Optimal Targeting in Super-Modular Games

Giacomo Como, *Member, IEEE*, Stéphane Durand, and Fabio Fagnani, *Member, IEEE*

***Abstract*—We study an optimal targeting problem for super-modular games with binary actions and finitely many players. The considered problem consists in the selection of a subset of players of minimum size such that, when the actions of these players are forced to a controlled value while the others are left to repeatedly play a best response action, the system will converge to the greatest Nash equilibrium of the game. Our main contributions consist in showing that the problem is NP-complete and in proposing an efficient iterative algorithm for its solution with provable probabilistic convergence properties. We discuss in detail the special case of network coordination games and its relation with the graph-theoretic notion of cohesiveness. Finally, through numerical simulations we compare the efficacy of our approach with respect to naive heuristics based on classical network centrality measures.**

***Index Terms*—optimal targeting; network intervention; network games; super-modular games; strategic complements; Nash equilibrium selection.**

I. Introduction

In a game with multiple Nash equilibria, what is the minimum number of players to target in order to force the system to move from an original Nash equilibrium A to a desired Nash equilibrium B? This paper deals with such a problem for the class of super-modular games with binary actions and where the two Nash equilibria A and B are, respectively, the least and the greatest ones in the game. In this paper, we show that the problem is NP-complete in general and we propose the design of an iterative randomized algorithm for an efficient solution.

The considered problem can be framed in the more general setting of studying minimal intervention strategies needed to drive a multi-agent system governed by agents' myopic utility maximization to a desired configuration. In applications where the goal is to achieve a social optimum, such interventions are often modeled as perturbations of the utility functions that lead to a modification of the Nash equilibria of the game. This viewpoint is natural for instance in analyzing the effect of taxes or subsidies in economic models or prices and tolls in transportation systems. More recently, a similar approach has been proposed in the context of network quadratic games [2] to model incentive interventions for instance in school and economic systems.

Some of the results in the paper appeared in preliminary form in [1].

The authors are with the Department of Mathematical Sciences "G.L. Lagrange," Politecnico di Torino, 10129 Torino, Italy (e-mail: {giacomo.como; stephane.durand; fabio.fagnani}@polito.it).

Giacomo Como is also with the Department of Automatic Control, Lund University, BOX 118, 22100 Lund, Sweden, where he is a member of the excellence centre ELLIIT.

This work was partially supported by a MIUR grant Dipartimento di Eccellenza 2018–2022 [CUP: E11G18000350001], a MIUR Research Project PRIN 2017 "Advanced Network Control of Future Smart Grids" (http://vectors.dieti.unina.it), and by the Compagnia di San Paolo.

A different viewpoint, that is the one considered in this paper, is that of individuating a subset of nodes (desirably small) that, when suitably controlled, lead the entire system to the desired equilibrium. The minimum cardinality of this set can also be interpreted as a measure of resilience of the system's equilibrium: the larger it is, the more energy is needed by an external intervention to destabilize it. In the context of binary actions $\{0, 1\}$ considered in this paper, the control action simply amounts to force the set of chosen players, originally playing action 0 state, to play action 1. This well models situations where action 1 indicates the use of a certain technology or the adoption of a new product and the control action corresponds for instance to a marketing intervention where, a certain item is offered for free to the targeted individuals.

Super-modular games have received a great deal of attention in the recent years as the basic way to model strategic complementarity effects [3]. Their numerous applications include modeling of social and economic behaviors like adoption of a new technology, participation to an event, or provision of a public good effort. In certain cases, games that are not super-modular can be made such through a change of variables (see, e.g., the discussion in [4] on the log super-modularity of Cournot oligopoly models). Super-modular games are typically endowed with multiple Nash equilibria that admit a Pareto ordering and the problem of the minimal effort needed to push the system from a lower to a greater equilibrium is natural and relevant in all these applicative contexts.

A fundamental example of super-modular games is that of network coordination games. These are analyzed in detail in [5] where the key graph-theoretic concept of cohesiveness of a set of players is introduced and then used in order to characterize all Nash equilibria. Moreover, the question whether an initial seed of influenced players (that maintain action 1 in all circumstances) is capable of propagating to the whole network is addressed in the same paper and an equivalent characterization of this spreading phenomenon is also expressed in terms of the related notion of uniform non-cohesiveness. More results in this direction are proved in [6] where the authors determine upper bounds to the maximum range of the spreading phenomenon as a function of the cardinality of the original seed. Other studies of this model include [7] for complete networks as well as [8] and in [9] for random graphs and random activation thresholds.

This contagion phenomenon is the content of our analysis in the more general framework of super-modular games. A subset of nodes from which propagation is successful is called a sufficient control sets and our goal is to find such sets of minimum possible cardinality. We notice that the condition proposed in [5] is computationally quite demanding and in practice it cannot be used directly to solve the optimization problem even for medium size games. Indeed, even determining whether a given subset of players is a sufficient control set requires a number of checks growing exponentially in the cardinality of the complement of such set.

The complementary problem of determining (for network coordination games) what is the maximum possible spreading of the state 1 starting from an initial seed of a given number k of targeted players was studied in the seminal paper [10]. While the problem in [10] and ours are related, they are independent, in the sense that solving one does not provide a solution of the other. Another point worth stressing out is that, in their setting, the authors of [10] consider players equipped with random independent activation thresholds and chose to optimize the expected size of the maximum spreading. They prove that such functional is sub-modular and then design a greedy algorithm for obtaining sub-optimal solutions. The introduced randomness is actually crucial in for the results obtained in [10], as the functional considered is not sub-modular for deterministic choices of thresholds. This lack of sub-modularity is actually a key feature of network coordination games where the utility functions present a threshold behavior and make it unfeasible to try to approximate our targeting problem by iteratively adding target nodes in a greedy way. In this regard, the spreading phenomenon we are considering is also quite different than the one analyzed in [11] where the underlying activation process is based on pairwise contagion and not on thresholds.

A targeting intervention problem, related to the one studied in [10], is considered in [12]. There, the authors consider the problem of a firm selling a good to a set of individuals organized through a social network. The firm, in order to maximize its profit, chooses a set of individuals on which to concentrate its advertising efforts or other marketing strategies relative to that specific good. The role of the social network is either of propagating information (in a gossip pairwise style) regarding the good so as to push other people to buy it, or rather to model a positive externality effect where the utility of an individual to buy that product depends on the number of neighbors already using it. This second instance is particularly related to the problem studied in [10] with the important difference that here authors model the network in a mean field fashion only considering the degree distribution.

A different targeting intervention problem is studied in [13] where authors consider network quadratic games and individuate the k most influential players by studying how the aggregate output decreases when this set of players is removed from the network. In the same context of network quadratic games, an optimal pricing problem similar to the one considered in [12] has been analyzed in [14]. The general problem of determining the best set of nodes to exert the most effective control in a networked system has recently appeared in other contexts. In [15]–[17] this is studied in the context of controllability problems for general linear network systems. In [18]–[20], the authors focus on the problem of the optimal position of stubborn influencers in voter models or in linear opinion dynamics. The effect of stubborn agents on the equilibrium behavior of such models is also analyzed in [21] and [22].

In this paper, for arbitrary finite super-modular games with binary action set, we study the problem of finding subsets of players with minimum cardinality such that if the actions of these players are forced to 1, then the game admits an improvement path from every strategy profile to the all-1 strategy profile. We shall refer to such subsets of players as optimal sufficient control sets. Our main contribution is threefold. First, we show how the optimal sufficient control set problem admits a simple solution in the special case of network coordination games in complete graphs. Second, we prove that, in contrast, finding optimal sufficient control sets is an NP-complete problem for general super-modular games, and in fact also just for network coordination games on general graphs. This is shown by reducing the considered problem to the well known 3-SAT problem. Third, we design an iterative randomized search algorithm with provable convergence properties towards sufficient control sets of minimum cardinality for general finite super-modular games with binary action set. The core of the algorithm is a time-reversible Markov chain over the family of all sufficient control sets that starts with the full set, moves through all of them in an ergodic way, and concentrates its mass on those of minimum cardinality.

The rest of the paper is organized as follows. In the final part of this section we report some basic notation conventions. Section II is dedicated to the formal introduction of the problem and in particular of the concept of sufficient control sets. Here we introduce the important notion of monotone improvement path (appeared for other purposes in [23] and [24]) and we give an equivalent (but more operative) characterization of sufficient control sets. In Section III, we focus on the special case of network coordination games: first, we consider the special case of an arbitrary graph and homogeneous thresholds and show how our problem is related to the notion of cohesiveness and then we analyse the special case of heterogeneous network coordination games on the complete graph. Section IV is dedicated to the complexity analysis: we show that the problem is equivalent to an instance of the 3-SAT problem and thus NP-complete. In Section V we present and analyze a distributed algorithm to find optimal sufficient control sets. In Section VI, we report some numerical simulation results comparing the performance of this algorithm with that of certain heristics. Finally, Section VII summarizes the main results of the paper and presents a discussion of possible directions for future research.

We conclude this introduction with a few notational conventions to be adopted throughout the paper. Vectors are indicated in bold-face letters, e.g., $\mathbf{x}$. We define the binary vectors δ_i with entries $(\delta_i)_i = 1$ and $(\delta_i)_j = 0$ for every $j \neq i$. For a subset $\mathcal{S} \subseteq \{1, \dots, n\}$, we put $\mathbb{1}_{\mathcal{S}} = \sum_{i \in \mathcal{S}} \delta_i$. Every $\mathbf{x}$ in $\{0, 1\}^n$ can be written as $\mathbf{x} = \mathbb{1}_S$ for some $S \subseteq \{1, \dots, n\}$. We use the notation $\mathbb{1}$ for the all-1 vector.

II. Problem formulation and basic properties

We consider strategic form games with a finite set of players $\mathcal{V} = \{1, \dots, n\}$ whereby each player i choses her action x_i from a binary set $\mathcal{A} = \{0, 1\}$. Let $\mathcal{X} = \mathcal{A}^n$ denote the (strategy) profile space, whose elements $\mathbf{x}$ will be referred to as (strategy) profiles. We shall consider the standard partial order on $\mathcal{X}$, given by

$$\mathbf{x} \le \mathbf{y} \qquad \Longleftrightarrow \qquad x_i \le y_i, \quad \forall i \in \mathcal{V}\,. \tag{1}$$

As customary, given a strategy profile $\mathbf{x}$ in $\mathcal{X}$ and a player i, we indicate with $\mathbf{x}_{-i}$ the strategy profile of all players but i. Each player i is endowed with a utility function $u_i : \mathcal{X} \to \mathbb{R}$, so that

$$u_i(\mathbf{x}) = u_i(x_i, \mathbf{x}_{-i})\,,$$

denotes the utility of player i when she plays action x_i while the rest of the players' strategy profile is $\mathbf{x}_{-i}$. A game will be formally identified by the triple $(\mathcal{V}, \mathcal{A}, \{u_i\})$.

The best response coorespondance for a player i in $\mathcal{V}$ is the set-valued map

$$\mathcal{B}_i(\mathbf{x}_{-i}) = \mathrm{argmax}_{a\in\mathcal{A}}\, u_i(a, \mathbf{x}_{-i})\,,$$

while the set of pure strategy Nash equilibria is formally defined as

$$\mathcal{N} = \{\mathbf{x} \in \mathcal{X} \,|\, x_i \in \mathcal{B}_i(\mathbf{x}_{-i}),\, \forall i \in \mathcal{V}\}\,.$$

Throughout the paper, we shall consider games satisfying the following *increasing difference property* [25].

Assumption 1: For every player i in $\mathcal{V}$ and every two strategy profiles $\mathbf{x}, \mathbf{y}$ in $\mathcal{X}$ such that $\mathbf{x}_{-i} \ge \mathbf{y}_{-i}$,

$$u_i(1, \mathbf{x}_{-i}) - u_i(0, \mathbf{x}_{-i}) \ge u_i(1, \mathbf{y}_{-i}) - u_i(0, \mathbf{y}_{-i})\,. \tag{2}$$

Assumption 1 states that the marginal utility of increasing player i's action from $x_i = 0$ to $x_i = 1$ is a non-decreasing function of the strategy profile $\mathbf{x}_{-i}$ of all the other players. For finite games, as is our case, such increasing difference property is equivalent to *super-modularity* [26], [27], [28]. For this reason, we will refer to a game $(\mathcal{V}, \mathcal{A}, \{u_i\})$ satisfying property (1) as a finite super-modular game. In the economic literature, these are also referred to as games of *strategic complements* [29].

A standard result for super-modular games ensures that their set of pure strategy Nash equilibria is always nonempty and that there exist a minimal and a maximal Nash equilibria with respect to the partial order (1). Throughout the paper, we shall assume that such minimal and maximal pure strategy Nash equilibria are the all-0 profile and, respectively, the all-1 profiles. This assumption implies no effective loss of generality since the presence of players that maintain a strict preference for action 0 or action 1 independently from the actions played by the other players can be easily integrated in our framework by suitably modifying the other players' utilities.

