Proceedings of Machine Learning Research vol 168:1–15, 2022 4th Annual Conference on Learning for Dynamics and Control

Convergence Rates of Two-Time-Scale Gradient Descent-Ascent Dynamics for Solving Nonconvex Min-Max Problems

Thinh T. Doan THINHDOAN@VT.EDU
Department of Electrical and Computer Engineering
Virginia Tech, USA

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

Abstract

There are much recent interests in solving nonconvex min-max optimization problems due to its broad applications in many areas including machine learning, networked resource allocations, and distributed optimization. Perhaps, the most popular first-order method in solving min-max optimization is the so-called simultaneous (or single-loop) gradient descent-ascent algorithm due to the simplicity in its implementation. However, theoretical guarantees on the convergence of this algorithm are very sparse since it can diverge even in a simple bilinear problem.

In this paper, our focus is to characterize the finite-time performance (or convergence rates) of the continuous-time variant of simultaneous gradient descent-ascent algorithm. In particular, we derive the rates of convergence of this method under a number of different conditions on the underlying objective function, namely, two-sided Polyak-Łojasiewicz (PŁ), one-sided PŁ, nonconvex-strongly concave, and strongly convex-nonconcave conditions. Our convergence results improve the ones in prior works under the same conditions of objective functions. The key idea in our analysis is to use the classic singular perturbation theory and coupling Lyapunov functions to address the time-scale difference and interactions between the gradient descent and ascent dynamics. Our results on the behavior of continuous-time algorithm may be used to enhance the convergence properties of its discrete-time counterpart.

Keywords: Gradient Descent-Ascent Methods, Two-Time-Scale Dynamics, Singular Perturbation Theory, Min-Max Optimization.

1. Introduction

In this paper, we consider the following min-max optimization problems

$$\min_{x\in\mathbb{R}^m}\max_{y\in\mathbb{R}^n} f(x,y), \tag{1}$$

where $f:\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}$ is a nonconvex function w.r.t x for a fixed y and (possibly) nonconcave w.r.t y for a fixed x. The min-max problem has received much interests for years due to its broad applications in different areas including control, machine learning, and economics. In particular, many problems in these areas can be formulated as problem (1), for example, game theory Basar and Olsder (1998); Shapley (1953), stochastic control and reinforcement learning Altman (1999); Achiam et al. (2017), training generative adversarial networks (GANs) Goodfellow et al. (2020); Mescheder et al. (2017), adversarial and robust machine learning Kurakin et al. (2017); Qian et al. (2019), resource allocation over networks Liu et al. (2013), and distributed optimization Lan et al. (2020); Chang et al. (2020); to name just a few.

In the existing literature, there are two types of iterative first-order methods for solving problem (1), namely, nested-loop algorithms and single-loop algorithms. Nested-loop algorithms implement multiple inner steps in each iteration to solve the maximization problem either exactly or approximately. However, this approach is not applicable to the setting when $f(x, y)$ is nonconcave in y, since the maximization problem is NP-hard. Only finding a stationary point of the maximization problem is likely to affect the quality of solving the minimization problem.

On the other hand, single-loop algorithm simultaneously updates the iterates x and y by using the vanilla gradient descent and ascent steps at different time scales, respectively. As a result, this algorithm is applicable to more general settings and more practical due to its simplicity in implementation. However, single-loop algorithms may not converge in many settings, for example, they fail to converge even in a simple bilinear zero-sum game Balduzzi et al. (2018). Indeed, theoretical guarantees of these methods are very sparse.

Our focus in this paper is to study the continuous-time variant of the single-loop gradient descent-ascent method for solving problem (1). Considering the continuous-time variant will help us to have a better understanding about the behavior of this method through studying the convergence of the corresponding differential equations using Lyapunov theory. Such an understanding can then be used to enhance the analysis of the discrete-time algorithms, as recently observed in the single objective optimization counterpart Krichene et al. (2015); Raginsky and Bouvrie (2012); Su et al. (2014); Diakonikolas and Orecchia (2019). Our main contributions are summarized below.

Main Contributions. The focus of this paper is to study the performance of the continuous-time gradient descent-ascent dynamics in solving nonconvex min-max optimization problems. In particular, we derive the rates of convergence of this method under a number of different conditions on the underlying objective function, namely, two-sided Polyak-Łojasiewicz (PŁ), one-sided PŁ, nonconvex-strongly concave, and strongly convex-nonconcave conditions. These rates are summarized in Table 1 and presented in detail in Section 3, where we show that our results improve the ones in prior works under the same conditions of objective functions. The key idea in our analysis is to use the classic singular perturbation theory and coupling Lyapunov function of the fast and slow dynamics to address the time-scale difference and interactions between the gradient descent and ascent dynamics. Proper choices of step sizes allows us to derive improved convergence properties of the two-time-scale gradient descent-ascent dynamics.

1.1. Related Works

Convex-Concave Settings. Given the broad applications of problem (1), there are a large number of works to study algorithms and their convergence in solving this problem, especially in the context of convex-concave settings. Some examples include prox-method and its variant Nemirovski (2004); Malitsky (2015); Wang and Li (2020); Cherukuri et al. (2017), extragradient and optimistic gradient methods Korpelevich (1976); Mokhtari et al. (2020); Monteiro and Svaiter (2010); Golowich et al. (2020); Yoon and Ryu (2021); Dang and Lan (2015), and recently Hamiltonian gradient descent methods Mescheder et al. (2017); Balduzzi et al. (2018); Abernethy et al. (2021). Some algorithms in these settings have convergence rates matched with the lower bound complexity; see the recent work Yoon and Ryu (2021) for a detailed discussion.

Nonconvex-Concave Settings. Unlike the convex-concave settings, algorithmic development and theoretical understanding in the general nonconvex settings are very limited. Indeed, finding the global optimality of nonconvex-nonconcave problem is NP-hard, or at least as hard as solving a single nonconvex objective problem. As a result, the existing literature often aims to find a stationary point of f when the max problem is concave. For example, multiple-loop algorithms have been studied in Thekumparampil et al. (2019); Kong and Monteiro (2021); Rafique et al. (2021); Lin et al. (2020b); Nouiehed et al. (2019). Our work in this paper is closely related to the recent literature on studying single-loop algorithm Lin et al. (2020a); Lu et al. (2020); Yang et al. (2020); Xu et al. (2020); Zhang et al. (2020). While these works study discrete-time algorithms, we consider continuous-time counterpart. We will show that for some settings, our approach improves the existing convergence results.

