

# Controlling a Social Network of Individuals with Coevolving Actions and Opinions

Roberta Raineri, Giacomo Como, Fabio Fagnani, Mengbin Ye, and Lorenzo Zino

**Abstract**—In this paper, we consider a population of individuals who have actions and opinions, which coevolve, mutually influencing one another on a complex network structure. In particular, we formulate a control problem for this social network, in which we assume that we can inject into the network a committed minority—a set of stubborn nodes—with the objective of steering the population, initially at a consensus, to a different consensus state. Our study focuses on two main objectives: i) determining the conditions under which the committed minority succeeds in its goal, and ii) identifying the optimal placement for such a committed minority. After deriving general monotone convergence result for the controlled dynamics, we leverage these results to build a computationally-efficient algorithm to solve the first problem and an effective heuristics for the second problem, which we prove to be NP-complete. The proposed methodology is illustrated through academic examples, and demonstrated on a real-world case study.

## I. INTRODUCTION

Over the past decades, the the systems and control community have witnessed a growing interest in developing and analyzing mathematical models to study, forecast, and control complex social phenomena and collective human behavior [1]–[7]. Within this general effort, particular interest has been devoted to collective decision-making, whereby a population of individuals have to repeatedly make decisions on a specific action to take (often binary) on the basis of several factors, including their opinions on the considered action. For instance, this scenario often arises in different contexts of social change problems: individuals may decide whether to use a disposable cup or a reusable cup to have a coffee, or whether to use inclusive language or not when writing an email. In these contexts, empirical evidence and social psychology theories suggest that decision-making is deeply intertwined with opinion formation processes [8], [9]. This calls for the development of model paradigms able to integrate opinion formation processes within the decision-making model in a coevolutionary fashion.

The continuous-opinion discrete-action model is a first step in this direction. The seminal paper by Martins [10] and its main extensions [11], [12] built on classical opinion dynamics

models [2], and rely on the assumption that the opinion formation process entirely shapes the decision-making, whereby actions are a quantization of opinions. Despite relevant for many applications, this assumption limits the possibility to capture the presence of a misalignment between individuals’ personal opinions and their actions, which is often observed in real-world social systems. This is the case, e.g., of the phenomenon of unpopular norms [13], [14], whereby a community keeps exhibiting a collective behavior that is disapproved by the most of its members. This limitation was addressed in [15], [16], where a coevolutionary model of actions and opinions was proposed. This model is built by incorporating an opinion formation process within a game-theoretic framework used to model decision-making [17], [18], and allows individuals to simultaneously revise their (binary) actions and share their opinions on their support for the action on a complex network, accounting for social pressure, opinion influence, and self-consistency. The analysis of this model has shown its ability to reproduce several real-life phenomena, including the emergence and persistence of unpopular norms and polarization.

In this paper, we take a step further from the analysis of social systems to their control. In the literature, the two most common approaches for controlling social systems are to assume that one can either i) provide monetary or societal incentives to favor a desired opinion/action over the others [19]–[23] or ii) directly control a committed minority of stubborn agents. Here, we focus on the second type of control action, which has been extensively studied in the context of opinion dynamics [24]–[26] and decision-making [17], [27]–[30], but it is still unexplored for the coevolutionary model.

Building on the model proposed in [16], we focus on the control problem of unlocking a paradigm shift in a population by means of a committed minority. Specifically, we consider a population initially at a consensus in which all individuals select and support the same action. Then, we introduce a committed minority of stubborn agents with the goal of steering the entire population to a consensus on the opposite action. Stubborn agents may consistently select the opposite action, share opinions supporting it, or both. This problem is relevant to many real-life applications. For instance, unlocking a paradigm shift is key for social change, e.g., to favor the collective transition towards more sustainable practices. From the opposite perspective, understanding how a committed minority can steer an entire population to a desired collective behavior is key to guaranteeing robustness of social systems against malicious attacks [31].

By leveraging systems and control theoretic tools, we study the controlled model and we establish an array of properties

R. Raineri and L. Zino are with the Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy ({roberta.raineri,lorenzo.zino}@polito.it). G. Como and F. Fagnani are with the Department of Mathematical Sciences, Politecnico di Torino, Turin, Italy ({giacomo.como,fabio.fagnani}@polito.it). M. Ye is with the School of Computer and Mathematical Sciences, University of Adelaide, Adelaide, Australia (ben.ye@adelaide.edu.au). G. Como is also with the Department of Automatic Control, Lund University, Lund, Sweden. This work was partially supported by the Western Australian Government through the Premier’s Science Fellowship Program.

and theoretical results, including monotonic convergence to an equilibrium point. Building on these theoretical findings, we establish necessary and sufficient conditions for the set of stubborn nodes to unlock a paradigm shift, depending on the model parameters and on the network structure. Then, we deal with the problem of identifying the minimal committed minority needed to control the network. After demonstrating analytically that the problem is NP-complete, we use our theoretical results and we get inspiration from [30] to design an efficient iterative algorithm for its solution.

In details, the main contribution of this paper is four-fold. First, we incorporate a control action in the coevolutionary model [16] and we formulate two control problems: determining if a control action is sufficient to steer the population to the desired consensus (*effectiveness guarantee problem*), and identifying the minimal set of nodes to be controlled to achieve such a goal (*minimal control set identification problem*). Second, we prove an array of general properties for the controlled dynamics, including convergence, and we characterize the complexity of our research problems, demonstrating that the minimal control set identification problem is NP-complete, which limits the possibility to adopt classical heuristics to approximate its solution. Third, we propose an algorithm to solve the effectiveness guarantee problem and, after prove its effectiveness analytically, we use it to evaluate the impact of the model parameters on a synthetic case study, where a closed-form solution of the problem can be derived. Fourth, we propose an iterative algorithm for the minimal control set identification problem with probabilistic convergence properties, and we demonstrate its efficiency on a case study, whose network topology is reconstructed from real-world face-to-face contact data [32].

Some of the results of this paper appeared, in a preliminary form, in [33]. Here, besides expanding the introduction to frame our contribution within the related literature and applications, we extend our preliminary results along several lines. First, we generalize the control action, allowing only part of the state of agents to be controlled. Second, we demonstrate that the minimal control set identification problem is NP-complete and it is not sub-modular (Theorem 2 and Proposition 4). Third, we extend the algorithm proposed in [33] to solve the effectiveness guarantee problem to the more general control setting considered in this paper, and we prove its effectiveness in Theorem 3, which was missing in our previous publication. Fourth, we propose a novel algorithm to solve the minimal control set identification problem for general networks and, after proving its convergence in Theorem 4, we demonstrate it on a realistic case study.

The rest of the paper is organized as follows. In Section II, we introduce the controlled coevolutionary dynamics and we formulate our two research problems. In Section III, we present some general results on the uncontrolled and controlled coevolutionary dynamics. Sections IV and V are devoted to the solution of the effectiveness guarantees and the minimal control set problems, respectively. In Section VI, we present a numerical case study. Section VII concludes the paper.

## II. MODEL AND PROBLEM STATEMENT

*Notation.* We denote a vector  $\mathbf{x}$  with bold lowercase font, with  $x_i$  its  $i$ th entry; and a matrix  $\mathbf{A}$  with bold capital font, and  $a_{ij}$  the  $j$ th entry of its  $i$ th row. The all-1 column vector is denoted as  $\mathbf{1}$ , with appropriate dimension depending on the context. Given two vectors  $\mathbf{x}, \mathbf{y}$  with same dimension, we use  $\mathbf{x} \leq \mathbf{y}$  to denote  $x_i \leq y_i$ , for all entries  $i$ .

### A. (Uncontrolled) Coevolutionary Model

We consider a population  $\mathcal{V} = \{1, \dots, n\}$  of  $n$  individuals. Each  $i \in \mathcal{V}$  is associated with a two-dimensional state variable  $(x_i(t), y_i(t)) \in \{-1, +1\} \times [-1, +1]$ , with discrete time  $t$ :  $x_i(t) \in \{-1, +1\}$  represents the *action* of individual  $i$  at time  $t$ ,  $y_i(t) \in [-1, +1]$  their *opinion* on the action ( $y_i(t) = -1$  means that  $i$  is totally in favor of action  $-1$ ,  $y_i(t) = +1$  that  $i$  fully supports action  $+1$ ). Actions and opinions are gathered in vectors  $\mathbf{x}(t) \in \{-1, 1\}^n$  and  $\mathbf{y}(t) \in [-1, 1]^n$ , and the state of the system is fully represented by the joint  $2n$ -dimensional vector  $\mathbf{z}(t) := (\mathbf{x}(t), \mathbf{y}(t)) \in \{-1, 1\}^n \times [-1, 1]^n$ . Given an individual  $i \in \mathcal{V}$ , we define as  $\mathbf{z}_{-i} := (\mathbf{x}_{-i}, \mathbf{y}_{-i}) \in \{-1, 1\}^{n-1} \times [-1, 1]^{n-1}$  the  $(2n-2)$ -dimensional vector with the state of all other individuals. At each time step  $t$ , we define a set  $\mathcal{R}(t) \subseteq \mathcal{V}$ , and we assume that all individuals in this set simultaneously revise their state at time  $t$ .

**Assumption 1** (Revision sequence). *There exists a constant  $T < \infty$  such that  $\cup_{s=0}^{T-1} \mathcal{R}(t+s) = \mathcal{V}$ , for any  $t \geq 0$ .*

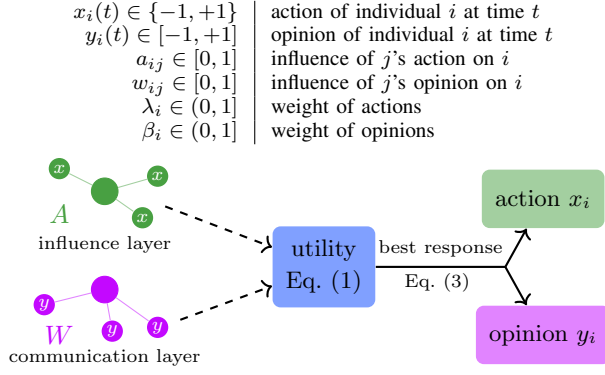
**Remark 1.** *Assumption 1 generalizes the synchronous and asynchronous update rules: for synchronous update rules,  $\mathcal{R}(t) = \mathcal{V}$  for all  $t$ ; for asynchronous update rules,  $\mathcal{R}(t)$  comprises a single individual, and Assumption 1 is imposed.*

At time  $t$ , each individual  $i \in \mathcal{R}(t)$  updates their state, aiming to maximize the utility function defined in [16], that accounts for three contributions: i) individuals' tendency to coordinate actions; ii) opinions exchanged with peers; and iii) an individual's tendency to have consistently between their action and opinion. Following [16], we define the utility that  $i$  receives for selecting an action and opinion pair  $\mathbf{z}_i = (x_i, y_i)$  when the state of the others is  $\mathbf{z}_{-i}$  as

$$u_i(\mathbf{z}_i, \mathbf{z}_{-i}) = \frac{\lambda_i(1-\beta_i)}{2} \sum_{j \in \mathcal{V}} a_{ij} [(1-x_j)(1-x_i + (1+x_j) \cdot (1+x_i))] - \beta_i(1-\lambda_i) \sum_{j \in \mathcal{V}} w_{ij} (y_i - y_j)^2 - \lambda_i \beta_i (x_i - y_i)^2, \quad (1)$$

where  $a_{ij} \in [0, 1]$  and  $w_{ij} \in [0, 1]$  are the influence of individual  $j$ 's action and opinion, respectively; and  $\lambda_i \in (0, 1]$  and  $\beta_i \in (0, 1]$  the weights given to actions observed and opinions exchanged, respectively. The quantities  $a_{ij}$  and  $w_{ij}$  are gathered into two matrices  $\mathbf{A}$  and  $\mathbf{W}$ , which induce a two-layer network  $\mathcal{G} = (\mathcal{V}, \mathcal{E}_A, \mathbf{A}, \mathcal{E}_W, \mathbf{W})$ , where  $\mathcal{E}_A$  are the edges on the influence layer on which individuals see others' actions and  $\mathcal{E}_W$  are the edges on the communication layer, on which individuals discuss about their opinions. All parameters are summarized in Table I. Before explicitly presenting the uncontrolled coevolutionary dynamics, we make some observations on the utility function in Eq. (1).

