

# Altruism Design in Networked Public Goods Games

Sixie Yu<sup>1</sup>, David Kempe<sup>2</sup> and Yevgeniy Vorobeychik<sup>1</sup>

<sup>1</sup>Washington University in St. Louis

<sup>2</sup>University of Southern California

{sixie.yu,yvorobeychik}@wustl.edu, david.m.kempe@gmail.com

## Abstract

Many collective decision-making settings feature a strategic tension between agents acting out of individual self-interest and promoting a common good. These include wearing face masks during a pandemic, voting, and vaccination. Networked public goods games capture this tension, with networks encoding strategic interdependence among agents. Conventional models of public goods games posit solely individual self-interest as a motivation, even though altruistic motivations have long been known to play a significant role in agents' decisions. We introduce a novel extension of public goods games to account for altruistic motivations by adding a term in the utility function that incorporates the perceived benefits an agent obtains from the welfare of others, mediated by an altruism graph. Most importantly, we view altruism not as immutable, but rather as a lever for promoting the common good. Our central algorithmic question then revolves around the computational complexity of modifying the altruism network to achieve desired public goods game investment profiles. We first show that the problem can be solved using linear programming when a principal can fractionally modify the altruism network. While the problem becomes in general intractable if the principal's actions are all-or-nothing, we exhibit several tractable special cases.

## 1 Introduction

Individuals in a collective decision-making environment often experience the following type of scenario. Each individual can decide whether or how much effort to invest for the common good; many others may benefit from the efforts, but the cost of the investment is incurred by the individual. Examples of such scenarios include decisions whether or not to wear a mask in a pandemic, vaccinate, or invest in security. The outcomes of such scenarios are often highly sub-optimal from a societal point of view: mask-wearing suggestions are flaunted, societies remain undervaccinated, and security measures are not taken. At the heart of this breakdown is that while individuals are “connected” in the sense that

their actions affect one another, they are often disconnected “socially,” in the sense that they do not experience the utility gain/loss of those affected by their actions. In economic terms, actions have *externalities* on other players, which are, by definition, not *internalized*.

Indeed, the outcomes of such scenarios tend to be significantly different when the individuals form a more tightly knit community. Within families, groups of friends, or small villages, individuals frequently take actions, at a cost to themselves, which primarily benefit others. Similarly, societies with a stronger sense of “duty” towards fellow citizens tend to witness more compliance in all of the above-mentioned examples. Not surprisingly then, campaigns to encourage individual effort (e.g., “Wear a mask — save a life!”) tend to appeal to notions of altruism and duty, attempting to get individuals to internalize some of their externalities, if only psychologically.

If the goal of campaigns is to encourage altruistic behavior, an important question is what type of campaign is most effective. Should a principal, aiming to achieve a societally desirable outcome, try to appeal to a generic sense of “duty towards your fellow citizens,” try to strengthen the social ties within a small neighborhood, or focus on building a few strong ties between some key individuals? Can the question of how best to *build* or *change* altruism in a society be approached algorithmically, and are the resulting questions tractable or intractable? This is the high-level question we investigate in the present paper.

The question of how to build altruism networks is meaningful in a variety of strategic settings. We focus on *networked public goods games* [Bramoullé and Kranton, 2007; Bramoullé *et al.*, 2014; Feldman *et al.*, 2013; Yu *et al.*, 2020], motivated by the real-world scenarios discussed earlier (e.g., encouraging mask wearing). In networked public goods games, the benefits of an individual's effort are reaped by those with whom the individual interacts, encoded by a network on the individuals.<sup>1</sup> Specifically, an individual's utility depends on 1) her own investment decision, and 2) the *aggregate* investment from her direct neighbors in the network.

In most conventional models of games, including public goods games, it is assumed that agents are driven solely by

<sup>1</sup>Public goods games can be viewed as the special case in which the network is complete.

their individual interests. This assumption is nearly always violated in behavioral studies of public goods games [Ledyard, 1997; Levine, 1998]. While there are many different ways to model altruistic behavior, one natural way was proposed by Ledyard [1997]: the utility of a player  $i$  is a linear combination of an *egocentric utility* term, which is the direct benefit to  $i$ , and an *altruistic* term, which is a sum of egocentric utilities of other players  $j$ , weighted by the strength  $a_{i,j}$  of altruism that  $i$  feels for  $j$ . In most prior work of this kind,  $a_{i,j}$  was modeled as a constant for all players  $i$  and  $j$ . A more general variation by Meier *et al.* [2008] considered an altruism network in vaccination games, but assumed that the altruism graph is identical with the graph representing strategic dependence, as well as that altruism weights are identical for all edges. Naturally, many settings call for more fine-grained models in which the weights can be different: for example, parents typically care more about their children’s welfare than that of strangers.

While the focus of past work on altruism in games has been on its equilibrium effects, our point of departure is to consider the altruism network itself as (partially) under the control of a principal. In other words, we view altruistic motivations as a lever that can be adjusted to promote the common good, for example, through public outreach campaigns, community meetings, and personal introductions. Specifically, we propose a model of modifying altruism networks, with the goal of inducing a target investment profile by the agents. We consider three variants of the altruism network: weighted, directed, and undirected. We show that even for very complex available actions, the problem can be solved efficiently using linear programming when the principal has fine-grained control over the extent to which actions are taken. When the principal can only control *which* actions are taken, the problem becomes NP-complete, even when each action affects only a single edge in the altruism network. However, when the altruism network is directed, we show that the problem is tractable in a broad array of special cases by reductions to the (tractable cases of) the KNAPSACK problem. We also leverage this connection to exhibit an FPTAS for the general case. When the altruism network is undirected, the hardness results apply even for much more restrictive special cases. However, we show that the problem is tractable when the benefits from investment are linear and uniform, by a non-trivial reduction to the problem of NETWORK DESIGN FOR DEGREE SETS (NDDS) introduced by Kempe *et al.* [2020], who also showed that it can be solved in polynomial time.