In this paper, we study the problem of finding subsets of players $\mathcal{S} \subseteq \mathcal{V}$ of minimal cardinality for which there exists an improvement path from $\mathcal{S}$ to the whole player set $\mathcal{V}$. This is formalized by the following definitions.

Definition 1: For a finite game with binary actions $(\mathcal{V}, \mathcal{A}, \{u_i\})$, a sequence of strategy profiles $(\mathbf{x}^k)_{k=0,\dots,m}$ is an *improvement path* from a set $\mathcal{S} \subseteq \mathcal{V}$ to a set $\mathcal{T} \subseteq \mathcal{V}$ if

1) $\mathbf{x}^0 = \mathbb{1}_{\mathcal{S}}$, $\mathbf{x}^m = \mathbb{1}_{\mathcal{T}}$;
2) for every $1 \le k \le m$ there exists i_k in $\mathcal{V}$ such that:
   - $\mathbf{x}^k_{-i_k} = \mathbf{x}^{k-1}_{-i_k}$ and $x^k_{i_k} \ne x^{k-1}_{i_k}$;
   - $u_{i_k}(\mathbf{x}^k) \ge u_{i_k}(\mathbf{x}^{k-1})$.

Definition 2 (Sufficient control set): For a finite game with binary actions $(\mathcal{V}, \mathcal{A}, \{u_i\})$:

- $\mathcal{S} \subseteq \mathcal{V}$ is a *sufficient control set* if there exists an improvement path from $\mathcal{S}$ to $\mathcal{V}$;
- a sufficient control set is *minimal* if none of its proper subsets is a sufficient control set;
- a sufficient control set $\mathcal{S} \subseteq \mathcal{V}$ is *optimal* if there exists no sufficient control set of strictly smaller cardinality.

Notice that sufficient control sets always exist, as the whole set of players $\mathcal{V}$ trivially is a sufficient control set. Also, observe that an optimal sufficient control set is necessarily minimal, but not vice versa. Our objective is to find optimal sufficient control sets: as we shall see in Section V, this can be performed by suitably exploring the space of minimal control sets.

A key fact is that, in dealing with the concept of sufficient control set, it is not restrictive to consider exclusively improvement paths where all action changes are from 0 to 1. Such improvement paths are formally defined below.

Definition 3 (Monotone Improvement path): For a finite game with binary actions $(\mathcal{V}, \mathcal{A}, \{u_i\})$, an improvement path $(\mathbf{x}^k)_{k=0,\dots,m}$ from a set $\mathcal{S} \subseteq \mathcal{V}$ to a set $\mathcal{T} \subseteq \mathcal{V}$ is *monotone* if for $1 \le k \le m$, there exists a player i_k in $\mathcal{T} \setminus \mathcal{S}$ such that $\mathbf{x}^k = \mathbf{x}^{k-1} + \delta_{i_k}$.

Remark 1: If there exists a monotone improvement path $(\mathbf{x}^k)_{k=0,\dots,m}$ from $\mathcal{S} \subseteq \mathcal{V}$ to $\mathcal{T} \subseteq \mathcal{V}$, then necessarily $\mathcal{S} \subseteq \mathcal{T}$, $m = |\mathcal{T} \setminus \mathcal{S}| = |\mathcal{T}| - |\mathcal{S}|$, and the players $i_1, \dots, i_m$ in $\mathcal{T} \setminus \mathcal{S}$ subsequently changing action from 0 to 1 are distinct. In fact, a monotone improvement path is completely specified such m-tuple of action-changing players.

The following result formalizes our previous claim.

Lemma 1: In a finite super-modular game with binary actions $(\mathcal{V}, \mathcal{A}, \{u_i\})$, a subset of players $\mathcal{S} \subseteq \mathcal{V}$ is a sufficient control set if and only if there exists a monotone improvement path from $\mathcal{S}$ to $\mathcal{V}$.

Proof: Clearly, if there exists a monotone improvement path from $\mathcal{S}$ to $\mathcal{V}$, then $\mathcal{S}$ is a sufficient control set.

Conversely, if $\mathcal{S}$ is a sufficient control set, then there exists a (not necessarily monotone) improvement path $(\mathbf{y}^k)_{k=0,\dots,l}$ from $\mathcal{S}$ to $\mathcal{V}$. For every player i in $\mathcal{V} \setminus \mathcal{S}$, let

$$k(i) = \min\{k = 1, \dots, l \mid \mathbf{y}^k = \mathbf{y}^{k-1} + \delta_i\}$$

be the first time that player i changes her action from 0 to 1 along the considered path. Now, let $m = n - |\mathcal{S}|$ and order the players in $\mathcal{V} \setminus \mathcal{S}$ as $i_1, \dots, i_m$ in such a way that $k(i_1) < k(i_2) < \dots < k(i_m)$. For $0 \le h \le m$, define

$$\mathbf{x}^h = \mathbb{1}_{\mathcal{S}} + \sum_{j=1}^{h} \delta_{i_j}\,,$$

and notice that $\mathbf{x}^h \ge \mathbf{y}^{k(i_h)}$. Then,

$$\begin{aligned} u_i(\mathbf{x}^h) - u_i(\mathbf{x}^{h-1}) &= u_i(1, \mathbf{x}^{h-1}_{-i_h}) - u_i(0, \mathbf{x}^{h-1}_{-i_h}) \\ &\ge u_i(1, \mathbf{y}^{k_{i_h}}_{-i_h}) - u_i(0, \mathbf{y}^{k_{i_h}}_{-i_h}) \\ &= u_i(\mathbf{y}^{k_{i_h}}) - u_i(\mathbf{y}^{k_{i_h}-1}) \\ &\ge 0\,, \end{aligned}$$

where the first inequality follows from the increasing difference property (since $\mathbf{x}^h_{-i_h} \geq \mathbf{y}^{k(i_h)}_{-i_h}$), and the last one from the fact that $(\mathbf{y}^k)_{k=0,\dots,l}$ is an improvement path. This shows that $(\mathbf{x}^k)_{k=0,\dots,m}$ is an improvement path from $\mathcal{S}$ to $\mathcal{V}$. By construction, it is monotone, thus proving the claim. ■

This new characterization of sufficient control sets allows for proving the following intuitive fact.

Proposition 1 (monotonicity for inclusion): In a finite supermodular game with binary actions $(\mathcal{V}, \mathcal{A}, \{u_i\})$, if $\mathcal{S} \subseteq \mathcal{V}$ is a sufficient control set, then every $\mathcal{T} \subseteq \mathcal{V}$ such that $\mathcal{S} \subseteq \mathcal{T}$ is also a sufficient control set.

Proof: For a sufficient control set $\mathcal{S}$, by Lemma 1 there exists a monotone improvement path $(\mathbf{x}^k)_{k=0,\dots,m}$ from $\mathcal{S}$ to $\mathcal{V}$. For $k = 1, \dots, m$, let i_k in $\mathcal{V} \setminus \mathcal{S}$ be such that $\mathbf{x}^k = \mathbf{x}^{k-1} + \delta_{i_k}$. Consider some $\mathcal{T} \supsetneq \mathcal{S}$. Let $l = |\mathcal{T}| - |\mathcal{S}|$ and $0 = k_0 < k_1 < k_2 < \dots < k_l \leq m$ be such that $i_{k_h} \in \mathcal{V} \setminus \mathcal{T}$ for $1 \leq h \leq l$. For $0 \leq h \leq l$, let $\mathbf{y}^h = \max\{\mathbb{1}_{\mathcal{T}}, \mathbf{x}^{k_h}\} \geq \mathbf{x}^{k_h}$. Then, a supermodularity argument analogous to the one in the proof of Lemma 1 shows that $(\mathbf{y}^k)_{k=0,\dots,m'}$ is a monotone improvement path from $\mathcal{T}$ to $\mathcal{V}$, thus proving the claim. ■

Remark 2: The notion of sufficient control set introduced in Definition 2 can be reinterpreted in terms of the asynchronous best response dynamics. Given a subset of players $\mathcal{S} \subseteq \mathcal{V}$, consider the Markov chain X^t on the strategy profile space $\mathcal{X}$ whose transitions are described as follows. At every discrete time, a player, among those in $\mathcal{V} \setminus \mathcal{S}$, is chosen uniformly at random and updates her action choosing uniformly at random among the actions of her current best response to the other players' strategy profile. Notice that the existence of an improvement path from $\mathcal{S}$ to $\mathcal{V}$ is equivalent to that, for every initial state X^0 such that $X^0_i = 1$ for all i in $\mathcal{S}$, the Markov chain X^t reaches the all-1 profile $\mathbb{1}$ in finite time with positive probability.

Actually, more is true. For a superset of players $\mathcal{S}' \supseteq \mathcal{S}$, consider the strategy profile $\mathbf{x} = \mathbb{1}_{\mathcal{S}'}$. If $\mathcal{S}$ is a sufficient control set, it follows from Proposition 1 that also $\mathcal{S}'$ is a sufficient. This implies that there exists a monotone improvement path from $\mathcal{S}'$ to $\mathcal{V}$ and thus X^t will also reach $\mathbb{1}$ from $\mathbf{x}$ in finite time with positive probability. If the all-1 strategy profile $\mathbb{1}$ is a strict Nash equilibrium (in the sense that all players have, in that profile, a best response consisting of the singleton 1) then this argument proves that $\mathcal{S} \subseteq \mathcal{V}$ is a sufficient control set if and only if the corresponding Markov chain X^t is absorbed in $\mathbb{1}$ in finite time with probability one. In the more general case where there are indifferent players for whom $u_i(0, \mathbf{x}_{-i}) = u_i(1, \mathbf{x}_{-i})$ for all $\mathbf{x}_{-i}$, then the condition on the Markov chain is replaced by the existence of a set of strategy profiles containing $\mathbb{1}$ on which the Markov chain X^t gets trapped in finite time with probability one and within which it moves ergodically.

III. Optimal targeting in network coordination games

A notable example of super-modular games with binary actions is provided by network coordination games. In this section, after reviewing network coordination games, we study the optimal targeting problem for two classes of them. We first study coordination games on arbitrary undirected networks where the players have homogeneous thresholds: in this case, we characterize best responses and highlight the relationship between the optimal control set problem and the notion of cohesiveness [5]. The second class we consider is that of coordination games on a complete networks with heterogeneous thresholds: in this case, we show that the optimal targeting problem admits a relatively simple analytical solution in the spirit of Granovetter's seminal work [7].

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a (finite, weighted, directed) graph, whereby $\mathcal{V}$ is the set of nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of directed links, and W in $\mathbb{R}_+^{\mathcal{V} \times \mathcal{V}}$ is the weight matrix, such that $W_{ij} > 0$ if and only if there is a link (i, j) in $\mathcal{E}$ directed from its tail node i to its head node j, in which case W_{ij} represents the weight of link (i, j). Let $w_i = \sum_{j \neq i} W_{ij}$ denote the out-degree of a node i in $\mathcal{V}$. We assume that $\mathcal{G}$ contains no self-loops and no sinks, i.e., that $W_{ii} = 0$ and $w_i > 0$ for every i in $\mathcal{V}$. We refer to the graph $\mathcal{G}$ as simple if $W_{ij} = W_{ji} \in \{0, 1\}$: in this case the weight matrix W is completely determined by link set $\mathcal{E}$ and the graph can be simply denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.

A *network coordination game* on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ is a game $(\mathcal{V}, \mathcal{A}, \{u_i\})$ with binary action set $\mathcal{A} = \{0, 1\}$ and utilities

$$u_i(\mathbf{x}) = \sum_{j \neq i} W_{ij} \left((1 - x_i)(1 - x_j) + x_i x_j\right) + c_i x_i \,, \tag{3}$$

where the constant c_i in $[-w_i, w_i]$ models a possible bias of player i towards action 0 (if $c_i < 0$) or action 1 (if $c_i > 0$). In fact, the best response correspondances are given by

$$\mathcal{B}_i(\mathbf{x}_{-i}) = \begin{cases} \{0\} & \text{if} \quad \dfrac{1}{w_i} \displaystyle\sum_{j \neq i} W_{ij} x_j < \theta_i \\ \{0, 1\} & \text{if} \quad \dfrac{1}{w_i} \displaystyle\sum_{j \neq i} W_{ij} x_j = \theta_i \\ \{1\} & \text{if} \quad \dfrac{1}{w_i} \displaystyle\sum_{j \neq i} W_{ij} x_j > \theta_i \end{cases} \tag{4}$$

where

$$\theta_i = \frac{w_i - c_i}{2 w_i} \tag{5}$$

is the threshold of player i in $\mathcal{V}$. In the special case when the graph is simple and $c_i = 0$ (so that the threshold is $\theta_i = 1/2$) for every player i in $\mathcal{V}$, this is also known as the *majority game*.

