Shifting Opinions in a Social Network Through Leader Selection
Yuhaoy Yi, Timothy Castiglia, and Stacy Patterson

Abstract—We study French-DeGroot opinion dynamics in a social network with two polarizing parties. We consider a network in which the leaders of one party are given, and we pose the problem of selecting the leader set of the opposing party so as to shift the average opinion to a desired value. This problem generalizes the intensely studied problem of influence maximization. When each party has only one leader, we express the average opinion in terms of the Laplacian matrix of the network. The analysis shows balance of influence between the two leader nodes. We show that the problem of selecting at most \( k \) completely stubborn leaders to shift the average opinion is \( \text{NP}- \text{hard} \). Then, we reduce the problem to a problem of submodular maximization with submodular knapsack constraint and additional cardinality constraint, and propose a greedy algorithm with upper bound search to approximate the optimum solution. We also conduct experiments in random networks and real-world networks to show the effectiveness of the algorithm.

Index Terms—Social Network, French-DeGroot model, Balance of Opinions, Optimization, Approximation Algorithm.

I. INTRODUCTION

Social networks have become increasingly influential in shaping people’s opinions. Within this field, the problem of effectively shifting opinions in a social network has received great interest in last two decades [1, 2, 3, 4, 5, 6, 7, 8, 9]. Much of the existing work studies the problem of choosing nodes to maximize the influence of a particular opinion, or equivalently shifting the average opinion of the network to an extreme opinion. However, fine-grained optimization of the average opinion has not been well studied. In this paper we study the problem of shifting the average opinion of a network to a given value, which generalizes the intensely studied influence maximization problem.

We study the French-DeGroot opinion model with two polarizing parties. The French-DeGroot model [10, 11] is one of the most popular models for opinion dynamics. In the model, the social network is modeled by a graph, with nodes representing individuals. Each node (representing an individual) has a real scalar-valued state that represents the individual’s opinion. Each node updates its state continuously by comparing its state and the states of its neighbors. We consider a variation on this model where the nodes consists of leaders nodes, defined as the nodes with external reference values, and follower nodes, defined as those without external information.

We assume that there are two opposing parties, Party A and Party B, with opinion 0 and 1, respectively, and a node’s opinion represents its preference for the parties. Take an election as an example, a node with opinion 0 is an unquestioned supporter of Party A, while a node with opinion 1 is a firm supporter of Party B. A node with opinion in \((0, 1)\) has varying levels of preferences for Party A and B. A node with opinion \( \frac{1}{2} \) has the same preference for both parties. Each party controls a set of nodes as their opinion leaders. The leader nodes can be fully or partially controlled by each party. If a leader node is fully controlled, its opinion is set to the constant opinion value of that party and never changes over time. We call such a leader-follower system a completely stubborn leader system. If a leader node is partially controlled, it receives a constant input from the corresponding party as a reference value, and it adjusts its state according to the reference value and the states of its neighbors. It therefore exhibits some stubbornness to the reference value. We define such leader-follower systems as partially stubborn leader systems.

We consider the problem of selecting a leader set for party 1, given the leader set for party 0, so as to shift the average opinion of the social network towards a target value. For example, the target value could be 1 in order to maximize the influence of the party with opinion 1, or \( \frac{1}{2} \) in order to balance the opinions in the network. As an example, to shift the average opinion in an online forum that currently has only supporters for Party A, Party B can cultivate its supporters to change the opinions in the network.

We begin by analyzing the two proposed models, and we propose a concept of domination score to characterize the balance of influence between leaders of two parties. This analysis relates the models to the concepts of information centrality and effective resistance in a network. We also identify the optimal solution to the leader selection problem for each model when a single leader is chosen for each party. Next, we study the general problem of choosing a leader set for party 1 with a given cardinality. For completely stubborn leader systems, we prove the \( \text{NP}- \text{hardness} \) of the problem by a reduction from the vertex cover problem on 3-regular graphs. We also show the monotonicity and submodularity of the average steady-state opinion as a function of the leader set of party 1, for both completely stubborn and partially stubborn leader systems. Then, we propose an algorithm for the leader selection problems with provable approximation guarantees. Our algorithm searches an appropriate upper bound for a greedy routine that solves a submodular cost submodular knapsack (SCSK) problem.
Related work: In last two decades, many works considered the French-DeGroot model with leaders accessing the same reference value [12, 13, 14, 15, 19]. In such systems, leader selection problems have been formulated for different objectives such as maximizing the convergence error [19] or minimizing the total deviation from the reference value of the system in the presence of additional noises on followers [13, 14, 15]. These combinatorial optimization problems are often intractable. For example, the leader selection problem proposed in [13] has been proven to be NP-hard in [20]. Various approaches have been proposed to address these problems, including convex relaxation heuristics [14] and greedy algorithms [19, 21] with constant approximation ratios.

Another line of works consider leaders with different reference values, in particular two group of leaders with polarizing opinions. In this case, the steady-state opinions of all nodes fall in to the interval of leader states [17, 18]. In such systems, different leader selection problems have also been studied. [4] investigated the problem of single leader placement to maximize its influence. The work [8] studied a problem of choosing leaders to maximize influence of the leader set in a French-DeGroot model where leaders have specified stubbornness, and [5] investigated a similar maximization problem. Both works proved the monotonicity and submodularity of the average opinion in a French-DeGroot opinion network with partially stubborn leaders. [3] studied the influence maximization problem in Friedkin-Johnsen model, which is related to a French-DeGroot opinion network with completely stubborn leaders in special cases but not equivalent in general. This work proved the submodularity of the average opinion in their model as a function of leader nodes and the NP-hardness of the average opinion maximization problem. Typical greedy algorithms were applied to these problems due to submodularity of the objective functions. In contrast, our work studies the problem of shifting the average opinion of the network to any specified problem. This problem includes the influence maximization problem as a special case. In addition, we show that our problem cannot be directly treated as submodular maximization problem with a cardinality constraint. Thus, a more sophisticated optimization algorithm is needed.

Apart from influence maximization, other leader selection problems have also been investigated recently. [9] studied the diversity of opinions in the network characterized by entropy of the distribution of opinions in the network, while the analysis is restricted to special families of graphs. Another recent work [7] studies a problem of minimizing the disagreement and polarization in a Friedkin-Johnsen model.

Paper outline: The reminder of the paper is organized as follows. In Section II we introduce basic notations and concepts. In Section III, we present the system model and the problem formulation. In Section IV, we give an explicit form of the steady-state opinion vector using the Laplacian of an augmented graph, and we show how this relates to the balance of the leader nodes’ influence in a network. We also prove the hardness of the investigated problem in completely stubborn leader systems. In Section V, we propose a greedy algorithm with an upper bound search and provide provable bounds on the approximation ratio of the algorithm. Section VI presents experimental results. Finally, we conclude in Section VII.

II. Preliminaries

In this section, we introduce the notation of a graph and its matrix representations. Further, we review useful concepts of resistance distance and information centrality, which are used as analytical tools in this paper.

Vectors and Matrices: We use $e_u$ to denote the $u$-th canonical basis vector of $\mathbb{R}^n$. The vector $b_{u,v}$ is defined as $b_{u,v} \overset{def}{=} e_u - e_v$. $1_n$ represents the all-one vector with length $n$, and $0_n (0_{p \times q})$ represents the all-zero vector (or matrix) with length $n$ (or size $p \times q$). We also use these notations without specifying the sizes if they are implied in context. Apart from these exceptions ($e_u$, $b_{u,v}$, $1_n$ and $0_n$), a vectors or matrices with subscripts denotes the vector or submatrix with indices specified by the subscripts. For example, given a vector $x$, $x_i$ is its $i$-th entry, and $x_T$ is a vector consists of entries $x_i$ for all $i \in I$. For a matrix $X$, $X_{i,j}$ is the $(i,j)$-th entry of $X$ and $X_{I,J}$ is the submatrix of $X$ consisting of the entries of $X$ whose rows are in $I$ and columns are in $J$. In addition, we use $I$ to denote the identity matrix, and we use $X^\dagger$ to denote the Moore Penrose pseudoinverse of the matrix $X$.

Graph and its Matrix Representation: We denote an undirected graph as $G = (V, E, w)$, where $V$ and $E$ are the node set and edge set of the graph, respectively, with $|V| = n$ and $|E| = m$. We let $e = (u, v) \in E$ represent an edge between nodes $u$ and $v$, and $w : E \rightarrow \mathbb{R}^+$ is the edge weight function. We denote $N_v$ as the set of nodes adjacent to node $v$. In addition, for an undirected graph $G = (V, E, w)$, and a subset of nodes $V \subseteq V$, we denote the subgraph supported on $V$ as $G[V] = (V, E, \omega)$, where $E = \{e = (u, v) \in E : u, v \in V\}$ and $\omega(e) = w(e)$ for all $e \in E$. Further, we define the plus operation on graphs as follows. For two graphs $G_1 = (V_1, E_1, w_1)$ and $G_2 = (V_2, E_2, w_2)$, let $H = (U, M, \omega) = G_1 + G_2$ be a new graph with $U = V_1 \cup V_2$, $M = E_1 \cup E_2$, and $\omega : M \rightarrow \mathbb{R}^+$ the new edge weight function defined as $\omega(e) = w_1(e)$ if $e \in (V_1 \setminus V_2)$, $\omega(e) = w_2(e)$ if $e \in (V_2 \setminus V_1)$, and $\omega(e) = w_1(e) + w_2(e)$ if $e \in (V_2 \cap V_1)$.

