Asymptotic stability of saddle points under the saddle-point dynamics

Ashish Cherukuri Jorge Cortés

***Abstract*— This paper considers continuously differentiable functions of two vector variables that have (possibly a continuum of) min-max saddle points. We study the asymptotic convergence properties of the associated saddle-point dynamics (gradient-descent in the first variable and gradient-ascent in the second one). We identify a suite of complementary conditions under which the set of saddle points is asymptotically stable under the saddle-point dynamics. Our first set of results is based on the convexity-concavity of the function defining the saddle-point dynamics to establish the convergence guarantees. For functions that do not enjoy this feature, our second set of results relies on properties of the linearization of the dynamics and the function along the proximal normals to the saddle set. We also provide global versions of the asymptotic convergence results. Various examples illustrate our discussion.**

I. Introduction

It is well known that the trajectories of the gradient dynamics of a continuously differentiable function with bounded sublevel sets converge asymptotically to its set of critical points, see e.g. [1]. This fact, however, is not in general true for the saddle-point dynamics (gradient descent in one variable and gradient ascent in the other) of a continuously differentiable function of two variables, see e.g. [2], [3]. In this paper, our aim is to investigate conditions under which the above statement is true for the case where the critical points are min-max saddle points and they possibly form a continuum. Our motivation comes from the applications of saddle-point (or primal-dual) dynamics to find solutions of equality constrained optimization problems and Nash equilibria of zero-sum games.

Literature review: In constrained optimization problems, the pioneering works [2], [4] popularized the use of the primal-dual dynamics to arrive at the saddle points of the Lagrangian. For inequality constrained problems, this discontinuous dynamics uses saddle-point information of the Lagrangian together with a projection operator on the dual variables to preserve their nonnegativity. Recent works have further explored the convergence analysis of such dynamics, both in continuous [5], [6] and discrete [7] time. The work [8] proposes instead a convergent dynamics continuous in the state that builds on first- and second-order information of the Lagrangian. The recent work [9] studies the asymptotic properties of the saddle-point dynamics when the trajectories of the dynamics do not converge to the saddle points but instead show oscillatory behavior.

A. Cherukuri and J. Cortés are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, {acheruku,cortes}@ucsd.edu.

In the context of distributed control and multi-agent systems, an important motivation to study saddle-point dynamics comes from network optimization problems where the objective function is an aggregate of each agents' local objective function and the constraints are given by a set of conditions that are locally computable at the agent level. Because of this structure, the saddle-point dynamics of the Lagrangian for such problems is inherently amenable to a distributed implementation. This observation explains the emerging body of work that looks, from this perspective, at problems in distributed convex optimization [10], [11], [12], distributed linear programming [13], and applications to power networks [14], [15], [16] and bargaining problems [17]. In game theory, it is natural to study the convergence properties of saddle-point dynamics to find the Nash equilibria of two-person zero-sum games [18], [19]. The works [20], [21] depart from the class of gradient-like dynamical systems to propose dynamics based on the best-response map (thus requiring the solution of an optimization problem at each evaluation) to converge to the set of saddle points. A majority of these works assume the function whose saddle points are sought to be convex-concave in its arguments. Our focus here instead is on the asymptotic stability of the saddle-point dynamics for a wider class of functions, without any nonnegativity preserving projection on the individual variables. Moreover, we explicitly allow for the possibility of a continuum of saddle points, instead of isolated ones, and wherever feasible, on establishing convergence of the trajectories to a point in the set. The issue of asymptotic convergence, even in the case of standard gradient systems, is a delicate one when equilibria are a continuum [22].

Statement of contributions: We study the asymptotic convergence properties of the saddle-point dynamics of continuously differentiable functions of two (vector) variables. The dynamics consists of gradient descent in the first variable and gradient ascent in the second variable. Under the assumption that the set of min-max saddle points of the function is nonempty, our contributions consist of identifying conditions on the function such that the trajectories of the saddle-point dynamics provably converge to the set of saddle points, and possibly to a point in the set. Our first contribution is regarding functions that are locally convex-concave on the set of saddle points. We show that asymptotic stability of the set of saddle points is guaranteed if either the convexity or concavity properties are strict, and convergence is pointwise. Furthermore, motivated by equality constrained optimization problems, we show that the same conclusions on convergence hold for functions that depend linearly on one of its arguments if the strictness requirement is dropped. Our second contribution pertains to functions that lack the convex-concave property. We study the linearization of the dynamics and use results from center manifold theory to identify conditions on the eigenvalues of the linearization at the saddle points to guarantee convergence. Our third contribution is motivated by the observation that there exist continuously differentiable functions that do not fall into the hypotheses of our first two contributions but still enjoy the desired asymptotic convergence guarantees. We reason with the variation of the function and its Hessian along the proximal normal directions to the set of saddle points. Specifically, we assume polynomial bounds for these variations and derive an appropriate relationship between these bounds that ensure asymptotic convergence. When discussing each contribution, we extend the conditions to obtain global convergence wherever feasible. Various examples illustrate the application of our results. For reasons of space, the proofs are omitted and will appear elsewhere.

