}] \ge b_i$ for every unit vector $v \in \mathbb{R}^{m_i}$. Then, Stackelberg learning converges to strict saddle points of the game on a set of measure zero.*

We also give asymptotic convergence results. These results, combined with the non-convergence guarantee in Theorem 6, provide a broad convergence analysis for this class of learning dynamics. Theorem G.3 in Appendix G.3 provides a global convergence guarantee in general-sum games to the stable critical point, which may or may not be a DSE, under assumptions on the global asymptotic stability of critical points of the continuous time limiting singularly perturbed dynamical system. In zero-sum games, we know that the only critical points of the continuous time limiting system are DSE. Hence, Corollary G.2 in Appendix G.3 gives a global convergence guarantee in zero-sum games to the DSE under identical assumptions.

Relaxing these assumptions, the following proposition provides a local convergence result which ensures that sample points asymptotically converge to locally asymptotic trajectories of the continuous time limiting singularly perturbed system, and thus to stable DSE.

Theorem 7. *Consider a general sum game (f_1, f_2) with $f_i \in C^q(X, \mathbb{R})$, $q \ge 2$ for $i = 1, 2$ and where, without loss of generality, player 1 is the leader and $\gamma_{1,k} = o(\gamma_{2,k})$. Given a DSE x^*, let $B_p(x^*) = B_{p_1}(x_1^*) \times B_{p_2}(x_2^*)$ with $p_1, p_2 > 0$ on which $\det(D_2^2 f_2(x)) \neq 0$. Suppose $x_0 \in B_p(x^*)$. If $r(x_1) \in B_{p_2}(x_2^*)$ is a locally asymptotically stable critical point of $\dot{x}_2 = -D_2 f_2(x)$ uniformly in x_1 and the dynamics $\dot{x}_1 = -Df_1(x_1, r(x_1))$ have a locally asymptotically stable critical point in $B_{p_1}(x_1^*)$, then $x_k \to x^*$ almost surely.*

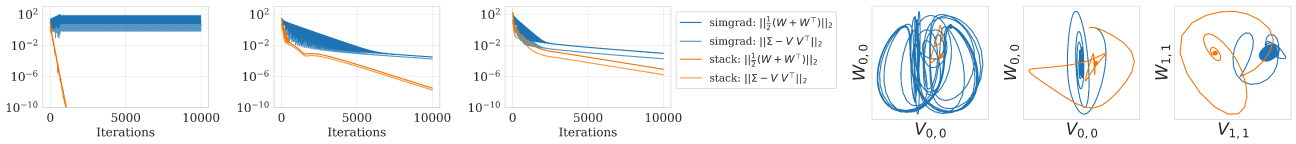
5. Experiments

We now present experiments showing the role of DSE in the optimization landscape of GANs and the empirical benefits of training GANs with Stackelberg learning compared to simultaneous gradient descent (simgrad). All detailed experiment information is given in Appendix H.

Example 1: Learning a Covariance Matrix. We consider a data generating process of $x \sim \mathcal{N}(0, \Sigma)$, where the covariance Σ is unknown and the objective is to learn it using a Wasserstein GAN. The discriminator is configured to be the set of quadratic functions defined as $D_W(x) = x^\top W x$ and the generator is a linear function of random input noise $z \sim \mathcal{N}(0, I)$ defined by $G_V(z) = Vz$. The matrices $W \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{m\times m}$ are the parameters of the discriminator and the generator, respectively. The Wasserstein GAN cost for the problem $f(V, W) = \sum_{i=1}^m \sum_{j=1}^m W_{ij}(\Sigma_{ij} - \sum_{k=1}^m V_{ik}V_{jk})$. We consider the generator to be the leader minimizing $f(V, W)$. The discriminator is the follower and it minimizes a regularized cost function defined by $-f(V, W) + \frac{\eta}{2}\operatorname{Tr}(W^\top W)$, where $\eta \ge 0$ is a tunable regularization parameter. The game is formally defined by the costs $(f_1, f_2) = (f(V, W), -f(V, W) + \frac{\eta}{2}\operatorname{Tr}(W^\top W))$, where player 1 is the leader and player 2 is the follower. In equilibrium, the generator picks V^* such that $V^*(V^*)^\top = \Sigma$ and the discriminator selects $W^* = 0$. Further details are given in Appendix C from Daskalakis et al. (2018).

