

The role of strong convexity-concavity in the convergence and robustness of the saddle-point dynamics

Ashish Cherukuri Enrique Mallada Steven Low Jorge Cortés

Abstract—This paper studies the projected saddle-point dynamics for a twice differentiable convex-concave function, which we term saddle function. The dynamics consists of gradient descent of the saddle function in variables corresponding to convexity and (projected) gradient ascent in variables corresponding to concavity. We provide a novel characterization of the omega-limit set of the trajectories of these dynamics in terms of the diagonal Hessian blocks of the saddle function. Using this characterization, we establish global asymptotic convergence of the dynamics under local strong convexity-concavity of the saddle function. If this property is global, and for the case when the saddle function takes the form of the Lagrangian of an equality constrained optimization problem, we establish the input-to-state stability of the saddle-point dynamics by providing an ISS Lyapunov function. Various examples illustrate our results.

I. INTRODUCTION

Saddle-point dynamics and its variations have been used extensively in the design and analysis of distributed feedback controllers and optimization algorithms in several applications, including power networks, network flow problems, and zero-sum games. The analysis of the global convergence of these dynamics typically relies on some global strong/strict convexity-concavity property of the saddle function that defines the dynamics. The main aim of this paper is to refine this analysis by unveiling two ways in which convexity-concavity of the saddle function plays a role. First, we show that local strong convexity-concavity is enough to conclude global asymptotic convergence, thus generalizing previous results that rely on global strong/strict convexity-concavity to conclude global convergence. Second, we show that global strong convexity-concavity in turn implies a stronger form of convergence, that is, input-to-state stability (ISS) of the dynamics. This property has important implications for the practical implementation of the saddle-point dynamics and its robustness against a wide variety of disturbances.

Literature review

The analysis of the convergence properties of (projected) saddle-point dynamics to the set of saddle points goes back to [1], [2], motivated by applications in nonlinear programming and optimization. These works employ direct meth-

ods, examining the approximate evolution of the distance of the trajectories to the saddle point along the dynamics and concluding attractivity by showing it to be decreasing. Motivated by network optimization problems, more recent work [3], [4] has employed indirect, LaSalle-type arguments to analyze the asymptotic convergence properties of these dynamics. For this class of problems, the aggregate nature of the objective function and the local computability of the constraints make the saddle-point dynamics corresponding to the Lagrangian naturally distributed. Many other works exploit this strategy to solve network optimization problems for various applications, e.g., distributed convex optimization [4], [5], distributed linear programming [6], bargaining problems [7], and power network [8], [9], [10]. Another area of application is game theory, where saddle-point dynamics is applied to find the Nash equilibria of two-person zero-sum games [11], [12]. In the context of distributed optimization, the recent work [13] employs a (strict) Lyapunov function approach to ensure asymptotic convergence of saddle-point-like dynamics. The work [14] examines the asymptotic behavior of the saddle-point dynamics when the set of saddle points is not asymptotically stable and, instead, the trajectories exhibit oscillatory behavior. Our previous work has established global asymptotic convergence of the saddle-point dynamics [15] and the projected saddle-point dynamics [16] under global strict convexity-concavity assumptions. The works mentioned above require similar or stronger global assumptions on the convexity-concavity properties of the saddle function to ensure convergence. By contrast, here we are able to guarantee global convergence under local strong convexity-concavity assumptions. Furthermore, we show that under global strong convexity-concavity, the saddle-point dynamics is ISS by identifying an ISS Lyapunov function.

Statement of contributions

We start with the definition of the projected saddle-point dynamics for a differentiable convex-concave function, which we refer to as saddle function. The dynamics has three components: gradient descent, projected gradient ascent, and gradient ascent of the saddle function (where each gradient is with respect to a subset of the arguments of the function). Our contributions highlight the effect of the convexity-concavity properties of the saddle function in the convergence analysis, namely, that local properties imply global convergence and global properties imply robustness. Our first contribution is a novel characterization of the omega-limit set of the trajectories of the projected saddle-

Ashish Cherukuri and Jorge Cortés are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, {acheruku, cortes}@ucsd.edu, Enrique Mallada is with the Department of Electrical and Computer Engineering, Johns Hopkins University mallada@hopkins.edu, and Steven Low is with the Computing and Mathematical Sciences and the Electrical Engineering Department, California Institute of Technology, slow@caltech.edu

point dynamics in terms of the diagonal Hessian blocks of the saddle function. To this end, we use the distance to a saddle point as a LaSalle function, express the Lie derivative of this function in terms of Hessian blocks, and show it is negative using second-order properties of the saddle function. Building on this characterization, our second contribution establishes global asymptotic convergence of the projected saddle-point dynamics to a saddle point assuming only local strong convexity-concavity of the saddle function. If the convexity-concavity property is global, and for the case of saddle functions of the form of a Lagrangian of an equality constrained optimization problem, our third contribution establishes the input-to-state stability properties of the saddle-point dynamics with respect to the set of saddle points. We illustrate our results in several examples. Proofs are omitted for reasons of space and will appear elsewhere.

Organization

Section II contains notation and preliminaries. Section III introduces the projected saddle-point dynamics and the problem statement. Section IV analyzes the asymptotic convergence of the projected saddle-point dynamics using local properties of the saddle function. Section V studies the input-to-state stability of the saddle-point dynamics. Finally, Section VI summarizes our conclusions and ideas for future work.