The weighted Laplacian matrix of a graph $G$ is defined as $L^G = D - A$, where $A$ is the adjacency matrix with $A_{u,v} = w(e)$ for $e = (u, v) \in E$ and $A_{u,u} = 0$ for $(u, v) \notin E$, and $D$ is the degree diagonal matrix, where $D_{u,u} = \sum_v A_{u,v}$ and $D_{u,v} = 0$ if $u \neq v$. From the definition, it is clear that $L^G = \sum_{(u,v)\in E} w(u,v) b_{u,v} b_{u,v}^\dagger$. We sometimes use $L^G$ and $L$ interchangeably when context is clear.

Effective Resistance and Information Centrality: Given an undirected graph $G$ we define an electrical network $\overline{G}$. In $\overline{G}$, every edge $e$ of $G$ is replaced by a resistor of resistance $1/w(e)$, and the resistors are connected if the edges are incident. Then, the effective resistance between node $u$ and $v$ in graph $G$ (or electrical graph $\overline{G}$) is defined as the voltage difference between vertices $u$ and $v$ in $\overline{G}$ when unit current is injected from $u$ and extracted from $v$. We recall the following lemma relating to effective resistance.
Lemma II.1 (Effective Resistance [22]). In a connected undirected electrical network defined by \( G = (V, E, w) \), the effective resistance between nodes \( u \) and \( v \) is
\[
R_{u,v}^G = ((L^G)_{u,v} - 2((L^G)_{v,u} + ((L^G)_{u,u})_{u,u}).
\]

We further recall the related definition of information centrality.

Definition II.2 (Information Centrality [23]). In a connected undirected graph \( G = (V, E, w) \), the information centrality of a vertex \( u \) is defined by
\[
\theta^G(u) = \frac{1}{n} \sum_{v \in V} R_{u,v}^G.
\]

From Lemma II.1 we obtain
\[
\sum_{u \in V} R_{u,v}^G = n \cdot (L^G)_{v,v} + \text{Tr}(L^G).
\]

III. Problem Formulation

We consider an undirected connected graph \( G = (V, E, w) \). Nodes represent individuals in the social network, and edges model social links between nodes. Edge weights represent the strengths of the social links. Each node \( v \) has a scalar-valued state \( x_v \in \mathbb{R} \), which represents its opinion. The node set can be divided into a leader set \( S \) and a follower set \( F \). The leader set \( S \) can be further divided into two disjoint sets \( S_0 \) and \( S_1 \), which are leader sets for two parties, namely party 0 and party 1. All leaders of party 0 have access to reference value 0, and all leaders of party 1 have access to reference value 1. Followers update their states according to a diffusion law, while leaders exhibit stubbornness with respect to their reference values.

A. System Dynamics

We consider the French-DeGroot opinion model with completely stubborn leaders and a variation of this model where leaders are partially stubborn. The two considered models differ in how the leaders use their external reference values.

In the completely stubborn leader system, leaders initialize their states with 0 (for \( v \in S_0 \)) or 1 (for \( v \in S_1 \)), and their states remain unchanged over time. The dynamics of a leader node \( v \) is characterized by \( \dot{x}_v(t) = 0 \). A follower node \( v \) begins with an arbitrary initial state \( x_v(0) = x_v^0 \), and it updates its state by the dynamics
\[
\dot{x}_v(t) = -\sum_{u \in N_v} w(v,u)(x_v(t) - x_u(t)).
\]

We partition the state vector \( x \) as
\[
x = (x_S^\top, x_F^\top)^\top,
\]
where \( x_S \) is associated with the leaders and \( x_F \) is associated with the followers. Similarly, we partition the Laplacian matrix \( L^G \) and adjacency matrix \( A \) into blocks as
\[
L^G = \begin{pmatrix} L_{S,S} & L_{S,F} \\ L_{S,F} & L_{F,F} \end{pmatrix},
\]
and
\[
A = \begin{pmatrix} A_{S,S} & A_{S,F} \\ A_{S,F} & A_{F,F} \end{pmatrix}.
\]

Then, the dynamics of the leaders and the followers can be written as
\[
\dot{x}_S(t) = 0 \\
\dot{x}_F(t) = -L_{F,F}x_F - L_{F,S}x_S.
\]

In the system described by (1) and (2), the steady-state values of the leader nodes are
\[
\dot{x}_S = x_S^0 \\
\dot{x}_F = -(L_{F,F})^{-1}L_{F,S}x_S.
\]

We note that \( L_{F,S}x_S \) can be viewed as the sum of columns of \( L_{F,S} \) that correspond to 1-leaders (columns of 0-leaders are weighted by 0).

In the partially stubborn leader system, each leader updates its state according to its current state, the states of its neighbors, and its reference value. In other words, a leader has some specified stubbornness to its reference value. The system can start from any initial state and the dynamics is given by
\[
\dot{x}_v = -\sum_{u \in N_v} w(v,u)(x_v(t) - x_u(t)) + \kappa_v(1 - x_v(t)), v \in S_0,
\]
\[
\dot{x}_v = -\sum_{u \in N_v} w(v,u)(x_v(t) - x_u(t)) + \kappa_v(1 - x_v(t)), v \in S_1,
\]
\[
\dot{x}_v = -\sum_{u \in N_v} w(v,u)(x_v(t) - x_u(t)), v \in F.
\]

where the value \( \kappa_v \) is the weight that a leader put on its reference value. We also refer to it as the stubbornness of the node. The dynamics can be expressed more compactly as
\[
\dot{x} = -(L^G + E^S K)x + E^S_1 K1,
\]
where \( E^S \) is the diagonal matrix with \( E^S_{v,v} = 1 \) for \( v \in S \) and \( E^S_{v,u} = 0 \) otherwise; \( E^S_1 \) is defined similarly with non-zero entries for \( v \in S_1 \). The matrix \( K \) is diagonal with \( K_{v,v} = \kappa_v \), the stubbornness of vertex \( v \) if chosen as leader.

For system (5), \( -(L^G + E^S K) \) is Hurwitz for a non-empty leader set \( S \), the system converges to a single stable steady-state [24]. Letting \( \dot{x}_F(t) = 0 \), we obtain the steady-state of the followers
\[
\dot{x}_F = -(L_{F,F})^{-1}L_{F,S}x_S.
\]

In this paper, we study the average opinion of all nodes in the network.

Definition III.1. In either completely stubborn leader systems or partially stubborn leader systems, given the leader set \( S_0 \), the average opinion \( \mu \) of a network as a function of leader set \( S_1 \) is defined as
\[
\mu(S_1) \equiv \frac{1}{n} \sum_{v \in V} \dot{x}_v.
\]
Besides the above definition, \( \mu(S_1) \) has an interesting interpretation in a slightly modified random opinion model based on the French-DeGroot model. We can model the unknown factors in the system by treating \( \hat{x}_v \) as the success probability of a Bernoulli random variable \( X_v \) of taking the value 1. In the social network, we assume that \( X_v = 1 \) indicates the event that node (individual) \( v \) chooses opinion 1 over opinion 0, and \( X_v = 0 \) indicates the event that \( v \) chooses opinion 0 over opinion 1. We recall that \( n = |V| \) for both the completely and partially stubborn leader systems. We then assume \( X_1, X_2, \ldots, X_v, \ldots, X_n \) to be \( n \) mutually independent Bernoulli random variables associated with corresponding nodes in the network. In particular \( X_v \) is defined by

\[
\begin{align*}
\Pr(X_v = 1) &= \hat{x}_v, \\
\Pr(X_v = 0) &= 1 - \hat{x}_v,
\end{align*}
\]

for all \( v \in V \), and therefore \( E[X_v] = \hat{x}_v \).

We are interested in the fraction of nodes that choose opinion 1 over opinion 0. We define the random variable \( \bar{X} := \frac{1}{n} \sum_v X_v \). Since \( X_v \) are independent bounded random variables, \( \bar{X} \) concentrates at

\[
\mu(S_1) = \frac{1}{n} \sum_v \hat{x}_v. \tag{8}
\]

According to the Hoeffding’s inequality,

\[
\Pr\left( |\bar{X} - \mu(S_1)| \geq \sqrt{\frac{\ln n}{n}} \right) \leq \frac{2}{n^2}, \tag{9}
\]

which indicates that \( \mu(S_1) \) determines the ratio of population that choose opinion 1 over 0 in a large network, with a diminishing error bound and a diminishing probability that this bound is violated.

### B. Leader Selection Problems

In a completely stubborn leader system where the set \( S_0 \) with opinion 0 is given, we define the problem of choosing at most \( k \) leaders for opinion 1 as leader set \( S_1 \), such that the average opinion of all nodes (including leaders and followers) \( \mu(S_1) \) is closest to a given value \( \alpha \). Formally, the problem is defined as follows.

**Problem 1 (Completely Stubborn Leader Selection).** Given a connected undirected graph \( \mathcal{G} = (\mathcal{V}, \mathcal{E}, w) \), an opinion 0 leader set \( S_0 \neq \emptyset \), a specified value \( \alpha \in [0, 1] \), a candidate set \( \mathcal{Q} \subseteq \mathcal{V}\backslash S_0 \), \( |\mathcal{Q}| = q \), and an integer \( 1 \leq k \leq q \), we aim to find the node set \( S_1 \subseteq \mathcal{Q}, |S_1| \leq k \) such that

\[
S_1 \in \arg\min_{P \subseteq \mathcal{Q}, |P| \leq k} \| \mu(P) - \alpha \|. \tag{10}
\]

We define a similar problem for the partially stubborn leader system.