TABLE I: Models variables and parameters.

Fig. 1: Schematic of the update rule for individuals  $i \in \mathcal{R}(t)$ .

**Proposition 1.** A game with the utility function in Eq. (1) is supermodular.

*Proof.* Being the domain compact and the utility function upper-semicontinuous, to prove supermodularity, we need to check that Eq. (1) has increasing differences [34], i.e., that given  $\mathbf{z}_i' \geq \mathbf{z}_i$  and  $\mathbf{z}_{-i}' \geq \mathbf{z}_{-i}$ , it holds that  $\Delta_i(\mathbf{z}_i', \mathbf{z}_i, \mathbf{z}') := u_i(\mathbf{z}_i', \mathbf{z}') - u_i(\mathbf{z}_i, \mathbf{z}') \geq \Delta_i(\mathbf{z}_i', \mathbf{z}_i, \mathbf{z}_{-i})$ . Using Eq. (1), we compute

$$\Delta_i(\mathbf{z}_i', \mathbf{z}_i, \mathbf{z}_{-i}) = \lambda_i(1 - \beta_i) \sum_{j \in \mathcal{V}} a_{ij}(x_i' - x_i)x_j + 2\beta_i(1 - \lambda_i) \sum_{j \in \mathcal{V}} w_{ij}y_j(y_i' - y_i) + \psi(\mathbf{z}_i', \mathbf{z}_i), \quad (2)$$

where  $\psi(\mathbf{z}_i', \mathbf{z}_i)$  is a function that depends only on  $\mathbf{z}_i'$  and  $\mathbf{z}_i$ . Hence, being  $x_i' \geq x_i$  and  $y_i' \geq y_i$ , Eq. (2) is monotonic nondecreasing in  $x_j$  and  $y_j$ ,  $j \in \mathcal{V}$ . Hence,  $\Delta_i(\mathbf{z}_i', \mathbf{z}_i, \mathbf{z}_{-i}') \geq \Delta_i(\mathbf{z}_i', \mathbf{z}_i, \mathbf{z}_{-i})$  for any  $\mathbf{z}_{-i}' \geq \mathbf{z}_{-i}$ , yielding the claim.  $\square$

In Eq. (1), we enforce  $\lambda_i > 0$  and  $\beta_i > 0$  to guarantee a nontrivial coupling between the two variables. In the limit case in which one of these parameters is equal to 0, the coevolutionary model would reduce to a simpler (and well-known) dynamics, as commented in the following.

**Remark 2.** The utility in Eq. (1) generalizes classical network coordination games [17], [18] (obtained in the limit case  $\beta_i \rightarrow 0$ ) and the French-DeGroot opinion dynamics model [2], [35] (in the limit  $\lambda_i \rightarrow 0$ ). See, [16] for more details.

We are now ready to present the coevolutionary dynamics, in which agents who activate seek to maximize their utility function in Eq. (1). Consequently, for each  $i \in \mathcal{V}$ , the action and opinion are revised as follows:

$$x_i(t+1), y_i(t+1) = \begin{cases} \operatorname{argmax}_{\mathbf{z}_i \in \{-1, 1\} \times [-1, 1]} u_i(\mathbf{z}_i, \mathbf{z}_{-i}) & i \in \mathcal{R}(t), \\ x_i(t), y_i(t) & i \notin \mathcal{R}(t), \end{cases} \quad (3)$$

with the convention that, when the  $\operatorname{argmax}_{\mathbf{z}_i} u_i(\mathbf{z}_i, \mathbf{z}_{-i})$  comprises multiple elements, we set  $x_i(t+1) = x_i(t)$ . In other words, each individual  $i \in \mathcal{R}(t)$  performs a joint best-response with respect to Eq. (1), as illustrated in Fig. 1.

In Proposition 1, we proved that the game is supermodular. This is an important property of games and, in some scenarios (e.g., asynchronous update rules), it is sufficient to guarantee

convergence of the best-response dynamics [18], [36]. However, for general update sequences, supermodularity is not sufficient to guarantee convergence—a well-known case is the synchronous network coordination game [18], where the best-response update rule may lead to permanent oscillations. Therefore, we need to establish some further results on the coevolutionary dynamics to study its convergence and characterize its equilibria. Using [16], [33], we derive the following closed-form expression for Eq. (3).

**Proposition 2.** Individual  $i \in \mathcal{R}(t)$  updates their state as:

$$x_i(t+1) = s(\mathbf{z}(t)), \quad (4a)$$

$$y_i(t+1) = (1 - \lambda_i) \sum_{j \in \mathcal{V}} w_{ij}y_j(t) + \lambda_i s(\mathbf{z}(t)), \quad (4b)$$

where

$$s(\mathbf{z}(t)) = \begin{cases} +1 & \text{if } \delta_i(\mathbf{z}(t)) > 0, \\ -1 & \text{if } \delta_i(\mathbf{z}(t)) < 0, \\ x_i(t) & \text{if } \delta_i(\mathbf{z}(t)) = 0, \end{cases} \quad (5)$$

with

$$\delta_i(\mathbf{z}(t)) = 2\beta_i(1 - \lambda_i) \sum_{j \in \mathcal{V}} w_{ij}y_j(t) + (1 - \beta_i) \sum_{j \in \mathcal{V}} a_{ij}x_j(t). \quad (6)$$

From Proposition 2, we derive the following observation.

**Proposition 3** (Proposition 3 from [33]). The (uncontrolled) coevolutionary dynamics in Eq. (3) has at least two equilibria:  $\mathbf{x} = \mathbf{y} = -1$  and  $\mathbf{x} = \mathbf{y} = 1$ , being the unique equilibria in which the action vector is at a consensus ( $x_i = x_j$ ,  $\forall i, j \in \mathcal{V}$ ).

### B. Controlled dynamics and problem statement

We study a scenario in which, at time  $t = 0$ , the population is at one of the two consensus equilibria. Without any loss in generality, we assume  $\mathbf{x}(0) = \mathbf{y}(0) = -1$ . Starting from this initial consensus, our goal is to steer the system to the opposite one, i.e.,  $\mathbf{x} = \mathbf{y} = +1$ . To achieve such a goal, we consider this problem from the perspective of a policymaker/designer, and assume that our control lever is in the form of directly controlling the state of a subset of agents by setting their opinion and/or action to  $+1$  for all  $t \geq 1$ , yielding the following assumption, which is illustrated in Fig. 2.

**Assumption 2** (Controlled dynamics). Consider a two-layer network  $\mathcal{G} = (\mathcal{V}, \mathcal{E}_A, \mathbf{A}, \mathcal{E}_W, \mathbf{W})$  with  $\mathbf{A}$  and  $\mathbf{W}$  stochastic and irreducible. Given  $\mathcal{C}^X$  the set of controlled actions, and  $\mathcal{C}^Y$  the set of controlled opinions, there holds

$$\begin{cases} x_i(t) = +1 & \forall i \in \mathcal{C}^X, \forall t \geq 1, \\ y_j(t) = +1 & \forall j \in \mathcal{C}^Y, \forall t \geq 1, \\ x_i(0) = -1 & \forall i \in \mathcal{V} \setminus \mathcal{C}^X, \\ y_j(0) = -1 & \forall j \in \mathcal{V} \setminus \mathcal{C}^Y. \end{cases} \quad (7)$$

For  $i \notin \mathcal{C}^X$ , if  $i \in \mathcal{R}(t)$ , then  $x_i(t+1)$  follows Eq. (4a); for  $i \notin \mathcal{C}^Y$ , if  $i \in \mathcal{R}(t)$ , then  $y_i(t+1)$  follows Eq. (4b).

**Remark 3.** We identify three scenarios of particular interest:

- 1) **opinion control**, in which one controls only individuals' opinions ( $\mathcal{C}^X = \emptyset$ );

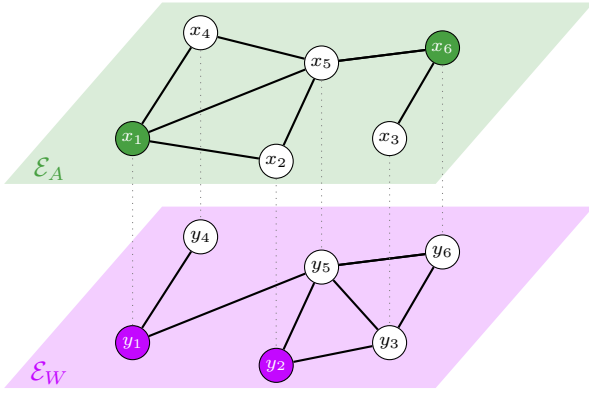


Fig. 2: Example of a setting that satisfies Assumption 2. The top layer represents individuals' action (white for  $-1$ , green for  $+1$ ), the bottom layer represents individuals' opinion (shades from white for  $-1$  to violet for  $+1$ ). Control sets are  $\mathcal{C}^X = \{1, 6\}$  and  $\mathcal{C}^Y = \{1, 2\}$ .

- 2) **action control**, in which one controls only individuals' actions ( $\mathcal{C}^Y = \emptyset$ )
- 3) **joint control**, in which one controls both variables ( $\mathcal{C}^X = \mathcal{C}^Y$ ), which is the special case considered in our preliminary work [33].

These three scenarios reflect different possible real-world interventions, which act only on opinions, on actions, or on both, capturing technical limitations or constraints which may prevent a policymaker from controlling both layers.

Hereafter, we will refer to a *controlled coevolutionary dynamics* as a coevolutionary dynamics with utility function in Eq. (1), under Assumptions 1 and 2. The goal of the controller, i.e., to lead all the agents to the desired consensus, can be formalized by first defining the objective function

$$\phi(\mathcal{C}^X, \mathcal{C}^Y) := \mathbb{P}[\exists T < \infty : \mathbf{x}(t) = \mathbf{1}, \forall t \geq T], \quad (8)$$

i.e., the probability (over the probability space generated by the revision sequence) that all individuals definitively switch their action to  $+1$  in finite time when the control sets are  $(\mathcal{C}^X, \mathcal{C}^Y)$ . The controller's goal is achieved iff  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$ . Hence, we formalize the following research problem.

