TARGETING INFLUENCE IN A HARMONIC OPINION MODEL*

ZACHARY M. BOYD†, NICOLAS FRAIMAN‡, JEREMY L. MARZUOLA§, PETER J. MUCHA¶, AND BRAXTON OSTING‖

Abstract. Influence propagation in social networks is a central problem in modern social network analysis, with important societal applications in politics and advertising. A large body of work has focused on cascading models, viral marketing, and finite-horizon diffusion. There is, however, a need for more developed, mathematically principled *adversarial models*, in which multiple, opposed actors strategically select nodes whose influence will maximally sway the crowd to their point of view.

In the present work, we develop and analyze such a model based on harmonic functions and linear diffusion. Our general problem is known to be NP-hard and that the objective function is monotone and submodular; consequently, we can greedily approximate the solution within a constant factor. Introducing and analyzing a convex relaxation, we show that the problem can be approximately solved using smooth optimization methods. We illustrate the effectiveness of our approach on a variety of example networks.

Key words. opinion dynamics, social network, maximizing influence, zealot, graph Laplace equation, convex relaxation

MSC codes. 35J05, 05C50, 49M41, 65K10

1. Introduction.

On social media platforms (and in real life), ideas, information, and opinions propagate through connections in the social network; network members may influence their connections, e.g., swaying the political viewpoint of friends, disseminating information about current events, or increasing product awareness in followers. To be concrete, let's consider a particular product for which there are $k = 2$ competing brands: a blue brand and a red one. Each member of the network will have some degree of preference between the brands and also have varying influence on their connections in the network. Indeed, *influencers* emerge on social media platforms and have the ability to widely market their preferred color product to their many connections. Additionally, a *zealot* is an individual that is uncompromising or inflexible in their (typically extreme) idea/opinion. An *authority* for the blue or red brand wants to most broadly advertise it within the network. They may not be a node in the network themselves, but instead try to influence the network by persuading key nodes to join their side. Companies and political parties are examples of authorities. We suppose that authorities have a budget to strategically target members of the network and convert them to become zealots. For example, they may sponsor them or otherwise incentivize them to disseminate their viewpoint. Intuitively, an authority might target members of the network with many connections, but it is not so sim-

*Submitted to the editors DATE.

Funding: Z.M.B. and P.J.M. were supported by ARO MURI award W911NF-18-1-0244. Z.M.B. was also supported by NSF DMS-2137511. J.L.M. acknowledges support from the NSF through grant DMS-2307384 and FRG grant DMS-2152289. P.J.M. was also supported by the NSF through grant BCS-2140024. B.O. acknowledges support from NSF DMS-1752202 and DMS-2136198.

†Department of Mathematics, Brigham Young University, Provo, UT 84602 (zachboyd@byu.edu).

‡Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 (fraiman@email.unc.edu).

§Department of Mathematics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 (marzuola@math.unc.edu).

¶Department of Mathematics, Dartmouth College, Hanover, NH 03755 (peter.j.mucha@dartmouth.edu).

‖Department of Mathematics, University of Utah, Salt Lake City, UT 84112 (osting@math.utah.edu).

ple. For example, choosing a powerful influencer may not be very effective if most of their connections already agree with the authority's point of view. We refer to the problem of authorities targeting members within a social network so as to maximize the influence of a particular opinion as the *targeting influence problem*. The targeting influence problem appears in many guises beyond the targeted product marketing example given here; indeed, swaying public opinion is a natural goal in both politics and advertising. As an additional example, it is paramount for various authorities to combat "fake news" by disseminating factual information to influential members of a network.

1.1. Contributions and results. In this paper, *we analyze a model for targeting influence in a social network in settings where there are a multitude of extreme opinions, with our analysis including the introduction and study of a convex relaxation of the targeting influence problem.* The opinion of each member of the network lies in an opinion space, taken to be the convex hull of some extreme opinions; see subsection 2.1. The opinion of each individual is influenced by their (directed) connections in the social network. We use a simple, harmonic model based on DeGroot learning [15] and label propagation [50] to describe these opinion dynamics; see subsection 2.2. This type of combinatorial optimization problem has been studied in a variety of methods in the past in the network literature, see for instance the surveys [6, 30, 43]. Our main goal will be to introduce a viable approximation method that is built off of a regularized notion of influence.

In subsection 2.3, we introduce our measure for the influence for a particular opinion and an optimization problem that describes how an authority might target influence within the social network; see (2.4). The measure of influence of each opinion is simply the sum of the opinions of the individuals in the network. Similar to the work in [9], we emphasize that in this framework we can work with directed opinion networks. In subsection 2.4, we show that each authority can group together the opposing opinions/ideas when considering how to maximize their own influence, thus resulting in a reduced *grouped opponent targeting influence problem*; see (2.5). In subsection 2.5, we review the probabilistic interpretation of harmonic functions on graphs in terms of the hitting time of an associated random walk. In subsection 2.6, we describe how the problem, in the case of an undirected network, can be interpreted in terms of the graph Dirichlet energy.

In section 3, we recall some hardness properties of our opinion model that were originally established in [21]. As it is instructive to understand the graph dynamics relevant to the model, we will re-prove using slightly different techniques that there exists a $\varepsilon > 0$ such that it is NP-hard to obtain a $(1-\varepsilon)$-approximation to the targeting influence problem; see Theorem 3.2. The proof uses a gap-introducing reduction to the vertex cover problem on cubic (3-regular) graphs, which is APX-hard. To recall, APX problems are optimization problems that have a polynomial-time constant-factor approximation algorithm. Some problems in APX have a polynomial-time approximation scheme (PTAS) and can be approximated to any factor. However, the PCP Theorem (probabilistically checkable proofs) [5] established that there are problems in APX which cannot be approximated efficiently by a PTAS. These so-called APX-hard problems have approximation-preserving reductions from every problem in APX. Therefore, APX-hard problems do not have a PTAS and cannot be approximated to an arbitrary factor. We also give an alternative proof that the objective function in the influence targeting problem is monotone and submodular and hence admits a $1 - 1/e$ approximation algorithm; see Proposition 3.4. The proof relies on the probabilistic interpretation of harmonic functions on graphs.

In section 4, we introduce our key contribution in this work. Namely, we develop an approximation strategy for solutions of the NP-hard targeting influence problem. We show that the objective function for the relaxed optimization problem (4.3) is concave and compute the gradient and Hessian; see Theorem 4.3. In subsection 4.1, we consider the relaxed optimization problem and prove that, in the case when the graph has a symmetry, the optimal solution also has the same symmetry; see Proposition 4.6.

In section 5, we describe a numerical implementation of our method and present the outcomes of numerous experiments to illustrate our results and test how well the relaxed optimization problem predicts the solution to the original (APX-hard) problem. For simplicity, we consider a game between two opinion authorities to capture the largest proportion of influence, where each authority alternatively selects one vertex to convert to their opinion. In general, we see that the relaxed problem highlights similar features of the graph as the unrelaxed problem giving a reasonable and computationally efficient strategy for targeting influence.

We conclude in section 6 with a discussion.

1.2. Previous work. There are a wide variety of interrelated opinion dynamics, belief propagation, and consensus formation models that have been introduced and analyzed. Generally in these models, the opinions of members in a social network evolve in time via network interactions. The member opinions can be modeled as discrete or continuous and can account for more than two opinions. The interactions between members can be modeled using Bayesian or non-Bayesian approaches; here we focus on non-Bayesian approaches, which can further be characterized as dynamical/stochastic models in discrete/continuous time. A review of this vast literature is beyond the scope of this paper (see, e.g., [2, 41]), but non-Bayesian opinion dynamics models include or are motivated by cascading models [24, 25, 26], percolation models, the Vicsek swarming/flocking model [46], the Kuramoto synchronization model [29], the DeGroot model [15] and its many variants (as described in, e.g., [13], the label propagation model [50], voter models [42], and bounded confidence models [7]). In the present work, we consider the steady state of a specific form of the DeGroot dynamical model, which is closely related to the label propagation model.

Within opinion dynamics models a wide variety of questions are posed and studied. What is the steady state of the dynamics: does the population reach consensus or are opinions polarized/fragmented? How do opinion dynamics depend on the topology of the social network or on the way in which interactions between members are modeled? How do we identify authoritative members of the network [28, 40]? If influencers/zealots/opinion authorities/forceful agents are introduced, how does this effect opinion dynamics [3, 20, 35, 45] and how does this change with network topology [27]? We view the targeting influence problem as an optimal control/inverse problem for this last question.

There are several previous works that look at targeting influence or "influence maximization," as considered in this paper. To our knowledge, this type of question was first mathematically posed in [16] who modeled influence by a Markov random field. Subsequent work also analyzed this type of question when opinion dynamics is an independent cascade model [8, 36].

In the unregularized version of our framework, there are several recent results that we recall here. We first mention the work [17], where the authors formulate a binary opinion/idea targeting influence problem equivalent to our grouped opponent targeting influence problem (2.5). In particular, the opinion dynamics model is taken to be the steady-state of the DeGroot model. This problem is reformulated and approximately solved via the alternating direction method of multipliers (ADMM). Interestingly, the authors also consider perturbations to the network connections to enhance the influence of zealots. Secondly, a related version of the model is presented in the paper [21], where the authors demonstrate that the classical version of our model is NP-hard using similar tools to what we present here. Lastly, we mention that directed network versions of opinion models were analyzed in [9].

We also mention the work [48, 49], in which the authors consider a targeting influence problem where the opinion dynamics model is also taken to the steady-state of the DeGroot model. However, whereas we consider the measure of influence of each opinion as the sum of the opinions of the individuals in the network, [48, 49] introduces a measure which counts the voters who prefer each opinion. This “competition” between binary opinions/ideas might be more realistic in an election, for example. The authors introduce an Influence Matrix (IM) criterion to predict the bias of each non-zealot individual and hence the result of the competition. They compare the IM criterion with seven centrality-based criteria and extensive numerical experiments suggest that the IM criterion is effective at predicting the competition result.

Finally, the targeting influence problem bears some resemblance to the problem of active learning for semi-supervised learning tasks. Here, one selects maximally informative vertices/edges of a graph (representing data) to be included as labeled data for the semi-supervised learning task. This problem has recently been studied in a variety of settings that use similar regularization and/or greedy methods; see, for example, [22, 23, 31, 32, 33, 34, 38, 51].

Notation. Write $[k] = \{1, 2, \dots, k\}$ for $k \in \mathbb{N} \setminus \{0\}$. Denote by $e_\ell \in \mathbb{R}^k$ the characteristic vector for coordinate $\ell \in [k]$. Denote by $e_V \in \mathbb{R}^{|V|}$ the ones vector. Denote the unit simplex by $\Delta_k = \{u \in \mathbb{R}^k \colon u \geq 0,\ \sum_{\ell \in [k]} u_\ell = 1\} = \operatorname{conv}\left(\{e_\ell\}_{\ell \in [k]}\right)$. For vectors or matrices a and b, $a \odot b$ represents the Hadamard product. The outer product of two vectors a and b is written $a \otimes b$.

