

Gradient Descent-Ascent Provably Converges to Strict Local Minmax Equilibria with a Finite Timescale Separation

Tanner Fiez, Lillian Ratliff

University of Washington
fiez@uw.com, ratliff@uw.edu

Abstract

We study the role that a finite timescale separation parameter τ has on gradient descent-ascent in non-convex, non-concave zero-sum games where the learning rate of player 1 is denoted by γ_1 and the learning rate of player 2 is defined to be $\gamma_2 = \tau\gamma_1$. We show there exists a finite timescale separation parameter τ^* such that x^* is a stable critical point of gradient descent-ascent for all $\tau \in (\tau^*, \infty)$ if and only if it is a strict local minmax equilibrium. Moreover, we provide an explicit construction for computing τ^* along with corresponding convergence rates. The convergence results we present are complemented by a non-convergence result: given a critical point x^* that is not a strict local minmax equilibrium, there exists a finite timescale separation τ_0 such that x^* is unstable for all $\tau \in (\tau_0, \infty)$. Finally, we extend the results to gradient penalty regularization methods for generative adversarial networks and empirically demonstrate on CIFAR-10 and CelebA the significant impact timescale separation has on training performance.

1 Introduction

In this paper we study learning in zero-sum games of the form

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} f(x_1, x_2)$$

where the objective function of the game f is assumed to be sufficiently smooth and potentially non-convex and non-concave in the strategy spaces X_1 and X_2 respectively with each X_i a precompact subset of \mathbb{R}^{n_i} . This general problem formulation has long been fundamental in game theory (Başar and Olsder 1998) and recently it has become central to machine learning with applications in generative adversarial networks (Goodfellow et al. 2014), robust supervised learning (Madry et al. 2018; Sinha, Namkoong, and Duchi 2018), reinforcement and multi-agent reinforcement learning (Rajeswaran, Mordatch, and Kumar 2020; Zhang, Yang, and Başar 2019), imitation learning (Ho and Ermon 2016), constrained optimization (Cherukuri, Gharesifard, and Cortes 2017), and hyperparameter optimization (MacKay et al. 2019; Lorraine, Vicol, and Duvenaud 2020).

The gradient descent-ascent learning dynamics are widely studied as a potential method for efficiently computing

equilibria in such problems. However, in non-convex, non-concave zero-sum games, a number of past works highlight issues of convergence to critical points devoid of game theoretic meaning, where common notions of ‘meaningful’ equilibria include the local Nash and local minmax/Stackelberg concepts.¹ Indeed, it has been shown gradient descent-ascent with a shared learning rate is prone to reaching critical points that are neither a differential Nash equilibrium nor a differential Stackelberg equilibrium (Daskalakis and Panageas 2018; Mazumdar, Ratliff, and Sastry 2020; Jin, Netrapalli, and Jordan 2020). While an important negative result, it does not rule out the prospect that gradient descent-ascent may be able to guarantee equilibrium convergence as it fails to account for a key structural parameter of the dynamics, namely the ratio of learning rates between the players.

Motivated by the observation that the order of play between players is fundamental to the definition of the game, the role of *timescale separation* in gradient descent-ascent has recently been explored theoretically (Heusel et al. 2017; Chasnov et al. 2019; Jin, Netrapalli, and Jordan 2020). On the empirical side, it has been widely demonstrated that timescale separation in gradient descent-ascent is crucial to improving the solution quality when training generative adversarial networks (Goodfellow et al. 2014; Arjovsky, Chintala, and Bottou 2017; Heusel et al. 2017). Denoting γ_1 as the learning rate of the player 1, the learning rate of player 2 can be redefined as $\gamma_2 = \tau\gamma_1$ where $\tau = \gamma_2/\gamma_1 > 0$ is the learning rate ratio. Toward understanding the effect of timescale separation, Jin, Netrapalli, and Jordan (2020) show the stable critical points of gradient descent-ascent coincide with the set of differential Stackelberg equilibrium as $\tau \rightarrow \infty$. In other words, all ‘bad critical points’ (critical points lacking game-theoretic meaning) become unstable and all ‘good critical points’ (game-theoretically meaningful equilibria) remain or become stable as $\tau \rightarrow \infty$. While a promising theoretical development, it does not lead to a practical, implementable learning rule or necessarily provide an explanation for the satisfying performance in applications of gradient descent-ascent with a *finite* timescale separation. Importantly, it leaves open the problem of fully character-

¹Following past works, we refer to strict local Nash equilibrium and strict local minmax/Stackelberg equilibrium as differential Nash equilibrium and differential Stackelberg equilibrium from here on, respectively.

izing gradient descent-ascent as a function of the timescale separation.

Contributions. We show that gradient descent-ascent converges to a critical point for a range of finite learning rate ratios if and only if the critical point is a differential Stackelberg equilibrium. Furthermore, we show that all other critical points are unstable for a range of finite learning rate ratios. To our knowledge, this is the first guarantee of its kind for an implementable first-order method. Moreover, the technical results in this work rely on tools that have not appeared in the machine learning and optimization communities analyzing games. Finally, we extend these results to gradient penalty regularization methods in generative adversarial networks, thereby providing theoretical guarantees for a common combination of heuristics used in practice, and empirically demonstrate the benefits and trade-offs of regularization and timescale separation on image datasets.

This paper is a condensed version of Fiez and Ratliff (2020). The proofs of the results we present along with further experimental results and comparison to related work with discussion can be found in full form within Fiez and Ratliff (2020).

2 Preliminaries

A two-player zero-sum continuous game is defined by a collection of costs (f_1, f_2) where $f_1 \equiv f$ and $f_2 \equiv -f$ with $f \in C^r(X, \mathbb{R})$ for some $r \geq 2$ and where $X = X_1 \times X_2$ with each X_i a precompact subset of \mathbb{R}^{n_i} for $i \in \{1, 2\}$ and $n = n_1 + n_2$. Each player $i \in \mathcal{I}$ seeks to minimize their cost $f_i(x_i, x_{-i})$ with respect to their choice variable x_i where x_{-i} is the vector of all other actions x_j with $j \neq i$. 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 $D_i^2 f_i$ as the partial derivative of $D_i f_i$ with respect to x_i .

Equilibrium. There are natural equilibrium concepts depending on the order of play: the (local) Nash equilibrium concept in the case of simultaneous play and the (local) Stackelberg (equivalently minmax in zero-sum games) equilibrium concept in the case of hierarchical play (Başar and Olsder 1998). Formal local equilibrium definitions are provided by Fiez and Ratliff (2020) in Definitions 1 and 2, while here we characterize the different equilibrium notions in terms of sufficient conditions on player costs as is typical in the machine learning and optimization literature (see, e.g., Daskalakis and Panageas 2018; Mazumdar, Ratliff, and Sastry 2020; Jin, Netrapalli, and Jordan 2020; Goodfellow 2016; Fiez, Chasnov, and Ratliff 2020; Wang, Zhang, and Ba 2020; Berard et al. 2020).

The following definition is characterized by sufficient conditions for a local Nash equilibrium.

