Distributed convergence to Nash equilibria by adversarial networks with undirected topologies

Bahman Gharesifard Jorge Cortés

Abstract— **This paper considers a class of strategic scenarios in which two undirected networks of agents have opposing objectives with regards to the optimization of a common objective function. In the resulting zero-sum game, individual agents collaborate with neighbors in their respective network and have only partial knowledge of the state of the agents in the other one. We synthesize a distributed saddle-point algorithm that is implementable via local interactions and establish its convergence to the set of Nash equilibria for a class of strictly concave-convex and locally Lipschitz objective functions. Our algorithm synthesis builds on a continuous-time optimization strategy for finding the set of minimizers of a sum of convex functions in a distributed way. As a byproduct, we show that this strategy can be itself cast as a saddle-point dynamics and use this fact to establish its asymptotic convergence properties. The technical approach combines tools from algebraic graph theory, nonsmooth analysis, set-valued dynamical systems, and game theory.**

I. INTRODUCTION

This paper considers a class of strategic scenarios in which two undirected networks of agents are involved in a zero-sum game. We assume that the objective function can be decomposed as a sum of concave-convex functions and networks have opposing objectives regarding its optimization. Agents collaborate with neighbors in their own network and have partial information about the state of the agents in the other one. Our aim is to design a distributed coordination algorithm that can be used by the networks to converge to the set of Nash equilibria. Potential applications include collective bargaining, competitive social networks, and collaborative pursuit-evasion.

Literature review: The present work has connections with the literature on distributed optimization and zero-sum games. The distributed optimization of a sum of convex functions has been intensively studied in recent years, see e.g. [1], [2], [3], [4]. These works build on consensus-based dynamics [5], [6], [7], [8] to find the solutions of the optimization problem in a variety of scenarios and are designed in discrete time. An exception is the recent work [9] on continuous-time distributed optimization. Regarding zero-sum games, the works [10], [11], [12] study the convergence of discrete-time subgradient dynamics to a saddle point. Continuous-time best-response dynamics for zero-sum games converges to the set of Nash equilibria for both convex-concave [13] and quasiconvex-quasiconcave [14] functions. Under strict convexity-concavity assumptions, continuous-time subgradient flow dynamics converges to a saddle point [15], [10]. Asymptotic convergence is also guaranteed when the Hessian of the objective function is positive definite in one argument and the function is linear in the other [10], [16].

Statement of contributions: Our contributions are twofold. First, we show that the solutions to the continuous-time distributed optimization problem for a sum of locally Lipschitz (i.e., not necessarily differentiable) convex functions correspond to saddle points of a certain objective function. This function is convex in one of its arguments and linear in the other. We also show that the dynamics proposed in [9] to solve the optimization problem exactly corresponds to the saddle point dynamics of the objective function when the network is undirected. This allows us to establish its asymptotic convergence properties using a set-valued version of the LaSalle Invariance Principle. Second, we introduce the problem of distributed convergence to Nash equilibria for two networks engaged in a strategic scenario. The networks' objectives are to either maximize or minimize a common objective function which can be written as a sum of concave-convex functions. Individual agents collaborate with neighbors in their respective network and have partial knowledge of the state of the agents in the other network. Building on our analysis of the distributed optimization problem, we synthesize a consensus-based saddle-point strategy for the adversarial network, which we term the distributed Nash seeking dynamics. We show that, for a class of strictly concave-convex and locally Lipschitz objective functions, the proposed dynamics is guaranteed to converge to the Nash equilibria. For reasons of space, some proofs are omitted and will appear elsewhere.

II. PRELIMINARIES

Let $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{R}_{\geq 0}$, $\mathbb{Z}_{\geq 1}$ denote the set of real, integer, nonnegative real, and positive integer numbers, respectively. We let $\mathsf{I}_{d\times d}$ denote the identity matrix in $\mathbb{R}^{d\times d}$, $d \in \mathbb{Z}_{\geq 1}$ and use $\mathbf{1}_d$ to denote the vector $(1,\dots,1) \in \mathbb{R}^d$. We denote by $||\cdot||$ the Euclidean norm on $\mathbb{R}^d$. We denote the set of subsets of $\mathbb{R}^d$ by $\mathfrak{B}(\mathbb{R}^d)$. If $v = (v_1,\dots,v_{d_1})^T \in \mathbb{R}^{d_1}$ and $w = (w_1,\dots,w_{d_2})^T \in \mathbb{R}^{d_2}$, We denote by $v \otimes w = (v_1 w,\dots,v_{d_1} w)^T \in \mathbb{R}^{d_1\times d_2}$ the Kronecker product of v and w. We call $\imath : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$, $d_1, d_2 \in \mathbb{Z}_{\geq 1}$, $d_2 \geq d_1$, a canonical inclusion if it consists of mapping each vector in $\mathbb{R}^{d_1}$ to a vector in $\mathbb{R}^{d_2}$ by adding zeros to the rest of its components. A function $f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$, is *concave-convex* if it is concave in its first argument and convex in

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

the second one [17]. Let $f : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$, where $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$ and $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$ are closed and convex, be a concave-convex function. Then a point $(x_1^*, x_2^*) \in \mathsf{X}_1 \times \mathsf{X}_2$ is called a *saddle point* of f if for all $x_1 \in \mathsf{X}_1$ and $x_2 \in \mathsf{X}_2$ we have

$$f(x_1, x_2^*) \leq f(x_1^*, x_2^*) \leq f(x_1^*, x_2).$$

A. Nonsmooth analysis

We briefly recall some notions of nonsmooth analysis [18]. A function $f : \mathbb{R}^d \to \mathbb{R}$ is *locally Lipschitz* at $x \in \mathbb{R}^d$ if there exists a neighborhood $\mathcal{U}$ of x and $C_x, \in \mathbb{R}_{\geq 0}$ such that $|f(y) - f(z)| \leq C_x ||y - z||$, for all $y, z \in \mathcal{U}$. By Rademacher's Theorem [18], locally Lipschitz functions are differentiable almost everywhere. Let us denote by Ω_f the set of points that f fails to be differentiable. The *generalized gradient* of f is then defined by

$$\partial f(x) = \mathrm{co}\Big\{ \lim_{k\to\infty} df(x_k) \mid x_k \to x, x_k \notin \Omega_f \cup S \Big\},$$

where S is any set of measure zero. A point $x \in \mathbb{R}^d$ such that $0 \in \partial f(x)$ is called a *critical point* of f. A function $f : \mathbb{R}^d \to \mathbb{R}$ is called *regular* at $x \in \mathbb{R}$ if for all $v \in \mathbb{R}^d$ the right directional derivative of f, in the direction of v, exists at x and coincides with the generalized directional derivative of f at x in the direction of v. We refer the interested reader to [18] for definitions of these notions. A locally Lipschitz function at x which is convex is always regular [18, Proposition 2.3.6]. If f is a convex function which is Lipschitz in a neighborhood $\mathcal{U}$ of $x \in \mathbb{R}^d$, then for all $\xi \in \partial f(x)$ and $x' \in \mathcal{U}$

$$f(x') - f(x) \geq \xi \cdot (x' - x). \tag{1}$$

We call this property the first order condition of convexity.