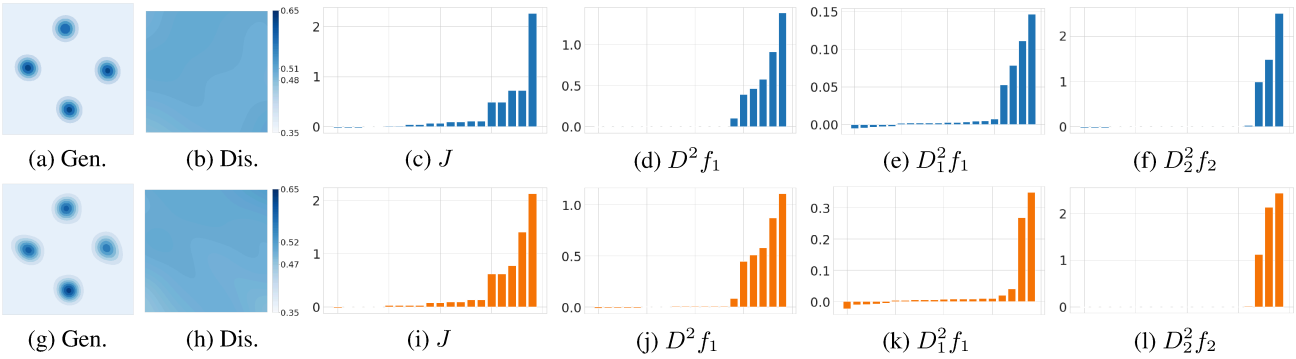
We compare the deterministic gradient update for Stackelberg learning with simultaneous learning, and analyze the distance from equilibrium as a function of time. We plot $\|\Sigma - VV^\top\|_2$ for the generator's performance and $\|\frac{1}{2}(W + W^\top)\|_2$ for the discriminator's performance in Fig. 2 for varying dimensions m with learning rates $\gamma_1 = \gamma_2/2 = 0.015$ and fixed regularization terms $\eta = m/5$. We observe that Stackelberg learning converges to an equilibrium in fewer iterations. For zero-sum games, our theory provides reasoning for this behavior since at any critical point the eigenvalues of the game Jacobian are purely real. This is in contrast to simultaneous gradient descent, whose Jacobian can admit complex eigenvalues, known to cause rotational forces in the dynamics. While there may be imaginary eigenvalues in the Jacobian of Stackelberg dynamics in general-sum games, this example demonstrates empirically that the Stackelberg dynamics mitigate rotations.

GAN training details. We now train GANs in which each



(a) $m = 3$ (b) $m = 9$ (c) $m = 25$ (d) $m = 3$ (e) $m = 9$ (f) $m = 25$

Figure 2. Stackelberg learning more effectively estimates covariance Σ as compared to `simgrad`. The generator (leader) and discriminator (follower) optimize for $V, W \in \mathbb{R}^{m \times m}$, respectively. Using learning rates $\gamma_1 = \gamma_2/2 = 0.015$ and regularization $\eta = m/5$, errors are shown in (a)–(c). Observe the cycling of `simgrad` (blue) in (d)–(f) for learning trajectory plots of elements of W and V.



(a) Gen. (b) Dis. (c) J (d) $D^2 f_1$ (e) $D_1^2 f_1$ (f) $D_2^2 f_2$

(g) Gen. (h) Dis. (i) J (j) $D^2 f_1$ (k) $D_1^2 f_1$ (l) $D_2^2 f_2$

Figure 3. The generator and discriminator performances for `simgrad` and Stackelberg are shown in (a)–(b) and (g)–(h), respectively. We show the 5 smallest and 15 largest real eigenvalues parts of relevant game objects in (c)–(f) for `simgrad` and (i)–(l) for Stackelberg.