II. PRELIMINARIES

This section introduces our notation and preliminary notions on convex-concave functions, discontinuous dynamical systems, and input-to-state stability.

A. Notation

Let \mathbb{R} , $\mathbb{R}_{\geq 0}$ denote the set of real, nonnegative real numbers, respectively. We let $\|\cdot\|$ denote the 2-norm on \mathbb{R}^n and the respective induced norm on $\mathbb{R}^{n \times m}$. Given $x, y \in \mathbb{R}^n$, x_i denotes the i -th component of x , and $x \leq y$ denotes $x_i \leq y_i$ for $i \in \{1, \dots, n\}$. 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 $a, b \in \mathbb{R}$, we let

$$[a]_b^+ = \begin{cases} a, & \text{if } b > 0, \\ \max\{0, a\}, & \text{if } b = 0. \end{cases}$$

For vectors $a, b \in \mathbb{R}^n$, $[a]_b^+$ denotes the vector whose i -th component is $[a_i]_b^+$, for $i \in \{1, \dots, n\}$. Given a set $\mathcal{S} \subset \mathbb{R}^n$, we denote by $\text{cl}(\mathcal{S})$, $\text{int}(\mathcal{S})$, and $|\mathcal{S}|$ its closure, interior, and cardinality, respectively. The distance of a point $x \in \mathbb{R}^n$ to the set $\mathcal{S} \subset \mathbb{R}^n$ in 2-norm is $\|x\|_{\mathcal{S}} = \inf_{y \in \mathcal{S}} \|x - y\|$. The projection of x onto a closed set \mathcal{S} is defined as the set $\text{proj}_{\mathcal{S}}(x) = \{y \in \mathcal{S} \mid \|x - y\| = \|x\|_{\mathcal{S}}\}$. When \mathcal{S} is convex, $\text{proj}_{\mathcal{S}}(x)$ is a singleton for any $x \in \mathbb{R}^n$. For a matrix $A \in \mathbb{R}^{n \times n}$, we use $A \succeq 0$, $A \succ 0$, $A \preceq 0$, and $A \prec 0$ to denote that A is positive semidefinite, positive definite, negative semidefinite, and negative definite, respectively. For

a symmetric matrix $A \in \mathbb{R}^{n \times n}$, $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the minimum and maximum eigenvalue of A . For a real-valued function $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $(x, y) \mapsto F(x, y)$, we denote by $\nabla_x F$ and $\nabla_y F$ the partial derivative of F with respect to the first and second arguments, respectively. Higher-order derivatives follow the convention $\nabla_{xy} F = \frac{\partial^2 F}{\partial x \partial y}$, $\nabla_{xx} F = \frac{\partial^2 F}{\partial x^2}$, and so on. A function $\alpha : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is class \mathcal{K} if it is continuous, strictly increasing, and $\alpha(0) = 0$. The set of unbounded class \mathcal{K} functions are called \mathcal{K}_{∞} functions. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is class \mathcal{KL} if for any $t \in \mathbb{R}_{\geq 0}$, $x \mapsto \beta(x, t)$ is class \mathcal{K} and for any $x \in \mathbb{R}_{\geq 0}$, $t \mapsto \beta(x, t)$ is continuous, decreasing with $\beta(t) \rightarrow 0$ as $t \rightarrow \infty$.

B. Saddle points and convex-concave functions

Here, we review notions on convexity, concavity, and saddle points from [17]. A function $f : \mathcal{X} \rightarrow \mathbb{R}$ is *convex* if

$$f(\lambda x + (1 - \lambda)x') \leq \lambda f(x) + (1 - \lambda)f(x'),$$

for all $x, x' \in \mathcal{X}$ and all $\lambda \in [0, 1]$. A convex differentiable function f satisfies the following *first-order convexity condition*

$$f(x') \geq f(x) + (x' - x)^{\top} \nabla f(x),$$

for all $x, x' \in \mathcal{X}$. A twice differentiable function f is *locally strongly convex* at $x \in \mathcal{X}$ if f is convex and $\nabla^2 f(x) \succeq mI$ for some $m > 0$. Moreover, a twice differentiable f is *strongly convex* if $\nabla^2 f(x) \succeq mI$ for all $x \in \mathcal{X}$ for some $m > 0$. A function $f : \mathcal{X} \rightarrow \mathbb{R}$ is *concave*, *locally strongly concave*, or *strongly concave* if $-f$ is convex, locally strongly convex, or strongly convex, respectively. A function $F : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ is *convex-concave* (on $\mathcal{X} \times \mathcal{Y}$) if, given any point $(\tilde{x}, \tilde{y}) \in \mathcal{X} \times \mathcal{Y}$, $x \mapsto F(x, \tilde{y})$ is convex and $y \mapsto F(\tilde{x}, y)$ is concave. When the space $\mathcal{X} \times \mathcal{Y}$ is clear from the context, we refer to this property as F being convex-concave in (x, y) . A function F is *locally strongly convex-concave* at a saddle point (x, y) if it is convex-concave and either $\nabla_{xx} F(x, y) \succeq mI$ or $\nabla_{yy} F(x, y) \preceq -mI$ for some $m > 0$. A point $(x_*, y_*) \in \mathcal{X} \times \mathcal{Y}$ is a *saddle point* of F on the set $\mathcal{X} \times \mathcal{Y}$ if $F(x_*, y) \leq F(x_*, y_*) \leq F(x, y_*)$, for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The set of saddle points of a convex-concave function F is convex.

