Optimal targeting in supermodular games∗

Stephane Durand†, Giacomo Como‡ and Fabio Fagnani§

Department of Mathematical Sciences “G.L. Lagrange”, Politecnico
di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy.

May 12, 2020

1 Introduction

In a binary (0 – 1) coordination game over a graph, where initially all players
are in the Nash equilibrium 0, what is the minimum number of players that if
forced to 1 will force the system to the Nash equilibrium of all 1’s? This paper
deals with such a problem for the more general class of super-modular games
with binary actions, namely those games that exhibit the so-called increasing
difference property. Our contribution is twofold: we show that the problem is
NP-complete and we propose the design of an iterative algorithm for an efficient
solution.

The problem considered can be framed in the more general setting of study-
ing minimal interventions strategies needed to make a multi-agent system gov-
erned by agents’ myopic utility maximization, to drive from a Nash equilibrium
to a desired another one. In game theory, typically, interventions have been
modeled as perturbations of the utility functions (e.g. taxes and prices in eco-
nomic models or tolls in transportation systems). Here we instead take a dif-
ferent viewpoint: that of individuating a subset of nodes (hopefully small) that
if suitably controlled will lead the entire system to the desired configuration.
The minimum cardinality of this set can also be interpreted as a measure of
resilience of the system: the larger it is, the more difficult is for an external
shock to destabilize it.

The problem of determining the best set of nodes to exert the most effective
control in a networked system has recently appeared in other contexts: for

∗This work was partially supported by MIUR grant Dipartimenti di Eccellenza 2018–2022
[CUP: E11G18000350001], the Swedish Research Council, and by the Compagnia di San Paolo.
†Email: stephane.durand@polito.it
‡Email: giacomo.como@polito.it. G. Como is also with the Department of Automatic
Control, Lund University, BOX 118, SE-22100, Lund, Sweden.
§Email: fabio.fagnani@polito.it
instance in [13] and [11] authors study the problem of the optimal position of stubborn influencers in linear opinion dynamics.

Super-modular games have received a great attention in the recent years as the basic way to model strategic complementarity effects [3]. Its variegated applications include modeling of social and economic behaviors like adopt a new technology, participate in an event, provide a public good effort. They 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 higher equilibrium is natural and relevant in all these applicative contexts.

The binary coordination game is analyzed in detail in [7] where the key concept of cohesiveness of a set of players is introduced and then used in characterizing all NE’s. Moreover, the question if an initial seed of influenced players (that maintain action 1 in all circumstances) is capable of propagating to the all network is addressed in the same paper and an equivalent characterization of this spreading phenomenon is also expressed in terms of cohesiveness. This contagion phenomenon is exactly what we want to analyze: subset of nodes from which propagation is successful will be called sufficient control sets and our goal is to find such sets of minimum possible cardinality.

The condition proposed in [7] is computationally quite demanding and can not be used to directly solve our optimization problem. Indeed, even to determine if a single set is a sufficient control set, it requires a number of check growing exponentially in the cardinality of the complement of such set.

The complementary problem of understanding, given an integer k, what is the maximum possible spreading of the state 1, starting from an initial seed of k influenced players, was studied in a seminal paper by [4]. While their problem 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 is that, in their setting, [4] consider players equipped with random independent activation thresholds and take as functional to be optimized the average size of the maximum spreading. They prove that such functional is sub-modular and then they design a greedy algorithm for obtaining sub-optimal solutions. The randomness that they introduce is actually crucial in their approach, as the functional considered would not be sub-modular for deterministic choices of thresholds. In this way it can not directly applied to the deterministic scenario considered in this paper.

After proving that the proposed problem is NP-complete (reducing it to the 3-SAT problem), we design an iterative search randomized algorithm with provable properties of convergence towards sufficient control sets of minimum cardinality. The core of the algorithm is a time-reversible Markov chains 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.

We conclude this introduction with a brief outline of the paper. In the final part of this section we report some basic notation used throughout the paper. Section 2 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 [1, 2]) and we give an equivalent (but more operative) characterization of sufficient control sets. Section 3 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 4 we present and analyze a distributed algorithm to find optimal sufficient control sets and we present some simulation results. Finally, conclusive Section 5 ends the paper.

1.1 Notation

Vectors are indicated in bold-face letters $\mathbf{x}, \mathbf{y}, \mathbf{z}$. For $\mathbf{x}, \mathbf{y}$ are two vectors of the same dimension, the notation $\mathbf{x} \leq \mathbf{y}$ indicates that $\mathbf{x}$ is lower or equal component-wise than $\mathbf{y}$. We define as usual the binary vectors $\delta_i: (\delta_i)_j = 1$ and $(\delta_i)_j = 0$ for every $j \neq i$. If $S \subseteq \{1, \ldots, n\}$, we put $1_S = \sum_{i \in S} \delta_i$. Every $\mathbf{x} \in \{0, 1\}^n$ can be written as $\mathbf{x} = 1_S$ for some $S \subseteq \{1, \ldots, n\}$. We call such a subset $S$ the support of $\mathbf{x}$ and we denote it $S_\mathbf{x}$. We use the notation 0 and 1 to denote, respectively, the vector of all 0’s and the vector of all 1’s of any possible dimension.

