



Opinion Dynamics Optimization by Varying Susceptibility to Persuasion via Non-Convex Local Search

REDIET ABEBE, UC Berkeley

T.-H. HUBERT CHAN, The University of Hong Kong

JON KLEINBERG, Cornell University

ZHIBIN LIANG, The University of Hong Kong

DAVID PARKES, Harvard University

MAURO SOZIO, LTCI, Télécom ParisTech University

CHARALAMPOS E. TSOURAKAKIS, Boston University

A long line of work in social psychology has studied variations in people's susceptibility to persuasion—the extent to which they are willing to modify their opinions on a topic. This body of literature suggests an interesting perspective on theoretical models of opinion formation by interacting parties in a network: in addition to considering interventions that directly modify people's intrinsic opinions, it is also natural to consider interventions that modify people's susceptibility to persuasion.

In this work, motivated by this fact, we propose an influence optimization problem. Specifically, we adopt a popular model for social opinion dynamics, where each agent has some fixed innate opinion, and a resistance that measures the importance it places on its innate opinion; agents influence one another's opinions through an iterative process. Under certain conditions, this iterative process converges to some equilibrium opinion vector. For the unbudgeted variant of the problem, the goal is to modify the resistance of any number of agents (within some given range) such that the sum of the equilibrium opinions is minimized; for the budgeted variant, in addition the algorithm is given upfront a restriction on the number of agents whose resistance may be modified.

We prove that the objective function is in general non-convex. Hence, formulating the problem as a convex program as in an early version of this work (Abebe et al., KDD'18) might have potential correctness issues. We instead analyze the structure of the objective function, and show that any local optimum is also a global optimum, which is somehow surprising as the objective function might not be convex. Furthermore, we combine the iterative process and the local search paradigm to design very efficient algorithms that can solve the unbudgeted variant of the problem optimally on large-scale graphs containing millions of nodes. Finally, we propose and evaluate experimentally a family of heuristics for the budgeted variant of the problem.

This research was partially supported by a grant from the PROCORE France–Hong Kong Joint Research Scheme sponsored by the Research Grants Council of Hong Kong and the Consulate General of France in Hong Kong under the project F-HKU702/16. T.-H. Hubert Chan was partially supported by the Hong Kong RGC under the grants 17200817 and 17201220. Jon Kleinberg acknowledges support from a Simons Investigator grant, a Vannevar Bush Faculty Fellowship, and the AFOSR grant FA9550-19-1-0183. Charalampos E. Tsourakakis acknowledges support from Intesa Sanpaolo Innovation Center; the funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' addresses: R. Abebe, UC Berkeley, Berkeley, CA 94720; email: rabebe@fas.harvard.edu; T.-H. Hubert Chan and Z. Liang, The University of Hong Kong, Pokfulam 999077, Hong Kong; email: hubert@cs.hku.hk and liangzb@connect.hku.hk; J. Kleinberg, Cornell University, Ithaca, NY 14853; email: kleinberg@cornell.edu; D. Parkes, Harvard University, Cambridge, MA 02138; M. Sozio, LTCI, Télécom ParisTech University, 13 Avenue de la Vauve, Palaiseau 91120, France; email: parkes@eecs.harvard.edu; C. E. Tsourakakis, Boston University, Boston, MA 02215; email: ctsourak@bu.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

1556-4681/2021/07-ART33 \$15.00

<https://doi.org/10.1145/3466617>

CCS Concepts: • **Human-centered computing** → **Social network analysis**; • **Mathematics of computing** → **Graph algorithms**;

Additional Key Words and Phrases: Opinion dynamics, susceptibility to persuasion, non-convex local search

ACM Reference format:

Rediet Abebe, T.-H. Hubert Chan, Jon Kleinberg, Zhibin Liang, David Parkes, Mauro Sozio, and Charalampos E. Tsourakakis. 2021. Opinion Dynamics Optimization by Varying Susceptibility to Persuasion via Non-Convex Local Search. *ACM Trans. Knowl. Discov. Data.* 16, 2, Article 33 (July 2021), 34 pages.

<https://doi.org/10.1145/3466617>

1 INTRODUCTION

A rich line of empirical work in development and social psychology has studied people’s susceptibility to persuasion. This property measures the extent to which individuals are willing to modify their opinions in reaction to the opinions expressed by those around them, and it is distinct from the opinions they express. Research in the area has ranged from adolescent susceptibility to peer pressure related to risky and antisocial behavior [4, 21, 23, 46, 51] to the role of susceptibility to persuasion in politics [25, 42, 44]. Individuals’ susceptibility to persuasion can be affected by specific strategies and framings aimed at increasing susceptibility [12, 13, 15, 37, 41, 49, 50]. For instance, if it is known that an individual is receptive to *persuasion by authority*, one can adopt a strategy that utilizes arguments from official sources and authority figures to increase that individual’s susceptibility to persuasion with respect to a particular topic.

Modifying network opinions has far-reaching implications including product marketing, public health campaigns, the success of political candidates, and public opinions on issues of global interest. In recent years, there has also been work in Human–Computer Interaction focusing on *persuasive technologies*, which are designed with the goal of changing a person’s attitude or behavior [24, 33, 37]. This work has shown that not only do people differ in their susceptibility to persuasion, but that persuasive technologies can also be adapted to each individual to change their susceptibility to persuasion. Despite the long line of empirical work emphasizing the importance of individuals’ susceptibility to persuasion, to our knowledge, theoretical studies of opinion formation models have not focused on interventions at the level of susceptibility. Social influence studies have considered interventions that directly act on the opinions themselves, where the opinions can take discrete (e.g., [1, 7, 22, 32, 36, 38]) or real [30, 43] values.

In this work, we adopt an opinion formation model introduced by the work of DeGroot [20] and Friedkin and Johnsen [26], and we initiate a study of the impact of interventions at the level of susceptibility. In this model, each agent i is endowed with an *innate opinion* s_i in $[0, 1]$, where 0 and 1 are polar opposites of opinions regarding a certain topic. Each agent also has a parameter representing susceptibility to persuasion, which we will call the *resistance parameter* $\alpha_i \in (0, 1]$. The innate opinion s_i reflects the intrinsic position of agent i on a certain topic. A higher value on the resistance parameter α_i means that the agent is less willing to conform with the opinions of neighbors in the social network. According to the opinion dynamics model, the final opinion of each agent i is a function of the social network,¹ the set of innate opinions, and the resistance parameters, determined by computing the *equilibrium* state of a dynamic process of opinion updating. We study the following natural question:

¹We shall see that the social network is characterized by how the agents interact with one another, which is mathematically captured by a row stochastic matrix that need not be symmetric.

Problem 1. Given an opinion dynamics model, and a set of agents, each of whom has an innate opinion that reflects the agent's intrinsic position on a topic, and a range for the resistance parameter measuring the agent's propensity for changing their opinion, how should we set the agents' resistance parameter in order to minimize the total sum of opinions at equilibrium?

Unbudgeted vs Budgeted Variants. In the unbudgeted variant, we are allowed to modify the resistance of any number of agents. For the budgeted variant, we are given some initial resistance vector and a budget k , the resistance of at most k agents can be changed.

We emphasize that the algorithm has global information on how the agents interact with one another and is allowed to change only the resistance parameters of the agents and do not directly change their innate opinions. In the minimization problem, the goal of the algorithm is to bias the average equilibrium opinions of the agents toward 0.

After reading the formal problem definition in Section 3, the reader can check that the problem is easy if the resistance of each agent can be picked from the closed interval $[0, 1]$, at least for the case when the underlying network is strongly connected. For minimizing the equilibrium opinions, it suffices to make the agent with the minimum innate opinion the most resistant (setting its resistance to 1) and everyone else totally compliant (setting its resistance to 0). Similarly, for the maximization problem, it suffices to make the agent with the maximum innate opinion the most resistant, and the rest of the nodes totally compliant. The problem is non-trivial if the resistance α_i of each agent i can take value from some interval $[l_i, u_i]$, where $0 < l_i < u_i < 1$. We discuss the model and Problem 1 in greater detail in Section 3.