2. Opinion dynamics and targeting influence.

2.1. A model for opinions and influencers. We consider a particular topic and assume that there are $k \geq 2$ (fixed) extreme opinions on the topic. Each individual (represented by a vertex in the network) has an opinion on the topic, and we model each individual's opinion as a convex combination of the extreme opinions. To this end, we represent the extreme opinions by the coordinate vectors $e_\ell \in \mathbb{R}^k$. The *opinion of individual* i is represented as the vector $u(i) = (u_1, \dots, u_k)(i) \in \Delta_k$. Thus, the *opinion space* is the unit simplex, $\Delta_k \subset \mathbb{R}^k_+$, the convex hull of the extreme opinions.

2.2. A harmonic model for opinion dynamics. Here we describe a simple model for opinion dynamics; this model is related to the DeGroot model [15] and the label propagation model [50], as well as other models in opinion dynamics. We call attention to the introduction of [13], which delineates the literature on numerous variants of the modeling we leverage here, including the introduction of continuous-time modeling [1] and the inclusion of zealots [19, 44], as well as many other models with rich behaviors. For our present purposes, we introduce the model we consider through a continuous-time dynamics; but we are ultimately focused only on the steady states resulting from that model in formulating our influence maximization problem (though, in so doing, we recognize that models yielding qualitatively different steady states might also be interesting to study in this manner). We will assume that the individuals are connected via a social network. Within the network, there are zealots with (fixed) extreme opinions and other members with opinions that are determined by their social connections. Roughly speaking, the opinion of the non-zealot individuals will be the average of the opinions of the individuals networked to the individual.

Let $G = (V, E)$ be a strongly connected, directed graph[1] that represents the social network. Set $n = |V|$ and enumerate the nodes so that we may identify V with $[n]$. Let $Z_\ell \subset V$, $\ell \in [k]$ be disjoint vertex sets representing zealots, where Z_ℓ is the set of zealots for opinion $\ell \in [k]$. Consequently, we set

$$u(i) = e_\ell, \qquad\qquad i \in Z_\ell,\ \ell \in [k].$$

The vertex subset $Z := \amalg_{\ell\in[k]} Z_\ell$ is the *collection of all zealots*. Note that for every vertex $i \in Z$, $u(i)$ is the characteristic vector for the disjoint partition $Z = \amalg_{\ell\in[k]} Z_\ell$.

Since G is a directed graph, individual i may be influenced by j even if j is not influenced by i. We use the convention that a directed edge (i, j) represents that i is influenced by j. That means edges are directed in the opposite direction in which opinions spread in the network. Let $A \in \mathbb{R}^{n\times n}$ be the directed adjacency matrix,

$$A_{i,j} = \begin{cases} 1 & i \text{ is influenced by } j \\ 0 & \text{otherwise} \end{cases}.$$

We will assume without loss of generality that $A_{ii} = 0$. Let $d_i = \sum_j A_{ij}$ be the out-degree of vertex $i \in V$ and define the diagonal matrix, $D \in \mathbb{R}^{n\times n}$ of out-degrees by $D_{ii} = d_i$, with $D_{ij} = 0$ whenever $i \neq j$. Denote the (out-degree) graph Laplacian for G by $L = D - A$. In the above construction, influence along edges is considered binary (edge weights are 1 or 0), but our analysis is unaffected if we instead consider weighted adjacency matrices where influence can be weighted differently along edges, i.e. $A = [a_{ij}]$ for $a_{ij} \geq 0$, and the corresponding weighted graph Laplacian $L = D - A$ where now $d_i = \sum_j a_{ij}$.

We assume a specific case of the DeGroot model, wherein the opinions evolve according to linear diffusion, satisfying

$$\frac{d}{dt}u(i) = -(Lu)(i) = -\sum_j L_{ij}u(j), \qquad i \in V \setminus Z, \tag{2.1a}$$

$$u(i) = e_\ell, \qquad i \in Z_\ell,\ \ell \in [k], \tag{2.1b}$$

with appropriate initial conditions. As $t \to \infty$, the opinions approach a steady state. The opinions of the non-zealot vertices are then defined to equal this steady state, satisfying

$$Lu(i) = 0, \qquad i \in V \setminus Z, \tag{2.2a}$$

$$u(i) = e_\ell, \qquad i \in Z_\ell,\ \ell \in [k]. \tag{2.2b}$$

Note that equation (2.2a) is written using vector notation assuming the order of operations that applies the Laplacian before evaluation for the ith node and is performed separately for each opinion index; that is, $Lu(i)$ should be read as $(Lu_\ell)(i) = 0$ for each $\ell \in [k]$. We interpret u to be a harmonic (vector-valued) vertex function satisfying Dirichlet boundary conditions, i.e., a *graph Laplace problem*. In particular,

[1] We recall that a *strongly connected* directed graph has a path between each ordered pair (both from and to) of vertices in the graph.

since $L = D - A$, we have that non-zealot i's opinion is the average of their outgoing neighbors $N(i)$ opinions,

$$u(i) = \frac{1}{d_i} \sum_{j \in N(i)} u(j), \qquad i \in V \setminus S,$$

where d_i is the out-degree of node i. Because of this local averaging, $u(i)$ is necessarily in the opinion space Δ_k if its neighbors' opinions are also in that opinion space. In (2.2), the zealots (vertices in the boundary Z) have opinion in the opinion space Δ_k. The following lemma shows that $u(i)$ for every $i \in V$ is in the opinion space Δ_k.

LEMMA 2.1. *Let $G = (V, E)$ be a strongly connected, directed graph, $Z \subset V$, and $f \colon Z \to \Delta_k$. Suppose $u \colon V \to \mathbb{R}^k$ solves*

$$\begin{aligned} Lu(i) &= 0, && i \in V \setminus Z, \\ u(i) &= f(i), && i \in Z. \end{aligned}$$

Then $u(i) \in \Delta_k$ for every $i \in V$.

Proof. Looking at each component $\ell \in [k]$ individually, by the graph maximum principle, we have that $u_\ell(i) \geq 0$ for all $i \in V$ and $\ell \in [k]$. Defining $v = \sum_{\ell \in [k]} u_\ell$, we have that $v \colon V \to \mathbb{R}$ solves the Laplace problem

$$\begin{aligned} Lv(i) &= 0, && i \in V \setminus Z, \\ v(i) &= 1, && i \in Z. \end{aligned}$$

The unique solution is given by $v = 1$, which gives that $u(i) \in \Delta_k$ for every $i \in V$. ◻

As an aside, we note that the simplex opinion space constraint holds not only for the steady state solution but also for the underlying dynamics in (2.1) when all individual opinions are initiated on the unit simplex.

Limitations of the harmonic model for opinion dynamics in the context of the targeting influence problem will be discussed in section 6.

2.3. Measuring and targeting influence. Suppose that we have opinions given by $u \colon V \to \Delta_k$. We measure the influence of opinion $m \in [k]$ by

$$\mathcal{I}_m(u) := \frac{1}{|V|} \sum_{i \in V} u_m(i) = \frac{1}{|V|} \|u_m\|_{\ell^1(V)}. \tag{2.3}$$

Note that the total influence of all opinions is given by

$$\sum_{m \in [k]} \mathcal{I}_m(u) = \frac{1}{|V|} \sum_{i \in V} \sum_{m \in [k]} u_m(i) = 1.$$

We thus interpret $\mathcal{I}_m(u)$ to be *the proportion of influence that opinion $m \in [k]$ has on the network*.

We can now formulate the targeting influence problem. Let the current zealot set be given by $Z = \amalg_{\ell \in [k]} Z_\ell$. Consider the authority for opinion $m \in [k]$. We suppose that this authority wants to augment their zealot set $Z_m \to Z_m \cup T$ such that $T \subset V \setminus Z$ has fixed size $|T| = t$ (corresponding to a fixed budget) as to maximize $\mathcal{I}_m$, the proportion of influence that opinion m has on the network. We formulate the *targeting influence problem* as

$$
\begin{aligned}
&\text{(2.4a)} && \max_{T\subset V} \mathcal{I}_m(u) && \\
&\text{(2.4b)} && \text{s.t. } Lu(i)=0, && i\in V\setminus(Z\cup T),\\
&\text{(2.4c)} && u(i)=e_\ell, && i\in Z_\ell,\ \ell\in[k],\\
&\text{(2.4d)} && u(i)=e_m, && i\in T,\\
&\text{(2.4e)} && |T|=t,\ T\cap Z=\varnothing. &&
\end{aligned}
$$

Intuitively, if the authority wants to have a large effect on the population, it should choose $T\subset V\setminus Z$ to consists of influencers within the population.

Limitations in our formulation of the targeting influence problem (2.4) will be discussed in section 6.

2.4. Reduction to a grouped opponent targeting influence problem. We now explain how, for fixed authority $m\in[k]$, the targeting influence problem (2.4) reduces to an us-vs-them targeting influence problem.

We observe that $\mathcal{I}_m$, defined in (2.3), only depends directly on u_m and not u_ℓ, $\ell\neq m$. Additionally, writing $v\equiv u_m$, the constraint set also separates and the equation for $v\colon V\to[0,1]$ can be written

$$
\begin{aligned}
Lv(i)&=0, && i\in V\setminus(Z\cup T),\\
v(i)&=1, && i\in Z_m\cup T,\\
v(i)&=0, && i\in Z_\ell,\ \ell\neq m.
\end{aligned}
$$

These three equations replace the constraints (2.4b)-(2.4d). Abusing notation we write $\mathcal{I}_m(v)=\frac{1}{|V|}\|v\|_{\ell^1(V)}$. For fixed $m\in[k]$, (2.4) can be equivalently expressed as

$$
\begin{aligned}
&\text{(2.5a)} && \max_{T\subset V} \mathcal{I}_m(v) && \\
&\text{(2.5b)} && \text{s.t. } Lv(i)=0, && i\in V\setminus(Z\cup T),\\
&\text{(2.5c)} && v(i)=1, && i\in Z_m\cup T,\\
&\text{(2.5d)} && v(i)=0, && i\in Z_\ell,\ \ell\neq m,\\
&\text{(2.5e)} && |T|=t,\ T\cap Z=\varnothing. &&
\end{aligned}
$$

Thus, authority m can group together the opposing opinions/ideas when consider how to maximize their own influence. We refer to (2.5) as the *grouped opponent targeting influence problem*. The advantage of (2.5) over (2.4) is that it involves a scalar-valued field, v, on V rather than a vector-valued field, u, on V.

Remark 2.2. Equation (2.5) is equivalent to [17, Eq. (1)], but presented with different language.

2.5. Probabilistic interpretation. The network opinion can be related to the hitting probabilities of the zealot sets for a random walk. Let X_s be a random walk on the network, that is $\mathbb{P}\left(X_{s+1}=j\mid X_s=i\right)=A_{ij}/d_i$. We write $\mathbb{P}_i\left(\,\cdot\,\right)=\mathbb{P}\left(\,\cdot\mid X_0=i\right)$ as shorthand for the random walk starting from node i. For a subset of vertices U define its hitting time as $\tau(U)=\inf\{s\geq 0\colon X_s\in U\}$ for a given random walk realization. That is, importantly, the $\tau(U)$ is a random variable dependent on the X_t random walk and selected subset U.