player is parameterized by a neural network. The generator is always taken to be the leader and the discriminator the follower in this set of experiments. Moreover, for both Stackelberg learning and simultaneous gradient descent we pass the gradient information of each player into the Adam optimizer (Kingma & Ba, 2015). To ensure the follower's Hessian is well-conditioned in the leader update, we regularize the implicit map of the follower so that the leader gradient is given by $\omega_{S,1} = D_1 f_1(x) + Dr_\eta(x)^\top D_2 f_1(x)$ where $Dr_\eta(x)^\top = -D_{21} f_2(x)^\top \left(D_2^2 f_2(x) + \eta I\right)^{-1}$ and η is the regularization parameter. We also employ regularization in the follower's implicit map when computing eigenvalues of $D^2 f_1(x)$ to determine whether an approximate critical point is in a neighborhood of a `DSE`. We provide details on the derivation of the regularized leader update along with a notion of a regularized `DSE` and specifics on the eigenvalue computation in Appendix H.1 and H.2.

Example 2: Learning a Mixture of Gaussians. We train a GAN to learn a mixture of Gaussian distribution. The generator and discriminator networks have two and one hidden layers, respectively; each hidden layer has 32 neurons. We train using a batch size of 256, a latent dimension of 16, with decaying learning rates. For both the diamond and circle configurations, 10 initial seeds were simulated for each set of learning dynamics and behavior was generally consistent across them for both algorithms. The experiments were run for 60,000 batches and the eigenvalues evaluated at that stopping point. We show detailed information for the best run of each algorithm in terms of KL-divergence and in Appendix H.4.1 examine all runs.

Diamond configuration. This experiment uses the saturating GAN objective and Tanh activations. In Fig. 3a–3b and Fig. 3g–3h we show a sample of the generator and the discriminator for `simgrad` and the Stackelberg dynamics at the end of training. Each learning rule converges so that the generator can create a distribution that is close to the ground truth and the discriminator is nearly at the optimal probability throughout the input space. In Fig. 3c–3f and Fig. 3i–3l, we show eigenvalues from the game that present a deeper view of the convergence behavior. We observe from the eigenvalues of J that both sets of dynamics converge to neighborhoods of points that are stable for the simultaneous dynamics and they appear to be in a neighborhood of a `DSE` since the eigenvalues of $D^2 f_1$ and $D_2^2 f_2$ are nearly all positive. Interestingly, however, since the eigenvalues of $D_1^2 f_1$ are nearly all zero and not all positive and this was consistent across the runs, it appears that the result may reflect the realizable assumption (cf. Sec. 3) as well as convergence to a `DSE` that is not a `DNE`. Given the good generator and discriminator performance, it is worth further empirical investigation to determine if `DSE` that are not `DNE` are desirable in GANs and if successful methods reach them.