**Problem 1** (Effectiveness guarantees). *Given a network  $\mathcal{G}$ , consider a controlled evolutionary dynamics on the network under Assumptions 1 and 2 with specified parameters. For given control sets  $(\mathcal{C}^X, \mathcal{C}^Y)$ , compute  $\phi(\mathcal{C}^X, \mathcal{C}^Y)$ .*

Solving Problem 1 would allow us to determine whether controlling the opinion and/or action of some nodes is sufficient to guarantee convergence to the desired consensus state. At this stage, a second question naturally follows: what is the minimal set of individuals that one should control in order to guarantee that the goal is achieved? Clearly, when controlling an individual, technical limitations may prevent from controlling both actions and opinions, as discussed in Remark 3. Hence, when formulating the problem, we introduce two additional constraints to denote the set of nodes whose action and opinion can be controlled as  $\mathcal{V}^X$  and  $\mathcal{V}^Y$ , respectively, allowing us to define the problem as follows.

**Problem 2** (Minimal control set). *Given a network  $\mathcal{G}$ , consider a controlled evolutionary dynamics on the network under Assumptions 1 and 2 with specified model parameters. Determine the solution to the following optimization problem*

$$\begin{aligned} \arg \min_{\mathcal{C}^X \subseteq \mathcal{V}, \mathcal{C}^Y \subseteq \mathcal{V}} \quad & |\mathcal{C}^X \cup \mathcal{C}^Y| \\ \text{s.t.} \quad & \phi(\mathcal{C}^X, \mathcal{C}^Y) = 1, \\ & \mathcal{C}^X \subseteq \mathcal{V}^X, \mathcal{C}^Y \subseteq \mathcal{V}^Y, \end{aligned} \quad (9)$$

where  $\mathcal{V}^X \subseteq \mathcal{V}$  and  $\mathcal{V}^Y \subseteq \mathcal{V}$  are constraints on the nodes whose action and opinion can be controlled, respectively.

**Remark 4.** By setting  $\mathcal{V}^X$  and  $\mathcal{V}^Y$ , one can enforce a specific form for the solution. In fact, by setting  $\mathcal{V}^X = \emptyset$  or  $\mathcal{V}^Y = \emptyset$ , we obtain solutions of Problem 2 with opinion or action control, respectively, i.e., the first two scenarios discussed in Remark 3. On the contrary, if the same constraints are imposed on the two sets ( $\mathcal{V}^X = \mathcal{V}^Y$ ), if a solution to Eq. (9) exists, then there is necessarily a solution with joint control ( $\mathcal{C}^X = \mathcal{C}^Y$ ), as it will be clear in the next section, after Corollary 1.

### III. MAIN PROPERTIES OF THE CONTROLLED DYNAMICS

In general, the uncontrolled coevolutionary dynamics requires restrictive assumptions to guarantee convergence, such as homogeneous parameters, symmetric layers, and presence of self-loops [16] or asynchronous update rules (see Proposition 1). For the controlled dynamics, instead, only the mild and general conditions on the activation sequence (Assumption 1) and on the stochasticity and irreducibility of the weight matrices (Assumption 2) are needed, as proved in the following.

**Theorem 1.** *Consider a controlled coevolutionary dynamics under Assumptions 1–2. Then, there exists an equilibrium  $(\mathbf{x}^*, \mathbf{y}^*)$  such that the action vector  $\mathbf{x}(t)$  converges to  $\mathbf{x}^*$  in finite time, and the opinion vector  $\mathbf{y}(t)$  converges to  $\mathbf{y}^*$  asymptotically. Moreover, both the opinion and action vectors are monotonically nondecreasing functions of time, i.e.,  $\mathbf{x}(t+1) \geq \mathbf{x}(t)$  and  $\mathbf{y}(t+1) \geq \mathbf{y}(t)$ , for all  $t \geq 0$ .*

*Proof.* The result follows the same line of arguments used in [33, Theorem 1]. Namely, we first prove that if actions are monotonically non-decreasing up to time  $\tau$  (i.e.,  $\mathbf{x}(t) \geq \mathbf{x}(t-1)$  for all  $t \leq \tau$ ), the also opinions are monotonically non-decreasing up to time  $\tau$  (i.e.,  $\mathbf{y}(t) \geq \mathbf{y}(t-1)$  for all  $t \leq \tau$ ). Second, we prove that the actions are indeed monotonically non-decreasing, yielding that both actions and opinion are monotonically non-decreasing, and thus they converge due to the monotone convergence theorem [37]. These two properties were proved in [33] for joint control. One can derive that the same monotonicity properties hold true under the more general Assumption 2. Explicit computations, omitted due to space limitations, follow the arguments in [33, Lemmas 1–2].  $\square$

Theorem 1 guarantees that under the general hypotheses of Assumption 2 the controlled coevolutionary dynamics converges and that actions converge in finite time. Moreover, it also guarantees monotonicity of the trajectory of the state vector  $\mathbf{z}(t)$ . As a consequence, if  $i \notin \mathcal{C}^X$  switches to action  $+1$  at a certain time, then  $i$  will never flip back. This observation will be fundamental to development of a systematic approach to

addressing our research problems, as presented in Sections IV and V. Finally, it is worth noticing that the hypothesis in Assumption 2 on the uncontrolled node dynamics are key for obtaining monotonicity (which then yields convergence). In fact, from a general initial condition, one may observe non-monotone trajectories (see, e.g., [16]). Besides monotonicity, we bring attention to the following property of the controlled dynamics, which simplifies our research problems.

**Lemma 1.** *Let us consider control sets  $(C^X, C^Y)$ . If  $\phi(C^X, C^Y) = 1$ , then all control sets  $(\bar{C}^X, \bar{C}^Y)$  such that  $C^X \subseteq \bar{C}^X$  and  $C^Y \subseteq \bar{C}^Y$  satisfy  $\phi(\bar{C}^X, \bar{C}^Y) = 1$ . If  $\phi(C^X, C^Y) = 0$ , then all control sets  $(\bar{C}^X, \bar{C}^Y)$  such that  $\bar{C}^X \subseteq C^X$  and  $\bar{C}^Y \subseteq C^Y$  satisfy  $\phi(\bar{C}^X, \bar{C}^Y) = 0$ .*

*Proof.* To prove the first claim, we observe from Eq. (2) that  $\delta_i(z)$  is a monotonic nondecreasing function of  $z$ . Consequently, from Eqs. (4a–4b),  $x(t+1)$  and  $y(t+1)$  are monotonic nondecreasing functions of  $z(t)$ . Let us define the controlled coevolutionary dynamics  $z(t)$ , with initial condition  $z(0)$  according to Eq. (7). Let us consider control sets  $\bar{C}^X \supset C^X$  and  $\bar{C}^Y \supset C^Y$ , and corresponding dynamics  $\bar{z}(t)$  with initial conditions defined according to according to Eq. (7). Then, fixed any common activation sequence for  $z(t)$  and  $\bar{z}(t)$ , the monotonicity properties described in the above yields  $\bar{z}(t) \geq z(t)$ . Finally, for almost every activation sequence we have that  $x(t) \rightarrow 1$ , being  $\phi(C^X, C^Y) = 1$ . Consequently, also  $\bar{x}(t) \rightarrow 1$  for all configurations (except for a set of measure zero), yielding the claim. The second claim is proved following a similar (symmetric) argument.  $\square$

An immediate consequence of Lemma 1 is the following.

**Corollary 1.** *If Problem 2 admits a solution, then there is always a solution such that, letting  $\mathcal{C} := C^X \cup C^Y$ , then  $C^X = \mathcal{C} \cap \mathcal{V}^X$  and  $C^Y = \mathcal{C} \cap \mathcal{V}^Y$ .*

*Proof.* Let  $C^X, C^Y$  be an optimal solution of Eq. (9). Then, we define  $\mathcal{C} := C^X \cup C^Y$ ,  $\hat{C}^X := \mathcal{C} \cap \mathcal{V}^X$ , and  $\hat{C}^Y := \mathcal{C} \cap \mathcal{V}^Y$ . Clearly,  $\hat{C}^X \supseteq C^X$  and  $\hat{C}^Y \supseteq C^Y$ . Hence, by Lemma 1, since  $\phi(C^X, C^Y) = 1$ , also  $\phi(\hat{C}^X, \hat{C}^Y) = 1$ . Moreover,  $C^X \subseteq \mathcal{V}^X$  and  $C^Y \subseteq \mathcal{V}^Y$  by construction, so  $(\hat{C}^X, \hat{C}^Y)$  is a feasible solution of Eq. (9). Finally, we observe that  $\hat{C}^X \cup \hat{C}^Y = \mathcal{C}$ . Hence,  $|\hat{C}^X \cup \hat{C}^Y| = |C^X \cup C^Y| = |\mathcal{C}|$ , and  $(\hat{C}^X, \hat{C}^Y)$  is an optimal solution of Eq. (9), yielding the claim.  $\square$

In plain words, Corollary 1 states that, if it is possible to control both the action and the opinion of a given individual, then it is always optimal to either control both variables or to control none. As a consequence of Lemma 1 and Corollary 1, Problem 2 can be reduced to an optimization problem over a single control set, as stated in the following.

**Corollary 2.** *Let*

$$C^* = \arg \min_{C \subseteq \mathcal{V}} |C| \quad \text{s.t.} \quad \phi(C \cap \mathcal{V}^X, C \cap \mathcal{V}^Y) = 1. \quad (10)$$

*Then, an optimal solution of Problem 2 is given by  $C^X = C^* \cap \mathcal{V}^X$  and  $C^Y = C^* \cap \mathcal{V}^Y$ .*

It is worth noticing that, despite the useful properties of the controlled coevolutionary dynamics demonstrated in the

above, the problem of controlling the dynamics and determining the minimal control sets is inherently complex, as stated in the following result, with the proof in Appendix A.

**Theorem 2.** *Problem 2 is NP-complete.*

Moreover, we can show that the objective function in Eq. (8) is not submodular, hindering the possibility to easily derive sub-optimal solutions via greedy algorithms [38], as is done for related control problems on social networks [27].