A. Homogeneous network coordination games

In this subsection, we focus on the special case when the players all have the same threshold $\theta_i = \theta$ in $[0, 1]$. Sufficient control sets in this case can be equivalently formulated in terms of the graph-theoretic notion of cohesiveness introduced in [5]. Specifically, a subset of nodes $\mathcal{S} \subseteq \mathcal{V}$ is called *α-cohesive* in a graph $\mathcal{G}$ if

$$\sum_{j \in \mathcal{S}} W_{ij} \geq \alpha w_i \,, \qquad \forall i \in \mathcal{S} \,. \tag{6}$$

For a simple graph, the above means that every node in $\mathcal{S}$ has at least a fraction α of its neighbors within $\mathcal{S}$ (equivalently,

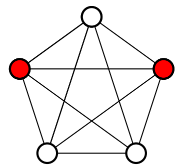


Fig. 1. An optimal sufficient control set for the complete graph with n nodes has size $\lfloor n/2 \rfloor$.

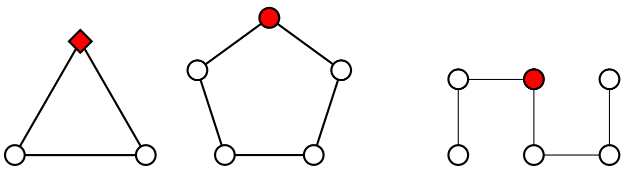


Fig. 2. Optimal sufficient control sets for simple graphs with maximum degree 2 are of size 1.

at most a fraction $1-\alpha$ outside $\mathcal{S}$). Considerations in [5] and in [3] yield the following characterization of sufficient control sets. Define a subset $\mathcal{S} \subseteq \mathcal{V}$ *uniformly no more than θ-cohesive* if no subset $\mathcal{S}' \subseteq \mathcal{S}$ is θ'-cohesive for some $\theta' > \theta$. The following is a consequence of this definition and explicitly proven in [3] (see Proposition 4 therein).

Proposition 2: Consider a network coordination game on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ where all players have the same threshold θ. Then, $\mathcal{S} \subseteq \mathcal{V}$ is a sufficient control set if and only if $\mathcal{V} \setminus \mathcal{S}$ is uniformly no more than $(1-\theta)$-cohesive.

This reformulation of the concept of sufficient control set is of limited interest from the computational point of view. Indeed, checking that the set $\mathcal{V} \setminus \mathcal{S}$ is uniformly no more than $(1-\theta)$-cohesive involves an analysis of all possible subsets of $\mathcal{V} \setminus \mathcal{S}$. Nevertheless, this characterization can be used to analyze special cases. Examples of sufficient control sets for the special case of the majority game for specific simple connected graphs are presented in Figures 1, 2,a nd 4 as well as in Example 1.

Example 1: Let $\mathcal{G}$ be a tree. Then, the set of the leaf nodes is always a sufficient control set. Indeed let $\mathcal{S}$ be any subset of the nodes not containing leaves and consider a path (a walk with no repeated nodes) of maximum length all consisting of nodes in $\mathcal{S}$, say $(i_1, \ldots, i_l)$. Notice that i_1 can not have other neighbors in $\mathcal{S}$ otherwise the path could be extendable. On the other hand, since i_1 is not a leaf in the tree, it must have degree at least 2, namely, at least one neighbor outside of $\mathcal{S}$. This implies that $\mathcal{S}$ can not be θ-cohesive for $\theta > 1/2$. We conclude using again Proposition 2. In general, such sets are not optimal. Indeed, the argument above shows that also the set of nodes that are neighbors of the leaves is a sufficient control set, typically of smaller cardinality than the set of leaves. An example is reported in Figure 3.

The examples considered above show that optimal sufficient control sets for the majority game may exhibit different relative sizes depending on the considered graph. In complete graphs, their size is a constant fraction of the number n of players and we expect the same to hold in very well connected graphs as for instance random Erdös-Rényi graphs. This conjecture is corroborated by numerical simulations presented in Section VI. In contrast, for more loosely connected graphs

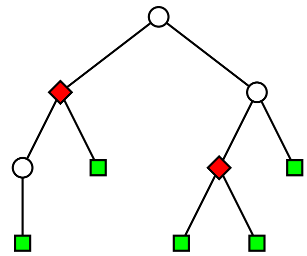


Fig. 3. Two examples of sufficient control sets for a tree: the one consisting of the leaves in green squares and the one consisting of the leaves' neighbors in red diamonds. The second one is optimal.

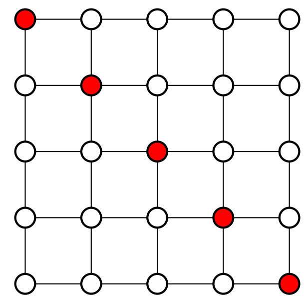


Fig. 4. An optimal sufficient control set for a 2-dimensional grid of order $n = m^2$ has size $m = \sqrt{n}$.

(trees, grids), the size of optimal sufficient control sets scales as a negligible fraction of the size n.

B. Heterogeneous coordination game on the complete graph

In this subsection, we focus on network coordination games on the complete graph, whereby $W_{ij} = 1$ for every $i \neq j$ in $\mathcal{V}$. In contrast with the previous subsection, we shall allow for full heterogeneity of the players' thresholds, that in this case are given by

$$\theta_i = \frac{n - c_i - 1}{2(n-1)}\,, \qquad i \in \mathcal{V}\,. \tag{7}$$

Our results show that optimal sufficient control sets can be completely characterized in terms of the threshold distribution function

$$F(z) = \frac{1}{n}\left|\{i \in \mathcal{V} : \theta_i \leq z\}\right|\,, \qquad z \in [0,1]\,. \tag{8}$$

First, we have the following technical result.

Lemma 2: On an heterogeneous network coordination game on the complete graph with threshold distribution $F(z)$, the empty set $\emptyset$ is a sufficient control set if and only if

$$F(z) \geq z\,, \qquad \forall z \in [0,1]\,. \tag{9}$$

Proof: We start with a general consideration that will be used to prove both implications. Fix $\mathcal{S} \subseteq \mathcal{V}$ and let $\mathbf{x} = \mathbb{1}_{\mathcal{S}}$. Put $n_1 = |\mathcal{S}|$. It follows from (4) that for every player i such that $x_i = 0$, it holds $\mathcal{B}_i(\mathbf{x}_{-i}) = \{0\}$ if only if

$$\theta_i > \frac{n_1}{n-1}\,. \tag{10}$$

Now, suppose that the emptyset $\emptyset$ is not a sufficient control set and let $\mathcal{S} \subsetneq \mathcal{V}$ be a set of maximum cardinality such that there exists a monotone improvement path from $\emptyset$ to $\mathcal{S}$. Put $x = \mathbb{1}_{\mathcal{S}}$ and observe that $n_1 = |\mathcal{S}| \leq n-1$ and that and that (10) holds true for every player i such that $x_i = 0$. Therefore,

$$n - n_1 \leq \left| \left\{ i \in \mathcal{V} : \theta_i > \frac{n_1}{n-1} \right\} \right| = n \left(1 - F\left(\frac{n_1}{n-1} \right) \right) .$$

By dividing both sides by n and rearranging terms, we obtain

$$F\left(\frac{n_1}{n-1}\right) \leq \frac{n_1}{n} < \frac{n_1}{n-1} . \tag{11}$$

This implies that (9) does not hold true.

Suppose instead that (9) does not hold true and let z in $[0,1]$ be such that $F(z) < z$. By the way F is defined, there exists n_1 in $\{0, 1, \dots, n-1\}$ such that $F(z) = n_1/n$. Observe that $n_1 = nF(z) < zn$ implies that $n_1 \leq zn - 1$ and, consequently,

$$\frac{n_1}{n-1} \leq \frac{zn-1}{n-1} \leq z .$$

Then, by monotonicity of the threshold distribution function we get

$$F\left(\frac{n_1}{n-1}\right) \leq F(z) = \frac{n_1}{n} . \tag{12}$$

Let $\mathcal{S}$ be a set consisting of n_1 players with the least possible threshold and let $\mathbf{x} = \mathbb{1}_{\mathcal{S}}$. It then follows from (12) that each player i playing $x_i = 0$ has threshold satisfying (10) and hence, as observed at the beginning of this proof, it is such that $\mathcal{B}_i(x_{-i}) = \{0\}$. This implies that there cannot be a monotone improvement path from $\mathcal{S}$ to $\mathcal{V}$. As a consequence, $\mathcal{S}$ is not a sufficient control set and neither is the empty set $\emptyset$ by Proposition 1. ■

Remark 3: Lemma 2 may be related to Granovetter's seminal work on threshold models [7]. Specifically, [7] studies a model collective behavior whereby n individuals are equipped each with threshold value θ_i in $[0,1]$ and iteratively update their binary action $x_i(t)$ for $t \geq 0$ according to the threshold rule

$$x_i(t+1) = \begin{cases} 1 & \text{if} \quad x_i(t) = 1 \text{ or } \frac{1}{n}\sum_{j\neq i} x_j(t) \geq \theta_i \\ 0 & \text{if} \quad x_i(t) = 0 \text{ and } \frac{1}{n}\sum_{j\neq i} x_j(t) < \theta_i . \end{cases} \tag{13}$$

Let now $z(t) = \frac{1}{n}\sum_{j=1}^{n} x_j(t)$ be the fraction of individuals choosing action 1 at time t and let $z(0) = 0$. Then it can be directly verified that $z(t)$ satisfies the recursion

$$z(t+1) = F\left(z(t)\right) , \qquad t \geq 0 , \tag{14}$$

where $F : [0,1] \to [0,1]$ is the threshold distribution function as defined in (8). Since $F(z)$ is non-decreasing and such that $F(1) = 1$, standard dynamical system arguments imply that $z(t) \to 1$ as t grows large if and only if

$$F\left(z\right) > z , \qquad \forall z \in [0,1) . \tag{15}$$

Notice that the threshold distribution function $F(z)$ is right-continuous and, for a finite population of n players as is our case, it is piecewise constant on $[0,1]$. This implies that, if $F(z^*) = z^*$ for some z^* in $[0,1)$, then, necessarily, $F(z) = z^* < z$ for z in a right neighborhood of z^*, thus showing that (9) and (15) are equivalent. It then follows from Lemma 2 that the empty set $\emptyset$ is a sufficient control set in the heterogeneous coordination game on the complete graph if and only if Granovetter's dynamical system (13) converges to the all-1 configuration when started from the all-0 one.

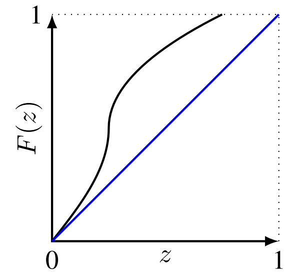


Fig. 5. If the graph of the threshold distribution function $F(z)$ is never below the one of the identity function z, then the empty set $\emptyset$ is a sufficient control set (Lemma 2).

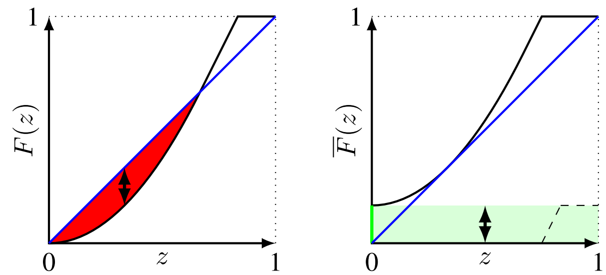


Fig. 6. IIf the graph of the threshold distribution function $F(z)$ is below the one of the identity function for some z in $[0,1]$, then the empty set $\emptyset$ is not a sufficient control set (Lemma 2). On the other hand, a set $\mathcal{S}$ of M players with maximal threshold, where M is defined as in (16) is an optimal sufficient control set (Proposition 3). Forcing these players to play action 1 corresponds to lowering their threshold to 0, leading to a modified threshold distribution function $\overline{F}(z)$ whose graph is obtained by shifting the one of $F(z)$ (and saturating it at 1) by just enough to keep it above the one of the identity function.

As an application of Lemma 2, we obtain the following characterization of the optimal sufficient control sets for heterogeneous network coordination games on the complete graph.

Proposition 3: Consider a heterogeneous network coordination game on the complete graph with threshold distribution $F(z)$. Then, the minimal size of a sufficient control set is

$$M = \left\lceil n \cdot \sup_{0 \leq z \leq 1} \left[z - F(z) \right]_+ \right\rceil . \tag{16}$$

In particular, every $\mathcal{S}$ consisting of M players i in $\mathcal{V}$ with the M largest thresholds θ_i gives an optimal sufficient control set.