LEMMA 2.3. *Given zealot sets Z_ℓ for $\ell\in[k]$, if u is the solution to* (2.2) *we have*

$$u_m(i)=\mathbb{P}_i\left(\tau(Z)=\tau(Z_m)\right).$$

Proof. All we need to check is that the right hand side is harmonic and satisfies the boundary conditions. For $i \in V \setminus Z$, conditioning on the first step of the random walk (using the Markov property and $\mathbb{P}_i\left(X_1 = j\right) = 1/d_i$ for $j \in N(i)$) we have

$$\begin{aligned}\mathbb{P}_i\left(\tau(Z) = \tau(Z_m)\right) &= \sum_{j \in N(i)} \mathbb{P}_i\left(\tau(Z) = \tau(Z_m) \mid X_1 = j\right) \mathbb{P}_i\left(X_1 = j\right) \\ &= \frac{1}{d_i} \sum_{j \in N(i)} \mathbb{P}_j\left(\tau(Z) = \tau(Z_m)\right).\end{aligned}$$

Moreover, for the boundary nodes we have

$$\mathbb{P}_i\left(\tau(Z) = \tau(Z_m)\right) = \begin{cases} 1 & \text{if } i \in Z_m, \\ 0 & \text{if } i \in Z \setminus Z_m. \end{cases}$$

Therefore, $\left(\mathbb{P}_i\left(\tau(Z) = \tau(Z_\ell)\right)\right)_{\ell \in [k]}$ satisfies equation (2.2). ∎

Remark 2.4. Note that $\tau(Z) = \min_{\ell \in [k]} \tau(Z_\ell)$. For fixed m define $Z' = \cup_{\ell \neq m} Z_\ell$. We have that the following two events coincide $\left\{\tau(Z) = \tau(Z_m)\right\} = \left\{\tau(Z_m) < \tau(Z')\right\}$.

One way to interpret this result is to say that i does a random walk until it meets one of the zealot sets and it then picks the corresponding opinion. Then, u is the distribution of the opinion of node i.

2.6. Energy interpretation in the undirected case. We consider the case in which the graph is undirected, i.e., $A_{ij} = A_{ji}$. We define the Dirichlet energy of the (vector-valued) graph vertex function $u\colon V \to \Delta_k$ by $E(u) = \frac{1}{2} \sum_{\ell \in [k]} \langle u_\ell, L u_\ell \rangle$. The opinions of the non-zealot vertices are then defined as the solution to the energy minimization problem,

$$\operatorname{argmin}_u \; E(u) \tag{2.6a}$$

$$\text{s.t. } u(i) = e_\ell, \qquad i \in Z_\ell, \; \ell \in [k]. \tag{2.6b}$$

The Dirichlet conditions (2.6b) state that the opinions of the zealots are fixed at their extreme values. The opinions of non-zealots take values as to minimize the Dirichlet energy of the opinion field u. Mathematically, (2.6) defines a harmonic field with image in the unit simplex. Again, the energy function and constraints in (2.6) are separable, and we can solve for each component of the opinion field $u\colon V \to \Delta_k$ independently.

3. Hardness and approximation of the targeting influence problem.

3.1. Hardness of the targeting influence problem.

Here we show that the targeting influence problem (2.4) is NP-hard to approximate by a reduction to the vertex cover problem on cubic (3-regular) graphs. These results first appeared in [21], but we give a simpler set of proofs below. We recall that a *vertex cover* is a set of vertices that includes at least one endpoint of every edge of the graph.

Remark 3.1. Establishing the hardness of the problem is most interesting for the case where the budget $t = |T|$ grows with $n = |V|$. Our problem (and the Vertex Cover problem) is fixed-parameter tractable, i.e., for bounded t one could solve by doing an exhaustive search of all subsets of size t which is polynomial in n (though only really feasible for very small t). The NP-hardness proof holds as you try, for instance, to have a number of influencers that are a fixed percentage of the total number of members of the network. While admittedly this may not be realistic in terms of resource allocation, the proof gives an understanding of how the computational complexity grows with network size in this capacity.

THEOREM 3.2. *There exists an* $\varepsilon > 0$ *such that it is NP-hard to obtain a* $(1-\varepsilon)$*-approximation to the influence targeting problem* (2.4) *when finding* $t(n)$ *influencers in a graph of size* n *with* $t(n) \to \infty$ *as* $n \to \infty$.

Proof of Theorem 3.2. We will use a gap-introducing reduction to the vertex cover problem in cubic graphs (where all vertices have degree exactly 3). Since this problem is APX-hard [4] by the PCP Theorem [5] there is a $\delta > 0$ and $t = t(n)$ so that it is NP-hard to decide the promise problem of whether a cubic graph with n vertices (and $m = 3n/2$ edges) has a vertex cover of size $\leq t$ (a `Yes`-instance) or if any subset of t vertices leaves at least δm edges uncovered (a `No`-instance).

Suppose you are given a cubic graph $G = (V, E)$ on n vertices that is a `Yes/No`-instance. Extend the graph G to G' by adding an extra node z and edges from each vertex v of G to z. Consider the influence maximization problem $\max_A J(A)$ with $|A| = t$ and $J(A) = \frac{1}{n+1}\sum_{i\in V} v(i)$ where $Lv(i) = 0$ for all $i \in V \setminus A$, $v(z) = 0$ and $v(i) = 1$ for all $i \in A$.

Note that since G is 3-regular any vertex cover has size at least $m/3 = n/2$ and at most n. Therefore we can assume $t = (1+\gamma)n/2$ for $\gamma \in [0, 1]$.

If G is a `Yes`-instance then by picking A to be the vertex cover on t vertices we can lower bound the optimal value by $J(A)$. For each $i \in V \setminus A$ we have $v(i) = \mathbb{P}_i\big(\tau(A) \leq \tau(z)\big) = 3/4$. This is because G is 3-regular and each one of those edges has its other endpoint in A and we added an edge to the vertex z, so the random walk hits A before z in one step with probability $3/4$. Therefore $\sum_{i\in V} v(i) = t + (3/4)(n-k)$. Using that $t = (1+\gamma)n/2$ we have that $(n+1)J(A) = (1+\gamma)n/2 + (3/4)(1-\gamma)n/2$ which simplifies to give

$$\max_A J(A) \geq J(A) = \left(\frac{7+\gamma}{8}\right)\frac{n}{n+1}.$$

If G is a `No`-instance then any subset A of t vertices leaves at least $\delta m = \delta 3n/2$ edges uncovered. Since G is 3-regular, there must be at least δn vertices with some uncovered edge. This is because there are $\delta 3n$ endpoints of the uncovered edges, if there were fewer than δn vertices by the pigeonhole principle one vertex would have to cover more than 3 of these endpoints which is impossible by the 3-regularity of G.

If i is one of these δn vertices then it has a neighbor j that is not in A, the random walk can move from i to j and then z, thus $v(i) = \mathbb{P}_i\big(\tau(A) \leq \tau(z)\big) \leq 1 - 1/4 - (1/4)^2 = 11/16$. Then, for any A we have that $(n+1)J(A) \leq t + (3/4)(n - t - \delta n) + (11/16)\delta n$. Therefore, using $t = (1+\gamma)n/2$ we can bound

$$\max_A J(A) \leq \left(\frac{7+\gamma}{8} - \frac{\delta}{16}\right)\frac{n}{n+1}.$$

If we were able to approximate the optimal value with arbitrary accuracy we would be able to determine if the graph G is a `Yes` or a `No` instance. ∎

3.2. Approximation of the targeting influence problem. For fixed $m \in [k]$ and a zealot set $Z = \amalg_{\ell\in[k]} Z_\ell$, we consider function $F_m \colon 2^{V\setminus Z} \to \mathbb{R}$, which takes a vertex subset $T \subset V \setminus Z$ and returns the value

$$F_m(T) := \mathcal{I}_m(v_T) = \frac{1}{|V|}\|v_T\|_{\ell^1(V)},$$

where v_T satisfies (2.5b)-(2.5d). We will prove that F_m is a monotone and submodular set function. Recall that f is a *monotone function* if for any $T_1, T_2 \subseteq V \setminus Z$ satisfying $T_1 \supseteq T_2$, we have $f(T_1) \geq f(T_2)$. We say that f is a *submodular set function* if for all $T \subseteq V \setminus Z$ and $x, y \in V \setminus Z$,

$$f(T \cup \{x\}) - f(T) \geq f(T \cup \{x, y\}) - f(T \cup \{y\}).$$

Intuitively, we think of submodularity of a function as giving diminishing returns as elements (vertices) are added to the set T.

For $T \subseteq V \setminus Z$, write $v_i(T)$ as the solution to the following equation evaluated at vertex $i \in V$

$$\begin{aligned} Lv(i) &= 0, && i \in V \setminus (Z \cup T), \\ v(i) &= 1, && i \in Z_m \cup T, \\ v(i) &= 0, && i \in Z_\ell,\ \ell \neq m. \end{aligned}$$

We first establish the following

LEMMA 3.3. *$v_i \colon 2^{V \setminus Z} \to \mathbb{R}$ is a monotone submodular set function.*

Proof. Fix an arbitrary subset $T \neq \varnothing$. Let $Z' = \cup_{\ell \neq m} Z_\ell$ and let $A = Z_m \cup T$. Recall the hitting time for a subset of vertices U by a random walk is defined as $\tau(U) = \inf\{s \geq 0 : X_s \in U\}$. By the probabilistic interpretation in subsection 2.5, we have that $v_i(T) = \mathbb{P}_i\left(\tau(A) < \tau(Z')\right)$, where $\mathbb{P}_i$ gives the distribution of the random walk started at i.

To prove that it is monotone note that if $T_1 = T$ and $T_2 = T \cup R$ we have

$$v_i(T_1) = \mathbb{P}_i\left(\tau(A) < \tau(Z')\right) \leq \mathbb{P}_i\left(\tau(A \cup R) < \tau(Z')\right) = v_i(T_2),$$

where the inequality follows from the fact that $\tau(A \cup R) \leq \tau(A)$.

Now we prove submodularity. Since $\tau(A \cup \{x\}) \leq \tau(A)$ we have that

$$\begin{aligned} v_i(T \cup \{x\}) - v_i(T) &= \mathbb{P}_i\left(\tau(A \cup \{x\}) < \tau(Z')\right) - \mathbb{P}_i\left(\tau(A) < \tau(Z')\right) \\ &= \mathbb{P}_i\left(\tau(A \cup \{x\}) < \tau(Z') \leq \tau(A)\right). \end{aligned} \tag{3.1}$$

Similarly, since $\tau(A \cup \{x, y\}) \leq \tau(A \cup \{y\})$ we have

$$\begin{aligned} v_i(T \cup \{x, y\}) - v_i(T \cup \{y\}) &= \mathbb{P}_i\left(\tau(A \cup \{x, y\}) < \tau(Z')\right) - \mathbb{P}_i\left(\tau(A \cup \{y\}) < \tau(Z')\right) \\ &= \mathbb{P}_i\left(\tau(A \cup \{x, y\}) < \tau(Z') \leq \tau(A \cup \{y\})\right). \end{aligned} \tag{3.2}$$

The event $\left\{\tau(A \cup \{x, y\}) < \tau(Z') < \tau(A \cup \{y\})\right\}$ is the random walk hitting x before Z' while avoiding $A \cup \{y\}$, this is contained in $\left\{\tau(A \cup \{x\}) < \tau(Z') \leq \tau(A)\right\}$ which is hitting x before Z' while avoiding just A. Therefore (3.2) is smaller than (3.1). That is

$$v_i(T \cup \{x\}) - v_i(T) \geq v_i(T \cup \{x, y\}) - v_i(T \cup \{y\}). \qquad \square$$

PROPOSITION 3.4. *The function $F_m \colon 2^{V \setminus Z} \to \mathbb{R}$ is a monotone and submodular set function satisfying*

- $F_m(V \setminus Z) = \frac{|V \setminus Z| + |Z_m|}{|V|} \leq 1.$

- $F_m(\varnothing) = \frac{1}{|V|}\|v_\varnothing\|_1 \geq 0$ *where* $v_\varnothing$ *satisfies*

$$\begin{aligned} Lv(i) &= 0, && i \in V \setminus Z, \\ v(i) &= 1, && i \in Z_m, \\ v(i) &= 0, && i \in Z_\ell,\ \ell \neq m. \end{aligned}$$

Proof. This follows from Lemma 3.3 and $\mathcal{I}_m(T) = F_m(v_T) = \frac{1}{|V|}\sum_{i\in V}(v_T)_i$. ∎

COROLLARY 3.5. *A greedy algorithm gives a* $1 - 1/e$ *approximation to the maximization of* F_m *over sets of a given size.*

Proof. The problem of maximizing a monotone submodular function f subject to a cardinality constraint admits a $1-1/e$ approximation algorithm. The approximation is given by a greedy method, which starts with the empty set S_0, and in step i adds the element of maximum increase to obtain S_i. Nemhauser et al. [37] proved that $f(S_k) \geq (1 - 1/e)\max_{|S|\leq k} f(S)$. ∎

We will further explore the computational complexity of implementing these greedy combinatorial algorithms in Section 5.