**Proposition 4.** *The function  $\phi(C^X, C^Y)$  in Eq. (8) is not submodular with respect to any of its two variables.*

*Proof.* We build a counterexample. Consider 2 nodes connected by a link with  $w_{12} = a_{12} = w_{21} = a_{21} = 1/3$  and  $w_{11} = a_{11} = w_{22} = a_{22} = 2/3$ , and synchronous activation rule  $(\mathcal{R}(t) = \mathcal{V}, \text{ for all } t)$ . We consider  $\mathcal{C}_1^X = \{1\}$ ,  $\mathcal{C}_2^X = \{2\}$ , and  $\mathcal{C}^Y = \mathcal{V}$ . Clearly, it holds  $\mathcal{C}_1^X \cup \mathcal{C}_2^X = \mathcal{V}$ , which implies that  $\phi(\mathcal{C}_1^X \cup \mathcal{C}_2^X, \mathcal{C}^Y) = \phi(\mathcal{V}, \mathcal{V}) = 1$ . On the other hand, for control sets  $\mathcal{C}_1^X, \mathcal{C}^Y$ , we immediately observe that the only state that can change is  $x_2(t)$ . At the first time instant  $t$  at which  $2 \in \mathcal{R}(t)$ , which occurs in finite time due to Assumption 1, individual 2 switches to +1 iff  $\delta_2(z(t)) = 2\beta_2(1 - \lambda_2) - (1 - \beta_2) > 0$ , according to Proposition 2. By symmetry, a similar condition holds for individual 1 when  $\mathcal{C}_2^X$ , exchanging the role of individuals 1 and 2. This implies that

$$\phi(\mathcal{C}_i^X, \mathcal{C}^Y) = \begin{cases} 1 & \text{if } \lambda_{3-i} \leq \frac{3\beta_{3-i}-1}{2\beta_{3-i}}, \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

Hence, if we set  $\beta_1 = \beta_2 = 1/2$  and  $\lambda = 2/3$ , then  $\phi(\mathcal{C}_1^X, \mathcal{C}^Y) = \phi(\mathcal{C}_2^X, \mathcal{C}^Y) = 0$ , which implies that the condition  $\phi(\mathcal{C}_1^X, \mathcal{C}^Y) + \phi(\mathcal{C}_2^X, \mathcal{C}^Y) \geq \phi(\mathcal{S} \cup \mathcal{T}, \mathcal{C}^Y) + \phi(\mathcal{S} \cap \mathcal{T}, \mathcal{C}^Y)$  required by submodularity [34] is not satisfied by the first variable. Similar, we can build a counterexample to prove that the function is not submodular also with respect to the second variable, yielding the claim.  $\square$

#### IV. EFFECTIVENESS GUARANTEES PROBLEM

The results from the previous section call for the development of algorithms able to solve our research problem in an efficient way. We start from Problem 1. In our preliminary work [33], we proposed an algorithm and conjectured that this algorithm can solve Problem 1 in the simplified setting of complete control,  $C^X = C^Y$  (see Remark 3). Here, we propose a refined version of the algorithm that accounts for the more general setting in Assumption 2, while also reducing the total computational complexity. Then, we rigorously prove the conjecture in the general scenario, demonstrating that the proposed algorithm solves Problem 1 in polynomial time. Our procedure is based on the following iterative scheme.

We consider the general case in which we control the action for nodes in  $C^X$  and the opinion for nodes in  $C^Y$  as described in Assumption 2. At iteration  $k = 1$ , we initialize the algorithm by defining  $\mathcal{A}(1) = C^X$ , which involve only the nodes whose action is controlled. At each step of the algorithm  $k$ , we construct a candidate equilibrium with action vector  $\hat{x}$  with

$$\hat{x}_i = \begin{cases} +1 & \text{if } i \in \mathcal{A}(k), \\ -1 & \text{if } i \notin \mathcal{A}(k), \end{cases} \quad (12)$$

and opinion vector  $\hat{\mathbf{y}}$ , computed by solving the linear system

$$\hat{y}_i = \begin{cases} (1 - \lambda_i) \sum_{j \in \mathcal{V}} w_{ij} \hat{y}_j + \lambda_i \hat{x}_i & \text{if } i \notin \mathcal{C}^Y, \\ +1 & \text{if } i \in \mathcal{C}^Y, \end{cases} \quad (13)$$

which has a unique solution, as we will prove later. Observe that solving Eq. (13) requires inverting a matrix that is independent of the variables. This operation can be optimized by computing it in advance, before running the iterations.

Then, we will demonstrate in Theorem 3 below the following properties. First, we show that given an action vector  $\hat{\mathbf{x}}$ , there exists a unique  $\hat{\mathbf{y}}$  such that  $\hat{\mathbf{z}} = (\hat{\mathbf{x}}, \hat{\mathbf{y}})$  is a candidate equilibrium of the controlled coevolutionary dynamics. To check whether  $\hat{\mathbf{z}}$  is an actual equilibrium, we check if any individual  $i$  who plays action  $-1$  at  $\hat{\mathbf{z}}$  would switch to  $+1$ . According to Proposition 2, this can be checked by computing the sign of  $\delta_i(\hat{\mathbf{z}})$  for all  $i \notin \mathcal{A}(k)$ . If all  $\delta_i(\hat{\mathbf{z}}) \leq 0$ , then no individual will switch action, and the candidate  $\hat{\mathbf{z}}$  is indeed the equilibrium reached by the system. Otherwise, we will prove that all individuals with  $\delta_i(\hat{\mathbf{z}}) > 0$  will eventually switch to  $+1$ . Hence,  $\hat{\mathbf{z}}$  is not an equilibrium of the controlled coevolutionary dynamics, and we need to consider other potential equilibria where also those individuals with  $\delta_i(\hat{\mathbf{z}}) > 0$  switch to action  $+1$ . To this aim, we increase the iteration index  $k$  by 1, and we enlarge the set  $\mathcal{A}(k)$  by incorporating these individuals into  $\mathcal{A}(k-1)$ , and we iterate the procedure, until the termination criterion  $\mathcal{A}(k) = \mathcal{A}(k-1)$ , which implies that no more individuals would change action. According to this procedure, we get a non-decreasing sequence of sets. When the termination criterion is met, the algorithm returns  $\mathcal{A}_f$ .

This algorithm, for which a computationally-improved pseudo-code is reported in Algorithm 1, offers a tool to solve Problem 1 in a polynomial time, as summarized in the following statement, whose proof is reported in Appendix B.

---

**Algorithm 1:** Equilibrium computation

---

**Data:**  $\mathbf{A}, \mathbf{W}, \mathcal{C}^X, \mathcal{C}^Y, \lambda_i$  and  $\beta_i$ , for all  $i \in \mathcal{U}$   
**Result:**  $\mathcal{A}_f := \mathcal{A}(k)$ , i.e., individuals with  $x^* = +1$   
 $k \leftarrow 1$ ;  $\mathcal{A}(0) \leftarrow \emptyset$ ;  $\mathcal{A}(1) \leftarrow \mathcal{C}^X$ ;  $\hat{y}_i \leftarrow +1 \ \forall i \in \mathcal{C}^Y$ ;  
 $\mathbf{M} \leftarrow (\mathbf{I} - [(\mathbf{I} - \text{diag}(\boldsymbol{\lambda}))\mathbf{W}])^{-1}$ ;  
**while**  $\mathcal{A}(k) \neq \mathcal{A}(k-1)$  **do**  
    Define  $\hat{\mathbf{x}}$  using Eq. (12);  
     $\hat{y}_i \leftarrow (\mathbf{M} \text{diag}(\boldsymbol{\lambda}) \hat{\mathbf{x}})_i$  for all  $i \notin \mathcal{C}^Y$ ;  
     $k \leftarrow k + 1$ ;  $\mathcal{A}(k) \leftarrow \mathcal{A}(k-1)$ ;  
    **check for**  $i \in \mathcal{V}$  &  $i \notin \mathcal{A}(k)$  **do**  
        **if**  $\delta_i(\hat{\mathbf{x}}, \hat{\mathbf{y}}) > 0$  **then**  
             $\mathcal{A}(k) \leftarrow \mathcal{A}(k) \cup \{i\}$ ;  
        **end**  
    **end**  
**end**

---

**Theorem 3.** Under Assumptions 1 and 2, Algorithm 1 solves Problem 1 in time  $O(n^3)$ . In fact, given control sets  $(\mathcal{C}^X, \mathcal{C}^Y)$  and output  $\mathcal{A}_f$  of Algorithm 1, then

$$\phi(\mathcal{C}^X, \mathcal{C}^Y) = \begin{cases} 1 & \text{if } \mathcal{A}_f = \mathcal{V}, \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

Moreover, the equilibrium reached by a controlled coevolutionary dynamics that satisfies Assumptions 1 and 2 with control

sets  $(\mathcal{C}^X, \mathcal{C}^Y)$  is  $(\mathbf{x}^*, \mathbf{y}^*)$ , with  $\mathbf{x}^*$  defined as in Eq. (12) with  $\mathcal{A}(k) = \mathcal{A}_f$  and  $\mathbf{y}^*$  the solution of Eq. (13) given  $\mathbf{x}^*$ .

In most practical scenarios, only a few iterations are needed for convergence, since  $\mathcal{A}(k)$  often increases by more than one individual at each iteration, further reducing the computational effort needed. Moreover, one can leverage symmetry properties of the network structure to reduce the dimension of the system, as illustrated in the following example.

#### A. Complete Graph

We illustrate a motivational example showing how our results can be used to analyze the coevolutionary dynamics and the possibility to control it. Specifically, we consider the case of a complete graph (including self-loops) with homogeneous parameters and weights, in which, due to symmetry reasons, the strategy adopted guarantees a solution for both Problems 1 and 2. In order to check if a candidate control set succeeds in complete network controllability we use Algorithm 1.

**Assumption 3** (Homogeneous complete graph). Let  $\mathcal{G}$  be a two-layer network with  $a_{ij} = w_{ij} = \frac{1}{n-1}$  and  $a_{ii} = w_{ii} = 0$ ,  $\forall i \neq j \in \mathcal{V}$ . Moreover, let  $\lambda_i = \lambda$  and  $\beta_i = \beta$ ,  $\forall i \in \mathcal{V}$ .

In [33], we have analysed the scenario of a complete graph with complete control (see Remark 3). Here, we focus on the other two scenarios of interest discussed in Remark 3. More precisely, we assume that we are able to control a certain number of individuals (whose position is irrelevant, due to the network completeness) and besides complete control, we also consider the two scenarios of opinion control and action control. Using Algorithm 1 and Theorem 3, we establish the following result.

**Proposition 5.** Consider a coevolutionary dynamics that satisfies Assumptions 1–2 on a complete graph  $\mathcal{G}$  with  $n$  nodes that satisfies Assumption 3. Given  $\mathcal{C} \subseteq \mathcal{V}$ , let  $\gamma = \frac{|\mathcal{C}|}{n-1}$ . Then,

i) for opinion control,  $\phi(\emptyset, \mathcal{C}) = 1$  iff it holds

$$\frac{\beta[3(1-\lambda)\gamma + 2\lambda\gamma(1-\lambda) + \lambda(2\lambda-1)]}{\gamma + \lambda - \lambda\gamma} > 1; \quad (15)$$

ii) for action control,  $\phi(\mathcal{C}, \emptyset) = 1$  iff  $\gamma > 1/2$ ;

iii) for joint control,  $\phi(\mathcal{C}, \mathcal{C}) = 1$  iff it holds

$$2\beta(1-\lambda)\left(\frac{\gamma-\lambda+\lambda\gamma}{\gamma+\lambda-\lambda\gamma}\right) + (1-\beta)(2\gamma-1) > 0. \quad (16)$$

*Proof.* In all cases, we apply Algorithm 1 and Theorem 3.

i) We start with  $\mathcal{A}(1) = \emptyset$ . The candidate equilibrium  $\hat{\mathbf{z}}$  has action vector  $\hat{\mathbf{x}} = -1$  and opinion vector  $\hat{\mathbf{y}}$  with  $\hat{y}_i = +1$  for all  $i \in \mathcal{C}$ , and by solving Eq. (13) for all  $i \notin \mathcal{C}$ . By symmetry, all  $i \notin \mathcal{C}$  have necessarily  $\hat{y}_i = +1$ , solution of  $\hat{y}_i = (1-\lambda)(\gamma + (1-\gamma)\hat{y}_i) - \lambda$ , which yields  $\hat{y}_i = \frac{(1-\lambda)\gamma - \lambda}{1 - (1-\lambda)(1-\gamma)}$ . Substituting this in  $\delta_i(\hat{\mathbf{z}})$ , we get that  $\delta_i(\hat{\mathbf{z}}) > 0$  iff the condition in Eq. (15) is satisfied. In this scenario,  $\mathcal{A}(2) = \mathcal{V}$ ; otherwise,  $\mathcal{A}(2) = \mathcal{A}(1)$ . In both cases the algorithm terminates, and Theorem 3 yields claim i).