Other Settings. We also want to mention some related literature in game theory Loizou et al. (2020); Zhang et al. (2019); Cen et al. (2021); Perolat et al. (2018); Zhang et al. (2021); Daskalakis et al. (2020), two-time-scale stochastic approximation Borkar (2008); Konda and Tsitsiklis (2004); Dalal et al. (2020); Doan and Romberg (2019); Gupta et al. (2019); Doan (2021b); Kaledin et al. (2020); Mokkadem and Pelletier (2006); Doan (2021c, 2020), and reinforcement learning Bhatnagar and Lakshmanan (2012); Paternain et al. (2019); Ding et al. (2020); Zeng et al. (2021), two-time-scale optimization , and decentralized optimization . These works study different variants of two-time-scale methods mostly for solving a single optimization problem, and often aim to find global optimality (or fixed points) using different structure of the underlying problems (e.g., Markov structure in stochastic games and reinforcement learning or strong monotonicity in stochastic approximation). Thus, the techniques therein may not be applicable to the context of problem (1) considered in the current paper.

Notation. Given any vector x we use $\|x\|$ to denote its 2-norm. We denote by $\nabla_x f(x,y)$ and $\nabla_y f(x,y)$ the partial gradients of f with respect to x and y, respectively.

2. Two-Time-Scale Gradient Descent-Ascent Dynamics

For solving problem (1), we are interested in studying two-time-scale gradient descent-ascent dynamics (GDAD), where we implement simultaneously the following two differential equations

$$\begin{aligned}\dot{x}(t) &= \frac{d}{dt}x(t) = -\alpha\nabla_x f(x(t),y(t)),\\ \dot{y}(t) &= \frac{d}{dt}y(t) = \beta\nabla_y f(x(t),y(t)),\end{aligned} \tag{2}$$

Here, α, β are two step sizes, whose values will be specified later. In the convex-concave setting, one can choose $\alpha = \beta$. However, as observed in Heusel et al. (2017), choosing different step sizes achieves a better convergence in the context of nonconvex problem. Indeed, we will choose $\alpha \ll \beta$ since in our settings studied in the following sections, the maximization problem is often easier to solve than the minimization problem. In this case, the dynamic of $y(t)$ is implemented at a faster time scale (using larger step sizes) than $x(t)$ (using smaller step sizes). The time-scale difference is loosely defined as the ratio $\alpha/\beta \ll 1$. Thus, one has to design these two step sizes properly so that the method converges as fast as possible.

Technical Approach. The convergence analysis of (2) studied in this paper is mainly motivated by the classic singular perturbation theory Kokotović et al. (1999), explained as follows. Since y is implemented at a faster time scale than x, we consider $x(t) = x$ being fixed in $\dot{y}$ and separately study the stability of the system $\dot{y}$ using Lyapunov theory. Let V_2 be the Lyapunov function corresponding to $\dot{y}$. When $\dot{y}$ converges to an equilibrium y (e.g., $\nabla_y f(x,y) = 0$), one can fix $y(t) = y$ and study the stability of $\dot{x}$. Let V_1 be the corresponding Lyapunov function of $\dot{x}$. We note that V_1 and V_2 both depend on x and y, as a result, their time derivatives are coupled through the dynamics in (2). Addressing this coupling and the time-scale difference between the two dynamics is the key idea in our approach. To do that, we will consider the following Lyapunov function

$$V(x,y) = V_1(x,y) + \frac{\gamma\alpha}{\beta}V_2(x,y), \tag{3}$$

where α/β represents the time-scale difference, while the constant γ will be properly chosen to eliminate the impact of x on the convergence of y and vice versa. Proper choices of these constants will also help us to derive the convergence rates of (2). Similar approach has been used in different settings of two-time-scale methods, see for example Chow and Kokotovic (1985); Bıyık and Arcak (2008); Doan (2020); Dutta et al. (2021).

We conclude this section by introducing two assumptions for our analysis studied later.

Assumption 1 *The function $f(\cdot,\cdot)$ has Lipschitz continuous gradients for each variable, i.e., there exist positive constants L_x, L_y, and L_{xy} such that for all $x_1, x_2 \in \mathbb{R}^m, y_1, y_2 \in \mathbb{R}^n$ we have*

$$\begin{aligned} \|\nabla_x f(x_1,y_1) - \nabla_x f(x_2,y_2)\| &\leq L_x\|x_1 - x_2\| + L_{xy}\|y_1 - y_2\|,\\ \|\nabla_y f(x_1,y_1) - \nabla_y f(x_2,y_2)\| &\leq L_{xy}\|x_1 - x_2\| + L_y\|y_1 - y_2\|. \end{aligned} \tag{4}$$

Assumption 2 *Given any x the problem $\max_y f(x,y)$ has a nonempty solution set $\mathcal{Y}^\star(x)$, i.e., there exists $y^\star(x) \in \mathcal{Y}^\star(x)$ such that*

$$y^\star(x) = \arg\max_{y\in\mathbb{R}^n} f(x,y), \quad \textit{where } f(x, y^\star(x)) \textit{ is finite.}$$

Table 1: Convergence rates of GDAD for solving (1) given some accuracy $\epsilon > 0$. The abbreviations NCvex, NCave, SCvex, SCave, and PŁ stand for nonconvex, nonconcave, strongly convex, strongly concave, and Polyak-Łojasiewicz condtions, respectively. Condition number κ is defined in (9), and R is the size of compact set used in Nouiehed et al. (2019).