C. Discontinuous dynamical systems

Here we present notions on discontinuous dynamical systems [18], [19]. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be Lebesgue measurable and locally bounded. Consider the differential equation

$$\dot{x} = f(x). \quad (1)$$

A map $\gamma : [0, T) \rightarrow \mathbb{R}^n$ is a (*Caratheodory*) *solution* of (1) on the interval $[0, T)$ if it is absolutely continuous on $[0, T)$ and satisfies $\dot{\gamma}(t) = f(\gamma(t))$ almost everywhere in $[0, T)$. We use the terms *solution* and *trajectory* interchangeably. A set $\mathcal{S} \subset \mathbb{R}^n$ is *invariant* under (1) if every solution starting from any point in \mathcal{S} remains in \mathcal{S} . For a solution γ of (1)

defined on the time interval $[0, \infty)$, the *omega-limit* set $\Omega(\gamma)$ is defined by

$$\Omega(\gamma) = \{y \in \mathbb{R}^n \mid \exists \{t_k\}_{k=1}^\infty \subset [0, \infty) \text{ with } \lim_{k \rightarrow \infty} t_k = \infty \\ \text{and } \lim_{k \rightarrow \infty} \gamma(t_k) = y\}.$$

If the solution γ is bounded, then $\Omega(\gamma) \neq \emptyset$ by the Bolzano-Weierstrass theorem [20, p. 33]. Given a continuously differentiable function $V : \mathbb{R}^n \rightarrow \mathbb{R}$, the *Lie derivative of V along (1)* at $x \in \mathbb{R}^n$ is $\mathcal{L}_f V(x) = \nabla V(x)^\top f(x)$. The next result is a simplified version of [18, Proposition 3].

Proposition 2.1: (Invariance principle for discontinuous Caratheodory systems): Let $\mathcal{S} \in \mathbb{R}^n$ be compact and invariant. Assume that, for each point $x_0 \in \mathcal{S}$, there exists a unique solution of (1) starting at x_0 and that its omega-limit set is invariant too. Let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable map such that $\mathcal{L}_f V(x) \leq 0$ for all $x \in \mathcal{S}$. Then, any solution of (1) starting at \mathcal{S} converges to the largest invariant set in $\text{cl}(\{x \in \mathcal{S} \mid \mathcal{L}_f V(x) = 0\})$.

D. Input-to-state stability

Here, we review the notion of input-to-state stability (ISS) following [21]. Consider a system

$$\dot{x} = f(x, u), \quad (2)$$

where $x \in \mathbb{R}^n$ is the state, $u : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^m$ is the input that is measurable and locally essentially bounded, and $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is locally Lipschitz. Assume that starting from any point in \mathbb{R}^n , the trajectory of (2) is defined on $\mathbb{R}_{\geq 0}$ for any given control. Let $\text{Eq}(f) \subset \mathbb{R}^n$ be the set of equilibrium points of the unforced system. Then, the system (2) is *input-to-state stable* (ISS) with respect to $\text{Eq}(f)$ if there exists $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}$ such that each trajectory $t \mapsto x(t)$ of (2) satisfies

$$\|x(t)\|_{\text{Eq}(f)} \leq \beta(\|x(0)\|_{\text{Eq}(f)}, t) + \gamma(\|u\|_\infty)$$

for all $t \geq 0$, where $\|u\|_\infty = \text{ess sup}_{t \geq 0} \|u(t)\|$ is the essential supremum (see [20, p. 185] for the definition) of u . This notion captures the graceful degradation of the asymptotic convergence properties of the unforced system as the size of the disturbance input grows. One convenient way of showing ISS is by finding an ISS-Lyapunov function for the dynamics. An *ISS-Lyapunov function* with respect to the set $\text{Eq}(f)$ for system (2) is a differentiable function $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ such that

(i) there exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ such that for all $x \in \mathbb{R}^n$,

$$\alpha_1(\|x\|_{\text{Eq}(f)}) \leq V(x) \leq \alpha_2(\|x\|_{\text{Eq}(f)}); \quad (3)$$

(ii) there exists a continuous, positive definite function $\alpha_3 : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ and $\gamma \in \mathcal{K}_\infty$ such that

$$\nabla V(x)^\top f(x, v) \leq -\alpha_3(\|x\|_{\text{Eq}(f)}) \quad (4)$$

for all $x \in \mathbb{R}^n$, $v \in \mathbb{R}^m$ for which $\|x\|_{\text{Eq}(f)} \geq \gamma(\|v\|)$.

Proposition 2.2: (Existence of ISS-Lyapunov function im-

plies ISS): If the system (2) admits an ISS-Lyapunov function, then it is ISS.