Proof: First observe that a subset of players $\mathcal{S} \subseteq \mathcal{V}$ is a sufficient control set for the network coordination game with utilities (3) if and only if $\emptyset$ is a sufficient control set for the modified network coordination game with utilities

$$\overline{u}_i(x) = \sum_{j \neq i} ((1-x_i)(1-x_j) + x_i x_j) + (n-1)x_i , \tag{17}$$

for every i in $\mathcal{S}$, and

$$\sum_{j \neq i} ((1-x_i)(1-x_j) + x_i x_j) + c_i x_i , \tag{18}$$

for every i in $\mathcal{V} \setminus \mathcal{S}$, whereby all the players i in $\mathcal{S}$ have modified threshold $\overline{\theta}_i = 0$ and the rest of the players j in $\mathcal{V} \setminus \mathcal{S}$

have the same threshold $\overline{\theta}_j = \theta_j$. Let $\overline{F}(z)$ be the threshold distribution function of this modified game and observe that

$$0 \leq \overline{F}(z) - F(z) \leq |\mathcal{S}|/n\,, \qquad \forall z \in [0,1]\,. \tag{19}$$

We now show that any subset $\mathcal{S} \subseteq \mathcal{V}$ such that $|\mathcal{S}| < M$ can not be a sufficient control set. If $M = 0$ there is nothing to prove. Assume now that $M \geq 1$ and notice that

$$n \cdot \sup_{0 \leq z' \leq 1} [z' - F(z')]_+ = M - 1 + \epsilon \tag{20}$$

for some $\epsilon > 0$. Since $\sup_{0 \leq z' \leq 1} [z' - F(z')]_+ > 0$, we have that

$$\sup_{0 \leq z' \leq 1} [z' - F(z')]_+ = \sup_{0 \leq z' \leq 1} \{z' - F(z')\}\;. \tag{21}$$

If $|\mathcal{S}| < M$, then (19), (20), and (21) imply that

$$\begin{aligned} 0 &\leq n\left(\overline{F}(z) - F(z)\right) \\ &\leq |\mathcal{S}| \\ &\leq M - 1 \\ &= n \cdot \sup_{0 \leq z' \leq 1} \{z' - F(z')\} - \varepsilon\,, \end{aligned}$$

for every $0 \leq z \leq 1$. This yields

$$z - \overline{F}(z) \geq z - F(z) - \sup_{0 \leq z' \leq 1} \{z' - F(z')\} + \varepsilon/n\,,$$

for every $z \in [0,1]$. Taking the sup on both sides of the above, we finally obtain

$$\sup_{0 \leq z \leq 1} \left\{z - \overline{F}(z)\right\} \geq \varepsilon/n > 0\,.$$

Then, Lemma 2 implies that $\emptyset$ is not a sufficient control set for the modified network coordination game with utilities (17)–(18), hence $\mathcal{S}$ is not a sufficient control set for the original game.

To complete the proof, we now consider a set $\mathcal{S}$ of M players with the highest thresholds. In this case,

$$\begin{aligned} \overline{F}(z) &= \min\{1, F(z) + M/n\} \\ &\geq \min\left\{1, F(z) + \sup_{0 \leq z' \leq 1} [z' - F(z')]_+\right\} \\ &\geq z\,, \end{aligned}$$

for every $0 \leq z \leq 1$. It then follows from Lemma 2 that $\emptyset$ is a sufficient control set for the modified network coordination game with utilities (17)–(18), thus showing that $\mathcal{S}$ is a sufficient control set for the original game. ■

A graphical illustration of Lemma 2 and Proposition 3 is proposed in Figures 5 and 6. If the threshold distribution function of a heterogeneous network coordination game on the complete graph is such that $F(z) \geq z$ for all z in $[0,1]$, then the empty set is a sufficient control set. Otherwise, if $F(z^*) > z^*$ for some z^* in $[0,1]$, then the size of an optimal sufficient control set (16) is given by the minimum integer M such that $\overline{F}(z) = \min\{F(z) + M/n, 1\} \geq z$ for z in $[0,1]$ and any subset of M players with the largest thresholds is a sufficient control set.

In particular, Proposition 3 implies that the problem of finding optimal sufficient sets in for a heterogeneous network coordination game on the complete graph allows for a simple solution. In fact, such solution heavily relies on the symmetry structure of the complete graph. As we shall see in Section IV, finding optimal sufficient control sets is a computationally hard problem for network coordination games on general graphs.

IV. Complexity of finding a sufficient control set

In this section, we study the complexity of finding sufficient control sets for arbitrary super-modular games and prove that it is an NP-complete problem [30, Section 7.4]. Formally, given a binary super-modular game and a positive integer n we define SCS to be the logical proposition "there exists a sufficient control set of size less than or equal to s for the game". The main result of this section is the following.

Theorem 1: The problem SCS is NP-complete.

In order to prove Theorem 1, we will first show that SCS belongs to the complexity class NP (c.f., [30, Definition 7.19]) and then that it is NP-hard.

Lemma 3: The problem SCS belongs to NP.

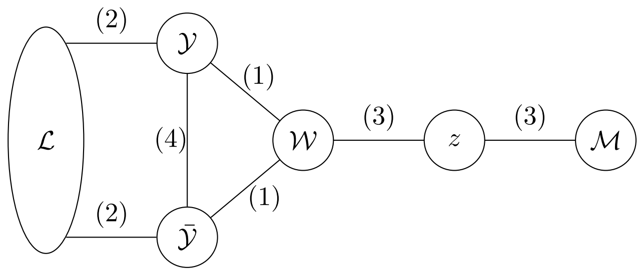
Proof: We show that, given an instance of a finite binary-action super-modular game and a witness consisting in subset of players $\mathcal{S} \subseteq \mathcal{V}$, checking if $\mathcal{S}$ is a sufficient control set can be done in a time growing proportionally to the square of $n - s$, where $n = |\mathcal{V}|$ and $s = |\mathcal{S}|$. In fact, this can be achieved by an iterative algorithm that starts with time index $t = 0$ and profile $\mathbf{x}(0) = \mathbb{1}_{\mathcal{S}}$ and then proceeds as follows. If there exists at least one player i in $\mathcal{V}$ such that

$$x_i(t) = 0\,, \qquad 1 \in \mathcal{B}_i(x_{-i}(t))\,, \tag{22}$$

then arbitrarily chose one such player i, increase the time index t by one unit and define the new profile $\mathbf{x}(t)$ with $x_i(t) = 1$ and $x_{-i}(t) = x_{-i}(t-1)$. Otherwise, if no player i satisfying (22) exists, then halt and return the current value of the time index t. Since, by Proposition 1, every superset of a sufficient control set is itself a sufficient control set, we have that $\mathcal{S}$ is a sufficient control set if and only if the algorithm defined above terminates with $t = n - s$. Clearly, the number of steps of the algorithm is at most $n - s$ and at the t-th step, it is necessary to compute the best responses of at most $n - t$ players, so that the algorithm effectively requires at most $\sum_{t=0}^{n-s-1}(n-t) = (n-s)(n-s+1)/2$ best response computations. This proves that the problem belongs to the complexity class NP. ■

We will now prove that SCS is NP-hard by showing that the 3-SAT problem [30, Ch. 7.2] can be reduced, in polynomial time, to a particular instance of SCS. Consider any instance $I = (X, C)$ of the 3-SAT problem, consisting of a set of variables $X = \{x_1, x_2 \cdots x_{s-1}\}$ and clauses $C = \{c_1, c_2, \cdots c_m\}$, such that in every clause in C exactly three, possibly negated, variables from X appear. Then, we associate to I a simple graph $\mathcal{G}_I = (\mathcal{V}_I, \mathcal{E}_I)$ of order $|\mathcal{V}_I| = 2s + 5m$ and size $|\mathcal{E}_I| = s + 8m$ as follows. The node set $\mathcal{V}_I$ is the union of the following six disjoint sets of nodes:

- A set $\mathcal{W} = \{w_1, w_2, \ldots, w_m\}$, whose elements correspond each to a clause in C;
- A set $\mathcal{Y} = \{y_1, y_2, \ldots, y_{s-1}\}$, whose elements correspond each to a variable in I, with the interpretation that y_i encodes x_i if x_i is true;



- A set $\bar{\mathcal{Y}} = \{\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_{s-1}\}$, whose elements correspond each to a variable in I, with the interpretation that $\bar{y}_i$ encodes x_i if x_i is false;
- A single node z, whose role will be to break possible ties;
- Two sets of leaves $\mathcal{L}$ and $\mathcal{M}$, of cardinality $|\mathcal{L}| = 3m$ and $|\mathcal{M}| = m + 1$.

Links in $\mathcal{E}_I$ only connect pairs of nodes belonging to different sets and in particular:

(1) A node w_j in $\mathcal{W}$ is connected to a node y_i in $\mathcal{Y}$ if and only if the variable x_i appears in the clause c_j and to a node $\bar{y}_i$ in $\bar{\mathcal{Y}}$ if and only if the variable $\bar{x}_i$ appears in the clause c_j;

(2) For each clause containing the variable x_i, node y_i in $\mathcal{Y}$ is connected to a different node in $\mathcal{L}$, and for each clause containing the variable $\overline{x}_i$, node $\overline{y}_i$ in $\overline{\mathcal{Y}}$ is connected to a different node in $\mathcal{L}$, in such a way that the elements of $\mathcal{L}$ are each connected to exactly one element either of $\mathcal{Y}$ or of $\overline{\mathcal{Y}}$;

(3) The node z is connected to every element of $\mathcal{W}$ and of $\mathcal{M}$;

(4) For every $i = 1, \ldots, s-1$, node y_i is connected to the corresponding node $\bar{y}_i$.

There is a total of $3m$ links of type (1), $3m$ links of type (2), $2m+1$ links of type (3), and $s-1$ links of type (4). Nodes in $\mathcal{L}$ and $\mathcal{M}$ all have degree 1, nodes in $\mathcal{W}$ all have degree 4, node z has degree $2m+1$, while the degree of a node y_i in $\mathcal{Y}$ (respectively $\overline{y}_i$ in $\overline{\mathcal{Y}}$) is 1 plus twice the number of clauses the variable x_i (respectively, $\overline{x}_i$) appears in.

Now, we shall consider the majority game on the graph $\mathcal{G}_I$, whereby each player in $\mathcal{V}_I$ has action set $\{0, 1\}$ and the utility of player i is equal to the number of her neighbors that play the same action as her. We then ask the question "is there a sufficient control set of size less than or equal to s for this game?" We will now show that the answer to this question is true if and only if the instance of 3-SAT is satisfiable.

Lemma 4: Let $I = (X, C)$ be an instance of the 3-SAT problem, and let $\mathcal{G}_I = (\mathcal{V}_I, \mathcal{E}_I)$ be the simple graph defined above. If I is satisfiable with a solution x^* in $\{0,1\}^{s-1}$, then

$$\mathcal{S} = \{z\} \cup \{y_i : x_i^* = 1\} \cup \{\overline{y}_i : x_i^* = 0\}$$

is a sufficient control set of size s for the majority game on $\mathcal{G}_I$.

Proof: Since I is satisfied by x^*, for every clause c_j in C there exists i in $\{1, \ldots, s-1\}$ such that either x_i appears in c_j and $x_i^* = 1$ or $\overline{x}_i$ appears in c_j and $\overline{x}_i^* = 1$. Thus, in the graph $\mathcal{G}_I$, all clause-related nodes in $\mathcal{W}$ have at least one neighbor in $(\mathcal{Y} \cup \overline{\mathcal{Y}}) \cap \mathcal{S}$. Since they are all connected to z in $\mathcal{S}$ also, and have all degree 4 in $\mathcal{G}_I$, this implies that there exists a monotone improvement path from $\mathcal{S}$ to $\mathcal{S} \cup \mathcal{W}$.

Now, consider a variable x_i in X and let m_i be the number of clauses it appears in. Then, notice that, if the corresponding node y_i in $\mathcal{Y}$ does not belong to $\mathcal{S}$, it necessarily has one neighbor in $\mathcal{S}$ ($\overline{y}_i$) as well as m_i neighbors in $\mathcal{W}$ (those corresponding to the clauses it belongs to). Since its degree in $\mathcal{G}_I$ is exactly $2m_i + 1$, this implies that $\mathcal{S} \cup \mathcal{W} \cup \mathcal{Y}$ can be reached by a monotone improvement path from $\mathcal{S} \cup \mathcal{W}$, hence from $\mathcal{S}$. Analogously, one proves that $\mathcal{S} \cup \mathcal{W} \cup \mathcal{Y} \cup \overline{\mathcal{Y}}$ can be reached by a monotone improvement path from $\mathcal{S}$.

Finally, since every remaining node in $\mathcal{L} \cup \mathcal{M}$ is of degree one and connected to a node in $\mathcal{Y} \cup \overline{\mathcal{Y}} \cup \{z\}$, we get that the monotone improvement path from $\mathcal{S}$ can be extended to reach the whole node set $\mathcal{V}_I$, thus proving that $\mathcal{S}$ is a sufficient control set. ■

We will now show that the converse of Lemma 4 holds true.

Lemma 5: Let $I = (X, C)$ be an instance of the 3-SAT problem, and let $\mathcal{G}_I = (\mathcal{V}_I, \mathcal{E}_I)$ be the simple graph defined above. If there is a sufficient control set $\mathcal{S}$ of size s for the majority game on $\mathcal{G}_I$, then I is solvable.

Proof: We will first show that there exists a sufficient control set $\mathcal{S}'$ of the same size s containing z and exactly one node among y_i and $\bar{y}_i$ for each $1 \leq i \leq s-1$.

First, for every $i = 1, \ldots, s-1$, at least one node among y_i, $\bar{y}_i$, and the leaves in $\mathcal{L}$ connected to them must be in $\mathcal{S}$. Otherwise, neither of the pair $\{y_i, \bar{y}_i\}$ can be converted before the other preventing any improvement path. For the same reason at least one element among z and the leaves in $\mathcal{M}$ must be in $\mathcal{S}$.