**Problem 2 (Partially Stubborn Leader Selection).** Given a connected undirected graph \( \mathcal{G} = (\mathcal{V}, \mathcal{E}, w) \), an opinion 0 (partially stubborn) leader set \( S_0 \neq \emptyset \), a stubbornness function of 0 leader nodes \( \kappa_0 : S_0 \rightarrow \mathbb{R}^+ \), a specified value \( \alpha \in [0, 1] \), a candidate set \( \mathcal{Q} \subseteq \mathcal{V}\backslash S_0 \), \( |\mathcal{Q}| = q \), another stubbornness function \( \kappa_1 : \mathcal{Q} \rightarrow \mathbb{R}^+ \), and a integer \( 1 \leq k \leq q \), find the node set \( S_1 \subseteq \mathcal{Q}, |S_1| \leq k \) such that

\[
S_1 \in \arg\min_{P \subseteq \mathcal{Q}, |P| \leq k} \| \mu(P) - \alpha \|. \tag{11}
\]

For both Problem 1 and 2, Influence maximization corresponds to the degenerate case of \( \alpha = 1 \).

### IV. Analysis

In this section, we give analytical solutions for Problems 1 and 2 for the case where \( k = 1 \). We also present hardness results for the case where \( k > 1 \).

Our analysis utilizes an augmented graph to give analytic expressions for the average opinion of the network with single leader for each party. We next define this augmented graph.

#### A. Opinions in Augmented Graphs

We note that the dynamics of both the completely stubborn leader system and the partially stubborn leader system can be fully characterized by a system defined in an augmented graph. For these two different kinds of systems, we construct the corresponding augmented graphs in different ways.

![Fig. 1: An example of constructing an augmented graph from a completely stubborn leader system. Nodes \( u \) and \( v \) in \( \mathcal{G} \) become the combined leader \( s'_0 \) in \( \mathcal{G}' \), and nodes \( i \) and \( j \) in \( \mathcal{G} \) become the combined leader \( s'_1 \) in \( \mathcal{G}' \). Edges without labels are weighted 1; otherwise, edges are labeled with their weights.](image)

The system described by (1) and (2) is equivalent to a system in which all nodes in \( S_0 \) are identified as a single completely stubborn leader \( s'_0 \), and all nodes in \( S_1 \) are identified as a single completely stubborn leader node \( s'_1 \). We refer to \( s'_0 \) and \( s'_1 \) as combined leaders. We denote the augmented graph by \( \mathcal{G}' = (\mathcal{V}', \mathcal{E}', w') \), where \( \mathcal{V}' = F \cup \{s'_0\} \cup \{s'_1\}, \mathcal{E}' = \{(u,v) : u,v \in F\} \cup \{(u,s'_0) : (N_u \cap S_0) \neq \emptyset\} \cup \{(u,s'_1) : (N_u \cap S_1) \neq \emptyset\}, and w'(u,v) = w(u,v) \) if \( u,v \in F, w'(u,s'_0) = \sum_{v \in (S_0 \cap N_u)} w(u,v), and w'(u,s'_1) = \sum_{v \in (S_1 \cap N_u)} w(u,v). In addition, we define \( s'_0 = \{s'_0, s'_1\} \) and \( F' = \mathcal{V}' \backslash S_0 \). Note that \( F' = F \) in this case. Figure 1 shows an example of constructing an augmented graph for a completely stubborn leader system.

We denote the Laplacian matrix of \( \mathcal{G} \) as \( L_\mathcal{G} \). Then the dynamics of \( F' \) in the system defined on the augmented graph is expressed by

\[
\hat{x}_{F'}(t) = -L_{\mathcal{G}'}, \hat{x}_{F'}, + E^{S_1}K'1, \tag{12}
\]

where \( E^{S_1} \) and \( K' \) are diagonal; \( E^{S_1}_{v,v} = 1 \) if \( v, s'_0 \) \( \in \mathcal{E}' \) and otherwise \( E_{v,v} = 0 \), and \( K'_{v,v} = w'(v, s'_0) \) if \( v, s'_0 \) \( \in \mathcal{E}' \), \( K'_{v,v} = w'(v, s'_1) \) if \( v, s'_1 \) \( \in \mathcal{E}' \), and \( K'_{v,v} = 0 \) otherwise.
shows an example of constructing an augmented graph from a partially stubborn leader system.

The system described by (5) is equivalent to a system in which two virtual completely stubborn leaders $s'_0$ and $s'_1$ are added to graph $G$, and all nodes in the original network $G$ are treated as followers. We define the augmented graph as $G' = (V', E')$, where $V' = V \cup \{s'_0\} \cup \{s'_1\}$, and $E' = E \cup \{(u, s'_0) : u \in S_0\} \cup \{(u, s'_1) : u \in S_1\}$, if $(u, v) \in E$, $w'(u, v) = w(u, v)$ if $(u, v) \in \mathcal{E}$, and $w'(u, s'_0) = \kappa_u$ if $u \in S_0$ and $w'(u, s'_1) = \kappa_u$ if $u \in S_1$. We again define $S' = \{s'_0, s'_1\}$ and $F' = V \setminus S'$; in this case $F' = V$. Figure 2 shows an example of constructing an augmented graph for a partially stubborn leader system. With this augmented graph, the dynamics of the system is also described by (12).

By constructing the corresponding augmented graphs, we can study both completely stubborn and partially stubborn leader systems using a unified framework. We remark that this does not mean the systems are equivalent. Choosing leaders in different system model leads to different augmented graphs and hence different steady-states, although system (2) approaches system (5) as $K_{v,v} \to +\infty$ for all $v \in (S_0 \cup S_1)$.

For both the completely stubborn and partially stubborn leader systems, the combined leaders (or virtual leaders) $s'_1$ and $s'_0$ are the only nodes that directly use reference values as their states in the augmented graph. Their steady states are

\[ x_{s'_0} = 0, \quad x_{s'_1} = 1. \]

The steady states of all remaining nodes satisfy

\[ L_{G', F'}^{S'} \tilde{x}_{G'} = E^{S'} K' \tilde{x}_{G'}. \]

**Proposition IV.1.** For either a completely stubborn leader system or a partially stubborn leader system, we consider its augmented graph $G'$. For any node $v \in V'$ in an undirected augmented graph, the steady state value $\tilde{x}_v$ is given by

\[ \tilde{x}_v = \left( b_{v, s'_0}^\top L_{G'}^{S'} \right)^+ b_{s'_1, s'_0}. \]

**Proof.** Substituting $v$ for $s'_0$ and $s'_1$, we immediately obtain $x_{s'_0} = 0$ and $x_{s'_1} = 1$.

Then we consider the $v$-th entry of (15). Plugging (16) into left hand side of (15) we get

\[ \left( L_{G', F'}^{G'} \tilde{x}_{G'} \right)_v = \left( E_{v, v}^{S} K_{v, v} b_{v, s'_0}^\top + \sum_{u \in (N_v \cap F')} w(v, u) \left( b_{v, s'_0}^\top - b_{u, s'_0}^\top \right) \left( L_{G'}^{S'} \right)^{\top} b_{s'_1, s'_0} \right) b_{s'_1, s'_0}. \]

Since for $v \not\in \{s'_0, s'_1\}$,

\[ e_v^\top L_{G'}^{S'} \left( L_{G'}^{S'} \right)^{\top} b_{s'_1, s'_0} = e_v^\top \left( I - \frac{1}{n} \mathbf{1} \mathbf{1}^\top \right) b_{s'_1, s'_0} = 0, \]

which means for $v \not\in \{s'_0, s'_1\}$

\[ e_v^\top \sum_{(i,j) \in \mathcal{E}} w(i, j) b_{i,j} b_{i,j}^\top \left( L_{G'}^{S'} \right)^{\top} b_{s'_1, s'_0} = 0. \]

Again this is equivalent to

\[ \left( \sum_{u \in N_v} w(v, u) \left( b_{v, s'_0}^\top - b_{u, s'_0}^\top \right) \left( L_{G'}^{S'} \right)^{\top} b_{s'_1, s'_0} \right) b_{s'_1, s'_0} = 0. \]

We let

\[ c \eqdef \frac{L_{G'}^{S'} b_{s'_1, s'_0}}{b_{s'_1, s'_0}}, \]

then for $v$ that satisfies $v \not\in E'$, we have

\[ \left( E_{v, v}^{S} K_{v, v} b_{v, s'_0} + \sum_{u \in (N_v \cap F')} w(v, u) \left( b_{v, s'_0}^\top - b_{u, s'_0}^\top \right) \right) c = 0, \]

If $(v, s'_1) \in E'$,

\[ \left( E_{v, v}^{S} K_{v, v} b_{v, s'_0} + \sum_{u \in (N_v \cap F')} w(v, u) \left( b_{v, s'_0}^\top - b_{u, s'_0}^\top \right) \right) c = K'_{v, v} b_{v, s'_0} c. \]

By comparing to the right hand side of (15), we conclude that (16) satisfies all equations required.

Since $L_{G', F'}^{G'}$ is full rank and $E^{S'} K' 1$ is not zero, the system of equations has a unique solution.

**B. Single Leader for Each Party**

For completely stubborn leader systems, if $|S_1| = |S_0| = 1$, the augmented graph $G'$ is the same as the original graph $G$. We let the leaders in $G$ be denoted $s_0$ and $s_1$ for parties with opinion 0 and 1, respectively. Then, by Proposition IV.1,

\[ \mu(S_1) = \frac{\left( L_{G_1}^{S_1} - L_{G_0}^{S_0} \right)_{s_0, s_1} + \left( L_{G_0}^{S_0} - L_{G_0}^{S_0} \right)_{s_1, s_0}}{\left( L_{G_0}^{S_0} - L_{G_0}^{S_0} \right)_{s_0, s_1} + \left( L_{G_0}^{S_0} - L_{G_0}^{S_0} \right)_{s_1, s_0}}. \]

Intuitively, we can view this expression as the influence of node $s_0$ over $s_1$, normalized by the effective resistance between them. We quantify this influence with the following definition.