Our problem of designing an altruism graph to achieve target equilibrium outcomes is, indeed, conceptually related to Kempe *et al.* [2020], who study the problem of designing the *strategic* network in networked public goods games. The main rationale for shifting focus to designing *altruism graphs* is that strategic networks are often difficult to change. For example, in a pandemic, it is difficult to directly affect contacts among individuals, as these are ultimately the products of individual choices (e.g., even lockdowns may be ineffective if individuals are non-compliant, except through levels of enforcement that are often viewed as unacceptable by the population). In contrast, it can be significantly easier to try to impact decisions indirectly by evoking altruistic motivations

in people. From a technical perspective, the problem of altruism design impacts utilities linearly, in contrast to the design of strategic networks; however, it is also distinct from the linear special case in Kempe *et al.* [2020], where the marginal impact of each neighbor on a player’s utility is identical, in contrast to altruism design, where these differ. An extended version of the paper with complete proofs is available at: <https://arxiv.org/abs/2105.00505>.

## 1.1 Related Work

Our work is related to four lines of research: graphical games, altruism modeling, mechanism and market design, and network design. Graphical games encode sparsity in the interdependence of player utility functions using a graph [Kearns *et al.*, 2001; Shoham and Leyton-Brown, 2008], with networked public goods games an important class of such models [Bramoullé and Kranton, 2007; Galeotti *et al.*, 2010; Grossklags *et al.*, 2008; Yu *et al.*, 2020]. A conventional assumption in such games is that agents act to exclusively promote their own interest. However, considerable experimental evidence exists that even games with this payoff structure elicit altruistic motivations among human subjects [Dong *et al.*, 2016; Levine, 1998]. This, in turn, led to a series of approaches to model altruism in a variety of games, including public goods games, which are of direct interest here [Ledyard, 1997; Dong *et al.*, 2016], inoculation games [Meier *et al.*, 2008], routing games [Chen and Kempe, 2008], and congestion games [Chen *et al.*, 2011; Chen *et al.*, 2014].

A typical way that altruism is captured in prior literature is by either adding a social welfare term to utility functions [Ledyard, 1997], or introducing a parameter that governs the extent to which agents care about their social network neighbors [Meier *et al.*, 2008]. Our model is distinct in that it allows altruism to be relationship-dependent, a property we model by an altruism network. Moreover, our goal is to *modify* an altruism network to achieve a target equilibrium (e.g., one that maximizes social welfare).

Mechanism and market design also aim to change the parameters of a game to induce desirable equilibrium outcomes [Nisan *et al.*, 2007; Haeringer, 2018; Dughmi, 2017]. We introduce altruism network design as a novel lever for aligning incentives with public good.

The last relevant line of research is network design. A particularly related thread in network design is to study the effects of network modification on equilibrium outcomes or welfare [Kempe *et al.*, 2020; Bramoullé and Kranton, 2007; Galeotti *et al.*, 2010]. Another related thread is to alter a network in order to effect a variety of different outcomes for different types of games [Sheldon *et al.*, 2010; Chen *et al.*, 2016; Ghosh and Boyd, 2006; Tong *et al.*, 2012; Bredereck and Elkind, 2017; Sina *et al.*, 2015; Matteo Castiglioni, 2020; Amelkin and Singh, 2019; Garimella *et al.*, 2018].

## 2 Networked Public Goods Games and Altruism

We study altruism in *binary networked public goods games* (BNPGs), which are an important variant of *public goods games* studied extensively in prior literature [Bramoullé and

Kranton, 2007; Galeotti *et al.*, 2010; Grossklags *et al.*, 2008; Suri and Watts, 2011; Dong *et al.*, 2016; Kempe *et al.*, 2020; Yu *et al.*, 2020]. We begin by formally describing BNPGs and pure strategy Nash equilibria, the solution concept we focus on. We then discuss a natural model of altruism in games, and its application to the specific case of BNPGs.

## 2.1 Binary Networked Public Goods Games

A *binary networked public goods* (BNPG) game is characterized by the following:

1. A simple, undirected, and loop-free graph  $H = (V, E_H)$  in which the nodes  $V = \{1, 2, \dots, n\}$  are the agents/players, and the edges  $E_H$  represent the interdependencies among the players' payoffs.
2. A binary strategy space  $\{0, 1\}$  for each player  $i$ .<sup>2</sup> We interpret the choice of strategy 1 as *investing* in a public good, while choosing 0 is interpreted as non-investment. The action of player  $i$  is denoted by  $x_i$ , and the joint pure strategy profile of all players by  $\mathbf{x} = (x_1, x_2, \dots, x_m)$ . We use  $\mathbf{x}_{-i}$  to denote a strategy profile that omits player  $i$ 's strategy.
3. A non-decreasing utility function  $U_i(x_i, \mathbf{x}_{\mathcal{N}_i^{(H)}})$  for each player  $i$ , where  $\mathcal{N}_i^{(H)} = \{j \mid (i, j) \in E_H\}$  is the set of  $i$ 's neighbors in the graph  $H$ .