In case when neither y_i nor $\bar{y}_i$ belong to $\mathcal{S}$, removing the leaf connected to them that is in $\mathcal{S}$ and adding its sole neighbor (either y_i or $\bar{y}_i$) let the control set stay sufficient and preserves its size. We construct $\mathcal{S}'$ in this way replacing leaves with variable nodes and finally applying the same substitution idea to include the node z removing a leaf connected to it. This gives us a sufficient control set of the same size as before, using no leave and having exactly z and one among each pair of variables.

Because of the structure of the graph and since $\mathcal{S}'$ contains no leaves in $\mathcal{L} \cup \mathcal{M}$, in any monotone improvement path from $\mathcal{S}'$ to $\mathcal{V}_I$, a node in $(\mathcal{Y} \cup \overline{\mathcal{Y}}) \setminus \mathcal{S}'$ can only appear after all nodes in $\mathcal{W}$ have already appeared. Since all nodes in $\mathcal{W}$ have degree 4, this says that each of them must have at least two neighbors in $\mathcal{S}'$, one of them being z. This implies that every node in $\mathcal{W}$ must have at least one neighbor in $\mathcal{S}' \setminus \{z\} \subseteq \mathcal{Y} \cup \overline{\mathcal{Y}}$.

Consider now the candidate solution x^* in $\{0,1\}^{s-1}$ that has $x_i^* = 1$ if and only if y_i in $\mathcal{S}'$. Then, it follows from the argument above that for every clause c_j there exists i in $\{1, \ldots, s-1\}$ such that either x_i appears in c_j and $x_i^* = 1$ or $\overline{x}_i$ appears in c_j and $\overline{x}_i^* = 1$. This proves that the instance I is solvable. ■

Lemma 4 and Lemma 5 thus show that starting from an instance of the 3-SAT, we could build an instance of the SCS problem in polynomial time and of polynomial size, solvable if and only if 3-SAT instance is solvable, with a polynomial way to convert the solution back to the 3-Sat format. This shows that SCS is NP-hard. Together with Lemma 3, this shows that SCS is an NP-complete problem.

Remark 4: Notice that we have in fact proven that the 3-SAT problem can be reduced to the SCS problem for the majority game on an arbitrary graph. Hence, not only is the SCS problem on general super-modular games with binary actions NP-complete, but also its special case for the majority game on general graphs is NP-complete.

V. A DISTRIBUTED OPTIMAL TARGETING ALGORITHM

The characterization of sufficient control sets through monotone improvement paths (Lemma 1) suggests the possibility that such sets may be searched for by starting from the all-1 profile $\mathbb{1}$ and iteratively replacing 1's with 0's in the attempt to follow backwards a monotone improvement path.

In order to capture this intuition, in this section we introduce a family of discrete-time Markov chains $(Z_t^\varepsilon)_{t\geq 0}$ on the strategy profile space $\mathcal{X}$, parameterized by a scalar value ε in $[0,1]$. We will then prove that, for every $0<\varepsilon\leq 1$, the Markov chain $(Z_t^\varepsilon)_{t\geq 0}$ is time-reversible and irreducible on the set of sufficient control sets, whereas, as ε vanishes, its stationary distribution concentrates on the family of optimal sufficient control sets.

The dynamics of the Markov chain Z_t^ε are described as follows: at every discrete time $t=0,1\ldots$, given that $Z_t^\varepsilon=\mathbf{z}$, a player i is chosen uniformly at random from the whole player set $\mathcal{V}$. Then, if $u_i(1,\mathbf{z}_{-i})<u_i(0,\mathbf{z}_{-i})$, the state is not changed, i.e., $Z_{t+1}^\varepsilon=\mathbf{z}$. Otherwise, if $u_i(1,\mathbf{z}_{-i})\geq u_i(0,\mathbf{z}_{-i})$, then, if the current action of player i is $z_i=1$ it is changed to 0 with probability 1, while if her current action is $z_i=0$, it is changed to 1 with probability ε and kept the same with probability $1-\varepsilon$.

Precisely, for a strategy profile $\mathbf{x}$ in $\mathcal{X}$, let

$$n_0(\mathbf{x})=|\{i\in\mathcal{V}:\,x_i=0,\,u_i(1,\mathbf{x}_{-i})\geq u_i(0,\mathbf{x}_{-i})\}|\,,$$

$$n_1(\mathbf{x})=|\{i\in\mathcal{V}:\,x_i=1,\,u_i(1,\mathbf{x}_{-i})\geq u_i(0,\mathbf{x}_{-i})\}$$

and

$$\alpha_\varepsilon(\mathbf{x})=\frac{\varepsilon n_0(\mathbf{x})+n_1(\mathbf{x})}{n}\,.$$

Then, let Z_t^ε be a Markov chain with state space $\mathcal{X}$ and transition probabilities

$$P^{\epsilon}_{\mathbf{x},\mathbf{y}}=\begin{cases}1/n & \text{if}\quad \mathbf{y}=\mathbf{x}-\delta_i \text{ and } u_i(\mathbf{y})\leq u_i(\mathbf{x})\\ \varepsilon/n & \text{if}\quad \mathbf{y}=\mathbf{x}+\delta_i \text{ and } u_i(\mathbf{y})\geq u_i(\mathbf{x})\\ 1-\alpha_\varepsilon(\mathbf{x}) & \text{if}\quad \mathbf{x}=\mathbf{y}\\ 0 & \text{otherwise}\,,\end{cases}\tag{23}$$

for every $\mathbf{x},\mathbf{y}$ in $\mathcal{X}$, where we recall that δ_i stands for a vector of all 0's except for a 1 in the i-th entry.

Notice that, for $\varepsilon=0$, only transitions from action 1 to action 0 are allowed. In this case, the Markov chain Z_t^0 has absorbing states. Let

$$\mathcal{Z}=\{x\in\mathcal{X}\mid\mathbb{P}(\exists t_0\geq 0\,:\,Z_{t_0}^0=x\mid Z_0^0=\mathbb{1})>0\}\tag{24}$$

be the set of all states (i.e., strategy profiles) that are reachable by the Markov chain Z_t^0 when started from $Z_0^0=\mathbb{1}$ and

$$\mathcal{Z}_\infty=\{x\in\mathcal{X}\mid\mathbb{P}(\exists t_0\geq 0\,:\,Z_t^0=x\,\forall t\geq t_0\mid Z_0^0=\mathbb{1})>0\}\tag{25}$$

the set of absorbing states reachable by Z_t^0 from $Z_0^0=\mathbb{1}$. Then, we have the following result:

Proposition 4: For a finite super-modular game with binary actions $(\mathcal{V},\mathcal{A},\{u_i\})$, let $\mathcal{Z}$ and $\mathcal{Z}_\infty$ be defined as in (24) and (25), respectively. Then,

(i) $\mathcal{S}\subseteq\mathcal{V}$ is a sufficient control set if and only if $\mathbb{1}_\mathcal{S}\in\mathcal{Z}$;
(ii) if $\mathcal{S}$ is a minimal sufficient control set then $\mathbb{1}_\mathcal{S}\in\mathcal{Z}_\infty$;
(iii) $\mathbf{x}$ in $\mathcal{X}$ is an absorbing state for Z_t^0 if and only if $n_1(\mathbf{x})=0$.

Proof: (i) By definition, $\mathbf{x}=\mathbb{1}_\mathcal{S}$ belongs to the set of reachable states $\mathcal{Z}$ if and only if there exists a sequence of strategy profiles $(\mathbf{y}^k)_{k=0,\ldots,l}$, such that $\mathbf{y}^0=\mathbb{1}$, $\mathbf{y}^l=\mathbb{1}_\mathcal{S}$, and

$$\mathbf{y}^k=\mathbf{y}^{k-1}-\delta_{i_k}\,,\qquad u_{i_k}(\mathbf{y}^k)\leq u_{i_k}(\mathbf{y}^{k-1})\qquad 1\leq k\leq l\,.\tag{26}$$

Notice that (26) is equivalent to say that the reversed path $(\mathbf{x}^k)_{k=0,\ldots,l}$ with $\mathbf{x}^k=\mathbf{y}^{l-k}$ for $0\leq k\leq l$ is a monotone improvement path from $\mathcal{S}$ to $\mathcal{V}$. By Lemma 1, this is equivalent to say that $\mathcal{S}$ is a sufficient control set.

(ii) If $\mathcal{S}$ is a minimal sufficient control set, we know from point (i) that the strategy profile $\mathbb{1}_\mathcal{S}$ belongs to the set $\mathcal{Z}$ of reachable states. By contradiction, if $\mathbb{1}_\mathcal{S}$ did not belong to the set of reachable absorbing states $\mathcal{Z}_\infty$, then, from $\mathbf{x}=\mathbb{1}_\mathcal{S}$, the Markov chain Z_t^0 could reach, in one step, a different state $\mathbf{x}'=\mathbb{1}_{\mathcal{S}'}$ with $\mathcal{S}'\subsetneq\mathcal{S}$, thus contradicting the minimality assumption on $\mathcal{S}$.

(iii) Since $\alpha_0(\mathbf{x})=n_1(\mathbf{x})/n$, by (23) we have $P^0_{\mathbf{x},\mathbf{x}}=1$ if and only if $n_1(\mathbf{x})=0$. ∎

Point (i) of Proposition 4 implies that the problem of finding optimal sufficient control sets can be equivalently stated as the problem of finding strategy profiles $\mathbf{x}$ in $\mathcal{Z}$ of minimal l_1-norm $||\mathbf{x}||_1=\sum_k x_k$, i.e., that $\mathcal{S}$ is an optimal sufficient control set if and only if

$$|\mathcal{S}|=\min_{\mathbf{x}\in\mathcal{Z}}||\mathbf{x}||_1\,.\tag{27}$$

On the other hand, point (ii) implies that we can actually restrict the minimization in (27) to the set $\mathcal{Z}_\infty$ of those absorbing states of the Markov chain Z_t^0 that are reachable from the all-1 strategy profile. However, as Example 2 below shows, the set $\mathcal{Z}_\infty$ may contain strategy profiles corresponding to sufficient control sets that are suboptimal and, possibly, not even minimal.

Example 2: Consider the majority game on the ring graph with four nodes $\{1,2,3,4\}$. Then, $\mathbf{z}^1=(1,0,1,0)$ in $\mathcal{Z}_\infty$ corresponds to the sufficient control set $\mathcal{S}=\{1,3\}$, but it is not minimal since $\{1\}$ is also a sufficient control set.

As a consequence, by simply simulating the Markov chain Z_t^0 started from $Z_0^0=\mathbb{1}$, we are not guaranteed to reach an optimal sufficient control set. To overcome this issue, we will instead use the Markov chain Z_t^ε with $\varepsilon>0$, which, as shown below, is time-reversible and ergodic on whole set $\mathcal{Z}$ of reachable strategy profiles and, hence, it does not get trapped in non-optimal control sets, and at the same time has a stationary distribution concentrating on the set of optimal control sets as the parameter ε vanishes.

Theorem 2: For a finite super-modular game with binary actions $(\mathcal{V}, \mathcal{A}, \{u_i\})$, let $\mathcal{Z}$ be defined as in (24). Then, for $0 < \varepsilon \leq 1$, the Markov chain Z_t^ε with transition probabilities (23)

(i) keeps the set $\mathcal{Z}$ invariant, namely, if Z_0^ε belongs to $\mathcal{Z}$, then Z_t^ε belongs to $\mathcal{Z}$ for every $t \geq 0$;

(ii) is time-reversible and ergodic on the set $\mathcal{Z}$;

(iii) has stationary probability distribution

$$\mu_{\mathbf{x}}^\varepsilon := \frac{1}{K_\varepsilon} \varepsilon^{||\mathbf{x}||_1}\,, \qquad \mathbf{x} \in \mathcal{Z}\,, \tag{28}$$

where
$$K_\varepsilon = \sum_{\mathbf{x} \in \mathcal{Z}} \varepsilon^{||\mathbf{x}||_1}\,.$$

In particular,
$$\lim_{\varepsilon \downarrow 0} \mu^\varepsilon = \mu\,,$$
where μ is the uniform probability distribution on the set of profiles corresponding to optimal sufficient control sets.

Proof: (i) Let $\mathbf{x}$ in $\mathcal{Z}$ be strategy profile that is reachable from the all-1 profile by the Markov chain Z_t^0 and $\mathbf{y}$ in $\mathcal{X}$ a strategy profile such that $P_{\mathbf{x},\mathbf{y}}^\varepsilon > 0$. We will prove that $\mathbf{y}$ belongs to $\mathcal{Z}$.

If $\mathbf{y} = \mathbf{x} - \delta_i$ for some player i in $\mathcal{V}$, then (23) implies that $P_{\mathbf{x},\mathbf{y}}^\varepsilon = 1/n > 0$, so that in particular $P_{\mathbf{x},\mathbf{y}}^0 = 1/n > 0$, thus showing that the strategy profile $\mathbf{y}$ belongs to $\mathcal{Z}$.