OBJECTIVES	PRIOR WORKS	THIS PAPER
PŁ& PŁ	$\mathcal{O}\left(\kappa^3\log(\frac{1}{\epsilon})\right)$ YANG ET AL. (2020)	$\mathcal{O}\left(\kappa^2\log(\frac{1}{\epsilon})\right)$- THEOREM 1
NCVEX & PŁ	$\mathcal{O}\left(R^2 L_{xy}\log(\frac{1}{\epsilon})\epsilon^{-2}\right)$ NOUIEHED ET AL. (2019)	$\mathcal{O}\left(L_{xy}^2\epsilon^{-2}\right)$- THEOREM 2
NCVEX & SCAVE	$\mathcal{O}\left(L_{xy}^2\epsilon^{-2}\right)$ XU ET AL. (2020)	$\mathcal{O}\left(L_{xy}^2\epsilon^{-2}\right)$- THEOREM 3
SCVEX & NCAVE	$\mathcal{O}\left(L_{xy}^2\epsilon^{-2}\right)$ XU ET AL. (2020)	$\mathcal{O}\left(L_{xy}^2\epsilon^{-2}\right)$- THEOREM 4

3. Main Results

In this section, we present the main results of this paper, where we derive the convergence rates of GDAD under different conditions on $f(x,y)$. Our results are summarized in Table 1. First, our approach improves the analysis in Yang et al. (2020), where we show in Section 3.1 that for two-sided PŁ functions the convergence of GDAD only scales with κ^2 instead of κ^3 studied in Yang et al. (2020). Our result addresses the conjecture raised in Yang et al. (2020), where the authors state that such an improvement may not be possible. Second, our analysis achieves a better result than the one in Nouiehed et al. (2019) for the case of one-sided PL function by a factor of $\log(1/\epsilon)$. We note that a nested-loop is studied in Nouiehed et al. (2019) while GDAD is a single-loop method. Finally, our result is the same as the one in Xu et al. (2020) when $f(x,y)$ is either strongly concave in y for fixed x. In Section 3.4, we will show that this observation also holds when $f(x,y)$ is either strongly convex in x and nonconcave in y. Note that as compared to the analysis in Xu et al. (2020), we use a simpler analysis and simpler choice of step sizes to achieve these results. Due to space limitation, some proofs of technical lemmas in this paper can be found in Doan (2021a).

3.1. Two-Sided Polyak–Łojasiewicz Conditions

We first study the convergence rates of GDAD when f satisfies a two-sided Polyak–Łojasiewicz (PŁ) condition, which is considered in Yang et al. (2020) and stated here for convenience.

Definition 1 (Two-Sided PŁ Conditions) *A continuously differentiable function $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ is called to satisfy two-sided PŁ conditions if there exist two positive constants μ_x and μ_y such that $\mu_x, \mu_y \leq \min\{L_x, L_y, L_{xy}\}$ the following conditions hold for all $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$:*

$$\begin{aligned} 2\mu_x[f(x,y) - \min_x f(x,y)] &\leq \|\nabla_x f(x,y)\|^2,\\ 2\mu_y[\max_y f(x,y) - f(x,y)] &\leq \|\nabla_y f(x,y)\|^2. \end{aligned} \tag{5}$$

The two-sided PŁ condition, which we will assume to hold in this subsection, is a generalized variant of the popular PŁ condition, proposed by Polyak (1963) as a sufficient condition to guarantee that the classic gradient descent method converges exponentially to the optimal value of an unconstrained minimization problem. As shown in Karimi et al. (2016), the PŁ condition also implies the quadratic growth condition, i.e., given any x we have

$$\max_{z\in\mathbb{R}^m} f(x,z) - f(x,y) \geq \frac{\mu_y}{2}\|\mathcal{P}_{\mathcal{Y}^\star(x)}[y] - y\|^2, \quad \forall y \in \mathbb{R}^m, \tag{6}$$

where we assume that $\mathcal{Y}^\star(x)$ is a nonempty solution set of $\max_y f(x,y)$ and $\mathcal{P}_{\mathcal{Y}^\star(x)}[y]$ is the projection of y to this set. More discussions on PŁ condition can be found in Karimi et al. (2016), while some examples of functions satisfying the two-sided PŁ condition are given in Yang et al. (2020).

Our focus in this section is to show that GDAD converges exponentially to the global min-max solution $(x^\star, y^\star)$ of f under the two-sided PŁ condition. To do that, we consider the following assumption and lemmas, which are useful for our analysis considered later. We first consider an assumption on the existence of $(x^\star, y^\star)$, a global min-max solution of f.

Assumption 3 *There exists a global min-max solution $(x^\star, y^\star)$ of f, i.e.,*

$$x^\star = \arg\min_{x\in\mathbb{R}^m} f(x, y^\star) \quad \textit{and} \quad y^\star = \arg\max_{y\in\mathbb{R}^n} f(x^\star, y).$$

Next, we consider the following lemma about the Lipschitz continuity of the gradient of $f(x, y^\star(x))$, which is a variant of the well-known Danskin lemma Bertsekas (1999)[Proposition B.25] and studied in Nouiehed et al. (2019)[Lemma A.5].

Lemma 1 *Suppose that Assumptions 1– 3 hold. Then, the function $\max_y f(x,y)$ is differentiable and its gradient $\nabla_x f(x, y^\star(x))$ is Lipschitz continuous with a constant $L_x + \frac{L_{xy}}{\mu_y}$.*

Finally, for our analysis we consider the following two Lyapunov functions

$$V_1(x) = \max_{y\in\mathbb{R}^n} f(x,y) - \min_{x\in\mathbb{R}^m}\max_{y\in\mathbb{R}^n} f(x,y) \quad \text{and} \quad V_2(x,y) = \max_y f(x,y) - f(x,y),$$

where it is obvious to see that V_1 and V_2 are nonnegative. The time derivatives of V_1 and V_2 over $\dot{x}$ and $\dot{y}$ are given in the following lemma, where its proof can be found in Doan (2021a).