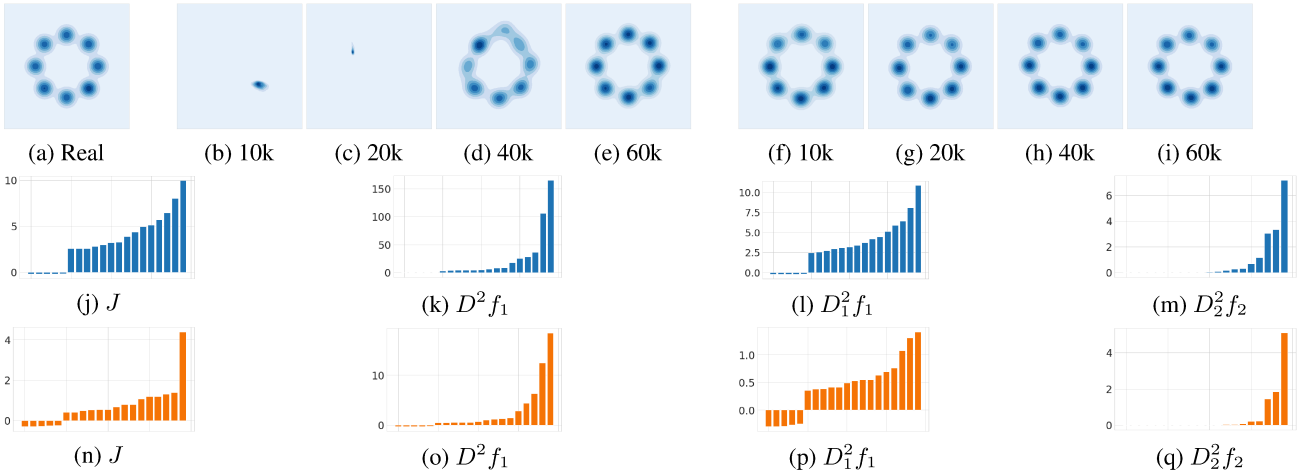


Figure 4. Stackelberg learning improves learning stability: `simgrad` generator in (b)–(e) and Stackelberg learning generator in (f)–(i). We show the 5 smallest and 15 largest real eigenvalue parts of relevant game objects in (j)–(m) for `simgrad` and (n)–(q) for Stackelberg.

(a) Real images. (b) Fake images. (c) Inception scores.

Figure 5. Stackelberg learning on the MNIST dataset.

Circle configuration. We demonstrate improved performance and stability when using Stackelberg learning dynamics in this example. We use ReLU activation functions and the non-saturating objective and show the performance in Fig. 4 along the learning path for the `simgrad` and Stackelberg learning dynamics. The former cycles and performs poorly until the learning rates have decayed enough to stabilize the training process. The latter converges quickly to a solution that nearly matches the ground truth distribution. We observed this behavior consistently across the runs. In a similar fashion as in the covariance example, the leader update is able to reduce rotations. We show the eigenvalues after training and see that for this configuration, `simgrad` converges to a neighborhood of a `DNE` and the Stackelberg dynamics converge again to the neighborhood of a `DSE` that is not a `DNE`. This provides further evidence that `DSE` may be easier to reach, and can provide suitable performance.

Example 3: MNIST GAN. To demonstrate that the Stackelberg learning dynamics can scale to high dimensional problems, we train a GAN on the MNIST dataset using the DCGAN architecture (Radford et al., 2015) adapted to handle 28×28 images. We simulate 10 random seeds and in Fig. 5c show the mean Inception score along the training process along with the standard error of the mean. The Inception score is calculated using a LeNet classifier following (Berard et al., 2020). We show a real sample in Fig. 5a and a fake sample in Fig. 5a after 7500 batches from the run with the fifth highest inception score. The Stackelberg learning dynamics are able to converge to a solution that generates realistic handwritten digits and get close to the maximum inception score in a stable manner. The primary purpose of this example is to show that the learning dynamics including second order information and an inverse is not an insurmountable problem for training with millions of parameters. We detail how the update can be computed efficiently using Jacobian-vector products and the conjugate gradient algorithm in Appendix H.2.

6. Conclusion

We study learning dynamics in Stackelberg games. This class of games pertains to any application in which there is an order of play. However, the problem has not been extensively analyzed in the way the learning dynamics of simultaneous play games have been. Consequently, we are able to give novel convergence results and draw connections to existing work focused on learning Nash equilibria.

Acknowledgements

This work was supported by NSF CAREER award 1844729 and a DoD NDSEG Fellowship.

References

Abraham, R., Marsden, J. E., and Ratiu, T. *Manifolds, Tensor Analysis, and Applications*. Springer, 2nd edition, 1988. doi: 10.1007/978-1-4612-1029-0.

Adolphs, L., Daneshmand, H., Lucchi, A., and Hofmann, T. Local saddle point optimization: A curvature exploitation approach. In *International Conference on Artificial Intelligence and Statistics*, pp. 486–495, 2019.

Argyros, I. K. A generalization of ostrowski's theorem on fixed points. *Applied Mathematics Letters*, 12:77–79, 1999.

Balduzzi, D., Racaniere, S., Martens, J., Foerster, J., Tuyls, K., and Graepel, T. The mechanics of n-player differentiable games. In *International Conference on Machine Learning*, pp. 354–363, 2018.