Proposition 2.1: **(Properties of generalized gradient):** Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz function at $x \in \mathbb{R}^d$. Then

(i) the map $\partial f : \mathbb{R}^d \to \mathfrak{B}(\mathbb{R}^d)$ is upper semicontinuous and locally bounded at $x \in \mathbb{R}^d$.
(ii) $\partial f(x)$ is nonempty, compact, and convex.

Proposition 2.2: **(Generalized gradient of finite sum of locally Lipschitz functions):** Suppose $\{f^i\}_{i=1}^n$, $n \in \mathbb{Z}_{\geq 1}$, is a collection of functions which are Lipschitz in a neighborhood of $x \in \mathbb{R}^d$. Then $\partial(\sum_i f^i)(x) \subseteq \sum_i \partial f^i(x)$. The equality holds when f^i is regular, for all $i \in \{1, \dots, n\}$.

B. Set-valued dynamical systems

We recall some background on continuous-time set-valued dynamical systems from [19], [20]. Let $x : \mathbb{R}_{\geq 0} \to \mathsf{X}$ be a curve on $\mathsf{X} \subset \mathbb{R}^d$, $d \in \mathbb{Z}_{\geq 1}$. A time-invariant set-valued dynamical systems is a differential inclusion

$$\dot{x}(t) \in \Psi(x(t)) \ a.e., \tag{2}$$

where $t \in \mathbb{R}_{\geq 0}$ and $\Psi : \mathbb{R}^d \to \mathfrak{B}(\mathbb{R}^d)$ is a set-valued map. If $0 \in \Psi(x)$, we call x an *equilibrium* of (2). A solution to this dynamical systems is any absolutely continuous curve $x : \mathbb{R}_{\geq 0} \to \mathsf{X}$ which satisfies (2).

Proposition 2.3: **(Existence of solutions for differential inclusions):** Let Ψ be upper semicontinuous with nonempty, compact, and convex values. Then for any initial condition, locally, there exists, a solution to (2).

The LaSalle Invariance Principle for set-valued continuous-time systems is helpful to establish the asymptotic stability properties of systems of the form (2). A set $W \subset \mathsf{X}$ is *weakly positively invariant* with respect to Ψ if for any $x \in W$, there exists $\tilde{x} \in \mathsf{X}$ such that $\tilde{x} \in \Psi(x)$. The set W is *strongly positively invariant* with respect to Ψ if $\Psi(x) \subset W$, for all $x \in W$. The *set-valued Lie derivative* of a differentiable function $V : \mathbb{R}^d \to \mathbb{R}$ with respect to Ψ at $x \in \mathbb{R}^d$ is defined by $\bar{\mathcal{L}}_\Psi V(x) = \{v \cdot \nabla V(x) \mid v \in \Psi(x)\}$.

Theorem 2.4: **(Set-valued LaSalle Invariance Principle):** Let $W \subset \mathsf{X}$ be a strongly positively invariant under (2) and $V : \mathsf{X} \to \mathbb{R}$ a continuously differentiable function. Suppose the evolutions of (2) are bounded and $\max \bar{\mathcal{L}}_\Psi V(x) \leq 0$ or $\bar{\mathcal{L}}_\Psi V(x) = \emptyset$, for all $x \in W$. If $S_{\Psi,V} = \{x \in \mathsf{X} \mid 0 \in \bar{\mathcal{L}}_\Psi V(x)\}$, then any solution $x(t)$, $t \in \mathbb{R}_{\geq 0}$, starting in W converges to the largest weakly positively invariant set M contained in $\bar{S}_{\Psi,V} \cap W$. When M is a finite collection of points, then the limit of each solution equals one of them.

C. Graph theory

A *directed graph*, or simply *digraph*, is a pair $\mathcal{G} = (V, E)$, where V is a finite set called the vertex set and $E \subseteq V \times V$ is the edge set. When E is unordered, we call $\mathcal{G}$ an *undirected graph* or simply a *graph*. We say that an edge $(u, v) \in E$ is *incident away from* u (or an *out-edge* of u) and *incident toward* v (or an *in-edge* of v), and we call u an *in-neighbor* of v and v an *out-neighbor* of u. We denote the set of in-neighbors and out-neighbors of v, respectively, with $\mathcal{N}_{\mathcal{G}}^{\text{in}}(v)$ and $\mathcal{N}_{\mathcal{G}}^{\text{out}}(v)$. For a graph, these two sets are equal and we call members of this set, denoted by $\mathcal{N}_{\mathcal{G}}(v)$, neighbors of v. A *bipartite digraph* is a digraph whose vertices can be divided into two disjoint sets V_1 and V_2 such that every edge can be written as (v_1, v_2) or (v_2, v_1), where $v_1 \in V_1$ and $v_2 \in V_2$. Here, we focus our attention only on the algebraic properties of graphs. A graph is called *connected* if there exists a path between any two vertices. A *weighted graph* is a triplet $\mathcal{G} = (V, E, \mathsf{A})$, where (V, E) is a graph and $\mathsf{A} \in \mathbb{R}_{\geq 0}^{n \times n}$ is the *adjacency matrix* of $\mathcal{G}$. The adjacency matrix has the property that $a_{ij} > 0$ if $(v_i, v_j) \in E$ and $a_{ij} = 0$, otherwise. The weighted degree v_i, $i \in \{1, \dots, n\}$ is $d^{\text{w}}(v_i) = \sum_{j=1}^n a_{ij}$. The *weighted degree matrix* D is the diagonal matrix defined by $(\mathsf{D})_{ii} = d^{\text{w}}(i)$, for all $i \in \{1, \dots, n\}$. The *Laplacian* is $\mathsf{L} = \mathsf{D} - \mathsf{A}$. For an undirected graph, $\mathsf{L}\mathbf{1}_n = \mathbf{1}_n^T \mathsf{L} = 0$ and $\mathsf{L} = \mathsf{L}^T$ and is positive semidefinite [21]. When $\mathcal{G}$ is connected, the zero eigenvalue is simple.