2 Controlled super-modular games

We consider strategic form games with a finite set of players $\mathcal{V} = \{1, \ldots, n\}$ whereby each player is endowed with a binary action set $\mathcal{A} = \{0, 1\}$. Let $\mathcal{X} = \mathcal{A}^n$ denote the strategy profile space, whose elements $\mathbf{x} \in \mathcal{X}$ will be referred to as strategy profiles or configurations. We shall endow the configuration space $\mathcal{X}$ with the partial order

$$\mathbf{x} \leq \mathbf{y} \iff x_i \leq y_i, \quad \forall i \in \mathcal{V}. \quad (1)$$

The $i$-th entry $x_i$ of a strategy profile $\mathbf{x}$ in $\mathcal{X}$ represents the action played by player $i$. 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 $i$ and we consequently write $\mathbf{x} = (x_i, \mathbf{x}_{-i})$.

Each player $i$ is endowed with a utility function $u_i : \mathcal{X} \to \mathbb{R}$ and, following the aforementioned convention, $u_i(\mathbf{x}) = u_i(x_i, \mathbf{x}_{-i})$ indicates the utility of player $i$ when she plays action $x_i$ while the rest of the players are playing $\mathbf{x}_{-i}$. The best response function for a player $i$ in $\mathcal{V}$ is then defined as

$$B_i(\mathbf{x}) = \arg\max_{a \in \mathcal{A}} u_i(a, \mathbf{x}_{-i}),$$

while the set of Nash equilibria will be denoted by

$$\mathcal{N} = \{ \mathbf{x} \in \mathcal{X} \mid x_i \in B_i(\mathbf{x}) \forall i \in \mathcal{V} \}.$$

Throughout the paper, we shall consider games of this type, indicated as usual as triples $(\mathcal{V}, \mathcal{A}, \{u_i\})$, and satisfying the following assumption.
Assumption 1. For every player \( i \) in \( V \) and every two strategy profiles \( x, y \) in \( X \) such that \( x_{-i} \succeq y_{-i} \),

\[
    u_i(1, x_{-i}) - u_i(0, x_{-i}) \geq u_i(1, y_{-i}) - u_i(0, y_{-i}),
\]

(2)

Assumption 1 is known as the increasing difference property [5] and for finite games it is known to be equivalent to supermodularity [9, 12, 10]. In the economic language, these are also referred to as games of strategic complements [6]. For games with binary action sets, as in our case, it amounts to say 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 \( x_{-i} \) of all the other players.

Network coordination games form a notable example of super-modular games and will be formally reviewed later on in the paper.

A standard result for super-modular games ensures that their set of Nash equilibria is always nonempty and contains, in particular, a minimal and a maximal element (with respect to the partial order (1)). Without loss of generality, we can assume that those extremal elements are the all-0 configuration \( \emptyset \) and, respectively, the all-1 configuration \( \emptyset \). The presence of players that maintain a strict preference independently on the actions chosen by the other players can be easily integrated in said other players’ utility function.

Definition 1. Given a game \( (V, A, \{u_i\}) \), a sequence of configurations \( x^k \in X \), for \( k = 0, \ldots, m \) is called an improvement path from the set \( S \subseteq V \) to the set \( T \subseteq V \) for the game if

1. \( x^0 = 1_S \), \( x^m = 1_T \)

2. for every \( k = 1, \ldots, m - 1 \) there exists \( i_k \in V \setminus S \) such that

   - \( x_{i_k}^{k+1} = x_{i_k}^k \) and \( x_{i_k}^{k+1} \neq x_{i_k}^k \)
   - \( u_{i_k}(x^{k+1}) \geq u_{i_k}(x^k) \)

In this paper we study the problem of finding sets \( S \subseteq V \) of minimal cardinality for which there exists an improvement path from \( S \) to the whole of \( V \). This is formalized by the following definition.

Definition 2 (Sufficient control set). For a super-modular game \( (V, A, \{u_i\}) \),

- \( S \subseteq V \) is a sufficient control set if there exists an improvement path from \( S \) to \( V \).

- A sufficient control set is optimal if there exists no sufficient control set of strictly smaller cardinality.

Notice that there are always sufficient control sets, as the whole set of players \( V \) is a trivial one. Our objective is to find optimal sufficient 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.

4
Definition 3 (Monotone Improvement path). Let $\mathcal{S} \subseteq \mathcal{V}$. Given a game $(\mathcal{V}, \mathcal{A}, \{u_i\})$, an improvement path $x^k \in \mathcal{X}$, for $k = 0, \ldots, m$, from the set $\mathcal{S} \subseteq \mathcal{V}$ to the set $\mathcal{T} \subseteq \mathcal{V}$ is called monotone if there exists a sequence of distinct nodes $i_k \in \mathcal{T} \setminus \mathcal{S}$ for $k = 0, \ldots, m - 1$ such that $x^{k+1} = x^k + \delta_{i_k}$ for $k = 0, \ldots, m - 1$.

Remark 1. In reference to the above definition, notice that a monotone improvement path is completely specified by the sequence of nodes $i_k \in \mathcal{T} \setminus \mathcal{S}$. Notice that $\mathcal{T} \setminus \mathcal{S} = \{i_1, \ldots, i_m\}$ and thus $m = |\mathcal{T} \setminus \mathcal{S}|$.

The following result formalizes our original claim.