Organization: Section II introduces notation and basic preliminaries. Section III presents the saddle-point dynamics and the problem statement. Section IV deals with the case of convex-concave functions. For the case when this property does not hold, Section V relies on linearization techniques and proximal normals to establish convergence. Finally, Section VI gathers our conclusions and ideas for future work.

II. PRELIMINARIES

This section introduces basic preliminaries on proximal calculus and saddle points. We start with some notational conventions. Let $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ be the set of real and nonnegative real numbers, respectively. Given two sets $\mathcal{A}_1, \mathcal{A}_2 \subset \mathbb{R}^n$ the set $\mathcal{A}_1 + \mathcal{A}_2$ represents the sum $\{x+y \mid x \in \mathcal{A}_1, y \in \mathcal{A}_2\}$. We denote by $\|\cdot\|$ the 2-norm on $\mathbb{R}^n$ and also the induced 2-norm on $\mathbb{R}^{n\times n}$. Let $B_\delta(x)$ represent the open ball centered at $x \in \mathbb{R}^n$ of radius $\delta > 0$. Given $x \in \mathbb{R}^n$, x_i denotes the i-th component of x. For vectors $u \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$, the vector $(u; w) \in \mathbb{R}^{n+m}$ denotes their concatenation. For $A \in \mathbb{R}^{n\times n}$, we use $A \succeq 0$ (resp. $A \preceq 0$) to denote the fact that A is positive (resp. negative) semidefinite. The range and the null spaces of A are represented by $\mathrm{range}(A)$ and $\mathrm{null}(A)$, respectively. The eigenvalues of A are $\lambda_i(A)$ for $i \in \{1, \dots, n\}$. If A is symmetric, $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ represent the maximum and minimum eigenvalues, respectively. A set $\mathcal{S} \subset \mathbb{R}^n$ is *path connected* if for any two points $a, b \in \mathcal{S}$ there exists a continuous map $\gamma : [0, 1] \to \mathcal{S}$ such that $\gamma(0) = a$ and $\gamma(1) = b$. A set $\mathcal{S}_c \subset \mathcal{S} \subset \mathbb{R}^n$ is an *isolated path connected component* of $\mathcal{S}$ if it is path connected and there exists an open neighborhood $\mathcal{U}$ of $\mathcal{S}_c$ in $\mathbb{R}^n$ such that $\mathcal{U} \cap \mathcal{S} = \mathcal{S}_c$. For a real-valued function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, we denote the partial derivative of F with respect to the first argument by $\nabla_x F$ and with respect to the second argument by $\nabla_z F$. The higher-order derivatives follow the convention $\nabla_{xz} F = \frac{\partial^2 F}{\partial x \partial z}$, $\nabla_{xx} F = \frac{\partial^2 F}{\partial x^2}$, and so on. The restriction of $f : \mathbb{R}^n \to \mathbb{R}^m$ to a subset $\mathcal{S} \subset \mathbb{R}^n$ is denoted by $f_{|\mathcal{S}}$. The Jacobian of a continuously differentiable map $f : \mathbb{R}^n \to \mathbb{R}^m$ at $x \in \mathbb{R}^n$ is denoted $Df(x) \in \mathbb{R}^{m\times n}$. Finally, for a real-valued function $V : \mathbb{R}^n \to \mathbb{R}$ and $\alpha > 0$, we denote the sublevel set of V by $V^{-1}(\leq\alpha) = \{x \in \mathbb{R}^n \mid V(x) \leq \alpha\}$.

A. Proximal calculus

We present here a few notions on proximal calculus following [23]. Given a closed set $\mathcal{E} \subset \mathbb{R}^n$ and a point $x \in \mathbb{R}^n \backslash \mathcal{E}$, the distance from x to $\mathcal{E}$ is,

$$d_{\mathcal{E}}(x) = \min_{y \in \mathcal{E}} \|x - y\|. \tag{1}$$

We let $\mathrm{proj}_{\mathcal{E}}(x)$ denote the set of points in $\mathcal{E}$ that are closest to x, i.e., $\mathrm{proj}_{\mathcal{E}}(x) = \{y \in \mathcal{E} \mid \|x - y\| = d_{\mathcal{E}}(x)\} \subset \mathcal{E}$. For $y \in \mathrm{proj}_{\mathcal{E}}(x)$, the vector $x - y$ is a *proximal normal direction* to $\mathcal{E}$ at y and any nonnegative multiple $\zeta = t(x-y)$, $t \geq 0$ is called a *proximal normal* (P-normal) to $\mathcal{E}$ at y. The distance function $d_{\mathcal{E}}$ might not be differentiable in general (unless $\mathcal{E}$ is convex), but is globally Lipschitz and regular [23, p. 23]. For a locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$, the *generalized gradient* $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is

$$\partial f(x) = \mathrm{co}\{\lim_{i\to\infty} \nabla f(x_i) \mid x_i \to x, x_i \notin S \cup \Omega_f\},$$

where co denotes convex hull, $S \subset \mathbb{R}^n$ is any set of measure zero, and Ω_f is the set (of measure zero) of points where f is not differentiable. In the case of the distance function, one can compute [23, p. 99] the generalized gradient to be,

$$\partial d_{\mathcal{E}}(x) = \mathrm{co}\{x - y \mid y \in \mathrm{proj}_{\mathcal{E}}(x)\}. \tag{2}$$

