Robust distributed linear programming

Dean Richert Jorge Cortés

Abstract—**This paper presents a robust, distributed algorithm to solve general linear programs. The algorithm design builds on the characterization of the solutions of the linear program as saddle points of a modified Lagrangian function. We show that the resulting continuous-time saddle-point algorithm is provably correct but, in general, not distributed because of a global parameter associated with the nonsmooth exact penalty function employed to encode the inequality constraints of the linear program. This motivates the design of a discontinuous saddle-point dynamics that, while enjoying the same convergence guarantees, is fully distributed and scalable with the dimension of the solution vector. We also characterize the robustness against disturbances and link failures of the proposed dynamics. Specifically, we show that it is integral-input-to-state stable but not input-to-state stable. The latter fact is a consequence of a more general result, that we also establish, which states that no algorithmic solution for linear programming is input-to-state stable when uncertainty in the problem data affects the dynamics as a disturbance. Our results allow us to establish the resilience of the proposed distributed dynamics to disturbances of finite variation and recurrently disconnected communication among the agents. Simulations in an optimal control application illustrate the results.**

I. Introduction

Linear optimization problems, or simply linear programs, model a broad array of engineering and economic problems and find numerous applications in diverse areas such as operations research, network flow, robust control, microeconomics, and company management. In this paper, we are interested in both the synthesis of distributed algorithms that can solve standard form linear programs and the characterization of their robustness properties. Our interest is motivated by multi-agent scenarios that give rise to linear programs with an intrinsic distributed nature. In such contexts, distributed approaches have the potential to offer inherent advantages over centralized solvers. Among these, we highlight the reduction on communication and computational overhead, the availability of simple computation tasks that can be performed by inexpensive and low-performance processors, and the robustness and adaptive behavior against individual failures. Here, we consider scenarios where individual agents interact with their neighbors and are only responsible for computing their own component of the solution vector of the linear program. We study the synthesis of provably correct, distributed algorithms that make the aggregate of the agents' states converge to a solution of the linear program and are robust to disturbances and communication link failures.

Literature review. Linear programs play an important role in a wide variety of applications, including perimeter patrolling [1], task allocation [2], [3], operator placement [4], process control [5], routing in communication networks [6], and portfolio optimization [7]. This relevance has historically driven the design of efficient methods to solve linear optimization problems, see e.g., [8], [9], [10]. More recently, the interest on networked systems and multi-agent coordination has stimulated the synthesis of distributed strategies to solve linear programs [11], [12], [13] and more general optimization problems with constraints, see e.g., [14], [15], [16] and references therein. The aforementioned works build on consensus-based dynamics [17], [18], [19], [20] whereby individual agents agree on the global solution to the optimization problem. This is a major difference with respect to our work here, in which each individual agent computes only its own component of the solution vector by communicating with its neighbors. This feature makes the messages transmitted over the network independent of the size of the solution vector, and hence scalable (a property which would not be shared by a consensus-based distributed optimization method for the particular class of problems considered here). Some algorithms that enjoy a similar scalability property exist in the literature. In particular, the recent work [21] introduces a partition-based dual decomposition algorithm for network optimization. Other discrete-time algorithms for non-strict convex problems are proposed in [22], [23], but require at least one of the exact solutions of a local optimization problem at each iteration, bounded feasibility sets, or auxiliary variables that increase the problem dimension. The algorithms in [24], [25], [26] on the other hand only achieves convergence to an approximate solution of the optimization problem. Closer to our approach, although without equality constraints, the works [27], [28] build on the saddle-point dynamics of a smooth Lagrangian function to propose algorithms for linear programming. The resulting dynamics are discontinuous because of the projections taken to keep the evolution within the feasible set. Both works establish convergence in the primal variables under the assumption that the solution of the linear program is unique [28] or that Slater's condition is satisfied [27], but do not characterize the properties of the final convergence point in the dual variables, which might indeed not be a solution of the dual problem. We are unaware of works that explicitly address the problem of studying the robustness of linear programming algorithms, particularly from a systems and control perspective. This brings up another point of connection of the present treatment with the literature, which is the body of work on robustness of dynamical systems against disturbances. In particular, we explore the properties of our proposed dynamics with respect to notions such as robust asymptotic stability [29], input-to-state stability (ISS) [30], and integral input-to-state stability (iISS) [31]. The term 'robust optimization' often employed in the literature, see e.g. [32], refers instead to worst-case optimization problems where uncertainty in the data is

The authors are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA 92093, USA, {drichert,cortes}@ucsd.edu

explicitly included in the problem formulation. In this context, 'robust' refers to the problem formulation and not to the actual algorithm employed to solve the optimization.

Statement of contributions. We consider standard form linear programs, which contain both equality and non-negativity constraints on the decision vector. Our first contribution is an alternative formulation of the primal-dual solutions of the linear program as saddle points of a modified Lagrangian function. This function incorporates an exact nonsmooth penalty function to enforce the inequality constraints. Our second contribution concerns the design of a continuous-time dynamics that find the solutions of standard form linear programs. Our alternative problem formulation motivates the study of the saddle-point dynamics (gradient descent in one variable, gradient ascent in the other) associated with the modified Lagrangian. It should be noted that, in general, saddle points are only guaranteed to be stable (and not necessarily asymptotically stable) for the corresponding saddle-point dynamics. Nevertheless, in our case, we are able to establish the global asymptotic stability of the (possibly unbounded) set of primal-dual solutions of the linear program and, moreover, the pointwise convergence of the trajectories. Our analysis relies on the set-valued LaSalle Invariance Principle and, in particular, a careful use of the properties of weakly and strongly invariant sets of the saddle-point dynamics. In general, knowledge of the global parameter associated with the nonsmooth exact penalty function employed to encode the inequality constraints is necessary for the implementation of the saddle-point dynamics. To circumvent this need, we propose an alternative discontinuous saddle-point dynamics that does not require such knowledge and is fully distributed over a multi-agent system in which each individual computes only its own component of the solution vector. We show that the discontinuous dynamics share the same convergence properties of the regular saddle-point dynamics by establishing that, for sufficiently large values of the global parameter, the trajectories of the former are also trajectories of the latter. Two key advantages of our methodology are that it (i) allows us to establish global asymptotic stability of the discontinuous dynamics without establishing any regularity conditions on the switching behavior and (ii) sets the stage for the characterization of novel and relevant algorithm robustness properties. This latter point bring us to our third contribution, which pertains the robustness of the discontinuous saddle-point dynamics against disturbances and link failures. We establish that no continuous-time algorithm that solves general linear programs can be input-to-state stable (ISS) when uncertainty in the problem data affects the dynamics as a disturbance. As our technical approach shows, this fact is due to the intrinsic properties of the primal-dual solutions to linear programs. Nevertheless, when the set of primal-dual solutions is compact, we show that our discontinuous saddle-point dynamics possesses an ISS-like property against small constant disturbances and, more importantly, is integral input-to-state stable (iISS) – and thus robust to finite energy disturbances. Our proof method is based on identifying a suitable iISS Lyapunov function, which we build by combining the Lyapunov function used in our LaSalle argument and results from converse Lyapunov theory. We conclude that one cannot expect better disturbance rejection properties from a linear programming algorithm than those we establish for our discontinuous saddle-point dynamics. These results allow us to establish the robustness of our dynamics against disturbances of finite variation and communication failures among agents modeled by recurrently connected graphs. Simulations in an optimal control problem illustrate the results.

Organization. Section II introduces basic preliminaries. Section III presents the problem statement. Section IV proposes the discontinuous saddle-point dynamics, establishes its convergence, and discusses its distributed implementation. Sections V and VI study the algorithm robustness against disturbances and communication link failures, respectively. Simulations illustrate our results in Section VII. Finally, Section VIII summarizes our results and ideas for future work.

II. Preliminaries

Here, we introduce notation and basic notions on nonsmooth analysis and dynamical systems. This section may be safely skipped by the reader familiar with these areas.

A. Notation and basic notions

The set of real numbers is $\mathbb{R}$. For $x \in \mathbb{R}^n$, $x \geq 0$ (resp. $x > 0$) means that all components of x are nonnegative (resp. positive). For $x \in \mathbb{R}^n$, we define $\max\{0, x\} = (\max\{0, x_1\}, \dots, \max\{0, x_n\}) \in \mathbb{R}^n_{\geq 0}$. We let $\mathbb{1}_n \in \mathbb{R}^n$ denote the vector of ones. We use $\|\cdot\|$ and $\|\cdot\|_\infty$ to denote the 2- and ∞-norms in $\mathbb{R}^n$. The Euclidean distance from a point $x \in \mathbb{R}^n$ to a set $A \subset \mathbb{R}^n$ is denoted by $\|\cdot\|_A$. The set $\mathbb{B}(x, \delta) \subset \mathbb{R}^n$ is the open ball centered at $x \in \mathbb{R}^n$ with radius $\delta > 0$. The set $A \subset \mathbb{R}^n$ is convex if it fully contains the segment connecting any two points in A.

A function $V : \mathbb{R}^n \to \mathbb{R}$ is positive definite with respect to $A \subset \mathbb{R}^n$ if (i) $V(x) = 0$ for all $x \in A$ and $V(x) > 0$ for all $x \notin A$. If $A = \{0\}$, we refer to V as positive definite. $V : \mathbb{R}^n \to \mathbb{R}$ is radially unbounded with respect to A if $V(x) \to \infty$ when $\|x\|_A \to \infty$. If $A = \{0\}$, we refer to V as radially unbounded. A function V is proper with respect to A if it is both positive definite and radially unbounded with respect to A. A set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ maps elements in $\mathbb{R}^n$ to subsets of $\mathbb{R}^n$. A function $V : X \to \mathbb{R}$ defined on the convex set $X \subset \mathbb{R}^n$ is convex if $V(kx + (1-k)y) \leq kV(x) + (1-k)V(y)$ for all $x, y \in X$ and $k \in [0, 1]$. V is concave iff $-V$ is convex. Given $\rho \in \mathbb{R}$, we define $V^{-1}(\leq \rho) = \{x \in X \mid V(x) \leq \rho\}$. The function $L : X \times Y \to \mathbb{R}$ defined on the convex set $X \times Y \subset \mathbb{R}^n \times \mathbb{R}^m$ is convex-concave if it is convex on its first argument and concave on its second. A point $(\bar{x}, \bar{y}) \in X \times Y$ is a saddle point of L if $L(x, \bar{y}) \geq L(\bar{x}, \bar{y}) \geq L(\bar{x}, y)$ for all $(x, y) \in X \times Y$.

Comparison functions are useful to formalize stability properties. The class of $\mathcal{K}$ functions is composed by functions $[0, \infty) \to [0, \infty)$ that are continuous, zero at zero, and strictly increasing. The subset of unbounded class $\mathcal{K}$ functions are called class $\mathcal{K}_\infty$. A class $\mathcal{KL}$ function $[0, \infty) \times [0, \infty) \to [0, \infty)$ is class $\mathcal{K}$ in its first variable and continuous, decreasing, and converging to zero in its second variable.

An undirected graph is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \dots, n\}$ is a set of vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is a set of edges. Given a matrix $A \in \mathbb{R}^{m \times n}$, we call a graph *connected with respect to* A if for each $\ell \in \{1, \dots, m\}$ such that $a_{\ell,i} \neq 0$ and $a_{\ell,j} \neq 0$, it holds that $(i, j) \in \mathcal{E}$.

B. Nonsmooth analysis

Here we review some basic notions from nonsmooth analysis following [33]. A function $V : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz at $x \in \mathbb{R}^n$ if there exist $\delta_x > 0$ and $L_x \geq 0$ such that $|V(y_1) - V(y_2)| \leq L_x \|y_1 - y_2\|$ for $y_1, y_2 \in \mathbb{B}(x, \delta_x)$. If V is locally Lipschitz at all $x \in \mathbb{R}^n$, we refer to V as locally Lipschitz. If V is convex, then it is locally Lipschitz. A locally Lipschitz function is differentiable almost everywhere. Let $\Omega_V \subset \mathbb{R}^n$ be then the set of points where V is not differentiable. The generalized gradient of a locally Lipschitz function V at $x \in \mathbb{R}^n$ is

$$\partial V(x) = \operatorname{co}\Big\{ \lim_{i \to \infty} \nabla V(x_i) : x_i \to x, x_i \notin S \cup \Omega_V \Big\},$$

where $\operatorname{co}\{\cdot\}$ denotes the convex hull and $S \subset \mathbb{R}^n$ is any set with zero Lebesgue measure. A critical point $x \in \mathbb{R}^n$ of V satisfies $0 \in \partial V(x)$. For a convex function V, the first-order condition of convexity states that $V(y) \geq V(x) + (y - x)^T g$ for all $g \in \partial V(x)$ and $x, y \in \mathbb{R}^n$. For $V : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, we use $\partial_x V(x, y)$ and $\partial_y V(x, y)$ to denote the generalized gradients of the maps $x' \mapsto V(x', y)$ at x and $y' \mapsto V(x, y')$ at y, respectively.

A set-valued map $F : X \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is upper semi-continuous if for all $x \in X$ and $\varepsilon \in (0, \infty)$ there exists $\delta_x \in (0, \infty)$ such that $F(y) \subseteq F(x) + \mathbb{B}(0, \varepsilon)$ for all $y \in \mathbb{B}(x, \delta_x)$. Conversely, F is lower semi-continuous if for all $x \in X$, $\varepsilon \in (0, \infty)$, and any open set A intersecting $F(x)$ there exists a $\delta \in (0, \infty)$ such that $F(y)$ intersects A for all $y \in \mathbb{B}(x, \delta)$. If F is both upper and lower semi-continuous then it is continuous. Also, F is locally bounded if for every $x \in X$ there exist $\varepsilon \in (0, \infty)$ and $M > 0$ such that $\|z\| \leq M$ for all $z \in F(y)$ and all $y \in \mathbb{B}(x, \varepsilon)$.