Our Contributions. In this work, we make the following key contributions:

- *Opinion Dynamics with Varying Susceptibility to Persuasion.* Based on previous models on opinion dynamics, we introduce an influence optimization problem that focuses on interventions at the level of susceptibility.
- *Analysis of the Unbudgeted Problem Structure.* We prove that the objective function is in general neither convex nor concave. We analyze the mathematical structure of the problem in Section 4. Perhaps the most important technical insight in this article is that we show (in Lemma 4.9) that if the current vector solution is not optimal, then there exists a coordinate that can be flipped such that the objective will be strictly improved. This shows that an optimal vector can be found by a simple local search algorithm.
- *Local Search with Irrevocable Updates.* In general, local search could still take exponential time to find an optimal solution, for instance, the simplex algorithm for linear programming. Indeed, we show (in Lemma 5.1) that the local search algorithm will change the resistance of each agent at most once, which implies that an optimal solution can be found in polynomial time.
- *Efficient Local Search on Large-Scale Graphs.* Typically, in local search, the objective function needs to be evaluated at the current solution in each step. However, since the objective function involves matrix inverse, its evaluation will be too expensive when the dimension of the matrix is in the order of millions. Instead, we use the iterative process of the opinion dynamics model itself to approximate the equilibrium vector. We have developed several update strategies for local search. For *conservative* or *opportunistic* updates, one always makes sure that the error of the estimated equilibrium vector is small enough before any coordinate of the resistance vector is flipped. For *optimistic* update, one might flip a coordinate of the resistance vector even before the estimated equilibrium vector is accurate enough. However,

this might introduce mistakes which need to be corrected later. Nevertheless, experiments show that mistakes are rarely made by the optimistic update strategy. In any case, for all three update strategies, an optimal vector will be returned when the local search terminates. Our approaches are scalable and can run on networks with millions of nodes. We report the experimental results in Section 8. In particular, using multiple number of threads, the optimistic update strategy can solve the problem optimally on networks with up to around 65 million nodes.

- *Scalable Heuristics for the Budgeted Problem.* We provide a family of efficient heuristics for the budgeted version of our problem, and a detailed experimental evaluation on large-scale real-world networks.

Comparison with Previous Versions. A preliminary version [2] of this work presented the problem, but it was overlooked that the objective function might not be convex or concave. A subsequent work [10] rectified this issue, and showed that local search can be performed efficiently to reach the optimal solution, even if the objective function is non-convex. The current presentation combines results from the aforementioned two works [2, 10]. We have also included a more detailed Section 7 on heuristic algorithms for the budgeted version of the problem, and the related experiments in Section 9.

Other Budgeted Variants. The aforementioned budgeted variant can be thought of as an L_0 -budget because given an initial $\hat{\alpha} \in [0, 1]^N$ and a budget k , the algorithm can only choose α such that $\|\alpha - \hat{\alpha}\|_0 = |\{i \in N : \alpha_i \neq \hat{\alpha}_i\}| \leq k$. Naturally, one can consider other norms to bound the budget. For instance, one can consider an L_p -budget, for which the constraint becomes $\|\alpha - \hat{\alpha}\|_p \leq k$. In a subsequent work [11], it is shown that the L_1 -budgeted variant is also NP-hard.

Paper Organization. We review some related work in Section 2. The formal problem definition is given in Section 3. The structural properties of the problem are given in Section 4 (where the more technical proofs are deferred to Section 10), and our local search framework for the unbudgeted variant is given in Section 5. We proved the NP-hardness of the budgeted variant in Section 6, and we consider the heuristic algorithms in Section 7. The experiments for the unbudgeted variant are in Section 8, and those for the budgeted variant are in Section 9. Finally, we mention some open direction in the conclusion.

2 RELATED WORK

To our knowledge, we are the first to consider an optimization framework based on opinion dynamics with varying susceptibility to persuasion. In the following, we review briefly some work that lies close to ours.

Susceptibility to Persuasion. Asch’s conformity experiments are perhaps the most famous study on the impact of agents’ susceptibility to change their opinions [6]. This study shows how agents have different propensities for conforming with others. These propensities are modeled in our context by the set of parameters α . Since the work of Asch, there have been various theories on peoples’ susceptibility to persuasion and how these can be affected. A notable example is Cialdini’s Six Principles of Persuasion, which highlight reciprocity, commitment and consistency, social proof, authority, liking, and scarcity, as key principles which can be utilized to alter people’s susceptibility to persuasion [12, 13]. This framework, and others, have been discussed in the context of altering susceptibility to persuasion in a variety of contexts. Crowley and Hoyer [15], and McGuire [41] discuss the “optimal arousal theory”, i.e., how novel stimuli can be utilized for persuasion when discussing arguments.

Opinion Dynamics Models. Opinion dynamics model social learning processes. DeGroot introduced an opinion dynamics model in his seminal work on consensus formation [20]. A set of n individuals in society start with initial opinions on a subject. Individual opinions are updated using the average of the neighborhood of a fixed social network. Friedkin and Johnsen [26] extended the DeGroot model to include both disagreement and consensus by mixing each individual's *innate belief* with some weight into the averaging process. This has inspired a lot of follow-up work, including [3, 9, 17, 29, 30].

Optimization and Opinion Dynamics. Bindel et al. use the Friedkin–Johnsen model as a framework for understanding the price of anarchy in society when individuals selfishly update their opinions in order to minimize the stress they experience [9]. They also consider network design questions: given a budget of k edges, and a node u , how should we add those k edges to u to optimize an objective related to the stress? Gionis, Terzi, and Tsaparas [30] use the same model to identify a set of target nodes whose innate opinions can be modified to optimize the sum of expressed opinions. Musco, Musco, and Tsourakakis adopt the same model to understand which graph topologies minimize the sum of disagreement and polarization [43].

Yildiz et al. [52] considered the placement of stubborn agents to maximize expected influence under the classical voter model.

Inferring Opinions and Conformity Parameters. While the expressed opinion of an agent is readily observable in a social network, both the agent's innate opinion and conformity parameter are hidden, and this leads to the question of inferring them. Such inference problems have been studied by Das et al. [16, 18]. Specifically, Das et al. give a near-optimal sampling algorithm for estimating the true average innate opinion of the social network and justify the algorithm both analytically and experimentally [18]. Das et al. view the problem of susceptibility parameter estimation as a problem in constrained optimization and give efficient algorithms, which they validate on real-world data [16].

Non-Convex Optimization. In general, optimizing a non-convex function under non-convex constraints is NP-hard. However, in many cases, one can exploit the structure of the objective function or constraints to devise polynomial-time algorithms; see the survey by Jain and Kar [34] on non-convex optimization algorithms encountered in machine learning. Indeed, variants of the gradient descent have been investigated to escape saddle points by Jin et al. [35], who also gave examples of problems where all local optima are also global optima; some examples are tensor decomposition [27], dictionary learning [48], phase retrieval [47], matrix sensing [8, 45], and matrix completion [28]. However, all these problems involve some quadratic loss functions, whose structures are totally different from our objective functions which involve matrix inverse.

Hartman [31] considered the special case that the objective function is the difference of two convex functions. Strelakovsky devised a local search method to optimize such objective functions. Even though the objective functions in our problem are somewhere convex and somewhere concave (see Figure 1), it is not immediately clear if they can be expressed as differences of convex functions.

3 MODEL

We consider a set V of *agents*, where each agent $i \in V$ is associated with an *innate opinion* $s_i \in [0, 1]$, where higher values correspond to more favorable opinions toward a given topic and a parameter measuring an agent's susceptibility to persuasion $\alpha_i \in (0, 1]$, where higher values signify agents who are less susceptible to changing their opinion. We call α_i the *resistance parameter*.