Definition 1 (Differential Nash Equilibrium, (Ratliff, Burden, and Sastry 2013)). *The joint strategy $x \in X$ is a differential Nash equilibrium if $D_1 f(x) = 0$, $-D_2 f(x) = 0$, $D_1^2 f(x) > 0$, and $D_2^2 f(x) < 0$.*

The Jacobian of the vector of individual gradients $g(x) = (D_1 f(x), -D_2 f(x))$ is defined by

$$J(x) = \begin{bmatrix} D_1^2 f(x) & D_{12} f(x) \\ -D_{12}^T f(x) & -D_2^2 f(x) \end{bmatrix}. \quad (1)$$

Let $S_1(\cdot)$ denote the Schur complement of (\cdot) with respect to the $n_2 \times n_2$ block in (\cdot) . The following definition is characterized by sufficient conditions for a local Stackelberg equilibrium.

Definition 2 (Differential Stackelberg Equilibrium (Fiez, Chasnov, and Ratliff 2020)). *The joint strategy $x \in X$ is a differential Stackelberg equilibrium if $D_1 f(x) = 0$, $-D_2 f(x) = 0$, $S_1(J(x)) > 0$, $D_2^2 f(x) < 0$.*

Learning Dynamics. We study agents seeking equilibria of the game via a learning algorithm and consider arguably the most natural learning rule in zero-sum continuous games: gradient descent-ascent (GDA). Moreover, we investigate this learning rule with *timescale separation* between the players. Let $\tau = \gamma_2/\gamma_1$ be the *learning rate ratio* and define $\Lambda_\tau = \text{blockdiag}(I_{n_1}, \tau I_{n_2})$ where I_{n_i} is a $n_i \times n_i$ identity matrix. The τ -GDA dynamics with $g(x) = (D_1 f(x), -D_2 f(x))$ are given by

$$x_{k+1} = x_k - \gamma_1 \Lambda_\tau g(x_k). \quad (2)$$

3 Stability of Continuous Time GDA with Timescale Separation

To characterize the convergence of τ -GDA, we begin by studying its continuous time limiting system

$$\dot{x} = -\Lambda_\tau g(x). \quad (3)$$

The Jacobian of the system from (3) is given by $J_\tau(x) = \Lambda_\tau J(x)$ where $J(x)$ is defined in (1). Observe that critical points (x such that $g(x) = 0$) are shared between τ -GDA and (3). Thus, by analyzing the stability of the continuous time system around critical points as a function of the timescale separation τ using the Jacobian $J_\tau(x)$, we can draw conclusions about the stability and convergence of the discrete time system τ -GDA. Recall that a critical point x^* is locally exponentially stable for $\dot{x} = -\Lambda_\tau g(x)$ if and only if $\text{spec}(-J_\tau(x^*)) \subset \mathbb{C}_-^\circ$ (Khalil 2002, Theorem 4.15), (or, equivalently, $\text{spec}(J_\tau(x^*)) \subset \mathbb{C}_+^\circ$) where \mathbb{C}_-° and \mathbb{C}_+° denote the open left-half and right-half complex plane, respectively. In what follows, we show that differential Stackelberg equilibria are the only critical points which are stable for a range of finite learning rate ratios,² whereas the remainder of critical points are unstable for a range of finite learning rate ratios.

3.1 Necessary and Sufficient Conditions for Stability

To motivate our main stability result, the following example shows the existence of a differential Stackelberg which is unstable for $\tau = 1$, but is stable all for $\tau \in (\tau^*, \infty)$ where τ^* is finite.

Example 1. *Consider the quadratic zero-sum game defined by the cost*

$$f(x_1, x_2) = \frac{v}{2}(-x_{11}^2 + \frac{1}{2}x_{12}^2 - 2x_{11}x_{21} - \frac{1}{2}x_{21}^2 + x_{12}x_{22} - x_{22}^2)$$

²Note that differential Nash are a subset of differential Stackelberg (Jin, Netrapalli, and Jordan 2020; Fiez, Chasnov, and Ratliff 2020).

where $v > 0$ and $x_1, x_2 \in \mathbb{R}^2$. The unique critical point $x^* = (0, 0)$ is a differential Stackelberg equilibrium since $g(x^*) = 0$, $S_1(J(x^*)) = \text{diag}(v, \frac{v}{4}) > 0$, and $D_2^2 f(x^*) = -\text{diag}(\frac{v}{2}, v) < 0$. Moreover, $\text{spec}(-J_\tau(x^*)) = \{\frac{-v}{4}(2\tau + 1 \pm \sqrt{4\tau^2 - 8\tau + 1}), \frac{-v}{4}(\tau - 2 \pm \sqrt{\tau^2 - 12\tau + 4})\}$. Observe that for any $v > 0$, x^* is unstable for $\tau = 1$ since $\text{spec}(-J_\tau(x^*)) \not\subset \mathbb{C}_-^\circ$, but x^* is stable for a range of learning rates since $\text{spec}(-J_\tau(x^*)) \subset \mathbb{C}_-^\circ$ for all $\tau \in (2, \infty)$.

In other words, GDA fails to converge to the equilibrium but a finite timescale separation is sufficient to remedy this problem. We now fully characterize this phenomenon. To provide some background, we remark it is known (see Appendix F of Fiez and Ratliff (2020) and Kokotovic, O'Reilly, and Khalil (1986, Chap. 2)) that the spectrum of $-J_\tau(x^*)$ asymptotically splits as $\tau \rightarrow \infty$ such that n_1 eigenvalues tend to fixed positions defined by the eigenvalues of $-S_1(J(x^*))$, while the remaining n_2 eigenvalues tend to infinity at a linear rate τ along asymptotes defined by the eigenvalues of $D_2^2 f(x^*)$. This fact directly results in the connection between critical points of ∞ -GDA and differential Stackelberg equilibrium. In contrast, we determine exactly the range of τ such that the spectrum of $-J_\tau(x)$ remains in \mathbb{C}_-° .

Theorem 1. Consider a zero-sum game $(f_1, f_2) = (f, -f)$ defined by $f \in C^r(X, \mathbb{R})$ for some $r \geq 2$. Suppose that x^* is such that $g(x^*) = 0$ and $S_1(J(x^*))$ and $D_2^2 f_2(x^*)$ are non-singular. There exists a $\tau^* \in (0, \infty)$ such that $\text{spec}(-J_\tau(x^*)) \subset \mathbb{C}_-^\circ$ for all $\tau \in (\tau^*, \infty)$ if and only if x^* is a differential Stackelberg equilibrium.

As a direct consequence of Theorem 1, τ -GDA converges locally asymptotically for any sufficiently small $\gamma(\tau)$ and for all $\tau \in (\tau^*, \infty)$ if and only if x^* is a differential Stackelberg equilibrium; for a formal statement, see Corollary 2 of Fiez and Ratliff (2020). To our knowledge, this is the only result showing τ -GDA converges to differential Stackelberg equilibria for a range of finite learning rate ratios. We now give an outline of the proof technique as it requires technical tools novel to this community.