Theorem 1. In a super-modular game $(\mathcal{V}, \mathcal{A}, \{u_i\})$, if $\mathcal{S} \subseteq \mathcal{V}$ is a sufficient control set, then there exists a monotone improvement path from $\mathcal{S}$ to $\mathcal{V}$.

Proof. If $\mathcal{S}$ is a sufficient control set, then there exists an improvement path $y^0, y^1, \ldots, y^T$ in $\mathcal{A}_\mathcal{S}$ from $\mathcal{S}$ to $\mathcal{V}$. For every player $i$ in $\mathcal{V} \setminus \mathcal{S}$, define

$$k(i) = \min\{k = 1, \ldots, T \mid y^k = y^{k-1} + \delta_i\}$$

that is the first time that player $i$ changes her action from 0 to 1 along the path $(y^k)$. Now, let $m = |\mathcal{V}| - |\mathcal{S}|$ and order the players in $\mathcal{V} \setminus \mathcal{S}$ as $i_1, \ldots, i_m$ in such a way that $k(i_1) < k(i_2) < \cdots < k(i_m)$. Then, for every $h = 0, 1, \ldots, m$, define

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

and notice that $x^{h-1} \leq y^{k(i_h)-1}$. Using the increasing difference property we now obtain that

$$u_{i_h}(x^h) - u_{i_h}(x^{h-1}) \geq u_{i_h}(y^{k(i_h)}) - u_{i_h}(y^{k(i_h)-1}) \geq 0,$$

for every $h = 1, \ldots, m$. This shows that $x^0, x^1, \ldots, x^m$ is an improvement path from $\mathcal{S}$ to $\mathcal{V}$. By construction, this improvement path is also monotone, thus proving the claim.

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

Proposition 1 (monotonicity for inclusion). A super-set of a sufficient control set is a sufficient control set.

Proof. Assume that $\mathcal{S}$ is a sufficient control set and let $\mathcal{S}' \supseteq \mathcal{S}$. Because of Theorem 1 there exists a monotone improvement path $x^k$ from $\mathcal{S}$ to $\mathcal{V}$ with associated sequence of points $(i_k)$ for $k = 1, \ldots, m = n - |\mathcal{S}|$ such that $x^{k+1} = x^k + \delta_{i_k}$ for each $k$. Consider the subsequence of points $i_{k_1}, i_{k_2}, \ldots, i_{k_m'}$ that are in $\mathcal{V} \setminus \mathcal{S}'$ and put $y^h = \max\{1_\mathcal{S}, x^{k_h}\}$. By construction, we have that $y^h \geq x^{k_{h+1}-1}$. By the increasing difference property 2 and the fact that $x^h$ is a monotone improvement path from $\mathcal{S}$ it follows that, for every $h$, putting $i = i_{k_{h+1}}$,

$$u_i(y^{h+1}) - u_i(y^h) = u_i(1, y^h) - u_i(0, y^h) \geq u_i(1, x^{k_{h+1}-1}) - u_i(0, x^{k_{h+1}-1}) = u_i(x^{k_{h+1}}) - u_i(x^{k_{h+1}-1}) \geq 0$$
Remark 2. The definition of sufficient control set can be reinterpreted in terms of the asynchronous best response dynamics. Given subset $S \subseteq V$, consider the Markov chain $X^t$ on the configuration space $\mathcal{X}$ such that $X^0 = 1_S$ and whose transitions are described as follows. At every discrete time, a player, among those in $V \setminus S$, is chosen uniformly at random and it updates its played action choosing randomly among elements that maximize their current best response. Notice that the existence of an improvement path from $S$ to $V$ is equivalent to say that $X^t$ will reach the absorbing state $1$ with positive probability. Actually more is true. Consider any node $x \in \mathcal{X}$ that is reachable by the Markov chain and notice that necessarily $x = 1_S$ for some subset $S' \supseteq S$. If $S$ is a sufficient control set, it follows from Proposition 1 that also $S'$ is sufficient. This implies that there exists a monotone improvement path from $S'$ to $V$ and thus $X^t$ will also reach $1$ from $x$ with positive probability. This argument proves that $S \subseteq V$ is a sufficient control set if and only if the corresponding Markov chain $X^t$ will be absorbed in $1$ in finite time with probability $1$.

2.1 The majority game on a network

Coordination games over a network form a notable example of super-modular games. In this subsection we present and discuss a famous instance of them, the majority game. It will constitute the key example considered throughout the paper.

We assume players in $V$ to be connected by a simple undirected graph $G = (V, E)$ (e.g. $E \subseteq V \times V$ is such that $(i, i) \notin E$ for any $i$ and $(i, j) \in E$ iff $(j, i) \in E$). Recalling that $A = \{0, 1\}$ and $\mathcal{X} = A^n$ we define the majority game on $G$ as the game where each player $i \in V$ has utility $\lambda^i : \mathcal{X} \to \mathbb{R}$ given by

$$\lambda^i(x) = |\{j \in N_i | x_j = x_i\}|$$

where $N_i$ is the neighborhood of player $i$ in $G$. In other words, $\lambda^i(x)$ is the number of neighbors of $i$ with which $i$ is in agreement.

The set $\mathcal{N}$ of Nash equilibria depends on the topology of the graph $G$, note however that in this case $0$ and $1$ are always Nash equilibria.