As is common in the literature on public goods games [Bramoullé and Kranton, 2007], we assume that each player's (egocentric) utility function  $U_i$  only depends on the total investment by  $i$ 's network neighbors. To formalize this, we define  $n_i^{(H, \mathbf{x})} = \sum_{j \in \mathcal{N}_i^{(H)}} x_j$  as the number of  $i$ 's neighbors who invest under  $\mathbf{x}$ . We omit  $H$ ,  $\mathbf{x}$ , or both from this notation when they are clear from the context. Each player  $i$ 's utility function then has the following form:

$$U_i(\mathbf{x}) = U_i(x_i, n_i^{(\mathbf{x})}) = g_i(x_i, n_i^{(\mathbf{x})}) - c_i x_i. \quad (1)$$

The second term ( $-c_i x_i$ ) captures the cost incurred by player  $i$  from investing. As is standard in the public goods games literature, each  $g_i$  is assumed to be a non-negative and non-decreasing function of both of its arguments, capturing the positive externality that  $i$  experiences from her neighbors' (and her own) investments. Observe that each function  $g_i$  can be represented using  $O(n)$  values, so the entire BNPG game (including the graph structure) can be represented using  $O(n^2)$  values.

We will consider *pure-strategy Nash equilibria* (PSNE) of BNPGs. A pure strategy profile  $\mathbf{x}^*$  is a PSNE if for all  $i$ ,  $U_i(x_i, n_i^{(H, \mathbf{x}^*)}) \geq U_i(1 - x_i, n_i^{(H, \mathbf{x}^*)})$ . We write  $\mathcal{E}(\mathcal{G})$  for the set of all PSNEs of the game  $\mathcal{G}$ .

<sup>2</sup>The public goods game literature extensively considers both binary and continuous decisions, and both are natural candidates for considering network design. Here, we focus on binary strategies due to their applicability to decisions such as vaccinations or mask wearing.

## 2.2 Altruistic Motivations in BNPGs

A natural way to model other-regarding utilities is to define a player's utility as a linear combination of her egocentric utility, defined by Equation (1), and the egocentric utilities of other players.

To formalize this, we can think of the matrix  $A = (a_{i,j})_{i,j}$  as encoding an *altruism network*. This captures the central motivation of our work, discussed in the introduction: that the agents who are *affected* by the actions of  $i$  may not be the same as the agents that  $i$  *cares about*.<sup>3</sup> The resulting utility function of a player  $i$  in our BNPG game model with altruism is

$$U_i^{(A)}(\mathbf{x}) = g_i(x_i, n_i) - c_i x_i + \sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j} g_j(x_j, n_j). \quad (2)$$

We denote the BNPG with altruism network  $A$  by BNPG( $A$ ).

We note two points about this model: First, the altruistic term of player  $i$ 's utility does not include a term for the *investment cost* of player  $j$ , only the *utility*. This is inconsequential, as investment decisions by  $j$  are not under  $i$ 's control, so from  $i$ 's point of view, these cost terms are constants. Second, the other-regarding terms have no component in which  $j$ 's utility due to  $i$ 's egocentric payoff is recursively considered. Such utilities may be harder to observe, and from a modeling perspective, they can be transformed into the case we study here; see [Bergstrom, 1999].

In some of the results in later sections, we will specifically want to stress the *network* aspect of altruism. In that case, we will assume that we are given a (directed or undirected) *altruism graph*  $G$ , and that  $a_{i,i} = 1$  for all  $i$ ,  $a_{i,j} = a$  for all  $i, j$  for which  $G$  contains the edge  $(i, j)$ , and  $a_{i,j} = 0$  for all  $i, j$  for which  $G$  contains no edge. In other words, the altruism strength of the edges of  $G$  is uniform. When  $G$  is undirected, we will refer to the case as *symmetric altruism*; when  $G$  is directed, we call it *asymmetric altruism*.

## 3 Modifying Altruism Networks

As discussed in the introduction, a major problem with public goods situations is that equilibria can be far from optimal because individuals may not fully internalize the impact of their actions on others. Therefore, a principal who seeks to steer the network to a better equilibrium might wish for agents to consider others in their decisions. We consider situations in which the principal can increase or decrease the salience of others in these settings, for example, through introductions, advertising, or community meetings.<sup>4</sup> We now formally model this problem as modifying an altruism network to achieve socially desired outcomes.

Our model is as follows. The principal aims to induce a particular *target investment profile*  $\mathbf{x}^*$ . This allows us to cleanly capture a broad variety of design goals, such as maximizing welfare or achieving fairness, while focusing on the

<sup>3</sup>In the context of vaccination games, Meier *et al.* [2008] study the special case in which the friendship network is the *same* as the network of who may transmit a disease to whom.