**Definition IV.2.** In a connected undirected graph $G$, the domination score of node $u$ over $v$ is defined as

\[ D_{u,v}^G = (L_{G'})_{u,v}^\top - (L_{G'})_{v,u}^\top. \]
Then from the definition of domination score and the definition of effective resistance in Lemma II.1, we obtain
\[ \mu(S_1) = \frac{D_{\hat{G}}^{s_1,s_0}}{R_{\hat{G}}^{s_0,s_1}}. \tag{26} \]

**Theorem IV.3.** For completely stubborn leader systems, if \(|S_0| = |S_1| = 1\),
\[ |\mu(S_1) - \alpha| = \frac{|(1-\alpha)D_{\hat{G}}^{s_1,s_0} - \alpha D_{\hat{G}}^{s_0,s_1}|}{R_{\hat{G}}^{s_0,s_1}}. \tag{27} \]

The proof of Theorem IV.3 follows directly from (24), Definition IV.2, and Lemma II.1. The numerator is the absolute value of a weighted average of \(D_{\hat{G}}^{s_1,s_0}\) and \(-D_{\hat{G}}^{s_0,s_1}\). Therefore, Theorem IV.3 shows a weighted balance between domination score of \(s_0\) over \(s_1\) and the domination score of \(s_1\) over \(s_0\), which decides the deviation of the average opinion from \(\alpha\).

Theorem IV.3 indicates that for Problem 1, if \(|S_0| = |S_1| = 1\), given the leader \(s_0\), it suffices to find a node \(s_1\) such that \((1-\alpha)D_{\hat{G}}^{s_1,s_0} = \alpha D_{\hat{G}}^{s_0,s_1}\) to shift the average opinion to \(\alpha\).

For partially stubborn leader systems, the vector \(\hat{x}\) is given by (6). We do not apply the augmented graph analysis in this case because \(\hat{G} \neq \hat{G}'.\) We instead interpret \(\hat{x}\) using properties of \(\hat{G}'.\) Fortunately, when we choose one leader for each party, \(E^{s_1}\) is a rank-1 matrix, and \(E^S = E^{s_0} + E^{s_1}\) is a rank-2 matrix. Applying the rank-1 update of matrices twice leads to the following theorem.

**Theorem IV.4.** For partially stubborn leader systems, if \(|S_0| = |S_1| = 1\),
\[ |\mu(S_1) - \alpha| = \frac{|(1-\alpha)(D_{\hat{G}}^{s_1,s_0} + 1/\kappa_0) - \alpha (D_{\hat{G}}^{s_0,s_1} + 1/\kappa_1)|}{R_{\hat{G}}^{s_0,s_1} + 1/\kappa_0 + 1/\kappa_1}. \]

**Proof.** According to the Sherman-Morrison formula,
\[ \hat{x}_v = e_v^\top (L + E^{s_0} \kappa_0)^{-1} - \frac{\kappa_1 (L + E^{s_0} \kappa_0)^{-1} e^{s_1} (L + E^{s_0} \kappa_0)^{-1} e^{s_1}}{1 + \kappa_1 e^{s_1} (L + E^{s_0} \kappa_0)^{-1} e^{s_1}} e^{s_1} \kappa_1. \tag{28} \]

Let us then consider \((L + E^{s_0} \kappa_0)^{-1}\). Since \(L\) is a singular matrix, the Sherman-Morrison formula cannot be applied in this case. Instead we apply the rank-1 update given in [25].
\[ (L + E^{s_0} \kappa_0)^{-1} = L^1 - (L^1 e^{s_0}) 1^\top 1 - (L^1 e^{s_0}) 1^\top \]
\[ + (1/\kappa_0 + e^{s_0} L^1 e^{s_0}) 11^\top. \tag{29} \]

Plugging (29) into (28), we arrive at
\[ \hat{x}_v = \frac{b^0_{s_0} L^1 b_{s_1,s_0} + 1/\kappa_0}{1/\kappa_1 + 1/\kappa_0 + b_{s_1,s_0} L^1 b_{s_1,s_0}}. \tag{30} \]

Similar to the completely stubborn leader case we have
\[ \mu(S_1) = \frac{D_{\hat{G}}^{s_1,s_0} + 1/\kappa_0}{R_{\hat{G}}^{s_0,s_1} + 1/\kappa_0 + 1/\kappa_1}, \tag{31} \]
and we obtain the desired result.

As we observed in Theorem IV.3 for completely stubborn leader systems, for partially stubborn leader systems, Theorem IV.4 also shows the balancing behavior of domination scores in the social network, which decides the deviation of average opinion from \(\alpha\). In addition, Theorem IV.4 indicates that for Problem 2, if \(|S_0| = |S_1| = 1\), given the leader \(s_0\), it suffices to find a node \(s_1\) such that \((1-\alpha)(D_{\hat{G}}^{s_1,s_0} + 1/\kappa_0) = \alpha (D_{\hat{G}}^{s_0,s_1} + 1/\kappa_1)\) to shift the average opinion to \(\alpha\). Assuming \(\kappa_1 = \kappa_2\), then the condition is the same as what we have derived in the completely stubborn leader system.

The balancing behaviors shown in Theorem IV.3 and IV.4 exhibit interesting results when \(\alpha = 1/2\). In particular, Theorems IV.3 and IV.4 imply the following corollaries.

**Corollary IV.5.** For completely stubborn leader systems, when \(\alpha = 1/2\) and \(|S_0| = |S_1| = 1\),
\[ |\mu(S_1) - 1/2| = \frac{|\theta^{\hat{G}}(s_0) - 1 - \theta^{\hat{G}}(s_1) - 1|}{2R_{\hat{G}}^{s_0,s_1}}. \tag{32} \]

**Corollary IV.6.** For partially stubborn leader systems, when \(\alpha = 1/2\) and \(|S_0| = |S_1| = 1\),
\[ |\mu(S_1) - 1/2| = \frac{|\theta^{\hat{G}}(s_0) - 1 + 1/\kappa_0 - \theta^{\hat{G}}(s_1) - 1 - 1/\kappa_1|}{2(R_{\hat{G}}^{s_0,s_1} + 1/\kappa_0 + 1/\kappa_1)}. \tag{33} \]

These corollaries shows the role information centrality of leader nodes when the objective is to balance the opinions in the opinion network. If \(s_1\) has the same information centrality as \(s_0\) (when \(\kappa_1 = \kappa_0\) for partially stubborn leader systems), then \(\mu(S_1) = 1/2\), and so the opinion network is balanced. If there is no such an \(s_1\), then its beneficial to find a node \(s_1\) such that \(|\theta^{\hat{G}}(s_1) - \theta^{\hat{G}}(s_0)|\) is small while \(R_{\hat{G}}^{s_0,s_1}\) is relatively large.

**C. Hardness of Choosing Optimal \(k\) Leaders**

We now prove the hardness of Problem 1. The hardness of Problem 2 remains an open question.

Although [3] studied a different problem, we can find instances of Problem 1 that are equivalent to the hard instances in the problem studied in [3]. Then a reduction follows from the vertex cover problem in 3-regular graphs, which is an NP-complete problem [26].

**Problem 3 (Vertex Cover on 3 Regular Graphs).** Given an undirected connected 3-regular graph \(G = (V,E)\) and an integer \(k\), decide whether there is a vertex set \(S_1 \subseteq V\) such that \(|S_1| \leq k\) and \(|S_1|\) is a vertex cover of graph \(G\).

We give a decision version of Problem 1 as follows.

**Problem 4 (Completely Stubborn Leader Selection Decision Problem).** Given a connected undirected graph \(G = (V,E)\), an edge weight function \(w : E \rightarrow \mathbb{R}^+\), an opinion 0 leader set \(S_0 \neq \emptyset\), two real numbers \(\alpha, \beta \in [0,1]\), a candidate set \(Q \subseteq \mathbb{R}\), \(|Q| = q\), and an integer \(1 \leq k \leq q\), decide whether there is a leader set \(S_1 \subseteq Q\) with opinion 1 with at most \(k\) nodes, such that the average opinion of all nodes (including leaders and followers) \(\mu(S_1) = \frac{1}{|S_1|} \sum_{v=1}^{n} \hat{x}_v\), satisfies \(|\mu(S_1) - \alpha| \leq \beta\).

**Lemma IV.7.** Given an instance of problem 4, it is NP-hard to decide if there is a set \(S_1\), \(|S_1| \leq k\), such that \(|\mu(S_1) - \alpha| \leq \beta\)
Proof. Let \( \mathcal{F} = (V, E, w) \) be a graph consisting of a star graph \( S_n \), plus a 3-regular subgraph \( \mathcal{F}[V] = (V, E, \omega) \) supported on \( n - 1 \) leaves of \( S_n \). Edges in \( S_n \) are weighted 3 and edges in \( \mathcal{F}[V] \) are weighted 1. Then, we can construct an instance of Problem 1 by letting \( S_0 = \{ s_0 \} \) be the central node of \( S_n \), and the candidate set \( Q \) be the node set \( V = V \setminus \{ s_0 \} \), and \( k \) be any integer that satisfies \( 1 \leq k \leq q \).

Completeness: If \( |S_1| = k \) and \( S_1 \) is a vertex cover of the 3-regular graph \( G = \mathcal{F}[V] \), then we consider the steady-state of the followers \( \hat{x}_F \) given by (4). In this case, \( \hat{x}_F = \text{diag}(\{6, \ldots, 6\})^{-1} [3, \ldots, 3]^T = [\frac{1}{2}, \ldots, \frac{1}{2}]^T \). There are \( n - 1 - k \) followers nodes; thus, we have \( \mu(S_1) = \frac{1}{n} (\frac{1}{2} \cdot (n - 1 - k) + k) = \frac{n - 1 + k}{2n} \).