D. Zero-sum games

We recall some game theoretic notions from [22]. An n-player game is a triplet $\mathbf{G} = (P, \mathsf{X}, U)$, where P is the set of players, $n = |P| \in \mathbb{Z}_{\geq 2}$, $\mathsf{X} = \mathsf{X}_1 \times \ldots \times \mathsf{X}_n$, $\mathsf{X}_i \subset \mathbb{R}^{d_i}$ is the set of (pure) strategies of player $v_i \in P$, $d_i \in \mathbb{Z}_{\geq 1}$, and $U = (u_1, \dots, u_n)$, where $u_i : \mathsf{X} \to \mathbb{R}$ is the payoff function of player v_i, $i \in \{1, \dots, n\}$. The game $\mathbf{G}$ is called a *zero-sum game* if $\sum_{i=1}^n u_i(x) = 0$, for all $x \in$

X. If $x_i \in \mathsf{X}_i$, we denote by x_{-i} the strategy set of all players except v_i. An outcome $x^* \in \mathsf{X}$ is called a (pure) *Nash equilibrium* of $\mathbf{G}$ if for all $i \in \{1, \dots, n\}$ and all $x_i \in \mathsf{X}_i$ we have $u_i(x_i^*, x_{-i}^*) \geq u_i(x_i, x_{-i}^*)$. One can extend this notion to *mixed* Nash equilibria by assigning probabilities to pure strategies [22]. In this paper, we focus on a particular class of two-players zero-sum games which have at least one pure Nash equilibrium. The following well-known result, characterizes this class of games.

Theorem 2.5: **(Minmax theorem):** Let $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$ and $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$, $d_1, d_2 \in \mathbb{Z}_{\geq 1}$, be nonempty, closed, bounded, and convex. If $u : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$ is continuous and the sets $\{x' \in \mathsf{X}_1 \mid u(x', y) \geq \alpha\}$, and $\{x' \in \mathsf{X}_2 \mid u(x, y') \leq \alpha\}$ are convex for all $x \in \mathsf{X}_1$, $y \in \mathsf{X}_2$, and $\alpha \in \mathbb{R}$, then $\max_x \min_y u(x, y) = \min_y \max_x u(x, y)$.

III. Problem statement

Consider two networks Σ_1 and Σ_2 composed of agents $\{v_1, \dots, v_{n_1}\}$ and agents $\{w_1, \dots, w_{n_2}\}$, respectively. Throughout this paper, Σ_1 and Σ_2 are undirected graphs. The state of Σ_1, denoted x_1, belongs to $\mathsf{X}_1 \subset \mathbb{R}^{d_1}$, $d_1 \in \mathbb{Z}_{\geq 1}$. Likewise, the state of Σ_2, denoted x_2, belongs to $\mathsf{X}_2 \subset \mathbb{R}^{d_2}$, $d_2 \in \mathbb{Z}_{\geq 1}$. We assume that X_1 and X_2 are compact and convex. Here we do not get into the details of what these states represent (as a particular case, the network state could correspond to the collection of the states of agents in it). In addition, each agent v_i in Σ_1 has an estimate $x_1^i \in \mathbb{R}^{d_1}$ of what the network state is, which may differ from the actual value x_1. Similarly, each agent w_j in Σ_2 has an estimate $x_2^j \in \mathbb{R}^{d_2}$ of what the network state is. For convenience, we let $\boldsymbol{x}_1 = (x_1^1, \dots, x_1^{n_1})^T$ and $\boldsymbol{x}_2 = (x_2^1, \dots, x_2^{n_2})^T$ denote the vector of agent estimates about the state of the respective networks. Within each network, neighboring agents can share their estimates. Networks can also obtain information about each other. This is modeled by means of a bipartite directed graph Σ_{eng}, called *engagement* topology, with disjoint vertex sets $\{v_1, \dots, v_{n_1}\}$ and $\{w_1, \dots, w_{n_2}\}$. According to this model, an agent in Σ_1 obtains information from its in-neighbors in Σ_{eng} about their estimates of the state of Σ_2, and vice versa. Figure 1 illustrates this concept.

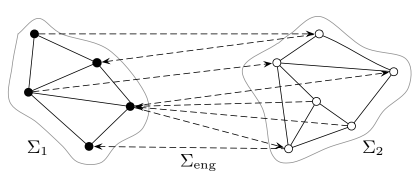

Fig. 1. Two networks Σ_1 and Σ_2 engaged in a strategic scenario. The edges of the engagement topology Σ_{eng} are dashed.

For each $i \in \{1, \dots, n_1\}$, let $f_1^i : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$ be a differentiable concave-convex function only available to agent $v_i \in \Sigma_1$. Similarly, let $f_2^j : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$ be a differentiable concave-convex function only available to agent $w_j \in \Sigma_2$, $j \in \{1, \dots, n_2\}$. The networks Σ_1 and Σ_2 are engaged in a zero-sum game with payoff function $U : \mathsf{X}_1 \times \mathsf{X}_2 \to \mathbb{R}$

$$U(x_1, x_2) = \sum_{i=1}^{n_1} f_1^i(x_1, x_2) = \sum_{j=1}^{n_2} f_2^j(x_1, x_2), \qquad (3)$$

where Σ_1 wishes to maximize U, while Σ_2 wishes to minimize it. The objective of the networks is therefore to settle upon a Nash equilibrium, i.e., to solve the following maxmin problem

$$\max_{x_1 \in \mathsf{X}_1} \min_{x_2 \in \mathsf{X}_2} U(x_1, x_2). \qquad (4)$$

We refer to the this zero-sum game as the 2-*network zero-sum game* and denote it by $\mathbf{G}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$. Interestingly, several problems in sensor networks such as estimation, localization, or routing [23], [24] can be cast into this framework when intelligent adversaries are present.