4. A relaxation strategy for the targeting influence problem. In subsection 2.3, we formulated a model for the targeting influence problem (2.4). This problem asks how an authority for opinion $m \in [k]$ would target network members in $V \setminus Z$ to maximize the proportion of influence that opinion m has on the network. In subsection 2.4, we showed that this problem can be recast as a grouped opponent targeting influence problem (2.5), where the authority does not have to consider all other opinions/ideas, but can lump them into a single group. In section 3, we recalled that there exists an $\varepsilon > 0$, such that a $(1-\varepsilon)$-approximation is NP-hard, but that the problem can be $(1 - 1/e)$-approximated with a polynomial-time algorithm. In this section we identify a convex relaxation of the NP-hard targeting influence problem.

For fixed $T \subset V \setminus Z$, $\varepsilon > 0$, and $m \in [k]$ we introduce the problem

$$\begin{aligned} &(4.1a) && Lu(i) + \varepsilon^{-1} e_T \odot (u - e_m)(i) = 0, && i \in V \setminus Z, \\ &(4.1b) && u(i) = e_\ell, && i \in Z_\ell,\ \ell \in [k]. \end{aligned}$$

Here, we have used the notation $e_T = \begin{cases} 1 & i \in T \\ 0 & i \notin T \end{cases}$ to be the characteristic vector for the set $T \subset V \setminus Z$ and $\odot$ to be the Hadamard product. Throughout, we thus take e_V to mean the vector of all ones.

LEMMA 4.1. *As* $\varepsilon \downarrow 0$*, the solution,* u_ε *of* (4.1) *converges to the solution* u *of the graph Laplace problem* (2.4b)–(2.4d)*; in particular,* $u_\varepsilon(i) \to e_m$ *for* $i \in T$*.*

Lemma 4.1 can be proved by considering the difference between solutions u_ε and u, and applying a discrete maximum principle on $V \setminus Z$.

For $\varepsilon > 0$, a relaxed formulation of the targeting influence problem (2.4) is then given by

$$\begin{aligned} &(4.2a) && \max_{\phi\colon V\to\mathbb{R}} \mathcal{I}_m(u) \\ &(4.2b) && \text{s.t. } Lu(i) + \varepsilon^{-1}\phi \odot (u - e_m)(i) = 0, && i \in V \setminus Z, \\ &(4.2c) && u(i) = e_\ell, && i \in Z_\ell,\ \ell \in [k], \\ &(4.2d) && \phi \geq 0,\ \phi|_Z = 0,\ \sum_{i\in V}\phi(i) = 1. \end{aligned}$$

Here, we have replaced the characteristic function $e_T\colon V \to \{0,1\}^n$ for $T \subset V \setminus Z$, $|T| = t$ by the function $\phi\colon V \to \mathbb{R}_+$, that is constrained to be supported on the set $V \setminus Z$ and satisfies $\sum_{i\in V} \phi(i) = 1$.

Remark 4.2. In (4.2d), we have constrained ϕ to satisfy $\|\phi\|_{\ell^1(V)} = 1$. Comparing to (2.4e), one might think that this constraint should be $\|\phi\|_{\ell^1(V)} = t$. However, this problem only depends on the ratio ε/t, so we can set $t = 1$ for convenience. This invariance signifies a shortcoming of the relaxed model; we do not expect that ϕ is supported on exactly t vertices.

Finally, as described in subsection 2.4, we can reduce this to an equivalent, us-vs-them binary problem. Writing $v = u_m$ (i.e. we will now let v represent the scalar problem associated to one component of the vector u) and again abusing notation by writing $\mathcal{I}_m(v) = \frac{1}{|V|}\|v\|_{\ell^1(V)}$, we have

$$
\begin{aligned}
&(4.3a) && \max_{\phi\colon V\to\mathbb{R}} \mathcal{I}_m(v), && \\
&(4.3b) && \text{s.t. } Lv(i) + \varepsilon^{-1}\phi \odot (v - e_V)(i) = 0, && i \in V \setminus Z,\\
&(4.3c) && v(i) = 1, && i \in Z_m,\\
&(4.3d) && v(i) = 0, && i \in Z_\ell,\ \ell \neq m,\\
&(4.3e) && \phi \geq 0,\ \phi|_Z = 0,\ \sum_{i\in V}\phi(i) = 1. &&
\end{aligned}
$$

Equation (4.3) is a relaxed version of the targeting influence problem that we will solve computationally. If $\phi^\star$ is the solution to (4.3), we obtain a set $T \subset V \setminus Z$ by choosing the largest t values of $\phi^\star$, arbitrarily breaking ties if necessary.

The following theorem shows that the objective function is strictly concave in ϕ and gives expressions for the gradient and Hessian. To state the theorem, we use the following notation. The function ϕ is constrained to have support on $Z^c = V \setminus Z$, so it is useful to partition the variables according to the sets $\{Z, Z^c\}$. For $v \in \mathbb{R}^{|V|}$, we write $v_c = v|_{Z^c}$ to denote the entries corresponding to Z^c. For $L \in \mathbb{R}^{|V|\times|V|}$, we write $L_{c,c}$ to denote the principle submatrix corresponding to $Z^c \times Z^c$. For ease of notation, for $L \in \mathbb{R}^{n\times n}$ and $\phi \in \mathbb{R}^n$, we simplify the writing of graph Schrödinger operators of the form $L + \operatorname{diag}(\phi) \in \mathbb{R}^{n\times n}$ by using the notation $L + \phi$.

THEOREM 4.3. *Let $G = (V,E)$ be a strongly connected, directed graph. For $\phi\colon V \to \mathbb{R}$ satisfying the constraints in* (4.3e)*, let v_ϕ denote the solution to equation* (4.3b) *satisfying the Dirichlet boundary conditions* (4.3c) *and* (4.3d)*. The map $\phi \mapsto \mathcal{I}_m(v_\phi)$ is strictly concave and twice differentiable. The gradient is given by*

$$\nabla_\phi \mathcal{I}_m\,|_c = \frac{1}{|V|}(e_V - v_c) \odot w_c \tag{4.4}$$

where $w_c\colon V \setminus Z \to \mathbb{R}$ is defined by $w_c := \varepsilon^{-1}\left[L_{c,c} + \varepsilon^{-1}\phi_c\right]^{-T} e$. The Hessian is given by

$$\nabla^2_\phi \mathcal{I}_m\,|_{c,c} = -\frac{2}{\varepsilon|V|}\ \mathit{sym}\left(\left[L_{c,c} + \varepsilon^{-1}\phi_c\right]^{-1} \odot \left[w_c \otimes (e_V - v_c)\right]\right), \tag{4.5}$$

where $\mathit{sym}(A) = (A + A^T)/2$.

Proof. We first compute the gradient and Hessian. We additionally need the notation that the submatrix of L corresponding to $V \setminus Z \times Z_m$ is denoted by $L_{c,m}$.

We can then rewrite (4.3b)-(4.3d) for the unknown v_c as

$$L_{c,c}v_c + L_{c,m}e_V + \varepsilon^{-1}\phi_c \odot (v_c - e_V) = 0. \tag{4.6}$$

We can write the solution v_c in terms of the Schur complement

$$v_c = [L_{c,c} + \varepsilon^{-1}\phi_c]^{-1}\left(\varepsilon^{-1}\phi_c - L_{c,m}e_V\right), \tag{4.7}$$

which is uniquely solvable since $\phi > 0$ and on any proper subset Z^c, the matrix $L_{c,c}$ is positive definite since the graph is strongly connected.

Differentiating (4.6) with respect to ϕ_j, with $j \in Z^c$ we obtain

$$\left(L_{c,c} + \varepsilon^{-1}\phi_c\right)\frac{\partial v_c}{\partial \phi_j} = \varepsilon^{-1} e_j \odot (e_V - v_c)$$

giving that

$$\frac{\partial v_c}{\partial \phi_j} = \varepsilon^{-1}[\left(L_{c,c} + \varepsilon^{-1}\phi_c\right)]^{-1}\left(e_j \odot (e_V - v_c)\right).$$

Thus,

$$\begin{aligned}
\frac{\partial \mathcal{I}_m}{\partial \phi_j}(v) &= \frac{1}{|V|}\sum_{i\in Z^c}\frac{\partial v}{\partial \phi_j}(i) = \frac{1}{|V|}\left\langle e_V, \frac{\partial v_c}{\partial \phi_j}\right\rangle \\
&= \varepsilon^{-1}\frac{1}{|V|}\left\langle e_V, \left[L_{c,c} + \varepsilon^{-1}\phi_c\right]^{-1}\left(e_j \odot (e_V - v_c)\right)\right\rangle \\
&= \varepsilon^{-1}\frac{1}{|V|}\left\langle (e_V - v_c) \odot \left[L_{c,c} + \varepsilon^{-1}\phi_c\right]^{-T} e_V, e_j\right\rangle \\
&= \frac{1}{|V|}\left\langle (e_V - v_c) \odot w_c, e_j\right\rangle.
\end{aligned}$$

This gives that $\nabla_\phi \mathcal{I}_m \mid_c = \frac{1}{|V|}(e_V - v_c) \odot w_c$, as desired.