B. Saddle points

Here, we provide basic definitions pertaining to the notion of saddle points. A point $(x_*, z_*) \in \mathbb{R}^n \times \mathbb{R}^m$ is a *(local) min-max saddle point* of a continuously differentiable function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ if there exist open neighborhoods $\mathcal{U}_{x_*} \subset \mathbb{R}^n$ of x_* and $\mathcal{U}_{z_*} \subset \mathbb{R}^m$ of z_* such that

$$F(x_*, z) \leq F(x_*, z_*) \leq F(x, z_*), \tag{3}$$

for all $z \in \mathcal{U}_{z_*}$ and $x \in \mathcal{U}_{x_*}$. The point (x_*, z_*) is a *global min-max saddle point* of F if $\mathcal{U}_{x_*} = \mathbb{R}^n$ and $\mathcal{U}_{z_*} = \mathbb{R}^m$. Min-max saddle points are a particular case of the more general notion of saddle points. We focus here on min-max saddle points motivated by problems in constrained optimization and zero-sum games, whose solutions correspond to min-max saddle points. With a slight abuse of terminology, throughout the paper we refer to them simply as saddle points. We denote by $\mathrm{Saddle}(F)$ the set of saddle points of F. From (3), for $(x_*, z_*) \in \mathrm{Saddle}(F)$, the point $x_* \in \mathbb{R}^n$ (resp. $z_* \in \mathbb{R}^m$) is a local minimizer (resp. local maximizer) of the map $x \mapsto F(x, z_*)$ (resp. $z \mapsto F(x_*, z)$). Each saddle point is a critical point of F, i.e., $\nabla_x F(x_*, z_*) = 0$ and $\nabla_z F(x_*, z_*) = 0$. Additionally, if F is twice continuously differentiable, then $\nabla_{xx} F(x_*, z_*) \preceq 0$ and $\nabla_{zz} F(x_*, z_*) \succeq 0$. Also, if $\nabla_{xx} F(x_*, z_*) \prec 0$ and $\nabla_{zz} F(x_*, z_*) \succ 0$, then the inequalities in (3) are strict.

A function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is *locally convex-*

concave at a point $(\tilde{x}, \tilde{z}) \in \mathbb{R}^n \times \mathbb{R}^m$ if there exists an open neighborhood $\mathcal{U}$ of $(\tilde{x}, \tilde{z})$ such that for all $(\bar{x}, \bar{z}) \in \mathcal{U}$, the functions $x \mapsto F(x, \bar{z})$ and $z \mapsto F(\bar{x}, z)$ are convex over $\mathcal{U} \cap (\mathbb{R}^n \times \{\bar{z}\})$ and concave over $\mathcal{U} \cap (\{\bar{x}\} \times \mathbb{R}^m)$, respectively. If in addition, either $x \mapsto F(x, \tilde{z})$ is strictly convex in an open neighborhood of $\tilde{x}$, or $z \mapsto F(\tilde{x}, z)$ is strictly concave in an open neighborhood of $\tilde{z}$, then F is *locally strictly convex-concave* at $(\tilde{x}, \tilde{z})$. F is locally (resp. locally strictly) convex-concave on a set $\mathcal{S} \subset \mathbb{R}^n \times \mathbb{R}^m$ if it is so at each point in $\mathcal{S}$. F is *globally convex-concave* if in the local definition $\mathcal{U} = \mathbb{R}^n \times \mathbb{R}^m$. Finally, F is *globally strictly convex-concave* if it is globally convex-concave and for any $(\bar{x}, \bar{z}) \in \mathbb{R}^n \times \mathbb{R}^m$, either $x \mapsto F(x, \bar{z})$ is strictly convex or $z \mapsto F(\bar{x}, z)$ is strictly concave.

III. Problem statement

Here we formulate the problem statement. Given a continuously differentiable function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, we consider its saddle-point dynamics (i.e., gradient-descent in one argument and gradient-ascent in the other),

$$\dot{x} = -\nabla_x F(x, z), \tag{4a}$$

$$\dot{z} = \nabla_z F(x, z). \tag{4b}$$

When convenient, we use the shorthand notation $X_{\text{sp}} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ to refer to this dynamics. Our aim is to provide conditions on F under which the trajectories of its saddle-point dynamics (4) locally asymptotically converge to its set of saddle points, and possibly to a point in the set. We are also interested in identifying conditions to establish global asymptotic convergence. Throughout our study, we assume that the set $\text{Saddle}(F)$ is nonempty. Our forthcoming discussion is divided in two threads, one for the case of convex-concave functions, cf. Section IV, and another for the case of general functions, cf. Section V. To deal with the latter, we use linearization techniques and also reason with proximal normals to the set of saddle points. In each case, illustrative examples show the applicability of the results.