IV. Continuous-time distributed optimization on undirected networks

Here, we review the continuous-time solution to the optimization problem proposed in [9] for undirected graphs. Consider a network composed by $n \in \mathbb{Z}_{\geq 1}$ agents $v_1, \dots, v_n$ whose communication topology is described by a connected graph $\mathcal{G}$. For each, $i \in \{1, \dots, n\}$, let $f^i : \mathbb{R}^d \to \mathbb{R}$ be locally Lipschitz and convex, and only available to agent v_i. The network objective is to solve the following optimization problem in a distributed way,

$$\text{minimize } f(x) = \sum_{i=1}^{n} f^i(x). \qquad (5)$$

Let $x^i \in \mathbb{R}^d$ denote the estimate of agent v_i about the value of the solution to (5) and define $\boldsymbol{x}^T = ((x^1)^T, \dots, (x^n)^T) \in \mathbb{R}^{nd}$. Next, we provide an alternative formulation of (5).

Lemma 4.1: Let $\mathsf{L} \in \mathbb{R}^{n \times n}$ be the Laplacian of $\mathcal{G}$ and define $\mathbf{L} = \mathsf{L} \otimes \mathsf{I}_d \in \mathbb{R}^{nd \times nd}$. The problem (5) on $\mathbb{R}^d$ is equivalent to the following problem on $\mathbb{R}^{nd}$,

$$\text{minimize } \tilde{f}(\boldsymbol{x}) = \sum_{i=1}^{n} f^i(x^i), \quad \text{subject to } \mathbf{L}\boldsymbol{x} = \mathbf{0}_{nd}. \qquad (6)$$

Proof: The proof follows by noting that (i) $\tilde{f}(\mathbf{1}_n \otimes x) = f(x)$ for all $x \in \mathbb{R}^d$ and (ii) since $\mathcal{G}$ is connected, $\mathbf{L}\boldsymbol{x} = \mathbf{0}_{nd}$ if and only if $\boldsymbol{x} = \mathbf{1}_n \otimes x$, for some $x \in \mathbb{R}^d$. ■

The formulation (6) is appealing because it brings together the estimates of each agent about the value of the solution to the original optimization problem. It is worth mentioning that $\tilde{f}$ is Lipschitz and convex. Moreover, from Proposition 2.2, the elements of its generalized gradient are of the form $\tilde{g}_{\boldsymbol{x}} = (g_{x^1}^1, \dots, g_{x^n}^n) \in \boldsymbol{\partial}\tilde{f}(\boldsymbol{x})$, where $g_{x^i}^i \in \partial f^i(x^i)$, for $i \in \{1, \dots, n\}$. Since $\tilde{f}$ is convex and the constraints in (6) are linear, the constrained optimization problem is feasible [25].

Proposition 4.2: **(Solutions of the distributed optimization problem as saddle points):** Let $\mathcal{G}$ be connected, and define $F : \mathbb{R}^{nd} \times \mathbb{R}^{nd} \to \mathbb{R}$ by

$$F(\boldsymbol{x}, \boldsymbol{z}) = \tilde{f}(\boldsymbol{x}) + \boldsymbol{z}^T \mathbf{L}\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^T \mathbf{L}\boldsymbol{x}. \qquad (7)$$

Then F is locally Lipschitz and convex in its first argument and linear in its second, and

(i) if $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ is a saddle point of F, then so is $(\boldsymbol{x}^*, \boldsymbol{z}^* + \mathbf{1}_n \otimes a)$, for any $a \in \mathbb{R}^d$.
(ii) if $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ is a saddle point of F, then $\boldsymbol{x}^*$ is a solution of (6).
(iii) if $\boldsymbol{x}^*$ is a solution of (6), there exists $\boldsymbol{z}^*$ with $\mathbf{L}\boldsymbol{z}^* \in -\boldsymbol{\partial} \tilde{f}(\boldsymbol{x}^*)$ such that $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ is a saddle point of F.

Proof: First, note that for $\mathcal{G}$ undirected, F is convex in its first argument and linear in the second. The statement (i) is immediate. To show (ii), using that $\mathcal{G}$ is connected, one can see that the saddle points of F are of the form $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ with $\boldsymbol{x}^* = \mathbf{1}_n \otimes x^*$, $x^* \in \mathbb{R}^d$, and $\mathbf{L}\boldsymbol{z}^* \in -\partial \tilde{f}(\boldsymbol{x}^*)$. The last inclusion implies that there exist $g^i_{x^*} \in \partial f^i(x^*)$, $i \in \{1, \dots, n\}$, such that $\mathbf{L}\boldsymbol{z}^* = -(g^1_{x^*}, \dots, g^n_{x^*})^T$. Noting that

$$(\mathbf{1}_n^T \otimes \mathsf{I}_d)\mathbf{L} = (\mathbf{1}_n^T \otimes \mathsf{I}_d)(\mathsf{L} \otimes \mathsf{I}_d) = \mathbf{1}_n^T \mathsf{L} \otimes \mathsf{I}_d = \mathbf{0}_{d \times dn},$$

we deduce $\mathbf{0}_d = (\mathbf{1}_n^T \otimes \mathsf{I}_d)\mathbf{L}\boldsymbol{z}^* = -\sum_{i=1}^n g^i_{x^*}$. As a result, using Proposition 2.2, $\boldsymbol{x}^*$ is a solution of (6). Finally, (iii) follows by noting $\boldsymbol{x}^* = \mathbf{1}_n \otimes x^*$ and the fact that $0 \in \partial f(x^*)$ implies that there exists $\boldsymbol{z}^* \in \mathbb{R}^{nd}$ with $\mathbf{L}\boldsymbol{z}^* \in -\boldsymbol{\partial} \tilde{f}(\boldsymbol{x}^*)$, yielding that $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ is a saddle point of F. ■

Since $\mathcal{G}$ is undirected, the gradient of F in (7) is distributed over $\mathcal{G}$. Given Proposition 4.2, it is natural to consider the saddle-point dynamics of F to solve (5),

$$\dot{\boldsymbol{x}} + \mathbf{L}\boldsymbol{x} + \mathbf{L}\boldsymbol{z} \in -\partial \tilde{f}(\boldsymbol{x}), \tag{8a}$$
$$\dot{\boldsymbol{z}} = \mathbf{L}\boldsymbol{x}. \tag{8b}$$

Note that (8) is a set-valued dynamical system. Using Propositions 2.1 and 2.3, one can guarantee the existence of solutions. From Proposition 4.2, if $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ is an equilibrium of (8), then $\boldsymbol{x}^*$ is a solution to (6). The next result states that the dynamics (8) leads the network to agree on a global minimum of f when $\mathcal{G}$ is undirected and f is the sum of locally Lipschitz convex functions.