The agents interact with one another in discrete time steps. As a simplification to help readers to get some intuition on the problem, consider an (unweighted) undirected² graph $G = (V, E)$. The opinion dynamics evolve in discrete time according to the following model [20, 26] in a graph:

$$z_i(t+1) = \alpha_i s_i + (1 - \alpha_i) \cdot \frac{\sum_{j \in N(i)} z_j(t)}{\deg(i)}. \quad (1)$$

Here, $N(i) = \{j : \{i, j\} \in E\}$ is the set of neighbors of i in some graph, and $\deg(i) = |N(i)|$ is the degree of node i . However, in this work, we consider the general case where the interaction between agents is captured by a row stochastic matrix³ $P \in [0, 1]^{V \times V}$ (i.e., each entry of P is non-negative and every row sums to 1, but P needs not be symmetric). By denoting $A = \text{Diag}(\alpha)$ as the diagonal matrix with $A_{ii} = \alpha_i$ and I as the identity matrix, we can equivalently rewrite Equation (1) as

$$z^{(t+1)} := As + (I - A)Pz^{(t)}. \quad (2)$$

Equating $z^{(t)}$ with $z^{(t+1)}$, one can see that the equilibrium opinion vector is given by $z = [I - (I - A)P]^{-1}As$, which exists under very mild conditions. For instance, in the proof of Lemma 10.1, we see that a sufficient (but not necessary) condition for the equilibrium opinion vector to exist uniquely is that P is irreducible and at least one $i \in V$ has $\alpha_i > 0$. In the rest of this article, we always call P the interaction matrix.

We quantify Problem 1 as follows. The objective is to choose a resistance vector α to minimize the sum of equilibrium opinions $\langle \mathbf{1}, z \rangle = \mathbf{1}^\top z$, i.e., the goal is to drive the average opinion toward 0. Observe that one can also consider maximizing the sum of equilibrium opinions. To see that the minimization and maximization problems are equivalent, consider the transformation $x \mapsto 1 - x$ on the opinion space $[0, 1]$ that is applied to the innate opinions and expressed opinions in every time step. Then, it follows that Equation (2) still holds after the transformation. Hence, the minimization problem in the original opinion space is equivalent to the maximization problem in the transformed space. Either one is an optimization problem in which the goal is to drive the average opinion to one of the polar opposites $\{0, 1\}$. Indeed, to get readers to be familiar with both interpretations, the hardness is proved via the maximization variant in Section 6, while the algorithms are given via the minimization variant in Sections 5 and 7.

Definition 3.1 (Opinion Susceptibility Problem). Given a set V of agents with innate opinions $s \in [0, 1]^V$ and interaction matrix $P \in [0, 1]^{V \times V}$, suppose for each $i \in V$, its resistance is restricted to some interval $\mathcal{I}_i := [l_i, u_i] \subseteq [0, 1]$ where we assume that $0 < l_i < u_i < 1$.

The objective is to choose $\alpha \in \mathcal{I}_V := \times_{i \in V} \mathcal{I}_i \subseteq [0, 1]^V$ such that the following objective function is minimized:

$$f(\alpha) := \mathbf{1}^\top [I - (I - A)P]^{-1}As,$$

where $A = \text{Diag}(\alpha)$ is the diagonal matrix with $A_{ii} = \alpha_i$. Observe that the assumption $\alpha > 0$ ensures that the above inverse exists.

Unbudgeted vs Budgeted Variants. In Definition 3.1, we are allowed to modify the resistance of any agent, and this is known as the unbudgeted variant. We also consider the budgeted variant: given

²An unweighted graph $G = (V, E)$ is undirected if E is a collection of $2[-]$ subsets of V .

³Given sets U and W , we use the notation U^W to denote the collection of all functions from W to U . Each such function can also be interpreted as a vector (or a matrix if W itself is a Cartesian product), where each coordinate is labeled by an element in W and takes a value in U . As an example, a member of $[0, 1]^{V \times V}$ is a matrix whose rows and columns are labeled by elements of V . The alternative notation $[0, 1]^{n \times n}$ implicitly assumes a linear ordering on V , which does not have any importance in our case and would simply be an artefact of the notation.

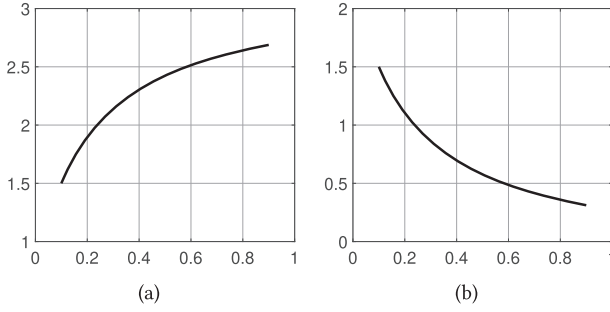


Fig. 1. Cross-sections of objective function.

some initial resistance vector and a budget k , the resistance of at most k agents can be changed. In this article, we focus on efficient algorithms that optimally solve the unbudgeted variant. In Section 6, we prove that the budgeted variant is NP-hard, and we propose efficient heuristics that scale to large networks. Designing algorithms with solid approximation guarantees for the budgeted variant is an interesting open problem.

Technical Assumption. To simplify our proofs, we assume that the interaction matrix P corresponds to an irreducible random walk. Irreducibility is satisfied if P arises from a connected graph.

4 STRUCTURAL PROPERTIES OF OBJECTIVE FUNCTION

In this section, we investigate the properties of the objective function f in Definition 3.1; we assume that the interaction matrix P and the innate opinion vector s are fixed, and f is a function on the resistance vector α .

Non-Convex Objective. Contrary to the claim in a preliminary version of this work (see [2]), the objective f in Definition 3.1 is in general not a convex function of α . In fact, the following example shows that it might be neither convex nor concave. Consider three vertices $V = \{1, 2, 3\}$, where the innate vector s and the interaction matrix P are given by $s = \begin{bmatrix} 1 \\ 0.5 \\ 0 \end{bmatrix}$ and $P = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}$.

Suppose we fix $\alpha_2 = \alpha_3 = 0.1$ and consider the objective as a function of α_1 as $g(\alpha_1) = \mathbf{1}^\top [I - (I - A)P]^{-1}As$, where $A = \text{Diag}(\alpha_1, \alpha_2, \alpha_3)$. Then, the plot of g in Figure 1(a) shows that it is not convex. Moreover, suppose this time we fix $\alpha_1 = \alpha_2 = 0.1$ and consider the objective as a function of α_3 as $h(\alpha_3) = \mathbf{1}^\top [I - (I - A)P]^{-1}As$. Then, the plot of h in Figure 1(b) shows that it is not concave.

Fortunately, we can still exploit some properties of the function. As we shall see, the main result of this section is Corollary 4.10, which states that every local optimum (which will be defined formally) is a global optimum. This enables us to use variants of the local search method to solve the problem optimally.

4.1 Marginal Monotonicity

As in [2], we show that when one chooses the resistance α_i for each agent $i \in V$, it suffices to consider the extreme points $\{l_i, u_i\}$. Our approach explicitly analyzes the partial derivative $\frac{\partial f(\alpha)}{\partial \alpha_i}$ which plays a major role in the local search algorithm that we later develop.

Intuition: Guidance by Current Equilibrium Vector. Observe that given the innate opinion vector s and irreducible interaction matrix P , for some resistance vector $\alpha \in (0, 1)^V$, the equilibrium opinion vector is given by $z(\alpha) = [I - (I - A)P]^{-1}As$, where $A = \text{Diag}(\alpha)$. For some $i \in V$, if the innate opinion s_i is larger than its equilibrium $z_i(\alpha)$, this suggests that by being more stubborn,

agent i should be able to increase its equilibrium opinion. In other words, one would expect $\frac{\partial z_i(\alpha)}{\partial \alpha_i}$ and $s_i - z_i(\alpha)$ to have the same sign. However, what is surprising is that in Lemma 4.2, we shall see that even for any $j \in V$, $\frac{\partial z_j(\alpha)}{\partial \alpha_i}$ and $s_i - z_i(\alpha)$ have the same sign.

Notation. For any $\alpha \in \mathbb{R}^V$ and $K \subseteq V$, let $\alpha_{-K} \in \mathbb{R}^V$ denote the vector such that $\alpha_{-K}(i) = \alpha(i)$ if $i \notin K$, and $\alpha_{-K}(i) = 0$ if $i \in K$, i.e., the coordinates K of α are replaced with 0. Similarly, given $\alpha \in \mathbb{R}^V$, we denote $A_{-K} = \text{Diag}(\alpha_{-K})$.