Sufficient control sets in this case can be equivalently formulated in terms of the cohesiveness concept introduced in [7] and recalled below. A subset $C \subseteq V$ is called $q$-cohesive in $G$ if for every $i \in C$, it holds that

$$|\{j \in C | (i, j) \in E\}| \geq q|N_i|$$

Considerations in [7] show that $S$ is a sufficient control set if and only if $V \setminus S$ does not contain any subset that is strictly more than $1/2$-cohesive.

Below we present examples of sufficient control sets for the majority game and specific topologies of the network.

Example 1. • Consider a clique graph with $n$ nodes. Then, the sets of size $\left\lfloor \frac{n}{2} \right\rfloor - 1$ are exactly the optimal sufficient control sets.
• Consider a graph where every node has degree at most 2. Then, any set of nodes consisting of just one node per connected component is an optimal sufficient control set. An example is depicted in Figure 2.

Figure 2: An optimal sufficient control set for a graph with nodes of degree at most 2

• Consider a tree. For such graph, the set of the leaves is always a sufficient control set. This can be shown by depicting the tree as a rooted tree and constructing recursively, starting from leaves, a monotone improvement path to the whole of the graph. In general, this is not minimal. The argument above shows that also the set of nodes that are neighbors of leaves is a sufficient control set, typically of smaller cardinality than the previous one. An example is reported in Figure 3 where we can see in green the set of leaves and in red the leaves’ neighbors that is also sufficient and actually optimal for this particular tree (though not in general).

Figure 3: Two examples of sufficient control sets for a tree: the one in red is optimal

• A d-dimensional grid graph is a graph in which nodes can be described by an integer vector of size d, such that two nodes in the graphs are neighbors
if and only if the difference between their vector is a unit vector. They can be represented by the points of integer coordinates in a $d$ dimensional space, with the neighbors of any one being the nodes closest to it. 

with $\mathcal{V} = \{0, \ldots, k-1\}^d$ and $\mathcal{E} = \{(a, b) \in \mathcal{V} \times \mathcal{V} \mid \sum h |a_h - b_h| = 1\}$. Put $S_l = \{(a_1, \ldots, a_d) \in \mathcal{V} \mid \sum a_i = l\}$. We claim that $S_{k-1}$ is a sufficient control set. To see this, notice that any $a \in S_l$ has exactly $d$ neighbors in $S_{l+1}$ if $l < k-1$. Similarly, any $a \in S_l$ has exactly $d$ neighbors in $S_{l-1}$ if $l > k-1$. Considering that the degree of every node is at most $2d$ in $G$, a straightforward induction argument then allows to construct a monotone improvement path from $S_{k-1}$ to the whole of $\mathcal{V}$. It can be checked directly that this control set is optimal for $d = 1$ and $d = 2$, while is not for $d \geq 3$.

Figure 4: Example for $d = 2$ and $k = 5$

The examples considered above show that optimal sufficient control set can exhibit different relative sizes. In networks like cliques, 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 Erdos-Renji graphs. This conjecture is corroborated by numerical simulations presented in Section... On the contrary, for more loosely connected graphs (trees, grids), the size of optimal sufficient control sets scales as a negligible fraction of the size $n$.

2.2 Coordination game with heterogeneous thresholds

We here analyze in detail an example of a different coordination game where no network structure is present the heterogeneity instead comes from the utility functions of the players. For this game, we give a full theoretical characterization of the sufficient control sets.

We consider a set $\mathcal{V} = \{1, \ldots, n\}$ of players and we assume each player in $i \in \mathcal{V}$ to have a utility function $\lambda_i : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} = \mathcal{A}^\mathcal{V}$, given by

$$
\lambda_i(x) = |\{j \in \mathcal{V} \mid x_j = x_i\}| + c_ix_i
$$

where $c_i \in \mathbb{R}$. The term $c_ix_i$, when $c_i \neq 0$, introduces a bias in the set of actions. The strategy $1$ is preferable if $c_i > 0$, while $0$ is preferable if instead
We first prove that

\[ \text{Proof.} \]

Then the set \( \bar{S} \) considers any ordering of the nodes in \( \mathcal{V} \). Let \( \theta \) be the values of the thresholds of the various players and denote by \( \mathcal{V}_i \) the set of players having threshold \( \theta \leq \theta_i \) and \( p_i = |\mathcal{V}_i|/n \) the corresponding fraction. Notice that \( F(\theta_i) = p_{i-1} \) for \( i > 1 \) and \( F(\theta_1) = 0 \). By the way, \( M \) was defined as the maximum of \( \theta \) such that \( \bar{S} \cup \mathcal{V}_1 \) is the first monotone improvement path from \( \bar{S} \cup \mathcal{V}_1 \) to \( \bar{S} \cup \mathcal{V}_2 \). Iterating this, we eventually obtain a monotone improvement path to the whole of \( \mathcal{V} \).

To prove that \( \bar{S} \) is optimal, we show that any subset \( S \subseteq \mathcal{V} \) such that \( |S| < |\mathcal{V}| \) can not be a sufficient control set. This yields \( |S|/n < M \). Using the definition of \( M \), it shows that there exists \( \bar{\theta} \in [0, 1] \) such that \( \bar{\theta} - F(\bar{\theta}) > |S|/n \) or also \( \bar{\theta} > |S|/n + F(\bar{\theta}) \). In particular, \( F(\bar{\theta}) < 1 \): there are players in the population having threshold \( \bar{\theta} > \bar{\theta} \).