Soundness: If \( S_1 \) is not a vertex cover of graph \( G \), then the follower node set is not an independent set. So, the matrix \( L_{F,F} \) is a block diagonal matrix with each block associated with a connected component of graph \( G[V \setminus S] \). Let \( T \subseteq V \setminus S \), \( |V| \geq 1 \) be the node set of a connected component. Following the analysis given in the proof of [3, Theorem 4.1], we obtain \( \hat{x}_u < \frac{1}{2} \) for any \( u \in T \).

Next, we give a polynomial reduction form VC3 to CSLSD: \( p : \{ G = (V, E), k \} \rightarrow \{ G = (V, E, w), Q, k, \alpha, \beta \} \). For any given 3-regular graph \( G \) with \( n - 1 \) nodes, we construct a weighted graph \( G = G + S_n \), with all edges in the original graph weighted 1 and all edges in the star \( S_n \), weighted 3. Let \( Q = V \), \( k \) be the same integer, \( \alpha \) be any constant \( t \) greater or equal to \( c = \frac{n - 1 + k}{2n} \), and \( \beta = t - c \). Then \( \mu(G = (V, E), k) = (G + S_n, V, k, t - c) \) is a reduction from VC3 to CSLSD.

Lemma IV.7 immediately implies the following theorem

**Theorem IV.8.** The Completely Stubborn Leader Selection problem for shifting social opinion, described in Problem 1, is NP-hard.

We note that in both Problem 1 and 2, \( \mu(S_1) \) as a function of \( S_1 \) is monotone and submodular.

**Theorem IV.9.** For both completely stubborn leader systems and partially stubborn leader systems, the set function \( \mu(S_1) \) is monotone and submodular.

The monotonicity and submodularity of \( \mu(S_1) \) for partially stubborn systems follow in a straightforward manner from similar results in [5], [8]. We are unaware of prior analogous results for completely stubborn systems. We give simple proofs for both cases in the appendix. Our proofs are based on analyzing the escape probabilities of random walkers in the network.

**V. Algorithm**

In this section, we present an algorithm for shifting the average opinion of a network, with a given set of leaders \( S_0 \), to a specified value \( \alpha \), by selecting a set of nodes to be leaders in \( S_1 \). This algorithm can be used for both Problems 1 and 2.

We start by recalling the concept of \( \epsilon \)-approximation [27]:

**Definition V.1.** Given two numbers \( a, b \in \mathbb{R}, a, b \geq 0 \), if

\[
\exp(-\epsilon)a \leq b \leq \exp(\epsilon)a,
\]

we say \( a \) is an \( \epsilon \)-approximation of \( b \), and we denote this by \( a \approx \epsilon b \).

Note that \( a \approx \epsilon b \) if and only if \( b \approx \epsilon a \).

Based on the definition of \( \epsilon \)-approximation, we can also define an \( \epsilon \)-approximation algorithm.

**Definition V.2.** In an optimization problem with cost function \( f(X) \), if there exists an algorithm \( A \) that returns a solution \( S \) that satisfies

\[
f(S) \approx \epsilon, OPT
\]

for any instance of the problem, where \( OPT \) is the optimum value for the given instance, then we call algorithm \( A \) an \( \epsilon \)-approximation algorithm for the problem.

We propose an algorithm that can be used to find an approximate solution for either Problem 1 or Problem 2.

The intuition behind our algorithm is to consider these problems as submodular cost submodular knapsack (SCSK) constraint maximization problems [28], [29]. An SCSK constrained maximization problem is defined as

\[
\text{maximize } f(X) \text{ subject to } g(X) \leq b.
\]

for submodular functions \( f \) and \( g \), and upper bound \( b \in \mathbb{R} \). Problems 1 and 2 can be interpreted as special cases of SCSK with additional cardinality constraints:

\[
\text{maximize } \mu(S_1) \text{ subject to: } S_1 \subseteq Q, \mu(S_1) \leq b, |S_1| \leq k.
\]

Our algorithm is motivated by an approach in [29] for the general SCSK problem. We approximate the optimum \( \mu(S_1) \) for Problem 1 or 2 by imposing an upper bound for the submodular function \( \mu \) and then applying a submodular maximization algorithm to the bounded problem. Specifically, we find an appropriate upper bound constraint \( \mu(S_1) \leq b \), such that a greedy algorithm for maximizing \( \mu(S_1) \) with upper bound \( b \) leads to an approximation algorithm for optimum solution \( S^* \), where

\[
S^* \in \arg \min_{P \subseteq Q, |P| \leq k} |\mu(P) - \alpha|
\]

is an optimal solution for Problem 1 or 2, respectively.

We apply a greedy algorithm to problem (34). For an upper bound \( b \), the algorithm Greedy returns a solution \( S_b \). We can compare different upper bounds by the solutions Greedy returns. The bound \( b_1 \) is a better upper bound than \( b_2 \) if \( \mu(S_{b_1}) \leq b_1 \leq \mu(S_{b_2}) \leq b_2 \). We further define the best upper bound input for algorithm Greedy as \( b^* \), or formally,

\[
b^* \in \arg \min_{b \in [\alpha, 1]} |\mu(S_b) - \alpha|.
\]

We use a modified binary search to converge to the best upper bound \( b^* \) for Greedy. In the next subsection, we describe both the bound search algorithm and the routine Greedy.
A. Bounded Search Approximation Algorithm

We first describe the algorithm in terms of Problem 1. We describe the changes of the algorithm in order to solve Problem 2 in the end of the subsection.

Algorithm 1: $P = \text{BoundSearch}(G, Q, \alpha, c, \delta)$

| Input: | $G = (V, E, w)$: A connected graph. $Q$: A candidate node set with $|Q| = q$. $\kappa$: stubbornness function for nodes $\alpha$: objective. $\delta$: precision parameter: $0 < \delta \leq 1/2$ |
| Output: | $P$: A subset of $Q$ satisfying constraint $c$. |
| $P \leftarrow \emptyset$ |
| $b_{\text{min}} \leftarrow \alpha$; $b_{\text{max}} \leftarrow 1$ |
| $b \leftarrow 1$ // current upper bound |
| $\hat{b} \leftarrow \hat{b}$ // current best upper bound |
| $d_{\text{min}} \leftarrow \alpha$ // minimal $|\mu - \alpha|$ so far |
| $t \leftarrow 0$ |
| $t \leftarrow t + 1$ |
| $(S, \mu) = \text{Greedy}(G, Q, \hat{b}, k)$ |
| $d \leftarrow |\mu - \alpha|$ |
| if $d < d_{\text{min}}$ then |
| $P \leftarrow S$; $d_{\text{min}} \leftarrow d$; $\hat{b} \leftarrow \hat{b}$ |
| if $\mu > \alpha$ then |
| while $\mu \leq (b_{\text{min}} + b_{\text{max}})/2$ do |
| $b_{\text{max}} \leftarrow (b_{\text{min}} + b_{\text{max}})/2$ |
| else |
| $b_{\text{min}} \leftarrow \hat{b}$ |
| if $(\alpha + d) < b_{\text{max}}$ then |
| $b_{\text{max}} \leftarrow (\alpha + d)$ |
| $\hat{b} \leftarrow b_{\text{max}}$; continue |
| $\hat{b} \leftarrow (b_{\text{min}} + b_{\text{max}})/2$ |
| while $b_{\text{max}} > \exp(\delta)/2$ |
| return $P$ |

Our algorithm, BoundSearch, is given in Algorithm 1. The algorithm takes as input a graph $G$, a candidate vertex set $Q$, an objective opinion $\alpha$, a cardinality constraint $k$, and a precision parameter $\delta$ for binary search.

The bound $\hat{b}$ is initialized with value 1, and the algorithm searches $b^*$ in the interval $[b_{\text{min}}, b_{\text{max}}]$ that might include a better upper bound than $\hat{b}$, the current best bound found by the algorithm that leads to the smallest $|\mu(S_b) - \alpha|$. We update $b_{\text{min}}$ and $b_{\text{max}}$ until $b_{\text{min}} \approx \delta b_{\text{max}}$, and $b^*, \hat{b}, b \in [b_{\text{min}}, b_{\text{max}}]$, we obtain $\hat{b} \approx \delta b^*$. Since $\hat{b}$ is the current best upper bound found by the algorithm, it is as good as any upper bound outside of $[b_{\text{min}}, b_{\text{max}}]$, that is, for any $b \notin [b_{\text{min}}, b_{\text{max}}]$, $|\mu(S_b) - \alpha| \leq |S_b - \alpha|$. In Algorithm 2, we present the greedy routine $P = \text{Greedy}(G, Q, \hat{b}, k)$ for the constrained submodular maximization described in (34). The algorithm chooses the node that most increases $\mu(P)$ without violating the upper bound from the candidate set in each iteration, deletes it from the candidate set, and adds it to the current leader set if adding the node does not violate any of the constraint.

Algorithm 2: $(P, \mu) = \text{Greedy}(G, Q, \hat{b}, k)$

| Input: | $G = (V, E, w)$: A connected graph. $Q$: A candidate node set with $|Q| = q$. $\hat{b}$: SCSK upper bound of $\mu$. $k$: an integer $1 \leq k \leq q$. |
| Output: | $S$: A subset of $Q$ satisfying cardinality constraint $k$. $\gamma$: a value $\mu(P)$. |
| $P \leftarrow \emptyset$ |
| while $|Q| > 0$ and $|P| < k$ do |
| $s \leftarrow \arg \max_{u \in Q} \mu(P \cup \{u\})$ |
| if $\mu(P \cup \{s\}) \leq \hat{b}$ then |
| $P \leftarrow (P \cup \{s\})$ |
| $Q \leftarrow (Q \setminus \{s\})$ |
| $\gamma \leftarrow \mu(P)$ |
| return $(P, \gamma)$ |

To analyze Algorithm 2, we introduce the concept of the minimum cover number.