In Definition 3.1, observe that the inverse $[I - (I - A)P]^{-1}$ is involved in the objective function $f(\alpha)$, where $A = \text{Diag}(\alpha)$. Since we wish to analyze the effect on $f(\alpha)$ of changing only a subset of coordinates in α , the next lemma will be used for simplifying matrix arithmetic involving the computation of inverses. Its proof is deferred to Section 10.

LEMMA 4.1 (INVERSE ARITHMETIC). *Suppose $\alpha \in [0, 1]^V$ has at least one non-zero coordinate and P is an irreducible interaction matrix. Denote $A := \text{Diag}(\alpha)$. Then, the inverse $M = [I - (I - A)P]^{-1}$ exists, and every entry of M is positive. Moreover, for every $k \in V$, we have*

- (1) $(PM)_{kk} = \frac{M_{kk} - 1}{1 - \alpha_k} > 0$;
- (2) $(PM)_{kj} = \frac{M_{kj}}{1 - \alpha_k} > 0$, for each $j \neq k$.

LEMMA 4.2 (SIGN OF PARTIAL DERIVATIVE). *In the Opinion Susceptibility Problem in Definition 3.1, given the innate opinion vector s and irreducible interaction matrix P , recall that $z(\alpha) := [I - (I - A)P]^{-1}As$, where $A = \text{Diag}(\alpha)$. Then, for any $\alpha \in (0, 1)^V$ and any $i, k \in V$, the two values $\frac{\partial z_k(\alpha)}{\partial \alpha_i}$ and $s_i - z_i(\alpha)$ have exactly the same sign in $\{-, 0, +\}$.*

In particular, this implies that $\frac{\partial f(\alpha)}{\partial \alpha_i} = \sum_{k \in V} \frac{\partial z_k(\alpha)}{\partial \alpha_i}$ also has the same sign as $s_i - z_i(\alpha)$.

PROOF. By the definition of the inverse of a matrix B , we have $BB^{-1} = I$. The partial derivative with respect to a variable t is $\frac{\partial B}{\partial t}B^{-1} + B\frac{\partial B^{-1}}{\partial t} = 0$. Hence, we have $\frac{\partial B^{-1}}{\partial t} = -B^{-1}\frac{\partial B}{\partial t}B^{-1}$. Applying the above result with $B = I - (I - A)P$ and $t = \alpha_i$ and denoting $M = [I - (I - A)P]^{-1}$, we get $\frac{\partial M}{\partial \alpha_i} = -M\mathbf{e}_i\mathbf{e}_i^\top PM$. Considering $z(\alpha) = MAs$, we have $\frac{\partial z(\alpha)}{\partial \alpha_i} = \frac{\partial M}{\partial \alpha_i}As + M\mathbf{e}_i\mathbf{e}_i^\top s$. Replacing $\frac{\partial M}{\partial \alpha_i}$, we obtain for any $i, k \in V$:

$$\frac{\partial z_k(\alpha)}{\partial \alpha_i} = -\mathbf{e}_k^\top M\mathbf{e}_i\mathbf{e}_i^\top PMAs + \mathbf{e}_k^\top M\mathbf{e}_i\mathbf{e}_i^\top s = M_{ki} \cdot [s_i - \mathbf{e}_i^\top Pz(\alpha)].$$

By Lemma 4.1, we know that every entry of M is positive. Thus, the sign of $\frac{\partial z_k(\alpha)}{\partial \alpha_i}$ is the same as that of the scalar $s_i - \mathbf{e}_i^\top Pz(\alpha)$.

Recalling $M = [I - (I - A)P]^{-1}$, we have $[I - (I - A)P]M = I \Rightarrow (I - A)PM = M - I \Rightarrow PM = (I - A)^{-1}(M - I)$ where $(I - A)^{-1}$ exists since $\alpha_j < 1$ for each $j \in V$.

Next, since $z(\alpha) = MAs$, we have

$$Pz(\alpha) = PMAs = (I - A)^{-1}(M - I)As = (I - A)^{-1}[z(\alpha) - As].$$

Finally, replacing $Pz(\alpha)$, we have

$$\begin{aligned} s_i - \mathbf{e}_i^\top Pz(\alpha) &= s_i - \mathbf{e}_i^\top (I - A)^{-1}[z(\alpha) - As] \\ &= s_i - \frac{1}{1 - \alpha_i}[z_i(\alpha) - \alpha_i s_i] \\ &= \frac{1}{1 - \alpha_i}[s_i - z_i(\alpha)]. \end{aligned}$$

Since $1 - \alpha_i > 0$, it follows that $\frac{\partial z_k(\alpha)}{\partial \alpha_i}$ and $s_i - z_i(\alpha)$ have exactly the same sign in $\{-, 0, +\}$, as required. \square

The next lemma shows that the sign of the partial derivatives with respect to coordinate i is actually independent of the current value α_i . Its proof is deferred to Section 10.

LEMMA 4.3 (SIGN OF PARTIAL DERIVATIVE INDEPENDENT OF COORDINATE VALUE). *Referring to Lemma 4.2. For any $\alpha \in (0, 1)^V$ and any $i \in V$, denote $M = [I - (I - A_{-\{i\}})P]^{-1}$. Then, $\frac{\partial f(\alpha)}{\partial \alpha_i}$ has the same sign in $\{-, 0, +\}$ as $s_i - \sum_{j \neq i} M_{ij} \alpha_j s_j$, which is independent of α_i .*

COROLLARY 4.4 (EXTREME POINTS ARE SUFFICIENT). *In Definition 3.1, for any $i \in V$, fixing the resistance values of all other agents except i , the objective $f(\alpha)$ is a monotone function in α_i . This implies that to minimize f , it suffices to consider the extreme points $\alpha_i \in \{l_i, u_i\}$, for each $i \in V$.*

PROOF. If $\alpha_i \mapsto f(\alpha)$ is a monotone function (that is either monotonically increasing or decreasing), then the function achieves its minimum at either $\alpha_i = l_i$ or $\alpha_i = u_i$.

Applying this argument to every $i \in V$ gives the result. \square

4.2 Local vs Global Optimum

As shown in Corollary 4.4, it suffices to choose the resistance vector α from the extreme points in Definition 3.1. Lemma 4.2 readily gives a method to decide, given a current choice of α , whether the objective f can be decreased by changing the resistance of some agent. In Lemma 4.9, we show that if α is not a global minimum, then such an agent must exist. As we shall see, this implies that a local search method can find a global minimum.

Given α and $\alpha' \in \mathbb{R}^V$, denote $\Delta(\alpha, \alpha') := \{i \in V : \alpha_i \neq \alpha'_i\}$ as the set of coordinates at which the vectors differ.

Definition 4.5 (Local Minimzer). Given an objective function $f : \mathcal{I}_V \rightarrow \mathbb{R}$, a vector $\alpha \in \mathcal{I}_V$ is a local minimizer of f , if for all $\alpha' \in \mathcal{I}_V$ such that $\Delta(\alpha, \alpha') = 1$, $f(\alpha) \leq f(\alpha')$.

Notation. When we wish to consider the effect of changing the resistance of only two agents $i \neq k \in V$, we write $f(\alpha) = f(\alpha_i, \alpha_k)$, assuming that $\alpha_{-\{i,k\}}$ is fixed.

Lemmas 4.6 and 4.7 give some technical results involving changing the resistance of two agents. Their proofs are deferred to Section 10.