IV. Convergence analysis for convex-concave functions

This section presents conditions for the asymptotic stability of saddle points under the saddle-point dynamics (4) that rely on the convexity-concavity properties of the function.

A. Stability under strict convexity-concavity

Our first result provides conditions that guarantee the local asymptotic stability of the set of saddle points.

Proposition 4.1: (Local asymptotic stability of the set of saddle points via convexity-concavity): For $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ continuously differentiable and locally strictly convex-concave on $\text{Saddle}(F)$, each isolated path connected component of $\text{Saddle}(F)$ is locally asymptotically stable under the saddle-point dynamics X_{sp} and, moreover, the convergence of each trajectory is to a point.

The result above shows that each saddle point is stable and that each path connected component of $\text{Saddle}(F)$ is asymptotically stable. Note that each saddle point might not be asymptotically stable. However, if a component consists of a single point, then that point is asymptotically stable. Interestingly, a close look at the proof of Proposition 4.1 reveals that, if the assumptions hold globally, then the asymptotic stability of the set of saddle points is also global, as stated next.

Corollary 4.2: (Global asymptotic stability of the set of saddle points via convexity-concavity): For $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ continuously differentiable and globally strictly convex-concave, $\text{Saddle}(F)$ is globally asymptotically stable under the saddle-point dynamics X_{sp} and the convergence of trajectories is to a point.

Remark 4.3: (Relationship with results on primal-dual dynamics): Corollary 4.2 is an extension to more general functions and less stringent assumptions of the results stated for Lagrangian functions of constrained convex (or concave) optimization problems in [10], [2], [5] and cost functions of differential games in [19]. In [2], [5], for a concave optimization, the matrix $\nabla_{xx} F$ is assumed to be negative definite at every saddle point and in [10] the set $\text{Saddle}(F)$ is assumed to be singleton. The work [19] assumes a sufficient condition on the cost functions to guarantee convergence that in the current setup is equivalent to having $\nabla_{xx} F$ and $\nabla_{zz} F$ positive and negative definite, respectively. •

B. Stability under convexity-linearity or linearity-concavity

In this section, we study the asymptotic convergence properties of the saddle-point dynamics when the convexity-concavity of the function is not strict but, instead, the function depends linearly on its second argument. The analysis follows analogously for functions that are linear in the first argument and concave in the other. The consideration of this class of functions is motivated by equality constrained optimization problems.

Proposition 4.4: (Local asymptotic stability of the set of saddle points via convexity-linearity): For continuously differentiable $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, if

(i) F is locally convex-concave on $\text{Saddle}(F)$ and linear in z,
(ii) for each $(x_*, z_*) \in \text{Saddle}(F)$, there exists a neighborhood $\mathcal{U}_{x_*} \subset \mathbb{R}^n$ of x_* where, if $F(x, z_*) = F(x_*, z_*)$ with $x \in \mathcal{U}_{x_*}$, then $(x, z_*) \in \text{Saddle}(F)$,

then each isolated path connected component of $\text{Saddle}(F)$ is locally asymptotically stable under the saddle-point dynamics X_{sp} and, moreover, the convergence of trajectories is to a point.

The next result extends the conclusions of Proposition 4.4 globally when the assumptions hold globally.

Corollary 4.5: (Global asymptotic stability of the set of saddle points via convexity-linearity): For continuously differentiable $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, if

(i) F is globally convex-concave and linear in z,
(ii) for each $(x_*, z_*) \in \operatorname{Saddle}(F)$, if $F(x, z_*) = F(x_*, z_*)$, then $(x, z_*) \in \operatorname{Saddle}(F)$,

then $\operatorname{Saddle}(F)$ is globally asymptotically stable under the saddle-point dynamics X_{sp} and, moreover, convergence of trajectories is to a point.

Example 4.6: (Saddle-point dynamics for convex optimization): Consider the following convex optimization problem on $\mathbb{R}^3$,

$$\text{minimize} \quad (x_1 + x_2 + x_3)^2, \tag{5a}$$

$$\text{subject to} \quad x_1 = x_2. \tag{5b}$$

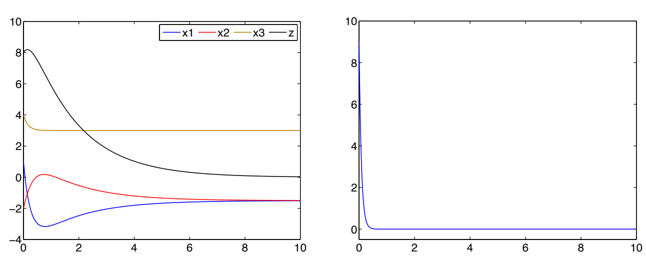
The set of solutions of this optimization is $\{x \in \mathbb{R}^3 \mid 2x_1 + x_3 = 0, x_2 = x_1\}$, with Lagrangian