<sup>4</sup>This is separate from, and in addition to, other channels, such as rewarding or punishing certain actions.

computational issues at the core of our specific problem of altruism design. To induce the outcome, the principal wants to (minimally) modify the altruism network  $A^{\text{in}}$  to  $A$  such that  $\mathbf{x}^*$  is a PSNE of the modified game  $\text{BNPG}(A)$ .<sup>5</sup>

As implied by the preceding discussions, the principal may have at his disposal a number of different actions, affecting the altruism between different sets of pairs of agents, in positive or negative ways. For example, a general appeal to watch out for one another may lead to a small increase in altruism between many pairs of individuals; a community meeting may lead to a stronger increase among a smaller subset, and a personal introduction may introduce one strong edge. We model such settings by assuming that there are  $K$  actions, with  $K$  polynomial in  $n$ . Each action  $k$  has associated with it a set  $\mathcal{S}^{(k)}$  of affected altruism edges, a cost  $\gamma^{(k)} \geq 0$ , and a sign  $\sigma^{(k)} \in \{-1, 1\}$ , which captures whether the action strengthens or weakens the edges in  $\mathcal{S}^{(k)}$ .<sup>6</sup> The edge set is encoded in the corresponding adjacency matrix  $M^{(k)}$ , with entries  $m_{i,j}^{(k)}$  which are 1 for  $(i, j) \in \mathcal{S}^{(k)}$  and 0 otherwise. The costs  $\gamma^{(k)}$  measure the monetary expense or effort/time needed to implement the corresponding activity, per unit of change in the altruism. The principal aims to solve the following problem:

**Definition 3.1** (Altruism Network Modifications (ANM)). *Given: an altruism network  $A^{\text{in}}$ , target investment profile  $\mathbf{x}^*$ , and actions  $\{\mathcal{S}^{(1)}, \dots, \mathcal{S}^{(K)}\}$  with signs  $\sigma^{(k)}$  and costs  $\gamma^{(k)}$ . Goal: choose a non-negative vector  $\mathbf{v} \in \mathbb{R}_{\geq 0}^K$  of minimum total cost  $\sum_{k=1}^K v_k \gamma^{(k)}$  such that the modified game with new altruism network*

$$A = A^{\text{in}} + \sum_{k=1}^K v_k \cdot \sigma^{(k)} \cdot M^{(k)} \quad (3)$$

has  $\mathbf{x}^*$  as a PSNE, i.e.,  $\mathbf{x}^* \in \mathcal{E}(\text{BNPG}(A))$ .

Here,  $\mathbf{v}$  is a vector capturing how much the principal spends on each of the available actions. We assume that the different actions are cumulative in their effects on any of the network's edges, and (partially) cancel out when they have opposite signs. Note that we could additionally truncate the entries of  $A$  so that  $0 \leq A \leq 1$ ; this is not consequential for our results, and we proceed with the slightly cleaner model above. The principal's spending on actions results in a modified altruism network, and his goal is to ensure that the target action profile  $\mathbf{x}^*$  becomes one of the equilibria of the modified game. The principal wants to achieve this goal at minimum cost (which is infinite if the problem is infeasible).

Since strategies are binary, a target profile  $\mathbf{x}^*$  can be equivalently represented by the set of agents who invest under this profile,  $\mathcal{I} = \mathcal{I}_{\mathbf{x}^*} = \{i \in V \mid x_i^* = 1\}$ . Whether or not players invest can be completely characterized using a collection

<sup>5</sup>There may be other PSNE of the game. We implicitly assume that the principal can suggest an equilibrium to the agents, who will follow the suggestion unless it is in their best interest to deviate.

<sup>6</sup>Typically, a principal would be more likely to want to strengthen the altruism network. However, one can easily imagine situations and construct instances in which a weakening of the network is necessary. We therefore aim for more generality in our model.

of inequalities. In these inequalities, what ultimately determines the decision is the agent's marginal value from investing. This marginal value has two elements: first, the agent's own marginal benefit,  $\Delta_{x_i} g_i(n_i) := g_i(1, n_i) - g_i(0, n_i)$ , and second, the marginal benefit that  $i$  can obtain from altruism towards  $j$ , which is  $\Delta_j^- := g_j(x_j, n_j) - g_j(x_j, n_j - 1)$  for agents  $i$  who invest, and  $\Delta_j^+ := g_j(x_j, n_j + 1) - g_j(x_j, n_j)$  for those agents  $i$  who do not. We further define  $\theta_i := c_i - \Delta_{x_i} g_i(n_i)$  for each agent  $i$ . We then obtain the following equivalent characterization of a PSNE of  $\text{BNPG}(A)$ , in which a set  $\mathcal{I}$  of players invest, and the rest do not:

$$\begin{aligned} \sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j} \Delta_j^- &\geq \theta_i & \text{if } i \in \mathcal{I} \\ \sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j} \Delta_j^+ &\leq \theta_i & \text{if } i \notin \mathcal{I}, \end{aligned} \quad (4)$$

with  $a_{i,j}$  the strength of  $i$ 's altruism for  $j$  under  $A$ .

The marginal benefit functions  $\Delta g_i := (\Delta_i^-, \Delta_i^+)$  are important parameters in our model; they can be restricted by putting limitations on  $g_i$ , e.g.,  $\Delta g_i$  is bounded by a polynomial in  $n$  iff  $g_i$  itself is. We consider three possible restrictions:

**general  $g_i$ :**  $\Delta_i^-$  and  $\Delta_i^+$  are arbitrary.

**polynomial  $g_i$ :** When the  $g_i$  are bounded by a polynomial in  $n$ ,  $\Delta_i^-$  and  $\Delta_i^+$  are also bounded by a polynomial in  $n$ .