ii) We start with  $\mathcal{A}(1) = \mathcal{C}$ . The corresponding candidate equilibrium  $\hat{\mathbf{z}}$  has  $\hat{\mathbf{x}}$  defined using Eq. (12). Conversely, for the opinion vector  $\hat{\mathbf{y}}$  we have to distinguish two different cases:  $\hat{y}_{\mathcal{C}}$



indicates the equilibrium value for those nodes whose action is controlled,  $\hat{y}_U$  the one for those nodes which are not controlled at all. The equilibrium vector is so defined as follows:

$$\begin{cases} \hat{y}_C &= (1 - \lambda)[\gamma \hat{y}_C + (1 - \gamma)\hat{y}_U] + \lambda \\ \hat{y}_U &= (1 - \lambda)[\gamma \hat{y}_C + (1 - \gamma)\hat{y}_U] - \lambda. \end{cases} \quad (17)$$

Solving Eq. (17) and substituting the solution in Eq. (2) for any  $i \notin \mathcal{C}$ , we obtain  $\delta_i(\hat{z}) = (2\gamma - 1)((1 - 2\lambda)\beta\lambda + 1)$ . It is easy to verify that  $((1 - 2\lambda)\beta\lambda + 1) > 0$  for any choice of  $\lambda$  and  $\beta$ , yielding the condition  $\gamma > 1/2$ . If  $\gamma > 1/2$ , then  $\delta_i(\hat{z}) > 0$ , and  $\mathcal{A}(2) = \mathcal{V}$ ; otherwise,  $\mathcal{A}(2) = \mathcal{A}(1) = \mathcal{C}$ . In both cases, the algorithm terminates, and Theorem 3 yields ii).

iii) We apply Theorem 3 to [33, Proposition 3].  $\square$

**Remark 5.** Proposition 5 solves both Problems 1 and 2 for a complete graph. In fact, given control sets  $(\mathcal{C}^X, \mathcal{C}^Y)$ , where  $\gamma = \frac{|\mathcal{C}^X \cup \mathcal{C}^Y|}{n-1}$ , Problem 1 is solved for those values of  $\lambda$  and  $\beta$  that satisfy the condition corresponding to the scenarios considered from Remark 3. On the other hand, given  $\gamma$ , the choice of the nodes to control is irrelevant, and re-writing Eq. (15) and Eq. (16) as conditions on  $\gamma$ , we can ultimately determine the minimum number of nodes to be controlled to guarantee  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$ , solving Problem 2.

From Proposition 5 we can draw some interesting conclusions. First, from item ii), we observe that for action control, the problem is exactly the same as controlling a majority on a complete network: we need more than 50% of nodes being controlled. On the contrary, the conditions become nontrivial when considering opinion control and complete control, since they depend on the model parameters. This suggests that, depending on the model parameters and on the choice of the controlled variables, one can facilitate the controllability of the network or increase its robustness against malicious attacks.

Figure 3 illustrates this nontrivial relation, with color intensity representing the minimal fraction of individuals  $\gamma$  needed to be controlled in order to guarantee  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$ . We highlight in particular how controlling only opinions has a great impact on the cardinality of the controlled set  $\mathcal{C}$ . Indeed, the impact of controlled opinions is directly proportional to parameter  $\beta$  and inversely proportional to  $\lambda$ . This clearly emerges from the color intensity in Figure 3. As expected, it increases for bigger values of  $\lambda$  and lower values of  $\beta$ . The joint control on actions and opinions makes network controllability easier, by reducing the percentage of agents needed to be controlled.

## V. MINIMAL CONTROL SET IDENTIFICATION

In the previous section, we have shown how Algorithm 1 can be used to solve our research problems for simple network structures such as a complete graph (Section IV-A) or a star graph (see [33]). However, in more realistic scenarios, the network does not have such a level of symmetry, and Algorithm 1 is not efficient to solve Problem 2. In fact, in order to find the optimal control sets  $\mathcal{C}^X$  and  $\mathcal{C}^Y$ , one should look for all pairs of subsets of  $\mathcal{V}$  to find those that guarantee that the constraint  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$  is satisfied. First, we observe that Corollary 2 allows us to simplify the problem (and the

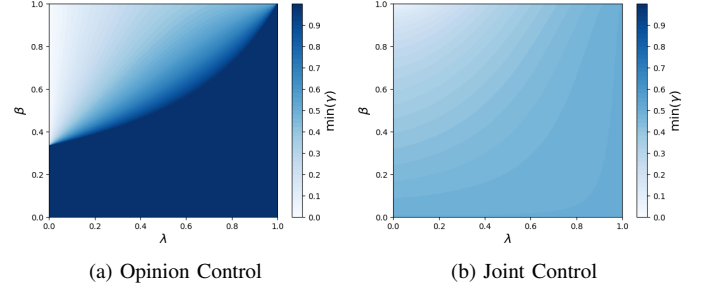


Fig. 3: Results for a complete graph. The color intensity represents the cardinality of the minimal control set  $\gamma = |\mathcal{C}|/(n-1)$  that solves Problem 2.

notation), since we simply need to find an optimal control set  $\mathcal{C}$ , and then define the solution of Problem 2 as  $(\mathcal{C}^X, \mathcal{C}^Y)$  with

$$\mathcal{C}^X = \mathcal{C} \cap \mathcal{V}^X \text{ and } \mathcal{C}^Y = \mathcal{C} \cap \mathcal{V}^Y. \quad (18)$$

However, even with this simplification, the determination of the minimal control set remains computationally challenging.

With an aim to reduce the computational cost to perform such a task, we now build on the optimal targeting algorithm proposed in [30] in order to design an algorithm able to solve Problem 2. Intuitively, the rationale behind the proposed methodology consists of starting from controlling all the possible individuals and moving backwards, removing individuals from the control set, until we find the minimal control set that allows us to reach the goal in Problem 2. However, such a naive approach might result in being stuck in a local minimum, not being able to reach the global optimum. For this reason, we employ a stochastic approach, which consists of defining a discrete-time Markov chain [39] that explores the space of control sets that are feasible solutions of Problem 2 in such a way that its invariant distribution will provably concentrate about the global optimal solutions of Eq. (10), allowing us to define an effective heuristics to solve Problem 2.

More precisely, given Problem 2 with constraints  $\mathcal{C}^X \subseteq \mathcal{V}^X$  and  $\mathcal{C}^Y \subseteq \mathcal{V}^Y$ , we define the set of all *controllable nodes* as

$$\mathcal{V}^* := \mathcal{V}^X \cup \mathcal{V}^Y, \quad (19)$$

and we let  $n^* := |\mathcal{V}^*|$  to be the number of controllable nodes. Then, we can define the space of all *potential control sets*, which is nothing but the power set of  $\mathcal{V}^*$ , i.e.,  $\mathcal{C} := \{\mathcal{C} : \mathcal{C} \subseteq \mathcal{V}^*\}$ . Moreover, we say that a potential control set  $\mathcal{C} \in \mathcal{C}$  is *admissible* if and only if  $\phi(\mathcal{C} \cap \mathcal{V}^X, \mathcal{C} \cap \mathcal{V}^Y) = 1$ . In other words, an admissible control set is a feasible solution of Eq. (10). We indicate the space of all admissible control sets as

$$\bar{\mathcal{C}} := \{\mathcal{C} \subseteq \mathcal{C} : \phi(\mathcal{C} \cap \mathcal{V}^X, \mathcal{C} \cap \mathcal{V}^Y) = 1\}, \quad (20)$$

which is clearly the set of all feasible solutions of Problem 2.

Our algorithm starts from the worst case in which we control all the controllable nodes, i.e., we set  $\mathcal{C} = \mathcal{V}^*$ . First, we need to check whether  $\mathcal{C}$  is admissible, i.e., we define  $\mathcal{C}^X$  and  $\mathcal{C}^Y$  as in Eq. (18), and check whether  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$ . This check is done by employing Algorithm 1. If  $\mathcal{C}$  is admissible, then it follows that  $(\mathcal{C}^X, \mathcal{C}^Y)$  is a feasible solution to Problem 2. Conversely, if  $\mathcal{C}$  is not admissible, then Problem 2 is unfeasible

and there is no need to proceed. Trivially, we observe that if  $\mathcal{V}^X = \mathcal{V}$  then  $\mathcal{C}$  is always admissible, being  $\mathcal{C}^X = \mathcal{V}$ , and  $\phi(\mathcal{V}, \mathcal{C}^Y) = 1$  for any choice of  $\mathcal{C}^Y$ .

If Problem 2 is feasible, we then adopt the following iterative procedure, which is detailed in Algorithm 2. At the  $k$ th iteration we start with the control set  $\mathcal{C}$ . We select a node  $r$ , uniformly at random among the controllable nodes, i.e.,  $r \in \mathcal{V}^*$ . Then, two cases are possible:

- 1) The node belongs to the control set ( $r \in \mathcal{C}$ ). In this case, if the set  $\mathcal{C} \setminus \{r\}$  is admissible (which is checked using Algorithm 1), then the control set is updated to  $\mathcal{C} \setminus \{r\}$ ; otherwise it remains  $\mathcal{C}$ ; or
- 2) The node does not belong to the control set ( $r \notin \mathcal{C}$ ). In this case, we introduce a probability  $\varepsilon \in (0, 1]$ , and the node  $r$  is added to the control set with probability  $\varepsilon$ , i.e., the control set is updated to  $\mathcal{C} \cup \{r\}$ ; otherwise,  $\mathcal{C}$  remains unchanged.

The iteration counter is thus increased to  $k+1$  and the process is repeated. As a design choice, we will consider a maximum number of iterations for the algorithm equal to  $T$ .