$$L(x, z) = (x_1 + x_2 + x_3)^2 + z(x_1 - x_2), \tag{6}$$

where $z \in \mathbb{R}$ is the Lagrange multiplier. The set of saddle points of L (which correspond to the set of primal-dual solutions to (5)) are $\operatorname{Saddle}(L) = \{(x, z) \in \mathbb{R}^3 \times \mathbb{R} \mid 2x_1 + x_3 = 0, x_1 = x_2, \text{ and } z = 0\}$. However, L is not strictly convex-concave and hence, it does not satisfy the hypotheses of Corollary 4.2. While L is globally convex-concave and linear in z, it does not satisfy assumption (ii) of Corollary 4.5. Therefore, to identify a dynamics that renders $\operatorname{Saddle}(L)$ asymptotically stable, we form the augmented Lagrangian

$$\tilde{L}(x, z) = L(x, z) + (x_1 - x_2)^2, \tag{7}$$

that has the same set of saddle points as L. Note that $\tilde{L}$ is globally convex-concave (this can be seen by computing its Hessian) and is linear in z. Moreover, given any $(x_*, z_*) \in \operatorname{Saddle}(L)$, we have $\tilde{L}(x_*, z_*) = 0$, and if $\tilde{L}(x, z_*) = \tilde{L}(x_*, z_*) = 0$, then $(x, z_*) \in \operatorname{Saddle}(L)$. By Corollary 4.5, the trajectories of the saddle-point dynamics of $\tilde{L}$ converge to a point in $\mathcal{S}$ and hence, solve the optimization problem (5). Figure 1 illustrates this fact. •



(a) (x, z) (b) $(x_1 + x_2 + x_3)^2$

Fig. 1. (a) Trajectory of the saddle-point dynamics of the augmented Lagrangian $\tilde{L}$ in (7) for the optimization problem (5). The initial condition is $(x, z) = (1, -2, 4, 8)$. The trajectory converges to $(-1.5, -1.5, 3, 0) \in \operatorname{Saddle}(L)$. (b) Evolution of the objective function of the optimization (5) along the trajectory. The value converges to the minimum, 0.

Remark 4.7: (Relationship with results on primal-dual dynamics – cont'd): The work [5, Section 4] considers concave optimization problems under inequality constraints where the objective function is not strictly concave. The paper studies a discontinuous dynamics based on the saddle-point information of an augmented Lagrangian combined with a projection operator that restricts the dual variables to the nonnegative orthant. For the concave optimization problem formulated with equality constraints, we have verified that the augmented Lagrangian satisfies the hypotheses of Corollary 4.5, implying that the dynamics X_{sp} renders the primal-dual optima of the problem asymptotically stable. •

V. Convergence analysis for general functions

We study here the convergence properties of the saddle-point dynamics associated to functions that are not convex-concave. Our first result gives conditions for local asymptotic stability based on the linearization of the dynamics and properties of the eigenvalues of the Jacobian at the saddle points.

Proposition 5.1: (Local asymptotic stability of manifold of saddle points via linearization): Given $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, let $\mathcal{S} \subset \operatorname{Saddle}(F)$ be a p-dimensional manifold of saddle points. Assume F is thrice continuously differentiable on a neighborhood of $\mathcal{S}$ and that the Jacobian of X_{sp} at each point in $\mathcal{S}$ has no eigenvalues in the imaginary axis other than 0, which is semisimple with multiplicity p. Then, $\mathcal{S}$ is locally asymptotically stable under the saddle-point dynamics X_{sp} and the trajectories converge to a point.

Next, we provide a sufficient condition under which, the Jacobian of X_{sp} for a function F that is linear in z does not have any eigenvalue on the imaginary axis excluding 0.

Lemma 5.2: (Sufficient condition for absence of imaginary eigenvalues of the Jacobian of X_{sp}): Let F be linear in z. Then, the Jacobian of X_{sp} at any $(x_*, z_*) \in \operatorname{Saddle}(F)$ has no eigenvalues on the imaginary axis except for 0 if $\operatorname{range}(\nabla_{zx} F(x_*, z_*)) \cap \operatorname{null}(\nabla_{xx} F(x_*, z_*)) = \{0\}$.

The following example illustrates an application of the above two results to a nonconvex constrained optimization problem.