Lemma 2 *Suppose that Assumptions 1– 3 hold. Then we have*

$$\dot{V}_1(x) \leq -\frac{\alpha}{2}\|\nabla_x f(x,y^\star(x))\|^2 + \frac{L_{xy}^2\alpha}{\mu_y}V_2(x,y). \tag{7}$$

$$\dot{V}_2(x,y) \leq -\beta\|\nabla_y f(x,y)\|^2 + \frac{3\alpha}{2}\|\nabla_x f(x,y^\star(x))\|^2 + \frac{5L_{xy}^2\alpha}{\mu_y}V_2(x,y). \tag{8}$$

As mentioned, the dynamics of $\dot{x}$ and $\dot{y}$ are implemented at different time scales, loosely defined as the ratio $\beta/\alpha > 1$. To capture such time-scale difference in our analysis, we will utilize the coupling Lyapunov function defined in (3). We denote by $\mu = \min\{\mu_x,\mu_y\}$ and the condition number

$$\kappa = \frac{L_{xy}}{\mu} \geq 1. \tag{9}$$

representing the condition number of $f(x,y)$. The convergence rate of GDAD under the two-sided PŁ condition is formally stated in the following theorem.

Theorem 1 *Suppose that Assumptions 1– 3 hold. Let γ,α,β be chosen as*

$$\gamma = \frac{L_{xy}^2}{\mu_y^2},\quad \alpha = \frac{\mu^2}{10\mu_x L_{xy}^2},\quad \beta = \frac{\mu^2}{\mu_x\mu_y^2}. \tag{10}$$

Then we have for all $t \geq 0$

$$V(x(t),y(t)) \leq e^{-\frac{t}{20\kappa^2}}V(x(0),y(0)). \tag{11}$$

Proof By (5) we have

$$\|\nabla_y f(x,y)\|^2 \geq 2\mu_y[\max_y f(x,y) - f(x,y)] = 2\mu_y V_2(x,y).$$

Thus, by using Lemma 2, (3), and the preceding relation we have

$$\begin{aligned}
\dot{V}(x(t),y(t)) &= \dot{V}_1(x(t)) + \frac{\gamma\alpha}{\beta}\dot{V}_2(x(t),y(t))\\
&\leq -\frac{\alpha}{2}\|\nabla_x f(x(t),y^\star(x(t)))\|^2 + \frac{L_{xy}^2\alpha}{\mu_y}V_2(x(t),y(t))\\
&\quad - 2\mu_y\gamma\alpha V_2(x(t),y(t)) + \frac{3\gamma\alpha^2}{2\beta}\|\nabla_x f(x(t),y^\star(x(t)))\|^2 + \frac{5L_{xy}^2\gamma\alpha^2}{\mu_y\beta}V_2(x(t),y(t))\\
&= -\frac{\alpha}{4}\|\nabla_x f(x(t),y^\star(x(t)))\|^2 - \frac{\mu_y\gamma\alpha}{2}V_2(x(t),y(t))\\
&\quad - \left(\frac{1}{2} - \frac{3\gamma\alpha}{\beta}\right)\frac{\alpha}{2}\|\nabla_x f(x(t),y^\star(x(t)))\|^2\\
&\quad - \left(\frac{3\mu_y\gamma}{2} - \frac{L_{xy}^2}{\mu_y} - \frac{5L_{xy}^2\gamma\alpha}{\mu_y\beta}\right)\alpha V_2(x(t),y(t)).
\end{aligned} \tag{12}$$

Using (10) , (5), and $y^\star(x) = \arg\max_y f(x,y)$ into (12) we obtain

$$\begin{aligned}\dot{V}(x(t),y(t)) &\leq -\frac{\alpha}{4}\|\nabla_x f(x(t),y^\star(x(t)))\|^2 - \frac{\mu_y\gamma\alpha}{2}V_2(x(t),y(t))\\ &\leq -\frac{\mu_x\alpha}{2}(\max_y f(x(t),y) - \min_x\max_y f(x,y)) - \frac{\mu_y\gamma\alpha}{2}V_2(x,y)\\ &\leq -\frac{\mu_x\alpha}{2}(V_1(x(t)) + \frac{\gamma\alpha}{\beta}V_2(x(t),y(t))) = -\frac{\mu_x\alpha}{2}V(x(t),y(t)),\end{aligned}$$

where the last inequality is due to $\mu_x\alpha = \frac{\mu^2}{10L_{xy}^2} \leq \mu_y\beta = \frac{\mu^2}{\mu_x\mu_y}$. Taking the integral on both sides of the equation above immediately gives (11). ∎

3.2. Nonconvex–Polyak-Łojasiewicz Conditions

We next consider an extension of the previous section, where we assume that the objective function $f(x,\cdot)$ satisfies the Polyak-Łojasiewicz condition given any x and $f(\cdot,y)$ is nonconvex given any y.

Assumption 4 (One-Sided PŁ Conditions) *We assume that $f:\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}$ is nonconvex in x for any fixed y and satisfies the PŁ condition in y for any fixed x, that is, there exists a positive constants μ_y such that the following condition hold for any $x\in\mathbb{R}^m$:*

$$2\mu_y[\max_y f(x,y) - f(x,y)] \leq \|\nabla_y f(x,y)\|^2. \tag{13}$$

Since f satisfies only one-sided PŁ condition, we are giving up the hope to find a global optimal solution of (1), as studied in Theorem 1. Instead, we will show that GDAD will return a stationary point of f, as studied in Nouiehed et al. (2019). Note that under Assumption 2 the result in Lemma 1 still holds since the work in Nouiehed et al. (2019) only assumes one-sided PŁ condition. In addition, since we relax the two-sided PŁ condition, we introduce the following two Lyapunov functions for our analysis studied later.

$$V_1(x,y) = f(x,y) - \min_{x\in\mathbb{R}^m}\min_{y\in\mathbb{R}^n} f(x,y) \quad\text{and}\quad V_2(x,y) = \max_y f(x,y) - f(x,y),$$

where it is obvious to see that V_1 and V_2 are nonnegative. The time derivatives of V_1 and V_2 over the trajectories $\dot{x}$ and $\dot{y}$ are given in the following lemma, whose proof can be found in Doan (2021a).