Proof Sketch of Theorem 1. The full proof is contained in Appendix E of Fiez and Ratliff (2020). The key tools used in this proof are a combination of Lyapunov stability and the notion of a *guard map* (cf. Saydy, Tits, and Abed (1990)), a new tool to the learning community. Recall that a matrix is exponentially stable if and only if there exists a symmetric positive definite $P = P^\top > 0$ such that $PJ_\tau(x^*) + J_\tau^\top(x^*)P > 0$ (Khalil 2002, Thm 4.15). Hence, given a positive definite $Q = Q^\top > 0$, $-J_\tau(x^*)$ is stable if and only if there exists a unique solution $P = P^\top$ to

$$\begin{aligned} \text{vec}(Q) &= (J_\tau^\top(x^*) \oplus J_\tau^\top(x^*))\text{vec}(P) \\ &= ((J_\tau^\top(x^*) \otimes I) + (I \otimes J_\tau^\top(x^*)))\text{vec}(P) \end{aligned} \quad (4)$$

where \otimes and \oplus denote the Kronecker product and Kro-

necker sum, respectively.³ The existence of a unique solution P occurs if and only if J_τ^\top and $-J_\tau^\top$ have no eigenvalues in common. Hence, using the fact that eigenvalues vary continuously, if we vary τ and examine the eigenvalues of the map $J_\tau^\top(x^*) \oplus J_\tau^\top(x^*)$, this tells us the range of τ for which $\text{spec}(-J_\tau(x^*))$ remains in \mathbb{C}_-° . This method of varying parameters and determining when the roots of a polynomial (or correspondingly, the eigenvalues of a map) cross the boundary of a domain uses a *guard map*; it provides a certificate that the roots of a polynomial lie in a particular guarded domain for a range of parameter values.

Formally, let \mathcal{X} be the set of all $n \times n$ real matrices or the set of all polynomials of degree n with real coefficients. Consider \mathcal{S} an open subset of \mathcal{X} with closure $\bar{\mathcal{S}}$ and boundary $\partial\mathcal{S}$. The map $\nu : \mathcal{X} \rightarrow \mathbb{C}$ is said to be a guardian map for \mathcal{S} if for all $x \in \bar{\mathcal{S}}$, $\nu(x) = 0 \iff x \in \partial\mathcal{S}$. Elements of $\mathcal{S}(\mathbb{C}_-^\circ) = \{A \in \mathbb{R}^{n \times n} : \text{spec}(A) \subset \mathbb{C}_-^\circ\}$ are (Hurwitz) stable. Given a pathwise connected set $U \subseteq \mathbb{R}$, the family $\{A(\tau) : \tau \in U\}$ is stable if and only if (i) it is nominally stable—i.e., $A(\tau_0) \in \mathcal{S}(\mathbb{C}_-^\circ)$ for some $\tau_0 \in U$ —and (ii) $\nu(A(\tau)) \neq 0$ for all $\tau \in U$ (Saydy, Tits, and Abed 1990, Prop. 1). The map $\nu(\tau) = \det(2(-J_\tau(x^*) \odot I)) = \det(-J_\tau(x^*) \oplus J_\tau(x^*))$ guards $\mathcal{S}(\mathbb{C}_-^\circ)$ where \odot is the *bialternate product* and is defined by $A \odot B = \frac{1}{2}(A \oplus B)$ for matrices A and B (cf. Govaerts (2000, Sec. 4.4.4)). For intuition, consider the case where each $x_1, x_2 \in \mathbb{R}$ so that

$$J_\tau(x^*) = \begin{bmatrix} a & b \\ -\tau b & \tau d \end{bmatrix} \in \mathbb{R}^{2 \times 2}.$$

It is known that $\text{spec}(-J_\tau(x^*)) \subset \mathbb{C}_-^\circ$ if $\det(-J_\tau(x^*)) > 0$ and $\text{tr}(-J_\tau(x^*)) < 0$ so that $\nu(\tau) = \det(-J_\tau(x^*)) \text{tr}(-J_\tau(x^*))$ is a guard map for the 2×2 stable matrices $\mathcal{S}(\mathbb{C}_-^\circ)$. Since the bialternate product generalizes the trace operator and $\det(-J_\tau(x^*)) = \tau^{n_2} \det(D_2^2 f(x^*)) \det(-S_1(J(x^*))) \neq 0$ for $\tau \neq 0$ by the assumptions ($\det(S_1(J(x^*))) \neq 0$ and $\det(D_2^2 f(x^*)) \neq 0$), a guard map in the general $n \times n$ case is $\nu(\tau) = \det(-J_\tau(x^*) \oplus J_\tau(x^*))$.

This guard map in τ is closely related to the vectorization in (4): for any symmetric positive definite $Q = Q^\top > 0$, there will be a symmetric positive definite solution $P = P^\top > 0$ of $-(J_\tau^\top(x^*) \oplus J_\tau^\top(x^*))\text{vec}(P) = \text{vec}(-Q)$ if and only if $\det(-J_\tau(x^*) \oplus J_\tau(x^*)) \neq 0$. Hence, to find the range of τ for which, given any $Q = Q^\top > 0$, the solution $P = P^\top$ is no longer positive definite, we need to find the value of τ such that $\nu(\tau) = \det(-J_\tau(x^*) \oplus J_\tau(x^*)) = 0$ —that is, where it hits the boundary $\partial\mathcal{S}(\mathbb{C}_-^\circ)$. Through algebraic manipulation, this problem reduces to an eigenvalue problem in τ , giving rise to an explicit construction of τ^* . \square

3.2 Sufficient Conditions for Instability

To motivate our main instability result, the following example shows a non-equilibrium critical point that is stable for $\tau = 1$, but is unstable for all $\tau \in (\tau_0, \infty)$ where τ_0 is finite.

³See Magnus (1988); Lancaster and Tismenetsky (1985) for more detail on the definition and properties of these mathematical operators, and Appendix E of Fiez and Ratliff (2020) for more detail directly related to their use in this paper.

Example 2. Consider the quadratic zero-sum game defined by the cost

$$f(x_1, x_2) = \frac{v}{4}(x_{11}^2 - \frac{1}{2}x_{12}^2 + 2x_{11}x_{21} + \frac{1}{2}x_{21}^2 + 2x_{12}x_{22} - x_{22}^2)$$

where $x_1, x_2 \in \mathbb{R}^2$ and $v > 0$. The unique critical point $x^* = (0, 0)$ is not a differential Stackelberg (nor Nash) equilibrium since $D_1^2 f(x^*) = \text{diag}(v/2, -v/4) \not\preceq 0$, $D_2^2 f(x^*) = \text{diag}(v/4, -v/2) \not\preceq 0$. Moreover, $\text{spec}(-J_\tau(x^*)) = \{\frac{-v}{8}(2\tau - 1 \pm \sqrt{4\tau^2 - 12\tau + 1}), \frac{-v}{8}(2 - \tau \pm \sqrt{\tau^2 - 12\tau + 4})\}$. Observe that for any $v > 0$, x^* is stable for $\tau = 1$ since $\text{spec}(-J_\tau(x^*)) \subset \mathbb{C}_-^\circ$, but x^* is unstable for a range of learning rates since $\text{spec}(-J_\tau(x^*)) \not\subset \mathbb{C}_-^\circ$ for all $\tau \in (2, \infty)$. This is not an artifact of the quadratic example: games can be constructed in which stable critical points lacking game-theoretic meaning become unstable for all $\tau > \tau_0$ even in the presence of multiple equilibria.