**uniform linear and separable  $g_i$ :** When all  $g_i$  are of the form  $g_i(x_i, n_i) = h_i(x_i) + \Delta \cdot n_i$  for some possibly idiosyncratic function  $h_i$  and some (common) constant  $\Delta$ , then  $\Delta_i^- = \Delta_i^+ = \Delta$  for all  $i$  and all values of  $n_i$ .

## 4 An LP for Fractional Modifications

We begin by considering the variant of the problem in which the principal can spend fractionally on each action, i.e.,  $\mathbf{v} \in \mathbb{R}_{\geq 0}^K$ . In that case, the principal's optimal strategy can be found using a straightforward linear program. The decision variables are simply the principal's investments  $v_k \geq 0$ . For ease of notation, we also define variables  $a_{i,j}$  for the altruism from  $i$  towards  $j$  resulting from the modifications. Given the available actions and costs, the relationship between  $a_{i,j}$  and  $v_k$  is exactly characterized by Equation (3), which is linear in the  $v_k$ . Next, notice that given the  $\text{BNPG}$  and the desired PSNE  $\mathbf{x}^*$ , the set  $\mathcal{I}$  and all relevant constants in Equation (4) (that is  $\Delta_i^-, \Delta_i^+, \theta_i$ ) can be immediately computed. Therefore, we obtain the following linear program for finding the optimal spending strategy for the principal:

$$\begin{aligned} \text{Minimize } & \sum_{k=1}^K v_k \cdot \gamma^{(k)} & \text{subject to} \\ a_{i,j} &= a_{i,j}^{\text{in}} + \sum_{k=1}^K v_k \cdot \sigma^{(k)} \cdot m_{i,j}^{(k)} & \text{for all } i, j \\ \sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j} \Delta_j^- &\geq \theta_i & \text{for all } i \in \mathcal{I} \\ \sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j} \Delta_j^+ &\leq \theta_i & \text{for all } i \notin \mathcal{I} \\ v_k &\geq 0 & \text{for all } k. \end{aligned} \quad (5)$$

Notice that we could fairly straightforwardly generalize the LP to even deal with more general actions, namely, allowing an action  $k$  to affect different edges in different ways by allowing  $M^{(k)}$  to have entries that are not all equal.

## 5 Graph-Based Modifications

While there are contexts in which a principal can precisely control the amount of effort invested in different actions, there are many others in which actions take more of an all-or-nothing nature. Indeed, many activities that increase altruism through making harms to others more salient, such as community meetings or public health advertising campaigns (e.g., to promote mask-wearing), are naturally discrete (a 1-second ad is not very effective) and, thus, have a discrete impact on the altruism graph. This motivates the special case in which all entries of  $v$  must be binary, corresponding to decisions whether or not the principal will add/remove the edge sets  $\mathcal{S}^{(k)}$ . As we show presently that even if we restrict  $\mathcal{S}^{(k)}$  to affect a single edge the problem now becomes NP-hard, we devote the sequel to this special case, and seek to identify what additional structure is sufficient to make the problem, or its approximation, tractable.

When  $\mathcal{S}^{(k)}$  is a singleton  $(i, j)$ , we can think of the input as a graph  $G^{\text{in}} = (V, E^{\text{in}})$  (instead of a weighted network). There is a cost  $\gamma^{(i,j)}$  associated with each (directed or undirected) node pair  $(i, j)$ . If  $(i, j) \in E^{\text{in}}$ , then this is the cost for removing  $(i, j)$ ; otherwise, it is the cost of adding  $(i, j)$ . Thus, implicitly,  $\sigma^{(i,j)} = 1$  if  $(i, j) \notin E^{\text{in}}$ , and  $\sigma^{(i,j)} = -1$  if  $(i, j) \in E^{\text{in}}$ . Adding/removing edges results in a new altruism network  $G = (V, E)$ . All off-diagonal non-zero entries of the altruism network  $A$  then have the same altruism value  $a$ , while all diagonal entries are set to 1.

We study two variants of this problem: 1) asymmetric altruism, that is, when the altruism graph is directed, and 2) symmetric altruism, where it is undirected. Typically, both will capture some important aspect of the real world: while altruism often aligns with actual social or kinship ties, it can also result from a general sense of responsibility or goodwill, which may not be reciprocated.

### 5.1 Asymmetric Altruism

We begin by formally showing that in this setting, ANM is in general intractable, even in the special case where we can only add or remove individual edges, rather than subsets of edges.

**Theorem 5.1.** *ANM with asymmetric altruism is NP-complete even when:*

1. *the sets  $\mathcal{S}^{(k)}$  are singletons and can only be added, i.e.,  $\sigma^{(k)} = 1$  for all  $k$ ,*
2. *the initial altruism network is empty, and*
3. *the target profile  $x^*$  has all agents investing.*

The proof is a direct reduction from the KNAPSACK problem and given in the Supplementary Material.

Next, we show that under mild additional assumptions, the problem becomes efficiently solvable. The key observation that enables our positive results is that for directed graphs, adding or removing an edge  $i \rightarrow j$  only affects agent  $i$ 's altruistic behavior. This allows us to decompose the problem into  $n$  independent subproblems only connected through a common budget constraint. Each subproblem can be naturally modeled as a KNAPSACK problem, and so long as the KNAPSACK problems are individually solvable, so is the overall

problem. More specifically, we distinguish two cases, based on whether  $i \in \mathcal{I}$ .