LEMMA 4.6. *For any $i, k \in V$ such that $i \neq k$, let $M = [I - (I - A_{-\{i\}})P]^{-1}$ and $R = [I - (I - A_{-\{i,k\}})P]^{-1}$. Then, for any $j \in V$, we have*

$$(1) M_{jk} = \frac{R_{jk}}{1 + \alpha_k R_{kk} - \alpha_k},$$

$$(2) M_{jh} = R_{jh} - \frac{\alpha_k R_{jk} R_{kh}}{1 + \alpha_k R_{kk} - \alpha_k}, \text{ for } h \neq k.$$

In particular, the quantity in Lemma 4.3 can be rewritten as follows:

$$s_i - \sum_{j \neq i} M_{ij} \alpha_j s_j = s_i - \sum_{j \neq i, k} R_{ij} \alpha_j s_j - \frac{\alpha_k R_{ik}}{1 + \alpha_k R_{kk} - \alpha_k} \left(s_k - \sum_{j \neq i, k} R_{kj} \alpha_j s_j \right).$$

LEMMA 4.7 (DIAGONAL ENTRY). *Suppose $\alpha \in (0, 1)^V$, recall that $A_{-\{i,k\}} := \text{Diag}(\alpha_{-\{i,k\}})$, and P corresponds to an irreducible interaction matrix. For any $i, k \in V$ such that $i \neq k$, let $R = [I - (I - A_{-\{i,k\}})P]^{-1}$, then $R_{ii} = \max_{j \in V} R_{ji}$. Moreover, $R_{ii} = R_{ki}$ if and only if $P_{kk} + P_{ki} = 1$.*

The following lemma gives the key insight for why local search works. Intuitively, it shows that there does not exist any discrete ‘‘saddle point’’. Even though its proof is technical, we still include it here because of its importance.

LEMMA 4.8 (SWITCHING LEMMA). Recall that f is defined in Definition 3.1 with an irreducible interaction matrix P , and assume $|V| \geq 3$. Suppose $\alpha, \beta \in (0, 1)^V$ such that $\Delta(\alpha, \beta) = \{i, k\}$ for some $i \neq k$. Moreover, suppose further that

$$\min\{f(\alpha_i, \alpha_k), f(\beta_i, \beta_k)\} < \min\{f(\alpha_i, \beta_k), f(\beta_i, \alpha_k)\}.$$

Then, we have

$$\max\{f(\alpha_i, \alpha_k), f(\beta_i, \beta_k)\} > \min\{f(\alpha_i, \beta_k), f(\beta_i, \alpha_k)\}.$$

PROOF. We prove the lemma by contradiction. Suppose

$$\max\{f(\alpha_i, \alpha_k), f(\beta_i, \beta_k)\} \leq \min\{f(\alpha_i, \beta_k), f(\beta_i, \alpha_k)\}.$$

Without loss of generality, suppose further that $f(\alpha_i, \alpha_k) \geq f(\beta_i, \beta_k)$. Then, we have

$$f(\alpha_i, \alpha_k) \leq \begin{cases} f(\beta_i, \alpha_k) \\ f(\alpha_i, \beta_k) \end{cases} \quad \text{and} \quad f(\beta_i, \beta_k) < \begin{cases} f(\beta_i, \alpha_k) \\ f(\alpha_i, \beta_k) \end{cases}.$$

We remark that it is important to distinguish between the strict and non-strict inequality. We use the notation f'_i to denote the partial derivative with respect to coordinate i .

From $f(\alpha_i, \alpha_k) \leq f(\beta_i, \alpha_k)$ and the fact that f is marginally monotone (Lemma 4.3) and $f'_i(x, \alpha_k)$ has the same sign in $\{-, 0, +\}$ for $x \in (0, 1)$, we have

$$f'_i(x, \alpha_k) \cdot (\alpha_i - \beta_i) \leq 0. \quad (3)$$

On the other hand, from the strict inequality $f(\beta_i, \beta_k) < f(\beta_i, \alpha_k)$, we know the partial derivative $f'_k(\beta_i, y)$ must have the same non-zero sign in $\{-, +\}$ for all $y \in (0, 1)$, again from Lemma 4.3. Therefore, we have

$$f'_k(\beta_i, y) \cdot (\alpha_k - \beta_k) > 0. \quad (4)$$

Similarly, $f(\alpha_i, \alpha_k) \leq f(\alpha_i, \beta_k)$ and $f(\beta_i, \beta_k) < f(\alpha_i, \beta_k)$ give that for all $x, y \in (0, 1)$, we have

$$f'_k(\alpha_i, y) \cdot (\alpha_k - \beta_k) \leq 0, \quad (5)$$

$$f'_i(x, \beta_k) \cdot (\alpha_i - \beta_i) > 0. \quad (6)$$

Next, using Lemma 4.3 and R as defined in Lemma 4.6, the above inequalities (3)–(6) become

$$\begin{aligned} & \left[s_i - \sum_{j \neq i, k} R_{ij} \alpha_j s_j - \frac{\alpha_k R_{ik}}{1 + \alpha_k R_{kk} - \alpha_k} \left(s_k - \sum_{j \neq i, k} R_{kj} \alpha_j s_j \right) \right] (\alpha_i - \beta_i) \leq 0, \\ & \left[s_k - \sum_{j \neq i, k} R_{kj} \beta_j s_j - \frac{\beta_i R_{ki}}{1 + \beta_i R_{ii} - \beta_i} \left(s_i - \sum_{j \neq i, k} R_{ij} \beta_j s_j \right) \right] (\alpha_k - \beta_k) > 0, \\ & \left[s_k - \sum_{j \neq i, k} R_{kj} \alpha_j s_j - \frac{\alpha_i R_{ki}}{1 + \alpha_i R_{ii} - \alpha_i} \left(s_i - \sum_{j \neq i, k} R_{ij} \alpha_j s_j \right) \right] (\alpha_k - \beta_k) \leq 0, \\ & \left[s_i - \sum_{j \neq i, k} R_{ij} \beta_j s_j - \frac{\beta_k R_{ik}}{1 + \beta_k R_{kk} - \beta_k} \left(s_k - \sum_{j \neq i, k} R_{kj} \beta_j s_j \right) \right] (\alpha_i - \beta_i) > 0. \end{aligned}$$

Recall that $\alpha_j = \beta_j$ for $j \neq i, k$. Hence, we denote

$$c_i := s_i - \sum_{j \neq i, k} R_{ij} \alpha_j s_j = s_i - \sum_{j \neq i, k} R_{ij} \beta_j s_j,$$

$$c_k := s_k - \sum_{j \neq i, k} R_{kj} \alpha_j s_j = s_k - \sum_{j \neq i, k} R_{kj} \beta_j s_j,$$

$$g_i(x) := \frac{x R_{ki}}{1 + x R_{ii} - x} \text{ and } g_k(x) := \frac{x R_{ik}}{1 + x R_{kk} - x}.$$

Then, we have

$$[c_i - g_k(\alpha_k) c_k](\alpha_i - \beta_i) \leq 0, \quad (7)$$

$$[c_k - g_i(\beta_i) c_i](\alpha_k - \beta_k) > 0, \quad (8)$$

$$[c_k - g_i(\alpha_i) c_i](\alpha_k - \beta_k) \leq 0, \quad (9)$$

$$[c_i - g_k(\beta_k) c_k](\alpha_i - \beta_i) > 0. \quad (10)$$

Observe that $c_i \neq 0$, otherwise (8) and (9) contradict each other. Similarly, $c_k \neq 0$, otherwise (7) and (10) contradict each other. We next argue that $c_i c_k > 0$.

From (7) and (10), we have

$$[c_i - g_k(\alpha_k) c_k][c_i - g_k(\beta_k) c_k] \leq 0. \quad (11)$$

If $c_i c_k < 0$, then the above expression will be positive, because $g_k(\cdot) \geq 0$ (we shall see that later). Hence, we conclude that c_i and c_k have the same sign.