III. PROBLEM STATEMENT

In this section, we provide a formal statement of the problem of interest. Consider a twice continuously differentiable function $F : \mathbb{R}^n \times \mathbb{R}_{\geq 0}^p \times \mathbb{R}^m \rightarrow \mathbb{R}$, $(x, y, z) \mapsto F(x, y, z)$, which we refer to as *saddle function*. With the notation of Section II-B, we set here $\mathcal{X} = \mathbb{R}^n$ and $\mathcal{Y} = \mathbb{R}_{\geq 0}^p \times \mathbb{R}^m$, and assume that F is convex-concave on $(\mathbb{R}^n) \times (\mathbb{R}_{\geq 0}^p \times \mathbb{R}^m)$. Let $\text{Saddle}(F)$ denote its set of saddle points, which we assume non-empty. We define the *projected saddle-point dynamics* for F as

$$\dot{x} = -\nabla_x F(x, y, z), \quad (5a)$$

$$\dot{y} = [\nabla_y F(x, y, z)]_y^+, \quad (5b)$$

$$\dot{z} = \nabla_z F(x, y, z). \quad (5c)$$

When convenient, we use the map $X_{p\text{-sp}} : \mathbb{R}^n \times \mathbb{R}_{\geq 0}^p \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m$ to refer to the dynamics (5). Note that the domain $\mathbb{R}^n \times \mathbb{R}_{\geq 0}^p \times \mathbb{R}^m$ is invariant under $X_{p\text{-sp}}$ (this follows from the definition of the projection operator) and its set of equilibrium points precisely corresponds to $\text{Saddle}(F)$ (this follows from the defining property of saddle points and the first-order condition for convexity-concavity of F). Thus, a saddle point (x_*, y_*, z_*) satisfies

$$\nabla_x F(x_*, y_*, z_*) = 0, \quad \nabla_z F(x_*, y_*, z_*) = 0, \quad (6a)$$

$$\nabla_y F(x_*, y_*, z_*) \leq 0, \quad y_*^\top \nabla_y F(x_*, y_*, z_*) = 0. \quad (6b)$$

Our interest in the dynamics (5) is motivated by two bodies of work in the literature: one that analyzes primal-dual dynamics, corresponding to (5a) together with (5b), for solving inequality constrained network optimization problems, see e.g., [2], [3], [22], [9]; and the other one analyzing saddle-point dynamics, corresponding to (5a) together with (5c), for solving equality constrained problems and finding Nash equilibrium of zero-sum games, see e.g., [15] and references therein. By considering (5a)-(5c) together, we seek to unify these two lines of work.

Our main objectives are to identify conditions that guarantee that the set of saddle points is globally asymptotically stable under the dynamics (5) and formally characterize the robustness properties using the concept of input-to-state stability.

IV. PROJECTED SADDLE-POINT DYNAMICS: LOCAL PROPERTIES IMPLY GLOBAL CONVERGENCE

Our first result of this section characterizes the omega-limit set of the trajectories of the projected saddle-point dynamics (5) in terms of second-order information of the saddle function.

Proposition 4.1: (Characterization of the omega-limit set of solutions of $X_{p\text{-sp}}$): Given a convex-concave function F , the set $\text{Saddle}(F)$ is stable under the projected saddle-point

dynamics $X_{p\text{-sp}}$ and the omega-limit set of every solution is contained in the largest invariant set \mathcal{M} in $\mathcal{E}(F)$, where

$$\begin{aligned} \mathcal{E}(F) = \{ & (x, y, z) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}^p \times \mathbb{R}^m \mid \\ & (x - x_*; y - y_*; z - z_*) \in \ker(\overline{H}(x, y, z, x_*, y_*, z_*)), \\ & \text{for all } (x_*, y_*, z_*) \in \text{Saddle}(F)\}, \end{aligned} \quad (7)$$

and

$$\begin{aligned} \overline{H}(x, y, z, x_*, y_*, z_*) &= \int_0^1 H(x(s), y(s), z(s)) ds, \\ (x(s), y(s), z(s)) &= (x_*, y_*, z_*) + s(x - x_*, y - y_*, z - z_*), \\ H(x, y, z) &= \begin{bmatrix} -\nabla_{xx}F & 0 & 0 \\ 0 & \nabla_{yy}F & \nabla_{yz}F \\ 0 & \nabla_{zy}F & \nabla_{zz}F \end{bmatrix}_{(x,y,z)}. \end{aligned} \quad (8)$$

The above result uses LaSalle Invariance principle by showing that the evolution of the following function is nonincreasing along the trajectories

$$V_1(x, y, z) = \frac{1}{2} (\|x - x_*\|^2 + \|y - y_*\|^2 + \|z - z_*\|^2). \quad (9)$$

The next result states that local strong convexity-concavity around a saddle point together with global convexity-concavity of the saddle function are enough to guarantee global convergence. The proof uses the above derived characterization of the omega-limit set.

Theorem 4.2: (Global asymptotic stability of the set of saddle points under $X_{p\text{-sp}}$): Given a convex-concave function F which is locally strongly convex-concave at a saddle point, the set $\text{Saddle}(F)$ is globally asymptotically stable under the projected saddle-point dynamics $X_{p\text{-sp}}$ and the convergence of trajectories is to a point.

The conclusion of the above result hold for slightly weaker set of conditions on the saddle function. In particular, F need only be twice continuously differentiable in a neighborhood of the saddle point and the local strong convexity-concavity can be relaxed to a condition on the line integral of Hessian blocks of F . We state next this stronger result.