Example 5.3: (Saddle-point dynamics for nonconvex optimization): Consider the following constrained optimization on $\mathbb{R}^3$,

$$\text{minimize} \quad (\|x\| - 1)^2, \tag{8a}$$

$$\text{subject to} \quad x_3 = 0.5. \tag{8b}$$

The optimizers are $\{x \in \mathbb{R}^3 \mid x_3 = 0.5, x_1^2 + x_2^2 = 0.75\}$. The Lagrangian $L : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$

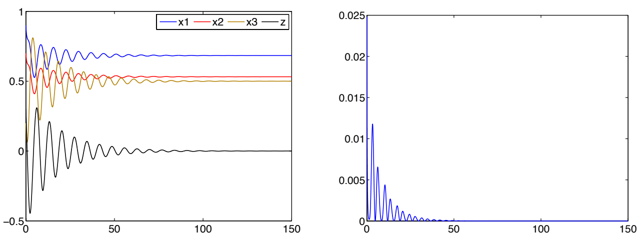
$$L(x, z) = (\|x\| - 1)^2 + z(x_3 - 0.5),$$

and its set of saddle points is the one-dimensional manifold $\operatorname{Saddle}(L) = \{(x, z) \in \mathbb{R}^3 \times \mathbb{R} \mid x_3 = 0.5, x_1^2 + x_2^2 = 0.75, z = 0\}$. The saddle-point dynamics of L takes the form

$$\dot{x} = -2\Big(1 - \frac{1}{\|x\|}\Big)x - [0, 0, z]^\top, \tag{9a}$$

$$\dot{z} = x_3 - 0.5. \tag{9b}$$

Note that $\mathrm{Saddle}(L)$ is nonconvex and that L is nonconvex in its first argument on any neighborhood of any saddle point. Therefore, results that rely on the convexity-concavity properties of L are not applicable to establish the asymptotic convergence of (9). This can, however, be established through Proposition 5.1 by observing that the Jacobian of X_{sp} at any point of $\mathrm{Saddle}(L)$ has 0 as an eigenvalue with multiplicity one and the rest of the eigenvalues have negative real parts as the hypotheses of Lemma 5.2 are met. Figure 2 illustrates in simulation the convergence of the trajectories to a saddle point. •



(a) (x, z) (b) $(x_1 + x_2 + x_3)^2$

Fig. 2. (a) Trajectory of the saddle-point dynamics (9) for the Lagrangian of the constrained optimization problem (8). The initial condition is $(x, z) = (0.9, 0.7, 0.2, 0.3)$. The trajectory converges to $(0.68, 0.53, 0.50, 0) \in \mathrm{Saddle}(L)$. (b) Evolution of the objective function of the optimization (8) along the trajectory. The value converges to the minimum, 0.

There are functions that do not satisfy the hypotheses of Proposition 5.1 whose saddle-point dynamics still seems to enjoy local asymptotic convergence properties. As an example, consider the function $F : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$,

$$F(x, z) = (\|x\| - 1)^4 - z^2\|x\|^2, \tag{10}$$

whose set of saddle points is the one-dimensional manifold $\mathrm{Saddle}(F) = \{(x, z) \in \mathbb{R}^2 \times \mathbb{R} \mid \|x\| = 1, z = 0\}$. The Jacobian of the saddle-point dynamics at any $(x, z) \in \mathrm{Saddle}(F)$ has -2 as an eigenvalue and 0 as the other eigenvalue, with multiplicity 2, which is greater than the dimension of $\mathcal{S}$ (and therefore Proposition 5.1 cannot be applied). Simulations show that the trajectories of the saddle-point dynamics asymptotically approach $\mathrm{Saddle}(S)$ if the initial condition is close enough to this set. Our next result allows us to formally establish this fact by studying the behavior of the function along the proximal normals to $\mathrm{Saddle}(F)$.

Proposition 5.4: (Asymptotic stability of manifold of saddle points via proximal normals): Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be twice continuously differentiable and $\mathcal{S} \subset \mathrm{Saddle}(F)$ be a closed set. Assume there exist constants $\lambda_M, k_1, k_2, \alpha_1, \beta_1 > 0$ and $L_x, L_z, \alpha_2, \beta_2 \geq 0$ such that, for every $(x_*, z_*) \in \mathcal{S}$ and every proximal normal $\eta = (\eta_x, \eta_z) \in \mathbb{R}^n \times \mathbb{R}^m$ to $\mathcal{S}$ at (x_*, z_*) with $\|\eta\| = 1$, it holds that the functions

$$[0, \lambda_M) \ni \lambda \mapsto F(x_* + \lambda\eta_x, z_*),$$
$$[0, \lambda_M) \ni \lambda \mapsto F(x_*, z_* + \lambda\eta_z),$$

are convex and concave, respectively, with

$$F(x_* + \lambda\eta_x, z_*) - F(x_*, z_*) \geq k_1\|\lambda\eta_x\|^{\alpha_1}, \tag{11a}$$

$$F(x_*, z_* + \lambda\eta_z) - F(x_*, z_*) \leq -k_2\|\lambda\eta_z\|^{\beta_1}, \tag{11b}$$

and, for all $\lambda \in [0, \lambda_M)$ and all $t \in [0, 1]$,

$$\begin{aligned}\|\nabla_{xz}F(x_*+t\lambda\eta_x, z_*+\lambda\eta_z) - \nabla_{xz}F(x_*+\lambda\eta_x, z_*+t\lambda\eta_z)\| \\ \leq L_x\|\lambda\eta_x\|^{\alpha_2} + L_z\|\lambda\eta_z\|^{\beta_2}.\end{aligned} \tag{12}$$