Reasoning by contradiction, if \( S \) was a sufficient control set, it would be valid because of Theorem 1. Consequently, there would exist a monotone improvement path \( x^k \) for \( k = 0, \ldots, m \) from \( S \) to \( \mathcal{V} \). Let \( i_1, i_2, \ldots \) be the corresponding sequence of players such that \( x_{i_k} = x_{i_{k-1}} + \delta_{x_{i_k}} \) and let \( \bar{k} \) be such that \( i_{\bar{k}} \) is the first player in the sequence whose threshold is not below \( \bar{\theta} \). Since, by construction, \( p(x_{i_{\bar{k}-1}}) \leq |S|/n + F(\bar{\theta}) < \bar{\theta} \), we deduce from (3) that \( B(x^k_{i_{\bar{k}}}) = \{0\} \) contradicting that \( x^k \) was a monotone improvement path. Proof is complete. \( \Box \)
3 Complexity of finding a valid control set

In this section, we study the complexity of finding valid control sets and prove that it is an NP-complete problem.

Formally, given a binary super-modular game and a positive integer \( n \) we define \( VCS \) to be the logical proposition "there exists a valid control set of size less then or equal to \( n + 1 \) for the game".

**Theorem 2.** The problem \( VCS \) is NP-complete.

In order to prove Theorem 2, we will first show that \( VCS \) is NP and then that it is NP-hard.

**Lemma 1.** The problem \( VCS \) is NP.

**Proof.** Given a finite binary-action supermodular game and a subset of players \( C \subseteq V \), checking if \( C \) is a valid control set can be done in a time growing proportionally to the square of the size of the complement of the set, simply by trying to switch other players’ strategies to one according to best response dynamics. If the set is valid, at least one players at 0 can switch, and the new set, being a superset of the previous one, is still valid. The algorithm of simply testing all player still at 0 until either none remain, or none change their choice is quadratic in worst case. Thus the problem is NP.

We will now prove that \( VCS \) is NP-hard by showing that the 3-SAT problem [8, Ch. 7.2] can be reduced to it in polynomial time to a particular instance of \( VCS \).

Towards this goal, let \( I = (X, C) \) be an instance of the 3-SAT problem, consisting of a set of variables \( X = \{x_1, x_2, \ldots, x_n\} \) and clauses \( C = \{c_1, c_2, \ldots, c_m\} \), such that in every every clause in \( C \) exactly three, possibly negated, variables from \( X \) appear. Then, we associate to \( I \) a simple graph \( G_I = (V_I, E_I) \) of order \( |V_I| = 2n + 5m + 2 \) and size \( |E_I| = n + 8m + 1 \) as follows. The node set \( V_I \) is the union of the following six disjoint sets of nodes:

- A set \( W = \{w_1, w_2, \ldots, w_m\} \), whose elements correspond each to a clause in \( C \);

- A set \( Y = \{y_1, y_2, \ldots, y_n\} \), whose elements correspond each to a variable in \( I \), with the interpretation that \( y_i \) encodes if \( x_i \) is true;

- A set \( \bar{Y} = \{\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n\} \), whose elements correspond each to a variable in \( I \), with the interpretation that \( \bar{y}_i \) encodes 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 \( E_I \) only connect pairs of nodes belonging to different sets and in particular:
A node \( w_j \) in \( C \) is connected to a node \( y_i \) in \( Y \) if and only if the variable \( x_i \) appears in the clause \( c_j \); 

(1b) A node \( w_j \) in \( C \) is connected to a node \( \overline{y}_i \) in \( \overline{Y} \) if and only if the variable \( \overline{x}_i \) appears in the clause \( c_j \); 

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

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

(4) For every \( i = 1, \ldots, n \), node \( y_i \) is connected to node \( \overline{y}_i \).

There is a total of \( 3m \) links of type 1 or \( 1b \) and \( 3m \) links of type 2, \( 2m + 1 \) links of type 3 and \( n \) links of type 4. On the other hand, nodes in \( L \) and \( M \) all have degree 1, nodes in \( W \) all have degree 4, node \( z \) has degree \( 2m + 1 \), while the degree of a node \( y_i \in Y \) (respectively \( \overline{y}_i \in \overline{Y} \)) 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 \( G_I \), whereby each player in \( V_I \) 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 valid control set of size less than or equal to \( n + 1 \) 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 2.** Let \( I = (X, C) \) be an instance of the 3-SAT problem, and let \( G_I = (V_I, E_I) \) be the simple graph defined above. If \( I \) is satisfiable with a solution \( x^* \in \{T, F\}^n \), then 

\[
S = \{z\} \cup \{y_i : x^*_i = 1\} \cup \{\overline{y}_i : x^*_i = 0\}
\]

is a valid control set of size \( n + 1 \) for the majority game on \( G_I \).

**Proof.** Since \( I \) is satisfied by \( x^* \), for every clause \( c_j \) in \( C \) there exists \( i \in \{1, \ldots, n\} \) 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 \( G_I \), all clause-related nodes in \( W \) have at least one neighbor in \((Y \cup \overline{Y}) \cap S \). Since they are all connected to \( z \in S \) also, and
have all degree 4 in \( \mathcal{G} \), this implies that there exists a monotone improvement path from \( S \) to \( S \cup \mathcal{W} \).