This example demonstrates a finite timescale separation can prevent convergence to critical points lacking game-theoretic meaning. We now characterize this behavior generally. Note that Theorem 1 implies that for any critical point which is not a differential Stackelberg equilibrium, there is no finite τ^* such that $\text{spec}(-J_\tau(x^*)) \subset \mathbb{C}_-^\circ$ for all $\tau \in (\tau^*, \infty)$. In particular, there exists at least one finite, positive value of τ such that $\text{spec}(-J_\tau(x^*)) \not\subset \mathbb{C}_-^\circ$. We can extend this result to answer the question of whether there exists a finite learning rate ratio τ_0 such that $-J_\tau(x^*)$ has at least one eigenvalue with strictly positive real part for all $\tau \in (\tau_0, \infty)$, thereby implying that x^* is unstable.

Theorem 2. Consider a zero-sum game $(f_1, f_2) = (f, -f)$ defined by $f \in C^r(X, \mathbb{R})$ for some $r \geq 2$. Suppose that x^* is such that $g(x^*) = 0$ and it is not a differential Stackelberg equilibrium. There exists a finite learning rate ratio $\tau_0 \in (0, \infty)$ such that $\text{spec}(-J_\tau(x^*)) \not\subset \mathbb{C}_-^\circ$ for all $\tau \in (\tau_0, \infty)$.

Proof Sketch. The full proof is provided in Appendix G of Fiez and Ratliff (2020). The key idea is to leverage the Lyapunov equation and Lemma 5 of Fiez and Ratliff (2020) to show that $-J_\tau(x^*)$ has at least one eigenvalue with strictly positive real part. Indeed, Lemma 5 of Fiez and Ratliff (2020), which is from Lancaster and Tismenetsky (1985), states that if $S_1(-J(x^*))$ has no zero eigenvalues, then there exists matrices $P_1 = P_1^\top$ and $Q_1 = Q_1^\top > 0$ such that $P_1 S_1(-J(x^*)) + S_1(-J(x^*)) P_1 = Q_1$ where P_1 and $S_1(-J(x^*))$ have the same *inertia*—that is, the number of eigenvalues with positive, negative and zero real parts, respectively, are the same. An analogous statement applies to $-D_2^2 f(x^*)$ with some P_2 and Q_2 . Since x^* is a non-equilibrium critical point, without loss of generality, let $S_1(-J(x^*))$ have at least one strictly positive eigenvalue so that P_1 does as well. Next, we construct a matrix P that is congruent to $\text{blockdiag}(P_1, P_2)$ and a matrix Q_τ such that $-P J_\tau(x^*) - J_\tau^\top(x^*) P = Q_\tau$. Since P and $\text{blockdiag}(P_1, P_2)$ are congruent, Sylvester’s law of inertia implies that they have the same number of eigenvalues with positive, negative, and zero real parts, respectively. Hence, P has at least one eigenvalue with strictly positive

real part. We then construct τ_0 via an eigenvalue problem such that for all $\tau > \tau_0$, $Q_\tau > 0$. Applying Lemma 5 of Fiez and Ratliff (2020) again, for any $\tau > \tau_0$, $-J_\tau(x^*)$ has at least one eigenvalue with strictly positive real part so that $\text{spec}(-J_\tau(x^*)) \not\subset \mathbb{C}_-^\circ$. \square

Unlike τ^* in Theorem 1, τ_0 in Theorem 2 is not tight in the sense that $-J_\tau(x^*)$ may become unstable for $\tau < \tau_0$ since, e.g., there are potentially many matrices P_1 and Q_1 that satisfy $S_1(J(x^*)) P_1 + P_1 S_1(J(x^*)) = Q_1$ and $S_1(J(x^*))$ and P_1 have the same inertia (and analogously for P_2, Q_2). The choice of these matrices impact the value of τ_0 . Hence, the question of finding the exact value of τ beyond which a spurious critical point of GDA is unstable remains open. Nonetheless, no result has appeared previously showing that GDA with a finite timescale separation avoids such critical points.

3.3 Regularization with Applications to Adversarial Learning

In this section, we focus on generative adversarial networks with regularization and using the theory developed so far extend the results to provide a stability guarantee for a range of regularization parameters and learning rate ratios. Consider the training objective

$$f(\theta, \omega) = \mathbb{E}_{p(z)}[\ell(D(G(z; \theta); \omega))] + \mathbb{E}_{p_{\mathcal{D}}(x)}[\ell(-D(x; \omega))] \quad (5)$$

where $D_\omega(x)$ and $G_\theta(z)$ are discriminator and generator networks, $p_{\mathcal{D}}(x)$ is the data distribution while $p(z)$ is the latent distribution, and $\ell \in C^2(\mathbb{R})$ is some real-value function.⁴ Nagarajan and Kolter (2017) show, under suitable assumptions, that gradient-based methods for training generative adversarial networks are locally convergent assuming the data distributions are absolutely continuous. However, as observed by Mescheder, Geiger, and Nowozin (2018), such assumptions not only may not be satisfied by many practical generative adversarial network training scenarios such as natural images, but often the data distribution is concentrated on a lower dimensional manifold. The latter characteristic leads to highly ill-conditioned problems and nearly purely imaginary eigenvalues.

Gradient penalties ensure that the discriminator cannot create a non-zero gradient which is orthogonal to the data manifold without suffering a loss. Introduced by Roth et al. (2017) and refined in Mescheder, Geiger, and Nowozin (2018), we consider training generative adversarial networks with one of two fairly natural gradient-penalties used to regularize the discriminator:

$$R_1(\theta, \omega) = \frac{\mu}{2} \mathbb{E}_{p_{\mathcal{D}}(x)}[\|\nabla_x D(x; \omega)\|^2] \quad \text{and} \\ R_2(\theta, \omega) = \frac{\mu}{2} \mathbb{E}_{p_\theta(x)}[\|\nabla_x D(x; \omega)\|^2],$$

where, by a slight abuse of notation, $\nabla_x(\cdot)$ denotes the partial gradient with respect to x of the argument (\cdot) when the argument is the discriminator $D(\cdot; \omega)$ in order prevent any

⁴For example, $\ell(x) = -\log(1 + \exp(-x))$ gives the original formulation of Goodfellow et al. (2014).

conflation between the notation $D(\cdot)$ elsewhere for derivatives. Let $h_1(\theta) = \mathbb{E}_{p_{\theta}(x)}[\nabla_{\omega} D(x; \omega)|_{\omega=\omega^*}]$ and $h_2(\omega) = \mathbb{E}_{p_{\mathcal{D}}(x)}[|D(x; \omega)|^2 + \|\nabla_x D(x; \omega)\|^2]$. Define *reparameterization manifolds* $\mathcal{M}_{\mathcal{G}} = \{\theta : p_{\theta} = p_{\mathcal{D}}\}$ and $\mathcal{M}_{\mathcal{D}} = \{\omega : h_2(\omega) = 0\}$ and let $T_{\theta^*}\mathcal{M}_{\mathcal{G}}$ and $T_{\omega^*}\mathcal{M}_{\mathcal{D}}$ denote their respective tangent spaces at θ^* and ω^* . As in Mescheder, Geiger, and Nowozin (2018), we make the following assumption.