Differentiating (4.6) a second time with respect to ϕ_k with $k \in Z^c$ yields

$$\left(L_{c,c} + \varepsilon^{-1}\phi_c\right)\frac{\partial^2 v_c}{\partial \phi_j \partial \phi_k} = -\varepsilon^{-1}\left(e_k \odot \frac{\partial v_c}{\partial \phi_j} + e_j \odot \frac{\partial v_c}{\partial \phi_k}\right).$$

Solving for $\frac{\partial^2 v_c}{\partial \phi_j \partial \phi_k}$ and using the definition of $\mathcal{I}_m$, we have that

$$\begin{aligned}
\frac{\partial^2 \mathcal{I}_m}{\partial \phi_j \partial \phi_k} &= \frac{1}{|V|}\left\langle e_V, \frac{\partial^2 v_c}{\partial \phi_j \partial \phi_k}\right\rangle \\
&= -\frac{1}{|V|}\left\langle w_c, e_k \odot \frac{\partial v_c}{\partial \phi_j} + e_j \odot \frac{\partial v_c}{\partial \phi_k}\right\rangle \\
&= -\frac{1}{|V|}\left(\langle w_c, e_k\rangle \left\langle e_k, \frac{\partial v_c}{\partial \phi_j}\right\rangle + \langle w_c, e_j\rangle \left\langle e_j, \frac{\partial v_c}{\partial \phi_k}\right\rangle\right),
\end{aligned}$$

where we have used the property that $\langle a, e_k \odot b\rangle = \langle a, e_k\rangle\langle b, e_k\rangle$. We then use this property again to compute

$$\begin{aligned}
\left\langle e_k, \frac{\partial v_c}{\partial \phi_j}\right\rangle &= \varepsilon^{-1}\left\langle \left[L_{c,c} + \varepsilon^{-1}\phi_c\right]^{-T} e_k, e_j \odot (e_V - v_c)\right\rangle \\
&= \varepsilon^{-1}\langle \Phi e_k, e_j\rangle \langle e_j, e_V - v_c\rangle,
\end{aligned}$$

where we have written $\Phi := \left[L_{c,c} + \varepsilon^{-T}\phi_c\right]^{-1}$. Substituting this above, we have

$$\frac{\partial^2 \mathcal{I}_m}{\partial\phi_j \partial\phi_k} = -\frac{1}{\varepsilon|V|}\left(\langle w_c, e_k\rangle \langle \Phi e_k, e_j\rangle \langle e_j, e_V - v_c\rangle + \langle w_c, e_j\rangle \langle \Phi e_j, e_k\rangle \langle e_k, e_V - v_c\rangle\right)$$
$$= \langle e_j, He_k\rangle,$$

where $H = -\frac{1}{\varepsilon|V|}\left(\Phi \odot [(e_V - v_c)\otimes w_c] + \Phi^T \odot [w_c \otimes (e_V - v_c)]\right)$ as desired.

We now prove strict concavity. Our argument follows the proof of [12, Prop. 2.4], using the Neumann series (i.e., $(I-T)^{-1} = \sum_{k=0}^{\infty} T^k$) in the Schur complement expression of the solution (4.7) and investigating convexity of each term. Let the degree matrix $D_{c,c}$ be the diagonal of $L_{c,c}$ and $d_{c,c}$ the corresponding vector along the diagonal. Writing $x = d_{c,c} + \varepsilon^{-1}\phi$ and $X = \mathrm{diag}(x) = D_{c,c} + \varepsilon^{-1}\mathrm{diag}(\phi)$ and noting $L_{c,m} = -A_{c,m}$, we have

$$\begin{aligned}
&(4.8a) && |V|\mathcal{I}_m(v_\phi) = \langle e_V, v\rangle \\
&(4.8b) && = \left\langle e_V,\ (X - A_{c,c})^{-1}\left(\varepsilon^{-1}\phi - L_{c,m}e_V\right)\right\rangle \\
&(4.8c) && = \left\langle e_V,\ \left(I - X^{-1}A_{c,c}\right)^{-1} X^{-1}(\varepsilon^{-1}\phi + A_{c,m}e_V)\right\rangle \\
&(4.8d) && = \sum_{k=0}^{\infty}\left\langle e_V,\ \left(X^{-1}A_{c,c}\right)^{k} X^{-1}(\varepsilon^{-1}\phi + A_{c,m}e_V)\right\rangle.
\end{aligned}$$

Let us reframe the problem in terms of the variable x. Then, following (4.8d), we define the function $E(x)$ such that

$$\begin{aligned}
|V|E(x) &:= \left\langle e_V,\ \left(I - X^{-1}A_{c,c}\right)^{-1} X^{-1}(x - d_{c,c} + A_{c,m}e_V)\right\rangle \\
&= \left\langle e_V, \sum_{k=0}^{\infty}(X^{-1}A_{c,c})^k(e_V - X^{-1}(d_{c,c} + A_{c,m}e_V))\right\rangle \\
&= \left\langle e_V, \sum_{k=0}^{\infty}(X^{-1}A_{c,c})^k e_V - \sum_{k=0}^{\infty}(X^{-1}A_{c,c})^k X^{-1}(d_{c,c} + A_{c,m}e_V)\right\rangle \\
&= \langle e_V, e_V\rangle + \left\langle e_V, \sum_{k=1}^{\infty}(X^{-1}A_{c,c})^k e_V - \sum_{k=0}^{\infty}(X^{-1}A_{c,c})^k X^{-1}(d_{c,c} + A_{c,m}e_V)\right\rangle \\
&= |Z^c| + \left\langle e_V, \sum_{k=0}^{\infty}(X^{-1}A_{c,c})^k X^{-1}A_{c,c}e_V - \sum_{k=0}^{\infty}(X^{-1}A_{c,c})^k X^{-1}(d_{c,c} + A_{c,m}e_V)\right\rangle \\
&= |Z^c| + \left\langle e_V, \sum_{k=0}^{\infty}(X^{-1}A_{c,c})^k X^{-1}\left(A_{c,c}e_V - d_{c,c} - A_{c,m}e_V\right)\right\rangle.
\end{aligned}$$

Notice that $A_{c,c}e_V - (d_{c,c} + A_{c,m})e_V$ is nonpositive because $A_{c,c}e_V$ is the degree of each node, counting only neighbors in c, whereas $d_{c,c}$ is the full degree.

We first show concavity of the function $E(x)$ defined above. Since the sum of convex functions is convex, it is enough to show that

$$e_V^T\left(X^{-1}A_{c,c}\right)^k X^{-1} f$$

is convex for each $k > 0$, for all nonnegative vectors f. For fixed k, and using the fact that e, f, and $A_{c,c}$ are nonnegative, the preceding term is a nonnegative linear combination of terms of the form

$$F(x) = \prod_i x_i^{-\alpha_i}.$$

Therefore, we seek to show that F is convex for any $\alpha = (\alpha_1, \cdots, \alpha_n)$ with each $\alpha_j \geq 0$. To prove this, we check whether the Hessian is semi-positive definite. Computing second derivatives gives

$$\frac{\partial^2 F}{\partial x_i \partial x_i}(x) = F(x)\alpha_i(\alpha_i + 1)x_i^{-2}$$

and

$$\frac{\partial^2 F}{\partial x_i \partial x_j}(x) = F(x)\alpha_i \alpha_j x_i^{-1} x_j^{-1}.$$

So the Hessian of F is

$$F(x)\left[(\alpha \odot x^{-1})(\alpha \odot x^{-1})^T + \operatorname{diag}(\alpha \odot x^{-2})\right].$$

The first term is positive semidefinite (PSD) since it is the outer product of a vector with itself and the second term is PSD since it is diagonal with positive entries. Thus, the Hessian of F is PSD.

To prove positive definiteness (rather than semi-definiteness), recognize that the $k = 0$ term contributes a term to the Hessian of the form DX^{-2}, which is strictly positive definite on the domain in question. This proves strict convexity. ∎

Remark 4.4. In subsection 2.6, we gave an energy interpretation in the case where the graph is undirected. For fixed $T \subset V \setminus Z$, $\varepsilon > 0$, and $m \in [k]$, define the relaxed energy

$$E_{m,\varepsilon}(u; T) = E(u) + \frac{1}{2\varepsilon}\sum_{i \in T} |u(i) - e_m|^2.$$

The minimizers of this energy satisfy (4.1). Furthermore, as $\varepsilon \downarrow 0$, the minimizer $u = u_\varepsilon$ will converge to the solution of the graph Laplace problem (2.4b)–(2.4d). In particular, $u_\varepsilon(i) \to e_m$ for $i \in T$.

4.1. Targeting influence on symmetric graphs. In this section, we consider the relaxation of the grouped opponent targeting influence problem (4.3) in the case when the graph has a symmetry and prove that the optimization shares the symmetry.

Recall that an *automorphism* of a graph (V, E) is a permutation $\pi \colon V \to V$ such that $(i, j) \in E$ if and only if $(\pi(i), \pi(j)) \in E$. An automorphism π induces a linear transformation P_π, on $\ell^2(V)$, defined by $(P_\pi(u)))_i = u_{\pi(i)}$, $u \in \ell^2(V)$, $i \in V$. The matrix $P_\pi \in \mathbb{R}^{|V|\times|V|}$ is a permutation matrix.

LEMMA 4.5. *Let $G = (V, E)$ be a strongly connected, directed graph. Suppose that π is an automorphism on (V, E) and let P_π be the induced linear transformation. Suppose $S \subset V$ is a permutation invariant set, i.e. $\pi S = S$. Let $f_i \in \ell^2(V \setminus S)$, $i = 1, 2$, and $f_3 \in \ell^2(S)$ and extend these functions by 0 to V and abuse notation by denoting the extensions by $f_i \in \ell^2(V)$, $i = 1, 2, 3$. Assume $f_i \geq 0$ and are invariant to P_π, $i = 1, 2, 3$, i.e., $P_\pi f_i = f_i$. Consider the equation*

$$Lv(i) + f_1(i)v(i) = f_2(i), \qquad i \in V \setminus S, \tag{4.9a}$$

$$v(i) = f_3(i), \qquad i \in S. \tag{4.9b}$$

Then the solution v is also invariant to P_π.

Proof. Since π is an automorphism, P_π commutes with the adjacency matrix A, the degree matrix D, and hence the graph Laplacian $L = D - A$. Multiplying both equations in (4.9) by P_π and using commutativity, we have

$$\begin{aligned} L(P_\pi v)(i) + (P_\pi f_1)(i)(P_\pi v)(i) &= (P_\pi f_2)(i), && i \in V \setminus S, \\ (P_\pi v)(i) &= (P_\pi f_3)(i), && i \in S, \end{aligned}$$

which shows that $P_\pi v$ is also a solution to (4.9). Here, we used that for vectors f, g, we have $P_\pi(f \odot g) = P_\pi f \odot P_\pi g$ since $(f \odot g)(i) = f(i)g(i)$. By uniqueness of the solution to (4.9), we have that $v = P_\pi v$. ∎

We similarly consider the relaxation of the grouped opponent targeting influence problem (4.3).