**Definition V.3.** The minimum cover number $k_{\mu,b}$ for set function $\mu(S), S \subseteq Q,$ and $b \in \mathbb{R}$ is defined as

$$k_{\mu,b} = \min\{|S| : \mu(S) \geq b\},$$

if there exists $S$ satisfying $\mu(S) \geq b$, otherwise $k_{\mu,b} = +\infty$.

Then we prove the approximation ratio of BoundSearch.

**Theorem V.4.** Consider a graph $G$, a candidate set $Q$, an objective $\alpha$, a cardinality constraint $|P| \leq k$, and a precision parameter $\delta > 0$. Let $S^*$ be an optimal solution for Problem 1 for these parameters. The algorithm $P = \text{BoundSearch}(G, Q, \alpha, k, \delta)$ returns a node set $P$ such that

$$\mu(P) \approx_{\alpha, \delta} \mu(S^*),$$

in which $\sigma = -\ln(1 - \zeta) + \delta$, and $\zeta = \max(1/e, 1/k_{\mu,\alpha})$.

**Proof.** We let $\hat{b}$ be the best bound found by the algorithm with smallest $|\mu(S_b) - \alpha|$. And Greedy with the best upper bound $b^*$ returns the result $\mu(S_{b^*})$, we recall that $b^*$ is defined by (35).

If $\mu(P \cup \{s\}) \leq \alpha$ is always satisfied during the execution, then $\mu(P \cup \{s\}) \leq \hat{b}$ is also always satisfied. Then the returned $S_b$ is the same as what we get from a greedy algorithm without the constraint $\mu(T) \leq \hat{b}$. We further define

$$\tilde{S} \in \arg \max_{T \subseteq Q, |T| \leq k} \mu(T),$$

therefore

$$\mu(\tilde{S}) \geq \mu(S_b) \geq \left(1 - \frac{1}{e} \right) \mu(\tilde{S}).$$

If $\alpha \geq \mu(\tilde{S})$, then $\mu(\tilde{S}) = \mu(S^*)$, we attain the guarantee $\mu(S_b) \approx_{\alpha, \delta} \mu(S^*)$, where $e^{-\gamma} = (1 - 1/e)$. If $\mu(S_b) \leq \alpha \leq |S^*|$, then...
\( \mu(\overline{S}) \), then \( \mu(S^*) \in [\mu(S_b), \mu(\overline{S})] \), which implies \( \mu(S_b) \approx_\gamma \mu(S^*) \), where \( e^{-\gamma} = (1 - 1/e) \).

If \( \mu(P \cup \{ s \}) \leq \alpha \) is first violated when we add the \((t + 1)\)th node, we define \( P_t \) as the set of changed nodes of size \( t \) in Greedy, therefore \( |P_t| = t \). We further define \( \rho(s_{t+1}) = \mu(P_t \cup \{ s_{t+1} \}) - \mu(P_t) \). From the submodularity of \( \mu(S) \) we know \( \rho(s_{t+1}) \leq \frac{1}{t+1} \mu(P_t \cup \{ s_{t+1} \}) \) holds for the greedy algorithm. Then \( \mu(P_t) = \mu(P_t \cup \{ s_{t+1} \}) - \rho(s_{t+1}) \geq \frac{t}{t+1} \mu(P_t \cup \{ s_{t+1} \}) \geq \frac{t}{t+1} \mu(P_t \cup \{ s_{t+1} \}) \). By letting \( \overline{b} = \mu(P_t \cup \{ s_{t+1} \}) \) (then by definition \( \overline{b} = \mu(S_b) = \mu(P_t \cup \{ s_{t+1} \}) \)), we attain \( \mu(S_b) \geq (1 - \frac{1}{t+1}) \mu(S_b) \). We further attain \( \mu(S^*), \mu(S_{b^*}) \in [\mu(S_b), \mu(S_b)] \), and \( b^* \in [\alpha, \overline{b}] \). Since \( \overline{b}, b^* \in [b_{min}, b_{max}] \) and \( b_{min} \approx_\delta b_{max}, b^* \approx_\delta b^* \) we attain \( b \in [\alpha, e^{\delta} \overline{b}] \) and therefore \( \mu(S_b) \in [\mu(S_b), e^{\delta} \mu(S_b)] \), so \( \mu(S_b) \approx_\gamma \mu(S^*) \), where \( e^{-\gamma} = (1 - 1/k_{\mu,\alpha}) e^{-\delta} \).

The guarantee given in Theorem V.4 can also be written as:

\[
(1 - \zeta)e^{-\delta} \mu(S^*) \leq \mu(P) \leq (1 - \zeta)^{-1} e^{\delta} \mu(S^*) .
\]

The BoundSearch algorithm can be applied to Problem 2 with the same approximation guarantee with the only difference that the stubbornness function \( \kappa \) is an input of the algorithm. The stubbornness function is also passed into Greedy to calculate the average opinion. Theorem V.4 holds for the corresponding algorithm \( P = \text{BoundSearch}([G, Q, \alpha, k, \kappa, \delta]) \), which calls Greedy\([G, Q, \overline{b}, k, \kappa]\).

**B. Complexity Analysis**

A naive implementation of the proposed algorithm runs in \( O(kn^{3} \log \frac{1}{\delta}) \) time, which is expensive for large graphs. Using blockwise inversion and rank-1 update of matrices we can improve the running time of BoundSearch to \( O(n^{3} \log \frac{1}{\delta}) \).

**Theorem V.5.** There exists an implementation of Algorithm 1 for a graph with \( n \) nodes that has running time \( O(n^{3} \log \frac{1}{\delta}) \).

**Proof.** We take the completely stubborn leader case as an example. In each execution of Line 3 of Algorithm 2, we need to calculate the sum of follower states, which is given by

\[
1^T (L_{F,F})^{-1} L_{F,S} x_S,
\]

for all \( P \cup \{ u \}, u \in Q \). \( P \) and \( Q \) are the current leader set of opinion 1 and the current candidate set. Calculating \( (L_{F,F})^{-1} \) when \( S_1 = \emptyset \) takes \( O(n^3) \) running time. \( L_{F,F} \) can be updated at iteration \( t + 1 \) by deleting the row and column associated with candidate node \( u \). From block matrix inversion, we obtain that its inverse can be updated by

\[
\left( (L_{F,F})^{-1} - e_u e_u^T \right) \frac{1}{(L_{F,F})^{-1}} - e_u^T \frac{1}{(L_{F,F})^{-1}} e_u .
\]

To calculate \( \mu(P_t) \), we do not need to find \( \left( L_{F,F} \right)^{-1} \) explicitly. It suffices to compare the value of

\[
\left( L_{(F,F)(\{u\}), (F,F)(\{u\})} \right)^{-1} L_{(F,F)(\{u\}), (S,S)(\{u\})} x_{S \cup \{u\}},
\]

for all \( u \) in the current candidate set. We note that \( e_u \{ e_u^T \} \) takes a column (row) of \( (L_{F,F})^{-1} \), and \( L_{(F,F)(\{u\}), (S,S)(\{u\})} x_{S \cup \{u\}} \) is a column vector. By the associative law, we compute the vector inner product first and find the updated \( \mu(S_1) \) for at most \( n \) candidates in \( O(n^2) \) time. The operations of taking the substractions do not change the complexity because for any candidate \( u \), these operations only take \( O(\{N_u\}) \) running time. So, in each execution of Line 3 of Algorithm 2, these operations can be done in \( O(m) \) total running time, where \( m \) is the number of edges in the graph. After we find the best choice \( s_{t+1} \) in step \( t + 1 \), we update \( (L_{F,F})^{-1} \) explicitly, which takes additional \( O(n^2) \) time. Therefore, execution of Line 3 of Algorithm 2 takes \( O(n^2) \) time. By using this simple acceleration, the complexity of Algorithm 2 is improved to \( O(n^3 + kn^2) = O(n^3) \).

Algorithm 1 calls Greedy \( O(\log \frac{1}{\delta}) \) times until \( b_{max} \approx_\delta b_{min} \). Since \( b_{max} - b_{min} \) decreases geometrically in Algorithm 1, the total running time of BoundSearch is \( O(n^3 \log \frac{1}{\delta}) \).

For the partially stubborn leader case, the the rank-1 update is obtained using the Sherman-Morrison formula. And, the running time of the Greedy routine is also \( O(n^3) \) by a similar implementation. We omit the details of the analysis.

**VI. EXPERIMENTS**

In this section, we present experiments to highlight the analytical results and to show the effectiveness of the proposed algorithm. In all experiments, the networks are undirected, and all edge weights are set to 1.

| \( \alpha \) | Optimum | DS | ER | Random |
|----------|--------|----|----|--------|
| 0.25     | 0.317  | 0.317 | 0.317 | 0.485 |
| 0.50     | 0.500  | 0.499 | 0.317 | 0.469 |
| 0.75     | 0.570  | 0.570 | 0.317 | 0.481 |
| 1.00     | 0.570  | 0.570 | 0.317 | 0.484 |

**TABLE I:** Average opinion in a completely stubborn leader system. The graph is an Erdős-Rényi graph with 200 nodes and connecting probability 0.1 with a fixed \( s_0 \) and a node \( s_1 \) chosen via various methods.

| \( \alpha \) | Optimum | DS&K | ER | Random |
|----------|--------|------|----|--------|
| 0.25     | 0.288  | 0.288 | 0.288 | 0.484 |
| 0.50     | 0.500  | 0.499 | 0.288 | 0.473 |
| 0.75     | 0.524  | 0.524 | 0.288 | 0.473 |
| 1.00     | 0.524  | 0.524 | 0.288 | 0.476 |

**TABLE II:** Average opinion in a partially stubborn leader system. The graph is an Erdős-Rényi graph with 200 nodes and connecting probability 0.01 with a fixed \( s_0 \) and a node \( s_1 \) chosen uniformly at random in experiment.