- If  $i \in \mathcal{I}$ , then the altruism edges originating with  $i$  must ensure that  $\sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j} \Delta_j^- \geq \theta_i$ . Let  $\phi_i = \sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j}^{\text{in}} \Delta_j^-$ . If  $\phi_i \geq \theta_i$ , then  $i$  will invest even without adding any edges. Otherwise, the principal will need to add edges out of  $i$  adding a total altruism term of  $\theta_i - \phi_i$ ; also, the principal will never want to remove edges out of  $i$ . Adding a directed edge  $i \rightarrow j$  can be thought of as putting an item with value  $a \cdot \Delta_j^-$  and cost/weight  $\gamma^{(i,j)}$  into a knapsack. Thus, the set of “items” available to add to agent  $i$  is  $\mathcal{P}_i = \{j \in \mathcal{N}_i^{(H)} \mid a_{i,j}^{\text{in}} = 0\}$ . The subproblem is then to select items from  $\mathcal{P}_i$  such that the total value is at least  $\theta_i - \phi_i$  while the total weight is minimized.
- Similarly, if  $i \notin \mathcal{I}$ , then the altruism edges originating with  $i$  must ensure that  $\sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j} \Delta_j^+ \leq \theta_i$ . Let  $\phi_i = \sum_{j \in \mathcal{N}_i^{(H)}} a_{i,j}^{\text{in}} \Delta_j^+$ . Analogously to the previous case, the principal now wants to *remove* edges such that the altruism is reduced by at least  $\phi_i - \theta_i$  (unless  $\phi_i \leq \theta_i$ , in which case nothing needs to be done). Again, this problem can be modeled as a KNAPSACK problem. The set of items available is  $\mathcal{P}_i = \{j \in \mathcal{N}_i^{(H)} \mid a_{i,j}^{\text{in}} = 1\}$ . The directed edge  $i \rightarrow j$  is modeled as an item with value  $a \cdot \Delta_j^+$  and cost/weight  $\gamma^{(i,j)}$ . The goal is to minimize the total weight subject to achieving total value at least  $\phi_i - \theta_i$ .

It is well known [Kleinberg and Tardos, 2005; Vazirani, 2001] that the KNAPSACK problem can be solved in polynomial time using Dynamic Programming when either the weights or the values are bounded by a polynomial in the number of items. Using standard rounding/scaling techniques, this approach also yields an FPTAS. By leveraging these algorithms, we obtain the corresponding results for our problem.

**General  $g_i$  and Polynomial Edge Costs.** When edge costs are bounded by a polynomial in  $n$ , which corresponds to polynomially bounded item weights in the KNAPSACK instances, which are therefore polynomial-time solvable with Dynamic Programming (DP). Applying DP for each agent  $i$  separately then yields a minimum-cost overall solution. We obtain the following:

**Proposition 5.2.** *Under asymmetric altruism, the problem ANM with general  $g_i$  and polynomially bounded edge costs is polynomial-time solvable.*

**Polynomial  $g_i$  and General Edge Costs.** When all the  $g_i$  are polynomially bounded, so are their differences, and hence the (scaled) item values  $\Delta_j^+$  and  $\Delta_j^-$ . Hence all the values of the “items” are polynomially bounded. Tractability then follows from the fact that the KNAPSACK problem is polynomial-time solvable (using Dynamic Programming) when item values are bounded by some polynomial in the length of the input. Again, this allows an algorithm to solve each subproblem in polynomial time, and then aggregate the optimal solutions. This gives rise to the following proposition:

**Proposition 5.3.** *Under asymmetric altruism, the problem ANM with polynomially bounded  $g_i$  and general edge costs is polynomial-time solvable.*

**An FPTAS for the General Case.** Finally, we can leverage the standard FPTAS for KNAPSACK to obtain an FPTAS for ANM with general asymmetric altruism. Specifically, given a parameter  $\epsilon$ , one can run the FPTAS with that parameter for each of the subproblems/agents  $i$  separately. The result for each  $i$  will be a set of edges to add/remove such that  $i$  invests iff  $x_i^* = 1$ , and the total cost of the modifications is within a factor  $(1 + \epsilon)$  of optimal. Adding all of these costs shows that the overall cost is within a factor  $(1 + \epsilon)$  of optimal. We obtain the following proposition:

**Proposition 5.4.** *Under asymmetric altruism, consider ANM with general  $g_i$  and general edge costs. Given  $\epsilon > 0$ , the optimal cost  $B^*$  can be approximated arbitrarily well, i.e., a solution of cost  $B \leq (1 + \epsilon)B^*$  can be found, with an algorithm which runs in time polynomial in  $n$  and  $1/\epsilon$ .*

## 5.2 Symmetric Altruism

Next, we turn to the setting in which altruism is reciprocal or symmetric: when an edge is added to the altruism network, it affects *both* incident agents. While many graph-theoretic questions are easier for undirected graphs than for directed ones, the ANM problem becomes harder. Intuitively, the reason is that while adding the edge  $(i, j)$  is very beneficial for  $i$ , it may be less so for  $j$ ; given the choice, adding a different edge out of  $j$  may be preferable. Under the symmetric altruism model, the principal does not have the fine-grained control of adding different edges out of  $i$  and  $j$ , and might have to “waste” one direction of the edge. The resulting “side-effects” of desirable edges must be more globally balanced. Indeed, we show that even special cases that are polynomial-time solvable in the asymmetric model become NP-hard in the symmetric model.