Basar, T. and Olsder, G. J. *Dynamic Noncooperative Game Theory*. Society for Industrial and Applied Mathematics, 2nd edition, 1998.

Benaïm, M. Dynamics of stochastic approximation algorithms. In *Seminaire de Probabilites XXXIII*, pp. 1–68, 1999.

Berard, H., Gidel, G., Almahairi, A., Vincent, P., and Lacoste-Julien, S. A closer look at the optimization landscapes of generative adversarial networks. In *International Conference on Learning Representations*, 2020.

Berger, T., Giribet, J., Pería, F. M., and Trunk, C. On a Class of Non-Hermitian Matrices with Positive Definite Schur Complements. *arXiv preprint arxiv:1807.08591*, 2018.

Bhatnagar, S., Prasad, H., and Prashanth, L. *Stochastic recursive algorithms for optimization: simultaneous perturbation methods*, volume 434. Springer, 2012.

Borkar, V. S. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, 2008.

Borkar, V. S. and Pattathil, S. Concentration bounds for two time scale stochastic approximation. In *Allerton Conference on Communication, Control, and Computing*, pp. 504–511. IEEE, 2018.

Broer, H. and Takens, F. Chapter 1 - preliminaries of dynamical systems theory. In Henk Broer, F. T. and Hasselblatt, B. (eds.), *Handbook of Dynamical Systems*, volume 3 of *Handbook of Dynamical Systems*, pp. 1 – 42. Elsevier Science, 2010.

Dai, B., Shaw, A., Li, L., Xiao, L., He, N., Liu, Z., Chen, J., and Song, L. Sbeed: Convergent reinforcement learning with nonlinear function approximation. In *International Conference on Machine Learning*, pp. 1125–1134, 2018.

Daneshmand, H., Kohler, J., Lucchi, A., and Hofmann, T. Escaping saddles with stochastic gradients. In *International Conference on Machine Learning*, pp. 1155–1164, 2018.

Danskin, J. M. The theory of max-min, with applications. *SIAM Journal on Applied Mathematics*, 14(4):641–664, 1966.

Danskin, J. M. *The Theory of Max-Min and its Application to Weapons Allocation Problems*. Springer, 1967.

Daskalakis, C. and Panageas, I. The limit points of (optimistic) gradient descent in min-max optimization. In *Advances in Neural Information Processing Systems*, pp. 9236–9246, 2018.

Daskalakis, C., Ilyas, A., Syrgkanis, V., and Zeng, H. Training gans with optimism. In *International Conference on Learning Representations*, 2018.

Foerster, J., Chen, R. Y., Al-Shedivat, M., Whiteson, S., Abbeel, P., and Mordatch, I. Learning with opponent-learning awareness. In *International Conference on Autonomous Agents and MultiAgent Systems*, pp. 122–130, 2018.

Fudenberg, D., Drew, F., Levine, D. K., and Levine, D. K. *The theory of learning in games*, volume 2. MIT press, 1998.

Gibson, C. G., Wirthmüller, K., du Plessis, A. A., and Looijenga, E. J. N. Topological stability of smooth mappings. In *Lecture Notes in Mathematics*, volume 552. Springer-Verlag, 1976.

Golubitsky, M. and Guillemin, V. *Stable Mappings and Their Singularities*. Springer-Verlag, 1973.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.

Hirsch, M. W. *Differential topology*. Springer New York, 1976.

Hirsch, M. W. *Differential Topology*, volume 33. Springer Science & Business Media, 2012.

Horn, R. and Johnson, C. *Topics in Matrix Analysis*. Cambridge University Press, 2011.

Jin, C., Netrapalli, P., and Jordan, M. I. What is local optimality in nonconvex-nonconcave minimax optimization? *arXiv preprint arXiv:1902.00618*, 2019.

Kelley, J. *General Topology*. Van Nostrand Reinhold Company, 1955.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

Knutson, A. and Tao, T. Honeycombs and sums of hermitian matrices. *Notices of the American Mathematical Society*, 2001.

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

Lee, J. *Introduction to smooth manifolds*. Springer, 2012.