Theorem 4.3: **(Asymptotic convergence of** (8) **on undirected networks):** Let $\mathcal{G}$ be a connected graph and consider the distributed optimization problem (5), where each f^i, $i \in \{1, \dots, n\}$ is locally Lipschitz and convex. Then, the projection onto the first component of any trajectory of (8) asymptotically converges to the set of solutions to (6). Moreover, if f has a finite number of critical points, the limit of the projection onto the first component of each trajectory is a solution of (6).

V. Nash seeking strategies on undirected adversarial networks

Consider the network $\Sigma_{\text{net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}})$ described in Section III. We propose a distributed dynamics which allows for agents in Σ_{net} to compute a Nash equilibrium of the zero-sum game $\mathbf{G}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$, when U is a strictly concave-convex Lipschitz continuous function.

A. Reformulation of the 2-network zero-sum game

We start by describing how agents in each network use the information obtained from their neighbors to compute the value of their own objective function. Based on these estimates, we introduce a reformulation of the $\mathbf{G}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$ which is instrumental for establishing our results. Each agent in Σ_1 has a function $\tilde{f}^i_1 : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2 n_2} \to \mathbb{R}$ with the following properties

(i) $\tilde{f}^i_1(x_1, \mathbf{1}_{n_2} \otimes x_2) = f^i_1(x_1, x_2)$, for $x_1 \in \mathbb{R}^{d_1}$, $x_2 \in \mathbb{R}^{d_2}$, and
(ii) for each $\boldsymbol{x}_2 \in \mathbb{R}^{d_2 n_2}$ and any canonical inclusion map $\imath_1 : \mathbb{R}^{d_2 |\mathcal{N}^{\text{in}}_{\Sigma_{\text{eng}}}(v_i)|} \to \mathbb{R}^{d_2 n_2}$,
$$\tilde{f}^i_1(x^i_1, \imath_1 \circ h^i_1(\boldsymbol{x}_2)) = \tilde{f}^i_1(x^i_1, \boldsymbol{x}_2), \tag{9}$$
where $h^i_1 : \mathbb{R}^{d_2 n_2} \to \mathbb{R}^{d_2 |\mathcal{N}^{\text{in}}_{\Sigma_{\text{eng}}}(v_i)|}$ projects $\boldsymbol{x}_2$ to the values received by v_i from its in-neighbors in Σ_{eng}.

Each agent w_j in Σ_2 has a function $\tilde{f}^j_2$ defined similarly. The collective payoff functions of the two networks are

$$\tilde{U}_1(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{i=1}^{n_1} \tilde{f}^i_1(x^i_1, \imath_1 \circ h^i_1(\boldsymbol{x}_2)), \tag{10a}$$
$$\tilde{U}_2(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{j=1}^{n_2} \tilde{f}^j_2(\imath_2 \circ h^j_2(\boldsymbol{x}_1), x^j_2). \tag{10b}$$

In general, the functions $\tilde{U}_1$ and $\tilde{U}_2$ need not be the same. However, note that $\tilde{U}_1(\mathbf{1}_{n_1} \otimes x_1, \mathbf{1}_{n_1} \otimes x_2) = \tilde{U}_2(\mathbf{1}_{n_1} \otimes x_1, \mathbf{1}_{n_1} \otimes x_2)$, for any $x_1 \in \mathbb{R}^{d_1}$, $x_2 \in \mathbb{R}^{d_2}$. Here, we assume that the individual payoff functions are assigned such that $\tilde{U}_1 = \tilde{U}_2$, and denote this common function by $\tilde{U}$. The merit of this assumption, as we will see in the next result, is that it allows us to still cast the problem as a (constrained) zero-sum game. The proof follows from an argument similar to the one in Lemma 4.1.

Lemma 5.1 (Reformulation of the 2-network zero-sum game): The problem (4) on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ is equivalent to the following problem on $\mathbb{R}^{n_1 d_1} \times \mathbb{R}^{n_2 d_2}$,

$$\max_{\boldsymbol{x}_1 \in \mathsf{X}_1^{n_1}} \min_{\boldsymbol{x}_2 \in \mathsf{X}_2^{n_2}} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \qquad \text{subject to} \tag{11a}$$
$$\mathbf{L}_1 \boldsymbol{x}_1 = \mathbf{0}_{n_1 d_1}, \quad \mathbf{L}_2 \boldsymbol{x}_2 = \mathbf{0}_{n_2 d_2}. \tag{11b}$$

We denote the constrained zero-sum game defined in (11) by $\tilde{\mathbf{G}}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$. Our objective is to design a coordination algorithm that is implementable with the local information available and leads them to find a Nash equilibrium of $\tilde{\mathbf{G}}_{\text{adv-net}}$ which corresponds to a Nash equilibrium of $\mathbf{G}_{\text{adv-net}}$. Achieving this goal, however, is nontrivial because individual agents, not networks themselves, are the decision makers. From the point of view of agents in each network, the objective is to agree on the states of both their own network and the other network, and that the resulting states correspond to a Nash equilibrium of $\mathbf{G}_{\text{adv-net}}$.