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 \( S \), it necessarily has one neighbor in \( S \) and \( m_i \) neighbors in \( \mathcal{W} \) (those corresponding to the clauses it belongs to).

Since its degree in \( \mathcal{G} \) is exactly \( 2m_i + 2 \), this implies that \( S \cup \mathcal{W} \cup \mathcal{Y} \) can be reached by a monotone improvement path from \( S \cup \mathcal{W} \), hence from \( S \). Analogously, one proves that \( S \cup \mathcal{W} \cup \mathcal{Y} \cup \overline{V} \) can be reached by a monotone improvement path from \( S \).

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

We will now show that the converse of Lemma 2 holds true.

**Lemma 3.** 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 valid control set \( S \) of size \( n + 1 \) for the majority game on \( \mathcal{G}_I \), then \( I \) is solvable.

**Proof.** We will first show that there exists a valid control set \( S' \) of the same size \( n + 1 \) containing \( z \) and exactly one node between \( y_i \) and \( \overline{y}_i \) for \( 1 \leq i \leq n \). We argue as follows. First, notice that, for every \( i = 1, \ldots, n \), at least one node among \( y_i, \overline{y}_i \), and the leaves in \( \mathcal{L} \) connected to them must be in \( S \) for, otherwise, it is easy to check that no improvement path would ever be able to reach the pair \( \{y_i, \overline{y}_i\} \). Similarly, at least one element among \( z \) and the leaves in \( \mathcal{M} \) must be in \( S \).

In case when neither \( y_i \) nor \( \overline{y}_i \) belong to \( S \), removing the leaf connected to them that is in \( S \) and adding its sole neighbor (either \( y_i \) or \( \overline{y}_i \)) maintains the control set validity and preserve its size. We construct \( S' \) 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.

Now observe that, since \( S' \) is a valid control set containing no leaves in \( \mathcal{L} \cup \mathcal{M} \), each node in \( \mathcal{W} \) must have at least two neighbors in \( S' \). Indeed, in any monotone improvement path from \( S' \), every node in \( (\mathcal{Y} \cup \overline{\mathcal{Y}}) \setminus S \) can only be added after each of its non-leave neighbors has been added, so that only nodes in \( S' \) can be used to affect the nodes in \( \mathcal{W} \); since these have degree 4, they must each have at least two neighbors in \( S' \). This implies that every node in \( \mathcal{W} \) must have one neighbor in \( S' \setminus \{z\} = \mathcal{Y} \cup \overline{\mathcal{Y}} \).

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

Lemma 2 and Lemma 3 thus show that starting from an instance of the 3-SAT, we could build an instance of the VCS problem in polynomial time and
of polynomial size, whose answer is the same as that of the 3-SAT. This shows that VCS is NP-hard.

4 A distributed algorithm for optimal control sets

The characterization of sufficient control sets through the concept of monotone improvement paths suggests the possibility that such sets may be searched for by starting from the all-1 configuration and iteratively replacing 1’s with 0’s in the attempt to follow backwards a monotone improvement path. In this section we introduce a family of discrete-time Markov chains \((Z^i_t)_{t \geq 0}\) on the configuration space \(X\), parameterized by a scalar parameter \(\varepsilon \in [0, 1]\), that capture this intuition. We will then prove that, for \(0 < \varepsilon \leq 1\) the Markov chain \((Z^i_t)_{t \geq 0}\) is reversible and that its stationary distribution concentrates on the family of optimal control sets as \(\varepsilon\) vanishes.

The dynamics of the Markov chain \(Z^i_t\) are described as follows: At every discrete time \(t = 0, 1, \ldots\), given that \(Z^i_t = z\), a node \(i\) is chosen at random uniformly from \(\mathcal{V}\). Then, if \(u_i(1, z_{-i}) < u_i(0, z_{-i})\), the state is not changed, i.e., \(Z^i_{t+1} = z\). Otherwise, if \(u_i(1, z_{-i}) \geq u_i(0, z_{-i})\), then if the current action of player \(i\) is \(z_i = 1\) it changes to 0 with probability 1, while if her current action is \(z_i = 0\), it changes to 1 with probability \(\varepsilon\). The transition probabilities of this Markov chain are then, for \(x, y \in X\),

\[
P^x_{xy} = \begin{cases} 
1/n & \text{if } y = x - \delta_i \text{ and } u_i(y) \leq u_i(x) \\
\varepsilon/n & \text{if } y = x + \delta_i \text{ and } u_i(y) \geq u_i(x) \\
0 & \text{otherwise} 
\end{cases}
\]  

Notice that, for \(\varepsilon = 0\), only transitions from 1 to 0 are allowed. In fact, in this case, the Markov chain \(Z^i_0\) has absorbing states. Specifically, let

\[Z_\infty = \{x \in X \mid P(\exists t_0 : Z^i_t = x \forall t \geq t_0 \mid Z^i_0 = 1) > 0\}\]  

(5)

and

\[Z = \{x \in X \mid P(\exists t_0 : Z^i_t = x \mid Z^i_0 = 1) > 0\}\]  

(6)

the sets of absorbing states and, respectively, of all states that are reachable by the Markov chain \(Z^i_0\) when started from \(Z^i_0 = 1\). Then, we have the following result.