On the other hand, if $\mathbf{y} = \mathbf{x} + \delta_i$ for some player i in $\mathcal{V}$, we argue as follows. Since $\mathbf{x}$ in $\mathcal{Z}$ is a strategy profile reachable by the Markov chain Z_t^0 from the all-1 profile, we can find a sequence of profiles $(\mathbf{x}^k)_{k=0,\ldots,l}$ such that $\mathbf{x}^0 = \mathbb{1}$ and $\mathbf{x}^l = \mathbf{x}$ and $P_{\mathbf{x}^{k-1},\mathbf{x}^k}^0 > 0$ for $1 \leq k \leq l$. From (23), this is equivalent to the existence of players $i_1, i_2, \ldots, i_l$ in $\mathcal{V}$ such that

$$\mathbf{x}^k = \mathbf{x}^{k-1} - \delta_{i_k}\,, \qquad u_{i_k}(\mathbf{x}^k) \leq u_{i_k}(\mathbf{x}^{k-1})\,, \qquad 1 \leq k \leq l\,. \tag{29}$$

Observe that $\mathbf{y} = \mathbf{x} + \delta_i$ implies that $x_i = 0$, so that that there exists a unique integer s in $\{1, \ldots, l\}$ such that $i_s = i$, hence

$$\mathbf{x}^s = \mathbf{x}^{s-1} - \delta_i\,, \qquad u_{i_k}(\mathbf{x}^k) \leq u_{i_k}(\mathbf{x}^{k-1}). \tag{30}$$

Now, define the sequence of profiles $(\mathbf{z}^k)_{k=0,\ldots,l-1}$ as

$$\mathbf{z}^k = \begin{cases} \mathbf{x}^k & \text{if} \quad 0 \leq k \leq s-1 \\ \mathbf{x}^{k+1} + \delta_i & \text{if} \quad s \leq k \leq l-1\,. \end{cases} \tag{31}$$

Observe that $\mathbf{z}^0 = \mathbb{1}$ and $\mathbf{z}^{l-1} = \mathbf{x} + \delta_i = \mathbf{y}$. We are going to show that

$$P_{\mathbf{z}^{k-1},\mathbf{z}^k}^0 > 0\,, \tag{32}$$

for every $1 \leq k \leq l-1$, thus proving that the strategy profile $\mathbf{y}$ is reachable from $\mathbb{1}$ by the Markov chain Z_t^0. In fact, for $1 \leq k \leq s-1$, we simply have $P_{\mathbf{z}^{k-1},\mathbf{z}^k}^0 = P_{\mathbf{x}^{k-1},\mathbf{x}^k}^0 > 0$, thus showing that (32) holds true in this case. On the other hand, it follows from the equalities in (29) and (30) that

$$\mathbf{z}^s = \mathbf{x}^{s+1} + \delta_i = \mathbf{x}^s + \delta_i - \delta_{i_{s+1}} = \mathbf{x}^{s-1} - \delta_{i_{s+1}} = \mathbf{z}^{s-1} - \delta_{i_{s+1}}\,,$$

and, for $s+1 \leq k \leq l-1$,

$$\mathbf{z}^k = \mathbf{x}^{k+1} + \delta_i = \mathbf{x}^k + \delta_i - \delta_{i_{k+1}} = \mathbf{z}^{k-1} - \delta_{i_{k+1}}\,,$$

so that

$$\mathbf{z}^k = \mathbf{z}^{k-1} - \delta_{i_{k+1}}\,, \qquad s \leq k \leq l-1\,. \tag{33}$$

Notice that, for $s \leq k \leq l-1$, $\mathbf{z}^k \geq \mathbf{x}^{k+1}$, so that (29) and the super-modularity property (2) imply that

$$\begin{aligned} 0 &\leq u_{i_{k+1}}(\mathbf{x}^k) - u_{i_{k+1}}(\mathbf{x}^{k+1}) \\ &= u_{i_{k+1}}(1, \mathbf{x}_{-i_{k+1}}^{k+1}) - u_{i_{k+1}}(0, \mathbf{x}_{-i_{k+1}}^{k+1}) \\ &\leq u_{i_{k+1}}(1, \mathbf{z}_{-i_{k+1}}^{k}) - u_{i_{k+1}}(0, \mathbf{z}_{-i_{k+1}}^{k}) \\ &= u_{i_{k+1}}(\mathbf{z}^{k-1}) - u_{i_{k+1}}(\mathbf{z}^k) \end{aligned}$$

Together with (33) and (23), the above implies that $P_{\mathbf{z}^{k-1},\mathbf{z}^k}^0 = 1/n > 0$ for $s \leq k \leq l-1$. Thus, (32) holds true for for every $1 \leq k \leq l-1$, showing that $\mathbf{y}$ belongs to $\mathcal{Z}$.

(ii) Notice that

$$\varepsilon^{||\mathbf{x}||_1} P_{\mathbf{x},\mathbf{y}}^\varepsilon = \varepsilon^{||\mathbf{y}||_1} P_{\mathbf{y},\mathbf{x}}^\varepsilon\,, \tag{34}$$

for every two strategy profiles $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{X}$, thus showing that the Markov chain Z_t^ε is time-reversible with respect to the stationary distribution $\mu_{\mathbf{x}}^\varepsilon$ defined in (28).

Since the positive probability transitions for the Markov chain Z_t^0 have also positive probability for the Markov chain Z_t^ε, we have that all profiles in $\mathcal{Z}$ can be reached from the all-1 profile $\mathbb{1}$ by the Markov chain Z_t^ε for every $0 < \varepsilon \leq 1$. Moreover, Equation (34) shows that a transition probability $P_{\mathbf{x},\mathbf{y}}^\varepsilon$ is positive if and only if the reverse transition $P_{\mathbf{y},\mathbf{x}}^\varepsilon$ is positive, thus $\mathbb{1}$ is reachable from every other profile in $\mathcal{Z}$. This shows that Z_t^ε is ergodic on the set $\mathcal{Z}$.

(iii) Ergodicity and Equation (34) imply that, for every $0 < \varepsilon \leq 1$, the unique stationary distribution of the Markov chain Z_t^ε restricted to $\mathcal{Z}$ has the form (28).

As ε vanishes, a direct check shows that the stationary distribution μ^ε converges to a uniform distribution on the set $\mathrm{argmin}_{\mathbf{x} \in \mathcal{Z}} ||\mathbf{x}||_1$. Using Proposition 4, $\mathrm{argmin}_{\mathbf{x} \in \mathcal{Z}} ||\mathbf{x}||_1$ coincides with the set of optimal sufficient control sets, thus completing the proof. ■

Theorem 2 suggests an iterative stochastic algorithm for the selection of an optimal sufficient control set: simply run the Markov chain Z_t^ε for some sufficiently small value of $\varepsilon > 0$. This can be considered a distributed algorithm in the sense that, for each iteration, once a uniform random player i is selected, the update rule only depends on the current action of player i and on the current action of the other players on which the utility u_i depends on.

In particular, notice that, by Proposition 4 (i), $\mathcal{Z}$ coincides with the space of sufficient control sets. Hence, Theorem 2 implies that, when started from $Z_0^\varepsilon = \mathbb{1}$, the Markov chain Z_t^ε explores the space of sufficient control sets $\mathcal{Z}$ in an ergodic fashion. Writing $Z_t^\varepsilon = \mathbb{1}_{\mathcal{S}_t}$, we have that the expected size of the random sufficient control set $\mathcal{S}_t$ in stationarity can be computed using (28) as

$$\lim_{t \to +\infty} E[|S_t|] = \frac{1}{K_\varepsilon} \sum_{l=M}^{n} l N_l \varepsilon^l\,, \qquad K_\varepsilon = \sum_{l=M}^{n} N_l \varepsilon^l\,,$$

where M is the size of a minimal control set, N_l is the number of sufficient control sets of size l for $l = M, \ldots, n$. As ε vanishes, such expected stationary size converges to the optimal sufficient control set size M.

However, because of the positive diagonal entries $P^{\varepsilon}_{\mathbf{x},\mathbf{x}} = 1-\alpha_{\varepsilon}(\mathbf{x})$ of the transition probability matrix of the Markov chain Z^{ε}_{n}, a straightforward implementation of this algorithm would remain idle form many iterations, the more so the closer to 1 such diagonal entries are. Notice that on the one hand

$$\lim_{\varepsilon\downarrow 0} P^{\varepsilon}_{\mathbf{x},\mathbf{x}} = 1 - n_1(\mathbf{x})/n\,, \qquad \forall \mathbf{x}\in\mathcal{Z}\,,$$

on the other hand, by Proposition 4 (iii), $n_1(\mathbf{x}) = 0$ if and only if $\mathbf{x}\in\mathcal{Z}_{\infty}$. Hence, in particular,

$$\lim_{\varepsilon\downarrow 0} P^{\varepsilon}_{\mathbf{x},\mathbf{x}} = 1\,, \qquad \forall \mathbf{x}\in\mathcal{Z}_{\infty}\,.$$

As a consequence, for small positive values of ε, the algorithm would tend to get stuck in small sufficient sets, waiting for either choosing one of the few remaining best-responding 1-players (if any, i.e., only if the state is not in $\mathcal{Z}_{\infty}$), or for choosing a 0-player with best response 1 and flipping her action with probability ε.

In order to speed up the algorithm, a possibility consists in zeroing the diagonal entries of the transition probability matrix and rescaling its off-diagonal entries in order to keep the row-sums equal to 1, c.f. [31]. Formally, this leads one to consider the modified Markov chain $\tilde{Z}^{\epsilon}_t$ on the same state space $\mathcal{X}$ and transition probabilities

$$\tilde{P}^{\epsilon}_{\mathbf{x},\mathbf{y}} = \begin{cases} 1/(n\alpha_{\varepsilon}(\mathbf{x})) & \text{if} \quad \mathbf{y}=\mathbf{x}-\delta_i \text{ and } u_i(\mathbf{y})\le u_i(\mathbf{x}) \\ \varepsilon/(n\alpha_{\varepsilon}(\mathbf{x})) & \text{if} \quad \mathbf{y}=\mathbf{x}+\delta_i \text{ and } u_i(\mathbf{y})\ge u_i(\mathbf{x}) \\ 0 & \text{otherwise}\,, \end{cases} \tag{35}$$

for all $\mathbf{x},\mathbf{y}$ in $\mathcal{X}$. Clearly, reachable states for the modified Markov chain $\tilde{Z}^{\epsilon}_t$ are the same as those for the original Markov chain Z^{ϵ}_t. On the other hand, typically $\tilde{Z}^{\epsilon}_t$ is not a reversible Markov chain. Nevertheless, we can compute its stationary distribution as stated in the following result. In it, we shall refer to a sufficient subset $\mathcal{U}\subseteq\mathcal{V}$ as quasi-optimal if

$$\mathcal{U} = \mathcal{S}\cup\{i\}\,, \tag{36}$$

for some optimal sufficient control set $\mathcal{S}$ and some player i in $\mathcal{V}\setminus\mathcal{S}$.

Corollary 1: For $0<\varepsilon\le 1$, consider the Markov chain $\tilde{Z}^{\varepsilon}_t$ with transition probabilities (35). Then:

(i) $\tilde{Z}^{\varepsilon}_t$ keeps the set $\mathcal{Z}$ invariant and is ergodic on it;
(ii) $\tilde{Z}^{\varepsilon}_t$ has stationary probability

$$\tilde{\mu}^{\varepsilon}_{\mathbf{x}} := \frac{1}{\tilde{K}_{\varepsilon}}\varepsilon^{||\mathbf{x}||_1}(n_1(\mathbf{x})+\varepsilon n_0(\mathbf{x}))\,, \qquad \mathbf{x}\in\mathcal{Z}\,, \tag{37}$$

where $\tilde{K}_{\varepsilon} = \sum_{\mathbf{x}\in\mathcal{Z}}(\varepsilon^{||\mathbf{x}||_1}n_1(\mathbf{x}) + \varepsilon^{||\mathbf{x}||_1+1}n_0(\mathbf{x}))$.
(iii) $\lim_{\varepsilon\downarrow 0}\mu^{\varepsilon} = \mu$, where $\tilde{\mu}$ is a probability distribution supported on the set of strategy profiles corresponding to optimal and quasi-optimal sufficient control sets whose entries are given by

$$\tilde{\mu}_{\mathbf{x}} = \begin{cases} n_0(\mathbf{x})/\tilde{K} & \text{if } \mathbf{x}=\mathbb{1}_{\mathcal{S}},\ \mathcal{S} \text{ optimal} \\ n_1(\mathbf{x})/\tilde{K} & \text{if } \mathbf{x}=\mathbb{1}_{\mathcal{S}},\ \mathcal{S} \text{ quasi-optimal} \\ 0 & \text{otherwise} \end{cases} \tag{38}$$

where

$$\tilde{K} = \sum_{\mathcal{S}\text{ optimal}} n_0(\mathbb{1}_{\mathcal{S}}) + \sum_{\mathcal{S}\text{ quasi-optimal}} n_1(\mathbb{1}_{\mathcal{S}})\,.$$

Proof: Point (i) follows from Theorem 2 (i) and (ii) and the observation that, for $\mathbf{x}\neq\mathbf{y}$ in $\mathcal{X}$, we have $\tilde{P}^{\varepsilon}_{\mathbf{x},\mathbf{y}}>0$ if and only if $P^{\varepsilon}_{\mathbf{x},\mathbf{y}}>0$.