Assumption 1. Consider a zero-sum game of the form given in (5) where $f \in C^2(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, \mathbb{R})$ and $G(\cdot; \theta)$ and $D(\cdot; \omega)$ are the generator and discriminator networks, respectively, and $x = (\theta, \omega) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Suppose that $x^* = (\theta^*, \omega^*)$ is an equilibrium. Then, (a) at (θ^*, ω^*) , $p_{\theta^*} = p_{\mathcal{D}}$ and $D(x; \omega^*) = 0$ in some neighborhood of $\text{supp}(p_{\mathcal{D}})$, (b) the function $\ell \in C^2(\mathbb{R})$ satisfies $\ell'(0) \neq 0$ and $\ell''(0) < 0$, (c) there are ϵ -balls $B_{\epsilon}(\theta^*)$ and $B_{\epsilon}(\omega^*)$ centered around θ^* and ω^* , respectively, so that $\mathcal{M}_{\mathcal{G}} \cap B_{\epsilon}(\theta^*)$ and $\mathcal{M}_{\mathcal{D}} \cap B_{\epsilon}(\omega^*)$ define C^1 -manifolds. Moreover, (i) if $w \notin T_{\theta^*}\mathcal{M}_{\mathcal{G}}$, then $w^{\top} \nabla_w h_1(\theta^*) w \neq 0$, and (ii) if $v \notin T_{\omega^*}\mathcal{M}_{\mathcal{D}}$, then $v^{\top} \nabla_{\omega}^2 h_2(\omega^*) v \neq 0$.

We note that as explained by Mescheder, Geiger, and Nowozin (2018), Assumption 1.c(i) implies that the discriminator is capable of detecting deviations from the generator distribution in equilibrium, and Assumption 1.c(ii) implies that the manifold $\mathcal{M}_{\mathcal{D}}$ is sufficiently regular and, in particular, its (local) geometry is captured by the second (directional) derivative of h_2 . Proposition 5 of Fiez and Ratliff (2020) provides necessary conditions on the network parameter dimensions for Assumption 1 to hold. Under Assumption 1, we show that x^* is a differential Stackelberg equilibrium, and characterize the learning rate ratio and regularization parameter range for which x^* is (locally) stable with respect to τ -GDA.

Theorem 3. Consider training a generative adversarial network via a zero-sum game with generator network G_{θ} , discriminator network D_{ω} , and loss $f(\theta, \omega)$ with regularization $R_j(\theta, \omega)$ (for either $j = 1$ or $j = 2$) and any regularization parameter $\mu \in (0, \infty)$ such that Assumption 1 is satisfied for an equilibrium $x^* = (\theta^*, \omega^*)$ of the regularized dynamics. Then, $x^* = (\theta^*, \omega^*)$ is a differential Stackelberg equilibrium. Furthermore, for any $\tau \in (0, \infty)$, $\text{spec}(-J_{(\tau, \mu)}(x^*)) \subset \mathbb{C}_{-}^{\circ}$.

4 Provable Convergence of GDA with Timescale Separation

In this section, we characterize the asymptotic convergence rate for τ -GDA to differential Stackelberg equilibria, and provide a finite time guarantee for convergence to an ϵ -approximate equilibrium. The asymptotic convergence rate result uses Theorem 1 to construct a finite $\tau^* \in (0, \infty)$ such that x^* is stable, meaning $\text{spec}(-J_{\tau}(x^*)) \subset \mathbb{C}_{-}^{\circ}$, and then for any $\tau \in (\tau^*, \infty)$, Lemmas 1 and 2 from Fiez and Ratliff (2020) imply a local asymptotic convergence rate.

Theorem 4. Consider a zero-sum game $(f_1, f_2) = (f, -f)$ defined by $f \in C^r(X, \mathbb{R})$ for $r \geq 2$ and let x^* be a differential Stackelberg equilibrium of the game. There exists a $\tau^* \in (0, \infty)$ such that for any $\tau \in (\tau^*, \infty)$

and $\alpha \in (0, \gamma)$, τ -GDA with learning rate $\gamma_1 = \gamma - \alpha$ converges locally asymptotically at a rate of $O((1 - \alpha/(4\beta))^{k/2})$ where $\gamma = \min_{\lambda \in \text{spec}(J_{\tau}(x^*))} 2\text{Re}(\lambda)/|\lambda|^2$, $\lambda_m = \arg \min_{\lambda \in \text{spec}(J_{\tau}(x^*))} 2\text{Re}(\lambda)/|\lambda|^2$, and $\beta = (2\text{Re}(\lambda_m) - \alpha|\lambda_m|^2)^{-1}$. Moreover, if x^* is a differential Nash equilibrium, $\tau^* = 0$ so that for any $\tau \in (0, \infty)$ and $\alpha \in (0, \gamma)$, τ -GDA with $\gamma_1 = \gamma - \alpha$ converges with a rate $O((1 - \alpha/(4\beta))^{k/2})$.

To build some intuition, consider a differential Stackelberg equilibrium x^* and its corresponding τ^* obtained via Theorem 1 so that for any fixed $\tau \in (\tau^*, \infty)$, $\text{spec}(-J_{\tau}(x^*)) \subset \mathbb{C}_{-}^{\circ}$. For the discrete time system $x_{k+1} = x_k - \gamma_1 \Lambda_{\tau} g(x_k)$, if γ_1 is chosen such that the spectral radius of the local linearization of the discrete time map is a contraction, then x_k locally (exponentially) converges to x^* (cf. Proposition 6, Appendix A, Fiez and Ratliff (2020)). With this in mind, we formulate an optimization problem to find the upper bound γ on the learning rate γ_1 such that for all $\gamma_1 \in (0, \gamma)$, $\rho(I - \gamma_1 J_{\tau}(x^*)) < 1$; indeed, let $\gamma = \min_{\gamma > 0} \{\gamma : \max_{\lambda \in \text{spec}(J_{\tau}(x^*))} |1 - \gamma\lambda| \leq 1\}$. The intuition is as follows. The inner maximization problem is over a finite set $\text{spec}(J_{\tau}(x^*)) = \{\lambda_1, \dots, \lambda_n\}$ where $J_{\tau}(x^*) \in \mathbb{R}^{n \times n}$. As γ increases away from zero, each $|1 - \gamma\lambda_i|$ shrinks in magnitude. The last λ_i such that $1 - \gamma\lambda_i$ hits the boundary of the unit circle in the complex plane gives us the optimal γ and the $\lambda_m \in \text{spec}(J_{\tau}(x^*))$ that achieves it. Examining the constraint, we have that for each λ_i , $\gamma(\gamma|\lambda_i|^2 - 2\text{Re}(\lambda_i)) \leq 0$ for any $\gamma > 0$. As noted this constraint will be tight for one of the λ , in which case $\gamma = 2\text{Re}(\lambda)/|\lambda|^2$ since $\gamma > 0$. Hence, by selecting $\gamma = \min_{\lambda \in \text{spec}(J_{\tau}(x^*))} 2\text{Re}(\lambda)/|\lambda|^2$, we have that $|1 - \gamma_1\lambda| < 1$ for all $\lambda \in \text{spec}(J_{\tau}(x^*))$ and any $\gamma_1 \in (0, \gamma)$. From here, one can use standard arguments from numerical analysis to show that for the choice of α and β , the claimed asymptotic rate holds.