**Proposition 3.** For a finite supermodular game with strategy profile space \(X = \{0, 1\}^n\), let \(Z\) and \(Z_\infty\) be defined as in (6) and (5), respectively. Then,

(i) \(S \subseteq \mathcal{V}\) is a sufficient control set iff \(1_S \in Z\);

(ii) if \(S\) is a minimal sufficient control set then \(1_S \in Z_\infty\).

**Proof.** (i) By definition, \(x = 1_S \in Z\) if and only if there exists a length-\(l\) path of configuration vectors \((y^k)_{0 \leq k \leq l}\), such that \(y^0 = 1\), \(y^l = 1_S\), and

\[y^k = y^{k-1} - \delta_i, \quad u_{ik}(y^k) \leq u_{ik}(y^{k-1}) \quad 0 \leq k \leq l.\]  

(7)
The above implies that the reversed path \((x^k)_{0 \leq k \leq l}\) with \(x^k = y^k - k\) for \(0 \leq k \leq l\), is a monotone improvement path from \(S\) to \(V\). Hence, \(S\) is a valid control set and Theorem 1 implies that it is a sufficient control set.

(ii) If \(S\) is a minimal sufficient control set, we know from point (i) that \(1_S \in Z\). Now if, by contradiction, \(1_S \notin Z_x\), then, from \(x = 1_S\), the Markov chain \(Z^0\) could reach, in one step, a different state \(x' = 1_{S'}\) with \(S' \subseteq S\), thus contradicting minimality of \(S\). \(\square\)

Proposition 3 implies that, in order to find an optimal control set, one may restrict the search to the set \(Z_\infty\) of absorbing states of the Markov chain \(Z^0\). Observe, however, that directly simulating the Markov chain \(Z^0\) started from \(Z^0 = 1\) would not be enough, since \(Z^0\) would be absorbed in finite time (in fact in a time smaller that \(|V|\)) in one state in \(Z_x\), but such a state may not be an optimal control set, and in fact not even a minimal control set, as the following example shows.

**Example 2.** Consider the majority game on the ring graph with four nodes. Then, \(Z^1 = (1, 0, 1, 0) \in Z_x\) is a sufficient control set, but it is not minimal since \(Z^2 = (1, 0, 0, 0) \in Z_x\) is also a sufficient control set.

To overcome this issue, we will instead use the Markov chain \(Z^\varepsilon\) with \(\varepsilon > 0\), which, as shown below, in \([1]\) turns out to be ergodic on the set \(Z\), and hence it does not get trapped in non-optimal or non-sufficient control sets, and to have stationary distribution concentrating on the set of optimal control sets as \(\varepsilon\) vanishes.

**Theorem 3.** For a finite supermodular game with strategy profile space \(X = \{0, 1\}^V\), let \(Z\) be defined as in \([6]\). Then, for \(\varepsilon > 0\), the Markov chain \(Z^\varepsilon\) with transition probabilities \(\{P\}\)

(i) is time-reversible and ergodic on the set of reachable states \(Z\);

(ii) its stationary probability

\[
\mu_x^\varepsilon := \frac{1}{K^\varepsilon} \varepsilon ||x||_1, \quad x \in Z, \tag{8}
\]

where \(K^\varepsilon = \sum_{x \in Z} \varepsilon ||x||_1\) converges in law to a probability measure \(\mu\) concentrated on the set of optimal sufficient control sets.

**Proof.** First observe that

\[
\varepsilon ||x||_1; P^\varepsilon_{x,y} = \varepsilon ||y||_1; P^\varepsilon_{y,x}, \tag{9}
\]

for every \(x, y \in X\). In particular, a transition probability \(P^\varepsilon_{x,y}\) is positive if and only if the reverse transition \(P^\varepsilon_{y,x}\) is positive. Thus, the set of reachable states from the starting state 1 is strongly connected. Clearly, the set of states reachable by the chain \(Z^\varepsilon\) from \(Z^0 = 1\) includes \(Z\) as every transition with positive probability for \(Z^0\) has also positive probability for \(Z^\varepsilon\).
In order to show the converse inclusion, let \( y \in \mathcal{X} \) be reachable by \( Z^*_i \) from \( Z^*_0 = 1 \) and let \( x^0 = 1, x^2, x^3 \cdots x^l = y \) be a length-\( l \) walk with positive probability for \( Z^*_i \). For every \( i \) such that \( y_i = 0 \), let \( T_i = \max \{ 0 \leq t \leq l : x^t_i = 1 \} \) the time of the latest change in the strategy of the \( i^{th} \) player along the path. Then, consider a new path \( z^0 = 1, z^2, z^3 \cdots z^m = y \), where \( m = n - \| y \|_1 \leq l \) and
\[
z^h = z^{h-1} - \delta_{i_h}, \quad 1 \leq h \leq l,
\]
where \( i_h \) is such that \( T_{i_h} \) is the \( h^{th} \) largest element of \( \{ T_i | y_i = 0 \} \). Observe that by construction
\[
x^{T_{i_h}} = x^{T_{i_h} + 1} + \delta_{i_h} \preceq z^h + \delta_{i_h} = z^{h-1}, \quad 0 \leq h \leq l - 1.
\]
On the other hand, it follows from (4) that
\[
u_{i_h}(0, x_{-i_h}^{T_{i_h}}) = u_{i_h}(x_{-i_h}^{T_{i_h}+1}) \leq u_{i_h}(x^{T_{i_h}}) = u_{i_h}(1, x_{-i_h}^{T_{i_h}}).
\]
Then, (10), the super-modularity property (2), and (11) imply that
\[
u_{i_h}(z^{h-1}) - u_{i_h}(z^h) = u_{i_h}(1, z^{h-1}) - u_{i_h}(0, z^h) \geq u_i(1, x_{-i_h}^{T_{i_h}}) - u_i(0, x_{-i_h}^{T_{i_h}}) \geq 0.
\]
It then follows from (12) and (4) that
\[P^0_{x_{-i_h}, z^h}, \quad \forall 1 \leq h \leq l,
\]
thus showing that \( y = z^l \) is reachable from \( z^0 \) by the Markov chain \( Z^*_i \). This proves that \( Z^*_i \) is ergodic on \( Z \) for \( \varepsilon \in \{ 0, 1 \} \). Together with (8), this also implies that \( Z^*_i \) is time-reversible with stationary distribution \( \mu^\varepsilon \). Clearly, as \( \varepsilon \to 0^+ \), the stationary distribution \( \mu^\varepsilon \) converges to a uniform distribution on \( \argmin_{x \in Z} \| x \|_1 \). Using proposition 3, the set \( \argmin_{x \in Z} \| x \|_1 \) is the set of optimal sufficient control sets.