---

**Algorithm 2: Optimal control set identification**


---

**Data:**  $\mathbf{A}, \mathbf{W}, \lambda, \beta, \varepsilon, n, T, \mathcal{V}^X$ , and  $\mathcal{V}^Y$   
**Result:**  $\hat{\mathcal{C}}$  such that  $(\hat{\mathcal{C}}^X, \hat{\mathcal{C}}^Y)$  solves Problem 2  
 $k \leftarrow 1$ ;  $\mathcal{C} \leftarrow \mathcal{V}^X \cup \mathcal{V}^Y$ ;  $\hat{\mathcal{C}} \leftarrow \emptyset$ ;  
**if**  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$  **then**  
     $\hat{\mathcal{C}} \leftarrow \mathcal{C}$ ;  
**while**  $k < T$  **do**  
     $k \leftarrow k + 1$ ; Choose at random a node  $r \in \mathcal{V}^X$ ;  
    **if**  $r \in \mathcal{C}$  **then**  
        **if** Algorithm 1 yields  $\mathcal{A}_f = \mathcal{V}$  **then**  
             $\mathcal{C} \leftarrow \mathcal{C} \setminus \{r\}$ ;  
            **if**  $|\mathcal{C}| < |\hat{\mathcal{C}}|$  **then**  
                 $\hat{\mathcal{C}} \leftarrow \mathcal{C}$ ;  
            **end**  
        **end**  
    **end**  
    **else**  
         $\mathcal{C} \leftarrow \mathcal{C} \cup \{r\}$  with probability  $\varepsilon$   
    **end**  
**end**

---

Before formally presenting the Markov chain induced by this iterative process and illustrating how this can be used to solve Problem 2, we offer here a simple example to elucidate the procedure described above. We consider a network with  $n = 4$  nodes and with  $\mathcal{V}^* = \mathcal{V}$ . We start, at an arbitrary iteration step, from a control set  $\mathcal{C} = \{1, 2, 3\}$ . Figure 4a illustrates the three elements of  $\mathcal{C}$  that can be reached from  $\mathcal{C}$ , depending on which node  $r$  is selected. Assume that Algorithm 1 prescribes that only the two sets on the left belong to  $\mathcal{C}$  and are thus admissible, while the third one is not admissible, and it is thus barred in Fig. 4a. If nodes  $r = 2$  or  $r = 3$  are selected, then we are in step 1) and the chain has a transition to  $\mathcal{C} \setminus \{2\}$  or  $\mathcal{C} \setminus \{3\}$ , respectively. If instead node  $r = 1$  is selected, we are in step 1) but the chain remains in

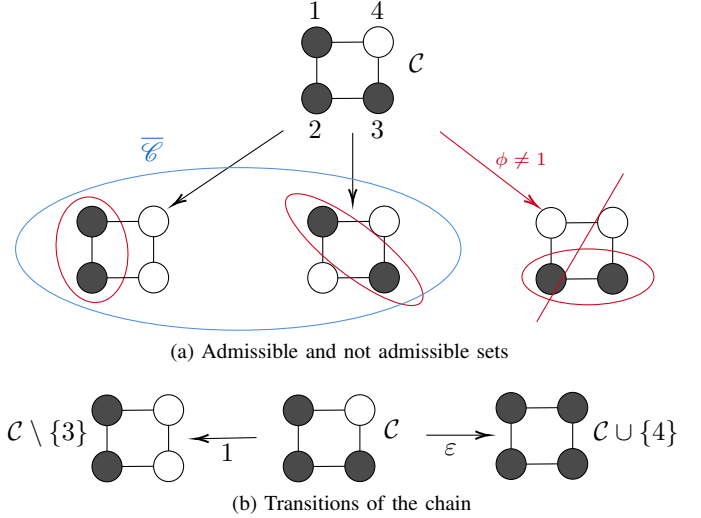


Fig. 4: Example of one iteration of the Markov chain from set  $\mathcal{C}$ . In (a), we highlight with a blue circle the set of admissible control sets. The red circles highlight all the candidate control sets. The last control set is not admissible and so it is not considered. In (b), we illustrate the transitions of the chain, described in Eq. (22). Nodes in the control set are denoted in black, nodes not in the control set in white.

$\mathcal{C}$ , since  $\mathcal{C} \setminus \{1\}$  is not admissible. Finally, if node  $r = 4$  is selected, we are in step 2) and the chain transitions to  $\mathcal{C} \cup \{4\}$  with probability  $\varepsilon$ , as illustrated in Fig. 4b.

**Proposition 6.** *The procedure described in Algorithm 2 induces a discrete-time Markov chain  $Z_\varepsilon(t)$ , which is defined on the space of admissible control sets  $\mathcal{C}$ , has initial state  $Z_\varepsilon(t) = \mathcal{V}^*$ , and, given any pair  $\mathcal{A}, \mathcal{B} \in \mathcal{C}$ , its transition probabilities are defined as*

$$P[Z_\varepsilon(t+1) = \mathcal{B} | Z_\varepsilon(t) = \mathcal{A}] = P_{\mathcal{A}, \mathcal{B}, \varepsilon}, \quad (21)$$

where

$$P_{\mathcal{A}, \mathcal{B}, \varepsilon} = \begin{cases} 1/n^* & \text{if } \mathcal{B} \subset \mathcal{A} \text{ and } |\mathcal{B}| = |\mathcal{A}| - 1, \\ \varepsilon/n^* & \text{if } \mathcal{B} \supset \mathcal{A} \text{ and } |\mathcal{B}| = |\mathcal{A}| + 1, \\ 1 - \alpha_\varepsilon(\mathcal{A}) & \text{if } \mathcal{B} = \mathcal{A}, \\ 0 & \text{otherwise,} \end{cases} \quad (22)$$

with  $\alpha_\varepsilon(\mathcal{A}) = \frac{\varepsilon(n^* - |\mathcal{A}|) + n_c(\mathcal{A})}{n^*}$ , where  $n_c(\mathcal{A})$  is the number of admissible configuration that can be reached by removing a node from  $\mathcal{A}$ , i.e.,  $n_c(\mathcal{A}) := |\{\mathcal{C} \in \mathcal{C} : \mathcal{C} = \mathcal{A} \setminus \{r\}, r \in \mathcal{A}\}|$ .

*Proof.* First, we observe that the iterative procedure described in Algorithm 2 explores only admissible control set. We proceed by induction. If Problem 2 is feasible, then we start from  $\mathcal{C} = \mathcal{V}^*$ , which is feasible. Then, if at the  $k$ th iteration the set  $\mathcal{C}$  is an admissible control set, we demonstrate that this holds true also at iteration  $k+1$ . In fact, if 1) occurs, then either  $\mathcal{C} \setminus \{r\}$  is admissible by construction or the control set remains unchanged; if 2) occurs, then the control set is updated to a superset of  $\mathcal{C}$  (possibly coinciding with  $\mathcal{C}$  with probability  $1 - \varepsilon$ ), which is admissible due to Lemma 1, yielding that Algorithm 2 induces a stochastic process with state space  $\mathcal{C}$ .

Second, let us denote by  $\mathcal{A}$  and  $\mathcal{B}$  the control set at the  $k$ th and  $(k+1)$ th iteration of Algorithm 2, respectively. Preliminary, we observe that, given  $\mathcal{A}, \mathcal{B}$  is independent



of the previous history of the process; hence, it is a Markov chain [39]. Then, if node  $r$  (selected uniformly at random among  $n^*$  nodes) is such that  $r \in \mathcal{A}$ , then  $\mathcal{B} = \mathcal{A} \setminus \{r\}$ . Hence, a generic set  $\mathcal{B} \subset \mathcal{A}$  with  $|\mathcal{B}| = |\mathcal{A}| - 1$  is reached with probability  $1/n^*$ , yielding the first line in Eq. (22). If the node  $r \notin \mathcal{A}$ , then  $\mathcal{B} = \mathcal{A} \cup \{r\}$  with probability  $\varepsilon$ . Hence, a generic set  $\mathcal{B} \supset \mathcal{A}$  with  $|\mathcal{B}| = |\mathcal{A}| + 1$  is reached with probability  $\varepsilon/n^*$ , yielding the second line in Eq. (22). No other state can be reached according to the algorithm. Therefore, we conclude that these probabilities match exactly those in Eq. (22), where the third line is simply obtained as the probability of the complementary event, yielding the claim.  $\square$

The Markov chain defined in Proposition 6 plays a key role in solving Problem 2, as claimed in the following statement.

**Theorem 4.** *Algorithm 2 induces a Markov chain whose invariant distribution  $\mu_\varepsilon \in [0, 1]^{\mathcal{C}}$  is such that  $\lim_{\varepsilon \searrow 0} \mu_\varepsilon = \mu$  where  $\mu$  is the uniform probability distribution on the set of solutions of Eq. (10) and, consequently, of Problem 2.*

*Proof.* We observe that the Markov chain  $Z_\varepsilon(t)$  with transition probabilities in Eq. (22) is ergodic, since every admissible configuration  $\mathcal{B} \in \mathcal{C}$  can be reached from any other one  $\mathcal{A} \in \mathcal{C}$  following a path of non-zero probability (trivially, first by adding nodes to  $\mathcal{A}$  until reaching  $\mathcal{V}^*$ , and then by removing nodes until reaching  $\mathcal{B}$ ) [39]. This path passes only through admissible control sets, due to Lemma 1. Hence, the chain converges to an invariant distribution.

To compute its invariant distribution, we follow the arguments from [30, Theorem 2], to conclude that  $Z_\varepsilon(t)$  has invariant distribution  $\mu_\varepsilon \in [0, 1]^{\mathcal{C}}$ , such that its generic component associated with  $\mathcal{C} \in \mathcal{C}$  is equal to  $[\mu_\varepsilon]_{\mathcal{C}} = \frac{1}{K_\varepsilon} \varepsilon^{|\mathcal{C}|}$ , where  $K_\varepsilon$  is a normalizing coefficient. Hence, for  $\varepsilon \searrow 0$ , it holds that the only non-zero components of  $\mu_\varepsilon$  are all equal and are those associated with the admissible control sets  $\mathcal{C}$  of minimal cardinality, yielding the claim.  $\square$

**Remark 6.** *The convergence result proved in Theorem 4 guarantees that, in the limit  $t \rightarrow \infty$ , the Markov chain defined in Proposition 6 concentrates about the solution(s) of Problem 2. From a practical point of view, in Algorithm 2 we incorporate a variable  $\hat{\mathcal{C}}$  to keep track of the minimal control set reached so far, so that it could be used also as an heuristics for admissible configurations space exploration.*

**Remark 7.** *The computations required to solve Algorithm 2 may be reduced. First, the most computationally complex operation is associated with the solution of Eq. (13) in Algorithm 1. However, this operation, which is performed by using matrix  $\mathbf{M}$ , is independent of the iteration. Hence, Algorithm 2 can be optimized by computing the matrix  $\mathbf{M}$  before starting the iterations and providing it as an input to Algorithm 1, so that the inversion of a matrix needs to be performed only once. Second, if  $\mathcal{V}^Y = \emptyset$ , i.e., we enforce action control, then in order to verify that a control set obtained by removing a node  $r$  from an admissible control set is admissible, we need not apply Algorithm 1 thoroughly. In fact, it is enough to stop the iteration in Algorithm 1 as soon as  $\delta_r(\hat{\mathbf{x}}, \hat{\mathbf{y}}) > 0$ , due to Lemma 1 and the monotonicity property of actions  $\mathbf{x}$ .*

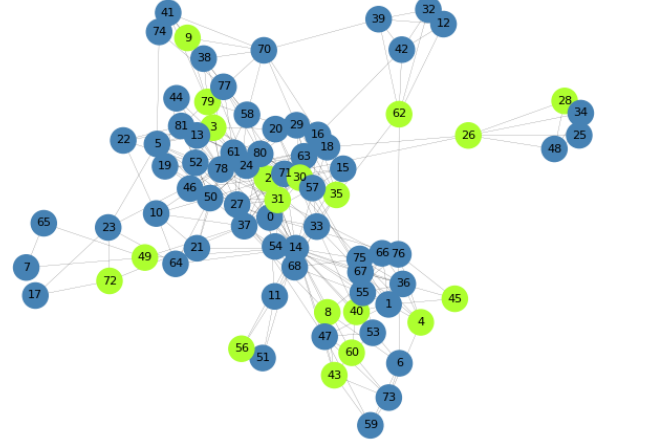


Fig. 5: The network used in Section VI. Green nodes are those identified by our algorithm as control nodes, in the scenario of joint control.