PROPOSITION 4.6. *Let π be an automorphism on the strongly connected, directed graph (V, E) with induced linear transformation P_π. Fix zealot sets $Z = \amalg_{\ell\in[k]} Z_\ell$. Suppose that Z_m and $\cup_{\ell\neq m} Z_\ell$ are permutation invariant sets. Fix $\varepsilon > 0$ and consider* (4.3)*. The maximizer $\phi^\star$ is invariant to P_π.*

Proof. Let $\phi^\star$ be the maximizer of (4.3) and let $v^\star$ be the corresponding harmonic function. Since V is finite, there exists an integer p, so that $P_\pi^p = I$. We consider the test function

$$\tilde{\phi} = \frac{1}{p} \sum_{j\in[p]} P_\pi^j \phi^*,$$

which is invariant to P_π. For each $j \in [p]$, denote the solution to (4.3b) with potential $\phi = P_\pi^j \phi^\star$ by v_j. We then use the linearity of the constraints to argue that $\tilde{v} = \frac{1}{p}\sum_{j\in[p]} v_j = \frac{1}{p}\sum_{j\in[p]} P_\pi^j v^\star$ is a solution for $\tilde{\phi}$. The strict concavity of $\mathcal{I}_m$ (Theorem 4.3) and Jensen's inequality then gives that

$$\mathcal{I}_m(\tilde{v}) > \frac{1}{p}\sum_{j\in[p]} \mathcal{I}_m(v_j) = \frac{1}{p}\sum_{j\in[p]} \mathcal{I}_m(v^\star) = \mathcal{I}_m(v^\star),$$

unless $\tilde{\phi} = \phi^*$. Here we use that the ℓ^1-norm is invariant to permutation of the argument. This strict inequality contradicts the optimality of $\phi^\star$. ∎

The results in Proposition 4.6 suggest that if a graph is symmetric with respect to an existing configuration of influencers, then the optimizer of the regularized problem (4.3) will be evenly distributed with respect to that symmetry operation. Hence, the optimizer of (2.5) may very well suggest more than the budgeted t vertices as optimal ways to optimize influence. We will see examples of this in Section 5. Choosing how to expend the budget of t nodes in this case is a topic we would like to explore in future work.

5. Computational methods and experiments.

In this section, we perform several numerical experiments to investigate and illustrate the targeting influence problem (2.5). We first describe our implementation of the two computational approaches to this NP-hard problem; see Algorithms 5.1 and 5.2. Both algorithms were implemented in MATLAB and run on a laptop computer.

5.1. Computational Methods.

We first implement a greedy algorithm to solve (2.5); see Algorithm 5.1. For each opinion m, we sweep through all non-zealot nodes, $i \in V \setminus Z$, and solve the version of (2.5) with Dirichlet boundary conditions appropriately chosen. For the relatively small size graphs considered here, we simply use

Algorithm 5.1 Greedy algorithm for approximately solving the grouped opponent targeting influence problem (2.5)

Input: directed graph, authority index $m \in [k]$, initial zealot sets $Z = \amalg_{\ell\in[k]} Z_\ell$, and an augmented zealot set size $t \in \mathbb{N} \setminus 0$.
Output: an updated zealot set $Z_m \cup T$ of size $|Z_m \cup T| = |Z_m| + t$.
Set $T = \varnothing$.
for $s \in [t]$ **do**
  for $i \in V \setminus (Z \cup T)$ **do**
    Set $\tilde{T} = T \cup \{i\}$ and solve

$$\begin{aligned} Lv(i) &= 0, && i \in V \setminus (Z \cup \tilde{T}),\\ v(i) &= 1, && i \in Z_m \cup \tilde{T},\\ v(i) &= 0, && i \in Z_\ell,\ \ell \neq m. \end{aligned}$$

    Compute the energy

$$E(i) = \mathcal{I}_m(v) = \frac{1}{|V|}\|v\|_{\ell^1(V)}.$$

  end for
  Compute $j = \text{argmax}_{i\in V\setminus(Z\cup T)}\ E(i)$, breaking ties randomly if needed.
  Set $T = T \cup \{j\}$.
end for

the MATLAB `backslash` operator to solve the Schur complement problem. After evaluating the energy for each non-zealot node, $E(i)$, we find the vertex which gives the largest energy and add it to the zealot set m. Since each iteration in s requires $O(n)$ linear solves each costing $O(n^3)$, the run time of this algorithm is $O(tn^4)$.

In Algorithm 5.2, we implement an algorithm based on the relaxation of the problem described in section 4. Given existing zealot sets, we use a quasi-Newton BGFS method in MATLAB via `fmincon` in order to approximate the unique optimal solution to (4.3). We solved (4.3a) using the MATLAB least squares function `lsqr`. Here, we dramatically speed up the numerical optimizations by including the gradient as computed in Theorem 4.3, which we verified by using the `Gradient Check` feature in the optimization solver. Each iteration in s requires the solution of a single optimization problem. The time complexity of computing the gradient is $O(n^3)$ and we perform $O(1)$ iterations to approximate the solution to the optimization problem. The run time for Algorithm 5.2 is then $O(tn^3)$, an $O(n)$ reduction over Algorithm 5.1. A choice of $\varepsilon > 0$ must be made and we discuss this choice in the examples below. Generally, we observe that for very small ε the algorithm prioritizes proximity to the existing zealot nodes, while for $\varepsilon \sim \|L\|^{-1}_{\text{Frobenius}}$, the graph dynamics play a more significant role.

In both algorithms, the initialization of zealot nodes can be done by, e.g., targeting a specific region of the graph, by random initial assignment, or by using a dynamic programming approach to identify the best possible first move for one of the label sets.

5.2. Computational experiments. In the general us-vs-them targeting influence problem (2.4) and its reduction to a group opponent problem (2.5), one may

Algorithm 5.2 Algorithm for approximately solving the grouped opponent targeting influence problem (2.5) utilizing the relaxation (4.3).

Input: directed graph, authority index $m \in [k]$, initial zealot sets $Z = \amalg_{\ell \in [k]} Z_\ell$, and an augmented zealot set size $t \in \mathbb{N} \setminus 0$.
Output: an updated zealot set $Z_m \cup T$ for of size $|Z_m \cup T| = |Z_m| + t$.
Set $T = \varnothing$.
for $s \in [t]$ **do**
Solve the convex optimization problem

$$\begin{aligned}
&\max_{\phi\colon V \to \mathbb{R}} \mathcal{I}_m(v) \\
&\quad \text{s.t. } Lv(i) + \varepsilon^{-1} \phi \odot (v - e_V)(i) = 0, && i \in V \setminus (Z \cup T), \\
&\qquad\quad v(i) = 1, && i \in Z_m \cup T, \\
&\qquad\quad v(i) = 0, && i \in Z_\ell,\ \ell \neq m, \\
&\qquad\quad \phi \geq 0,\ \phi|_{Z \cup T} = 0,\ \sum_{i \in V} \phi(i) = 1.
\end{aligned}$$

Compute $j = \operatorname{argmax}_{i \in V \setminus (Z \cup T)} \phi(i)$, breaking ties randomly if needed.
Set $T = T \cup \{j\}$.
end for

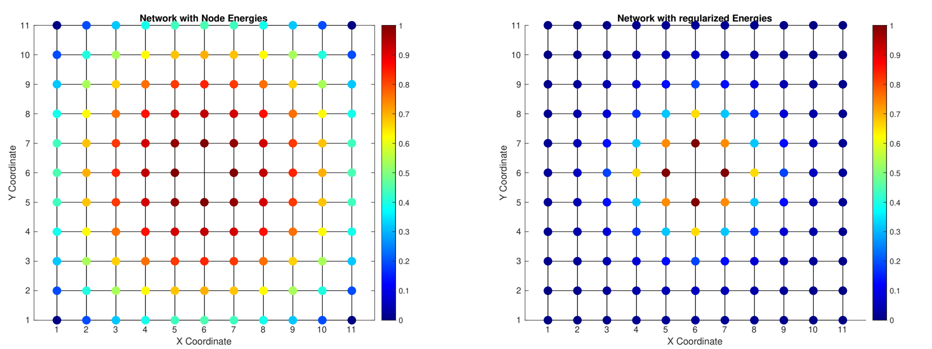

FIG. 1. *On an 11×11 square grid graph, we consider the case where first authority has chosen the center node and the second authority is choosing a vertex.* **(left)** *We plot the measure of influence $\tilde{\mathcal{I}}_2$ for the second authority for each available node.* **(right)** *We plot the $\tilde{\phi}$ value obtained as the solution to Algorithm 5.2 with $\varepsilon = .15$. The $\tilde{\phi}$ values respect the underlying symmetry and correctly predict that the optimal single-node choice is to pick one of the 4 nearest neighbors. See subsection 5.2.1.*

consider the case where multiple authorities $m \in [k]$ compete for non-zealot individuals in the network at the same time. For simplicity, in each experiment below, we consider a game between two opinion authorities to capture the largest proportion of influence. Each authority alternatively selects one vertex to convert to their opinion.

5.2.1. Square grid. In a first experiment, we consider an 11×11 square grid graph. The first opinion authority has chosen Z_1 to be the center node. In Figure 1, we compare the energy that the second opinion authority would realize choosing each node vs. the prediction via the relaxed problem (4.3). On the left of Figure 1, we plot

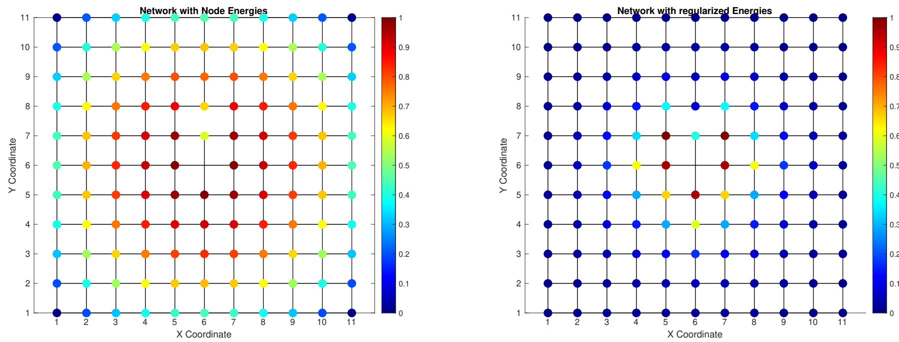

FIG. 2. *We again consider an* 11×11 *square grid graph, but now with a single edge missing. The left and right panels are as in Figure* 1. *See subsection* 5.2.2.

for each $i \in V \setminus Z_1$, a normalized energy

$$\tilde{\mathcal{I}}_2(i) := \frac{\mathcal{I}_2(i) - \min_j \mathcal{I}_2(j)}{\max_j \mathcal{I}_2(j) - \min_j \mathcal{I}_2(j)}, \qquad i \in V \setminus Z_1.$$

where v satisfies the boundary condition $v = 0$ on Z_1 and $v = 1$ on $Z_2 = \{i\}$. On the right, we solve (4.3) with $\varepsilon = .15$ and plot the normalized vertex function, normalized similar to above,

$$\tilde{\phi}(i) := \frac{\phi(i) - \min_j \phi(j)}{\max_j \phi(j) - \min_j \phi(j)}, \qquad i \in V \setminus Z_1.$$

As predicted by Proposition 4.6, this normalized vertex function $\tilde{\phi}$ has 4-fold symmetry. From the left plot, we see that it is advantageous to choose a neighboring node of Z_1 and the plot of $\tilde{\phi}$ in the right panel is a good predictor of this behavior.

5.2.2. Square grid with a defect. We repeat the first experiment (see subsection 5.2.1), but for an 11×11 square grid graph with a defect – the edge connecting vertex $(6, 7)$ with vertex $(6, 8)$ has been removed. The results are plotted in Figure 2. As predicted by Proposition 4.6, the vertex function $\tilde{\phi}$ has a horizontal reflection symmetry, but the vertical reflection symmetry has now been broken. From the left plot, we see that it is advantageous to choose a neighboring node of Z_1 that is below Z_1. The neighboring node above Z_1 is less advantageous; because of the missing edge, there are fewer edges for the opinion to propagate from that node. Again, the function from the relaxed problem plotted in the right does a good job predicting this asymmetry, and also highlighting that the nodes diagonal up to the right/left are preferable; while the specific rank order of these diagonal nodes relative to the nearest neighbors down, left and right differ slightly between the two panels, the impacts are very similar from each of these nodes.