Lee, J. D., Simchowitz, M., Jordan, M. I., and Recht, B. Gradient descent only converges to minimizers. In *Conference on Learning Theory*, pp. 1246–1257, 2016.

Letcher, A., Foerster, J., Balduzzi, D., Rocktäschel, T., and Whiteson, S. Stable opponent shaping in differentiable games. In *International Conference on Learning Representations*, 2019.

Maclaurin, D., Duvenaud, D., and Adams, R. Gradient-based hyperparameter optimization through reversible learning. In *International Conference on Machine Learning*, pp. 2113–2122, 2015.

Madry, A., Makelov, A., Schmidt, L., Tsipras, D., and Vladu, A. Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations*, 2018.

Mazumdar, E. and Ratliff, L. J. Local nash equilibria are isolated, strict local nash equilibria in 'almost all' zero-sum continuous games. In *IEEE Conference on Decision and Control*, 2019.

Mazumdar, E., Jordan, M., and Sastry, S. S. On finding local nash equilibria (and only local nash equilibria) in zero-sum games. *arxiv:1901.00838*, 2019.

Mazumdar, E., Ratliff, L. J., and Sastry, S. S. On gradient-based learning in continuous games. *SIAM Journal on Mathematics of Data Science*, 2(1):103–131, 2020.

Mertikopoulos, P., Lecouat, B., Zenati, H., Foo, C.-S., Chandrasekhar, V., and Piliouras, G. Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile. In *International Conference on Learning Representations*, 2019.

Mescheder, L., Nowozin, S., and Geiger, A. The numerics of gans. In *Advances in Neural Information Processing Systems*, pp. 1825–1835, 2017.

Metz, L., Poole, B., Pfau, D., and Sohl-Dickstein, J. Unrolled generative adversarial networks. In *International Conference on Learning Representations*, 2017.

Nagarajan, V. and Kolter, J. Z. Gradient descent gan optimization is locally stable. In *Advances in Neural Information Processing Systems*, pp. 5585–5595, 2017.

Ortega, J. M. and Rheinboldt, W. C. *Iterative Solutions to Nonlinear Equations in Several Variables*. Academic Press, 1970.

Panageas, I. and Piliouras, G. Gradient Descent Only Converges to Minimizers: Non-Isolated Critical Points and Invariant Regions. In *Innovations in Theoretical Computer Science*, 2017.

Pemantle, R. Nonconvergence to unstable points in urn models and stochastic approximations. *Annuals Probability*, 18(2):698–712, 04 1990.

Radford, A., Metz, L., and Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

Ratliff, L. J., Burden, S. A., and Sastry, S. S. Genericity and structural stability of non-degenerate differential nash equilibria. In *American Control Conference*, pp. 3990–3995. IEEE, 2014.

Ratliff, L. J., Burden, S. A., and Sastry, S. S. On the Characterization of Local Nash Equilibria in Continuous Games. *IEEE Transactions on Automatic Control*, 61(8):2301–2307, 2016.

Sastry, S. S. *Nonlinear Systems Theory*. Springer, 1999.

Schäfer, F. and Anandkumar, A. Competitive gradient descent. In *Advances in Neural Information Processing Systems*, pp. 7623–7633, 2019.

Shub, M. *Global Stability of Dynamical Systems*. Springer-Verlag, 1978.

Thoppe, G. and Borkar, V. A concentration bound for stochastic approximation via alekseev's formula. *Stochastic Systems*, 9(1):1–26, 2019.

Tretter, C. *Spectral Theory of Block Operator Matrices and Applications*. Imperial College Press, 2008.

Von Stackelberg, H. *Market structure and equilibrium*. Springer Science & Business Media, 2010.

Zaslavski, A. J. Necessary optimality conditions for bilevel minimization problems. *Nonlinear Analysis: Theory, Methods & Applications*, 75(3):1655–1678, 2012.

Zhang, C. and Lesser, V. Multi-agent learning with policy prediction. In *AAAI Conference on Artificial Intelligence*, 2010.

Zhang, K., Yang, Z., and Başar, T. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *arXiv preprint arXiv:1911.10635*, 2019.