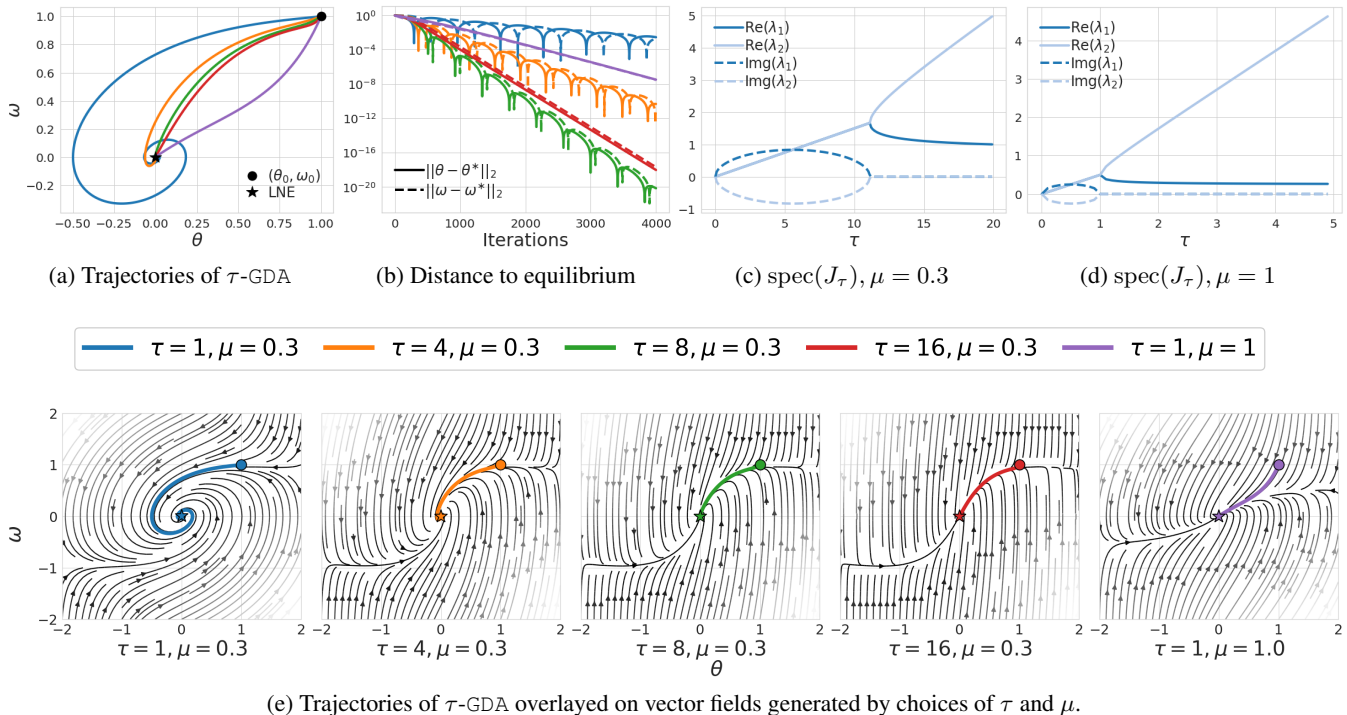
Theorem 4 directly implies a finite time convergence guarantee for obtaining an ϵ -differential Stackelberg equilibrium, that is, a point with an ϵ -ball around a differential Stackelberg equilibrium x^* .

Corollary 1. Given $\epsilon > 0$, under the assumptions of Theorem 4, τ -GDA obtains an ϵ -differential Stackelberg equilibrium in $\lceil (4\beta/\alpha) \log(\|x_0 - x^*\|/\epsilon) \rceil$ iterations for any $x_0 \in B_{\delta}(x^*)$ with $\delta = \alpha/(4L\beta)$ where L is the local Lipschitz constant of $I - \gamma J_{\tau}(x^*)$.

Moreover, the convergence rates and finite time guarantees extend to the gradient penalty regularized generative adversarial network described in the preceding section.

Corollary 2. Under the assumptions of Theorems 3 and 4, for any fixed $\mu \in (0, \infty)$ and $\tau \in (0, \infty)$, τ -GDA converges locally asymptotically at a rate of $O((1 - \alpha/(4\beta))^{k/2})$, and achieves an ϵ -equilibrium in $\lceil (4\beta/\alpha) \log(\|x_0 - x^*\|/\epsilon) \rceil$ iterations for any $x_0 \in B_{\delta}(x^*)$.

In Fiez and Ratliff (2020), the convergence analysis is extended to the stochastic setting in which agents have an unbiased estimator of their individual gradients.



(e) Trajectories of τ -GDA overlaid on vector fields generated by choices of τ and μ .

Figure 1: Figure 1a shows trajectories of τ -GDA for $\tau \in \{1, 4, 8, 16\}$ with regularization $\mu = 0.3$ and $\tau = 1$ with regularization $\mu = 1$. Figure 1b shows the distance from the equilibrium along the learning paths. Figures 1c and 1d show the trajectories of the eigenvalues for $J_\tau(\theta^*, \omega^*)$ as a function of τ with regularization set to $\mu = 0.3$ and $\mu = 1$, respectively where (θ^*, ω^*) is the unique critical point of the game. Figure 1e shows the trajectories of τ -GDA overlaid on the vector field generated by the respective timescale separation and regularization parameters. The shading of the vector field is dictated by its magnitude so that lighter shading corresponds to a higher magnitude and darker shading corresponds to a lower magnitude.

5 Experiments

We now present numerical experiments examining gradient descent-ascent with timescale separation. Fiez and Ratliff (2020) contains more experimental results and details.