4.1 Simulations

Below, we present some numerical simulations of the proposed algorithm for the case of the majority game on Erdös-Renyi random graphs. The Erdös-Renyi graph \( E(n, p) \) is a random undirected graph with \( n \) nodes where an edge between any pair of nodes is present with probability \( p \) and the presence of the various edges is an independent set of events. We consider \( n \) ranging up to 70 and two different models for \( p \). In the first case, we consider \( p = 0.4 \) is fixed and leads to a quite densely connected graph and one instead where \( p = 4 \log n \). This second choice leads to a more sparse graph that is typically connected \( \| \). We run the randomized algorithm \( Z^*_i \), with \( \varepsilon = 0.3 \), for a number of steps proportional to the square of the size of the graph (exactly \( 100n^2 \)) and the control set returned is the one of minimum cardinality during the walk. For small values of \( n \), an explicit comparison with the optimal solution, obtained through exhaustive search, proves the efficiency of our approach. Simulations
are reported in Figure 5. In Fig 6 we have made a comparison with respect to a naif heuristics that choose the highest degree nodes. Specifically, for each value of $n$, we have considered the set of the highest degree nodes in the graph of the same cardinality than the one found by our algorithm and we have plotted 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 the 30% and shows how the degree is not the right property to look at in the optimization of these control sets.

![Figure 5: Size of Control Sets for random graphs $E(n,p)$ with $p = 0.4$ (left) and $p = 4 \frac{\log n}{n}$ (right)](image)

![Figure 6: Coverage obtained by taking the $k$ highest degree node, with $k$ the size of the set found by the algorithm for random graphs $E(n,p)$ with $p = 0.4$ (left) and $p = 4 \frac{\log n}{n}$ (right)](image)

5 Conclusion

We have studied the problem of optimally control a set of players participating in a super-modular game in order to push their behavior from the minimal to the maximal Nash equilibrium. We have assumed the players to have the same binary action set and the control to consist in forcing action 1 in a subset of players. We have shown that the problem of selecting the minimum number of players to control to achieve the goal is NP complete and we have proposed a
computationally simple randomized algorithm with provable convergence properties.

The problem we have considered is an instance of a control problem in game models. Several directions for future research can be considered. For instance, in the context of super-modular games, one can consider non binary action sets and possibly more complex control actions that alter payoff’s functions of the players. Our techniques strongly use the super-modular assumption. Extensions to more general games (for instance best-shot public good games) are challenging and would require to develop different technical tools.

References

[1] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis. An efficient curing policy for epidemics on graphs. *IEEE Transactions on Network Science and Engineering*, 1(2):67–75, 2014.

[2] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis. When is a network epidemic hard to eliminate? *Mathematics of Operations Research*, 42(1-14), 2016.

[3] Matthew Jackson and Yves Zenou. Games on networks. *Handbook of Game Theory with Economic Applications*, 2015.

[4] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of SIGKDD’03*, pages 137–146, 2003.

[5] P. Milgrom and J. Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58(6):1255–1277, 1990.

[6] P. Milgrom and C. Shannon. Monotone comparative statics. *Econometrica*, 62(1):157–180, 1994.

[7] Stephen Morris. Contagion. *The Review of Economic Studies*, 2000.

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

[9] D. M. Topkins. Equilibrium points in nonzero-sum n-person submodular games. *SIAM Journal on Control and Optimization*, 17(6):773–787, 1979.

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

[11] L. Vassio, F. Fagnani, P. Frasca, and A. Ozdaglar. Message passing optimization of harmonic influence centrality. *IEEE Transactions on Control of Network Systems*, 1(1):109–120, 2014.
[12] X. Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19:305–321, 1990.

[13] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione. Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation*, 1(4):19–30, 2013.