Proposition 5.2: **(Characterization of the Nash equilibria of $\tilde{\mathbf{G}}_{\text{adv-net}}$):** Let F_1 and F_2 be defined by

$$F_1(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2) = -\tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) + \boldsymbol{z}_1^T \mathbf{L}_1 \boldsymbol{x}_1 + \frac{1}{2} \boldsymbol{x}_1^T \mathbf{L}_1 \boldsymbol{x}_1,$$
$$F_2(\boldsymbol{x}_2, \boldsymbol{z}_2, \boldsymbol{x}_1) = \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2) + \boldsymbol{z}_2^T \mathbf{L}_2 \boldsymbol{x}_2 + \frac{1}{2} \boldsymbol{x}_2^T \mathbf{L}_2 \boldsymbol{x}_2,$$

where $\mathbf{L}_i = \mathsf{L}_i \otimes \mathsf{I}_{d_i}$, for $i \in \{1,2\}$, with L_i the Laplacian of Σ_i. Then if $(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*)$ and $(\boldsymbol{x}_2^*, \boldsymbol{z}_2^*, \boldsymbol{x}_1^*)$ are saddle points of, respectively, F_1 and F_2

(i) so are $(\boldsymbol{x}_1^*, \boldsymbol{z}_1^* + \mathbf{1}_{n_1} \otimes a_1, \boldsymbol{x}_2^*)$ and $(\boldsymbol{x}_2^*, \boldsymbol{z}_2^* + \mathbf{1}_{n_2} \otimes a_2, \boldsymbol{x}_1^*)$, $a_1 \in \mathbb{R}^{d_1}$ and $a_2 \in \mathbb{R}^{d_2}$,

(ii) $(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*)$ is a Nash equilibrium of $\tilde{\mathbf{G}}_{\text{adv-net}}$,

Furthermore,

(iii) if $(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*)$ is a Nash equilibria of $\tilde{\mathbf{G}}_{\text{adv-net}}$ then there exists $\boldsymbol{z}_1^*, \boldsymbol{z}_2^*$ such that $(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*)$ and $(\boldsymbol{x}_2^*, \boldsymbol{z}_2^*, \boldsymbol{x}_1^*)$ are saddle points of, respectively, F_1 and F_2.

B. The distributed Nash seeking dynamics

Here, we introduce a dynamics to solve (11). Specifically, we design gradient dynamics to find the saddle points of F_1 and F_2 prescribed by Proposition 5.2. Consider the set-valued dynamical system $\Psi_{\text{Nash-undir}} : (\mathbb{R}^{d_1 n_1} \times \mathbb{R}^{d_2 n_2})^2 \rightrightarrows (\mathbb{R}^{d_1 n_1} \times \mathbb{R}^{d_2 n_2})^2$ given by

$$\dot{\boldsymbol{x}}_1 + \mathbf{L}_1 \boldsymbol{x}_1 + \mathbf{L}_1 \boldsymbol{z}_1 \in \boldsymbol{\partial}_{\boldsymbol{x}_1} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \tag{12a}$$

$$\dot{\boldsymbol{z}}_1 = \mathbf{L}_1 \boldsymbol{x}_1, \tag{12b}$$

$$\dot{\boldsymbol{x}}_2 + \mathbf{L}_2 \boldsymbol{x}_2 + \mathbf{L}_2 \boldsymbol{z}_2 \in -\boldsymbol{\partial}_{\boldsymbol{x}_2} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2), \tag{12c}$$

$$\dot{\boldsymbol{z}}_2 = \mathbf{L}_2 \boldsymbol{x}_2, \tag{12d}$$

where $\boldsymbol{x}_j, \boldsymbol{z}_j \in \mathbb{R}^{n_j d_j}$, $j \in \{1,2\}$. We refer to $\Psi_{\text{Nash-undir}}$ as the *undirected distributed Nash seeking dynamics*. Note that local solutions to this dynamics exist by virtue of Proposition 2.1 and 2.3. The next result captures the main contribution of this section.

Theorem 5.3: **(Asymptotic convergence of the undirected distributed Nash seeking dynamics):** Consider the zero-sum game $\mathbf{G}_{\text{adv-net}} = (\Sigma_1, \Sigma_2, \Sigma_{\text{eng}}, U)$, where

(i) Σ_1 and Σ_2 are connected and undirected,

(ii) $\tilde{U} : \mathsf{X}_1^{n_1} \times \mathsf{X}_2^{n_2} \to \mathbb{R}$, X_1 and X_2 compact convex subsets of, respectively, $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, is a Lipschitz continuous strictly concave-convex function, distributed over $(\Sigma_1, \Sigma_2, \Sigma_{\text{eng}})$ in the sense of (10).

Then the projection onto the first and third components of the solutions to (12) asymptotically converges to the solution of (11).

Proof: Throughout this proof, since (9) holds for both networks, without loss of generality and for simplicity, we assume that agents in Σ_1 have access to $\boldsymbol{x}_2$ and, similarly, agents in Σ_2 have access to $\boldsymbol{x}_1$. By Theorem 2.5, a solution to (11) exists. By the strict concavity-convexity properties, this solution is, in fact, unique. Let us denote this solution by $\boldsymbol{x}_1^* = \mathbf{1}_{n_1} \otimes x_1^*$ and $\boldsymbol{x}_2^* = \mathbf{1}_{n_2} \otimes x_2^*$. By Proposition 5.2(iii), there exists $\boldsymbol{z}_1^*$ and $\boldsymbol{z}_2^*$ such that $(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*, \boldsymbol{z}_2^*) \in \mathrm{Eq}(\Psi_{\text{Nash-undir}})$. First, note that given any initial condition $(\boldsymbol{x}_1^0, \boldsymbol{z}_1^0, \boldsymbol{x}_2^0, \boldsymbol{z}_2^0) \in \mathbb{R}^{2n_1 d_1} \times \mathbb{R}^{2n_2 d_2}$, the set $W_{\boldsymbol{z}_1^0, \boldsymbol{z}_2^0} = \{(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \mid (\mathbf{1}_{n_j}^T \otimes \mathsf{I}_{d_j}) \boldsymbol{z}_j = (\mathbf{1}_{n_j}^T \otimes \mathsf{I}_{d_j}) \boldsymbol{z}_j^0, \;\; j \in \{1,2\}\}$ is strongly positively invariant under (12). Consider the function $V : (\mathbb{R}^{d_1 n_1})^2 \times (\mathbb{R}^{d_2 n_2})^2 \to \mathbb{R}_{\geq 0}$ defined by

$$\begin{aligned} V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) = {} & \\ & \frac{1}{2}(\boldsymbol{x}_1 - \boldsymbol{x}_1^*)^T(\boldsymbol{x}_1 - \boldsymbol{x}_1^*) + \frac{1}{2}(\boldsymbol{z}_1 - \boldsymbol{z}_1^*)^T(\boldsymbol{z}_1 - \boldsymbol{z}_1^*) \\ & + \frac{1}{2}(\boldsymbol{x}_2 - \boldsymbol{x}_2^*)^T(\boldsymbol{x}_2 - \boldsymbol{x}_2^*) + \frac{1}{2}(\boldsymbol{z}_2 - \boldsymbol{z}_2^*)^T(\boldsymbol{z}_2 - \boldsymbol{z}_2^*). \end{aligned}$$