Theorem 4.3: (Global asymptotic stability of the set of saddle points under $X_{p\text{-sp}}$): Let F be convex-concave and continuously differentiable with locally Lipschitz gradient. Suppose there exists a saddle point (x_*, y_*, z_*) and a neighborhood of this point $\mathcal{U}_* \subset \mathbb{R}^n \times \mathbb{R}_{\geq 0}^p \times \mathbb{R}^m$ such that F is twice continuously differentiable on \mathcal{U}_* and either of the following is true

(i) for all $(x, y, z) \in \mathcal{U}_*$,

$$\int_0^1 \nabla_{xx}F(x(s), y(s), z(s)) ds \succ 0,$$

(ii) for all $(x, y, z) \in \mathcal{U}_*$,

$$\int_0^1 \begin{bmatrix} \nabla_{yy}F & \nabla_{yz}F \\ \nabla_{zy}F & \nabla_{zz}F \end{bmatrix}_{(x(s), y(s), z(s))} ds \prec 0,$$

where $(x(s), y(s), z(s))$ are given in (8). Then, $\text{Saddle}(F)$ is globally asymptotically stable under the projected saddle-

point dynamics $X_{p\text{-sp}}$ and the convergence of trajectories is to a point.

Next, we illustrate the application of the above result with an example.

Example 4.4: (Illustration of global asymptotic convergence): Consider $F : \mathbb{R}^2 \times \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{R}$ given as

$$F(x, y, z) = f(x) + y(-x_1 - 1) + z(x_1 - x_2), \quad (10)$$

where

$$f(x) = \begin{cases} \|x\|^4, & \text{if } \|x\| \leq \frac{1}{2}, \\ \frac{1}{16} + \frac{1}{2}(\|x\| - \frac{1}{2}), & \text{if } \|x\| \geq \frac{1}{2}. \end{cases}$$

Note that F is convex-concave on $(\mathbb{R}^2) \times (\mathbb{R}_{\geq 0} \times \mathbb{R})$ and $\text{Saddle}(F) = \{0\}$. Also, F is continuously differentiable on the entire domain and its gradient is locally Lipschitz. Finally, F is twice continuously differentiable on the neighborhood $\mathcal{U}_* = B_{1/2}(0) \cap (\mathbb{R}^2 \times \mathbb{R}_{\geq 0} \times \mathbb{R})$ of the saddle point 0 and hypothesis (i) of Theorem 4.3 holds on \mathcal{U}_* . Therefore, we conclude from Theorem 4.3 that the trajectories of the projected saddle-point dynamics of F converge globally asymptotically to the saddle point 0. Figure 1 shows an execution. •

Remark 4.5: (Comparison with the literature): Theorems 4.2 and 4.3 complement the available results in the literature concerning the asymptotic convergence properties of saddle-point [2], [15], [13] and primal-dual dynamics [3], [16]. The former dynamics corresponds to (5) when the variable y is absent and the later to (5) when the variable z is absent. For both these dynamics, existing global asymptotic stability results require assumptions on the global properties of F , in addition to the global convexity-concavity of F . For example, global strong convexity-concavity [2], global strict convexity-concavity, and its generalizations [15]. In contrast, the novelty of our results lies in establishing that certain local properties of the saddle function are enough to guarantee global asymptotic convergence of the projected saddle-point dynamics. •

V. INPUT-TO-STATE STABILITY OF THE SADDLE-POINT DYNAMICS

In this section we show that, when the saddle function possesses other global properties in addition to being convex-concave, cf. Remark 4.5, the associated saddle-point dynamics in fact enjoys robustness properties beyond just global asymptotic convergence. To this end, throughout the section, we consider saddle functions corresponding to the Lagrangian of an equality-constrained optimization problem,

$$F(x, z) = f(x) + z^\top (Ax - b), \quad (11)$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $f : \mathbb{R}^n \rightarrow \mathbb{R}$. The reason behind this focus is that, in this case, the dynamics (5) is smooth, and this allows to directly bring to bear the theory of input-to-state stability outlined in Section II-D. The projected saddle-point dynamics (5) for the class of saddle functions

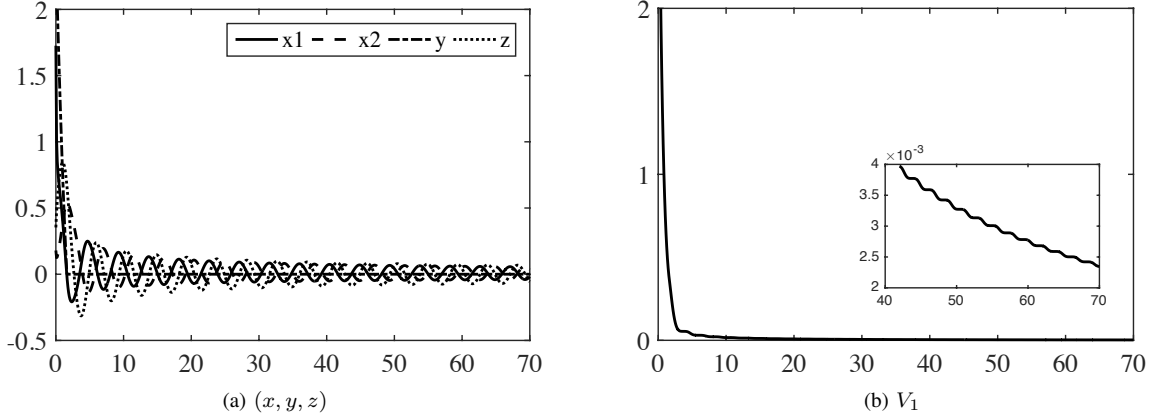


Fig. 1. Execution of the projected saddle-point dynamics (5) starting from $(1.7256, 0.1793, 2.4696, 0.3532)$ for Example 4.4. As guaranteed by Theorem 4.3, the trajectory converges to the unique saddle point 0 and the function V_1 defined in (9) decreases monotonically.

given in (11) takes the form

$$\dot{x} = -\nabla_x F(x, z) = -\nabla f(x) - A^\top z, \quad (12a)$$

$$\dot{z} = \nabla_z F(x, z) = Ax - b, \quad (12b)$$

corresponding to equations (5a) and (5c). We term these simply *saddle-point dynamics* and denote it as $X_{\text{sp}} : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$.