5.2.3. H-graph. We next consider an 'H' graph, illustrated in Figure 3. The graph is obtained by connecting two 5×10 square grid graphs by a single edge ("bridge") connecting the $(5, 5)$ vertex of one with the $(1, 5)$ vertex of the other (at (6,5) in the figure). In the top panels of Figure 3, the first opinion authority has chosen Z_1 to be the $(3, 5)$ node on the left subgraph. In the bottom panels, the first opinion authority has chosen Z_1 to be the $(3, 5)$ node on both the left and right

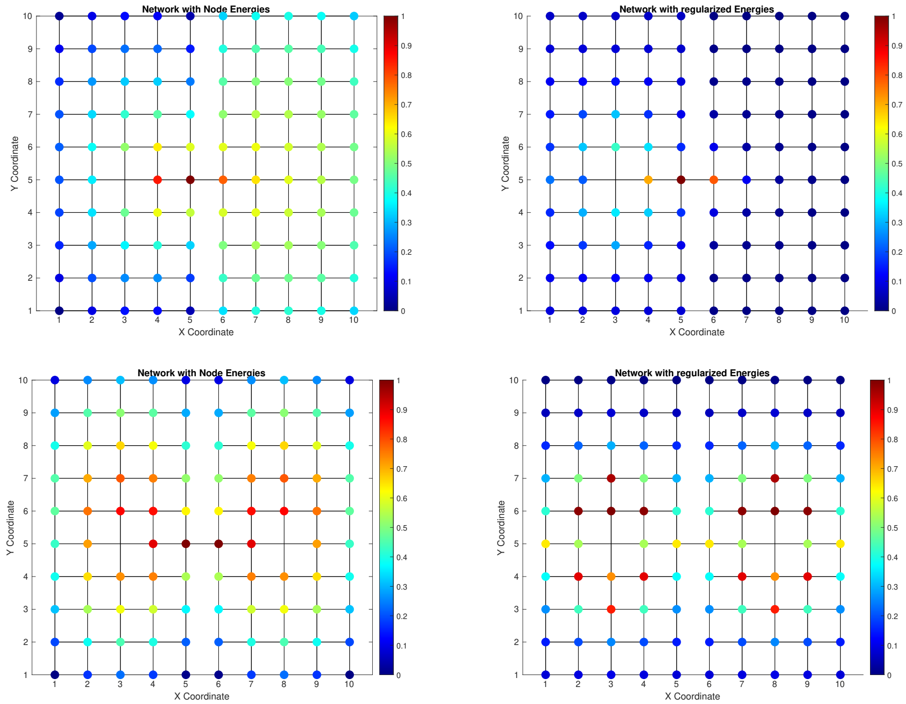

Fig. 3. *We consider an H-graph as described in subsection* 5.2.3 *and again consider the case where the first authority has chosen* Z_1 *and the second authority is choosing a vertex.* **(top)** *Here,* Z_1 *consists of a single node* $(3,5)$ *in the left subgraph. The left and right panels are as in Figure* 1. **(bottom)** *Here,* Z_1 *consists of a two nodes, one taken from the left and one taken from the right subgraphs. Again, The left and right panels are as in Figure* 1. *See subsection* 5.2.3 *for details.*

subgraphs. As before, in the left panels, we plot the energy that the second opinion authority would realize if they chosen each available node and, in the right panels, we plot the the prediction $\tilde{\phi}(i)$ via the relaxed problem (4.3) with $\varepsilon = .15$. In both cases, the vertices in the top half of the graph are slightly preferred to the bottom half. In the left panels, we see that there is a trade-off between proximity to the set Z_1 and the high betweenness centrality of the vertices near the bridge. Again, $\tilde{\phi}$ plotted in the right panels does a good job predicting this behavior.

In a second experiment on the H-graph, we consider varying the parameter $\varepsilon > 0$. We again consider the case where the first opinion authority has chosen Z_1 to be the $(3,5)$ nodes on both the left and right subgraphs. We now take $\varepsilon = .015$, a factor of 10 smaller than the previous value. We plot the function $\tilde{\phi}$ in Figure 4, which should be compared to the bottom right panel of Figure 3. We observe that as ε decreases, the function $\tilde{\phi}$ localizes near the set Z_1. In other words, as $\varepsilon \to 0$, ϕ will move to completely isolate the opinion class of Z_1. Since ϕ is not really constrained to consider a budget of a small number of nodes, this is the optimal way to minimize Z_1's influence. However, for ε chosen such that the global graph dynamics can be viewed, it can demonstrate further useful nodes for restricting spread of Z_1's opinion. An expanded version of our algorithm could actually be implemented to choose a response based upon a probability distribution set by the function ϕ with ε chosen at a scale that allows for spreading out amongst the graph.

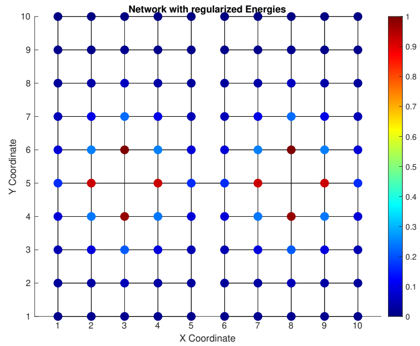

FIG. 4. *As in Figure 3, we consider an H-graph, but this time we plot $\tilde{\phi}$ for the value $\varepsilon = .015$, which is a factor of 10 smaller than the previous value. Compare to the bottom right figure of Figure 3. We observe that the algorithm correctly wants to cordon off the first authority opinion.*

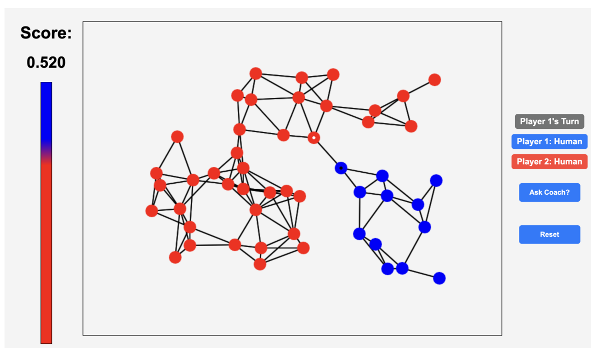

FIG. 5. *An illustration of the interactive game implementation of the targeting influence problem [11]. Here, we have an random geometric graph on 50 nodes. Two human players have each chosen a single node and Player 2 is currently winning; see subsection 5.2.4.*

5.2.4. Interactive game implementation. Finally, we implemented our algorithm, which can be found on github [10], in a graphical user interface at [11]. The graphical user interface is designed for non-experts and allows the user to select a graph type and size. The graph types available are spatial random graphs, tree graphs, ladder graphs, square lattices, hexagonal lattices, triangle lattices, and cycles. A version of the game played on an implementation of a random geometric graph with 50 vertices is shown in Figure 5. The user can also choose to have one or both players be played automatically. The automatic player can be set to random (easy), greedy (medium), and dynamic programming heuristics (hard) levels of play and are useful for developing intuition as well as for introducing the problem to popular audiences on a conceptual level.

6. Discussion. In this paper, we studied a model for targeting influence in a social network in settings where there are multiple extreme opinions — see (2.4) — along with introduction and study of a convex relaxation of the optimization problem. We first showed that each authority can group together the opposing opinions when considering how to maximize their own influence; see (2.5). For completeness, we recall results that the targeting influence problem is NP-hard to approximate; see Theorem 3.2. We also recalled that the objective function in the influence targeting problem is monotone and submodular and hence admits a $1 - 1/e$ approximation algorithm; see Proposition 3.4. In section 4, we developed an approximation strategy for solutions of the NP-hard targeting influence problem. We show that the objective function for the relaxed optimization problem (4.3) is strictly concave and compute the gradient and Hessian; see Theorem 4.3. We also proved that, in the case when the graph has a symmetry, the optimal solution to the relaxed problem also has the same symmetry; see Proposition 4.6. Finally, in section 5, we described a numerical implementation of our method and described the results of experiments to illustrate our results and test how well the relaxed optimization problem predicts the solution to the original problem.

There are several simplifications that we made in formulating the targeting influence problem (2.4), including the following.

1. Each social connection is parameterized by a single number (the link weight), whereas real social connections are highly complex. Indeed, we might even disapprove of the choices made by our acquaintances; this could be modeled by incorporating signed weights in the directed graph or allowing for multiple kinds of connections.
2. In the harmonic model for opinion dynamics used here, there is no time dependence. This is equivalent to assuming that the timescale of social opinion spread is faster than the timescale on which players can select and influence particular nodes, which may or may not be valid, depending on the situation. The assumption of linear dynamics on this short time scale is also not obviously true, and is one of the main departures of our contribution from previous approaches, such as cascading models. A very interesting line of future work would be to explore further when the various dynamics of opinion spread are appropriate.
3. We have modeled zealots to hold only extreme opinions. While this might be a reasonable assumption in some cases (for example, in advertising, sponsored influencers tend to push a particular product), in political elections it might be better to allow multiple preferences of zealots (for example, an activist might prefer anyone but candidate X).
4. We have assumed that an authority for a particular opinion can convert any chosen member to a zealot through sponsorship/incentives. But this might not be the case in all applications or the budget might have to account for different prices for conversion of different nodes. Our framework is extensible to such cases by incorporating costs for converting vertices, i.e., in (4.2d) or (4.3e), we could assign weights $w_i \geq 0$, $i \in V \setminus Z$ and require $\sum_{i \in V} w_i \phi(i) = 1$.
5. In the numerical examples, we have explored the case when the different authorities play a game, taking turns adding to their respective zealot sets. However, in practice, authorities compete for influence all at once.

Beyond addressing the model shortcomings outlined above, there are of course other interesting possible future directions for this work. In the game scenario where the authorities take turns choosing a single zealot to convert, we think it is an interesting question of whether the first player has an advantage. In the case of an undirected graph, we conjecture that the first player, if they play optimally, will always beat or tie the second player. In the case of a directed graph, this is not true. Consider a directed cycle graph where all orientations are the same. If the second player chooses the node adjacent and "downstream" to the node chosen by the first player, they will completely influence all nodes but one and win the turn.

In recent years, there have been many developments investigating the consistency of problems posed on geometric graphs in the limit as the number of sampled nodes tends to infinity (see, e.g., [14, 39, 47]). It would be interesting to consider this continuum limit here, where one might expect to obtain the Laplace problem on a fixed domain and authorities alternatively introduce boundary components where a Dirichlet condition is satisfied. A generalization that would allow Dirichlet data to be specified on lower dimensional sets is to consider the p-Laplacian for $p > 2$. A very recent paper [18] considers a continuum version of the continuous-time DeGroot model.

Acknowledgments. We thank Wesley Hamilton for working with them on a very early version of this idea and Alex McAvoy for helpful conversations about game theory interpretations of these models. We also thank Wilson Stoddard, Connor McBride, Taylor Larkins, Ian Goodwin, and other members of the Boyd Lab for help developing the interactive software for the game [10], which greatly helped develop our intuition. We gratefully acknowledge the Banff International Research Station and the Casa Mathematica Oaxaca for supporting our research efforts. We thank the anonymous referees for their helpful comments and, in particular, pointing us to useful references about the hardness of the unrelaxed model.