The function V is smooth. Next, we examine its set-valued Lie derivative along $\Psi_{\text{Nash-undir}}$. Let $\xi \in \bar{\mathcal{L}}_{\Psi_{\text{Nash-undir}}} V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2)$. By definition, there exists $v \in \Psi_{\text{Nash-undir}}(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2)$, given by

$$\begin{aligned} v = (& -\mathbf{L}_1 \boldsymbol{x}_1 - \mathbf{L}_1 \boldsymbol{z}_1 + g_{1,(\boldsymbol{x}_1, \boldsymbol{x}_2)}, \\ & -\mathbf{L}_2 \boldsymbol{x}_2 - \mathbf{L}_2 \boldsymbol{z}_2 - g_{2,(\boldsymbol{x}_1, \boldsymbol{x}_2)}, \mathbf{L}_1 \boldsymbol{x}_1, \mathbf{L}_2 \boldsymbol{x}_2), \end{aligned}$$

where $g_{1,(\boldsymbol{x}_1, \boldsymbol{x}_2)} \in \boldsymbol{\partial}_{\boldsymbol{x}_1} U(\boldsymbol{x}_1, \boldsymbol{x}_2)$ and $g_{2,(\boldsymbol{x}_1, \boldsymbol{x}_2)} \in \boldsymbol{\partial}_{\boldsymbol{x}_2} U(\boldsymbol{x}_1, \boldsymbol{x}_2)$, such that

$$\begin{aligned} \xi = {} & v \cdot \nabla V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \\ = {} & (\boldsymbol{x}_1 - \boldsymbol{x}_1^*)^T(-\mathbf{L}_1 \boldsymbol{x}_1 - \mathbf{L}_1 \boldsymbol{z}_1 + g_{1,(\boldsymbol{x}_1, \boldsymbol{x}_2)}) \\ & + (\boldsymbol{x}_2 - \boldsymbol{x}_2^*)^T(-\mathbf{L}_2 \boldsymbol{x}_2 - \mathbf{L}_2 \boldsymbol{z}_2 - g_{2,(\boldsymbol{x}_1, \boldsymbol{x}_2)}) \\ & + (\boldsymbol{z}_1 - \boldsymbol{z}_1^*)^T \mathbf{L}_1 \boldsymbol{x}_1 + (\boldsymbol{z}_2 - \boldsymbol{z}_2^*)^T \mathbf{L}_2 \boldsymbol{x}_2. \end{aligned}$$

Note that $-\mathbf{L}_1 \boldsymbol{x}_1 - \mathbf{L}_1 \boldsymbol{z}_1 + g_{1,(\boldsymbol{x}_1, \boldsymbol{x}_2)} \in -\boldsymbol{\partial}_{\boldsymbol{x}_1} F_1(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2)$, $\mathbf{L}_1 \boldsymbol{x}_1 \in \boldsymbol{\partial}_{\boldsymbol{z}_1} F_1(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2)$, $-\mathbf{L}_2 \boldsymbol{x}_2 - \mathbf{L}_2 \boldsymbol{z}_2 - g_{2,(\boldsymbol{x}_1, \boldsymbol{x}_2)} \in -\boldsymbol{\partial}_{\boldsymbol{x}_2} F_1(\boldsymbol{x}_1, \boldsymbol{z}_2, \boldsymbol{x}_2)$, and $\mathbf{L}_2 \boldsymbol{x}_2 \in \boldsymbol{\partial}_{\boldsymbol{z}_2} F_2(\boldsymbol{x}_2, \boldsymbol{z}_2, \boldsymbol{x}_1)$. Using the first-order convexity property of F_1 and F_2 in their first two arguments, one gets

$$\begin{aligned} \xi \leq {} & F_1(\boldsymbol{x}_1^*, \boldsymbol{z}_1, \boldsymbol{x}_2) - F_1(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2) + F_2(\boldsymbol{x}_2^*, \boldsymbol{z}_2, \boldsymbol{x}_1) \\ & - F_2(\boldsymbol{x}_2, \boldsymbol{z}_2, \boldsymbol{x}_1) + F_1(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2) - F_1(\boldsymbol{x}_1, \boldsymbol{z}_1^*, \boldsymbol{x}_2) \\ & + F_2(\boldsymbol{x}_2, \boldsymbol{z}_2, \boldsymbol{x}_1) - F_2(\boldsymbol{x}_2, \boldsymbol{z}_2^*, \boldsymbol{x}_1). \end{aligned}$$

By substituting each term in the right-hand side and using the fact that $(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*, \boldsymbol{z}_2^*) \in \mathrm{Eq}(\Psi_{\text{Nash-undir}})$,

$$\begin{aligned} \xi \leq {} & -\tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2) + \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*) - \boldsymbol{z}_1^* \mathbf{L}_1 \boldsymbol{x}_1 \\ & - \frac{1}{2} \boldsymbol{x}_1 \mathbf{L}_1 \boldsymbol{x}_1 - \boldsymbol{z}_2^* \mathbf{L}_2 \boldsymbol{x}_2 - \frac{1}{2} \boldsymbol{x}_2 \mathbf{L}_2 \boldsymbol{x}_2. \end{aligned}$$

By rearranging, we thus have $\xi \leq -F_2(\boldsymbol{x}_2, \boldsymbol{z}_2^*, \boldsymbol{x}_1^*) - F_1(\boldsymbol{x}_1, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*)$. Next, since $F_2(\boldsymbol{x}_1^*, \boldsymbol{z}_2^*, \boldsymbol{x}_2^*) + F_1(\boldsymbol{x}_2^*, \boldsymbol{z}_2^*, \boldsymbol{x}_1^*) = 0$, we conclude

$$\begin{aligned} \xi \leq {} & F_1(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*) - F_1(\boldsymbol{x}_1, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*) \\ & + F_2(\boldsymbol{x}_2^*, \boldsymbol{z}_2^*, \boldsymbol{x}_1^*) - F_2(\boldsymbol{x}_2, \boldsymbol{z}_2^*, \boldsymbol{x}_1^*), \end{aligned}$$

yielding that $\xi \leq 0$. As a result,

$$\max \bar{\mathcal{L}}_{\Psi_{\text{Nash-undir}}} V(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \leq 0.$$