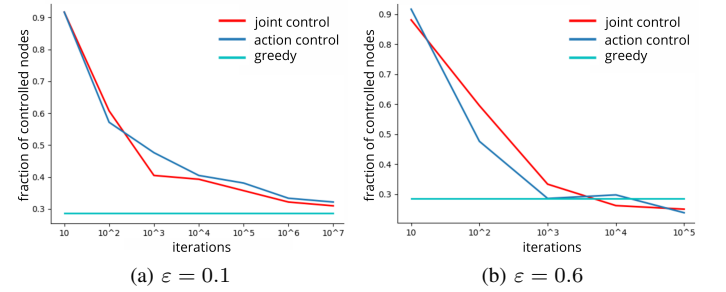


Fig. 6: Comparison of control strategies for different values of  $\varepsilon$ . Blue and red curves represent the fraction of agents to be controlled with joint and action control, respectively. In cyan, the results obtained with a greedy heuristics.

## VI. CASE STUDY

We demonstrate our approach on a real-world case study of a network associated with social contacts in a village in rural Malawi, whose dataset is available on Sociopatterns [32]. We use the largest connected component of the network, consisting of 84 individuals and 346 weighted undirected edges, illustrated in Fig. 5. We apply Algorithm 2 with the exploration parameter  $\varepsilon$  set to be  $\varepsilon = 0.1$  and  $\varepsilon = 0.6$ , and we keep track of how the quality of the optimal solution (in terms of fraction of nodes to be controlled) evolves as the number of iterations increases. The results are reported in Fig. 6.

First, we observe that increasing the number of iterations of the algorithm significantly reduces the fraction of nodes that need to be controlled. As expected, in the long run, joint control outperforms action control, i.e., controlling only the  $x$  values. However, since the performance difference is not substantial, controlling only the agents' actions, which one expects would require a weaker enforcement policy, could be a valid choice to reduce costs. On the contrary, opinion control is not sufficient to steer the system to the desired consensus.

Second, a remarkable result is highlighted in Fig. 6. Specifically, for a sufficiently large value of  $\varepsilon$  and number of algorithm iterations, Algorithm 2 outperforms a greedy heuristics algorithm that selects nodes with larger Bonacich centrality as control nodes. For instance, given  $\varepsilon = 0.6$ , after  $10^5$  algorithm

iterations the minimum control set cardinality required is reduced from 26.2% (in case of greedy algorithm, i.e cyan line in Fig. 6-(b)) to 23.8%. Moreover, the computational cost is not prohibitive since it required only 70s on a standard PC. This underlines not only the effectiveness of our algorithm, but also the importance of fine-tuning of its parameters based on the specific case study at hand, as the optimal parameters may vary significantly depending on the context.

## VII. CONCLUSION

In this paper, we formalized a novel control problem for social networks by incorporating a committed minority in a coevolutionary model of actions and opinions. By analyzing the controlled model, we established a general convergence result and we leverage it to tackle two research problems: i) determine whether the committed minority are able to steer the population to the desired state and ii) identify the minimal control set needed to achieve the goal. By developing algorithms to address these two questions, we offer a novel set of effective tools to assess the robustness of social systems against malicious attacks and assist policy makers in designing policies to promote social change.

The promising results presented in this paper outline several lines of future research. First, while this paper focuses on static control policies, one could conjecture that, after a critical mass is reached, one may uplift the control action. Consistently, future research may focus on designing dynamic control policies, towards minimizing the total control effort. Second, this paper focuses on the problem of steering a population to a consensus. In some applications, however, policy makers may want to favor diversity by reaching a non-consensus state. Extending our methodology to different control objectives is thus a key objective of our future research. Third, the improved performance of Algorithm 2 in identifying the optimal control set compared to classical heuristics suggests that our approach can be possibly extended to similar control problems for other multi-dimensional supermodular games.

## REFERENCES

- [1] N. E. Friedkin, "The Problem of Social Control and Coordination of Complex Systems in Sociology: A Look at the Community Cleavage Problem," *IEEE Control Syst. Mag.*, vol. 35, no. 3, pp. 40–51, 2015.
- [2] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part I," *Annu. Rev. Control*, vol. 43, pp. 65–79, 2017.
- [3] —, "A tutorial on modeling and analysis of dynamic social networks. Part II," *Annu. Rev. Control*, vol. 45, pp. 166–190, 2018.
- [4] A. Fontan and C. Altafini, "Multiequilibria analysis for a class of collective decision-making networked systems," *IEEE Trans. Control. Netw. Syst.*, vol. 5, no. 4, p. 1931–1940, 2018.
- [5] G. De Pasquale and M. E. Valcher, "Consensus for clusters of agents with cooperative and antagonistic relationships," *Automatica*, vol. 135, p. 110002, Jan. 2022.
- [6] A. Bizyaeva, A. Franci, and N. E. Leonard, "Nonlinear opinion dynamics with tunable sensitivity," *IEEE Trans. Autom. Control*, vol. 68, no. 3, p. 1415–1430, 2023.
- [7] C. Bernardo, C. Altafini, A. Proskurnikov, and F. Vasca, "Bounded confidence opinion dynamics: A survey," *Automatica*, vol. 159, p. 111302, 2024.
- [8] S. Gavrillets and P. J. Richerson, "Collective action and the evolution of social norm internalization," *Proc. Natl. Acad. Sci. USA*, vol. 114, no. 23, pp. 6068–6073, 2017.
- [9] B. Lindström, S. Jangard, I. Selbing, and A. Olsson, "The Role of a 'Common Is Moral' Heuristic in the Stability and Change of Moral Norms," *J. Exp. Psychol. Gen.*, vol. 147, no. 2, p. 228, 2018.
- [10] A. C. R. Martins, "Continuous opinions and discrete actions in opinion dynamics problems," *Int. J. Mod. Phys. C*, vol. 19, pp. 617–624, 2008.
- [11] F. Ceragioli and P. Frasca, "Consensus and disagreement: The role of quantized behaviors in opinion dynamics," *SIAM J. Control Optim.*, vol. 56, no. 2, pp. 1058–1080, 2018.
- [12] K. Tang, Y. Zhao, J. Zhang, and J. Hu, "Synchronous coda opinion dynamics over social networks," in *40th Chinese Control Conf.*, 2021, pp. 5448–5453.
- [13] D. Centola, R. Willer, and M. Macy, "The emperor's dilemma: A computational model of self-enforcing norms," *Am. J. Sociol.*, vol. 110, no. 4, pp. 1009–1040, 2005.
- [14] R. Willer, K. Kuwabara, and M. W. Macy, "The False Enforcement of Unpopular Norms," *Am. J. Sociol.*, vol. 115, pp. 451–490, 2009.
- [15] L. Zino, M. Ye, and M. Cao, "A two-layer model for coevolving opinion dynamics and collective decision-making in complex social systems," *Chaos*, vol. 30, no. 8, p. 083107, 2020.
- [16] H. D. Aghbolagh, M. Ye, L. Zino, Z. Chen, and M. Cao, "Coevolutionary dynamics of actions and opinions in social networks," *IEEE Trans. Autom. Control*, vol. 68, no. 12, pp. 7708–7723, 2023.
- [17] A. Montanari and A. Saberi, "The spread of innovations in social networks," *Proc. Natl. Acad. Sci. USA*, vol. 107, no. 47, pp. 20196–20201, 2010.
- [18] M. O. Jackson and Y. Zenou, "Games on Networks," in *Handbook of Game Theory with Economic Applications*. Elsevier, 2015, vol. 4, ch. 3, pp. 95–163.
- [19] P. Guo, Y. Wang, and H. Li, "Algebraic formulation and strategy optimization for a class of evolutionary networked games via semi-tensor product method," *Automatica*, vol. 49, no. 11, pp. 3384–89, 2013.
- [20] J. R. Riehl and M. Cao, "Towards optimal control of evolutionary games on networks," *IEEE Trans. Automat. Contr.*, vol. 62, pp. 458–462, 2017.
- [21] J. Riehl, P. Ramazi, and M. Cao, "Incentive-based control of asynchronous best-response dynamics on binary decision networks," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 2, pp. 727–736, 2018.
- [22] N. Quijano *et al.*, "The role of population games and evolutionary dynamics in distributed control systems: The advantages of evolutionary game theory," *IEEE Control Syst. Mag.*, vol. 37, pp. 70–97, 2017.
- [23] T. Başar, "Inducement of desired behavior via soft policies," *Int. Game Theory Rev.*, p. 2440002, 2024.
- [24] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate," *Automatica*, vol. 50, no. 12, pp. 3209–3215, 2014.
- [25] Y. Yi, T. Castiglia, and S. Patterson, "Shifting opinions in a social network through leader selection," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 3, pp. 1116–1127, 2021.
- [26] L. Wang, C. Bernardo, Y. Hong, F. Vasca, G. Shi, and C. Altafini, "Consensus in concatenated opinion dynamics with stubborn agents," *IEEE Trans. Autom. Control*, vol. 68, no. 7, pp. 4008–4023, 2023.
- [27] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the Spread of Influence through a Social Network," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2003, pp. 137–146.
- [28] D. Centola, J. Becker, D. Brackbill, and A. Baronchelli, "Experimental evidence for tipping points in social convention," *Science*, vol. 360, no. 6393, pp. 1116–1119, 2018.
- [29] M. Ye *et al.*, "Collective patterns of social diffusion are shaped by individual inertia and trend-seeking," *Nat. Comm.*, vol. 12, p. 5698, 2021.
- [30] G. Como, S. Durand, and F. Fagnani, "Optimal targeting in supermodular games," *IEEE Trans. Autom. Control*, vol. 67, no. 12, pp. 6366–6380, 2022.
- [31] C. M. Schneider, A. A. Moreira, J. S. Andrade, S. Havlin, and H. J. Herrmann, "Mitigation of malicious attacks on networks," *Proc. Natl. Acad. Sci. USA*, vol. 108, no. 10, p. 3838–3841, 2011.
- [32] L. Ozella *et al.*, "Using wearable proximity sensors to characterize social contact patterns in a village of rural Malawi," *EPJ Data Sci.*, vol. 10, no. 1, 2021.
- [33] R. Raineri, G. Como, F. Fagnani, M. Ye, and L. Zino, "On controlling a coevolutionary model of actions and opinions," *Proc. 63rd IEEE Conf. Decis. Control*, pp. 4550–4555, 2024.
- [34] D. M. Topkis, *Supermodularity and Complementarity*. Princeton University Press, 1998.
- [35] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Trans. Syst. Man Cybern. B*, vol. 39, no. 6, pp. 1393–1407, 2009.

- [36] W. H. Sandholm, *Population Games and Evolutionary Dynamics*. Cambridge University Press, 2010.
- [37] R. G. Bartle, *The Elements of Real Analysis*. Wiley, 1976.
- [38] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Math. Program.*, vol. 14, no. 1, p. 265–294, 1978.
- [39] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*. Providence RI, US: American Mathematical Society, 2006.

## APPENDIX

### A. Proof of Theorem 2

To prove that Problem 2 is NP-complete, we first show that it belongs to the NP-class. Then, we show that an instance of the problem is NP-hard, which implies NP-completeness.