REFERENCES

[1] R. P. Abelson, *Mathematical models in social psychology*, in Advances in Experimental Social Psychology, L. Berkowitz, ed., vol. 3, Academic Press, 1967, p. 1–54, https://doi.org/10.1016/S0065-2601(08)60341-X, https://www.sciencedirect.com/science/article/pii/S006526010860341X.

[2] D. Acemoglu and A. Ozdaglar, *Opinion dynamics and learning in social networks*, Dynamic Games and Applications, 1 (2011), pp. 3–49, https://doi.org/10.1007/s13235-010-0004-1, https://doi.org/10.1007/s13235-010-0004-1.

[3] D. Acemoglu, A. E. Ozdaglar, and A. ParandehGheibi, *Spread of (mis)information in social networks*, SSRN Electronic Journal, (2009), https://doi.org/10.2139/ssrn.1401648, http://dx.doi.org/10.2139/ssrn.1401648.

[4] P. Alimonti and V. Kann, *Some apx-completeness results for cubic graphs*, Theoretical Computer Science, 237 (2000), pp. 123–134.

[5] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, *Proof verification and the hardness of approximation problems*, Journal of the ACM (JACM), 45 (1998), pp. 501–555.

[6] S. Banerjee, M. Jenamani, and D. K. Pratihar, *A survey on influence maximization in a social network*, Knowledge and Information Systems, 62 (2020), pp. 3417–3455.

[7] C. Bernardo, C. Altafini, A. Proskurnikov, and F. Vasca, *Bounded confidence opinion dynamics: A survey*, Automatica, 159 (2024), p. 111302, https://doi.org/10.1016/j.automatica.2023.111302.

[8] S. Bharathi, D. Kempe, and M. Salek, *Competitive influence maximization in social networks*, in Internet and Network Economics: Third International Workshop, WINE 2007, San Diego, CA, USA, December 12-14, 2007. Proceedings 3, Springer, 2007, pp. 306–311.

[9] D. Bindel, J. Kleinberg, and S. Oren, *How bad is forming your own opinion?*, Games and Economic Behavior, 92 (2015), pp. 248–265.

[10] Z. Boyd, I. Goodwin, T. Larkins, C. McBride, and W. Stoddard, *Influence maximization.* GitHub, 2024, https://github.com/zboyd2/influence-maximization.

[11] Z. Boyd, I. Goodwin, T. Larkins, C. McBride, and W. Stoddard, *Web version of influence maximization game.* GitHub, 2024, https://influence-maximization-93d86e9cfc16.herokuapp.com/.

[12] Z. M. Boyd, N. Fraiman, J. L. Marzuola, P. J. Mucha, and B. Osting, *An escape time formulation for subgraph detection and partitioning of directed graphs*, SIAM Journal on Matrix Analysis and Applications, 45 (2024), pp. 685–711.

[13] H. Z. Brooks, P. S. Chodrow, and M. A. Porter, *Emergence of polarization in a sigmoidal bounded-confidence model of opinion dynamics*, SIAM Journal on Applied Dynamical Systems, (2024), p. 1442–1470, https://doi.org/10.1137/22M1527258.

[14] J. Calder, D. Slepčev, and M. Thorpe, *Rates of convergence for Laplacian semi-supervised learning with low labeling rates*, Research in the Mathematical Sciences, 10 (2023), p. 10.

[15] M. H. DeGroot, *Reaching a consensus*, Journal of the American Statistical Association, 69 (1974), p. 118, https://doi.org/10.2307/2285509, http://dx.doi.org/10.2307/2285509.
[16] P. Domingos and M. Richardson, *Mining the network value of customers*, in Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, KDD01, ACM, Aug. 2001, https://doi.org/10.1145/502512.502525, http://dx.doi.org/10.1145/502512.502525.
[17] M. Fardad, X. Zhang, F. Lin, and M. R. Jovanović, *On the optimal dissemination of information in social networks*, in 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), IEEE, 2012, pp. 2539–2544.
[18] G. Favre, G. Jankowiak, S. Merino-Aceituno, and L. Trussardi, *Continuum modeling for opinion formation on a network*, 2024, https://arxiv.org/abs/2406.18076.
[19] N. E. Friedkin and E. C. Johnsen, *Social influence and opinions*, The Journal of Mathematical Sociology, 15 (1990), pp. 193—-206, https://doi.org/10.1080/0022250X.1990.9990069.
[20] S. Galam and F. Jacobs, *The role of inflexible minorities in the breaking of democratic opinion dynamics*, Physica A: Statistical Mechanics and its Applications, 381 (2007), p. 366–376, https://doi.org/10.1016/j.physa.2007.03.034, http://dx.doi.org/10.1016/j.physa.2007.03.034.
[21] A. Gionis, E. Terzi, and P. Tsaparas, *Opinion maximization in social networks*, in Proceedings of the 2013 SIAM international conference on data mining, SIAM, 2013, pp. 387–395.
[22] D. Golovin and A. Krause, *Adaptive submodularity: Theory and applications in active learning and stochastic optimization*, Journal of Artificial Intelligence Research, 42 (2011), pp. 427–486.
[23] M. Ji and J. Han, *A variance minimization criterion to active learning on graphs*, in Artificial Intelligence and Statistics, PMLR, 2012, pp. 556–564.
[24] D. Kempe, J. Kleinberg, and É. Tardos, *Maximizing the spread of influence through a social network*, in Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2003, https://doi.org/10.1145/956750.956769, https://doi.org/10.1145%2F956750.956769.
[25] D. Kempe, J. Kleinberg, and É. Tardos, *Influential nodes in a diffusion model for social networks*, in Automata, Languages and Programming, Springer Berlin Heidelberg, 2005, pp. 1127–1138, https://doi.org/10.1007/11523468_91, https://doi.org/10.1007%2F11523468_91.
[26] D. Kempe, J. M. Kleinberg, and E. Tardos, *Maximizing the spread of influence through a social network.*, Theory of Computing, 11 (2015), p. 105–147.
[27] P. P. Klamser, M. Wiedermann, J. F. Donges, and R. V. Donner, *Zealotry effects on opinion dynamics in the adaptive voter model*, Physical Review E, 96 (2017), p. 052315.
[28] J. M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment*, Princeton University Press, Dec. 2011, p. 514–542, https://doi.org/10.1515/9781400841356.514, http://dx.doi.org/10.1515/9781400841356.514.
[29] Y. Kuramoto, *Self-entrainment of a population of coupled non-linear oscillators*, Springer, Berlin, Heidelberg, 1975, pp. 420–422.
[30] Y. Li, J. Fan, Y. Wang, and K.-L. Tan, *Influence maximization on social graphs: A survey*, IEEE Transactions on Knowledge and Data Engineering, 30 (2018), pp. 1852–1872.
[31] Y. Ma, R. Garnett, and J. Schneider, *σ-optimality for active learning on gaussian random fields*, Advances in Neural Information Processing Systems, 26 (2013).
[32] K. Miller, H. Li, and A. L. Bertozzi, *Efficient graph-based active learning with probit likelihood via gaussian approximations*, arXiv preprint arXiv:2007.11126, (2020).
[33] K. Miller and R. Murray, *Dirichlet active learning*, arXiv preprint arXiv:2311.05501, (2023).
[34] K. S. Miller and A. L. Bertozzi, *Model change active learning in graph-based semi-supervised learning*, Communications on Applied Mathematics and Computation, 6 (2024), pp. 1270–1298.
[35] M. Mobilia, A. Petersen, and S. Redner, *On the role of zealotry in the voter model*, Journal of Statistical Mechanics: Theory and Experiment, 2007 (2007), p. P08029–P08029, https://doi.org/10.1088/1742-5468/2007/08/p08029, http://dx.doi.org/10.1088/1742-5468/2007/08/p08029.
[36] E. Mossel and S. Roch, *Submodularity of influence in social networks: From local to global*, SIAM Journal on Computing, 39 (2010), p. 2176–2188, https://doi.org/10.1137/080714452, http://dx.doi.org/10.1137/080714452.
[37] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, *An analysis of approximations for maximizing submodular set functions—i*, Mathematical programming, 14 (1978), pp. 265–294.
[38] B. Osting, C. Brune, and S. J. Osher, *Optimal data collection for informative rankings*

expose well-connected graphs, Journal of Machine Learning Research, 15 (2014), pp. 2981–3012, http://jmlr.org/papers/v15/osting14a.html.

[39] B. Osting and T. H. Reeb, *Consistency of Dirichlet partitions*, SIAM Journal on Mathematical Analysis, 49 (2017), p. 4251–4274, https://doi.org/10.1137/16m1098309, http://dx.doi.org/10.1137/16m1098309.

[40] S. Pei, J. Wang, F. Morone, and H. A. Makse, *Influencer identification in dynamical complex systems*, Journal of Complex Networks, 8 (2019), https://doi.org/10.1093/comnet/cnz029, https://doi.org/10.1093%2Fcomnet%2Fcnz029.

[41] A. F. Peralta, J. Kertész, and G. Iñiguez, *Opinion dynamics in social networks: From models to data*, 2022, https://arxiv.org/abs/2201.01322.

[42] S. Redner, *Reality-inspired voter models: A mini-review*, Comptes Rendus Physique, 20 (2019), pp. 275–292, https://doi.org/10.1016/j.crhy.2019.05.004.

[43] S. S. Singh, D. Srivastva, M. Verma, and J. Singh, *Influence maximization frameworks, performance, challenges and directions on social network: A theoretical study*, Journal of King Saud University-Computer and Information Sciences, 34 (2022), pp. 7570–7603.

[44] M. Taylor, *Towards a mathematical theory of influence and attitude change*, Human Relations, 21 (1968), pp. 121–139, https://doi.org/10.1177/001872676802100202.

[45] G. Verma, A. Swami, and K. Chan, *The impact of competing zealots on opinion dynamics*, Physica A: Statistical Mechanics and its Applications, 395 (2014), pp. 310–331.

[46] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, *Novel type of phase transition in a system of self-driven particles*, Physical Review Letters, 75 (1995), p. 1226–1229, https://doi.org/10.1103/physrevlett.75.1226, http://dx.doi.org/10.1103/physrevlett.75.1226.

[47] A. Yuan, J. Calder, and B. Osting, *A continuum limit for the pagerank algorithm*, European Journal of Applied Mathematics, 33 (2021), p. 472–504, https://doi.org/10.1017/s0956792521000097, http://dx.doi.org/10.1017/s0956792521000097.

[48] J. Zhao, Q. Liu, L. Wang, and X. Wang, *Competition between two sets of agents on complex networks*, in 2015 34th Chinese Control Conference (CCC), IEEE, 2015, pp. 1340–1344.

[49] J. Zhao, Q. Liu, and X. Wang, *Competitive dynamics on complex networks*, Scientific reports, 4 (2014), p. 5858.

[50] D. Zhou, J. Huang, and B. Schölkopf, *Learning from labeled and unlabeled data on a directed graph*, in Proceedings of the 22nd international conference on Machine learning (ICML), ACM Press, 2005, https://doi.org/10.1145/1102351.1102482, https://doi.org/10.1145%2F1102351.1102482.

[51] M. Zuluaga, G. Sergent, A. Krause, and M. Püschel, *Active learning for multi-objective optimization*, in International conference on machine learning, PMLR, 2013, pp. 462–470.