Lemma 3 *Suppose that Assumptions 1, 2, and 4 hold. Then we have*

$$\dot{V}_1(x(t),y(t)) = -\alpha\|\nabla_x f(x(t),y(t))\|^2 + \beta\|\nabla_y f(x(t),y(t))\|^2. \tag{14}$$

$$\dot{V}_2(x(t),y(t)) \leq -\beta\|\nabla_y f(x(t),y(t))\|^2 + \frac{\alpha}{2}\|\nabla_x f(x(t),y(t))\|^2 + \frac{L_{xy}^2\alpha}{\mu_y}V_2(x(t),y(t)). \tag{15}$$

The convergence rate of GDAD under the nonconvex-PŁ condition is formally stated as follows.

Theorem 2 *Suppose that Assumptions 1, 2, and 4 hold. Let γ,α,β be chosen as*

$$\gamma = \frac{32L_{xy}^2}{\mu_y^2},\quad \alpha = \frac{1}{8L_{xy}^2},\quad \beta = \frac{1}{\mu_y^2}. \tag{16}$$

Then we have for all $T\geq 0$

$$\min_{0\leq t\leq T}\left\|\begin{array}{c}\nabla_x f(x(t),y(t))\\ \nabla_y f(x(t),y(t))\end{array}\right\| \leq \frac{4L_{xy}\sqrt{V_1(x(0),y(0)) + 4V_2(x(0),y(0))}}{\sqrt{T}}. \tag{17}$$

Proof By using (14), (15), and (3) we have

$$
\begin{aligned}
\dot{V}(x(t),y(t)) &= \dot{V}_1(x(t)) + \frac{\gamma\alpha}{\beta}\dot{V}_2(x(t),y(t))\\
&\leq -\alpha\|\nabla_x f(x(t),y(t))\|^2 + \beta\|\nabla_y f(x(t),y(t))\|^2\\
&\quad - \gamma\alpha\|\nabla_y f(x(t),y(t))\|^2 + \frac{\gamma\alpha^2}{8\beta}\|\nabla_x f(x(t),y(t))\|^2 + \frac{4L_{xy}^2\gamma\alpha^2}{\mu_y\beta}V_2(x,y)\\
&= -\frac{\alpha}{2}\|\nabla_x f(x(t),y(t))\|^2 - \frac{\gamma\alpha}{2}\|\nabla_y f(x(t),y(t))\|^2\\
&\quad - \frac{\gamma\alpha}{4}\|\nabla_y f(x(t),y(t))\|^2 + \beta\|\nabla_y f(x(t),y(t))\|^2\\
&\quad - \frac{\gamma\alpha}{4}\|\nabla_y f(x(t),y(t))\|^2 + \frac{4L_{xy}^2\gamma\alpha^2}{\mu_y\beta}V_2(x,y)\\
&\quad - \frac{\alpha}{2}\|\nabla_x f(x(t),y(t))\|^2 + \frac{\gamma\alpha^2}{8\beta}\|\nabla_x f(x(t),y(t))\|^2\\
&\leq -\frac{\alpha}{2}\|\nabla_x f(x(t),y(t))\|^2 - \frac{\gamma\alpha}{2}\|\nabla_y f(x(t),y(t))\|^2\\
&\quad - \frac{\gamma\alpha}{4}\|\nabla_y f(x(t),y(t))\|^2 + \beta\|\nabla_y f(x(t),y(t))\|^2\\
&\quad - \frac{\mu_y\gamma\alpha}{2}\Big(1 - \frac{4L_{xy}^2\alpha}{\mu_y^2\beta}\Big)V_2(x,y) - \frac{\alpha}{2}\Big(1 - \frac{\gamma\alpha}{4\beta}\Big)\|\nabla_x f(x(t),y(t))\|^2,
\end{aligned} \tag{18}
$$

where in the last inequality we use (13) to have

$$\|\nabla_y f(x,y)\|^2 \geq 2\mu_y[\max_y f(x,y) - f(x,y)] = 2\mu_y V_2(x,y).$$

Using (16) and the preceding relation into (18) gives

$$\dot{V}(x(t),y(t)) \leq -\frac{\alpha}{2}\|\nabla_x f(x(t),y(t))\|^2 - \frac{\gamma\alpha}{2}\|\nabla_y f(x(t),y(t))\|^2.$$

Taking the integral on both sides over $t \in [0,T]$ for some $T \geq 0$ and rearranging we obtain

$$\frac{\alpha}{2}\int_{t=0}^{T}\|\nabla_x f(x(t),y(t))\|^2 dt + \frac{\gamma\alpha}{2}\int_{t=0}^{T}\|\nabla_y f(x(t),y(t))\|^2 dt \leq V(x(0),y(0)),$$

which since $\gamma \geq 1$ and by using (16) immediately gives us (17). ■

3.3. Nonconvex–Strongly Concave Conditions

In this subsection, we study the rate of GDAD when the function $f(x,y)$ is nonconvex given any y and strongly concave given any x. In particular, we consider the following assumption.

Assumption 5 *The objective function $f(\cdot,y)$ is nonconvex for any given y and $f(x,\cdot)$ is strongly concave with constant $\mu_y > 0$ for any given x. The latter is equivalent to*

$$f(x,y_1) - f(x,y_2) - \langle\nabla f(x,y_2), y_1 - y_2\rangle \leq -\frac{\mu_y}{2}\|y_1 - y_2\|^2, \quad \forall y_1,y_2 \in \mathbb{R}^n. \tag{19}$$

For our analysis of in this section, we introduce the following two Lyapunov functions

$$V_1(x,y) = f(x,y) - \min_{(x,y)} f(x,y), \quad \text{and} \quad V_2(x,y) = \frac{1}{2}\|\dot{y}\|^2 = \frac{1}{2}\|\beta\nabla_y f(x,y)\|^2.$$

The time derivatives of V_1 and V_2 are given below, whose proof can be found in Doan (2021a).

Lemma 4 *Suppose that Assumptions 1 and 5 hold. Then we have*

$$\dot{V}_1(x(t),y(t)) \leq -\frac{1}{\alpha}\|\dot{x}(t)\|^2 + \frac{1}{\beta}\|\dot{y}(t)\|^2. \tag{20}$$

$$\dot{V}_2(x(t),y(t)) \leq L_{xy}\beta\|\dot{y}(t)\|\|\dot{x}(t)\| - \mu_y\beta\|\dot{y}(t)\|^2. \tag{21}$$

We next derive the convergence rate of GDAD under Assumption 5 in the following theorem, where we show that GDAD converges sublinear to a stationary point of f.