Lemma II.1 (Properties of the generalized gradient). *If $V : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz at $x \in \mathbb{R}^n$, then $\partial V(x)$ is nonempty, convex, and compact. Moreover, $x \mapsto \partial V(x)$ is locally bounded and upper semi-continuous.*

C. Set-valued dynamical systems

Our exposition on basic concepts for set-valued dynamical systems follows [34]. A time-invariant set-valued dynamical system is represented by the differential inclusion

$$\dot{x} \in F(x), \tag{1}$$

where $t \in \mathbb{R}_{\geq 0}$ and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set valued map. If F is locally bounded, upper semi-continuous and takes nonempty, convex, and compact values, then from any initial condition in $\mathbb{R}^n$, there exists an absolutely continuous curve $x : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, called solution, satisfying (1) almost everywhere. The solution is maximal if it cannot be extended forward in time. The set of equilibria of F is defined as $\{x \in \mathbb{R}^n \mid 0 \in F(x)\}$. A set $\mathcal{M}$ is strongly (resp. weakly) invariant with respect to (1) if, for each $x_0 \in \mathcal{M}$, $\mathcal{M}$ contains all (resp. at least one) maximal solution(s) of (1) with initial condition x_0. The set-valued Lie derivative of a differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ along the trajectories of (1) is

$$\mathcal{L}_F V(x) = \{\nabla V(x)^T v : v \in F(x)\}.$$

The following result helps establish the asymptotic convergence properties of (1).

Theorem II.2 (Set-valued LaSalle Invariance Principle). *Let $X \subset \mathbb{R}^n$ be compact and strongly invariant with respect to* (1)*. Assume $V : \mathbb{R}^n \to \mathbb{R}$ is differentiable and F is locally bounded, upper semi-continuous and takes nonempty, convex, and compact values. If $\mathcal{L}_F V(x) \subset (-\infty, 0]$ for all $x \in X$, then any solution of* (1) *starting in X converges to the largest weakly invariant set $\mathcal{M}$ contained in $\overline{\{x \in X : 0 \in \mathcal{L}_F V(x)\}}$.*

Differential inclusions are specially useful to handle differential equations with discontinuities. Specifically, let $f : X \subset \mathbb{R}^n \to \mathbb{R}^n$ be a piecewise continuous vector field and consider

$$\dot{x} = f(x). \tag{2}$$

The classical notion of solution is not applicable to (2) because of the discontinuities. Instead, consider the Krasovskii set-valued map associated to f, defined by $\mathscr{K}[f](x) := \bigcap_{\delta > 0} \overline{\operatorname{co}}\{f(\mathbb{B}(x, \delta))\}$, where $\overline{\operatorname{co}}\{\cdot\}$ denotes the closed convex hull. One can show that the set-valued map $\mathscr{K}[f]$ is locally bounded, upper semi-continuous and takes nonempty, convex, and compact values, and hence solutions exist to

$$\dot{x} \in \mathscr{K}[f](x) \tag{3}$$

starting from any initial condition. The solutions of (2) in the sense of Krasovskii are, by definition, the solutions of the differential inclusion (3).

III. Problem statement and equivalent formulation

This section introduces standard form linear programs and describes an alternative formulation that is useful later in fulfilling our main objective, which is the design of robust, distributed algorithms to solve them. Consider the following standard form linear program,

$$\min \quad c^T x \tag{4a}$$
$$\text{s.t.} \quad Ax = b, \quad x \geq 0, \tag{4b}$$

where $x, c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$. We only consider feasible linear programs with finite optimal value. The set of solutions to (4) is $\mathcal{X} \subset \mathbb{R}^n$. The dual formulation is

$$\max \quad -b^T z \tag{5a}$$
$$\text{s.t.} \quad A^T z + c \geq 0. \tag{5b}$$

The set of solutions to (5) is denoted by $\mathcal{Z} \subset \mathbb{R}^m$. We use x_* and z_* to denote a solution of (4) and (5), respectively. The following result is a fundamental relationship between primal and dual solutions of linear programs and can be found in many optimization references, see e.g., [10].

Theorem III.1 (Complementary slackness and strong duality). *Suppose that $x \in \mathbb{R}^n$ is feasible for* (4) *and $z \in \mathbb{R}^m$ is feasible for* (5)*. Then x is a solution to* (4) *and z is a solution to* (5) *if and only if $(A^T z + c)^T x = 0$. In compact form,*

$$\mathcal{X} \times \mathcal{Z} = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m \mid Ax = b,\ x \geq 0,\\ A^T z + c \geq 0,\ (A^T z + c)^T x = 0\}. \quad (6)$$

Moreover, for any $(x_*, z_*) \in \mathcal{X} \times \mathcal{Z}$, $c^T x_* = -b^T z_*$.

The equality $(A^T z + c)^T x = 0$ is called the *complementary slackness* condition whereas the property that $c^T x_* = -b^T z_*$ is called *strong duality*. One remarkable consequence of Theorem III.1 is that the set on the right-hand side of (6) is convex (because $\mathcal{X} \times \mathcal{Z}$ is convex). This fact is not obvious since the complementary slackness condition is not affine in the variables x and z. This observation will allow us to use a simplified version of Danskin's Theorem (see Lemma A.2) in the proof of a key result of Section V. The next result establishes the connection between the solutions of (4) and (5) and the saddle points of a modified Lagrangian function. Its proof can be deduced from results on penalty functions that appear in optimization, see e.g. [35], but we include it here for completeness and consistency of the presentation.

Proposition III.2 (Solutions of linear program as saddle points). *For $K \geq 0$, let $L^K : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be defined by*

$$L^K(x, z) = c^T x + \frac{1}{2}(Ax - b)^T(Ax - b) + z^T(Ax - b)\\ + K\mathbb{1}_n^T \max\{0, -x\}. \quad (7)$$

Then, L^K is convex in x and concave (in fact, linear) in z. Moreover,

(i) *if $x_* \in \mathbb{R}^n$ is a solution of* (4) *and $z_* \in \mathbb{R}^m$ is a solution of* (5)*, then the point (x_*, z_*) is a saddle point of L^K for any $K \geq \|A^T z_* + c\|_\infty$,*

(ii) *if $(\bar{x}, \bar{z}) \in \mathbb{R}^n \times \mathbb{R}^m$ is a saddle point of L^K with $K > \|A^T z_* + c\|_\infty$ for some $z_* \in \mathbb{R}^m$ solution of* (5)*, then $\bar{x} \in \mathbb{R}^n$ is a solution of* (4).

Proof: One can readily see from (7) that L^K is a convex-concave function. Let x_* be a solution of (4) and let z_* be a solution of (5). To show (i), using the characterization of $\mathcal{X} \times \mathcal{Z}$ described in Theorem III.1 and the fact that $K \geq \|A^T z_* + c\|_\infty$, we can write for any $x \in \mathbb{R}^n$,

$$\begin{aligned} L^K(x, z_*) &= c^T x + (Ax - b)^T(Ax - b) + z_*^T(Ax - b)\\ &\quad + K\mathbb{1}_n^T \max\{0, -x\},\\ &\geq c^T x + z_*^T(Ax - b) + (A^T z_* + c)^T \max\{0, -x\},\\ &\geq c^T x + z_*^T(Ax - b) - (A^T z_* + c)^T x,\\ &= c^T x + z_*^T A(x - x_*) - (A^T z_* + c)^T(x - x_*),\\ &= c^T x - c^T(x - x_*) = c^T x_* = L^K(x_*, z_*). \end{aligned}$$

The fact that $L^K(x_*, z) = L^K(x_*, z_*)$ for any $z \in \mathbb{R}^m$ is immediate. These two facts together imply that (x_*, z_*) is a saddle point of L^K.

We prove (ii) by contradiction. Let $(\bar{x}, \bar{z})$ be a saddle point of L^K with $K > \|A^T z_* + c\|_\infty$ for some $z_* \in \mathcal{Z}$, but suppose $\bar{x}$ is not a solution of (4). Let $x_* \in \mathcal{X}$. Since for fixed x, $z \mapsto L^K(x, z)$ is concave and differentiable, a necessary condition for $(\bar{x}, \bar{z})$ to be a saddle point of L^K is that $A\bar{x} - b = 0$. Using this fact, $L^K(x_*, \bar{z}) \geq L^K(\bar{x}, \bar{z})$ can be expressed as

$$c^T x_* \geq c^T \bar{x} + K\mathbb{1}_n^T \max\{0, -\bar{x}\}. \quad (8)$$

Now, if $\bar{x} \geq 0$, then $c^T x_* \geq c^T \bar{x}$, and thus $\bar{x}$ would be a solution of (4). If, instead, $\bar{x} \ngeq 0$,

$$\begin{aligned} c^T \bar{x} &= c^T x_* + c^T(\bar{x} - x_*),\\ &= c^T x_* - z_*^T A(\bar{x} - x_*) + (A^T z_* + c)^T(\bar{x} - x_*),\\ &= c^T x_* - z_*^T(A\bar{x} - b) + (A^T z_* + c)^T \bar{x},\\ &> c^T x_* - K\mathbb{1}_n^T \max\{0, -\bar{x}\}, \end{aligned}$$

which contradicts (8), concluding the proof. ■

The relevance of Proposition III.2 is two-fold. On the one hand, it justifies searching for the saddle points of L^K instead of directly solving the constrained optimization problem (4). On the other hand, given that L^K is convex-concave, a natural approach to find the saddle points is via the associated saddle-point dynamics. However, for an arbitrary function, such dynamics is known to render saddle points only stable, not asymptotically stable (in fact, the saddle-point dynamics derived using the standard Lagrangian for a linear program does not converge to a solution of the linear program, see e.g., [36], [28]). Interestingly [28], the convergence properties of saddle-point dynamics can be improved using penalty functions associated with the constraints to augment the cost function. In our case, we augment the linear cost function $c^T x$ with a quadratic penalty for the equality constraints and a nonsmooth penalty function for the inequality constraints. This results in the nonlinear optimization problem,

$$\min_{Ax=b} c^T x + \|Ax - b\|^2 + K\mathbb{1}_n^T \max\{0, -x\},$$

whose standard Lagrangian is equivalent to L^K. We use the nonsmooth penalty function to ensure that there is an *exact* equivalence between saddle points of L^K and the solutions of (4). Instead, the use of smooth penalty functions such as the logarithmic barrier function used in [16], results only in approximate solutions. In the next section, we show that indeed the saddle-point dynamics of L^K asymptotically converges to saddle points.

Remark III.3 (Bounds on the parameter K). It is worth noticing that the lower bounds on K in Proposition III.2 are characterized by certain dual solutions, which are unknown a priori. Nevertheless, our discussion later shows that this problem can be circumvented and that knowledge of such bounds is not necessary for the design of robust, distributed algorithms that solve linear programs. •

IV. Saddle-point dynamics for distributed linear programming

In this section, we design a continuous-time algorithm to find the solutions of (4) and discuss its distributed implementation in a multi-agent system. We further build on the elements of analysis introduced here to characterize the robustness properties of linear programming dynamics in the forthcoming sections. Building on the result in Proposition III.2, we consider the saddle-point dynamics (gradient descent in one argument, gradient ascent in the other) of the modified Lagrangian L^K. Our presentation proceeds by characterizing the properties of this dynamics and observing its limitations, leading up to the main contribution, which is the introduction of a discontinuous saddle-point dynamics amenable to distributed implementation.

The nonsmooth character of L^K means that its saddle-point dynamics takes the form of the following differential inclusion,

$$\dot{x} + c + A^T(z + Ax - b) \in -K\partial \max\{0, -x\}, \tag{9a}$$

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

For notational convenience, we use $F^K_{\text{sdl}} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m$ to denote the set-valued vector field which defines the differential inclusion (9). The following result characterizes the asymptotic convergence of (9) to the set of solutions to (4)-(5).

Theorem IV.1 (Asymptotic convergence to the primal-dual solution set). *Let $(x_*, z_*) \in \mathcal{X} \times \mathcal{Z}$ and define $V : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ as*

$$V(x, z) = \frac{1}{2}(x - x_*)^T(x - x_*) + \frac{1}{2}(z - z_*)^T(z - z_*).$$

For $\infty > K > \|A^T z_* + c\|_\infty$, it holds that $\mathcal{L}_{F^K_{\text{sdl}}} V(x, z) \subset (-\infty, 0]$ for all $(x, z) \in \mathbb{R}^n \times \mathbb{R}^m$ and any trajectory $t \mapsto (x(t), z(t))$ of (9) converges asymptotically to the set $\mathcal{X} \times \mathcal{Z}$.

Proof: Our proof strategy is based on verifying the hypotheses of the LaSalle Invariance Principle, cf. Theorem II.2, and identifying the set of primal-dual solutions as the corresponding largest weakly invariant set. First, note that Lemma II.1 implies that F^K_{sdl} is locally bounded, upper semi-continuous and takes nonempty, convex, and compact values. By Proposition III.2(i), (x_*, z_*) is a saddle point of L^K when $K \geq \|A^T z_* + c\|_\infty$. Consider the quadratic function V defined in the theorem statement, which is continuously differentiable and radially unbounded. Let $a \in \mathcal{L}_{F^K_{\text{sdl}}} V(x, z)$. By definition, there exists $v = (-c - A^T(z + Ax - b) - g_x, Ax - b) \in F^K_{\text{sdl}}(x, z)$, with $g_x \in K\partial \max\{0, -x\}$, such that

$$\begin{aligned} a = v^T \nabla V(x, z) = (x - x_*)^T(- c - A^T(z + Ax - b) - g_x) \\ + (z - z_*)(Ax - b). \end{aligned} \tag{10}$$

Since L^K is convex in its first argument, and $c + A^T(z + Ax - b) + g_x \in \partial_x L^K(x, z)$, using the first-order condition of convexity, we have

$$L^K(x, z) \leq L^K(x_*, z) + (x - x_*)^T\big(c + A^T(z + Ax - b) + g_x\big).$$

Since L^K is linear in z, we have $L^K(x, z) = L^K(x, z_*) + (z - z_*)^T(Ax - b)$. Using these facts in (10), we get

$$\begin{aligned} a &\leq L^K(x_*, z) - L^K(x, z_*) \\ &= L^K(x_*, z) - L^K(x_*, z_*) + L^K(x_*, z_*) - L^K(x, z_*) \leq 0, \end{aligned}$$

since (x_*, z_*) is a saddle point of L^K. Since a is arbitrary, we deduce that $\mathcal{L}_{F^K_{\text{sdl}}} V(x, z) \subset (-\infty, 0]$. For any given $\rho \geq 0$, this implies that the sublevel set $V^{-1}(\leq \rho)$ is strongly invariant with respect to (9). Since V is radially unbounded, $V^{-1}(\leq \rho)$ is also compact. The conditions of Theorem II.2 are then satisfied with $X = V^{-1}(\leq \rho)$, and therefore any trajectory of (9) starting in $V^{-1}(\leq \rho)$ converges to the largest weakly invariant set $\mathcal{M}$ in $\{(x, z) \in V^{-1}(\leq \rho) : 0 \in \mathcal{L}_{F^K_{\text{sdl}}} V(x, z)\}$ (note that for any initial condition (x_0, z_0) one can choose a ρ such that $(x_0, z_0) \in V^{-1}(\leq \rho)$). This set is closed, which can be justified as follows. Since F^K_{sdl} is upper semi-continuous and V is continuously differentiable, the map $(x, z) \mapsto \mathcal{L}_{F^K_{\text{sdl}}} V(x, z)$ is also upper semi-continuous. Closedness then follows from [37, Convergence Theorem]. We now show that $\mathcal{M} \subseteq \mathcal{X} \times \mathcal{Z}$. To start, take $(x', z') \in \mathcal{M}$. Then $L^K(x_*, z_*) - L^K(x', z_*) = 0$, which implies

$$\tilde{L}^K(x', z_*) - (Ax' - b)^T(Ax' - b) = 0, \tag{11}$$

where $\tilde{L}^K(x', z_*) = c^T x_* - c^T x' - z_*^T(Ax' - b) - K\mathbb{1}_n^T \max\{0, -x'\}$. Using strong duality, the expression of $\tilde{L}^K$ can be simplified to $\tilde{L}^K(x', z_*) = -(A^T z_* + c)^T x' - K\mathbb{1}_n^T \max\{0, -x'\}$. In addition, $A^T z_* + c \geq 0$ by dual feasibility. Thus, when $K \geq \|A^T z_* + c\|_\infty$, we have $\tilde{L}^K(x, z_*) \leq 0$ for all $(x, z) \in V^{-1}(\leq \rho)$. This implies that $(Ax' - b)^T(Ax' - b) = 0$ for (11) to be true, which further implies that $Ax' - b = 0$. Moreover, from the definition of $\tilde{L}^K$ and the bound on K, one can see that if $x' \not\geq 0$, then $\tilde{L}^K(x', z_*) < 0$. Therefore, for (11) to be true, it must be that $x' \geq 0$. Finally, from (11), we get that $\tilde{L}^K(x', z_*) = c^T x_* - c^T x' = 0$. In summary, if $(x', z') \in \mathcal{M}$ then $c^T x_* = c^T x'$, $Ax' - b = 0$, and $x' \geq 0$. Therefore, x' is a solution of (4). Now, we show that z' is a solution of (5). Because $\mathcal{M}$ is weakly invariant, there exists a trajectory starting from (x', z') that remains in $\mathcal{M}$. The fact that $Ax' = b$ implies that $\dot{z} = 0$, and hence $z(t) = z'$ is constant. For any given $i \in \{1, \dots, n\}$, we consider the cases (i) $x'_i > 0$ and (ii) $x'_i = 0$. In case (i), the dynamics of the ith component of x is $\dot{x}_i = -(c + A^T z')_i$ where $(c + A^T z')_i$ is constant. It cannot be that $-(c + A^T z')_i > 0$ because this would contradict the fact that $t \mapsto x_i(t)$ is bounded. Therefore, $(c + A^T z')_i \geq 0$. If $\dot{x}_i = -(c + A^T z')_i < 0$, then $x_i(t)$ will eventually become zero, which we consider in case (ii). In fact, since the solution remains in $\mathcal{M}$, without loss of generality, we can assume that (x', z') is such that either $x'_i > 0$ and $(c + A^T z')_i = 0$ or $x'_i = 0$ for each $i \in \{1, \dots, n\}$. Consider now case (ii). Since $x_i(t)$ must remain non-negative in $\mathcal{M}$, it must be that $\dot{x}_i(t) \geq 0$ when $x_i(t) = 0$. That is, in $\mathcal{M}$, we have $\dot{x}_i(t) \geq 0$ when $x_i(t) = 0$ and $\dot{x}_i(t) \leq 0$ when $x_i(t) > 0$. Therefore, for any trajectory $t \mapsto x_i(t)$ in $\mathcal{M}$ starting at $x'_i = 0$, the unique Krasovskii solution is that $x_i(t) = 0$ for all $t \geq 0$. As a consequence, $(c + A^T z')_i \in [0, K]$ if $x'_i = 0$. To summarize cases (i) and (ii), we have

- $Ax' = b$ and $x' \geq 0$ (primal feasibility),
- $A^T z' + c \geq 0$ (dual feasibility),
- $(A^T z' + c)_i = 0$ if $x'_i > 0$ and $x'_i = 0$ if $(A^T z' + c)_i > 0$ (complementary slackness),

which is sufficient to show that $z \in \mathcal{Z}$ (cf. Theorem III.1). Hence $\mathcal{M} \subseteq \mathcal{X} \times \mathcal{Z}$. Since the trajectories of (9) converge to $\mathcal{M}$, this completes the proof. ■

Using a slightly more complicated lower bound on the parameter K, we are able to show point-wise convergence of the saddle-point dynamics. We state this result next.

Corollary IV.2 (Point-wise convergence of saddle-point dynamics). *Let $\rho > 0$. Then, with the notation of Theorem IV.1, for*

$$\infty > K > \max_{(x,z) \in (\mathcal{X} \times \mathcal{Z}) \cap V^{-1}(\leq \rho)} \|A^T z + c\|_\infty, \qquad (12)$$

it holds that any trajectory $t \mapsto (x(t), z(t))$ of (9) starting in $V^{-1}(\leq \rho)$ converges asymptotically to a point in $\mathcal{X} \times \mathcal{Z}$.

Proof: If K satisfies (12), then in particular $K > \|A^T z_* + c\|_\infty$. Thus, $V^{-1}(\leq \rho)$ is strongly invariant under (9) since $\mathcal{L}_{F^K_{\text{sdl}}} V(x,z) \subset (-\infty, 0]$ for all $(x,z) \in V^{-1}(\leq \rho)$ (cf. Theorem IV.1). Also, $V^{-1}(\leq \rho)$ is bounded because V is quadratic. Therefore, by the Bolzano-Weierstrass theorem [38, Theorem 3.6], there exists a subsequence $(x(t_k), z(t_k)) \in V^{-1}(\leq \rho)$ that converges to a point $(\tilde{x}, \tilde{z}) \in (\mathcal{X} \times \mathcal{Z}) \cap V^{-1}(\leq \rho)$. Given $\varepsilon > 0$, let k^* be such that $\|(x(t_{k^*}), z(t_{k^*})) - (\tilde{x}, \tilde{z})\| \leq \varepsilon$. Consider the function $\tilde{V}(x,z) = \frac{1}{2}(x - \tilde{x})^T(x - \tilde{x}) + \frac{1}{2}(z - \tilde{z})^T(z - \tilde{z})$. When K satisfies (12), again it holds that $K \geq \|A^T \tilde{z} + c\|_\infty$. Applying Theorem IV.1 once again, $\tilde{V}^{-1}(\leq \rho)$ is strongly invariant under (9). Consequently, for $t \geq t_{k^*}$, we have $(x(t), z(t)) \in \tilde{V}^{-1}(\leq \tilde{V}(x(t_{k^*}), z(t_{k^*}))) = \overline{\mathbb{B}}\big((\tilde{x}, \tilde{z}), \|(x(t_{k^*}), z(t_{k^*})) - (\tilde{x}, \tilde{z})\|\big) \subset \overline{\mathbb{B}}((\tilde{x}, \tilde{z}), \varepsilon)$. Since ε can be taken arbitrarily small, this implies that $(x(t), z(t))$ converges to the point $(\tilde{x}, \tilde{z}) \in \mathcal{X} \times \mathcal{Z}$. ■

Remark IV.3 (Choice of parameter K). The bound (12) for the parameter K depends on (i) the primal-dual solution set $\mathcal{X} \times \mathcal{Z}$ as well as (ii) the initial condition, since the result is only valid when the dynamics start in $V^{-1}(\leq \rho)$. However, if the set $\mathcal{X} \times \mathcal{Z}$ is compact, the parameter K can be chosen independently of the initial condition since the maximization in (12) would be well defined when taken over the whole set $\mathcal{X} \times \mathcal{Z}$. We should point out that, in Section IV-A we introduce a discontinuous version of the saddle-point dynamics which does not involve K. •

A. Discontinuous saddle-point dynamics

Here, we propose an alternative dynamics to (9) that does not rely on knowledge of the parameter K and also converges to the solutions of (4)-(5). We begin by defining the *nominal flow function* $f^{\text{nom}} : \mathbb{R}^n_{\geq 0} \times \mathbb{R}^m \to \mathbb{R}^n$ by

$$f^{\text{nom}}(x,z) := -c - A^T(z + Ax - b).$$

This definition is motivated by the fact that, for $(x,z) \in \mathbb{R}^n_{>0} \times \mathbb{R}^m$, the set $\partial_x L^K(x,z)$ is the singleton $\{-f^{\text{nom}}(x,z)\}$. The *discontinuous saddle-point dynamics* is, for $i \in \{1, \dots, n\}$,

$$\dot{x}_i = \begin{cases} f^{\text{nom}}_i(x,z), & \text{if } x_i > 0, \\ \max\{0, f^{\text{nom}}_i(x,z)\}, & \text{if } x_i = 0, \end{cases} \qquad (13a)$$

$$\dot{z} = Ax - b. \qquad (13b)$$

When convenient, we use the notation $f_{\text{dis}} : \mathbb{R}^n_{\geq 0} \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ to refer to the discontinuous dynamics (13). Note that the discontinuous function that defines the dynamics (13a) is simply the positive projection operator, i.e., when $x_i = 0$, it corresponds to the projection of $f^{\text{nom}}_i(x,z)$ onto $\mathbb{R}_{\geq 0}$. We understand the solutions of (13) in the Krasovskii sense. We begin our analysis by establishing a relationship between the Krasovskii set-valued map of f_{dis} and the saddle-point dynamics F^K_{sdl} which allows us to relate the trajectories of (13) and (9).

Proposition IV.4 (Trajectories of the discontinuous saddle-point dynamics are trajectories of the saddle-point dynamics). *Let $\rho > 0$ and $(x_*, z_*) \in \mathcal{X} \times \mathcal{Z}$ be given and the function V be defined as in Theorem IV.1. Then, for any*

$$\infty > K \geq K_1 := \max_{(x,z) \in V^{-1}(\leq \rho)} \|f^{\text{nom}}(x,z)\|_\infty,$$

the inclusion $\mathscr{K}[f_{dis}](x,z) \subseteq F^K_{sdl}(x,z)$ holds for every $(x,z) \in V^{-1}(\leq \rho)$. Thus, the trajectories of (13) starting in $V^{-1}(\leq \rho)$ are also trajectories of (9).

Proof: The projection onto the i^{th} component of the Krasovskii set-valued map $\mathscr{K}[f_{\text{dis}}]$ is

$$\text{proj}_i(\mathscr{K}[f_{\text{dis}}](x,z)) = \begin{cases} \{f^{\text{nom}}_i(x,z)\}, \\ \qquad \text{if } i \in \{1, \dots, n\} \text{ and } x_i > 0, \\ [f^{\text{nom}}_i(x,z), \max\{0, f^{\text{nom}}_i(x,z)\}], \\ \qquad \text{if } i \in \{1, \dots, n\} \text{ and } x_i = 0, \\ \{(Ax - b)_i\}, \\ \qquad \text{if } i \in \{n+1, \dots, n+m\}. \end{cases}$$

As a consequence, for any $i \in \{n+1, \dots, n+m\}$, we have

$$\text{proj}_i(F^K_{\text{sdl}}(x,z)) = (Ax - b)_i = \text{proj}_i(\mathscr{K}[f_{\text{dis}}](x,z)),$$

and, for any $i \in \{1, \dots, n\}$ such that $x_i > 0$, we have

$$\begin{aligned} \text{proj}_i(F^K_{\text{sdl}}(x,z)) &= (-c - A^T(Ax - b + z))_i \\ &= \{f^{\text{nom}}_i(x,z)\} = \text{proj}_i(\mathscr{K}[f_{\text{dis}}](x,z)). \end{aligned}$$

Thus, let us consider the case when $x_i = 0$ for some $i \in \{1, \dots, n\}$. In this case, note that

$$\begin{aligned} \text{proj}_i(\mathscr{K}[f_{\text{dis}}](x,z)) &= [f^{\text{nom}}_i(x,z), \max\{0, f^{\text{nom}}_i(x,z)\}] \\ &\subseteq [f^{\text{nom}}_i(x,z), f^{\text{nom}}_i(x,z) + |f^{\text{nom}}_i(x,z)|], \\ \text{proj}_i(F^K_{\text{sdl}}(x,z)) &= [f^{\text{nom}}_i(x,z), f^{\text{nom}}_i(x,z) + K]. \end{aligned}$$

The choice $K \geq |f^{\text{nom}}_i(x,z)|$ for each $i \in \{1, \dots, n\}$ makes $\mathscr{K}[f_{\text{dis}}](x,z) \subseteq F^K_{\text{sdl}}(x,z)$. More generally, since $V^{-1}(\rho)$ is compact and f^{nom} is continuous, the choice

$$\infty > K \geq \max_{(x,z) \in V^{-1}(\rho)} \|f^{\text{nom}}(x,z)\|_\infty,$$

guarantees $\mathscr{K}[f_{\text{dis}}](x,z) \subseteq F^K_{\text{sdl}}(x,z)$ for all $(x,z) \in V^{-1}(\rho)$. By Theorem IV.1, we know that V is non-increasing along (9), implying that $V^{-1}(\leq \rho)$ is strongly invariant with respect to (9), and hence (13) too. Therefore, any trajectory of (13) starting in $V^{-1}(\leq \rho)$ is a trajectory of (9). ■

Note that the inclusion in Proposition IV.4 may be strict and that the set of trajectories of (9) is, in general, richer than the set of trajectories of (13). Figure 1 illustrates the effect that increasing K has on (9). From a given initial condition, at some point the value of K is large enough, cf. Proposition IV.4, to make the trajectories of (13) (which never leave $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$) also be a trajectory of (9).

Building on Proposition IV.4, the next result characterizes the asymptotic convergence of (13).

Corollary IV.5 (Asymptotic convergence of the discontinuous saddle-point dynamics). *The trajectories of* (13) *starting in* $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$ *converge asymptotically to a point in* $\mathcal{X} \times \mathcal{Z}$.

Proof: Let V be defined as in Theorem IV.1. Given any initial condition $(x_0, z_0) \in \mathbb{R}^n \times \mathbb{R}^m$, let $t \mapsto (x(t), z(t))$ be a trajectory of (13) starting from (x_0, z_0) and let $\rho = V(x_0, z_0)$. Note that $t \mapsto (x(t), z(t))$ does not depend on K because (13) does not depend on K. Proposition IV.4 establishes that $t \mapsto (x(t), z(t))$ is also a trajectory of (9) for any $K \geq K_1$. Imposing the additional condition that

$$\infty > K > \max\Big\{K_1, \max_{(x_*, z_*) \in (\mathcal{X} \times \mathcal{Z}) \cap V^{-1}(\leq \rho)} \|A^T z_* + c\|_\infty\Big\},$$

Corollary IV.2 implies that the trajectories of (9) (which include $t \mapsto (x(t), z(t))$ converge asymptotically to a point in $\mathcal{X} \times \mathcal{Z}$. ■

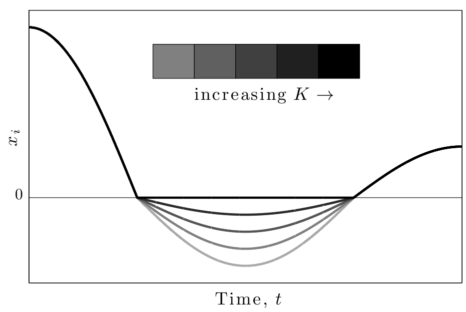


Fig. 1. Illustration of the effect that increasing K has on (9). For a fixed initial condition, the trajectory of (9) has increasingly smaller "incursions" into the region where $x_i < 0$ as K increases, until a finite value is reached where the corresponding trajectory of (13) is also a trajectory of (9).

One can also approach the convergence analysis of (13) from a switched systems perspective, which would require checking that certain regularity conditions hold for the switching behavior of the system. We have been able to circumvent this complexity by relying on the powerful stability tools available for set-valued dynamics to analyze (9) and by relating its solutions with those of (13). Moreover, the interpretation of the trajectories of (13) in the Krasovskii sense is instrumental for our analysis in Section V where we study the robustness against disturbances using powerful Lyapunov-like tools for differential inclusions.

Remark IV.6 (Comparison to existing dynamics for linear programming). Though a central motivation for the development of our linear programming algorithm is the establishment of various robustness properties which we study next, the dynamics (13) and associated convergence results of this section are both novel and have distinct contributions. The work [28] builds on the saddle-point dynamics of a smooth Lagrangian function to introduce an algorithm for linear programming. Instead of exact penalty functions, this approach uses projections to keep the evolution within the feasible set, resulting in a discontinuous dynamics in both the primal and dual variables. The work [27] employs a similar approach to deal with non-strictly convex programs under inequality constraints, where projection is used instead employed to keep nonnegative the value of the dual variables. These works establish convergence in the primal variables ([28] under the assumption that the solution of the linear program is unique, [27] under the assumption that Slater's condition is satisfied) to a solution of the linear program. In both cases, the dual variables converge to some unknown point which might not be a solution to the dual problem. This is to be contrasted with the convergence properties of the dynamics (13) stated in Corollary IV.5 which only require the linear program to be feasible with finite optimal value. •

B. Distributed implementation

An important advantage of the dynamics (13) over other linear programming methods is that it is well-suited for distributed implementation. To make this statement precise, consider a scenario where each component of $x \in \mathbb{R}^n$ corresponds to an independent decision maker or agent and the interconnection between the agents is modeled by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. To see under what conditions the dynamics (13) can be implemented by this multi-agent system, let us express it component-wise. First, the nominal flow function in (13a) for agent $i \in \{1, \dots, n\}$ is,

$$\begin{aligned} f_i^{\text{nom}}(x, z) &= -c_i - \sum_{\ell=1}^m a_{\ell,i}\Big[z_\ell + \sum_{k=1}^n a_{\ell,k}x_k - b_\ell\Big], \\ &= -c_i - \sum_{\{\ell \,:\, a_{\ell,i} \neq 0\}} a_{\ell,i}\Big[z_\ell + \sum_{\{k \,:\, a_{\ell,k} \neq 0\}} a_{\ell,k}x_k - b_\ell\Big], \end{aligned}$$

and the dynamics (13b) for each $\ell \in \{1, \dots, m\}$ is

$$\dot{z}_\ell = \sum_{\{i \,:\, a_{\ell,i} \neq 0\}} a_{\ell,i}x_i - b_\ell. \tag{14}$$

According to these expressions, in order for agent $i \in \{1, \dots, n\}$ to be able to implement its corresponding dynamics in (13a), it also needs access to certain components of z (specifically, those components z_ℓ for which $a_{\ell,i} \neq 0$), and therefore needs to implement their corresponding dynamics (14). We say that the dynamics (13) is *distributed over* $\mathcal{G}$ when the following holds

(D1) for each $i \in \mathcal{V}$, agent i knows
  a) $c_i \in \mathbb{R}$,
  b) every $b_\ell \in \mathbb{R}$ for which $a_{\ell,i} \neq 0$,
  c) the non-zero elements of every row of A for which the i^{th} component, $a_{\ell,i}$, is non-zero,

(D2) agent $i \in \mathcal{V}$ has control over the variable $x_i \in \mathbb{R}$,

(D3) $\mathcal{G}$ is connected with respect to A, and

(D4) agents have access to the variables controlled by neighboring agents.

Note that (D3) guarantees that the agents that implement (14) for a particular $\ell \in \{1, \dots, m\}$ are neighbors in $\mathcal{G}$.

Remark IV.7 (Scalability of the nominal saddle-point dynamics). A different approach to solve (4) is the following: reformulate the optimization problem as the constrained minimization of a sum of convex functions all of the form $\frac{1}{n}c^T x$ and use the algorithms developed in, for instance, [14], [15], [11], [12], [16], for distributed convex optimization. However, in this case, this approach would lead to agents storing and communicating with neighbors estimates of the entire solution vector in $\mathbb{R}^n$, and hence would not scale well with the number of agents of the network. In contrast, to execute the discontinuous saddle-point dynamics, agents only need to store the component of the solution vector that they control and communicate it with neighbors. Therefore, the dynamics scales well with respect to the number of agents in the network. •

V. Robustness against disturbances

Here we explore the robustness properties of the discontinuous saddle-point dynamics (13) against disturbances. Such disturbances may correspond to noise, unmodeled dynamics, or incorrect agent knowledge of the data defining the linear program. Note that the global asymptotic stability of $\mathcal{X} \times \mathcal{Z}$ under (13) characterized in Section IV naturally provides a robustness guarantee on this dynamics: when $\mathcal{X} \times \mathcal{Z}$ is compact, sufficiently small perturbations do not destroy the global asymptotic stability of the equilibria, cf. [29]. Our objective here is to go beyond this qualitative statement to obtain a more precise, quantitative description of robustness. To this end, we consider the notions of input-to-state stability (ISS) and integral-input-to-state stability (iISS). Section V-A shows that, when the disturbances correspond to uncertainty in the problem data, no dynamics for linear programming can be ISS. This motivates us to explore the weaker notion of iISS. Section V-B shows that (13) with additive disturbances is iISS.

Remark V.1 (Robust dynamics versus robust optimization). We make a note of the distinction between the notion of algorithm robustness, which is what we study here, and the term robust (or worst-case) optimization, see e.g., [32]. The latter refers to a type of problem formulation in which some notion of variability (which models uncertainty) is explicitly included in the problem statement. Mathematically,

$$\min \quad c^T x \quad \text{s.t.} \quad f(x, \omega) \leq 0, \ \forall \omega \in \Omega,$$

where ω is an uncertain parameter. Building on the observation that one only has to consider the worst-case values of ω, one can equivalently cast the optimization problem with constraints that only depend on x, albeit at the cost of a loss of structure in the formulation. Another point of connection with the present work is the body of research on stochastic approximation in discrete optimization, where the optimization parameters are corrupted by disturbances, see e.g. [39]. •

Without explicitly stating it from here on, we make the following assumption along the section:

(A) The solution sets to (4) and (5) are compact (i.e., $\mathcal{X} \times \mathcal{Z}$ is compact).

The justification for this assumption is twofold. On the technical side, our study of the iISS properties of (15) in Section V-B builds on a Converse Lyapunov Theorem [29] which requires the equilibrium set to be compact (the question of whether the Converse Lyapunov Theorem holds when the equilibrium set is not compact and the dynamics is discontinuous is an open problem). On the practical side, one can add box-type constraints to (4), ensuring that (A) holds.

We now formalize the disturbance model considered in this section. Let $w = (w_x, w_z) : \mathbb{R}_{\geq 0} \to \mathbb{R}^n \times \mathbb{R}^m$ be locally essentially bounded and enter the dynamics as follows,

$$\dot{x}_i = \begin{cases} f_i^{\text{nom}}(x, z) + (w_x)_i, & \text{if } x_i > 0, \\ \max\{0, f_i^{\text{nom}}(x, z) + (w_x)_i\}, & \text{if } x_i = 0, \end{cases} \ \forall i, \tag{15a}$$

$$\dot{z} = Ax - b + w_z. \tag{15b}$$

For notational purposes, we use $f_{\text{dis}}^w : \mathbb{R}^{2(n+m)} \to \mathbb{R}^{n+m}$ to denote (15). We exploit the fact that f^{nom} is affine to state that the additive disturbance w captures unmodeled dynamics, measurement and computation noise, and any error in an agent's knowledge of the problem data (b or c). For example, if agent $i \in \{1, \dots, n\}$ uses an estimate $\hat{c}_i$ of c_i when computing its dynamics, this can be modeled in (15) by considering $(w_x(t))_i = c_i - \hat{c}_i$. To make precise the correspondence between the disturbance w and uncertainties in the problem data, we provide the following convergence result when the disturbance is constant.

Corollary V.2 (Convergence under constant disturbances). *For constant $\overline{w} = (\overline{w}_x, \overline{w}_z) \in \mathbb{R}^n \times \mathbb{R}^m$, consider the* perturbed linear program*,*

$$\min \quad (c - \overline{w}_x - A^T \overline{w}_z)^T x \tag{16a}$$

$$\text{s.t.} \quad Ax = b - \overline{w}_z, \quad x \geq 0, \tag{16b}$$

and, with a slight abuse in notation, let $\mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w})$ be its primal-dual solution set. Suppose that $\mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w})$ is nonempty. Then each trajectory of (15) *starting in $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$ with constant disturbance $w(t) = \overline{w} = (\overline{w}_x, \overline{w}_z)$ converges asymptotically to a point in $\mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w})$.*

Proof: Note that (15) with disturbance $\overline{w}$ corresponds to the undisturbed dynamics (13) for the perturbed problem (16). Since $\mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w}) \neq \emptyset$, Corollary IV.5 implies the result. ■

A. No dynamics for linear programming is input-to-state stable

The notion of input-to-state stability (ISS) is a natural starting point to study the robustness of dynamical systems against disturbances. Informally, if a dynamics is ISS, then bounded disturbances give rise to bounded deviations from the equilibrium set. Here we show that any dynamics that (i) solve any feasible linear program and (ii) where uncertainties in the problem data (A, b, and c) enter as disturbances is not input-to-state stable (ISS). Our analysis relies on the properties of the solution set of a linear program. To make our discussion precise, we start by recalling the definition of ISS.

Definition V.3 (Input-to-state stability [30]). *The dynamics* (15) *is ISS with respect to* $\mathcal{X}\times\mathcal{Z}$ *if there exist* $\beta\in\mathcal{KL}$ *and* $\gamma\in\mathcal{K}$ *such that, for any trajectory* $t\mapsto(x(t),z(t))$ *of* (15), *one has*

$$\|(x(t),z(t))\|_{\mathcal{X}\times\mathcal{Z}}\leq\beta(\|(x(0),z(0)\|_{\mathcal{X}\times\mathcal{Z}},t)+\gamma(\|w\|_\infty),$$

for all $t\geq 0$. *Here,* $\|w\|_\infty:=\operatorname{esssup}_{s\geq 0}\|w(s)\|$ *is the essential supremum of* $w(t)$.

Our method to show that no dynamics is ISS is constructive. We find a constant disturbance such that the primal-dual solution set to some perturbed linear program is unbounded. Since any point in this unbounded solution set is a stable equilibrium by assumption, this precludes the possibility of the dynamics from being ISS.

Theorem V.4 (No dynamics for linear programming is ISS). *Consider the generic dynamics*

$$(\dot{x},\dot{z})=\Phi(x,z,v)\tag{17}$$

with disturbance $t\mapsto v(t)$. *Assume uncertainties in the problem data are modeled by* v. *That is, there exists a surjective function* $g=(g_1,g_2):\mathbb{R}^{n+m}\to\mathbb{R}^n\times\mathbb{R}^m$ *with* $g(0)=(0,0)$ *such that, for* $\bar{v}\in\mathbb{R}^{n+m}$, *the primal-dual solution set* $\mathcal{X}(\bar{v})\times\mathcal{Z}(\bar{v})$ *of the linear program*

$$\min\quad (c+g_1(\bar{v}))^Tx\tag{18a}$$

$$\text{s.t.}\quad Ax=b+g_2(\bar{v}),\quad x\geq 0.\tag{18b}$$

is the stable equilibrium set of $(\dot{x},\dot{z})=\Phi(x,z,\bar{v})$ *whenever* $\mathcal{X}(\bar{v})\times\mathcal{Z}(\bar{v})\neq\emptyset$. *Then, the dynamics* (17) *is not ISS with respect to* $\mathcal{X}\times\mathcal{Z}$.

Proof: We divide the proof in two cases depending on whether $\{Ax=b,x\geq 0\}$ is (i) unbounded or (ii) bounded. In both cases, we design a constant disturbance $v(t)=\bar{v}$ such that the equilibria of (17) contains points arbitrarily far away from $\mathcal{X}\times\mathcal{Z}$. This would imply that the dynamics is not ISS. Consider case (i). Since $\{Ax=b,x\geq 0\}$ is unbounded, convex, and polyhedral, there exists a point $\hat{x}\in\mathbb{R}^n$ and direction $\nu_x\in\mathbb{R}^n\setminus\{0\}$ such that $\hat{x}+\lambda\nu_x\in\operatorname{bd}(\{Ax=b,x\geq 0\})$ for all $\lambda\geq 0$. Here $\operatorname{bd}(\cdot)$ refers to the boundary of the set. Let $\eta\in\mathbb{R}^n$ be such that $\eta^T\nu_x=0$ and $\hat{x}+\varepsilon\eta\notin\{Ax=b,x\geq 0\}$ for any $\varepsilon>0$ (geometrically, η is normal to and points out of $\{Ax=b,x\geq 0\}$ at $\hat{x}$). Now that these quantities have been defined, consider the following linear program,

$$\min\quad \eta^Tx\quad\text{s.t.}\quad Ax=b,\quad x\geq 0.\tag{19}$$

Because g is surjective, there exists $\bar{v}$ such that $g(\bar{v})=(-c+\eta,0)$. In this case, the program (19) is exactly the program (18), with primal-dual solution set $\mathcal{X}(\bar{v})\times\mathcal{Z}(\bar{v})$. We show next that $\hat{x}$ is a solution to (19) and thus in $\mathcal{X}(\bar{v})$. Clearly, $\hat{x}$ satisfies the constraints of (19). Since $\eta^T\nu_x=0$ and points outward of $\{Ax=b,x\geq 0\}$, it must be that $\eta^T(\hat{x}-x)\leq 0$ for any $x\in\{Ax=b,x\geq 0\}$, which implies that $\eta^T\hat{x}\leq\eta^Tx$. Thus, $\hat{x}$ is a solution to (19). Moreover, $\hat{x}+\lambda\nu_x$ is also a solution to (19) for any $\lambda\geq 0$ since (i) $\eta^T(\hat{x}+\lambda\nu_x)=\eta^T\hat{x}$ and (ii) $\hat{x}+\lambda\nu_x\in\{Ax=b,x\geq 0\}$. That is, $\mathcal{X}(\bar{v})$ is unbounded. Therefore, there is a point $(x_0,z_0)\in\mathcal{X}(\bar{v})\times\mathcal{Z}(\bar{v})$, which is also an equilibrium of (17) by assumption, that is arbitrarily far from the set $\mathcal{X}\times\mathcal{Z}$. Clearly, $t\mapsto(x(t),z(t))=(x_0,z_0)$ is an equilibrium trajectory of (17) starting from (x_0,z_0) when $v(t)=\bar{v}$. The fact that (x_0,z_0) can be made arbitrarily far from $\mathcal{X}\times\mathcal{Z}$ precludes the possibility of the dynamics from being ISS.

Next, we deal with case (ii), when $\{Ax=b,x\geq 0\}$ is bounded. Consider the linear program

$$\max\quad -b^Tz\quad\text{s.t.}\quad A^Tz\geq 0.$$

Since $\{Ax=b,x\geq 0\}$ is bounded, Lemma A.1 implies that $\{A^Tz\geq 0\}$ is unbounded. Using an analogous approach as in case (i), one can find $\eta\in\mathbb{R}^m$ such that the set of solutions to

$$\max\quad \eta^Tz\quad\text{s.t.}\quad A^Tz\geq 0,\tag{20}$$

is unbounded. Because g is surjective, there exists $\bar{v}$ such that $g(\bar{v})=(-c,-b-\eta)$. In this case, the program (20) is the dual to (18), with primal-dual solution set $\mathcal{X}(\bar{v})\times\mathcal{Z}(\bar{v})$. Since $\mathcal{Z}(\bar{v})$ is unbounded, one can find equilibrium trajectories of (17) under the disturbance $v(t)=\bar{v}$ that are arbitrarily far away from $\mathcal{X}\times\mathcal{Z}$, which contradicts ISS. ■

Note that, in particular, the perturbed problem (16) and (18) coincide when

$$g(\overline{w})=g(\overline{w}_x,\overline{w}_z)=(-\overline{w}_x-A^T\overline{w}_z,-\overline{w}_z).$$

Thus, by Theorem V.4, the discontinuous saddle-point dynamics (15) is not ISS. Nevertheless, one can establish an ISS-like result for this dynamics under small enough and constant disturbances. We state this result next, where we also provide a quantifiable upper bound on the disturbances in terms of the solution set of some perturbed linear program.

Proposition V.5 (ISS of discontinuous saddle-point dynamics under small constant disturbances). *Suppose there exists* $\delta>0$ *such that the primal-dual solution set* $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ *of the perturbed problem* (16) *is nonempty for* $\overline{w}\in\overline{\mathbb{B}}(0,\delta)$ *and* $\cup_{\overline{w}\in\overline{\mathbb{B}}(0,\delta)}\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ *is compact. Then there exists a continuous, zero-at-zero, and increasing function* $\gamma:[0,\delta]\to\mathbb{R}_{\geq 0}$ *such that, for all trajectories* $t\mapsto(x(t),z(t))$ *of* (15) *with constant disturbance* $\overline{w}\in\overline{\mathbb{B}}(0,\delta)$, *it holds that*

$$\lim_{t\to\infty}\|(x(t),z(t))\|_{\mathcal{X}\times\mathcal{Z}}\leq\gamma(\|\overline{w}\|).$$

Proof: Let $\gamma:[0,\delta]\to\mathbb{R}_{\geq 0}$ be given by

$$\gamma(r):=\max\Big\{\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}}:(x,z)\in\bigcup_{\overline{w}\in\overline{\mathbb{B}}(0,r)}\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})\Big\}.$$

By hypotheses, γ is well-defined. Note also that γ is increasing and satisfies $\gamma(0)=0$. Next, we show that γ is continuous. By assumption, $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ is nonempty and bounded for every $\overline{w}\in\overline{\mathbb{B}}(0,\delta)$. Moreover, it is clear that $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ is closed for every $\overline{w}\in\overline{\mathbb{B}}(0,\delta)$ since we are considering linear programs in standard form. Thus, $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ is nonempty and compact for every $\overline{w}\in\overline{\mathbb{B}}(0,\delta)$. By [40, Corollary 11], these two conditions are sufficient for the set-valued map $\overline{w}\mapsto\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ to be continuous on $\overline{\mathbb{B}}(0,\delta)$. Since $r\mapsto\overline{\mathbb{B}}(0,r)$

is also continuous, [37, Proposition 1, pp. 41] ensures that the following set-valued composition map

$$r \mapsto \bigcup_{\overline{w} \in \overline{\mathbb{B}}(0,r)} \mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w})$$

is continuous (with compact values, by assumption). Therefore, [37, Theorem 6, pp. 53] guarantees then that γ is continuous on $\overline{\mathbb{B}}(0, \delta)$. Finally, to establish the bound on the trajectories, recall from Corollary V.2 that each trajectory $t \mapsto (x(t), z(t))$ of (15) with constant disturbance $\overline{w} \in \overline{\mathbb{B}}(0, \delta)$ converges asymptotically to a point in $\mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w})$. The distance between $\mathcal{X} \times \mathcal{Z}$ and the point in $\mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w})$ to which the trajectory converges is upper bounded by

$$\lim_{t\to\infty} \|(x(t), z(t))\|_{\mathcal{X}\times\mathcal{Z}} \leq \max\big\{\|(x, z)\|_{\mathcal{X}\times\mathcal{Z}} : (x, z) \in \mathcal{X}(\overline{w}) \times \mathcal{Z}(\overline{w})\big\} \leq \gamma(\|\overline{w}\|),$$

which concludes the proof. ∎

B. Discontinuous saddle-point dynamics is integral input-to-state stable

Here we establish that the dynamics (15) possess a notion of robustness weaker than ISS, namely, integral input-to-state stability (iISS). Informally, iISS guarantees that disturbances with small energy give rise to small deviations from the equilibria. This is stated formally next.

Definition V.6 (Integral input-to-state stability [31]). *The dynamics* (15) *is iISS with respect to the set $\mathcal{X} \times \mathcal{Z}$ if there exist functions $\alpha \in \mathcal{K}_\infty, \beta \in \mathcal{KL}$, and $\gamma \in \mathcal{K}$ such that, for any trajectory $t \mapsto (x(t), z(t))$ of* (15) *and all $t \geq 0$, one has*

$$\alpha(\|(x(t), z(t))\|_{\mathcal{X}\times\mathcal{Z}}) \leq \beta(\|(x(0), z(0)\|_{\mathcal{X}\times\mathcal{Z}}, t) + \int_0^t \gamma(\|w(s)\|)ds. \tag{21}$$

Our ensuing discussion is based on a suitable adaptation of the exposition in [31] to the setup of asymptotically stable sets for discontinuous dynamics. A useful tool for establishing iISS is the notion of iISS Lyapunov function, whose definition we review next.

Definition V.7 (iISS Lyapunov function). *A differentiable function $V : \mathbb{R}^{n+m} \to \mathbb{R}_{\geq 0}$ is an iISS Lyapunov function with respect to the set $\mathcal{X} \times \mathcal{Z}$ for dynamics* (15) *if there exist functions $\alpha_1, \alpha_2 \in \mathcal{K}_\infty, \sigma \in \mathcal{K}$, and a continuous positive definite function α_3 such that*

$$\alpha_1(\|(x, z)\|_{\mathcal{X}\times\mathcal{Z}}) \leq V(x, z) \leq \alpha_2(\|(x, z)\|_{\mathcal{X}\times\mathcal{Z}}), \tag{22a}$$

$$a \leq -\alpha_3(\|(x, z)\|_{\mathcal{X}\times\mathcal{Z}}) + \sigma(\|w\|), \tag{22b}$$

for all $a \in \mathcal{L}_{\mathscr{K}[f^w_{dis}]} V(x, z)$ and $x \in \mathbb{R}^n, z \in \mathbb{R}^m, w \in \mathbb{R}^{n+m}$.

Note that, since the set $\mathcal{X} \times \mathcal{Z}$ is compact (cf. Assumption (A)), (22a) is equivalent to V being proper with respect to $\mathcal{X} \times \mathcal{Z}$. The existence of an iISS Lyapunov function is critical in establishing iISS, as the following result states.

Theorem V.8 (iISS Lyapunov function implies iISS). *If there exists an iISS Lyapunov function with respect to $\mathcal{X} \times \mathcal{Z}$ for* (15)*, then the dynamics is iISS with respect to $\mathcal{X} \times \mathcal{Z}$.*

This result is stated in [31, Theorem 1] for the case of differential equations with locally Lipschitz right-hand side and asymptotically stable origin, but its extension to discontinuous dynamics and asymptotically stable sets, as considered here, is straightforward. We rely on Theorem V.8 to establish that the discontinuous saddle-point dynamics (15) is iISS. Interestingly, the function V employed to characterize the convergence properties of the unperturbed dynamics in Section IV is not an iISS Lyapunov function (in fact, our proof of Theorem IV.1 relies on the set-valued LaSalle Invariance Principle because, essentially, the Lie derivative of V is not negative definite). Nevertheless, in the proof of the next result, we build on the properties of this function with respect to the dynamics to identify a suitable iISS Lyapunov function for (15).

Theorem V.9 (iISS of saddle-point dynamics). *The dynamics* (15) *is iISS with respect to $\mathcal{X} \times \mathcal{Z}$.*

Proof: We proceed by progressively defining functions V_{euc}, $V^{\text{rep}}_{\text{euc}}$, V_{CLF}, and $V^{\text{rep}}_{\text{CLF}} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. The rationale for our construction is as follows. Our starting point is the squared Euclidean distance from the primal-dual solution set, denoted V_{euc}. The function $V^{\text{rep}}_{\text{euc}}$ is a reparameterization of V_{euc} (which remains radially unbounded with respect to $\mathcal{X} \times \mathcal{Z}$) so that state and disturbance appear separately in the (set-valued) Lie derivative. However, since V_{euc} is only a LaSalle-type function, this implies that only the disturbance appears in the Lie derivative of $V^{\text{rep}}_{\text{euc}}$. Nevertheless, via a Converse Lyapunov Theorem, we identify an additional function V_{CLF} whose reparameterization $V^{\text{rep}}_{\text{CLF}}$ has a Lie derivative where both state and disturbance appear. The function $V^{\text{rep}}_{\text{CLF}}$, however, may not be radially unbounded with respect to $\mathcal{X} \times \mathcal{Z}$. This leads us to the construction of the iISS Lyapunov function as $V = V^{\text{rep}}_{\text{euc}} + V^{\text{rep}}_{\text{CLF}}$.

We begin by defining the differentiable function V_{euc}

$$V_{\text{euc}}(x, z) = \min_{(x_*, z_*) \in \mathcal{X}\times\mathcal{Z}} \frac{1}{2}(x - x_*)^T(x - x_*) + \frac{1}{2}(z - z_*)^T(z - z_*).$$

Since $\mathcal{X} \times \mathcal{Z}$ is convex and compact, applying Theorem A.2 one gets $\nabla V_{\text{euc}}(x, z) = (x - x_*(x, z), z - z_*(x, z))$, where

$$(x_*(x, z), z_*(x, z)) = \operatorname*{argmin}_{(x_*, z_*) \in \mathcal{X}\times\mathcal{Z}} \frac{1}{2}(x - x_*)^T(x - x_*) + \frac{1}{2}(z - z_*)^T(z - z_*).$$

It follows from Theorem IV.1 and Proposition IV.4 that $\mathcal{L}_{\mathscr{K}[f_{\text{dis}}]} V_{\text{euc}}(x, z) \subset (-\infty, 0]$ for all $(x, z) \in \mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$. Next, similar to the approach in [31], define the function $V^{\text{rep}}_{\text{euc}}$ by

$$V^{\text{rep}}_{\text{euc}}(x, z) = \int_0^{V_{\text{euc}}(x,z)} \frac{dr}{1+\sqrt{2r}}.$$

Clearly, $V_{\rm euc}^{\rm rep}(x,z)$ is positive definite with respect to $\mathcal{X}\times\mathcal{Z}$. Also, $V_{\rm euc}^{\rm rep}(x,z)$ is radially unbounded with respect to $\mathcal{X}\times\mathcal{Z}$ because (i) $V_{\rm euc}(x,z)$ is radially unbounded with respect to $\mathcal{X}\times\mathcal{Z}$ and (ii) $\lim_{y\to\infty}\int_0^y \frac{dr}{1+\sqrt{2r}}=\infty$. In addition, for any $a\in\mathcal{L}_{\mathscr{K}[f_{\rm dis}^w]}V_{\rm euc}^{\rm rep}(x,z)$ and $(x,z)\in\mathbb{R}^n_{\geq 0}\times\mathbb{R}^m$, one has

$$a\leq\frac{\sqrt{2V_{\rm euc}(x,z)}\|w\|}{1+\sqrt{2V_{\rm euc}(x,z)}}\leq\|w\|. \tag{23}$$

Next, we define the function $V_{\rm CLF}$. Since $\mathcal{X}\times\mathcal{Z}$ is compact and globally asymptotically stable for (13) $(\dot x,\dot z)=\mathscr{K}[f_{\rm dis}^w](x,z)$ when $w\equiv 0$ (cf. Corollary IV.5) the Converse Lyapunov Theorem [29, Theorem 3.13] ensures the existence of a smooth function $V_{\rm CLF}:\mathbb{R}^{n+m}\to\mathbb{R}_{\geq 0}$ and class $\mathcal{K}_\infty$ functions $\tilde\alpha_1$, $\tilde\alpha_2$, $\tilde\alpha_3$ such that

$$\begin{aligned}\tilde\alpha_1(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})\leq V_{\rm CLF}(x,z)\leq\tilde\alpha_2(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}}),\\ a\leq-\tilde\alpha_3(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}}),\end{aligned}$$

for all $a\in\mathcal{L}_{\mathscr{K}[f_{\rm dis}]}V_{\rm CLF}(x,z)$ and $(x,z)\in\mathbb{R}^n_{\geq 0}\times\mathbb{R}^m$. Thus, when $w\not\equiv 0$, for $a\in\mathcal{L}_{\mathscr{K}[f_{\rm dis}^w]}V_{\rm CLF}(x,z)$ and $(x,z)\in\mathbb{R}^n_{\geq 0}\times\mathbb{R}^m$, we have

$$\begin{aligned}a&\leq-\tilde\alpha_3(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})+\nabla V_{\rm CLF}(x,z)w,\\ &\leq-\tilde\alpha_3(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})+\|\nabla V_{\rm CLF}(x,z)\|\cdot\|w\|,\\ &\leq-\tilde\alpha_3(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})+(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}}\\ &\quad+\|\nabla V_{\rm CLF}(x,z)\|)\cdot\|w\|,\\ &\leq-\tilde\alpha_3(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})+\lambda(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})\cdot\|w\|,\end{aligned}$$

where $\lambda:[0,\infty)\to[0,\infty)$ is given by

$$\lambda(r)=r+\max_{\|\eta\|_{\mathcal{X}\times\mathcal{Z}}\leq r}\|\nabla V_{\rm CLF}(\eta)\|.$$

Since $V_{\rm CLF}$ is smooth, λ is a class $\mathcal{K}$ function. Next, define

$$V_{\rm CLF}^{\rm rep}(x,z)=\int_0^{V_{\rm CLF}(x,z)}\frac{dr}{1+\lambda\circ\tilde\alpha_1^{-1}(r)}.$$

Without additional information about $\lambda\circ\tilde\alpha_1^{-1}$, one cannot determine if $V_{\rm CLF}^{\rm rep}$ is radially unbounded with respect to $\mathcal{X}\times\mathcal{Z}$ or not. Nevertheless, $V_{\rm CLF}^{\rm rep}$ is positive definite with respect to $\mathcal{X}\times\mathcal{Z}$. Then for any $a\in\mathcal{L}_{\mathscr{K}[f_{\rm dis}^w]}V_{\rm CLF}^{\rm rep}(x,z)$ and $(x,z)\in\mathbb{R}^n_{\geq 0}\times\mathbb{R}^m$ we have,

$$\begin{aligned}a&\leq\frac{-\tilde\alpha_3(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})+\nabla V_{\rm CLF}(x,z)w}{1+\lambda\circ\tilde\alpha_1^{-1}(V_{\rm CLF}(x,z))},\\ &\leq\frac{-\tilde\alpha_3(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})}{1+\lambda\circ\tilde\alpha_1^{-1}\circ\tilde\alpha_2(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})}+\frac{\lambda(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})}{1+\lambda(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})}\|w\|\\ &\leq-\rho(\|(x,z)\|_{\mathcal{X}\times\mathcal{Z}})+\|w\|,\end{aligned} \tag{24}$$

where ρ is the positive definite function given by

$$\rho(r)=\tilde\alpha_3(r)/(1+\lambda\circ\tilde\alpha_1^{-1}\circ\tilde\alpha_2(r)).$$

and we have used the fact that $\tilde\alpha_1^{-1}$ and $\tilde\alpha_2$ are positive definite. We now show that $V=V_{\rm euc}^{\rm rep}+V_{\rm CLF}^{\rm rep}$ is an iISS Lyapunov function for (15) with respect to $\mathcal{X}\times\mathcal{Z}$. First, (22a) is satisfied because V is positive definite and radially unbounded with respect to $\mathcal{X}\times\mathcal{Z}$ since (i) $V_{\rm euc}^{\rm rep}$ is positive definite and radially unbounded with respect to $\mathcal{X}\times\mathcal{Z}$ and (ii) $V_{\rm CLF}^{\rm rep}$ is positive definite with respect to $\mathcal{X}\times\mathcal{Z}$. Second, (22b) is satisfied as a result of the combination of (23) and (24). Since V satisfies the conditions of Theorem V.8, (15) is iISS. ∎

Based on the discussion in Section V-A, the iISS property of (15) is an accurate representation of the robustness of the dynamics, not a limitation of our analysis. A consequence of iISS is that the asymptotic convergence of the dynamics is preserved under finite energy disturbances [41, Proposition 6]. In the case of (15), a stronger convergence property is true under finite variation disturbances (which do not have finite energy). The following formalizes this fact.

Corollary V.10 (Finite variation disturbances). *Let* $w:\mathbb{R}_{\geq 0}\to\mathbb{R}^n\times\mathbb{R}^m$ *be such that* $\int_0^\infty\|w(s)-\overline{w}\|ds<\infty$ *for some* $\overline{w}=(\overline{w}_x,\overline{w}_z)\in\mathbb{R}^n\times\mathbb{R}^m$. *Assume* $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ *is nonempty and compact. Then each trajectory of* (15) *under the disturbance* w *converges asymptotically to a point in* $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$.

Proof: Let $f_{\rm dis,pert}^v$ be the discontinuous saddle-point dynamics derived for the perturbed program (16) associated to $\overline{w}$ with additive disturbance $v:\mathbb{R}_{\geq 0}\to\mathbb{R}^n\times\mathbb{R}^m$. By Corollary V.2, $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})\neq\emptyset$ is globally asymptotically stable for $f_{\rm dis,pert}^0$. Additionally, by Theorem V.9 and since $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ is compact, $f_{\rm dis,pert}^v$ is iISS. As a consequence, by [41, Proposition 6], each trajectory of $f_{\rm dis,pert}^v$ converges asymptotically to a point in $\mathcal{X}(\overline{w})\times\mathcal{Z}(\overline{w})$ if $\int_0^\infty\|v(s)\|ds<\infty$. The result follows by noting that $f_{\rm dis}^w$ with disturbance w is exactly $f_{\rm dis,pert}^v$ with disturbance $v=w-\overline{w}$ and that, by assumption, the latter satisfies $\int_0^\infty\|v(s)\|ds<\infty$. ∎

VI. Robustness in recurrently connected graphs

Here, we build on the iISS properties of the saddle-point dynamics (9) to study its convergence under communication link failures. As such, agents do not receive updated state information from their neighbors at all times and use the last known value of their state to implement the dynamics. The link failure model we considered is described by recurrently connected graphs (RCG), in which periods of communication loss are followed by periods of connectivity, formalized next.

Definition VI.1 (Recurrently connected graphs). *Given a strictly increasing sequence of times* $\{t_k\}_{k=0}^\infty\subset\mathbb{R}_{\geq 0}$ *and a base graph* $\mathcal{G}_b=(\mathcal{V},\mathcal{E}_b)$, *we call* $\mathcal{G}(t)=(\mathcal{V},\mathcal{E}(t))$ recurrently connected with respect to $\mathcal{G}_b$ and $\{t_k\}_{k=0}^\infty$ *if* $\mathcal{E}(t)\subseteq\mathcal{E}_b$ *for all* $t\in[t_{2k},t_{2k+1})$ *while* $\mathcal{E}(t)\supseteq\mathcal{E}_b$ *for all* $t\in[t_{2k+1},t_{2k+2})$, $k\in\mathbb{Z}_{\geq 0}$.

Intuitively, one may think of $\mathcal{G}_b$ as a graph over which (13) is distributed: during time intervals of the form $[t_{2k},t_{2k+1})$, links are failing and hence the network cannot execute the algorithm properly, whereas during time intervals of the form $[t_{2k+1},t_{2k+2})$, enough communication links are available to implement it correctly. In what follows, and for simplicity of presentation, we only consider the worst-case link failure scenario: i.e., if a link fails during the time interval $[t_{2k},t_{2k+1})$, it remains down during its entire duration. The results stated here also apply to the general scenarios where edges may fail and reconnect multiple times within a time interval.

In the presence of link failures, the implementation of the evolution of the z variables, cf. (14), across different agents would yield in general different outcomes (given that different agents have access to different information at different times). To avoid this problem, we assume that, for each $\ell \in \{1, \dots, m\}$, the agent with minimum identifier index,

$$j = \mathbb{S}(\ell) := \min\{i \in \{1, \dots, n\} : a_{\ell,i} \neq 0\},$$

implements the z_ℓ-dynamics and communicates this value when communication is available to its neighbors. Incidentally, only neighbors of $j = \mathbb{S}(\ell)$ need to know z_ℓ. With this convention in place, we may describe the network dynamics under link failures. Let $\mathbb{F}(k)$ be the set of failing communication edges for $t \in [t_k, t_{k+1})$. In other words, if $(i, j) \in \mathbb{F}(k)$ then agents i and j do not receive updated state information from each other during the whole interval $[t_k, t_{k+1})$. The nominal flow function of i on a RCG for $t \in [t_k, t_{k+1})$ is

$$\begin{aligned} f_i^{\mathrm{nom,RCG}}(x, z) = -c_i - \sum_{\substack{\ell=1 \\ (i,\mathbb{S}(\ell))\notin\mathbb{F}(k)}}^{m} a_{\ell,i} z_\ell - \sum_{\substack{\ell=1 \\ (i,\mathbb{S}(\ell))\in\mathbb{F}(k)}}^{m} a_{\ell,i} z_\ell(t_k) \\ - \sum_{\ell=1}^{m} a_{\ell,i} \Big[\sum_{\substack{j=1 \\ (i,j)\notin\mathbb{F}(k)}}^{n} a_{\ell,j} x_j + \sum_{\substack{j=1 \\ (i,j)\in\mathbb{F}(k)}}^{n} a_{\ell,j} x_j(t_k) - b_\ell \Big]. \end{aligned}$$

Thus the x_i-dynamics during $[t_k, t_{k+1})$ for $i \in \{1, \dots, n\}$ is

$$\dot{x}_i = \begin{cases} f_i^{\mathrm{nom,RCG}}(x, z), & \text{if } x_i > 0, \\ \max\{0, f_i^{\mathrm{nom,RCG}}(x, z)\}, & \text{if } x_i = 0. \end{cases} \tag{25a}$$

Likewise, the z-dynamics for $\ell \in \{1, \dots, m\}$ is

$$\dot{z}_\ell = \sum_{\substack{i=1 \\ (i,\mathbb{S}(\ell))\notin\mathbb{F}(k)}}^{n} a_{\ell,i} x_i + \sum_{\substack{i=1 \\ (i,\mathbb{S}(\ell))\in\mathbb{F}(k)}}^{n} a_{\ell,i} x_i(t_k) - b_\ell. \tag{25b}$$

It is worth noting that (25) and (13) coincide when $\mathbb{F}(k) = \emptyset$. The next result shows that the discontinuous saddle-point dynamics still converge under recurrently connected graphs.

Proposition VI.2 (Convergence of saddle-point dynamics under RCGs). *Let $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ be recurrently connected with respect to $\mathcal{G}_b = (\mathcal{V}, \mathcal{E}_b)$ and $\{t_k\}_{k=0}^{\infty}$. Suppose that* (25) *is distributed over $\mathcal{G}_b$ and $T^{\max}_{\mathrm{disconnected}} := \sup_{k\in\mathbb{Z}_{\geq 0}}(t_{2k+1} - t_{2k}) < \infty$. Let $t \mapsto (x(t), z(t))$ be a trajectory of* (25)*. Then there exists $T^{\min}_{\mathrm{connected}} > 0$ (depending on $T^{\max}_{\mathrm{disconnected}}$, $x(t_0)$, and $z(t_0)$) such that $\inf_{k\in\mathbb{Z}_{\geq 0}}(t_{2k+2} - t_{2k+1}) > T^{\min}_{\mathrm{connected}}$ implies that $\|(x(t_{2k}), z(t_{2k}))\|_{\mathcal{X}\times\mathcal{Z}} \to 0$ as $k \to \infty$.*

Proof: The proof method is to (i) show that trajectories of (25) do not escape in finite time and (ii) use a $\mathcal{KL}$ characterization of asymptotically stable dynamics [29] to find $T^{\min}_{\mathrm{connected}}$ for which $\|(x(t_{2k}), z(t_{2k}))\|_{\mathcal{X}\times\mathcal{Z}} \to 0$ as $k \to \infty$. To prove (i), note that (25) represents a switched system of affine differential equations. The modes are defined by all κ-combinations of link failures (for $\kappa = 1, \dots, |\mathcal{E}_b|$) and all κ-combinations of agents (for $\kappa = 1, \dots, n$). Thus, the number of modes is $d := 2^{|\mathcal{E}_b|+n}$. Assign to each mode a number in the set $\{1, \dots, d\}$. Then, for any given $t \in [t_k, t_{k+1})$, the dynamics (25) is equivalently represented as

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = P_{\sigma(t)} \begin{bmatrix} x \\ z \end{bmatrix} + q_{\sigma(t)}(x(t_k), z(t_k)),$$

where $\sigma : \mathbb{R}_{\geq 0} \to \{1, \dots, d\}$ is a switching law and $P_{\sigma(t)}$ (resp. $q_{\sigma(t)}$) is the flow matrix (resp. drift vector) of (25) for mode $\sigma(t)$. Let $\rho = \|(x(t_0), z(t_0))\|_{\mathcal{X}\times\mathcal{Z}}$ and define

$$\tilde{q} := \max_{\substack{p\in\{1,\dots,d\} \\ \|(x,z)\|_{\mathcal{X}\times\mathcal{Z}}\leq\rho}} \|q_p(x, z)\|, \quad \text{and} \quad \tilde{\mu} := \max_{p\in\{1,\dots,d\}} \mu(P_p),$$

where $\mu(P_p) = \lim_{h\to 0^+} \frac{\|I - hP_p\|^{-1}}{h}$ is the logarithmic norm of P_p. Both $\tilde{q}$ and $\tilde{\mu}$ are finite. Consider an arbitrary interval $[t_{2k}, t_{2k+1})$ where $\|(x(t_{2k}), z(t_{2k}))\|_{\mathcal{X}\times\mathcal{Z}} \leq \rho$. In what follows, we make use of the fact that the trajectory of an affine differential equation $\dot{y} = \mathcal{A}y + \beta$ for $t \geq t_0$ is

$$y(t) = e^{\mathcal{A}(t-t_0)} y(t_0) + \int_{t_0}^{t} e^{\mathcal{A}(t-s)} \beta ds. \tag{26}$$

Applying (26), we derive the following bound,

$$\begin{aligned} &\|(x(t_{2k+1}), z(t_{2k+1})) - (x(t_{2k}), z(t_{2k}))\| \\ &\leq \|(x(t_{2k}), z(t_{2k}))\| (e^{\tilde{\mu}(t_{2k+1}-t_{2k})} - 1) + \int_{t_{2k}}^{t_{2k+1}} e^{\tilde{\mu}(t_{2k+1}-s)} \tilde{q} ds, \\ &\leq (\rho + \tilde{q}/\tilde{\mu})(e^{\tilde{\mu} T^{\max}_{\mathrm{disconnected}}} - 1) =: M. \end{aligned}$$

In words, M bounds the distance that trajectories travel on intervals of link failures. Also, M is valid for all such intervals where $\|(x(t_{2k}), z(t_{2k}))\|_{\mathcal{X}\times\mathcal{Z}} \leq \rho$. Next, we address the proof of (ii) by designing $T^{\min}_{\mathrm{connected}}$ to enforce this condition. By definition, $\|(x(t_0), z(t_0))\|_{\mathcal{X}\times\mathcal{Z}} = \rho$. Thus, $\|(x(t_1), z(t_1)) - (x(t_0), z(t_0))\| \leq M$. Given that $\mathcal{X} \times \mathcal{Z}$ is globally asymptotically stable for (25) if $\mathbb{F}(k) = \emptyset$ (cf. Theorem V.9), [29, Theorem 3.13] implies the existence of $\beta \in \mathcal{KL}$ such that

$$\|(x(t), z(t))\|_{\mathcal{X}\times\mathcal{Z}} \leq \beta(\|(x(t_0), z(t_0))\|_{\mathcal{X}\times\mathcal{Z}}, t).$$

By [41, Proposition 7], there exist $\theta_1, \theta_2 \in \mathcal{K}_\infty$ such that $\beta(s, t) \leq \theta_1(\theta_2(s)e^{-t})$. Thus,

$$\begin{aligned} &\alpha(\|(x(t_2), z(t_2))\|_{\mathcal{X}\times\mathcal{Z}}) \\ &\quad \leq \theta_1(\theta_2(\|(x(t_1), z(t_1))\|_{\mathcal{X}\times\mathcal{Z}}) e^{-t_2+t_1}) \\ &\quad \leq \theta_1(\theta_2(\rho + M) e^{-t_2+t_1}). \end{aligned}$$

Consequently, if

$$t_2 - t_1 > T^{\min}_{\mathrm{connected}} := \ln\left(\frac{\theta_2(\rho + M)}{\theta_1^{-1}(\alpha(\rho))}\right) > 0,$$

then $\|(x(t_2), z(t_2))\|_{\mathcal{X}\times\mathcal{Z}} < \rho$. Repeating this analysis reveals that $\|(x(t_{2k+2}), z(t_{2k+2}))\|_{\mathcal{X}\times\mathcal{Z}} < \|(x(t_{2k}), z(t_{2k}))\|_{\mathcal{X}\times\mathcal{Z}}$ for all $k \in \mathbb{Z}_{\geq 0}$ when $t_{2k+2} - t_{2k+1} > T^{\min}_{\mathrm{connected}}$. Thus $\|(x(t_{2k}), z(t_{2k}))\|_{\mathcal{X}\times\mathcal{Z}} \to 0$ as $k \to \infty$ as claimed. ■

Remark VI.3 (More general link failures). Proposition VI.2 shows that, as long as the communication graph is connected with respect to A for a sufficiently long time after periods of failure, the discontinuous saddle-point dynamics converge. We have observed in simulations, however, that the dynamics is not robust to more general link failures such as when the communication graph is never connected with respect to A but its union over time is. We believe the reason is the lack of consistency in the $z-$dynamics for all time across agents in this case. •

VII. Simulations

Here we illustrate the convergence and robustness properties of the discontinuous saddle-point dynamics. We consider a finite-horizon optimal control problem for a network of agents with coupled dynamics and underactuation. The network-wide dynamics is open-loop unstable and the aim of the agents is to find a control to minimize the actuation effort and ensure the network state remains small. To achieve this goal, the agents use the discontinuous saddle-point dynamics (13). Formally, consider the finite-horizon optimal control problem,

$$\min \quad \sum_{\tau=0}^{T} \|x(\tau+1)\|_1 + \|u(\tau)\|_1 \tag{27a}$$

$$\text{s.t.} \quad x(\tau+1) = Gx(\tau) + Hu(\tau), \quad \tau = 0, \ldots T, \tag{27b}$$

where $x(\tau) \in \mathbb{R}^N$ and $u(\tau) \in \mathbb{R}^N$ is the network state and control, respectively, at time τ. The initial point $x_i(0)$ is known to agent i and its neighbors. The matrices $G \in \mathbb{R}^{N\times N}$ and $H = \operatorname{diag}(h) \in \mathbb{R}^{N\times N}$, $h \in \mathbb{R}^N$, define the network evolution, and the network topology is encoded in the sparsity structure of G. We interpret each agent as a subsystem whose dynamics is influenced by the states of neighboring agents. An agent knows the dynamics of its own subsystem and its neighbor's subsystem, but does not know the entire network dynamics. A solution to (27) is a time history of optimal controls $(u_*(0), \ldots, u_*(T)) \in (\mathbb{R}^N)^T$.

To express this problem in standard linear programming form (4), we split the states into their positive and negative components, $x(\tau) = x^+(\tau) - x^-(\tau)$, with $x^+(\tau), x^-(\tau) \geq 0$ (and similarly for the inputs $u(\tau)$). Then, (27) can be equivalently formulated as the following linear program,

$$\min \sum_{\tau=0}^{T}\sum_{i=1}^{N} x_i^+(\tau+1) + x_i^-(\tau+1) + u_i^+(\tau) + u_i^-(\tau) \tag{28a}$$

$$\begin{aligned} \text{s.t.} \ \ & x^+(\tau+1) - x^-(\tau) = G(x^+(\tau) - x^-(\tau)) \\ & \qquad + H(u^+(\tau) - u^-(\tau)), \quad \tau = 0, \ldots, T \end{aligned} \tag{28b}$$

$$x^+(\tau+1), x^-(\tau+1), u^+(\tau), u^-(\tau) \geq 0, \quad \forall \tau \tag{28c}$$

The optimal control for (27) at time τ is then $u_*(\tau) = u_*^+(\tau) - u_*^-(\tau)$, where the vector $(u_*^+(0), u_*^-(0), \ldots, u_*^+(T), u_*^-(T))$ is a solution to (28), cf. [42, Lemma 6.1].

We implement the discontinuous saddle-point dynamics (13) for problem (28) over the network of 5 agents described in Figure 2. To implement the dynamics (13), neighboring agents must exchange their state information with each other. In this example, each agent is responsible for $2(T+1) = 24$ variables, which is independent of the network size. This is in contrast to consensus-based distributed optimization algorithms, where each agent would be responsible for $2N(T+1) = 120$ variables, which grows linearly with the network size N. For simulation purposes, we implement the dynamics as a single program in MATLAB®, using a first-order (Euler) approximation of the differential equation with a stepsize of 0.01. The CPU time for the simulation is $3.1824s$ on a 64-bit 3GHz Intel® Core™ i7-3540M processor with 16GB of installed RAM.

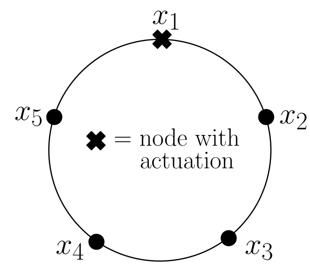
Note that, when implementing this dynamics, agent $i \in \{1, \ldots, 5\}$ computes the time history of its optimal control, $u_i^-(0), u_i^+(0), \ldots, u_i^-(T), u_i^+(T)$, as well as the time history of its states, $x_i^-(1), x_i^+(1), \ldots, x_i^-(T+1), x_i^+(T+1)$. With respect to the solution of the optimal control problem, the time history of states are auxiliary variables used in the discontinuous dynamics and can be discarded after the control is determined. Figure 3 shows the results of the implementation of (13) when a finite energy noise signal disturbs the agents' execution. Clearly (13) achieves convergence initially in the absence of noise. Then, the finite energy noise signal in Figure 3(b) enters each agents' dynamics and disrupts this convergence, albeit not significantly due to the iISS property of (15) characterized in Theorem V.9. Once the noise disappears in the agents' computation of its optimal control, convergence of the algorithm ensues. The constraint violation is plotted in Figure 3(c). Once the time history of optimal controls has been computed (corresponding to the steady-state values in Figure 3(a)), agent 1 implements it, and the resulting network evolution is displayed in Figure 3(d). Agent 1 is able to drive the system state to zero, despite it being open-loop unstable. Figure 4 shows the results of implementation in a recurrently connected communication graph and (13) still achieves convergence as characterized in Proposition VI.2. The link failure model here is a random number of random links failing during times of disconnection. The graph is repeatedly connected for $1s$ and then disconnected for $4s$ (i.e., the ratio $T_{\text{disconnected}}^{\max} : T_{\text{connected}}^{\min}$ is $4:1$). The fact that convergence is still achieved under this unfavorable ratio highlights the strong robustness properties of the algorithm.

VIII. Conclusions

We have considered a network of agents whose objective is to have the aggregate of their states converge to a solution of a general linear program. We proposed an equivalent formulation of this problem in terms of finding the saddle points of a modified Lagrangian function. To make an exact correspondence between the solutions of the linear program and saddle points of the Lagrangian we incorporate a nonsmooth penalty term. This formulation has naturally led us to study the associated saddle-point dynamics, for which we established the point-wise convergence to the set of solutions of the linear program. Based on this analysis, we introduced an alternative algorithmic solution with the same asymptotic convergence properties. This dynamics is amenable to distributed implementation over a multi-agent system, where each individual controls its own component of the solution vector and shares its value with its neighbors. We also studied the robustness against disturbances and link failures of this dynamics. We showed that it is integral-input-to-state stable but not input-to-state stable (and, in fact, no algorithmic solution for linear programming is). These results have allowed us to formally establish the resilience of our

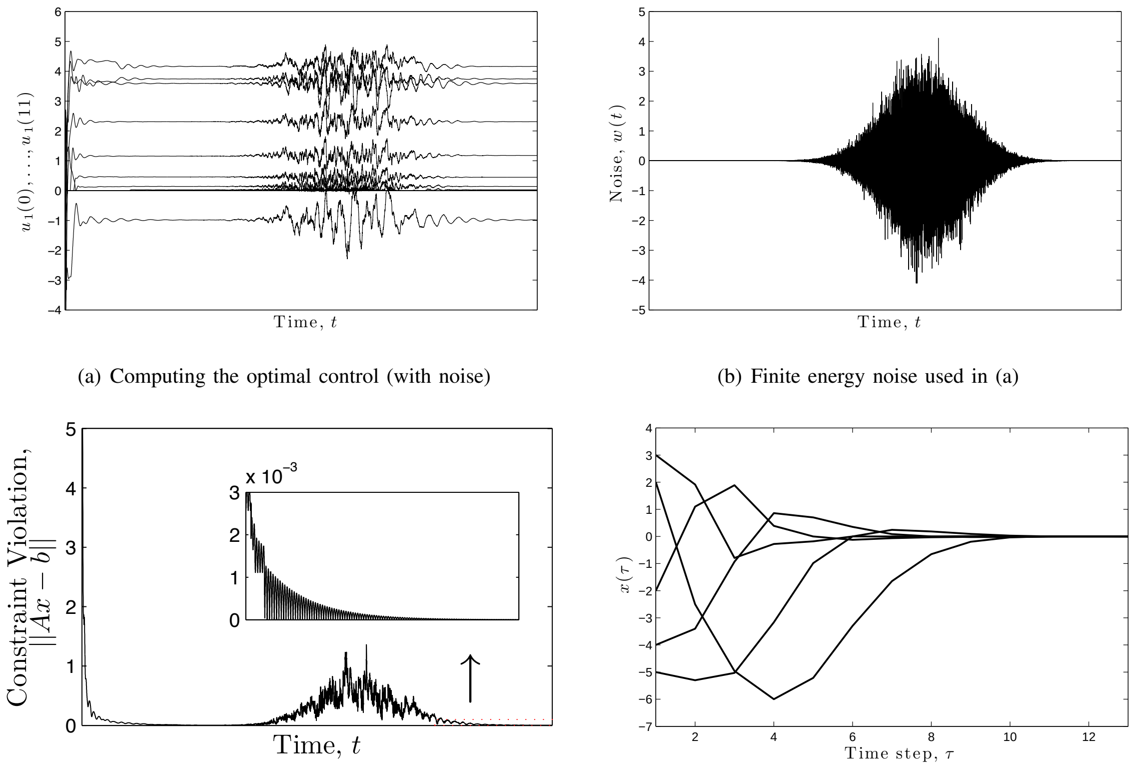
$$x(\tau+1) = \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0.7 \\ 0.7 & 0.5 & 0 & 0 & 0 \\ 0 & 0.7 & 0.5 & 0 & 0 \\ 0 & 0 & 0.7 & 0.5 & 0 \\ 0 & 0 & 0 & 0.7 & 0.5 \end{bmatrix} x(\tau) + \operatorname{diag}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right) u(\tau)$$

(a) Network dynamics



(b) Communication topology

Fig. 2. Network dynamics and communication topology of the multi-agent system. The network dynamics is underactuated and open-loop unstable but controllable. The presence of a communication link in (b) among every pair of agents whose dynamics are coupled in (a) ensures that the algorithm (13) is distributed over the communication graph.



(a) Computing the optimal control (with noise)

(b) Finite energy noise used in (a)

(c) Equality constraint violation in (a)

(d) Network evolution under optimal control found in (a)

Fig. 3. Plot (a) shows the trajectories of the discontinuous saddle-point dynamics (15) subject to the noise depicted in (b) for agent 1 as it computes its time history of optimal controls. Plot (c) shows the associated equality constraint violation. The asymptotic convergence of the trajectories appears to be exponential. The time horizon of the optimal control problem (28) is $T = 11$. The 12 trajectories in (a) and (b) represent agent 1's evolving estimates of the optimal controls $u_1(0), \ldots, u_1(11)$. The steady-state values achieved by these trajectories correspond to the solution of (27). Once determined, these controls are then implemented by agent 1 and result in the network evolution depicted in (d). The dynamics is initialized to a random point.

distributed dynamics to disturbances of finite variation and recurrently disconnected communication graphs. Future work will include the study of the convergence rate of the dynamics and its robustness properties under more general link failures, the synthesis of continuous-time computation models with opportunistic discrete-time communication among agents, and the extension of our design to other convex optimization problems. We also plan to explore the benefits of the proposed distributed dynamics in a number of engineering scenarios, including the smart grid and power distribution, bargaining and matching in networks, and model predictive control.

Acknowledgments

The authors would like to acknowledge helpful comments from Simon Niederlaender, in particular regarding the proofs of Theorem IV.1 and Corollary IV.2. This research was supported by Award FA9550-10-1-0499 and NSF award CMMI-1300272.

References

[1] R. Alberton, R. Carli, A. Cenedese, and L. Schenato, "Multi-agent perimeter patrolling subject to mobility constraints," in *American Control Conference*, (Montreal), pp. 4498–4503, 2012.

[2] D. P. Bertsekas, *Network Optimization: Continuous and Discrete Models*. Athena Scientific, 1998.

[3] M. Ji, S. Azuma, and M. Egerstedt, "Role-assignment in multi-agent coordination," *International Journal of Assistive Robotics and Mechatronics*, vol. 7, no. 1, pp. 32–40, 2006.

[4] B. W. Carabelli, A. Benzing, F. Dürr, B. Koldehofe, K. Rothermel, G. Seyboth, R. Blind, M. Burger, and F. Allgower, "Exact convex formulations of network-oriented optimal operator placement," in *IEEE Conf. on Decision and Control*, (Maui), pp. 3777–3782, Dec. 2012.

[5] D. R. Kuehn and J. Porter, "The application of linear programming techniques in process control," *IEEE Transactions on Applications and Industry*, vol. 83, no. 75, pp. 423–427, 1964.

[6] J. Trdlicka, Z. Hanzalek, and M. Johansson, "Optimal flow routing in multi-hop sensor networks with real-time constraints through linear programming," in *IEEE Conf. on Emerging Tech. and Factory Auto.*, pp. 924–931, 2007.

[7] W. F. Sharpe, "A linear programming algorithm for mutual fund portfolio selection," *Management Science*, vol. 13, no. 7, pp. 499–510, 1967.

[8] G. B. Dantzig, *Linear Programming and Extensions*. Princeton, NJ: Princeton University Press, 1963.

[9] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*, vol. 6 of *Optimization and Neural Computation*. Belmont, MA: Athena Scientific, 1997.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[11] M. Burger, G. Notarstefano, F. Bullo, and F. Allgower, "A distributed simplex algorithm for degenerate linear programs and multi-agent assignment," *Automatica*, vol. 48, no. 9, pp. 2298–2304, 2012.

[12] G. Notarstefano and F. Bullo, "Distributed abstract optimization via constraints consensus: Theory and applications," *IEEE Transactions on Automatic Control*, vol. 56, no. 10, pp. 2247–2261, 2011.

[13] G. Yarmish and R. Slyke, "A distributed, scalable simplex method," *Journal of Supercomputing*, vol. 49, no. 3, pp. 373–381, 2009.

[14] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.

[15] M. Zhu and S. Martínez, "On distributed convex optimization under inequality and equality constraints," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 151–164, 2012.

[16] 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.

[17] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[18] W. Ren and R. W. Beard, *Distributed Consensus in Multi-Vehicle Cooperative Control*. Communications and Control Engineering, Springer, 2008.

[19] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*. Applied Mathematics Series, Princeton University Press, 2009. Electronically available at http://coordinationbook.info.

[20] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Applied Mathematics Series, Princeton University Press, 2010.

[21] R. Carli and G. Notarstefano, "Distributed partition-based optimization via dual decomposition," in *IEEE Conf. on Decision and Control*, (Firenze), pp. 2979–2984, Dec. 2013.

[22] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.

[23] I. Necoara and J. Suykens, "Application of a smoothing technique to decomposition in convex optimization," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2674 – 2679, 2008.

[24] N. Garg and J. Konemann, "Faster and simpler algorithms for multicommodity flow and other fractional packing problems," in *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, (Palo Alto, CA), pp. 300–309, 1998.

[25] V. Nedelcu, I. Necoara, and D. Quoc, "Computational complexity of inexact gradient augmented Lagrangian methods: Application to constrained MPC," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 3109–3134, 2014.

[26] I. Necoara and V. Nedelcu, "Rate analysis of inexact dual first order methods," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1232–1243, 2014.

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

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

[29] C. Cai, A. R. Teel, and R. Goebel, "Smooth Lyapunov functions for hybrid systems part II: (pre)asymptotically stable compact sets," *IEEE Transactions on Automatic Control*, vol. 53, no. 3, pp. 734–748, 2008.

[30] E. D. Sontag, "Further facts about input to state stabilization," *IEEE Transactions on Automatic Control*, vol. 35, pp. 473–476, 1989.

[31] D. Angeli, E. D. Sontag, and Y. Wang, "A characterization of integral input-to-state stability," *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1082–1097, 2000.

[32] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Review*, vol. 53, no. 3, pp. 464–501, 2011.

[33] F. H. Clarke, *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts, Wiley, 1983.

[34] 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.

[35] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 1st ed., 2003.

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

[37] J. P. Aubin and A. Cellina, *Differential Inclusions*, vol. 264 of *Grundlehren der mathematischen Wissenschaften*. New York: Springer, 1984.

[38] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill, 1953.

[39] H. J. Kushner and G. G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, vol. 35 of *Applications of Mathematics: Stochastic Modelling and Applied Probability*. New York: Springer, 2nd ed., 2003.

[40] R. J. B. Wets, "On the continuity of the value of a linear program and of related polyhedral-valued multifunctions," *Mathematical Programming Study*, vol. 24, pp. 14–29, 1985.

[41] E. D. Sontag, "Comments on integral variants of ISS," *Systems & Control Letters*, vol. 34, no. 1-2, pp. 93–100, 1998.

[42] G. B. Dantzig, *Linear Programming: 1: Introduction*. New York: Springer, 1997.

[43] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 2nd ed., 1999.

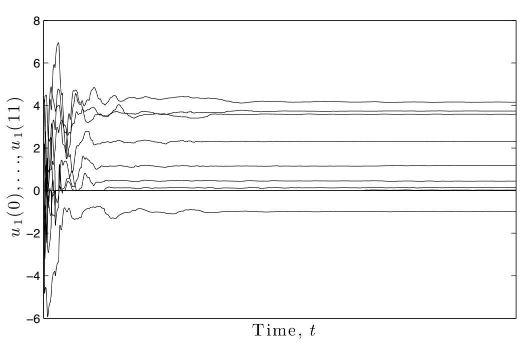


Fig. 4. The trajectories of the discontinuous saddle-point dynamics (15) under a recurrently connected communication graph where a random number of random links failed during periods of disconnection. The simulation parameters are the same as in Figure 3.

Appendix

The following is a technical result used in the proof of Theorem V.4.

Lemma A.1 (Property of feasible set). *If* $\{Ax = b, x \geq 0\}$ *is non-empty and bounded then* $\{A^T z \geq 0\}$ *is unbounded.*

Proof: We start by proving that there exists an $\nu \in \mathbb{R}^m$ such that $\{Ax = b + \nu, x \geq 0\}$ is empty. Define the vector $s \in \mathbb{R}^n$ component-wise as $s_i = \max_{\{Ax=b, x\geq 0\}} x_i$. Since $\{Ax = b, x \geq 0\}$ is compact and non-empty, s is finite. Next, fix $\varepsilon > 0$ and let $\nu = -A(s + \varepsilon \mathbb{1}_n)$. Note that $Ax = b + \nu$ corresponds to $A(x+s+\varepsilon\mathbb{1}_n) = b$, which is a shift by $s+\varepsilon\mathbb{1}_n$ in each component of x. By construction, $\{Ax = b + \nu, x \geq 0\}$ is empty. Then, the application of Farkas' Lemma [10, pp. 263] yields that there exists $\hat{z} \in \mathbb{R}^m$ such that $A^T \hat{z} \geq 0$ and $(b + \nu)^T \hat{z} < 0$ (in particular, $(b + \nu)^T \hat{z} < 0$ implies that $\hat{z} \neq 0$). For any $\lambda \in \mathbb{R}_{\geq 0}$, it holds that $A^T(\lambda \hat{z}) \geq 0$, and thus $\lambda \hat{z} \in \{A^T z \geq 0\}$, which implies the result. ■

The proof of Theorem V.9 makes use of the following result from [43, Proposition B.25].

Theorem A.2 (Danskin's Theorem). *Let* $Y \subset \mathbb{R}^m$ *be compact and convex. Given* $g : \mathbb{R}^n \times Y \to \mathbb{R}$*, suppose that* $x \mapsto g(x, y)$ *is differentiable for every* $y \in Y$*,* $\partial_x g$ *is continuous on* $\mathbb{R}^n \times Y$*, and* $y \mapsto g(x, y)$ *is strictly convex and continuous for every* $x \in \mathbb{R}^n$*. Define* $f : \mathbb{R}^n \to \mathbb{R}$ *by* $f(x) = \min_{y\in Y} g(x, y)$*. Then,* $\nabla f(x) = \partial_x g(x, y)|_{y=y_*(x)}$*, where* $y_*(x) = \operatorname{argmin}_{y\in Y} g(x, y)$*.*

Dean Richert received the B.Sc. and M.Sc. degree in electrical and computer engineering from the University of Calgary, Canada, in 2008 and 2010 respectively. In 2014, he received the Ph.D. degree in mechanical and aerospace engineering at the University of California, San Diego (UCSD). He was the recipient of a Jacob's Fellowship from UCSD as well as a Postgraduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. His research interests include cooperative control, distributed optimization, game theory, nonsmooth analysis, and set-valued dynamical systems. Currently, Dean works as a Sr. Algorithms Engineer at Cymer, Inc. (an ASML company).

Jorge Cortés received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, Urbana, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA, from 2004 to 2007. He is currently a Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. He is the author of Geometric, Control and Numerical Aspects of Nonholonomic Systems (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of Distributed Control of Robotic Networks (Princeton University Press, 2009). He is an IEEE Fellow and an IEEE Control Systems Society Distinguished Lecturer. His current research interests include distributed control, networked games, power networks, distributed optimization, spatial estimation, and geometric mechanics.