We also conclude that the trajectories of (12) are bounded. By Theorem 2.4, any trajectory starting from an initial condition $(\boldsymbol{x}_1^0, \boldsymbol{z}_1^0, \boldsymbol{x}_2^0, \boldsymbol{z}_2^0)$ converges to the largest positively invariant set M in $S_{\Psi_{\text{Nash-undir}}, V} \cap V^{-1}(\leq V(\boldsymbol{x}_1^0, \boldsymbol{z}_1^0, \boldsymbol{x}_2^0, \boldsymbol{z}_2^0))$. Let $(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \in M$. Because $M \subset S_{\Psi_{\text{Nash-undir}}, V}$, then $F_1(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*) - F_1(\boldsymbol{x}_1, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*) = 0$, i.e.,

$$-\tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*) + \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*) - \boldsymbol{z}_1^* \mathbf{L}_1 \boldsymbol{x}_1 - \frac{1}{2} \boldsymbol{x}_1^T \mathbf{L}_1 \boldsymbol{x}_1 = 0. \tag{13}$$

Define now $G_1 : \mathbb{R}^{n_1 d_1} \times \mathbb{R}^{n_1 d_1} \times \mathbb{R}^{n_2 d_2} \to \mathbb{R}$ by $G_1(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2) = F_1(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2) - \frac{1}{2}\boldsymbol{x}_1^T \mathbf{L}_1 \boldsymbol{x}_1$. G_1 is convex in its first argument and linear in its second, and that it has the same saddle points as F_1. As a result, $G_1(\boldsymbol{x}_1^*, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*) - G_1(\boldsymbol{x}_1, \boldsymbol{z}_1^*, \boldsymbol{x}_2^*) \leq 0$, or equivalently, $-\tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*) + \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*) - \boldsymbol{z}_1^* \mathbf{L}_1 \boldsymbol{x}_1 \leq 0$. Combining this with (13), we have that $\mathbf{L}_1 \boldsymbol{x}_1 = 0$ and $-\tilde{U}(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*) + \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2^*) = 0$. Since $\tilde{U}$ is strictly concave in its first argument $\boldsymbol{x}_1 = \boldsymbol{x}_1^*$. A similar argument establishes that $\boldsymbol{x}_2 = \boldsymbol{x}_2^*$. Using now the fact that M is positively invariant, one can deduce that $\mathbf{L}_j \boldsymbol{z}_j \in -\partial_{\boldsymbol{x}_j} \tilde{U}(\boldsymbol{x}_1, \boldsymbol{x}_2)$, for $j \in \{1, 2\}$, and thus $(\boldsymbol{x}_1, \boldsymbol{z}_1, \boldsymbol{x}_2, \boldsymbol{z}_2) \in \mathrm{Eq}(\Psi_{\text{Nash-undir}})$. ■

Remark 5.4: **(Comparison with the best-response dynamics):** The advantage of using the gradient flow is that it avoids the cumbersome computation of the best-response map. This, however, does not come for free. There are convex-concave functions for which the (distributed) gradient flow dynamics, unlike the best-response dynamics, fails to converge to the saddle point, see [16] for an example. •

Remark 5.5: **(Zero-sum games with more than two adversarial networks):** It is known that there are continuous-time zero-sum games with three players and strictly concave-convex payoff functions, for which even the best-response dynamics fails to converge, see [14]. This leaves little hope for extending the results of Theorem 5.3 to N-network zero-sum games, with $N \in \mathbb{Z}_{\geq 3}$. •

VI. Conclusions and future work

We have considered a class of strategic scenarios in which two networks of agents are involved in a zero-sum game. The networks' objectives are to either maximize or minimize a common objective function. Individual agents collaborate with neighbors in their respective network and have partial knowledge of the state of the agents in the other network. We have proposed a distributed saddle-point dynamics that is implementable by each network via local interactions. We have shown that, for a class of strictly concave-convex and locally Lipschitz objective functions, the proposed dynamics is guaranteed to converge to the Nash equilibria. Our algorithm synthesis builds on a continuous-time optimization strategy designed to find the set of minimizers of a sum of convex functions in a distributed way. As a byproduct of our study, we have shown that this strategy can be interpreted itself as a saddle-point dynamics and have established its convergence properties. Future areas of work include the extension of the convergence analysis for directed network topologies; the generalization of our results to not necessarily strict concave-convex functions; and the applications to various areas, including social networks, collective bargaining, and collaborative pursuit-evasion.

Acknowledgments

The first author is grateful to Dr. Abdol Reza Mansouri for useful discussions. This research was partially supported by AFOSR Award FA9550-10-1-0499.

References

[1] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[2] 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.

[3] B. Johansson, M. Rabi, and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM Journal on Control and Optimization*, vol. 20, no. 3, pp. 1157–1170, 2009.

[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] 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.

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

[7] 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.

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

[9] J. Wang and N. Elia, "Control approach to distributed optimization," in *Allerton Conf. on Communications, Control and Computing*, (Monticello, IL), pp. 557–561, Oct. 2010.

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

[11] D. Maistroskii, "Gradient methods for finding saddle points," *Matekon*, vol. 13, pp. 3–22, 1977.

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

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

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

[15] K. Arrow, L. Hurwitz, and H. Uzawa, *A Gradient Method for Approximating Saddle Points and Constrained Maxima*. United States Army Air Forces: Rand Corporation, 1951.

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

[17] R. T. Rockafellar, *Convex Analysis*. Princeton Landmarks in Mathematics and Physics, Princeton, NJ: Princeton University Press, 1997. Reprint of 1970 edition.

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

[19] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides*, vol. 18 of *Mathematics and Its Applications*. Kluwer Academic Publishers, 1988.

[20] J. P. Aubin and A. Cellina, *Differential Inclusions*. New York: Springer, 1994.

[21] N. Biggs, *Algebraic Graph Theory*. Cambridge University Press, 2 ed., 1994.

[22] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. SIAM, 2 ed., 1999.

[23] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *Symposium on Information Processing of Sensor Networks*, (Berkeley, CA), pp. 20–27, Apr. 2004.

[24] P. Wan and M. D. Lemmon, "Event-triggered distributed optimization in sensor networks," in *Symposium on Information Processing of Sensor Networks*, (San Francisco, CA), pp. 49–60, 2009.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.