We now prove (ii). Let D^{ε} in $\mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ be a diagonal matrix with entries $D^{\varepsilon}_{\mathbf{x}\mathbf{x}} = \alpha_{\varepsilon}(\mathbf{x})$. Then, the transition probability matrices P^{ε} and $\tilde{P}^{\varepsilon}$ of the Markov chains Z^{ε}_t and $\tilde{Z}^{\varepsilon}_t$, respectively, are related by the formula

$$P^{\varepsilon} = I - D^{\varepsilon} + D^{\varepsilon}\tilde{P}^{\varepsilon}\,,$$

which implies that $(P^{\varepsilon})'\mu = \mu$ if and only of $(\tilde{P}^{\varepsilon})'D^{\varepsilon}\mu = D^{\varepsilon}\mu$. This shows that the unique stationary distribution of Z^{ε}_t on $\mathcal{Z}$ satisfies

$$\tilde{\mu}_{\mathbf{x}} \propto \mu_{\mathbf{x}}\alpha_{\varepsilon}(\mathbf{x})\,,$$

thus proving (ii).

Finally, in order to prove (iii), we first make a number of observations. Let m be the minimum cardinality of a sufficient control set. First, consider an optimal sufficient control set $\mathcal{S}$ and let $\mathbf{x}=\mathbb{1}_{\mathcal{S}}$ be the associated strategy profile. Then, $\mathcal{S}$ is necessarily a minimal sufficient control set so that, by Proposition 4 (ii), $\mathbf{x}\in\mathcal{Z}_{\infty}$. Hence, $\mathbf{x}$ is an absorbing state of the Markov chain Z^{0}_t, so that $n_1(\mathbf{x})=0$ by Proposition 4 (iii). On the other hand, since for $0<\varepsilon\le 1$ the Markov chain Z^{ε}_t is ergodic on $\mathcal{Z}$, we necessarily have $n_0(\mathbf{x})\ge 1$ (since, if $n_0(\mathbf{x})=0$, then $\mathbf{x}$ would be an absorbing state for Z^{ε}_t). It then follows from (37) that

$$\tilde{\mu}^{\varepsilon}_{\mathbf{x}} = \frac{n_0(\mathbf{x})\varepsilon^{m+1}}{\tilde{K}_{\varepsilon}}\,, \qquad \forall\varepsilon\in(0,1]\,. \tag{39}$$

Now, consider a quasi-optimal sufficient control set $\mathcal{U}$ and let $\mathbf{x}=\mathbb{1}_{\mathcal{U}}$. Since (36) is satisfed for some optimal sufficient control set $\mathcal{S}$ and some player i in $\mathcal{V}\backslash\mathcal{S}$ such that $u_i(1,\mathbf{x}_{-i})\ge u_i(0,\mathbf{x}_{-i})$, we have that $n_1(\mathbf{x})>0$. Then, (37) yields

$$\tilde{\mu}_{\mathbf{x}} = \frac{n_1(\mathbf{x})\varepsilon^{m+1}+o(\varepsilon^{m+1})}{\tilde{K}_{\varepsilon}}\,, \qquad \text{as } \varepsilon\downarrow 0\,. \tag{40}$$

Finally, if we take $\mathbf{x}=\mathbb{1}_{\mathcal{T}}$ for any $\mathcal{T}$ that is neither optimal nor quasi-optimal, we necessarily have that $||\mathbf{x}||_i\ge m+1$ and that if $||\mathbf{x}||_i = m+1$, then $u_i(1,\mathbf{x}_{-i})<u_i(0,\mathbf{x}_{-i})$ for every i in $\mathcal{T}$. Hence, $n_i(\mathbf{x})=0$ and (37) implies that, in any case,

$$\tilde{\mu}_{\mathbf{x}} = \frac{o(\varepsilon^{m+1})}{\tilde{K}_{\varepsilon}}\,, \qquad \text{as } \varepsilon\downarrow 0\,. \tag{41}$$

The limit relation (38) then follows from (39)–(41). ■

In the following section, we shall refer to the Distributed Optimal Targeting (DOT) algorithm as an implementation of the modified Markov chain Z^{ε}_t.

VI. Numerical simulations

In this section, we briefly present some numerical simulations of the DOT algorithm proposed in Section V for the case of the network coordination game on a simple graph $\mathcal{G}$, as defined in Section III. Specifically, we apply our algorithm to

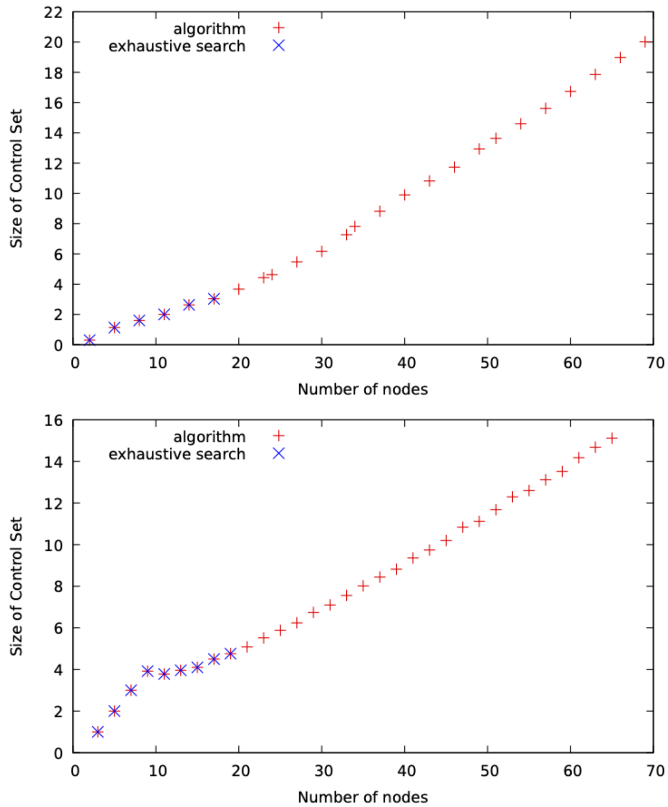


Fig. 7. Size of the optimal sufficient control sets for Erdös-Rényi random graphs $E(n, p)$ with $p = 0.4$ (left) and $p = \frac{4}{n} \log n$ (right).

determine optimal control sets for the majority game on Erdös-Rényi random graphs. The Erdös-Rényi random graph $E(n, p)$ is a random undirected graph with n nodes where undirected links between pairs of nodes are present with probability p in $[0, 1]$ independently from one another. We let the graph order n range up to 70 and consider two different regimes for the probability p. In the first regime, we set $p = 0.4$ to be a constant independent from the graph order n, thus leading to quite a densely connected graph. In contrast, in the second case, we set $p = \frac{4}{n} \log n$, leading to a more sparse graph that nevertheless remains connected with high probability as n grows large [32, Theorem 2.8.1].

First, we have run the DOT algorithm with parameter $\epsilon = 0.3$ for a number of steps proportional to the square of the graph order (precisely, $100n^2$) and the control set returned is the one of minimum cardinality encountered along the walk of the Markov chain. For small values of n, an explicit comparison with the optimal solution, obtained through exhaustive search, proves correctness of our approach for reasonable time horizons. Simulation results are reported in Figure 7.

Clearly, a fundamental parameter of our algorithm is the time horizon T over which the DOT algorithm is run. In Figure 8, for a specific sample of an Erdös-Rényi graph of order $n = 100$ we report the evolution in time of the cardinality of the smallest sufficient control set found so far by the DOT algorithm. In particular, we have run the simulation 100 times and plotted the average size of the smallest sufficient control set found so far as the blue curve, while the range between the minimal and maximal size of the current control set (among all the 100 simulations) is plotted as a light red interval. Notice that the average is very close to the minimum, thus suggesting small variability of the DOT algorithm.

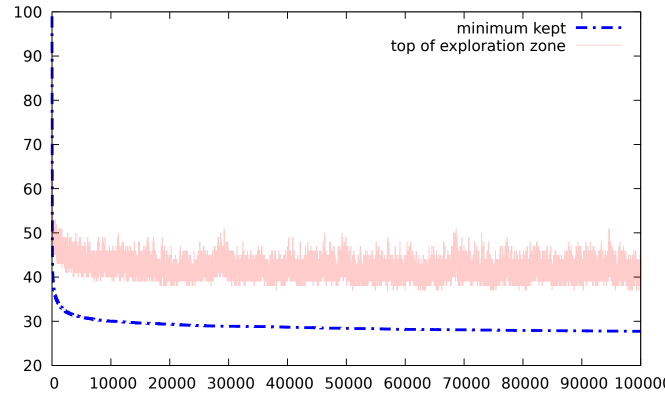


Fig. 8. For a specific sample of the Erdös-Rényi graph of order $n = 100$, and 100 independent simulation runs, the average cardinality of the smallest sufficient control set found so far by the algorithm is plotted in blue as a function of the number of transitions of the Markov chain ("time"). The light red interval corresponds to the range between the minimal and maximal size of the current control set (i.e., the current state of the Markov chain).

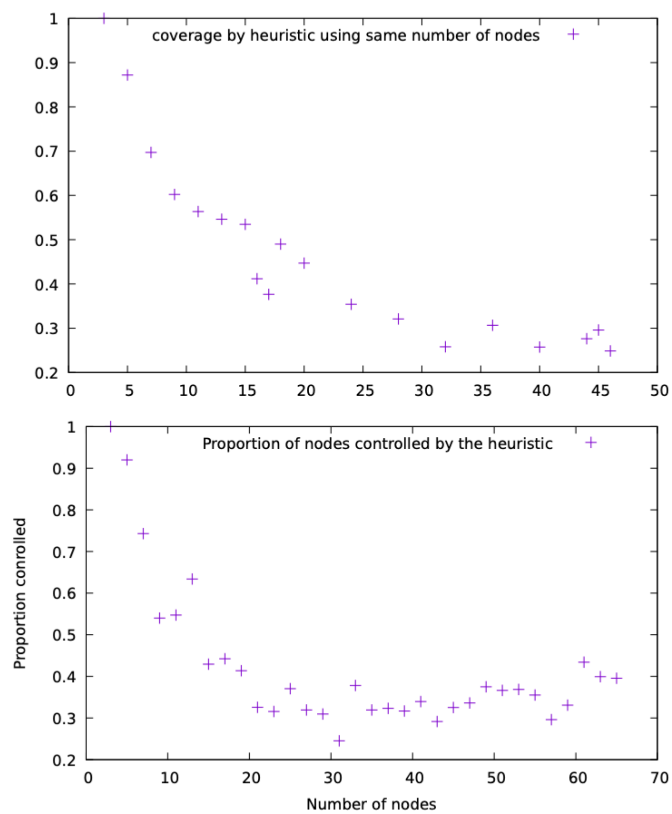


Fig. 9. Coverage obtained by taking the k highest degree node, with k the size of the set found by the Markov Chain algorithm for random graphs $E(n, p)$ with $p = 0.4$ (top) and $p = \frac{4}{n} \log n$ (bottom)

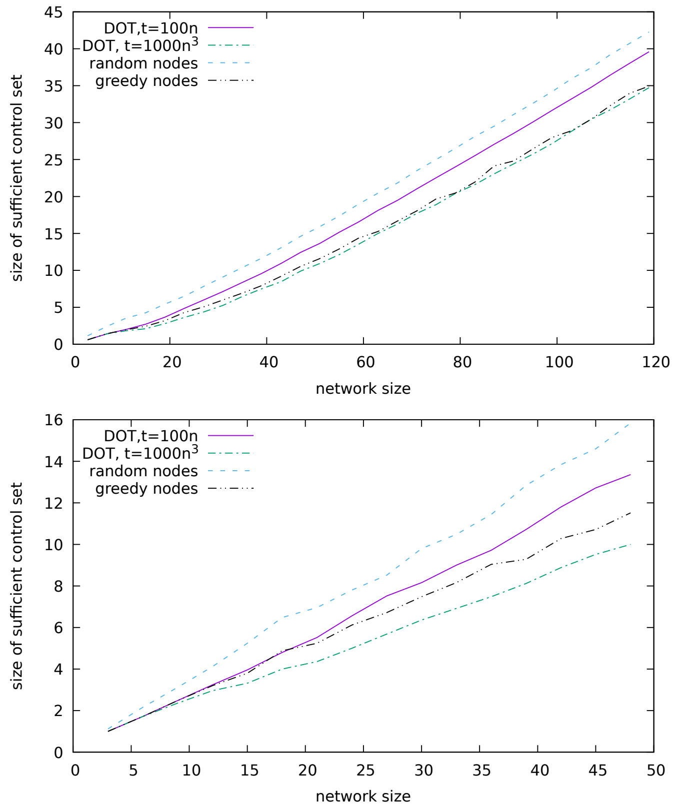


Fig. 10. For random graphs $E(n,p)$ with $p = 0.4$ (top) and $p = \frac{4}{n}\log n$ (bottom), comparison between the performance of the random-node, max-degree, and greedy heuristics with that of our Markov Chain algorithm with two different time horizons.