We first study the properties of \( \mu(S_1) \) when \( |S_0| = |S_1| = 1 \) with completely stubborn leaders and partially stubborn leaders for \( \alpha = 0.25, 0.5, 0.75 \), and 1. The leader \( s_0 \) is chosen uniformly at random in experiment.

For the stubborn leader system, we run an experiment on an Erdős-Rényi graph with 200 nodes with connecting probability 0.1. We find the average opinion of the network for the optimal solution to Problem 1, i.e., the optimal \( s_1 \) as given by Theorem IV.3. We also show the average opinion when \( s_1 \) is
using heuristics motivated by the theorem. The first heuristic, DS, is based on the domination score; we find the $s_1$ such that the resulting $\mu(\{s_1\})$ minimizes the numerator of (27). We also use a heuristic based on effective resistance (ER); here, $s_1$ is chosen so as to maximize the denominator of (27). Finally, we compute the average opinion for a randomly chosen $s_1$. The results of this experiment are shown in Table I.

We also conduct an experiment for partially stubborn system. We use a different Erdős-Rényi graph with 200 nodes and connecting probability 0.01. Partially stubborn leaders have uniform stubbornness $\kappa = 1$, and the other parameters are the same as the experiment for completely stubborn system. We find the optimal $s_1$ as well as the $s_1$ chosen by heuristics motivated by the numerator (DS&K) and denominator (ER) of the result given in Theorem IV.4. We note that the $s_1$ that minimizes denominator of the result in Theorem IV.3 also minimizes denominator of the result in IV.4. The results are shown in Table II.

Table I and II show that when $|S_0| = |S_1| = 1$, the domination score well captures the behavior of $\mu(\{s_1\})$. We have observed similar results in various Erdős-Rényi graph with different choices of a single leader $s_0$. This does not mean that the domination score alone determines the best choice of $s_1$. Sometimes the denominator of (27) affects the result. For example, in the graph given in Figure 3, we let $s_0$ be the blue node, $\alpha = 0.5$, and $s_1$ can be chosen from $\{a, b, c, d\}$. Node $a$ (or node $c$) minimizes the numerator of equation (27), which is 0.0875, while choosing $b$ leads to 0.1, which is larger. However, $\mu(\{a\}) = 0.64$ and $\mu(\{b\}) = 0.6$, therefore choosing $b$ makes the average opinion closer to $\alpha$.

![Figure 3: An example where Domination Score does not return a good result.](image)

Next, to show the effectiveness of our leader selection algorithm, we compare the result returned by our algorithm BoundSearch with the optimal value returned by brute-force search. We use an Erdős-Rényi graph with 30 nodes and connecting probability 0.1. We choose an $S_0$ leader set of size 3 at random. We run the BoundSearch algorithm for both completely stubborn and partially stubborn systems. With $\alpha \in \{0.25, 0.5, 0.75\}$. Partially stubborn leaders use uniform stubbornness $\kappa = 1$. The results are shown in Figure 4. In all cases, BoundSearch returns nearly optimal results.

![Figure 4: Average opinion of Optimum vs. average opinion of BoundSearch in an Erdős-Rényi graph with 30 nodes and connecting probability 0.1. $S_0$ leader sets are chosen randomly and $S_1$ leader sets are chosen by brute-force search and BoundSearch with different $k$ and $\alpha$ values.](image)

Finally, we explore the effect of varying the $\delta$ parameter in BoundSearch. We run BoundSearch on the Haggle graph [30] social contact graph. The Haggle graph is a multigraph, which we turn it into a simple graph by deleting all duplicate edges. We use the largest connected component of the graph which has 274 nodes and 2124 edges. We set $k = 15$ and $\alpha \in \{0.2, 0.35, 0.5\}$. We vary $\delta$ from 0.0001 to 0.25. For completely stubborn leaders, we have $|S_0| = 80$ and for partially stubborn leaders we have $|S_0| = 15$. Partial leaders use uniform stubbornness $\kappa = 1$. The results are shown in Figure 5. We observe that as $\delta$ decreases, the results from BoundSearch converge to a value close to $\alpha$.

![Figure 5: Effect of varying $\delta$ on BoundSearch at $\alpha \in \{0.20, 0.35, 0.50\}$. Experiment run using the Haggle graph.](image)

**VII. Conclusion**

We have studied two French-DeGroot opinion dynamics models where leaders have polarizing opinions. For both models, we showed analytic expressions for the steady-state opinion using the Laplacian matrix of an augmented graph. For the single leader case, we gave an explicit expression for steady-state opinion vector and analyzed the average opinion based on the expression. Then, we studied the problem of shifting the average steady-state opinion to a given value by selecting an opposing leader set with cardinality constraint. We gave both hardness results for this problem and algorithms with provable approximation ratios. We also presented experiments showing that our algorithm returns results close to optimal results in practice. Future work will focus on algorithms with better approximation ratio and running time, and the hardness of leader selection in partially stubborn leader systems.

**References**

[1] D. Kempe, J. M. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” in *Proc. of 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003, pp. 137–146.
We present simple proofs for the submodularity based on the escape probability interpretation of \( \hat{x} \).

### A. Steady-State Opinion Interpreted as Escape Probability

The entries of \( \hat{x}_F \) can be interpreted as the escape probability of a random walker [31] in a Markov chain with absorbing states define on graph \( G \). Consider an absorbing Markov chain \( P \) with \( S_0 \cup S_1 \) the set of absorbing states and \( F \) the set of non-absorbing states. Then the transition matrix has the form

\[
P = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix}.
\]  

where \( R = (D_{F,F})^{-1}A_{F,S} \) and \( Q = (D_{F,F})^{-1}A_{F,F} \).

Define a harmonic function \( y \) with boundary \( y_B = \hat{x}_S \). The interior \( y_D \) is determined by [31]

\[
y_D = (I - Q)^{-1}Ry_B.
\]

Then we obtain

\[
\hat{x}_F = y_D = -(L_{F,F})^{-1}L_{F,S}x_S.
\]

Combining with the boundary condition \( y_B = \hat{x}_S \), we obtain \( y = \hat{x} \). \( y \) defines the concept of escape probability explained below.

Let \( S_0 \) and \( S_1 \) be two sets of absorbing states in a Markov chain (36). We let \( \tau_0^G(S_1, -S_0) \) represent the event that in a Markov chain defined by graph \( G \), a random walker starts from node \( v \), hits any state \( u \in S_1 \) before it reaches any state \( u \in S_0 \). Then \( \hat{x}_v \) is the probability that \( \tau_0^G(S_1, -S_0) \) happens. We denote the escape probability as \( p_0^G(S_1, -S_0) \) defined as \( \Pr(\tau_0^G(S_1, -S_0)) \). This escape probability is given by the harmonic function \( y \) defined above (for example, see [31]). We have shown that \( y = \hat{x} \), so \( \hat{x}_v = p_0^G(S_1, -S_0) \). Similarly, we define \( \tau_1^G(S_0, -S_1) \) as the event that in the Markov chain defined by graph \( G \), a random walker starts from node \( v \), hits any state \( u \in S_0 \) before it reaches any state \( u \in S_1 \), and we also denote by \( p_1^G(S_0, -S_1) \) the probability that event \( \tau_1^G(S_0, -S_1) \) happens. Since a random walker is either absorbed by \( u \in S_0 \) or \( u \in S_1 \), \( p_0^G(S_1, -S_0) + p_1^G(S_0, -S_1) = 1 \).

### B. Completely Stubborn Leaders

In the considered leader-follower system with completely stubborn leaders, given fixed \( S_0 \), \( \mu(S_1) \) is defined as

\[
\mu(S_1) \overset{\text{def}}{=} \frac{1}{n} \sum_{v \in V} p_v^G(S_1, -S_0).
\]
To prove that \( \mu(S_1) \) is monotone and submodular, it suffices to show that \( p^G_v(S_1, -S_0) \) is monotone and submodular for all \( v \in \mathcal{V} \).

**Lemma A.1.** For any \( S_1 \subseteq T_1 \subseteq \mathcal{V}, S_0 \subseteq \mathcal{V}, \) and \( T_1 \cap S_0 = \emptyset \), for any \( v \in \mathcal{V} \)

\[
p^G_v(T_1, -S_0) \geq p^G_v(S_1, -S_0)
\]

**Proof.** We first consider \( S_1^{(0)} = S_1 \) and \( S_1^{(1)} = S_1 \cup \{u\} \), where \( u \in \{T_1\} \setminus S_1 \). For a random walker in graph \( G \) starting from node \( v \), we observe that

\[
p^G_v(S_1^{(1)}, -S_0) = p^G_v((S_1 \cup \{u\}), -S_0)
= p^G_v(S_1, -(S_0 \cup \{u\})) + p^G_v(\{u\}, -(S_0 \cup S_1))
\]

and

\[
p^G_v(S_1^{(0)}, -S_0) = p^G_v((S_1, -S_0)
= p^G_v(S_1, -(S_0 \cup \{u\})) + p^G_v(\{u\}, -(S_0 \cup S_1)) \cdot p^G_v(S_1, -S_0),
\]

by the Markov property. Therefore

\[
p^G_v(S_1^{(1)}, -S_0) - p^G_v(S_1^{(0)}, -S_0) = p^G_v(\{u\}, -(S_0 \cup S_1)) \cdot (1 - p^G_v(S_1, -S_0))
= p^G_v(\{u\}, -(S_0 \cup S_1)) \cdot p^G_v(S_1, -S_0) \geq 0.
\]