Theorem 3 *Suppose that Assumptions 1 and 5 hold. Let γ,α,β be chosen as*

$$\gamma = \mu_y L_{xy}^2, \quad \alpha = \frac{1}{L_{xy}^2}, \quad \beta = \frac{4}{\mu_y^2}. \tag{22}$$

Then we have for all $T \geq 0$

$$\min_{0\leq t\leq T}\left\| \begin{array}{c} \nabla_x f(x(t),y(t)) \\ \nabla_y f(x(t),y(t)) \end{array} \right\| \leq \frac{L_{xy}\sqrt{2V_1(x(0),y(0))}}{\sqrt{T}} + \frac{2L_{xy}\|\nabla_y f(x(0),y(0))\|}{\sqrt{\mu_y T}}. \tag{23}$$

Proof By using (20) and (21) we consider

$$\begin{aligned}
\dot{V}(x(t),y(t)) &= \dot{V}_1(x(t),y(t)) + \frac{\gamma\alpha}{\beta}\dot{V}_2(x(t),y(t)) \\
&\leq -\frac{1}{\alpha}\|\dot{x}(t)\|^2 + \frac{1}{\beta}\|\dot{y}(t)\|^2 + L_{xy}\gamma\alpha\|\dot{y}(t)\|\|\dot{x}(t)\| - \mu_y\gamma\alpha\|\dot{y}(t)\|^2 \\
&= -\frac{1}{2\alpha}\|\dot{x}(t)\|^2 - \frac{\mu_y\gamma\alpha}{4}\|\dot{y}(t)\|^2 - (\frac{\mu_y\gamma\alpha}{4} - \frac{1}{\beta})\|\dot{y}(t)\|^2 \\
&\quad - \frac{1}{2\alpha}\|\dot{x}(t)\|^2 + L_{xy}\gamma\alpha\|\dot{y}(t)\|\|\dot{x}(t)\| - \frac{\mu_y\gamma\alpha}{2}\|\dot{y}(t)\|^2 \\
&\leq -\frac{1}{2\alpha}\|\dot{x}(t)\|^2 - \frac{\mu_y\gamma\alpha}{4}\|\dot{y}(t)\|^2.
\end{aligned}$$

where in the last inequality we use (22). Using (2) we obtain

$$\begin{aligned}
\dot{V}(x(t),y(t)) &\leq -\frac{1}{2\alpha}\|\dot{x}(t)\|^2 - \frac{\mu_y\gamma\alpha}{4}\|\dot{y}(t)\|^2 \\
&= -\frac{\alpha}{2}\|\nabla_x f(x(t),y(t))\|^2 - \frac{4L_{xy}^2\alpha}{\mu_y^2}\|\nabla_y f(x(t),y(t))\|^2 \\
&\leq \frac{-\alpha}{2}\left(\|\nabla_x f(x(t),y(t))\|^2 + \|\nabla_y f(x(t),y(t))\|^2\right),
\end{aligned}$$

which when taking the integral on both sides over t from 0 to T and rearrange we obtain

$$\frac{\alpha}{2}\int_{t=0}^{T}\left(\|\nabla_x f(x(t),y(t))\|^2 + \|\nabla_y f(x(t),y(t))\|^2\right)dt \leq V(x(0),y(0)).$$

Thus, the preceding relation gives (23), i.e., for all $T > 0$

$$\begin{aligned}\min_{0\leq t\leq T}\left\|\begin{array}{c}\nabla_x f(x(t),y(t))\\ \nabla_y f(x(t),y(t))\end{array}\right\| &\leq \frac{\sqrt{2V(x(0),y(0))}}{\sqrt{\alpha T}} \leq \frac{\sqrt{2V_1(x(0),y(0))}}{\sqrt{\alpha T}} + \frac{\sqrt{2V_2(x(0),y(0))}}{\sqrt{\alpha T}}\\ &= \frac{L_{xy}\sqrt{2V_1(x(0),y(0))}}{\sqrt{T}} + \frac{2L_{xy}\|\nabla_y f(x(0),y(0))\|}{\sqrt{\mu_y T}}.\end{aligned}$$

■

3.4. Strongly Convex–Nonconcave Conditions

As mentioned, the single-loop GDA method is applicable to the convex-nonconcave min-max problem, while the nested-loop GDA method is not. In this section, we complete our analysis by studying the rate of GDAD when the function $f(x,y)$ is strongly convex given any y and nonconcave given any x. The analysis in this section is symmetric to the one in Section 3.3, so it is omitted here for brevity. More detail can be found in Doan (2021a).

Assumption 6 *The objective function $f(x,\cdot)$ is nonconcave for any given x and $f(\cdot,y)$ is strongly convex with constant $\mu_x > 0$ for any given y. The latter is equivalent to*

$$f(x_1,y) - f(x_2,y) - \langle\nabla f(x_2,y), x_1 - x_2\rangle \geq \frac{\mu_x}{2}\|x_1 - x_2\|^2, \quad \forall x_1,x_2\in\mathbb{R}^m. \tag{24}$$

The convergence rate of GDAD under Assumption 6 is presented in the following theorem, which basically is similar to the one in Theorem 3

Theorem 4 *Suppose that Assumptions 1 and 6 hold. Let γ,α,β be chosen as*

$$\gamma = \mu_x L_{xy}^2, \quad \alpha = \frac{4}{\mu_x^2}, \quad \beta = \frac{1}{L_{xy}^2}. \tag{25}$$

Then we have for all $T \geq 0$

$$\min_{0\leq t\leq T}\left\|\begin{array}{c}\nabla_x f(x(t),y(t))\\ \nabla_y f(x(t),y(t))\end{array}\right\| \leq \frac{L_{xy}\sqrt{2V_1(x(0),y(0))}}{\sqrt{T}} + \frac{2L_{xy}\|\nabla_x f(x(0),y(0))\|}{\sqrt{\mu_x T}}. \tag{26}$$

4. Concluding Remarks

In this paper, we consider two-time-scale gradient descent-ascent dynamics for solving nonconvex min-max optimization problems. Our main focus is to derive the convergence rates of this method for different settings of the underlying objective functions. Our techniques are mainly motivated by the classic singular perturbation, where we show that our analysis improves the existing results under the same conditions. A natural extension from this work is to provide a better analysis for the discrete-time variant of GDAD. Another interesting future direction is to consider the stochastic setting and its accelerated counterpart.