We now compare the performance of the DOT algorithm with other heuristic algorithms. First, we consider a naive heuristic algorithm selecting the maximal degree nodes, to be referred as the maximal degree heuristic (MDH) algorithm. In fact, the degree is the first and conceptually simplest measure of network centrality [33, p. 38]: e.g., in undirected graphs, the degree vector coincides, up to a scaling factor, with the stationary distribution of a random walk on the graph. In Figure 9 we make a comparison between the DOT and the MDH algorithms. Specifically, for each value of n, we consider a set of the highest degree nodes of the same cardinality as the one found by our DOT algorithm. We then plot the percentage of the graph nodes that would turn to 1 using that specific control set. When n is sufficiently large this percentage is around 30% and shows how the degree is not the right property to look at for the optimal targeting problem.

Finally, in Figure 10 we compare the performance of our DOT algorithm with that of the previously discussed MDH algorithm, as well as to that of a completely random node heuristic (RNH) algorithm —simply selecting nodes to be added to $\mathcal{S}$ uniformly at random until $\mathcal{S}$ becomes a sufficient control set— and of a "greedy heuristic" (GH) algorithm working as follows: Start with the all-zero action profile $\mathbf{x}^{(0)} = 0$, pick a node i_1 with highest degree in $\mathcal{G}$ and put $\mathbf{x}^{(0)}_{i_1} = 1$, then follow a monotone improvement path in the network coordination game until reaching an action profile $\mathbf{x}^{(1)}$; then pick a node i_2 with highest degree among those with $\mathbf{x}^{(1)}_{i_2} = 0$, put $\mathbf{x}^{(1)}_{i_2} = 1$ and follow a monotone improvement path in the network coordination until reaching an action profile $\mathbf{x}^{(2)}$; stop when the all-1 action profile $\mathbf{x}^{(k)} = \mathbb{1}$ ie reached and return the sufficient control set $\mathcal{S} = \{i_1, \ldots, i_k\}$. As Figure 10 shows, our DOT algorithm outperforms the RNH algorithm even for short simulation lengths, and does better than both the MDH and the GH algorithms (which have a similar performance) when the Markov chain is allowed to evolve over a long enough time horizon.

A comparison of the computational complexities of the aforementioned heuristics is in order. On the one hand, a straightforward implementation of either the RNH or the MDH algorithms requires order of n^3 computations since each time a node is added to $\mathcal{S}$ one needs to verify whether the obtained set $\mathcal{S}$ is a sufficient control set, which requires a number of iterations that is quadratic in n at worst. On the other hand, the complexity of the GH algorithm presented above is of the order of n^2, since the complexity of computing a monotone improvement path is of order n times the length of the path and the total length of the followed monotone improvement paths is less than n. In fact, our simulations show that the DOT algorithm displays a similar performance as such greedy heuristic when the length of the time horizon over which the Markov chain is allowed to run is of the order of the quadratic complexity of the GH algorithm and better performance when it is allowed to run for longer time horizons.

However, the "anytime" structure of our Markov chain based DOT algorithm provides an inherent advantage with respect to any of such heuristics, as it allows for the flexibility of getting a feasible solution consisting of a sufficient control set of non-increasing size at any time during its run, with the guarantee of convergence to the optimal one as the time horizon increases.

VII. Conclusion

In this paper, we have studied an optimal targeting problem for super-modular games with binary actions and finitely many players. The considered problem consists in the selection of a subset of players of minimum size such that, when the actions of these players are forced to a given common value, there exists an improvement path from every strategy profile to the pure strategy Nash equilibrium where all players play the same chosen action in the constrained super-modular game. Our main contributions consist in: (i) showing that this is an NP-complete problem; (ii) proposing a computationally simple randomized algorithm that provably selects an optimal solution with high probability.

Moreover, we have presented some numerical simulations results for the case of the majority game on Erdös-Rényi random graphs. We have compared the performance of our algorithm with that of an exhaustive search (for small problem sizes) and that of simple heuristics, including targeting players are those with the highest centrality in the graph. The first such comparison validates our theoretical results. The second comparison shows that the centrality-based heuristic performs as much as 70% worse than our algorithm in this problem, thus highlighting the relevance of our analysis.

The problem studied in this paper can be considered a particular instance of a control problem in a game-theoretic framework. Our results show how the structure of the game, in particular super-modularity, can be leveraged to get insight into the solution of the control problem.

Several directions for future research can be considered. In the context of super-modular games, two natural generalizations are of certain interest. The first one concerns the extension to games with non-binary action sets. Particularly interesting is the example of partnership games where each player's action is a nonnegative real number measuring the individual effort put in a certain common activity. A challenging step will be the design of a similar randomized algorithm to find optimal sufficient control sets in the continuous setting.

A second direction of research is that of considering more complex interventions, where the utilities of the controlled players are altered, rather than their actions directly forced to a desired one. A third direction would consist in studying the optimal sufficient control set problem for ensembles of large-scale random games. In particular, we believe that the local mean-field approaches such as those developed in [8], [34] could be a viable tool for developing results with probabilistic guarantees in sparse random graphical game settings.

As our approach strongly relies on super-modularity, extensions to more general classes of games is an even more challenging direction. In particular, network public goods games are another interesting family for which target intervention problems can be formulated and studied. We believe that, in spite of the difference with respect to super-modular games, randomized distributed algorithms, presumably based on different mechanisms, can play a role to solve optimal intervention problems.

References

[1] S. Durand, G. Como, and F. Fagnani, "Controlling network coordination games," in *IFAC-PapersOnLine*, vol. 53, no. 2, 2020, pp. 2862–2867.

[2] A. Galeotti, B. Golub, and S. Goyal, "Targeting interventions in networks," *Econometrica*, vol. 88, no. 6, pp. 2445–2471, 2020.

[3] M. O. Jackson and Y. Zenou, *Games on Networks*. Elsevier, 2015, vol. Handbook of Game Theory with Economic Applications, ch. 3, pp. 95–163.

[4] R. Amir, "Supermodularity and complementarity in economics: An elementary survey," *Southern Economic Journal*, vol. 71, no. 3, pp. 636–660, 2005.

[5] S. Morris, "Contagion," *The Review of Economic Studies*, vol. 67, no. 1, pp. 57–78, 2000.

[6] D. Acemoglu, A. Ozdaglar, and E. Yildiz, "Diffusion of innovations in social networks," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 2329–2334.

[7] M. Granovetter, "Threshold models of collective behavior," *American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.

[8] W. Rossi, G. Como, and F. Fagnani, "Threshold models of cascades in large-scale networks," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 2, pp. 158–172, 2019.

[9] M. Moharrami, V. Subramanian, M. Liu, and M. Lelarge, "Impact of community structure on cascades," in *Proceedings of the 2016 ACM Conference on Economics and Computation*, 2016, pp. 635–636.

[10] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of SIGKDD'03*, 2003, pp. 137–146.

[11] M. Akbarpour, S. Malladi, and A. Saberi, "Diffusion, seeding, and the value of network information," in *Proceedings of the 2018 ACM Conference on Economics and Computation*, 2018, p. 641.

[12] A. Galeotti and S. Goyal, "Influencing the influencers: a theory of strategic diffusion," *RAND Journal of Economics*, vol. 40, no. 3, pp. 509–532, 2009.

[13] C. Ballester, A. Calvó-Armengol, and Y. Zenou, "Who's who in networks. wanted: The key player," *Econometrica*, vol. 74, no. 5, pp. 1403–1417, 2006.

[14] O. Candogan, K. Bimpikis, and A. Ozdaglar, "Optimal pricing in networks with externalities," *Operations Research*, vol. 60, pp. 883–905, 2012.

[15] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, pp. 167–173, 2011.

[16] J. Gao, Y.-Y. Liu, R. D'Souza, and A.-L. Barabási, "Target control of complex networks," *Nature Communications*, vol. 5, no. 1, p. 5415, 2014.

[17] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, 2014.

[18] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione, "Binary opinion dynamics with stubborn agents," *ACM Transactions on Economics and Computation*, vol. 1, no. 4, pp. 19–30, 2013.

[19] L. Vassio, F. Fagnani, P. Frasca, and A. Ozdaglar, "Message passing optimization of harmonic influence centrality," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 109–120, 2014.

[20] M. Grabisch, A. Mandel, A. Rusinowska, and E. Tanimura, "Strategic influence in social networks," *Mathematics of Operations Research*, vol. 43, no. 1, pp. 29–50, 2018.

[21] D. Acemoğlu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 1–27, 2013.

[22] G. Como and F. Fagnani, "From local averaging to emergent global behaviors: The fundamental role of network interconnections," *Systems and Control Letters*, vol. 95, pp. 70–76, 2016.

[23] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis, "An efficient curing policy for epidemics on graphs," *IEEE Transactions on Network Science and Engineering*, vol. 1, no. 2, pp. 67–75, 2014.

[24] ——, "When is a network epidemic hard to eliminate?" *Mathematics of Operations Research*, vol. 42, no. 1-14, 2016.

[25] P. Milgrom and J. Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica*, vol. 58, no. 6, pp. 1255–1277, 1990.

[26] D. M. Topkins, "Equilibrium points in nonzero-sum n-person submodular games," *SIAM Journal on Control and Optimization*, vol. 17, no. 6, pp. 773–787, 1979.

[27] X. Vives, "Nash equilibrium with strategic complementarities," *Journal of Mathematical Economics*, vol. 19, pp. 305–321, 1990.

[28] D. M. Topkins, *Supermodularity and Complementarity*. Princeton University Press, 1998.

[29] P. Milgrom and C. Shannon, "Monotone comparative statics," *Econometrica*, vol. 62, no. 1, pp. 157–180, 1994.

[30] M. Sipser, *Introduction to the theory of computation*, 3rd ed. Cengage Learning, 2012.

[31] F. Pin, A. Busic, and B. Gaujal, "Acceleration of perfect sampling by skipping events," in *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools*, 2011, pp. 207–216.

[32] R. Durrett, *Random Graph Dynamics*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2006.

[33] M. Jackson, *Social and Economic Networks*. Princeton, New Jersey: Princeton University Press, 2008.

[34] C. Ravazzi, G. Como, M. Garetto, E. Leonardi, and A. Tarable, "Asynchronous semi-anonymous dynamics over large-scale networks," *https://arxiv.org/abs/2102.03840*, 2021.

Giacomo Como received the B.Sc., M.S., and Ph.D. degrees in Applied Mathematics from Politecnico di Torino, in 2002, 2004, and 2008, respectively.

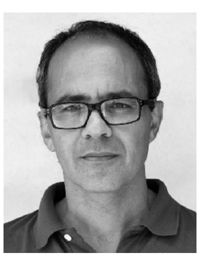
He has been a Professor at the Department of Mathematical Sciences, Politecnico di Torino, Italy, since 2016, and a Senior Lecturer at Lund University, Sweden, since 2011. He was a Visiting Assistant in Research at Yale University in 2006–2007 and a Postdoctoral Associate at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, from 2008 to 2011.

Prof. Como currently serves as Associate Editor of the *IEEE Transactions on Network Science and Engineering* and of the *IEEE Transactions on Control of Network Systems* and as chair of the IEEE-CSS Technical Committee on Networks and Communications. He was the IPC chair of the IFAC Workshop NecSys'15 and a semiplenary speaker at the International Symposium MTNS'16. He is recipient of the 2015 George S. Axelby Outstanding Paper Award. His research interests are in dynamics, information, and control in network systems with applications to cyber-physical systems, and infrastructure, social, and economic networks.

Stéphane Durand received his Master of Research in 2016 from the École National Supérieure de Lyon, France, and his Ph.D. in December 2018 from the University of Grenoble, France.

He has been a Postdoctoral Researcher at the Department of Mathematical Sciences, Politecnico di Torino, Italy, since 2019.

Dr. Durand's research interests are in game theory, network games, and Markov models.



Fabio Fagnani received a Laurea degree in Mathematics from the University of Pisa and the Scuola Normale Superiore, Pisa, Italy, in 1986 and a PhD degree in Mathematics from the University of Groningen, Groningen, The Netherlands, in 1991.

He was an Assistant Professor of Mathematical Analysis at the Scuola Normale Superiore from 1991 to 1998. In 1997, he was a Visiting Professor at the Massachusetts Institute of Technology (MIT), Cambridge, MA. Since 1998, he has been with Politecnico of Torino, where he is currently (since 2002) a Full Professor of Mathematical Analysis.

Prof. Fagnani acted as Coordinator of the PhD program in Mathematics for Engineering Sciences at Politecnico di Torino from 2006 to 2012. From 2012 to 2019, he served as the Head of the Department of Mathematical Sciences, Politecnico di Torino. His current research topics are cooperative algorithms and dynamical systems over graphs, inferential distributed algorithms, and opinion dynamics. He is an Associate Editor of the *IEEE Transactions on Automatic Control* and served in the same role for the *IEEE Transactions on Network Science and Engineering* and of the *IEEE Transactions on Control of Network Systems*.