From (8) and (9), we have

$$[c_k - g_i(\beta_i) c_i][c_k - g_i(\alpha_i) c_i] \leq 0. \quad (12)$$

Rearranging (11) and (12), we have

$$\left[\frac{c_i}{c_k} - g_k(\alpha_k) \right] \left[\frac{c_i}{c_k} - g_k(\beta_k) \right] \leq 0$$

and

$$\left[\frac{c_k}{c_i} - g_i(\beta_i) \right] \left[\frac{c_k}{c_i} - g_i(\alpha_i) \right] \leq 0.$$

Note that every entry of R is positive by Lemma 4.1 and we can easily prove $g_i(\cdot)$ and $g_k(\cdot)$ are both strictly increasing functions in $[0, 1]$. Since $\alpha, \beta \in (0, 1)^V$, the above two inequalities imply that

$$0 = g_k(0) < g_k(\min\{\alpha_k, \beta_k\}) \leq \frac{c_i}{c_k},$$

$$\frac{c_i}{c_k} \leq g_k(\max\{\alpha_k, \beta_k\}) < g_k(1) = \frac{R_{ik}}{R_{kk}} \leq 1,$$

and

$$0 = g_i(0) < g_i(\min\{\alpha_i, \beta_i\}) \leq \frac{c_k}{c_i},$$

$$\frac{c_k}{c_i} \leq g_i(\max\{\alpha_i, \beta_i\}) < g_i(1) = \frac{R_{ki}}{R_{ii}} \leq 1,$$

where $\frac{R_{ik}}{R_{kk}} \leq 1$ and $\frac{R_{ki}}{R_{ii}} \leq 1$ are from Lemma 4.7.

Notice that we get $0 < \frac{c_i}{c_k} < 1$ and $0 < \frac{c_k}{c_i} < 1$, which is a contradiction. Hence, the proof is completed. \square

LEMMA 4.9 (DESCENDING COORDINATE). *Let f be the function as defined in Definition 3.1. Suppose $\alpha, \beta \in (0, 1)^V$ such that $f(\alpha) > f(\beta)$. Then, there exists some $i \in \Delta(\alpha, \beta)$ and $\gamma \in (0, 1)^V$ such that $\Delta(\alpha, \gamma) = \{i\}$, $\Delta(\gamma, \beta) = \Delta(\alpha, \beta) \setminus \{i\}$, and $f(\alpha) > f(\gamma)$.*

In other words, by switching one coordinate (corresponding to i) of α to that of β , the objective function f decreases strictly.

PROOF. We prove the lemma by induction on $|\Delta(\alpha, \beta)|$. The base case $|\Delta(\alpha, \beta)| = 1$ is trivial.

We consider the inductive step with $|\Delta(\alpha, \beta)| = q$, for some $q \geq 2$. Given a list S of coordinates from $\Delta(\alpha, \beta)$, we use $\alpha_{[S]}$ to denote the resulting vector obtained from switching coordinates S of α to those of β .

For contradiction's sake, we assume that for all $j \in \Delta(\alpha, \beta)$, $f(\alpha_{[j]}) \geq f(\alpha)$; moreover, we pick $i \in \Delta(\alpha, \beta)$ such that $f(\alpha_{[i]})$ is minimized.

Observe that $f(\alpha_{[i]}) \geq f(\alpha) > f(\beta)$ and $|\Delta(\alpha_{[i]}, \beta)| = q - 1$. Therefore, by the induction hypothesis, there exists some $k \in \Delta(\alpha_{[i]}, \beta)$ such that $f(\alpha_{[i]}) > f(\alpha_{[i,k]})$.

Next, starting from α , we shall fix all coordinates in V except i and k , and we write the objective $f(x, y)$ as a function on only these two coordinates.

Observe that we have already assumed that

$$f(\alpha_i, \alpha_k) \leq \min\{f(\beta_i, \alpha_k), f(\alpha_i, \beta_k)\}. \quad (13)$$

Moreover, from above, we have $f(\alpha_{[i]}) > f(\alpha_{[i,k]})$, which translates to $f(\beta_i, \alpha_k) > f(\beta_i, \beta_k)$. Observe that we must have $f(\alpha_i, \beta_k) \leq f(\beta_i, \beta_k)$; otherwise, we have $f(\beta_i, \beta_k) < \min\{f(\beta_i, \alpha_k), f(\alpha_i, \beta_k)\}$, which, together with (13), will contradict Lemma 4.8.

Therefore, we have $f(\alpha_{[k]}) = f(\alpha_i, \beta_k) \leq f(\beta_i, \beta_k) < f(\beta_i, \alpha_k) = f(\alpha_{[i]})$, which contradicts the choice of $i \in \Delta(\alpha, \beta)$ to minimize $f(\alpha_{[i]})$. This completes the inductive step and also the proof of the lemma. \square

COROLLARY 4.10. *For the function f in Definition 3.1 every local minimizer is a global minimizer.*

PROOF. Suppose that α is a local minimizer, but there is some β with $f(\alpha) > f(\beta)$. Then, Lemma 4.9 implies there is some γ with $|\Delta(\alpha, \gamma)| = 1$ such that $f(\alpha) > f(\gamma)$, contradicting that α is a local minimizer. \square

5 EFFICIENT LOCAL SEARCH

We have explained in Section 3 that our optimization problem can be interpreted equivalently as either a minimization or a maximization problem. Since our algorithm can be seen as a variant of gradient descent and local search, it will be more intuitive to view it as a minimization problem.

In Section 4, we conclude in Corollary 4.4 that it suffices to consider the extreme points of the search space of resistance vectors. Moreover, Corollary 4.10 states that every local minimizer is a global minimizer. Since we know how to compute the sign of the partial derivative with respect to each coordinate using Lemma 4.2, we can design a simple local search algorithm to find a global minimizer.

However, it is possible that $O(2^n)$ extreme points are encountered before a global minimizer is reached. Fortunately, in this section, we will explore further properties of the objective function, and design a local search algorithm that encounters at most $O(n)$ extreme points before finding a global minimizer.

5.1 Irrevocable Updates

Local Search Strategy. We shall start with the upper bound resistance vector, i.e., for each $i \in V$, $\alpha_i = u_i$. This induces the corresponding equilibrium opinion vector $z(\alpha)$. According to Lemma 4.2, if there is some agent i such that $\alpha_i = u_i$ and $s_i - z_i(\alpha) > 0$, then we should flip α_i to the lower bound l_i . The following lemma shows that each α_i will be flipped at most once. Essentially, we show that we will never encounter the situation that there is some agent k such that $\alpha_k = l_k$ and $s_k - z_k(\alpha) < 0$, in which case we would have to switch α_k back to u_k .

LEMMA 5.1 (EACH COORDINATE FLIPPED AT MOST ONCE). *Starting from the upper bound resistance vector, the above local search strategy flips each α_i at most once.*

PROOF. We first show that for each agent $k \in V$, the quantity $s_k - z_k(\alpha)$ cannot decrease when α is modified according to the local search strategy. According to the strategy, α is modified because there is some agent i such that $\alpha_i = u_i$ and $s_i - z_i(\alpha) > 0$. By Lemma 4.2, $\frac{\partial z_k}{\partial \alpha_i} > 0$ for each $k \in V$. Hence, after α_i is switched from u_i to l_i , $z_k(\alpha)$ decreases, and the quantity $s_k - z_k(\alpha)$ increases.

Observe that if a coordinate α_k is ever flipped from u_k to l_k , this means that at that moment, we must have $s_k - z_k(\alpha) > 0$, which, as we have just shown, will stay positive after α is subsequently updated. \square

5.2 Approximating the Equilibrium Vector

Observe that in our local search algorithm, we need to compute the equilibrium opinion vector $z(\alpha) = [I - (I - A)P]^{-1}As$ for the current resistance vector α , where $A = \text{Diag}(\alpha)$. However, computing matrix inverse is an expensive operation. Instead, we approximate $z(\alpha)$ using the recurrence $z^{(0)} \in [0, 1]^V$ and $z^{(t+1)} := As + (I - A)Pz^{(t)}$. The following lemma gives an upper bound on the additive error for each coordinate.

LEMMA 5.2 (APPROXIMATION ERROR). *Suppose for some $\epsilon > 0$, for all $i \in V$, $\alpha_i \geq \epsilon$. Then, for every $t \geq 0$, $\|z(\alpha) - z^{(t)}\|_\infty \leq \frac{(1-\epsilon)^t}{\epsilon}$.*

PROOF. Using the Neumann series $[I - (I - A)P]^{-1} = \sum_{j=0}^{\infty} [(I - A)P]^j$, we have $z(\alpha) - z^{(t)} = \sum_{j=t}^{\infty} [(I - A)P]^j As - [(I - A)P]^t z^{(0)}$.