Then, $\mathcal{S}$ is locally asymptotically stable under the saddle-point dynamics X_{sp} if the following are true

$$\text{(a) either } L_x = 0 \text{ or } \alpha_1 \leq \alpha_2 + 1, \tag{13a}$$

$$\text{(b) either } L_z = 0 \text{ or } \beta_1 \leq \beta_2 + 1. \tag{13b}$$

Moreover, the convergence of the trajectories is to a point if every point of $\mathcal{S}$ is stable and the convergence is global if for every $\lambda_M \in \mathbb{R}_{\geq 0}$ there exist $k_1, k_2, \alpha_1, \beta_1 > 0$ such that along with $L_x = L_z = 0$ they satisfy the above hypotheses.

Intuitively, the hypotheses of Proposition 5.4 imply that along the proximal normal to the saddle set, the convexity (resp. concavity) in the x-coordinate (resp. z-coordinate) is 'stronger' than the influence of the x- and z-dynamics on each other, represented by the off-diagonal Hessian terms. When this coupling is absent (i.e., $\nabla_{xz}F \equiv 0$), the x- and z-dynamics are independent of each other and they function as individually aiming to minimize (resp. maximize) a function of one variable, thereby, reaching a saddle point. Note that the assumptions of Proposition 5.4 do not imply that F is locally convex-concave. As an example, we will see next that the F given in (10) satisfies the assumptions while it is not convex-concave in any neighborhood of any saddle point.

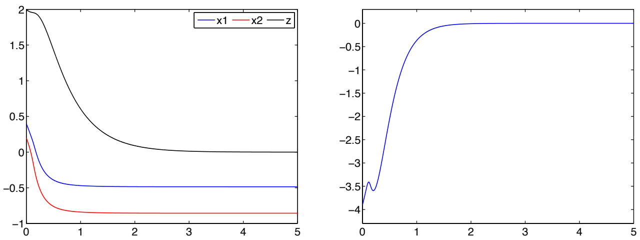
Example 5.5: (Convergence guarantee by proximal normal based analysis): Consider the function F defined in (10). Consider a saddle point $(x_*, z_*) = (\cos\theta, \sin\theta, 0) \in \mathrm{Saddle}(F)$, where $\theta \in [0, 2\pi)$. Let

$$\eta = (\eta_x, \eta_z) = ((a_1\cos\theta, a_1\sin\theta), a_2),$$

with $a_1, a_2 \in \mathbb{R}$ and $a_1^2 + a_2^2 = 1$, be a proximal normal to $\mathrm{Saddle}(F)$ at (x_*, z_*). Note that the function $\lambda \mapsto F(x_* + \lambda\eta_x, z_*) = (\lambda a_1)^4$ is convex, satisfying (11a) with $k_1 = 1$ and $\alpha_1 = 4$. The function $\lambda \mapsto F(x_*, z_* + \lambda\eta_z) = -(\lambda a_2)^2$ is concave, satisfying (11b) with $k_2 = 1$, $\beta_1 = 2$. Also, given any $\lambda_M > 0$, we can write

$$\begin{aligned}&\|\nabla_{xz}F(x_* + t\lambda\eta_x, z_* + \lambda\eta_z) - \nabla_{xz}F(x_* + \lambda\eta_x, z_* + t\lambda\eta_z)\| \\ &= \| -4(\lambda a_2)(1 + t\lambda a_1)\left(\begin{smallmatrix}\cos\theta \\ \sin\theta\end{smallmatrix}\right) + 4(t\lambda a_2)(1 + \lambda a_1)\left(\begin{smallmatrix}\cos\theta \\ \sin\theta\end{smallmatrix}\right)\|, \\ &\leq \|4(\lambda a_2)(1 + t\lambda a_1) - 4(t\lambda a_2)(1 + \lambda a_1)\|, \\ &\leq 8(1 + \lambda a_1)(\lambda a_2) \leq L_z(\lambda a_2),\end{aligned}$$

for $\lambda \leq \lambda_M$, where $L_z = 8(1 + \lambda_M a_1)$. This implies that $L_x = 0$, $L_z \neq 0$ and $\beta_2 = 1$. Therefore, the conditions (13) are satisfied and Proposition 5.4 establishes the asymptotic convergence of the saddle-point dynamics. Figure 3 illustrates this fact. Note that since $L_z \neq 0$, we cannot guarantee global convergence. •



(a) (x, z) (b) F

Fig. 3. (a) Trajectory of the saddle-point dynamics for the function defined by (10). The initial condition is $(x, z) = (0.1, 0.2, 4)$. The trajectory converges to $(0.49, 0.86, 0) \in \text{Saddle}(F)$. (b) Evolution of the function F along the trajectory. The value converges to 0, the value that the function takes on its saddle set.

Interestingly, Propositions 5.1 and 5.4 complement each other. The function (10) satisfies the hypotheses of Proposition 5.4 but not those of Proposition 5.1. Conversely, the Lagrangian of the optimization (8) satisfies the hypotheses of Proposition 5.1 but not those of Proposition 5.4.

VI. Conclusions

We have studied the asymptotic stability of the saddle-point dynamics associated to a continuously differentiable function. We have identified a set of complementary conditions under which the trajectories of the dynamics provably converge to the set of saddle points of the function and, wherever feasible, we have also established global stability guarantees. Our first class of convergence results is based on the convexity-concavity properties of the function defining the dynamics. When these properties are not met, our second class of results explore the existence of convergence guarantees using linearization techniques and the properties of the function along proximal normals to the set of saddle points. Several examples illustrate the applicability of our results, with special attention to finding primal-dual solutions of constrained optimization problems. Future work will characterize the robustness properties of the dynamics against disturbances, study the case of nonsmooth functions (where the associated saddle-point dynamics takes the form of a differential inclusion involving the generalized gradient of the function), and explore the application of our results to optimization problems with inequality constraints. We also plan to build on our results to synthesize distributed algorithmic solutions for networked optimization problems in power networks.

VII. Acknowledgements

This research was partially supported by NSF award ECCS-1307176. The authors would like to thank Prof. Bahman Gharesifard for multiple discussions on the saddle-point dynamics.

References

[1] W. M. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press, 1974.

[2] K. Arrow, L. Hurwitz, and H. Uzawa, *Studies in Linear and Non-Linear Programming*. Stanford, California: Stanford University Press, 1958.

[3] R. Dorfman, P. A. Samuelson, and R. Solow, *Linear programming in economic analysis*. New York, Toronto, and London: McGraw Hill, 1958.

[4] T. Kose, "Solutions of saddle value problems by differential equations," *Econometrica*, vol. 24, no. 1, pp. 59–70, 1956.

[5] D. Feijer and F. Paganini, "Stability of primal-dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, pp. 1974–1981, 2010.

[6] A. Cherukuri, E. Mallada, and J. Cortés, "Asymptotic convergence of primal-dual dynamics," *Systems & Control Letters*, 2015. Submitted.

[7] A. Nedić and A. Ozdgalar, "Subgradient methods for saddle-point problems," *Journal of Optimization Theory & Applications*, vol. 142, no. 1, pp. 205–228, 2009.

[8] H. B. Durr and C. Ebenbauer, "A smooth vector field for saddle point problems," in *IEEE Conf. on Decision and Control*, (Orlando, Florida), pp. 4654–4660, Dec. 2011.

[9] T. Holding and I. Lestas, "On the convergence of saddle points of concave-convex functions, the gradient method and emergence of oscillations," in *IEEE Conf. on Decision and Control*, (Los Angeles, CA), pp. 1143–1148, 2014.

[10] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *IEEE Conf. on Decision and Control*, (Orlando, Florida), pp. 3800–3805, 2011.

[11] B. Gharesifard and J. Cortés, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 781–786, 2014.

[12] G. Droge and M. Egerstedt, "Proportional integral distributed optimization for dynamic network topologies," in *American Control Conference*, (Portland, OR), pp. 3621–3626, June 2014.

[13] D. Richert and J. Cortés, "Robust distributed linear programming," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, 2015. To appear. Available at `http://carmenere.ucsd.edu/jorge`.

[14] X. Ma and N. Elia, "A distributed continuous-time gradient dynamics approach for the active power loss minimizations," in *Allerton Conf. on Communications, Control and Computing*, (Monticello, IL), pp. 100–106, Oct. 2013.

[15] X. Zhang and A. Papachristodoulou, "A real-time control framework for smart power networks with star topology," in *American Control Conference*, (Washington, DC), pp. 5062–5067, June 2013.

[16] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.

[17] D. Richert and J. Cortés, "Distributed bargaining in dyadic-exchange networks," *IEEE Transactions on Control of Network Systems*, 2014. Conditionally accepted.

[18] T. Başar and G. Oldser, *Dynamic Noncooperative Game Theory*. Academic Press, 1982.

[19] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "Characterization and computation of local nash equilibrium in continuous games," in *Allerton Conf. on Communications, Control and Computing*, (Monticello, IL), Oct. 2013.

[20] J. Hofbauer and S. Sorin, "Best response dynamics for continuous zero-sum games," *Discrete and Continuous Dynamical Systems Ser. B*, vol. 6, no. 1, pp. 215–224, 2006.

[21] E. N. Barron, R. Goebel, and R. R. Jensen, "Best response dynamics for continuous games," *Proceeding of the American Mathematical Society*, vol. 138, no. 3, pp. 1069–1083, 2010.

[22] P.-A. Absil and K. Kurdyka, "On the stable equilibrium points of gradient systems," *Systems & Control Letters*, vol. 55, pp. 573–577, 2006.

[23] F. H. Clarke, Y. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, vol. 178 of *Graduate Texts in Mathematics*. Springer, 1998.