Dirac-GAN: Regularization, Timescale Separation, and Convergence Rate. The Dirac-GAN (Mescheder, Geiger, and Nowozin 2018) consists of a univariate generator distribution $p_\theta = \delta_\theta$ and a linear discriminator $D(x; \omega) = \omega x$, where the real data distribution $p_{\mathcal{D}}$ is given by a Dirac-distribution concentrated at zero. The resulting zero-sum game is defined by the cost $f(\theta, \omega) = \ell(\theta\omega) + \ell(0)$ and the unique critical point $(\theta^*, \omega^*) = (0, 0)$ is a local Nash equilibrium. However, the eigenvalues of the Jacobian are purely imaginary regardless of the choice of timescale separation so that τ -GDA oscillates and fails to converge. This behavior is expected since the equilibrium is not hyperbolic and corresponds to neither a differential Nash equilibrium nor a differential Stackelberg equilibrium but it is undesirable nonetheless. The zero-sum game corresponding to the Dirac-GAN with regularization can be defined by the cost $f(\theta, \omega) = \ell(\theta\omega) + \ell(0) - \frac{\mu}{2}\omega^2$. The unique critical point remains unchanged, but for all $\tau \in (0, \infty)$ and $\mu \in (0, \infty)$ the equilibrium of the unregularized game is stable and corresponds to a differential Stackelberg equilibrium of the reg-

ularized game.

From Figures 1a and 1e, we observe that the impact of timescale separation with regularization $\mu = 0.3$ is that the trajectory is not as oscillatory since it moves faster to the zero line of $-D_2 f(\theta, \omega)$ and then follows along that line until reaching the equilibrium. We further see from Figure 1b that with regularization $\mu = 0.3$, τ -GDA with $\tau = 8$ converges faster to the equilibrium than τ -GDA with $\tau = 16$, despite the fact that the former exhibits some cyclic behavior in the dynamics while the latter does not. The eigenvalues of the Jacobian with regularization $\mu = 0.3$ presented in Figure 1c explains this behavior since the imaginary parts are non-zero with $\tau = 8$ and zero with $\tau = 16$, but the eigenvalue with the minimum real part is greater at $\tau = 8$ than at $\tau = 16$. This highlights that some oscillatory behavior in the dynamics is not always harmful for convergence and it can even speed up the rate of convergence. For $\mu = 1$ and $\tau = 1$, Figures 1a and 1b show that even though τ -GDA follows a direct path toward the equilibrium and does not cycle since the eigenvalues of the Jacobian are purely real, the trajectory converges slowly to the equilibrium. Indeed, for each regularization parameter, the eigenvalues of $J_\tau(\theta^*, \omega^*)$ split after becoming purely real and then converge toward the eigenvalues of $S_1(J(\theta^*, \omega^*))$ and $-\tau D_2^2 f(\theta^*, \omega^*)$. Since $S_1(J(\theta^*, \omega^*)) \propto 1/\mu$ and $-\tau D_2^2 f(\theta^*, \omega^*) \propto \tau\mu$, there is

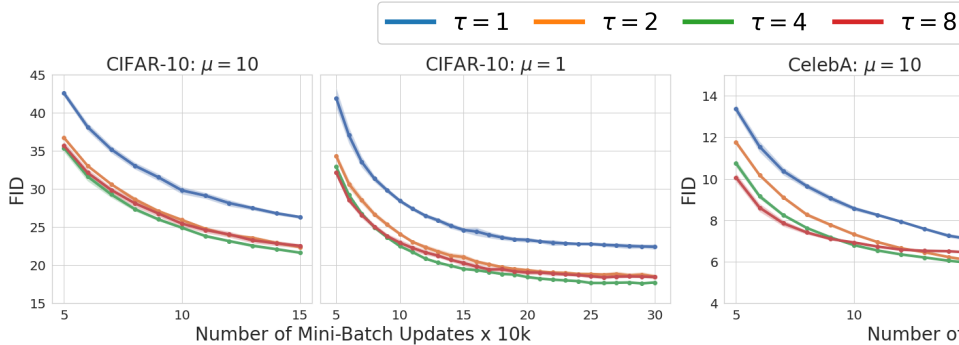


Figure 2: CIFAR-10 FID

$\mu \backslash \tau$	1	2	4	8
10	26.29	22.36	21.62	22.52
1	24.6/22.4	21.05/18.5	19.49/17.72	20.27/18.45

Figure 4: CIFAR-10 FID at 150k/300k mini-batch updates.

a trade-off between the choice of regularization μ and the timescale separation τ on the conditioning of the Jacobian matrix that dictates the convergence rate.

Generative Adversarial Networks: Image Datasets.

We build our experiments based on the methods and implementations of Mescheder, Geiger, and Nowozin (2018) and train with the non-saturating objective function and the R_1 gradient penalty. The network architectures for the generator and discriminator are both based on the ResNet architecture. The initial learning rate for the generator in all of our experiments is fixed to be $\gamma_1 = 0.0001$ and we decay the stepsizes so that at update k the learning rate of the generator is given by $\gamma_{1,k} = \gamma_1 / (1 + \nu)^k$ where $\nu = 0.005$ and the learning rate of the discriminator is $\gamma_{2,k} = \tau \gamma_{1,k}$. The batch size is 64, the latent data is drawn from a standard normal distribution of dimension 256, and the resolution of the images is $32 \times 32 \times 3$. We run RMSprop with parameter $\alpha = 0.99$ and retain an exponential moving average of the generator parameters to produce the model that is evaluated with parameter $\beta = 0.9999$.

The FID scores (Heusel et al. 2017) along the learning path and in numeric form at 150k/300k mini-batch updates for CIFAR-10 and CelebA with regularization parameters $\mu = 10$ and $\mu = 1$ are presented in Figures 2–5. We repeated each experiment 3 times and report the mean scores. For CIFAR-10 and with each regularization parameter, $\tau = 4$ converges fastest, followed by $\tau = 8$, then $\tau = 2$, and finally $\tau = 1$. For CelebA and regularization $\mu = 10$, the timescale parameters of $\tau = 2$ and $\tau = 4$ outperform $\tau = 1$ and $\tau = 8$ by a wide margin. A similar trend can be observed for regularization $\mu = 1$, but with $\tau = 8$ performing closer to $\tau = 2$ and $\tau = 4$.

The performance with regularization $\mu = 1$ is far superior to that with regularization $\mu = 10$ for each timescale parameter and with each dataset. Moreover, we generally see that timescale separation improves convergence until hitting a limiting value. Interestingly, these conclusions agree with

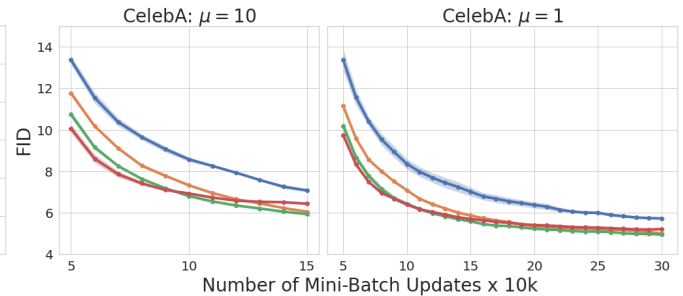


Figure 3: CelebA FID

$\mu \backslash \tau$	1	2	4	8
10	7.07	6.05	5.93	6.43
1	7.01/5.72	5.86/5.0	5.59/4.95	5.7/5.21

Figure 5: CelebA FID at 150k/300k mini-batch updates.

the insights from the simple Dirac-GAN experiment. This experiment reinforces that timescale separation is an important hyperparameter worth careful consideration in conjunction with regularization since the interplay between them can significantly impact convergence speed and final performance.