We next prove, by induction, that for any $x \in [0, 1]^V$, $\|[(I - A)P]^j x\|_\infty \leq (1 - \epsilon)^j$, for all $j \geq 0$. The base case $j = 0$ is trivial because every coordinate of x is between 0 and 1. For the inductive step, assume that for some $j \geq 0$, every coordinate of $y = [(I - A)P]^j x$ has magnitude at most $(1 - \epsilon)^j$. Since P is a row stochastic matrix, it follows that $\|Py\|_\infty \leq (1 - \epsilon)^j$; finally, since $\alpha_i \geq \epsilon$ for all $i \in V$, we have $\|[(I - A)P]y\|_\infty \leq (1 - \epsilon)^{j+1}$, completing the induction proof.

Finally, observing that both $\sum_{j=t}^{\infty} [(I - A)P]^j As$ and $[(I - A)P]^t z^{(0)}$ have non-negative coordinates, we have

$$\begin{aligned} \|z(\alpha) - z^{(t)}\|_\infty &\leq \max \left\{ \left\| \sum_{j=t}^{\infty} [(I - A)P]^j As \right\|_\infty, \left\| [(I - A)P]^t z^{(0)} \right\|_\infty \right\} \\ &\leq \sum_{j=t}^{\infty} (1 - \epsilon)^j = \frac{(1 - \epsilon)^t}{\epsilon}, \end{aligned}$$

as required. \square

5.3 Local Search Algorithm

Based on Lemmas 5.1 and 5.2, we give a local search framework in Algorithm 1. Observe that in line 1, we perturb the innate opinions s slightly to ensure that for each resistance vector α encountered, no coordinate of s and $z(\alpha)$ would coincide. The reason is that if for some $i \in V$, $s_i = z_i(\alpha)$, then no matter how many iterations we run, the estimation z_i of $z_i(\alpha)$ will satisfy $|s_i - z_i| \leq \text{err}(t)$, which means that the **while** loop in line 4 will not terminate. However, $s_i = z_i(\alpha)$ means that the objective function stays the same as coordinate i changes; this implies that the optimal solution will be reached eventually, but the algorithm might not know when to terminate.

ALGORITHM 1: Local Search Framework

Input: Innate opinions $s \in [0, 1]^V$; interaction matrix P ; for each agent $i \in V$, upper u_i and lower l_i bounds for resistance.

Output: Optimal resistance vector $\alpha \in \times_{i \in V} \{l_i, u_i\}$.

- 1 (*Technical step.*) Randomly perturb each coordinate of s slightly.
- 2 Initially, for each agent i , set $\alpha_i \leftarrow u_i$ to its upper bound; denote $\epsilon_\alpha := \min_{i \in V} \alpha_i$.
- 3 Pick arbitrary $z \in [0, 1]^V$, and set $t \leftarrow 0$; denote $\text{err}(t) := \frac{(1-\epsilon_\alpha)^t}{\epsilon_\alpha}$.
- 4 **while** $\exists i \in V : |s_i - z_i| \leq \text{err}(t) \vee (z_i < s_i \wedge \alpha_i = u_i)$ **do**
- 5 $z \leftarrow As + (I - A)Pz$, where $A = \text{Diag}(\alpha)$.
- 6 $t \leftarrow t + 1$.
- 7 (*Non-deterministic step.*) Pick arbitrary $L \subseteq V$ (that can be empty) such that for each $i \in L$, $z_i \leq s_i - \text{err}(t)$ and $\alpha_i = u_i$.
- 8 **if** $L \neq \emptyset$ **then**
- 9 **for each** $i \in L$ **do**
- 10 Set $\alpha_i \leftarrow l_i$ to its lower bound (and update ϵ_α).
- 11 $t \leftarrow 0$.
- 12 **return** Resistance vector α .

The **while** loop in line 4 combines local search to update α and estimation of the equilibrium vector $z(\alpha)$. Here are two general update strategies, which are both captured by the non-deterministic step in line 7:

- **Conservative Update.** The opinion vector z is iteratively updated in line 5 until all coordinates of z and s are sufficiently far apart. Then, for every coordinate α_i such that $\alpha_i = u_i$ and $z_i < s_i$, we flip α_i to the lower bound l_i . After we update α , we reset t to 0, and continue to iteratively update z . Whenever we update α and set t to 0, we say that a new *phase* begins; we use the convention that the initial phase is known as phase 0.
- **Opportunistic Update.** Instead of waiting for the approximation error of every coordinate of z to be small enough, we can update some coordinates α_i , if $\alpha_i = u_i$ and $z_i \leq s_i - \text{err}(t)$ are small enough. However, there is some tradeoff between waiting for the errors of all coordinates to be small enough and updating coordinates of α that are ready sooner. We might be tempted to update coordinates of α as soon as they are ready, but once α is modified, we might need to wait for many iterations for the estimation of $z(\alpha)$ to be accurate enough again. Therefore, we might also choose to wait for more coordinates of α to be ready before we flip them all at once. In Section 8, we will evaluate empirically different update strategies.

Optimistic Update. In both conservative and opportunistic updates, a coordinate α_i is flipped only when we know for sure that the current estimate z_i has small enough error with respect to the equilibrium $z_i(\alpha)$; hence, no mistake in flipping any α_i is ever made. However, our insight is that as the algorithm proceeds, the general trend is for every z_i to decrease.

The first intuition is that if we start with some $z^{(0)}$ such that every coordinate of $z^{(0)}$ is at least its equilibrium coordinate of $z(\alpha)$, then $z^{(t)}$ should converge to $z(\alpha)$ from above. The second observation is that every time we flip some α_i , this will not increase any coordinate of the equilibrium vector $z(\alpha)$, thereby preserving the condition that the current estimate $z^{(t)} \geq z(\alpha)$. Hence, without worrying about the accuracy of the current estimate z , we will simply flip coordinate α_i to l_i when z_i drops below s_i . However, it is possible that we might need to flip α_i back to u_i , if z_i increases in

ALGORITHM 2: Optimistic Local Search

Input: Innate opinions $s \in [0, 1]^V$; interaction matrix P ; for each agent $i \in V$, upper u_i and lower l_i bounds for resistance.

Output: Optimal resistance vector $\alpha \in \times_{i \in V} \{l_i, u_i\}$.

- 1 (Technical step.) Randomly perturb each coordinate of s slightly.
- 2 Initially, for each agent i , set $\alpha_i \leftarrow u_i$ to its upper bound; denote $\epsilon_\alpha := \min_{i \in V} \alpha_i$.
- 3 Pick $z = (1, 1, \dots, 1)$, and set $t \leftarrow 0$; denote $\text{err}(t) := \frac{(1-\epsilon_\alpha)^t}{\epsilon_\alpha}$.
- 4 **while** $\exists i \in V : |s_i - z_i| \leq \text{err}(t)$ **do**
- 5 $z \leftarrow As + (I - A)Pz$, where $A = \text{Diag}(\alpha)$.
- 6 $t \leftarrow t + 1$.
- 7 (Optimistic Candidates.) Set $L \leftarrow \{i \in V : z_i \leq s_i \wedge \alpha_i = u_i\}$.
- 8 (Rare Mistakes.) Set $J \leftarrow \{i \in V : z_i > s_i \wedge \alpha_i = l_i\}$.
- 9 **if** $L \cup J \neq \emptyset$ **then**
- 10 **for each** $i \in L$ **do**
- 11 Set $\alpha_i \leftarrow l_i$ to its lower bound (and update ϵ_α).
- 12 **for each** $i \in J$ **do**
- 13 Set $\alpha_i \leftarrow u_i$ to its upper bound (and update ϵ_α).
- 14 $t \leftarrow 0$.
- 15 **return** Resistance vector α .

the next iteration and becomes larger than s_i again. We shall see in Section 8 that this scenario is extremely rare. Specifically, in line 8 of Algorithm 2, the set J is (almost) always empty.