References

J. Abernethy, K. A. Lai, and A. Wibisono. Last-iterate convergence rates for min-max optimization: Convergence of hamiltonian gradient descent and consensus optimization. In *Proceedings of the 32nd International Conference on Algorithmic Learning Theory*, volume 132, pages 3–47, 2021.

J. Achiam, D. Held, A. Tamar, and P. Abbeel. Constrained policy optimization. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pages 22–31, 2017.

Eitan Altman. *Constrained Markov decision processes*. Chapman and Hall/CRC Press, 1999.

D. Balduzzi, S. Racaniere, J. Martens, J. Foerster, K. Tuyls, and T. Graepel. The mechanics of n-player differentiable games. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80, pages 354–363, 2018.

Tamer Basar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory, 2nd Edition*. Society for Industrial and Applied Mathematics, 1998.

D. Bertsekas. *Nonlinear Programming: 2nd Edition*. Cambridge, MA: Athena Scientific, 1999.

S. Bhatnagar and K. Lakshmanan. An online actor–critic algorithm with function approximation for constrained Markov decision processes. *Journal of Optimization Theory and Applications*, 2012.

Emrah Bıyık and Murat Arcak. Area aggregation and time-scale modeling for sparse nonlinear networks. *Systems & Control Letters*, 57(2):142–149, 2008.

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

Shicong Cen, Yuting Wei, and Yuejie Chi. Fast policy extragradient methods for competitive games with entropy regularization. *ArXiv*, abs/2105.15186, 2021.

T-H Chang, M. Hong, H-T Wai, X. Zhang, and S. Lu. Distributed learning in the nonconvex world: From batch data to streaming and beyond. *IEEE Signal Processing Magazine*, 37(3):26–38, 2020.

A. Cherukuri, B. Gharesifard, and J. Cortés. Saddle-point dynamics: Conditions for asymptotic stability of saddle points. *SIAM Journal on Control and Optimization*, 55(1):486–511, 2017.

J Chow and P Kokotovic. Time scale modeling of sparse dynamic networks. *Automatic Control, IEEE Transactions on*, 30(8):714–722, 1985.

G. Dalal, B. Szorenyi, and G. Thoppe. A tale of two-timescale reinforcement learning with the tightest finite-time bound. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04): 3701–3708, Apr. 2020.

Cong D. Dang and Guanghui Lan. On the convergence properties of non-euclidean extragradient methods for variational inequalities with generalized monotone operators. *Computational Optimization and Applications*, 60:277–310, 2015.

Constantinos Daskalakis, Dylan J Foster, and Noah Golowich. Independent policy gradient methods for competitive reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 33, pages 5527–5540, 2020.

Jelena Diakonikolas and Lorenzo Orecchia. The approximate duality gap technique: A unified theory of first-order methods. *SIAM Journal on Optimization*, 29(1):660–689, 2019.

Dongsheng Ding, Kaiqing Zhang, Tamer Basar, and Mihailo R Jovanovic. Natural policy gradient primal-dual method for constrained Markov decision processes. In *NeurIPS*, 2020.

T. T. Doan. Convergence rates of two-time-scale gradient descent-ascent dynamics for solving nonconvex min-max problems. Available at: https://www.dropbox.com/s/wxlg14nedcbsck4/Continuous_Time_Gradient_Descent_Ascent_for_Minmax_Problems.pdf?dl=0, 2021a.

T. T. Doan and J. Romberg. Linear two-time-scale stochastic approximation a finite-time analysis. In *57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 399–406, 2019.

Thinh T Doan. Nonlinear two-time-scale stochastic approximation: Convergence and finite-time performance. *arXiv preprint arXiv:2011.01868*, 2020.

Thinh T. Doan. Finite-time analysis and restarting scheme for linear two-time-scale stochastic approximation. *SIAM Journal on Control and Optimization*, 59(4):2798–2819, 2021b.

Thinh T Doan. Finite-time convergence rates of nonlinear two-time-scale stochastic approximation under Markovian noise. *arXiv preprint arXiv:2104.01627*, 2021c.

Amit Dutta, Nila Masrourisaadat, and Thinh T. Doan. Convergence rates of decentralized gradient methods over cluster networks. *arXiv preprint arXiv:2110.06992*, 2021.

Noah Golowich, Sarath Pattathil, Constantinos Daskalakis, and Asuman Ozdaglar. Last iterate is slower than averaged iterate in smooth convex-concave saddle point problems. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125, pages 1758–1784, 2020.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks. *Commun. ACM*, 63(11):139–144, 2020.

H. Gupta, R. Srikant, and L. Ying. Finite-time performance bounds and adaptive learning rate selection for two time-scale reinforcement learning. In *Advances in Neural Information Processing Systems*, 2019.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Maxim Kaledin, Eric Moulines, Alexey Naumov, Vladislav Tadic, and Hoi-To Wai. Finite time analysis of linear two-timescale stochastic approximation with Markovian noise. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125, pages 2144–2203, 2020.

Hamed Karimi, Julie Nutini, and Mark Schmidt. Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition. In Paolo Frasconi, Niels Landwehr, Giuseppe Manco, and Jilles Vreeken, editors, *Machine Learning and Knowledge Discovery in Databases*, pages 795–811, Cham, 2016. Springer International Publishing.

Petar Kokotović, Hassan K. Khalil, and John O'Reilly. *Singular Perturbation Methods in Control: Analysis and Design*. Society for Industrial and Applied Mathematics, 1999.

V. R. Konda and J. N. Tsitsiklis. Convergence rate of linear two-time-scale stochastic approximation. *The Annals of Applied Probability*, 14(2):796–819, 2004.

Weiwei Kong and Renato D. C. Monteiro. An accelerated inexact proximal point method for solving nonconvex-concave min-max problems. *SIAM Journal on Optimization*, 31(4):2558–2585, 2021.