Our aim in this section is to establish that the saddle-point dynamics is ISS with respect to the set $\text{Saddle}(F)$ when disturbance inputs affect it additively. Disturbance inputs can arise when implementing the dynamics as a controller of a physical system because of a variety of malfunctions, including errors in the gradient computation, noise in state measurements, and errors in the controller implementation. In such scenarios, the following result shows that the dynamics (12) exhibits a graceful degradation of its convergence properties, one that scales with the size of the disturbance.

Theorem 5.1: (ISS of saddle-point dynamics): Let the saddle function F be of the form (11), with f strongly convex, twice continuously differentiable, and satisfying $mI \preceq \nabla^2 f(x) \preceq MI$ for all $x \in \mathbb{R}^n$ and some constants $0 < m \leq M < \infty$. Then, the dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -\nabla_x F(x, z) \\ \nabla_z F(x, z) \end{bmatrix} + \begin{bmatrix} u_x \\ u_z \end{bmatrix}, \quad (13)$$

where $(u_x, u_z) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ is a measurable and locally essentially bounded map, is ISS with respect to $\text{Saddle}(F)$.

We omit the proof for space reasons, but provide here a brief outline. For notational convenience, we refer to (13) by $X_{\text{sp}}^p : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$. The proof essentially pconsists of establishing that the function $V_2 : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}$,

$$V_2(x, z) = \frac{\beta_1}{2} \|X_{\text{sp}}(x, z)\|^2 + \frac{\beta_2}{2} \|(x, z)\|_{\text{Saddle}(F)}^2 \quad (14)$$

with $\beta_1 > 0$, $\beta_2 = \frac{4\beta_1 M^4}{m^2}$, is an ISS-Lyapunov function with respect to $\text{Saddle}(F)$ for X_{sp}^p . The statement then directly

follows from Proposition 2.2.

Remark 5.2: (Relaxing global bounds on Hessian of f): The assumption of the Hessian of f in Theorem 5.1 is restrictive, but there are functions other than quadratic that satisfy it, see e.g. [23, Section 6], that satisfy it. We conjecture that the global upper bound on the Hessian can be relaxed by resorting to the notion of semiglobal ISS. •

The above result has the following consequence.

Corollary 5.3: (Lyapunov function for saddle-point dynamics): Let the saddle function F be of the form (11), with f strongly convex, twice continuously differentiable, and satisfying $mI \preceq \nabla^2 f(x) \preceq MI$ for all $x \in \mathbb{R}^n$ and some constants $0 < m \leq M < \infty$. Then, the function V_2 (14) is a Lyapunov function with respect to the set $\text{Saddle}(F)$ for the saddle-point dynamics (12).

Remark 5.4: (ISS with respect to $\text{Saddle}(F)$ does not imply bounded trajectories): Note that Theorem 5.1 bounds only the distance of the trajectories of (13) to $\text{Saddle}(F)$. Thus, if $\text{Saddle}(F)$ is unbounded, the trajectories of (13) can be unbounded under arbitrarily small constant disturbances. However, if matrix A has full row-rank, then $\text{Saddle}(F)$ is a singleton and the ISS property implies that the trajectory of (13) remains bounded under bounded disturbances. •

As pointed out in Remark 5.4, if $\text{Saddle}(F)$ is not unique, then the trajectories of the dynamics might not be bounded. We next look at a particular type of disturbance input under which this property is still guaranteed. Pick any $(x_*, z_*) \in \text{Saddle}(F)$ and define the function $\tilde{V}_2 : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}$ as

$$\tilde{V}_2(x, z) = \frac{\beta_1}{2} (\|X_{\text{sp}}(x, z)\|^2) + \frac{\beta_2}{2} (\|x - x_*\|^2 + \|z - z_*\|^2)$$

with $\beta_1 > 0$, $\beta_2 = \frac{4\beta_1 M^4}{m^2}$. One can show, following similar steps as those of proof of Theorem 5.1 that, the function \tilde{V}_2 is an ISS Lyapunov function with respect to the point (x_*, z_*) for the dynamics X_{sp}^p when the disturbance input to z -dynamics has the special structure $u_z = A\tilde{u}_z$, $\tilde{u}_z \in \mathbb{R}^n$. This type of disturbance is motivated by scenarios with measurement errors in the values of x and z used in (12)

and without any computation error of the gradient term in the z -dynamics. The following statement makes these facts precise.

Corollary 5.5: (ISS of saddle-point dynamics): Let the saddle function F be of the form (11), with f strongly convex, twice continuously differentiable, and satisfying $mI \preceq \nabla^2 f(x) \preceq MI$ for all $x \in \mathbb{R}^n$ and some constants $0 < m \leq M < \infty$. Then, the dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -\nabla_x F(x, z) \\ \nabla_z F(x, z) \end{bmatrix} + \begin{bmatrix} u_x \\ A\tilde{u}_z \end{bmatrix}, \quad (15)$$

where $(u_x, \tilde{u}_z) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^{2n}$ is measurable and locally essentially bounded input, is ISS with respect to every point of $\text{Saddle}(F)$.

The proof is analogous to that of Theorem 5.1. One arrives at the condition (4) for Lyapunov \tilde{V}_2 and dynamics (15). Corollary 5.5 implies that the trajectory of dynamics (15) remains bounded for bounded input even when the set $\text{Saddle}(F)$ is unbounded.

Example 5.6: (ISS property of saddle-point dynamics): Consider $F : \mathbb{R}^2 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ of the form (11) with

$$f(x) = \|x\|^2, \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad b = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}. \quad (16)$$

Then, $\text{Saddle}(F) = \{(x, z) \in \mathbb{R}^2 \times \mathbb{R}^3 \mid x = (1, 1), z = (-1, -1, -1) + \lambda(1, -1, -1), \lambda \in \mathbb{R}\}$ is a continuum of points. Note that $\nabla^2 f(x) = 2I$, thus, satisfying the assumption of bounds on the Hessian of f . By Theorem 5.1, the saddle-point dynamics for this saddle function F is input-to-state stable with respect to the set $\text{Saddle}(F)$. This fact is illustrated in Figure 2, which also depicts how the specific structure of the disturbance input in (15) affects the boundedness of the trajectories. •

VI. CONCLUSIONS

We have studied the global convergence and robustness properties of the projected saddle-point dynamics. We have provided a characterization of the omega-limit set in terms of the Hessian blocks of the saddle function. Building on this result, we have established the global asymptotic convergence of the projected saddle-point dynamics assuming only local strong convexity-concavity of the saddle function. When the strong convexity-concavity property is global, and for the case when the saddle function takes the form of a Lagrangian of an equality constrained optimization problem, we have established the input-to-state stability of the saddle-point dynamics by identifying an ISS Lyapunov function. Future work will generalize the ISS results to more general classes of saddle functions and exploit their application in the design of opportunistic state-triggered implementations of controllers for optimal frequency regulation in power networks.

ACKNOWLEDGMENTS

We would like to thank Simon K. Niederländer for discussions on Lyapunov functions for the saddle-point dynamics. This work was supported by NSF award ECCS-1307176 and ARPA-e Cooperative Agreement DE-AR0000695 (AC and JC), NSF CPS grant CNS 1544771 (EM), and NSF CNS grant 1545096 (SL).

REFERENCES

- [1] T. Kose, "Solutions of saddle value problems by differential equations," *Econometrica*, vol. 24, no. 1, pp. 59–70, 1956.
- [2] K. Arrow, L. Hurwicz, and H. Uzawa, *Studies in Linear and Non-Linear Programming*. Stanford, California: Stanford University Press, 1958.
- [3] D. Feijer and F. Paganini, "Stability of primal-dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, pp. 1974–1981, 2010.
- [4] 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.
- [5] B. Ghahserifard 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.
- [6] D. Richert and J. Cortés, "Robust distributed linear programming," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, pp. 2567–2582, 2015.
- [7] D. Richert and J. Cortés, "Distributed bargaining in dyadic-exchange networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 3, pp. 310–321, 2016.
- [8] X. Zhang and A. Papachristodoulou, "A real-time control framework for smart power networks: Design methodology and stability," *Automatica*, vol. 58, pp. 43–50, 2015.
- [9] 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.
- [10] T. Stegink, C. D. Persis, and A. J. van der Schaft, "A unifying energy-based approach to stability of power grids with market dynamics," *IEEE Transactions on Automatic Control*, 2016. Submitted.
- [11] B. Ghahserifard and J. Cortés, "Distributed convergence to Nash equilibria in two-network zero-sum games," *Automatica*, vol. 49, no. 6, pp. 1683–1692, 2013.
- [12] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "On the characterization of local Nash equilibria in continuous games," *IEEE Transactions on Automatic Control*, 2016. To appear.
- [13] S. K. Niederländer and J. Cortés, "Distributed coordination for nonsmooth convex optimization via saddle-point dynamics," *SIAM Journal on Control and Optimization*, 2016. Submitted.
- [14] 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.
- [15] A. Cherukuri, B. Ghahserifard, and J. Cortés, "Saddle-point dynamics: conditions for asymptotic stability of saddle points," *SIAM Journal on Control and Optimization*, 2015. Submitted.
- [16] A. Cherukuri, E. Mallada, and J. Cortés, "Asymptotic convergence of primal-dual dynamics," *Systems & Control Letters*, vol. 87, pp. 10–15, 2016.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.
- [18] A. Bacciotti and F. Ceragioli, "Nonpathological Lyapunov functions and discontinuous Caratheodory systems," *Automatica*, vol. 42, no. 3, pp. 453–458, 2006.
- [19] J. Cortés, "Discontinuous dynamical systems - a tutorial on solutions, nonsmooth analysis, and stability," *IEEE Control Systems Magazine*, vol. 28, no. 3, pp. 36–73, 2008.
- [20] S. Lang, *Real and Functional Analysis*. New York: Springer, 3 ed., 1993.
- [21] Y. Lin, E. Sontag, and Y. Wang, "Various results concerning set input-to-state stability," in *IEEE Conf. on Decision and Control*, (New Orleans, LA), pp. 1330–1335, 1995.

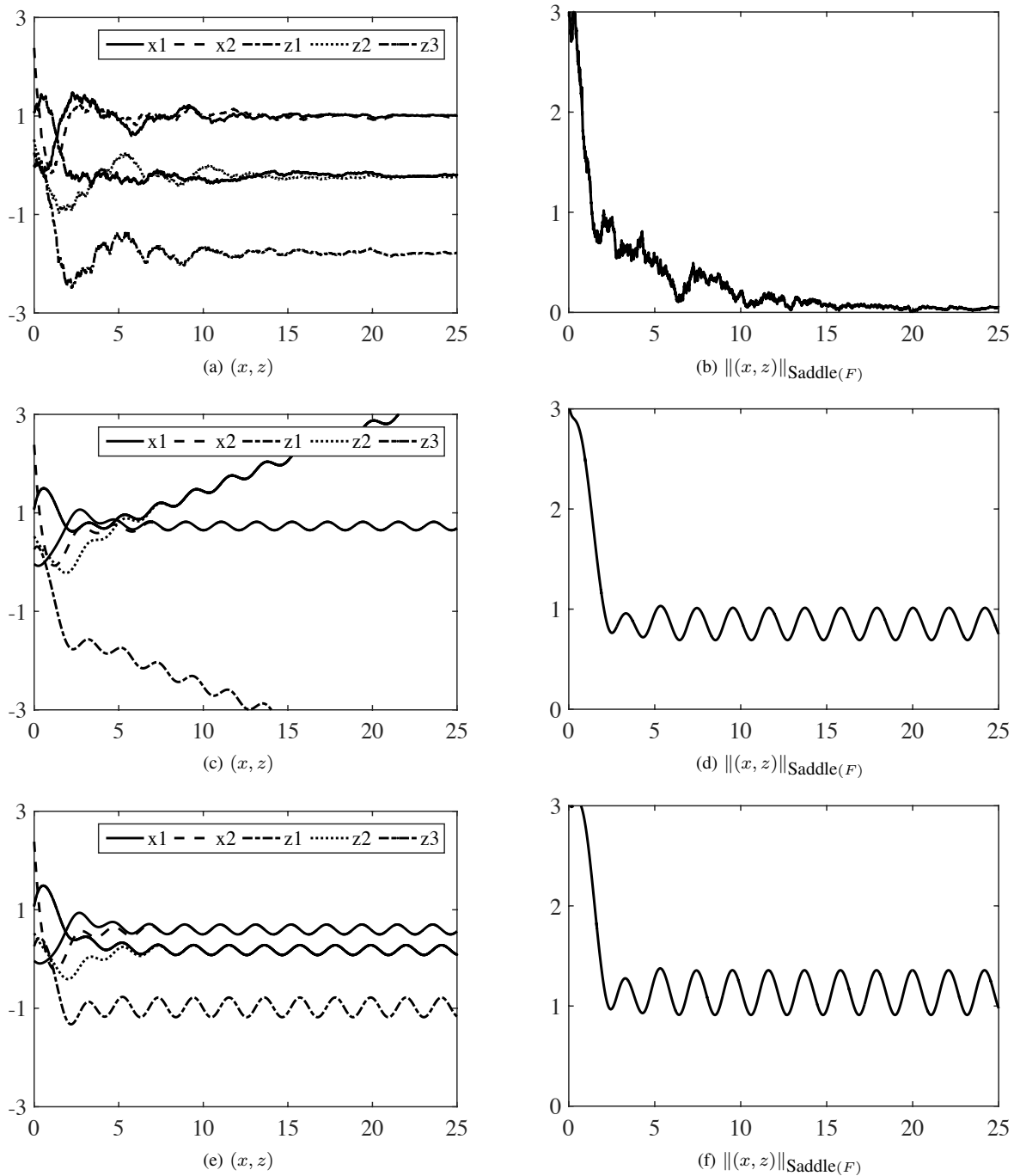


Fig. 2. Plots (a)-(b) show the ISS property, cf. Theorem 5.1, of the dynamics (13) for the saddle function F defined by (16). The initial condition is $x(0) = (-0.0377, 2.3819)$ and $z(0) = (0.2580, 0.5229, 1.0799)$ and the input u is exponentially decaying in magnitude. As shown in (a)-(b), the trajectory converges asymptotically to a saddle point as the input is vanishing. Plots (c)-(d) have the same initial condition but the disturbance input consists of a constant plus a sinusoid. The trajectory is unbounded under bounded input while the distance to the set of saddle points remains bounded, cf. Remark 5.4. Plots (e)-(f) have the same initial condition but the disturbance input to the z -dynamics is of the form (15). In this case, the trajectory remains bounded as the dynamics is ISS with respect to each saddle point, cf. Corollary 5.5.

- [22] E. Mallada, C. Zhao, and S. Low, "Optimal load-side control for frequency regulation in smart grids," 2014. Available at <http://arxiv.org/abs/1410.2931>.
- [23] S. S. Kia, J. Cortés, and S. Martínez, "Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication," *Automatica*, vol. 55, pp. 254–264, 2015.