6 NP-HARD BUDGETED OPINION SUSCEPTIBILITY PROBLEM

We now consider the setting where there is a constraint on the size of the target-set. That is, we want to identify a set $T \subseteq V$ of size k such that changing the resistance parameters of agents in T optimally maximizes the sum of equilibrium opinions. Even though the minimization and the maximization problems are equivalent, in the NP-hardness proof, the maximization variant will make the calculation slightly simpler, because in this case, the “default” opinion value is 0 and the algorithm tries to increase agents’ opinions. This means that under the maximization variant of the hardness proof, the calculation will have more opinion values with 0.

We use $\alpha^{(0)}$ to denote the given initial resistance vector. For $T \subseteq V$, we define $F(T) := \max\{f(\alpha) : \forall i \notin T, \alpha_i = \alpha_i^{(0)}\}$; observe that F is defined with respect to the initial resistance vector $\alpha^{(0)}$. The budgeted opinion optimization problem is to maximize $F(T)$ subject to the budget constraint $|T| \leq k$.

THEOREM 6.1. *The budgeted opinion optimization problem is NP-hard.*

PROOF. We give a reduction from the vertex cover problem for regular graphs. Given a d -regular graph $G = (V, E)$ and an integer K , the vertex cover problem asks whether there exists a set S of nodes with size at most K such that S is a vertex cover, i.e., every edge in E is incident to at least one node in S . For simplicity, we assume that \sqrt{d} is an integer.

We give a reduction from the above vertex cover problem to the decision version of the opinion optimization problem. In addition to a given graph G' , the innate opinion vector s , the initial resistance vector $\alpha^{(0)}$ and the budget k , an instance of the decision version of opinion maximization also has some threshold θ . The instance is “yes” iff there exists some node set T in G' with size at most k such that $F(T) \geq \theta$. To illustrate our ideas, we first give a reduction in which each agent’s

resistance parameter is in the range $[0, 1]$. Then, we show how to restrict the resistance to the range $[\epsilon, 1]$ for some small enough $\epsilon > 0$.

Reduction Construction. Suppose we are given an instance of the vertex cover problem for regular graphs. We construct an instance of the decision version of the opinion optimization problem. Define $G' = (V \cup V', E \cup E')$, where V and E come from original vertex cover problem. For $i \in V$, $s_i = 1$ and $\alpha^{(0)} = 0$; for $i' \in V'$, $s_{i'} = 0$, and we give more details on their initial resistance parameters. The additional nodes V' and edges E' are added as follows. Let $\sigma = 2n^2(\sqrt{d} + 1)$ (specified later).

For each $i \in V$, we add $(\sigma + 1)\sqrt{d}$ additional nodes V_i into V' :

- (i) \sqrt{d} flexible nodes. Each such node has degree 1 and is connected only to node i ; its initial resistance parameter in $\alpha^{(0)}$ is 0.
- (ii) $\sigma\sqrt{d}$ stubborn nodes. These nodes form \sqrt{d} cliques, each of which has size σ . In each clique, exactly one node is connected to i . All the stubborn nodes have initial resistance parameters in $\alpha^{(0)}$ being 1.

Observe that in G' , the degree of each node in V is $d + 2\sqrt{d}$. Finally, we set the budget $k = K$ and the threshold $\theta = (\sqrt{d} + 1)k + (\sqrt{d} + 1)(n - k)\frac{d}{d + \sqrt{d}}$. To complete the reduction proof, we show that there exists a vertex cover of size k in G iff there exists some $T \subseteq V \cup V'$ of size k such that $F(T) \geq \theta$.

Forward Direction. Suppose in G , there is some vertex cover $T \subseteq V$ with size k . We show that in G' , $F(T) \geq \theta$; we set $\alpha_i = 1$ for each $i \in T$, while the resistance parameters of all other nodes remain the same as in $\alpha^{(0)}$. We next analyze the equilibrium opinion of each node. Observe that all stubborn nodes in V' have equilibrium opinion 0.

For $i \in T$, node i has equilibrium opinion 1; moreover, all its \sqrt{d} flexible neighbors in V' will also have equilibrium opinion 1.

For $j \in V \setminus T$, we compute its equilibrium opinion z_j . Since T is a vertex cover, all d neighbors of j in V are in T and have equilibrium opinion 1. All \sqrt{d} flexible neighbors of j in V' have the same equilibrium opinion z_j , while the \sqrt{d} stubborn neighbors have opinion 0. Therefore, z_j satisfies the equation $z_j = \frac{d}{d + 2\sqrt{d}} \cdot 1 + \frac{\sqrt{d}}{d + 2\sqrt{d}} \cdot z_j + \frac{\sqrt{d}}{d + 2\sqrt{d}} \cdot 0$, which gives $z_j = \frac{d}{d + \sqrt{d}}$.

Therefore, we have $F(T) \geq (\sqrt{d} + 1)k + (\sqrt{d} + 1)(n - k)\frac{d}{d + \sqrt{d}} = \theta$.

Backward Direction. Suppose there is some $T \subseteq V \cup V'$ such that $|T| = k$ and $F(T) \geq \theta$. The goal is to show that there is a vertex cover with size k in G . Observe that the innate opinions of nodes in V are 1; hence, if we are allowed to change the resistance of a node $i \in V$, we should set $\alpha_i = 1$ to maximize the total equilibrium opinion.

We consider the following two cases:

1. Case $T \subseteq V$. i.e., all vertices in T are from V . We prove that T is a vertex cover in G by contradiction.

Assume that there exists an edge $\{i, j\} \in E$ such that both $i, j \notin T$. We derive an upper bound z for the equilibrium opinion of i and j . Observe that for node i , at most $(d - 1)$ of its neighbors are in T . Hence, we have $z \leq \frac{d-1}{d+2\sqrt{d}} \cdot 1 + \frac{\sqrt{d}+1}{d+2\sqrt{d}} \cdot z + \frac{\sqrt{d}}{d+2\sqrt{d}} \cdot 0$, which gives $z_j \leq \frac{d-1}{d+\sqrt{d}-1}$; the important fact is that $\frac{d-1}{d+\sqrt{d}-1} < \frac{d}{d+\sqrt{d}}$.

Observe that for any node in $V \setminus T$, its equilibrium opinion is maximized when all its neighbors in V are in T .

Hence, we have $F(T) \leq \theta'$, where $\theta' := (\sqrt{d} + 1)k + (\sqrt{d} + 1)(n - k - 2)\frac{d}{d+\sqrt{d}} + (\sqrt{d} + 1)2\frac{d-1}{d+\sqrt{d}-1} < \theta$, achieving the desired contradiction.

2. Case $T \setminus V \neq \emptyset$. In this case, we choose T of size k such that $F(T)$ is maximized; if there is more than one such T , we arbitrarily pick one such that $|T \cap V|$ is maximized. For contradiction's sake, we assume that $T \setminus V \neq \emptyset$ and $T \cap V$ is not a vertex cover of G . (We actually just need the weaker condition that $V \setminus T$ is non-empty.) We further consider the following cases:

(i) There is some flexible node u in $T \setminus V$. Suppose the degree-1 node u is connected to $i \in V$. If $i \notin T$, then one can consider $T' := T - u + i$; if $i \in T$, then pick any $j \in V \setminus T$ and consider $T' := T - u + j$.

In either case, it follows that $F(T') \geq F(T)$ and $|T' \cap V| > |T \cap V|$, achieving the desired contradiction.

(ii) There is some stubborn node u in $T \setminus V$. Suppose u is in the clique associated with $i \in T$. Observe that at most k nodes in the clique are in T . Hence, it follows that the equilibrium opinion of any stubborn node is at most $\frac{k+1}{\sigma}$.

Hence, for any $j \in V \setminus T$, its equilibrium opinion satisfies $z_j \leq \frac{d}{d+2\sqrt{d}} \cdot 1 + \frac{\sqrt{d}}{d+2\sqrt{d}} \cdot z_j + \frac{\sqrt{d}}{d+2\sqrt{d}} \cdot \frac{k+1}{\sigma}$. Since $\frac{k+1}{\sigma} \leq \frac{1}{2}$, we have $z_j \leq \frac{d+\frac{1}{2}\sqrt{d}}{d+\sqrt{d}}$.

Next, if $i \notin T$, then set $j := i$; otherwise, just pick any $j \in V \setminus T$; consider $T' := T - u + j$. