# Distributed Coordination for Separable Convex Optimization with Coupling Constraints

Simon K. Niederländer Jorge Cortés

*Abstract*— **This paper considers a network of agents described by an undirected graph that seek to solve a convex optimization problem with separable objective function and coupling equality and inequality constraints. Both the objective function and the inequality constraints are Lipschitz continuous. We assume that the constraints are compatible with the network topology in the sense that, if the state of an agent is involved in the evaluation of any given local constraint, this agent is able to fully evaluate it with the information provided by its neighbors. Building on the saddle-point dynamics of an augmented Lagrangian function, we develop provably correct distributed continuous-time coordination algorithms that allow each agent to find their component of the optimal solution vector along with the optimal Lagrange multipliers for the equality constraints in which the agent is involved. Our technical approach combines notions and tools from nonsmooth analysis, set-valued and projected dynamical systems, viability theory and convex programming.**

## I. Introduction

Distributed convex optimization problems appear in a wide range of disciplines, including network control systems, traffic flow optimization, sensor fusion, and power grid control. Recently, numerous algorithms that efficiently solve convex programs in a parallel or distributed fashion have been proposed. Decentralized optimization approaches often yield advantages over centralized solvers when the problem at hand requires inexpensive and low-performance computations, robustness against malfunctions, or ability to quickly react to changes. Motivated by network objectives and constraints that give rise to general nonsmooth convex programs with an inherent distributed structure, our objective in this paper is to synthesize continuous-time coordination algorithms that allow each agent to find their component of the optimal solution vector. In particular, we consider convex programs composed of an additively separable objective function and local coupling equality and inequality constraints with no differentiability assumptions imposed on their initial data. This setup substantially differs from consensus-based distributed optimization where agents agree on the entire optimal solution vector.

*Literature Review.* Our work is in context with the body of literature on convex optimization (as a general reference, see [1]). Motivated by large scale problems and systems with data naturally partitioned over a network [2], there has been vast interest on distributed convex optimization, including the works [3], [4], [5], [6], [7] and references therein. These works particularly build on consensus-based dynamics [8], [9], [10]. However, a direct transcription of those methods to the present setup would result in cooperative strategies where agents interact with their neighbors but operate over a vector that represents the entire network state, and are therefore not scalable. Considering convex optimization scenarios over networks that involve a separable objective function and local coupling constraints, various distributed methods based on saddle-point or primal-dual dynamics [11] have been recently proposed [12], [13], [14]. The work [12] studies partial primal-dual gradient algorithms for convex programs with equality constraints in the context of power networks. In contrast, the work [13] introduces primal-dual dynamics for convex programs that only possess inequality constraints. Although convergence in the primal variables is established, the dual variables converge to some unknown point which might not correspond to a dual optimal solution. Moreover, the inequality constraints are, in general, only feasible asymptotically. Our algorithm design builds on [14], which develops set-valued and discontinuous saddle-point(-like) dynamics specifically tailored for linear programs.

*Statement of Contributions.* We consider general nonsmooth convex optimization scenarios defined by separable convex objective functions with coupling equality (affine) and inequality (convex) constraints. Both the objective function and the inequality constraints are Lipschitz continuous. As a result of this generality in the problem statement, we face various technical challenges in both the design and convergence analysis of our distributed algorithms, particularly in what concerns the lack of differentiability of the problem data and the handling of the inequality constraints. Our technical approach relies on results from nonsmooth analysis as well as viability theory, and involves set-valued projection operations on dynamical systems. In contrast to existing works, we establish asymptotic stability of the set of primal-dual solutions of the convex program and, moreover, achieve point-wise convergence of solutions of the continuous-time algorithms developed. In addition, our proposed set-valued and discontinuous saddle-point-like dynamics guarantee feasibility with respect to the inequality constraints at any time. Simulations in a linear model predictive control example illustrate our results. The proofs are omitted for reasons of space and will appear elsewhere.

## II. Preliminaries

Let $\langle \cdot, \cdot \rangle$ denote the Euclidean inner product and let $\|\cdot\|_2$ and $\|\cdot\|_\infty$ denote the $\ell_2$- and $\ell_\infty$-norms in $\mathbb{R}^n$, respectively. Let $\mathbb{1}_n = (1, \ldots, 1) \in \mathbb{R}^n$. Given $x \in \mathbb{R}^n$, let $[x]^+ = (\max\{0, x_1\}, \ldots, \max\{0, x_n\}) \in \mathbb{R}^n_{\geq 0}$. Given a set $X \subset \mathbb{R}^n$, we denote its convex hull by $\overline{\text{co}}(X)$, its interior by

Simon Konrad Niederländer is with the Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany, Email: simon.niederlaender@ist.uni-stuttgart.de.

Jorge Cortés is with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, United States, Email: cortes@ucsd.edu.

$\text{int}(X)$, and its boundary by $\text{bd}(X)$. The closure of $X$ is denoted by $\text{cl}(X) = \text{int}(X) \cup \text{bd}(X)$. Let $\mathbb{B}(x,\delta) = \{y \in \mathbb{R}^n \mid \|y - x\|_2 < \delta\}$ denote the open ball.

A point $(x^*, \mu^*) \in X \times M$ is a *saddle point* of $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \supset X \times M \to \mathbb{R}$ if $\mathcal{L}(x^*, \mu) \leq \mathcal{L}(x^*, \mu^*) \leq \mathcal{L}(x, \mu^*)$ holds for all $(x, \mu) \in X \times M$. A *set-valued map* $F : \mathbb{R}^n \supset X \rightrightarrows \mathbb{R}^n$ maps elements of $X$ to subsets of $\mathbb{R}^n$. Let $\text{Ln} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued map that associates to each subset $F$ of $\mathbb{R}^n$ the set of least-norm elements of its closure $\text{cl}(F)$.

### A. Nonsmooth Analysis

We review here relevant basic notions from nonsmooth analysis following [15]. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *Lipschitz continuous at* $x \in \mathbb{R}^n$ if there exist $\delta_x \in \mathbb{R}_{>0}$ and $L_x \in \mathbb{R}_{\geq 0}$ such that $|f(y) - f(z)| \leq L_x \|y - z\|_2$ for all $y, z \in \mathbb{B}(x, \delta_x)$. If $f$ is Lipschitz continuous on $\mathbb{R}^n$, then it is said to be *Lipschitz continuous*. Note that a convex function is Lipschitz continuous (cf. Theorem 3.1.1, p. 16 in [16]).

Let $\Omega_f \subset \mathbb{R}^n$ be the set of points at which $f$ fails to be differentiable, and let $S \subset \mathbb{R}^n$ denote any other set of measure zero. The *generalized gradient* $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ of $f$ at $x \in \mathbb{R}^n$ is defined by

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

A set-valued map $F : \mathbb{R}^n \supset X \rightrightarrows \mathbb{R}^n$ is *upper semi-continuous* if for every $x \in X$ and $\varepsilon \in \mathbb{R}_{>0}$ there exists $\delta_x \in \mathbb{R}_{>0}$ such that $F(y) \subset F(x) + \mathbb{B}(0, \varepsilon)$ for all $y \in \mathbb{B}(x, \delta_x)$. $F$ is *locally bounded* if for every $x \in X$ there exist $\delta, \varepsilon \in \mathbb{R}_{>0}$ such that $\|z\|_2 \leq \varepsilon$ for all $z \in F(y)$ and all $y \in \mathbb{B}(x, \delta)$. Note that the set-valued map $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is locally bounded, upper semi-continuous, and takes nonempty, convex, and compact values (cf. Proposition 2.1.2 in [15]).

### B. Set-Valued and Projected Dynamical Systems

We present here basic notions on set-valued and projected dynamical systems as well as viability theory following [17], [18], [19]. Let $F : \mathbb{R}^n \supset X \rightrightarrows \mathbb{R}^n$ be a set-valued map. Consider the differential inclusion

$$\dot{x}(t) \in F(x(t)), \qquad x(t_0) = x_0 \in X. \tag{1}$$

A solution of (1) on an interval $[t_0, t_1] \subset \mathbb{R}$ is an absolutely continuous mapping $x : [t_0, t_1] \to X$ such that $\dot{x} \in F(x(t))$ for almost all (a.a.) $t \in [t_0, t_1]$. If the set-valued map $F$ is locally bounded, upper semi-continuous, and takes nonempty, convex, and compact values, then the existence of solutions of (1) starting from $x_0 \in X$ is guaranteed (cf. Theorem 3, p. 98 in [18]).

Given a Lipschitz continuous function $f : \mathbb{R}^n \supset X \to \mathbb{R}$, the *set-valued Lie derivative* $\mathcal{L}_F f : \mathbb{R}^n \supset X \rightrightarrows \mathbb{R}$ of $f$ at $x \in X$ with respect to (1) is defined by

$$(\mathcal{L}_F f)(x) = \big\{\psi \in \mathbb{R} \mid \exists \xi \in F(x) : \langle \xi, \pi \rangle = \psi, \forall \pi \in \partial f(x)\big\}.$$

Let $X \subset \mathbb{R}^n$ be a nonempty, closed, and convex set; called the *viability set*. Let the *distance function* $d_X : \mathbb{R}^n \to \mathbb{R}$ be defined by $d_X(x) = \inf_{y\in X} \|x - y\|_2$. The *tangent cone* and the *normal cone* of $X$ at $x \in X$ are, respectively,

$$T_X(x) = \text{cl} \bigcup_{\delta>0} \frac{1}{\delta}(X - x), \quad N_X(x) = \text{cl} \bigcup_{\lambda\geq 0} \lambda \partial d_X(x).$$

Note that if $x \in \text{int}(X)$, then $T_X(x) = \mathbb{R}^n$ and $N_X(x) = \{0\}$. A solution $x : [t_0, t_1] \to X$ of (1) starting from $x_0 \in X$ is *viable in* $X$ under $F$ if $x(t) \in X$ for all $t \in [t_0, t_1]$. Let $\text{proj}_X(v) = \text{argmin}_{y\in X} \|v - y\|_2$. The *orthogonal (set) projection* $\Pi_{T_X} : \mathbb{R}^n \rightrightarrows T_X$ of $F$ onto $T_X$ at $x \in X$ is

$$\Pi_{T_X}(x, F(x)) = \bigcup_{\xi\in F(x)} \lim_{\delta\searrow 0} \frac{\text{proj}_X(x + \delta\xi) - x}{\delta}.$$

Note that if $x \in \text{int}(X)$, then $\Pi_{T_X}(x, F(x))$ reduces to the set $F(x)$. Consider the differential inclusion

$$\dot{x}(t) \in \Pi_{T_X}(x, F(x))(t), \qquad x(t_0) = x_0 \in X. \tag{2}$$

In general, the projection operator $\Pi_{T_X}$ possesses no continuity properties and the values of $\Pi_{T_X}(x, F(x))$ are not necessarily convex [18]. Yet, the following theorem states conditions under which viable solutions of (2) exist [18].

**Theorem 2.1** (Existence of viable solutions)**.** *Let $X \subset \mathbb{R}^n$ be nonempty, closed and convex, and let the set-valued map $F : \mathbb{R}^n \supset X \rightrightarrows \mathbb{R}^n$ be locally bounded, upper semi-continuous with nonempty, convex and compact values. Then, for any $x_0 \in X$, there exists a viable solution $x : [t_0, t_1] \to X$ of*

$$\dot{x} \in F(x(t)) - N_X(x(t)), \qquad x(t_0) = x_0 \in X. \tag{3}$$

*Moreover, the set of viable solutions of* (2) *and* (3) *coincide.*

## III. Problem Statement

Consider a network of $n \in \mathbb{Z}_{>0}$ agents whose communication topology is represented by an undirected and connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, n\}$ is the set of vertices and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges. The state of agent $i \in \{1, \ldots, n\}$ is $x_i \in \mathbb{R}$ and its set of neighbors is $\mathcal{N}(i) = \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}\}$. The objective of the agents is to cooperatively solve the optimization problem

$$\inf_{x\in\mathbb{R}^n} \left\{ \sum_{i=1}^{n} f_i(x_i) \;\middle|\; h(x) = 0,\ g(x) \preceq 0 \right\}, \tag{4}$$

where $f_i : \mathbb{R} \to \mathbb{R}$ is a convex and Lipschitz continuous cost function associated with agent $i \in \{1, \ldots, n\}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ is convex and Lipschitz continuous, and $h : \mathbb{R}^n \to \mathbb{R}^p$ is affine, i.e., $h(x) = Ax - b$, with $A \in \mathbb{R}^{p\times n}$ and $b \in \mathbb{R}^p$, where $p \leq n$. The equality and inequality functions $h, g$ are understood component-wise, i.e., $h_\ell(x) = 0$ for all $\ell \in \{1, \ldots, p\}$, and $g_k(x) \leq 0$ for all $k \in \{1, \ldots, m\}$. The network state is denoted by $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and the aggregate objective function is $f(x) = \sum_{i=1}^n f_i(x_i)$. Let the feasibility set and the solution set of (4) be denoted by $C = \{x \in \mathbb{R}^n \mid h(x) = 0,\ g(x) \preceq 0\}$ and $X = \{x^* \in C \mid f(x^*) \leq f(x),\ \forall x \in C\}$. Let the viability set associated with (4) be defined by $G = \{x \in \mathbb{R}^n \mid g(x) \preceq 0\}$.

Throughout the paper, we assume that the convex optimization problem (4) is feasible, possesses finite primal-dual optimal values, and satisfies the Slater constraint qualification certificate [1]. Additionally, we assume that the constraints of (4) are compatible with the network topology described by $\mathcal{G}$. Formally, the constraint $g_k(x) \leq 0$ is compatible with $\mathcal{G}$ if $g_k$ can be expressed as a function of some components of the network state $x \in \mathbb{R}^{|\mathcal{V}|}$, say $\tilde{x} \in \mathbb{R}^{|\mathcal{U}_k|}$, where $\mathcal{U}_k \subset \mathcal{V}$ induces a complete undirected subgraph of $\mathcal{G}$. A similar definition can be stated for $h_\ell$.

## IV. Lagrangian Saddle-Point Characterization

In this section, we characterize primal-dual solutions of (4) in terms of saddle points of an augmented Lagrangian function. The standard Lagrangian function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}$ associated with (4) is $\mathcal{L}(x, \mu, \nu) = f(x) + \langle \mu, Ax - b\rangle + \langle \nu, g(x)\rangle$, where $\mu = (\mu_1, \dots, \mu_p) \in \mathbb{R}^p$ and $\nu = (\nu_1, \dots, \nu_m) \in \mathbb{R}^m_{\geq 0}$ are Lagrange multiplier. The Lagrangian dual problem of (4) takes the form

$$\sup_{(\mu,\nu)\in\mathbb{R}^p\times\mathbb{R}^m_{\geq 0}} \left\{ \phi(\mu,\nu) \;\middle|\; -A^\top \mu \in \partial f(x) + \sum_{k\in K(x)} \nu_k \partial g_k(x) \right\} \quad (5)$$

where the Lagrangian dual function $\phi : \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R} \cup \{-\infty, +\infty\}$ is defined by $\phi(\mu, \nu) = \inf_{x\in\mathbb{R}^n} \mathcal{L}(x, \mu, \nu)$, with $K(x) = \{k \in \{1, \dots, m\} \mid g_k(x) = 0\}$. Let the set of solutions of (5) be denoted by $M \times N \subset \mathbb{R}^p \times \mathbb{R}^m_{\geq 0}$. The following proposition establishes a relationship between the primal-dual solutions of (4), respectively (5), and the saddle points of an augmented Lagrangian function.

**Proposition 4.1** (Saddle-point characterization)**.** *Given $\kappa \in \mathbb{R}_{\geq 0}$. Let $\mathcal{L}^\kappa : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ be defined by*

$$\begin{aligned} \mathcal{L}^\kappa(x, \mu) = \; & f(x) + \frac{1}{2}\|Ax - b\|_2^2 + \langle \mu, Ax - b\rangle \\ & + \kappa \langle \mathbb{1}_m, [g(x)]^+ \rangle. \end{aligned} \quad (6)$$

*Then, the function $\mathcal{L}^\kappa$ is convex-concave, and*

- *(i) if $x^* \in X$ and $(\mu^*, \nu^*) \in M \times N$ are primal-dual solutions of* (4)*, respectively* (5)*, then $(x^*, \mu^*) \in X \times M$ is a saddle point of $\mathcal{L}^\kappa$ for any $\kappa \geq \|\nu^*\|_\infty$,*
- *(ii) if $(\tilde{x}, \tilde{\mu}) \in \mathbb{R}^n \times \mathbb{R}^p$ is a saddle point of $\mathcal{L}^\kappa$ with $\kappa > \|\nu^*\|_\infty$ for some $\nu^* \in N$, then $\tilde{x} \in \mathbb{R}^n$ is a primal solution of* (4)*.*

Proposition 4.1 motivates the search for saddle points of $\mathcal{L}^\kappa$ rather than directly solving the convex optimization problem (4). Since the augmented Lagrangian function $\mathcal{L}^\kappa$ is convex-concave, a natural approach to find the saddle points is via its associated saddle-point dynamics.

*Remark* 4.2 (Bound on $\kappa$)*.* Note that the lower bound on $\kappa$ in Proposition 4.1 is characterized by the dual solution $\nu^* \in N$ of (4) which is unknown a priori. However, our forthcoming discussion proposes fully distributed dynamics that do not incorporate any knowledge on $\kappa$. •

## V. Continuous-Time Distributed Optimization

This section presents a saddle-point algorithm that asymptotically converges to the set of primal-dual solutions of (4), respectively (5), given knowledge of a suitable lower bound on $\kappa$ (cf. Remark 4.2). Moreover, we propose a discontinuous saddle-point-like algorithm that enjoys the same convergence properties but does not rely on the global parameter $\kappa$, and is therefore amenable for distributed implementation.

### *A. Saddle-Point Dynamics*

Since the augmented Lagrangian function $\mathcal{L}^\kappa : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ defined in (6) is Lipschitz continuous, the saddle-point dynamics $(\dot{x}, \dot{\mu}) \in (-\partial_x \mathcal{L}^\kappa(x, \mu), +\partial_\mu \mathcal{L}^\kappa(x, \mu))$ (gradient descent in the primal variable $x$ and gradient ascent in the dual variable $\mu$) defined over $\mathbb{R}^n \times \mathbb{R}^p$ take the form

$$\begin{aligned} \dot{x}(t) + A^\top(Ax(t) - b + \mu(t)) \in & -\partial f(x(t)) \\ & - \kappa \sum_{k\in K(x)} \partial [g_k(x(t))]^+, \end{aligned} \quad (7a)$$

$$\dot{\mu}(t) = Ax(t) - b, \quad (7b)$$

for a.a. $t \in [t_0, +\infty)$ with initial condition $(x_0, \mu_0) \in \mathbb{R}^n \times \mathbb{R}^p$. Note that the existence of solutions $(x, \mu) : [t_0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^p$ of (7) is guaranteed by the properties of the generalized gradients $\partial_x \mathcal{L}^\kappa$ and $\partial_\mu \mathcal{L}^\kappa$. We understand the solutions of (7) in the sense of Krasovskii (see [20] for a definition). For notational convenience, we use the set-valued map $F^\flat : \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^p$ to refer to (7). The following theorem characterizes asymptotic convergence of the saddle-point dynamics (7) to a point in $X \times M$.

**Theorem 5.1** (Point-wise convergence)**.** *Let $\kappa \in \mathbb{R}_{\geq 0}$, $x^* \in X$, and $(\mu^*, \nu^*) \in M \times N$. Define $\mathcal{V} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}_{\geq 0}$ by*

$$\mathcal{V}(x, \mu) = \frac{1}{2}\|x - x^*\|_2^2 + \frac{1}{2}\|\mu - \mu^*\|_2^2. \quad (8)$$

*If $\kappa > \|\nu^*\|_\infty$, then $(\mathcal{L}_{F^\flat}\mathcal{V})(x, \mu) \subset (-\infty, 0]$ for all $(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^p$ and any solution $(x, \mu) : [t_0, +\infty) \to \mathbb{R}^n \times \mathbb{R}^p$ of* (7) *converges asymptotically to a point in the set of primal-dual solutions $X \times M$.*

### *B. Projected Saddle-Point-Like Dynamics*

Here, we propose discontinuous saddle-point-like dynamics that do not rely on a priori knowledge of the dual solution $\nu^* \in N$, as in Theorem 5.1, but also converge asymptotically to a point in the set of primal-dual solutions $X \times M$. Recall that $G \subset \mathbb{R}^n$ denotes the viability set associated with (4). Let the set-valued flow $F : G \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n$ be defined by

$$F(x, \mu) = -A^\top(Ax - b + \mu) - \partial f(x). \quad (9)$$

The definition is motivated by the fact that, for $(x, \mu) \in \mathrm{int}(G) \times \mathbb{R}^p$, we have $\partial_x \mathcal{L}^\kappa = -F(x, \mu)$. Consider the saddle-point-like dynamics defined over $G \times \mathbb{R}^p$,

$$\dot{x}(t) \in \begin{cases} F(x(t), \mu(t)), & \text{if } x \in \mathrm{int}(G), \\ \Pi_{T_G}(x, F(x, \mu))(t), & \text{if } x \in \mathrm{bd}(G), \end{cases} \quad (10a)$$

$$\dot{\mu}(t) = Ax(t) - b, \quad (10b)$$

for a.a. $t \in [t_0, +\infty)$ with initial condition $(x_0, \mu_0) \in G \times \mathbb{R}^p$. Note that the set-valued map $\Pi_{T_G}$ possesses no continuity properties and the values of $\Pi_{T_G}(x, F(x, \mu))$ are not necessarily convex. However, the existence of viable solutions $(x, \mu) : [t_0, +\infty) \to G \times \mathbb{R}^p$ of (10) is guaranteed by Theorem 2.1. For notational convenience, we use the set-valued map $F^\sharp : G \times \mathbb{R}^p \rightrightarrows G \times \mathbb{R}^p$ to refer to the discontinuous saddle-point-like dynamics (10).

Our objective here is to provide an alternative expression of the projection operator $\Pi_{T_G}$ that more clearly displays its amenability to distributed implementation. Let the set of outward normals to $G$ at $x \in \mathrm{bd}(G)$ be defined by

$$N_G^\sharp(x) = \left\{ \pi_g \in \sum_{k\in K(x)} \partial g_k(x) \;\middle|\; \|\pi_g\|_2 = 1 \right\} \subset N_G(x).$$

Following the idea of singleton-valued projected dynamics [19], the set-valued projection $\Pi_{T_G} : \mathbb{R}^n \rightrightarrows T_G$ of $F$ onto $T_G$ at $(x,\mu) \in G \times \mathbb{R}^p$ in (10a) can be expressed as

$$\Pi_{T_G}(x, F(x,\mu)) = \bigcup_{\xi \in F(x,\mu)} \xi - \max\left\{0, \langle \xi, \pi_g^*(\xi)\rangle\right\} \pi_g^*(\xi),$$

with $\pi_g^*(\xi)$ as the unique maximizer determined by the (sub-) optimization problem

$$\sup_{\pi_g \in \mathbb{R}^n} \left\{ \langle \xi, \pi_g \rangle \mid \pi_g \in N_G^\sharp(x) \right\}. \tag{11}$$

Note that if $(x,\mu) \in \operatorname{int}(G) \times \mathbb{R}^p$, then $\langle \xi, \pi_g^*(\xi)\rangle \leq 0$ and thus, $\Pi_{T_G}(x, F(x,\mu)) = F(x,\mu)$. The following lemma states existence and uniqueness of $\pi_g^*$.

**Lemma 5.2** (Existence and uniqueness)**.** *Given a point $(x,\mu) \in \operatorname{bd}(G) \times \mathbb{R}^p$. If there exists $\xi \in F(x,\mu)$ such that*

$$\sup_{\pi_g \in \mathbb{R}^n} \left\{ \langle \xi, \pi_g \rangle \mid \pi_g \in N_G^\sharp(x) \right\} > 0,$$

*then the maximizer $\pi_g^*(\xi)$ of (11) exists and is unique.*

*Remark* 5.3 (Relationship to slow solutions). A *slow solution* $x : [t_0, +\infty) \to G$ of (7a) is a solution satisfying

$$\dot{x}(t) \in \operatorname{Ln}\left(\Pi_{T_G}(x, F(x,\mu))\right)(t), \qquad x(t_0) = x_0 \in G,$$

for a.a. $t \in [t_0, +\infty)$. If the projection operator $\Pi_{T_G}$ satisfies the conditions in Theorem 2.1, and if, in addition, $\Pi_{T_G}$ is continuous (i.e., upper and lower semi-continuous), then the existence of slow solutions of (10) is guaranteed [18]. Therefore, given $(x,\mu) \in \operatorname{bd}(G) \times \mathbb{R}^p$, it suffices to study the singleton-valued map $\operatorname{Ln}\left(\Pi_{T_G}(x, F(x,\mu))\right)$. •

*Remark* 5.4 (Projection operator for continuously differentiable optimization). If $f, g \in \mathcal{C}^1$, then the projection operator in (10a) reduces to the singleton-valued map

$$\Pi_{T_G}(x, F(x,\mu)) = F(x,\mu) - \max\left\{0, \langle F(x,\mu), \pi_g^*(x)\rangle\right\} \pi_g^*(x),$$

since $F(x,\mu) = -A^\top(Ax - b + \mu) - \nabla f(x)$, and $\pi_g^*(x) = \sum_{k \in K(x)} \nu_k^* \nabla g_k(x) \in N_G^\sharp(x)$, with $\nu^*$ as the unique maximizer determined by the (sub-)optimization problem

$$\begin{aligned} \sup_{\nu \in \mathbb{R}^m} \Bigg\{ \Big\langle F(x,\mu), \sum_{k \in K(x)} \nu_k \nabla g_k(x) \Big\rangle \;\Big|\; \nu \succeq 0, \\ \Big\| \sum_{k \in K(x)} \nu_k \nabla g_k(x) \Big\|_2 = 1 \Bigg\}. \end{aligned} \tag{12}$$

Since $F$ is singleton-valued, it follows $\Pi_{T_G}(x, F(x,\mu)) = \operatorname{Ln}\left(\Pi_{T_G}(x, F(x,\mu))\right)$, and therefore, any solution $(x,\mu) : [t_0, +\infty) \to G \times \mathbb{R}^p$ of (10) is also a slow solution. •

We now establish convergence properties of (10) by means of a relationship between solutions of the saddle-point dynamics $F^\flat$ and the saddle-point-like dynamics $F^\sharp$.

**Lemma 5.5** (Relationship of solutions)**.** *For any $(x,\mu) \in G \times \mathbb{R}^p$, let $\kappa \in \mathbb{R}_{\geq 0}$ satisfy the inequality*

$$\kappa \geq \sup_{\xi \in F(x,\mu)} \max\left\{0, \langle \xi, \pi_g^*(\xi)\rangle\right\}, \tag{13}$$

*where $\pi_g^*(\xi) = \operatorname{argmax}_{\pi_g \in N_G^\sharp(x)} \langle \xi, \pi_g \rangle$. Then, the inclusion $F^\sharp(x,\mu) \subset F^\flat(x,\mu)$ holds for all $(x,\mu) \in G \times \mathbb{R}^p$.*

Note that the inclusion in Lemma 5.5 may be strict and, in general, the set of solutions of (7) is richer than the set of solutions of (10). Building on the previous lemma, our next contribution characterizes point-wise convergence of solutions of (10) to the set of primal-dual solutions $X \times M$.

**Theorem 5.6** (Point-wise convergence)**.** *Any viable solution $(x,\mu) : [t_0, +\infty) \to G \times \mathbb{R}^p$ of (10) starting from $(x_0, \mu_0) \in G \times \mathbb{R}^p$ converges asymptotically to a point in $X \times M$.*

### *C. Distributed Implementation*

In what follows, we show that the proposed saddle-point-like algorithm (10) is well-suited for distributed implementation. Consider the network model introduced in Section III. If $(x,\mu) \in \operatorname{int}(G) \times \mathbb{R}^p$, then each agent $i \in \{1, \ldots, n\}$ implements its dynamics defined by the flow (9), i.e.,

$$\dot{x}_i + \sum_{\{\ell : a_{\ell i} \neq 0\}} a_{\ell i} \Bigg( \sum_{\{j : a_{\ell j} \neq 0\}} a_{\ell j} x_j - b_\ell + \mu_\ell \Bigg) \in -\partial f_i(x_i),$$

and some dual dynamics defined by (10b), i.e.,

$$\dot{\mu}_\ell = \sum_{\{i : a_{\ell i} \neq 0\}} a_{\ell i} x_i - b_\ell,$$

where $\ell \in \{1, \ldots, m\}$. Hence, in order for agent $i$ to be able to implement its corresponding dynamics (10a), it also needs access to certain dual components $\mu_\ell$ for which $a_{\ell i} \neq 0$. However, if $(x,\mu) \in \operatorname{bd}(G) \times \mathbb{R}^p$, then each agent $i \in \{1, \ldots, n\}$ implements its dynamics defined by (10a), i.e.,

$$\dot{x}_i \in \bigcup_{\xi_i \in F_i(x,\mu)} \xi_i - \max\Bigg\{0, \sum_{\{j : \pi_{g_j}^* \neq 0\}} \xi_j \pi_{g_j}^* \Bigg\} \pi_{g_i}^*,$$

where $F_i(x,\mu)$ denotes the set of $i^{\text{th}}$ components of $F$, and $\pi_{g_i}^*$ is the $i^{\text{th}}$ component of the maximizer of (11). Hence, if the state of an agent is involved in the active inequality constraints, it needs to solve (11) and communicate the solution $\pi_g^*$ to its neighbors. Note that (11) requires only local information. We say the dynamics (10) are distributed over $\mathcal{G}$ when the following conditions are satisfied:

(C1) for each $i \in \{1, \ldots, n\}$, agent $i$ knows its own state $x_i \in \mathbb{R}$ and its local cost function $f_i$,

(C2) agent $i$ has access to its neighbors decision variables $x_j \in \mathbb{R}$, their local cost functions $f_j$, and
  - (i) the non-zero elements of every row of $A \in \mathbb{R}^{p \times n}$ for which $a_{\ell i} \neq 0$, and
  - (ii) every $b_\ell \in \mathbb{R}$ for which $a_{\ell i} \neq 0$, and
  - (iii) has knowledge of the active inequality functions $g_k$ in which it is involved.

## VI. Simulations

Consider a linear model predictive control setup for a network of agents with coupling constraints, whose aim is to compute a control input sequence that simultaneously minimizes the actuation effort and the network state. Formally,

$$\begin{aligned} \underset{\{u(k|t)\}_k}{\text{minimize}} \quad & \frac{1}{2} \sum_{k=t}^{t+N-1} \|x(k+1|t)\|_Q^2 + \|u(k|t)\|_R^2 \\ \text{subject to} \quad & x(k+1|t) = Sx(k|t) + Pu(k|t), x(t|t) = x(t), \\ & Cx(k|t) + Du(k|t) \leq \mathbb{1}_m, \forall k \in [t, t+N-1], \end{aligned}$$

where $x(k|t) \in \mathbb{R}^n$ and $u(k|t) \in \mathbb{R}^n$ are the network state and control input, predicted at time $t \in \mathbb{R}$ over the horizon $N$. Let $\|x\|_Q^2 = \langle x, Qx\rangle$, and let $Q, R \in \mathbb{R}^{n\times n}$ be such that $Q = Q^\top \succeq 0$ and $R = R^\top \succ 0$. The initial condition $x_i(t|t)$ is known to agent $i \in \{1, \dots, n\}$ and its neighbors. Note that the network topology is encoded in the sparsity structure of the matrices $S, P \in \mathbb{R}^{n\times n}$ and $C, D \in \mathbb{R}^{m\times n}$.

Each agent can be interpreted as a subsystem whose dynamics are influenced by the predicted states of neighboring agents. Here, every agent knows the dynamics of its own subsystem and those of its neighbors, but not the entire network dynamics. We consider only discrete-time systems $(S, P)$ that are controllable. The prediction horizon $N$ is chosen such that stability of the overall model predictive control setup is guaranteed. At every iteration step, a solution to the problem is a sequence of predicted open-loop optimal controls $u^*(\cdot|t) = \{u^*(k|t), \dots, u^*(t+N-1|t)\}$, where only $u^*(t) = u^*(t|t)$ is implemented to the network system.

To express the model predictive control problem as a convex optimization problem of the form (4), we introduce the vector of variables $x(\cdot|t) = (x(k+1|t), \dots, x(t+N|t), u(k|t), \dots, u(t+N-1|t)) \in \mathbb{R}^{2Nn}$. The equality constraints of (4) are determined by the matrix $A \in \mathbb{R}^{Nn\times 2Nn}$,

$$A = \left(\begin{array}{cccc|cccc} I_n & 0_n & \cdots & 0_n & -P & 0_n & \cdots & 0_n \\ -S & I_n & \cdots & 0_n & 0_n & -P & \cdots & 0_n \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0_n & 0_n & \cdots & I_n & 0_n & 0_n & \cdots & -P \end{array}\right),$$

and the vector $b = (Sx(t|t), 0, \dots, 0) \in \mathbb{R}^{Nn}$, where $I_n, 0_n \in \mathbb{R}^{n\times n}$ denote the identity and zero matrices, respectively. Similarly, the inequality constraint function $g : \mathbb{R}^{2Nn} \to \mathbb{R}^{Nm}$ is characterized by

$$g(x) = \left(\begin{array}{cccc|cccc} C & 0_m & \cdots & 0_m & D & 0_m & \cdots & 0_m \\ 0_m & C & \cdots & 0_m & 0_m & D & \cdots & 0_m \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0_m & 0_m & \cdots & C & 0_m & 0_m & \cdots & D \end{array}\right) x - \begin{pmatrix} \mathbb{1}_m \\ \mathbb{1}_m \\ \vdots \\ \mathbb{1}_m \end{pmatrix},$$

where $0_m \in \mathbb{R}^{m\times n}$. Note that the objective function in the model predictive control setup is separable only for diagonal weighting matrices $Q$ and $R$. Let $Q = R = I_n$. Hence, the problem can be equivalently stated as

$$\begin{aligned} \underset{x\in\mathbb{R}^{2Nn}}{\text{minimize}} \quad & f(x) = \frac{1}{2} \sum_{k=t}^{t+N-1} \sum_{i=1}^{n} x_i(k+1|t)^2 + u_i(k|t)^2 \\ \text{subject to} \quad & Ax = b, \quad g(x) \preceq 0. \end{aligned}$$

Consider the network dynamics and constraints as given in Figure 1(a) with a prediction horizon $N = 4$. The total number of primal-dual variables of algorithm (10) is $2Nn + N = 68$. In this setup, each agent $i$ is responsible for its own $2N = 8$ variables in $x$ and $\mu$, and also for the variables from its neighbors with respect to the network topology depicted in Figure 1(b). Note that agent 5 and its neighbors incorporate additional $N = 4$ variables related to the control input sequence over the prediction horizon.

When implementing the algorithm (10), we use a first-order Euler approximation with stepsize 0.008. For every iteration step, we use $\|Ax - b\|_2 \leq \varepsilon$ with $\varepsilon = 0.001$ as stopping criteria for the dynamics and perform a warm-start of the algorithm with the steady-state variables obtained from the previous iteration step. At every iteration step, the implementation of (10) requires agent 5 to compute its optimal control sequence over the prediction horizon, i.e., $\{u_5(k|t), \dots, u_5(t+N-1|t)\}$. In addition, all agents $i \in \{1, \dots, 8\}$ compute their predicted open-loop state evolutions over the horizon, i.e., $\{x_i(k+1|t), \dots, x_i(t+N|t)\}$.

Figure 2 shows the result of the iterative implementation of (10). Once the optimal control and network states over the prediction horizon are computed (corresponding to the steady-state values in Figures 2(d)-(e) for every iteration), they are implemented by their respective agents. This results in the predicted open-loop and closed-loop network evolution depicted in Figure 2(a) and the associated open-loop and closed-loop optimal control input as shown in Figure 2(b). Note that the projection operator used prevents the trajectories of (10) from violating the viability condition at any time. The equality constraint violation is plotted in Figure 2(c) and the LaSalle function is illustrated in 2(f). Clearly, the algorithm (10) converges asymptotically in every iteration step and thus, agent 5 steers the network state to zero under minimal actuation effort.

## VII. Conclusions

We have developed continuous-time coordination algorithms for networks of agents that seek to collectively solve a class of constrained convex optimization problems with an inherent distributed structure. Based on an augmented Lagrangian function, we have proposed saddle-point(-like) dynamics that converge point-wise to the set of solutions of the convex program. The dynamics are amenable for distributed implementation in the sense that each individual agent asymptotically finds its component of the optimal solution vector by sharing its state with its neighbors. Future work will seek to characterize the convergence rate of the proposed dynamics and its robustness properties against disturbances and noise.

## Acknowledgments

Research supported by the German National Academic Foundation, the Dr. Jürgen & Irmgard Ulderup Foundation, and NSF award CMMI-1300272.

## References


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

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

[3] 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.

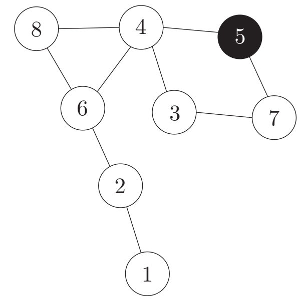
[4] M. Zhu and S. Martínez, "On distributed convex optimization under inequality and equality constraints," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 151–164, 2012.

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

[6] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson, "Subgradient methods and consensus algorithms for solving convex optimization problems," in *IEEE Conf. on Decision and Control*, Cancun, Mexico, 2008, pp. 4185–4190.

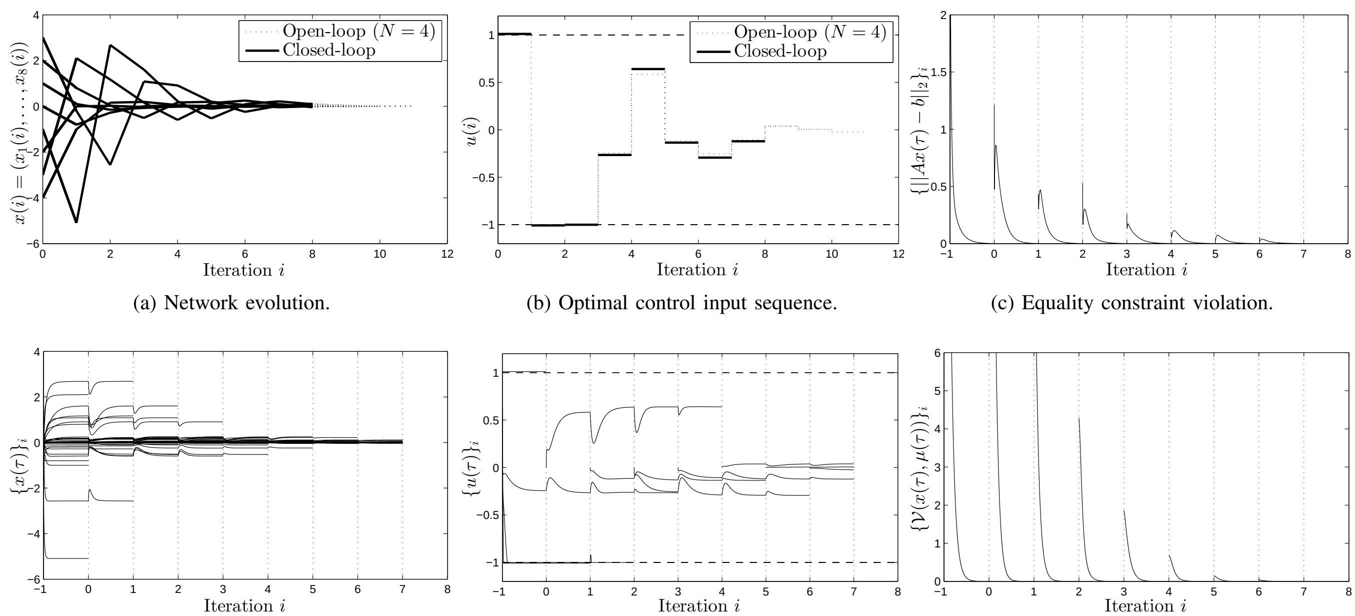
$$S = \begin{pmatrix} \frac{1}{10} & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{10} & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 \\ 0 & 0 & \frac{1}{10} & 0 & 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & \frac{6}{5} & \frac{1}{10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & \frac{1}{10} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{10} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{10} & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 & \frac{1}{5} & 0 & \frac{1}{10} \end{pmatrix}, \ P = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ C = 0_{2\times 8}, \ D = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

(a) Network and constraint matrices.




(b) Network topology.

Fig. 1. Network and constraint matrices (a) and associated graph topology (b) of the multi-agent system over which the algorithm (10) is fully distributed. Note that the tuple $(S, P)$ is controllable. The network of agents is controlled through the dynamics of agent 5 ((b) black node). Also, note that the choice of $D$ implies that the box constraints $\|(u(k|t), \dots, (t+N-1|t))\|_\infty \leq 1$ are satisfied.




(a) Network evolution. (b) Optimal control input sequence. (c) Equality constraint violation.

(d) Iterative computation of network states. (e) Iterative computation of optimal control. (f) LaSalle function.

Fig. 2. The predicted open-loop and closed-loop network evolution of the linear model predictive control setup is depicted in plot (a) over 8 iteration steps. Plot (b) shows the associated open-loop and closed-loop optimal control sequence of agent 5 with a prediction horizon of $N = 4$. At every iteration step, the network state evolution (d) and optimal control input sequence (e) are computed by algorithm (10). The steady-state values achieved by these trajectories at every iteration step correspond to the optimal network state evolution and optimal control input over the entire prediction horizon. Plot (c) shows the equality constraint violation and (f) depicts the evolution of the LaSalle function (8) at every iteration step. For illustrative purposes, the trajectories are shifted backwards by one iteration step $(i-1)$ to emphasize that the steady-state values are applied to the control system at iteration step $i$.


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

[8] 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.

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

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

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

[12] N. Li, L. Chen, C. Zhao, and S. H. Low, "Connecting automatic generation control and economic dispatch from an optimization view," in *American Control Conference*, Portland, OR, Jun. 2014, pp. 735–740.

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

[14] D. Richert and J. Cortés, "Robust distributed linear programming," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, 2015, to appear.

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

[16] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I*, ser. Grundlehren Text Editions. New York: Springer, 1993.

[17] 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.

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

[19] A. Nagurney and D. Zhang, *Projected Dynamical Systems and Variational Inequalities with Applications*, ser. International Series in Operations Research and Management Science. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996, vol. 2.

[20] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid Dynamical Systems: Modeling, Stability, and Robustness*. Princeton University Press, 2012.