Similarly, by defining a sequence of \( S_1^{(i)}, i = 1, \ldots, t \) such that \( t = |T_1 \setminus S_1| \) and \( S_1^{(i)} = T_1 \), we attain the following corollary

**Corollary A.2.** For any \( S_0 \subseteq \mathcal{V}, S_1 \subseteq T_1 \subseteq \mathcal{V}, \) and \( T_1 \cap S_0 = \emptyset \),

\[
p^G_v(S_0, -T_1) \leq p^G_v(S_0, -S_1).
\]

**Lemma A.3.** For any \( S_1 \subseteq T_1 \subseteq \mathcal{V}, S_0 \subseteq \mathcal{V}, \) \( T_1 \cap S_0 = \emptyset \), and \( u \in \mathcal{V} \setminus (T_1 \cup S_0) \),

\[
p^G_v(T_1 \cup \{u\}, -S_0) - p^G_v(T_1, -S_0) \leq p^G_v(S_1 \cup \{u\}, -S_0) - p^G_v(S_1, -S_0).
\]

**Proof.** For any \( v \in \mathcal{V}, \)

\[
p^G_v(T_1 \cup \{u\}, -S_0) - p^G_v(T_1, -S_0)
= p^G_v(\{u\}, -(T_1 \cup S_0)) \cdot p^G_v(S_0, -T_1)
\]

and

\[
p^G_v(S_1 \cup \{u\}, -S_0) - p^G_v(S_1, -S_0)
= p^G_v(\{u\}, -(S_1 \cup S_0)) \cdot p^G_v(S_0, -S_1)
\]

Using corollary A.2 we get the inequality in the lemma by comparing (44) and (45).

**C. Partially Stubborn Leaders**

In the considered leader-follower system with partially stubborn leaders, given fixed \( S_0, \) \( \mu(S_1) \) is defined as

\[
\mu(S_1) = \frac{1}{n} \sum_{v \in \mathcal{V}} p^G_v(\{s'_1\}, -\{s'_0\}),
\]

in which \( p^G_v(\{s'_1\}, -\{s'_0\}) \) represents the probability that a random walker in augmented graph \( G' \) starting from \( v \) reaches \( s'_1 \) before it reaches \( s'_0 \). To prove that \( \mu(S_1) \) is monotone and submodular, it suffices to show that \( p^G_v(\{s'_1\}, \{s'_0\}) \) is monotone and submodular for all \( v \in \mathcal{V} \).

**Lemma A.4.** For any \( S_1 \subseteq T_1 \subseteq \mathcal{V}, S_0 \subseteq \mathcal{V}, \) and \( T_1 \cap S_0 = \emptyset \), and the augmented graph \( \mathcal{H}' \) defined by \( G, S_0, \) and \( S_1 \), then \( \mathcal{H}' \) has the same node set as \( \mathcal{G}' \), the edge set of \( \mathcal{H}' \) consists of all edges in the edge set of \( \mathcal{G}' \), and all \( (u, s'_1), u \in (T_1 \setminus S_1) \). For any \( v \in \mathcal{V} \)

\[
p^H_v(\{s'_1\}, -\{s'_0\}) \geq p^G_v(\{s'_1\}, -\{s'_0\}).
\]

**Proof.** Let \( \mathcal{G} + (u, v) \) be the graph attained by adding an edge \( (u, v) \) to the graph \( \mathcal{G} \). We start by considering \( G(0) = \mathcal{G} \) and \( G(1) = \mathcal{G} + (u, s'_1), u \in (T_1 \setminus S_1) \). Let \( \xi_v(u, s'_1) \) be the event that a random walker in \( \mathcal{G}' \) starting from node \( v \) passes through edge \( (u, s'_1) \) before it reaches any absorbing state, and \( \xi'_v(u, s'_1) \) be the event that a random walker does not pass through \( (u, s'_1) \) before reaching an absorbing state.

\[
p^G_v(\{s'_1\}, -\{s'_0\}) = \Pr \left( \xi_v(u, s'_1) \mid \xi_v(u, s'_1) \right) \Pr \left( \xi'_v(u, s'_1) \right)
+ \Pr \left( \xi'_v(u, s'_1) \right) \Pr \left( \xi'_v(u, s'_1) \right)
\]

\[
= \Pr \left( \xi_v(u, s'_1) \right) \cdot \left(1 - \Pr \left( \xi'_v(u, s'_1) \right) \right).
\]

We note that

\[
\Pr \left( \xi_v(u, s'_1) \right) = \Pr \left( \xi_v(u, s'_1) \right),
\]

therefore

\[
p^G_v(\{s'_1\}, -\{s'_0\}) - p^G_v(\{s'_1\}, -\{s'_0\})
= \Pr \left( \xi_v(u, s'_1) \right) \cdot \left(1 - p^G_v(\{s'_1\}, -\{s'_0\}) \right)
= \Pr \left( \xi_v(u, s'_1) \right) \cdot p^G_v(\{s'_0\}, -\{s'_1\}) \geq 0.
\]

Similarly, by defining a sequence of \( G(i), i = 1, \ldots, t \) such that \( t = |T_1 \setminus S_1| \), we attain \( G(t) = \mathcal{H}' \) and

\[
p^G_v(\{s'_1\}, -\{s'_0\}) \geq p^G_v(\{s'_1\}, -\{s'_0\}) \]

holds for all \( i \in [t] \). This leads to the result in lemma A.4.
Corollary A.5. For any $S_0 \subseteq \mathcal{V}$, $S_1 \subseteq T_1 \subseteq \mathcal{V}$, and $T_1 \cap S_0 = \emptyset$, $\mathcal{G}'$ and $\mathcal{H}'$ have the same definitions as they are defined in Lemma A.4. then
\[
p_w^{\mathcal{H}'}(\{s'_0\}, -\{s'_1\}) \leq p_w^{\mathcal{G}'}(\{s'_0\}, -\{s'_1\}) .
\]

Lemma A.6. For any $S_1 \subseteq T_1 \subseteq \mathcal{V}$, $S_0 \subseteq \mathcal{V}$, $T_1 \cap S_0 = \emptyset$, and $u \notin (S_0 \cup T_1)$, we consider the augmented graph $\mathcal{G}'$ defined by $\mathcal{G}$, $S_0$, and $S_1$, and the augmented graph $\mathcal{H}'$ defined by $\mathcal{G}$, $S_0$, and $T_1$. Then $\mathcal{H}'$ has the same node set as $\mathcal{G}'$, the edge set of $\mathcal{H}'$ consists of all edges in the edge set of $\mathcal{G}'$, and all $(l, s_1), l \in (T_1 \setminus S_1)$. For any $v \in \mathcal{V}$ and $u \notin (S_0 \cup T_1)$,
\[
p_w^{\mathcal{H}'+(u, s_1)}(\{s'_1\}, -\{s'_0\}) - p_w^{\mathcal{H}'}(\{s'_1\}, -\{s'_0\})
\leq p_w^{\mathcal{G}'+(u, s_1)}(\{s'_1\}, -\{s'_0\}) - p_w^{\mathcal{G}'}(\{s'_1\}, -\{s'_0\}) .
\]

Proof. Following similar analysis as the proof of Lemma A.4, we obtain
\[
p_w^{\mathcal{H}'+(u, s_1)}(\{s'_1\}, -\{s'_0\}) - p_w^{\mathcal{H}'}(\{s'_1\}, -\{s'_0\})
= \Pr(\xi_{\mathcal{H}'+(u, s_1)}(u, s'_1)) \cdot p_w^{\mathcal{H}'}(\{s'_0\}, -\{s'_1\})
\]
\[
p_w^{\mathcal{G}'+(u, s_1)}(\{s'_1\}, -\{s'_0\}) - p_w^{\mathcal{G}'}(\{s'_1\}, -\{s'_0\})
= \Pr(\xi_{\mathcal{G}'+(u, s_1)}(u, s'_1)) \cdot p_w^{\mathcal{G}'}(\{s'_0\}, -\{s'_1\})
\]

Then we extend the definition of $\xi_v^{\mathcal{G}'}(u, s'_1)$ and denote $\xi_v^{\mathcal{G}'}(U, s'_1)$ as the event that a random walker in $\mathcal{G}'$ starting from $v$ passes through any edge $(u, s'_1)$, $u \in U$ before it reaches any absorbing state. Similarly we define $\xi_v^{\mathcal{H}'}(U, s'_1)$ as the event that the random walker reaches an absorbing state without passing through any $(u, s'_1)$, $u \in U$.
\[
\Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}(u, s'_1))
= \Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}(u, s'_1) | \xi_v^{\mathcal{H}'+(u, s'_1)}((T_1 \setminus S_1), s'_1)) \cdot \Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}((T_1 \setminus S_1), s'_1))
+ \Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}(u, s'_1) | \xi_v^{\mathcal{H}'+(u, s'_1)}((T_1 \setminus S_1), s'_1)) \cdot \Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}((T_1 \setminus S_1), s'_1))
\]
In addition,
\[
\Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}(u, s'_1) | \xi_v^{\mathcal{H}'+(u, s'_1)}((T_1 \setminus S_1), s'_1)) = 0
\]
and
\[
\Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}(u, s'_1) | \xi_v^{\mathcal{H}'+(u, s'_1)}((T_1 \setminus S_1), s'_1)) = \Pr(\xi_v^{\mathcal{G}'+(u, s'_1)}(u, s'_1)) ,
\]
the we attain
\[
\Pr(\xi_v^{\mathcal{H}'+(u, s'_1)}(u, s'_1)) \leq \Pr(\xi_v^{\mathcal{G}'+(u, s'_1)}(u, s'_1)) .
\] Applying Corollary A.5 and (53) to (49) and (50) leads to the result stated in Lemma A.6. □