**Theorem 5.5.** *ANM with symmetric altruism is NP-complete even when*

- all the sets  $\mathcal{S}^{(k)}$  are singletons,
- all agents invest under the target equilibrium  $\mathbf{x}^*$ ,
- all  $g_i$  are polynomially bounded,
- all edge costs  $\gamma^{(k)}$  are 1, and
- the graph  $H$  is a clique.

The proofs of this and the remaining results are in the Supplementary Material. Recall that for the asymmetric case, even just one of uniform (or even just polynomially bounded) edge costs and polynomially bounded  $g_i$  is enough to obtain tractability.

For the remainder of this section, we will focus on the special case when all agents invest under the target equilibrium  $\mathbf{x}^*$ . This is because for more general target equilibria, even deciding if there exists *any* altruism graph yielding this equilibrium is NP-complete. In other words, even the decision problem of a principal with infinite budget is NP-complete.

**Theorem 5.6.** *ANM with symmetric altruism for arbitrary target equilibria  $\mathbf{x}^*$  is NP-complete even when*

- all the sets  $\mathcal{S}^{(k)}$  are singletons,
- all  $g_i$  are polynomially bounded,

- all edge costs  $\gamma^{(k)}$  are 0 (i.e., the principal’s budget is infinite),
- the graph  $H$  is a clique.

Theorem 5.6 implies that no approximation guarantee can be attained in polynomial time for ANM when  $\mathbf{x}^*$  is an arbitrary action profile. However, we remark that when all agents invest under  $\mathbf{x}^*$ , there is a straightforward polynomial-time  $(2 + \epsilon)$ -approximation algorithm: the algorithm applies the FPTAS to the asymmetric version and adds a reciprocal edge whenever a directed edge is added. This leads to a blowup of a factor 2 compared to the FPTAS achieved in the asymmetric setting.

**Uniform Separable Linear Utility Functions.** When the agents have utility functions  $g_i$  that are separable in the arguments, and linear with common slope in the second argument (we term these uniform separable and linear, or USL), the marginal benefits are uniform and equal to a constant  $\Delta$ , i.e.,  $\Delta_i^- = \Delta_i^+ = \Delta$  for all  $i$  and all values of  $n_i$ . Such utility functions are commonly studied in public goods games [Suri and Watts, 2011]. We show that when the goal is for all players to invest, for USL utility functions and when all sets  $\mathcal{S}^{(k)}$  are single edges, the problem becomes tractable, even with all other parameters (edge costs, network structure, etc.) being fully general.

**Theorem 5.7.** *ANM with symmetric altruism is polynomial time solvable when*

- the utility functions  $g_i$  are USL,
- the sets  $\mathcal{S}^{(k)}$  are singleton,
- the target equilibrium  $\mathbf{x}^*$  has all agents investing.

The proof of Theorem 5.7 (given in the Supplementary Material) is through a non-trivial reduction to the MIN-COST PERFECT MATCHING problem via a connection to another related problem: NETWORK DESIGN FOR DEGREE SETS (NDDS) [Kempe *et al.*, 2020].

## 6 Conclusion

We consider how to change altruistic behavior of individuals so as to induce societally desirable outcomes. One major contribution of our work is to separately capture the strategic interdependencies and the altruistic behaviors of individuals. We propose a model of modifying the altruism network, with the goal of inducing a target investment profile by the individuals. A series of corresponding algorithmic results are exhibited, including hardness results even in very restrictive scenarios (e.g., each modification only affects a single edge), and tractability results in a broad array of special cases.

## Acknowledgments

This work was partially supported by the National Science Foundation (grants IIS-1905558, IIS-1939677, and 1903207) and Army Research Office (MURI grant W911NF1810208). The authors would like to thank the anonymous reviewers. YV would like to thank Amazon.