6 Conclusion

We prove gradient descent-ascent converges to a critical point for a range of finite learning rate ratios if and only if the critical point is a differential Stackelberg equilibrium. This answers a standing open question about the convergence of first order methods to local minimax equilibria. A key component of the proof is the construction of a (tight) finite lower bound on the learning rate ratio τ for which stability is guaranteed, and hence local asymptotic convergence of τ -GDA. This being said, the question of the size of the region of attraction remains open. Related, but distinct techniques handle the nonlinear system directly. The downside of this technique is that one needs to be able to construct Lyapunov functions for both the *boundary layer model* (the system that arises from treating the choice variable of the slow player as being ‘static’) and the *reduced order model* (the system that arises from plugging in the implicit mapping from the fast player’s action to the slow player’s action into the slow player’s dynamics). A convex combination of these functions provides a Lyapunov function for the original system $\dot{x} = -\Lambda_\tau g(x)$. The level sets of this combined Lyapunov function then give a sense of the region of attraction. In fact, one can optimize over the weighting in the convex combination in order to obtain improved estimates of the region of attraction. This is an interesting avenue to explore in the context of learning in games with intrinsic structure that can potentially be exploited to improve both the rate of convergence and the region on which convergence is guaranteed.

References

- Arjovsky, M.; Chintala, S.; and Bottou, L. 2017. Wasserstein generative adversarial networks. In *International Conference on Machine Learning (ICML)*, 214–223.
- Başar, T.; and Olsder, G. J. 1998. *Dynamic noncooperative game theory*. SIAM.
- Berard, H.; Gidel, G.; Almahairi, A.; Vincent, P.; and Lacoste-Julien, S. 2020. A Closer Look at the Optimization Landscapes of Generative Adversarial Networks. In *International Conference on Learning Representations (ICLR)*.
- Chasnov, B.; Ratliff, L. J.; Mazumdar, E.; and Burden, S. A. 2019. Convergence Analysis of Gradient-Based Learning in Continuous Games. *Conference on Uncertainty in Artificial Intelligence (UAI)*.
- Cherukuri, A.; Gharesifard, B.; and Cortes, J. 2017. Saddle-point dynamics: conditions for asymptotic stability of saddle points. *SIAM Journal on Control and Optimization* 55(1): 486–511.
- Daskalakis, C.; and Panageas, I. 2018. The limit points of (optimistic) gradient descent in min-max optimization. In *Advances in Neural Information Processing Systems*, 9236–9246.
- Fiez, T.; Chasnov, B.; and Ratliff, L. J. 2020. Implicit Learning Dynamics in Stackelberg Games: Equilibria Characterization, Convergence Analysis, and Empirical Study. *International Conference on Machine Learning (ICML)*.
- Fiez, T.; and Ratliff, L. 2020. Gradient Descent-Ascent Provably Converges to Strict Local Minmax Equilibria with a Finite Timescale Separation. *arXiv preprint arXiv:2009.14820*.
- Goodfellow, I. 2016. NIPS 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160*.
- Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2672–2680.
- Govaerts, W. J. F. 2000. *Numerical Methods for Bifurcations of Dynamical Equilibria*. Society for Industrial and Applied Mathematics.
- Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; and Hochreiter, S. 2017. Gans trained by a two time-scale update rule converge to a local nash equilibrium. In *Advances in Neural Information Processing Systems (NeurIPS)*, 6626–6637.
- Ho, J.; and Ermon, S. 2016. Generative adversarial imitation learning. In *Advances in neural information processing systems*, 4565–4573.
- Jin, C.; Netrapalli, P.; and Jordan, M. I. 2020. What is local optimality in nonconvex-nonconcave minimax optimization? *International Conference on Machine Learning (ICML)*.
- Khalil, H. 2002. *Nonlinear Systems*. Prentice Hall, 3rd edition.
- Kokotovic, P. V.; O’Reilly, J.; and Khalil, H. K. 1986. *Singular Perturbation Methods in Control: Analysis and Design*. Academic Press, Inc.
- Lancaster, P.; and Tismenetsky, M. 1985. *The theory of matrices: with applications*. Elsevier.
- Lorraine, J.; Vicol, P.; and Duvenaud, D. 2020. Optimizing millions of hyperparameters by implicit differentiation. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 1540–1552.
- MacKay, M.; Vicol, P.; Lorraine, J.; Duvenaud, D.; and Grosse, R. 2019. Self-tuning networks: Bilevel optimization of hyperparameters using structured best-response functions. *International Conference on Learning Representations (ICLR)*.
- Madry, A.; Makelov, A.; Schmidt, L.; Tsipras, D.; and Vladu, A. 2018. Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations*.
- Magnus, J. 1988. *Linear Structures*. Griffin.
- Mazumdar, E.; Ratliff, L. J.; and Sastry, S. S. 2020. On Gradient-Based Learning in Continuous Games. *SIAM Journal on Mathematics of Data Science* 2(1): 103–131.
- Mescheder, L.; Geiger, A.; and Nowozin, S. 2018. Which Training Methods for GANs do actually Converge? In *International Conference on Machine Learning (ICML)*, 3481–3490.
- Nagarajan, V.; and Kolter, J. Z. 2017. Gradient descent GAN optimization is locally stable. In *Advances in Neural Information Processing Systems (NeurIPS)*, 5585–5595.
- Rajeswaran, A.; Mordatch, I.; and Kumar, V. 2020. A Game Theoretic Framework for Model Based Reinforcement Learning. *International Conference on Machine Learning (ICML)*.
- Ratliff, L. J.; Burden, S. A.; and Sastry, S. S. 2013. Characterization and computation of local Nash equilibria in continuous games. In *Allerton Conference on Communication, Control, and Computing*, 917–924.
- Roth, K.; Lucchi, A.; Nowozin, S.; and Hofmann, T. 2017. Stabilizing training of generative adversarial networks through regularization. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018–2028.
- Saydy, L.; Tits, A. L.; and Abed, E. H. 1990. Guardian maps and the generalized stability of parametrized families of matrices and polynomials. *Mathematics of Control, Signals and Systems* 3(4): 345–371.
- Sinha, A.; Namkoong, H.; and Duchi, J. 2018. Certifiable distributional robustness with principled adversarial training. *International Conference on Learning Representations (ICLR)*.
- Wang, Y.; Zhang, G.; and Ba, J. 2020. On solving minimax optimization locally: A follow-the-ridge approach. *International Conference on Learning Representations (ICLR)*.
- Zhang, K.; Yang, Z.; and Başar, T. 2019. Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms. *arXiv preprint arXiv:1911.10635*.