G.M. Korpelevich. The extragradient method for finding saddle points and other problems. *Matecon*, 12:747–756, 1976.

Walid Krichene, Alexandre Bayen, and Peter L Bartlett. Accelerated mirror descent in continuous and discrete time. In *Advances in Neural Information Processing Systems*, volume 28, 2015.

Alexey Kurakin, Ian J. Goodfellow, and Samy Bengio. Adversarial machine learning at scale. In *5th International Conference on Learning Representations, ICLR*, 2017.

Guanghui Lan, Soomin Lee, and Yi Zhou. Communication-efficient algorithms for decentralized and stochastic optimization. *Mathematical Programming*, 180:237–284, 2020.

Tianyi Lin, Chi Jin, and Michael Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pages 6083–6093. PMLR, 13–18 Jul 2020a.

Tianyi Lin, Chi Jin, and Michael I. Jordan. Near-optimal algorithms for minimax optimization. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125, pages 2738–2779. PMLR, 09–12 Jul 2020b.

Ya-Feng Liu, Yu-Hong Dai, and Zhi-Quan Luo. Max-min fairness linear transceiver design for a multi-user mimo interference channel. *IEEE Transactions on Signal Processing*, 61(9):2413–2423, 2013.

N. Loizou, H. Berard, A. Jolicoeur-Martineau, P. Vincent, S. Lacoste-Julien, and I. Mitliagkas. Stochastic Hamiltonian gradient methods for smooth games. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pages 6370–6381, 2020.

Songtao Lu, Ioannis Tsaknakis, Mingyi Hong, and Yongxin Chen. Hybrid block successive approximation for one-sided non-convex min-max problems: Algorithms and applications. *IEEE Transactions on Signal Processing*, 68:3676–3691, 2020.

Yu. Malitsky. Projected reflected gradient methods for monotone variational inequalities. *SIAM Journal on Optimization*, 25(1):502–520, 2015.

Lars Mescheder, Sebastian Nowozin, and Andreas Geiger. The numerics of gans. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 1823–1833, 2017.

Aryan Mokhtari, Asuman E. Ozdaglar, and Sarath Pattathil. Convergence rate of O(1/k) for optimistic gradient and extragradient methods in smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 30(4):3230–3251, 2020.

A. Mokkadem and M. Pelletier. Convergence rate and averaging of nonlinear two-time-scale stochastic approximation algorithms. *The Annals of Applied Probability*, 16(3):1671–1702, 2006.

R. D. C. Monteiro and B. F. Svaiter. On the complexity of the hybrid proximal extragradient method for the iterates and the ergodic mean. *SIAM Journal on Optimization*, 20(6):2755–2787, 2010.

Arkadi Nemirovski. Prox-method with rate of convergence o(1/t) for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15(1):229–251, 2004.

Maher Nouiehed, Maziar Sanjabi, Tianjian Huang, Jason D Lee, and Meisam Razaviyayn. Solving a class of non-convex min-max games using iterative first order methods. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

S. Paternain, L. Chamon, M. Calvo-Fullana, and A. Ribeiro. Constrained reinforcement learning has zero duality gap. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Julien Perolat, Bilal Piot, and Olivier Pietquin. Actor-critic fictitious play in simultaneous move multistage games. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84, pages 919–928, 2018.

Boris Polyak. Gradient methods for the minimisation of functionals. *Ussr Computational Mathematics and Mathematical Physics*, 3:864–878, 12 1963.

Q. Qian, S. Zhu, J. Tang, R. Jin, B. Sun, and H. Li. Robust optimization over multiple domains. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:4739–4746, Jul. 2019.

Hassan Rafique, Mingrui Liu, Qihang Lin, and Tianbao Yang. Weakly-convex–concave min–max optimization: provable algorithms and applications in machine learning. *Optimization Methods and Software*, 0(0):1–35, 2021.

Maxim Raginsky and Jake Bouvrie. Continuous-time stochastic mirror descent on a network: Variance reduction, consensus, convergence. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 6793–6800, 2012.

Lloyd S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 1953.

W. Su, S. Boyd, and E. J. Candès. A differential equation for modeling nesterov's accelerated gradient method: Theory and insights. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, NIPS'14, page 2510–2518, 2014.

K. K Thekumparampil, P. Jain, P. Netrapalli, and S. Oh. Efficient algorithms for smooth minimax optimization. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Yuanhao Wang and Jian Li. Improved algorithms for convex-concave minimax optimization. In *Advances in Neural Information Processing Systems*, volume 33, pages 4800–4810, 2020.

Z. Xu, H-L Zhang, Y. Xu, and G. Lan. A unified single-loop alternating gradient projection algorithm for nonconvex-concave and convex-nonconcave minimax problems. *ArXiv*, abs/2006.02032, 2020.

Junchi Yang, Negar Kiyavash, and Niao He. Global convergence and variance reduction for a class of nonconvex-nonconcave minimax problems. In *Advances in Neural Information Processing Systems*, volume 33, pages 1153–1165, 2020.

Taeho Yoon and Ernest K Ryu. Accelerated algorithms for smooth convex-concave minimax problems with $o(1/k^2)$ rate on squared gradient norm. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pages 12098–12109, 2021.

S. Zeng, T. T. Doan, and J. Romberg. Finite-time complexity of online primal-dual natural actor-critic algorithm for constrained Markov decision processes. *ArXiv*, abs/2110.11383, 2021.

Jiawei Zhang, Peijun Xiao, Ruoyu Sun, and Zhiquan Luo. A single-loop smoothed gradient descent-ascent algorithm for nonconvex-concave min-max problems. In *Advances in Neural Information Processing Systems*, volume 33, pages 7377–7389, 2020.

K. Zhang, Z. Yang, and T. Basar. Policy optimization provably converges to nash equilibria in zero-sum linear quadratic games. In *NeurIPS*, volume 32, 2019.

Runyu Zhang, Zhaolin Ren, and Na Li. Gradient play in stochastic games: stationary points, convergence, and sample complexity. *ArXiv*, abs/2106.00198, 2021.