**Lemma 2.** *Problem 2 is NP.*

*Proof.* In order to prove that Problem 2 is NP, we need to demonstrate that, given an instance of the coevolutionary dynamics and a control set  $(\mathcal{C}^X, \mathcal{C}^Y)$ , we can check whether  $(\mathcal{C}^X, \mathcal{C}^Y)$  is a feasible solution of Eq. (9) in a polynomial time. In other words, if there exists an algorithm able to solve Problem 1 in polynomial time, then clearly one can determine whether  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$ , while checking whether  $\mathcal{C}^X$  is a subset of  $\mathcal{V}^X$  and  $\mathcal{C}^Y$  of  $\mathcal{V}^Y$  can be easily checked in time  $O(n)$  using, for instance, an hash function. To prove that there exists an algorithm to solve Problem 1 in polynomial time, we refer to Theorem 3, which proves that Algorithm 1 solves Problem 1 in time  $O(n^3)$ , yielding the claim.  $\square$

**Lemma 3.** *Determining the minimal control set for a best-response dynamics for a majority game is NP-hard, and is an instance of Problem 2.*

*Proof.* The majority game on a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with  $n$  agents is a game with binary action set  $\mathcal{A} = \{-1, 1\}$ , action profile  $\tilde{x} \in \{-1, 1\}^n$ . The utility that  $i$  would get for selecting action  $\tilde{c}$  is given by

$$u_i(\tilde{x}) = \frac{1}{2|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} ((1 - \tilde{x}_j)(1 - \tilde{c}) + (1 + \tilde{x}_j)(1 + \tilde{c})). \quad (23)$$

If set  $\tilde{\mathcal{C}}$  is controlled, then  $i \in \mathcal{C}$  has  $\tilde{x}_i(t) = +1$  for all  $t \geq 0$ ; while an individual  $i \in \mathcal{V} \setminus \mathcal{C}$  is selected uniformly at random at each time step to revise their action according to a best-response dynamics, ultimately yielding

$$\tilde{x}_i(t+1) = \begin{cases} +1 & \text{if } n_i^+(t) \geq \frac{1}{2} \\ -1 & \text{otherwise,} \end{cases} \quad (24)$$

where  $n_i^+(t) := \frac{1}{|\mathcal{N}_i|} |\{j \in \mathcal{N}_i : \tilde{x}_j(t) = +1\}|$ . Determining the minimal control set for a best-response dynamics for a majority game consists in determining the set  $\tilde{\mathcal{C}}$  that minimizes  $|\tilde{\mathcal{C}}|$  such that all individuals eventually reaches  $\tilde{x}_i = +1$  according to Eq. (24).

Let us map  $x_i = \tilde{x}_i$  and consider a two-layer network  $\mathcal{G} = (\mathcal{V}, \mathcal{E}_A, \mathbf{A}, \mathcal{E}_W, \mathbf{W})$  with  $a_{ij} = \frac{1}{|\mathcal{N}_i|}$  if  $i \in \mathcal{N}_i$  and  $a_{ij} = 0$  otherwise, and any  $\mathbf{W}$  stochastic and irreducible. Then, if we let  $\lambda_i = 1$  for all  $i \in \mathcal{V}$ , for any value of  $\beta_i$ , Proposition 2 yields exactly Eq. (24) for the update rule of  $x_i$ , while we get  $y_i(t+1) = x_i(t+1)$ . Hence, a set  $\tilde{\mathcal{C}}^X$  is a minimal control set for the majority game if and only if  $(\tilde{\mathcal{C}}^X, \emptyset)$  solves Problem 2 in the setting described in the above.

Finally, the fact that determining the minimal control set for a best-response dynamics for a majority game is NP-hard is well-known [30], yielding the claim.  $\square$

### B. Proof of Theorem 3

We start by proving the following result.

**Lemma 4.** *Given control sets  $(\mathcal{C}^X, \mathcal{C}^Y)$  and a vector  $\hat{x} \in \{-1, 1\}^n$ , let  $\hat{y}$  be the solution of Eq. (13) given  $\hat{x}$ . Then, configuration  $\hat{z} = (\hat{x}, \hat{y})$  is an equilibrium for the controlled coevolutionary dynamics under Assumptions 1–2 iff  $\hat{x}_i \delta_i(\hat{z}) \geq 0$ ,  $\forall i \notin \mathcal{C}^X$ . If there exists at least an individual  $i \notin \mathcal{C}^X$  with  $\hat{x}_i \delta_i(\hat{z}) < 0$ , there exist no equilibria with  $x = \hat{x}$ .*

*Proof.* Fixed the action vector  $\hat{x}$  and the opinion of  $j \in \mathcal{C}^Y$  to  $\hat{y}_j = 1$  (because of Assumption 2), the dynamics in Eq. (4b) for a generic individual  $i \notin \mathcal{C}^Y$  reduces to

$$\begin{aligned} y_i(t+1) &= (1 - \lambda_i) \left[ \sum_{j \notin \mathcal{C}^Y} w_{ij} y_j(t) + \sum_{j \in \mathcal{C}^Y} w_{ij} \hat{y}_j \right] + \lambda_i \hat{x}_i \\ &= (1 - \tau_i) \sum_{j \notin \mathcal{C}^Y} \bar{w}_{ij} y_j(t) + \tau_i u_i, \end{aligned} \quad (25)$$

with  $\tau_i = 1 - (1 - \lambda_i) \sum_{j \notin \mathcal{C}^Y} w_{ij}$ ,  $\bar{w}_{ij} = w_{ij} / \sum_{j \notin \mathcal{C}^Y} w_{ij}$  (with the convention that,  $\bar{w}_{ij} = 0$  if  $w_{ij} = 0$ ), and  $u_i = \frac{(1 - \lambda_i)(1 - \sum_{j \notin \mathcal{C}^Y} w_{ij}) + \lambda_i \hat{x}_i}{1 - (1 - \lambda_i) \sum_{j \notin \mathcal{C}^Y} w_{ij}}$ . This can be seen as the update rule of a Friedkin–Johnsen opinion dynamics model [2], which is known to converge under Assumption 1 to the unique solution of  $\hat{y}_i = (1 - \tau_i) \sum_{j \notin \mathcal{C}^Y} \bar{w}_{ij} \hat{y}_j + \tau_i u_i$ , which coincides with Eq. (13). See, [3] for more details. Hence,  $\hat{z} = (\hat{x}, \hat{y})$  is the only admissible equilibrium with action vector  $\hat{x}$ .

Now, we observe that  $\hat{z}$  is an equilibrium iff there are no individuals that would change their action according to Eq. (4a) when the system is at  $\hat{z}$ . This corresponds to verify that all individuals  $i \notin \mathcal{C}^X$  with  $\hat{x}_i = -1$  have  $\delta_i(\hat{z}) \leq 0$ , and all those with  $\hat{x}_i = 1$  have  $\delta_i(\hat{z}) \geq 0$ . In fact, if there exists  $i \notin \mathcal{C}^X$  with  $\hat{x}_i = -1$  and  $\delta_i(\hat{z}) > 0$ , then Assumption 1 guarantees that within a finite time-window,  $i$  activates and flips action to +1 (being  $\delta_i(\hat{z}) > 0$ ).  $\square$

Now, we use Lemma 4 to prove the following result.

**Lemma 5.** *The equilibrium reached by a controlled coevolutionary dynamics that satisfies Assumptions 1 and 2 with control sets  $(\mathcal{C}^X, \mathcal{C}^Y)$  is  $(x^*, y^*)$ , with  $x^*$  defined as in Eq. (12) with  $\mathcal{A}(k) = \mathcal{A}_f$  (output of Algorithm 1) and  $y^*$  solution of Eq. (13) given  $x^*$ .*

*Proof.* In the first iteration of the algorithm ( $k = 1$ ), Lemma 4 establishes that state  $\hat{z}$  defined using Eq. (12) with set  $\mathcal{A}(1)$  and Eq. (13) is an equilibrium iff  $\mathcal{A}(2) = \mathcal{A}(1)$ . Otherwise, we will now prove that individuals in  $\mathcal{A}(2) \setminus \mathcal{A}(1)$  will eventually switch action to +1. In fact, as long as  $x(t) = \hat{x}$ , then  $y_j(t)$  converges asymptotically to  $\hat{y}_j$  for all  $j \notin \mathcal{C}^Y$  (due to the observations made in the proof of Lemma 4). Hence,  $\delta_i(z(t))$  converges asymptotically to  $\delta_i(\hat{z}) > 0$ . By continuity,  $\exists \bar{t}$  such that  $\delta_i(z(t)) > 0$  for all  $t \geq \bar{t}$ , as long as  $x(t) = \hat{x}$ . Moreover, since  $\delta_i(z(t))$  is monotonically increasing in  $z$  and  $z(t)$  is monotonically increasing in  $t$ , then  $\delta_i(z(t))$  is a monotonically increasing function of time. This implies that  $\delta_i(z(t)) > 0$

for all  $t \geq \tilde{t}$ . This, together with Assumption 1, guarantees that  $i$  switches to  $+1$  (Proposition 2) and cannot switch back (Theorem 1), then  $x_i(t) = +1$  for all  $t \geq \tilde{t} + T$ .

If  $\mathcal{A}(2) = \mathcal{A}(1)$ , then the system necessarily converges to the equilibrium  $\hat{z}$ , yielding the claim. Otherwise,  $\hat{z}$  is not an equilibrium. In this case, all individuals in  $\mathcal{A}(2) \setminus \mathcal{A}(1)$  will necessarily switch action to  $+1$  in finite time. Hence, we re-iterate considering the set  $\mathcal{A}(2)$  and computing the corresponding  $\hat{z}$ , observing that, if  $i$  has switched to  $+1$ , Theorem 1 guarantees that  $i$  will never switch back, so we just need to check whether all  $i \in \mathcal{V} \setminus \mathcal{A}(2)$  have  $\delta_i(\hat{z}) \leq 0$  to get the terminal condition  $\mathcal{A}(3) = \mathcal{A}(2)$ , for which the system necessarily converges to the equilibrium  $\hat{z}$ . Otherwise, we re-iterate the process. Finally, in each iteration  $k$  in which the terminal condition is not met, the size of  $\mathcal{A}(k)$  increases by at least 1, implying that within at most  $k = n - |\mathcal{C}^X|$  iterations we would get  $\mathcal{A}(k) = \mathcal{V}$ , for which  $\hat{z} = (\mathbf{1}, \mathbf{1})$  is a trivial equilibrium, terminating the algorithm.  $\square$

Theorem 1 implies that a controlled coevolutionary dynamics always converge to an equilibrium. Lemma 5 implies that the equilibrium is independent of the activation sequence, but depends only on the model parameters and on the initial condition, which are determined by  $\mathcal{C}^X$  and  $\mathcal{C}^Y$ . Hence, fixed the parameters and given the  $\mathcal{C}^X$  and  $\mathcal{C}^Y$  either the system converges to  $x = \mathbf{1}$ , implying  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 1$  or to any other equilibrium, implying  $\phi(\mathcal{C}^X, \mathcal{C}^Y) = 0$ .

Finally, observe that Eq. (13) can be rewritten as  $\hat{\mathbf{y}} = (\mathbf{I} - (\mathbf{I} - \text{diag}(\boldsymbol{\lambda}))\mathbf{W})^{-1} \text{diag}(\boldsymbol{\lambda})\hat{\mathbf{x}}$ . The matrix  $\mathbf{M} := (\mathbf{I} - [(\mathbf{1} - \boldsymbol{\lambda})]\mathbf{W})^{-1}$  does not depend on the iteration step, thus it can be computed once at the beginning of the iterations (such procedure requires  $O(n^3)$  operations). Then, at each iteration of Algorithm 1, the dominant operation is the computation of  $\hat{\mathbf{y}}$  which requires  $O(n^2)$  operations. Since the number of iterations is at most  $n - |\mathcal{C}^X|$ , the total computational complexity of Algorithm 1 is  $O(n^3)$ , yielding the claim.