## References

- [Amelkin and Singh, 2019] Victor Amelkin and Ambuj K. Singh. Fighting opinion control in social networks via link recommendation. In *Proc. 25th Intl. Conf. on Knowledge Discovery and Data Mining*, pages 677–685, 2019.
- [Bergstrom, 1999] Theodore C. Bergstrom. Systems of benevolent utility functions. *J. of Public Economic Theory*, 1(1):71–100, 1999.
- [Bramoullé and Kranton, 2007] Yann Bramoullé and Rachel Kranton. Public goods in networks. *Journal of Economic Theory*, 135(1):478–494, 2007.
- [Bramoullé et al., 2014] Yann Bramoullé, Rachel Kranton, and Martin D’Amours. Strategic interaction and networks. *American Economic Review*, 104(3):898–930, 2014.
- [Bredereck and Elkind, 2017] Robert Bredereck and Edith Elkind. Manipulating opinion diffusion in social networks. In *Proc. 26th Intl. Joint Conf. on Artificial Intelligence*, pages 894–900, 2017.
- [Chen and Kempe, 2008] Po-An Chen and David Kempe. Altruism and selfishness in traffic routing. In *Proc. 9th ACM Conf. on Electronic Commerce*, pages 140–149, 2008.
- [Chen et al., 2011] Po-An Chen, Bart de Keijzer, David Kempe, and Guido Schäfer. The robust price of anarchy of altruistic games. In *Proc. 7th Workshop on Internet and Network Economics (WINE)*, pages 383–390, 2011.
- [Chen et al., 2014] Po-An Chen, Bart de Keijzer, David Kempe, and Guido Schäfer. Altruism and its impact on the price of anarchy. *ACM Trans. Economics and Computation*, 2(4):17, 2014.
- [Chen et al., 2016] Chen Chen, Hanghang Tong, B. Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, and Christos Faloutsos. Eigen-optimization on large graphs by edge manipulation. *ACM Transactions on Knowledge Discovery from Data*, 10(4):49:1–49:30, 2016.
- [Dong et al., 2016] Yali Dong, Boyu Zhang, and Yi Tao. The dynamics of human behavior in the public goods game with institutional incentives. *Scientific Reports*, 6(1):1–7, 2016.
- [Dughmi, 2017] Shaddin Dughmi. Algorithmic information structure design: A survey. *ACM SIGecom Exchanges*, 15(2):2–24, 2017.
- [Feldman et al., 2013] Michal Feldman, David Kempe, Brendan Lucier, and Renato Paes Leme. Pricing public goods for private sale. In *Proc. 14th ACM Conf. on Electronic Commerce*, pages 417–434, 2013.
- [Galeotti et al., 2010] Andrea Galeotti, Sanjeev Goyal, Matthew O. Jackson, Fernando Vega-Redondo, and Leeat Yariv. Network games. *The Review of Economic Studies*, 77(1):218–244, 2010.
- [Garimella et al., 2018] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Reducing controversy by connecting opposing views. In *Proc. 27th Intl. Joint Conf. on Artificial Intelligence*, pages 5249–5253, 2018.
- [Ghosh and Boyd, 2006] Arpita Ghosh and Stephen Boyd. Growing well-connected graphs. In *Proc. 45th IEEE Conference on Decision and Control (CDC)*, pages 6605–6611, 2006.
- [Grossklags et al., 2008] Jens Grossklags, Nicolas Christin, and John Chuang. Security and insurance management in networks with heterogeneous agents. In *Proc. 9th ACM Conf. on Electronic Commerce*, pages 160–169, 2008.
- [Haeringer, 2018] Guillaume Haeringer. *Market Design: Auctions and Matching*. The MIT Press, 2018.
- [Kearns et al., 2001] Michael Kearns, Michael L. Littman, and Satinder Singh. Graphical models for game theory. In *Proc. 17th Conf. on Uncertainty in Artificial Intelligence*, pages 253–260, 2001.
- [Kempe et al., 2020] David Kempe, Sixie Yu, and Yevgeniy Vorobeychik. Inducing equilibria in networked public goods games through network structure modification. In *Proc. 19th Intl. Conf. on Autonomous Agents and Multiagent Systems*, pages 611–619, 2020.
- [Kleinberg and Tardos, 2005] Jon Kleinberg and Eva Tardos. *Algorithm Design*. Addison-Wesley, 2005.
- [Ledyard, 1997] John Ledyard. Public goods: A survey of experimental research. In J. Kagel and A. Roth, editors, *Handbook of Experimental Economics*, pages 111–194. Princeton University Press, 1997.
- [Levine, 1998] David Levine. Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1:593–622, 1998.
- [Matteo Castiglioni, 2020] Nicola Gatti Matteo Castiglioni, Diodato Ferraioli. Election control in social networks via edge addition or removal. In *Proc. 34th AAAI Conf. on Artificial Intelligence*, 2020.
- [Meier et al., 2008] Dominic Meier, Yvonne Anne Oswald, Stefan Schmid, and Roger Wattenhofer. On the windfall of friendship: Inoculation strategies on social networks. In *Proc. 9th ACM Conf. on Electronic Commerce*, pages 294–301, 2008.
- [Nisan et al., 2007] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.
- [Sheldon et al., 2010] Daniel Sheldon, Bistra Dilikina, Adam N. Elmachtoub, Ryan Finseth, Ashish Sabharwal, Jon Conrad, Carla Gomes, David Shmoys, William Allen, Ole Amundsen, and William Vaughan. Maximizing the spread of cascades using network design. In *Proc. 26th Conf. on Uncertainty in Artificial Intelligence*, pages 517–526, 2010.
- [Shoham and Leyton-Brown, 2008] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2008.
- [Sina et al., 2015] Sigal Sina, Noam Hazon, Avinatan Hassidim, and Sarit Kraus. Adapting the social network to affect elections. In *Proc. 14th Intl. Conf. on Autonomous Agents and Multiagent Systems*, pages 705–713, 2015.
- [Suri and Watts, 2011] Siddharth Suri and Duncan J Watts. Cooperation and contagion in web-based, networked public goods experiments. *PLoS One*, 6(3):e16836, 2011.
- [Tong et al., 2012] Hanghang Tong, B. Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, and Christos Faloutsos. Gelling, and melting, large graphs by edge manipulation. In *Proc. 21st ACM Intl. Conf. on Information and Knowledge Management*, pages 245–254, 2012.
- [Vazirani, 2001] Vijay V. Vazirani. *Approximation Algorithms*. Springer, 2001.
- [Yu et al., 2020] Sixie Yu, Kai Zhou, P. Jeffrey Brantingham, and Yevgeniy Vorobeychik. Computing equilibria in binary networked public goods games. In *Proc. 34th AAAI Conf. on Artificial Intelligence*, 2020.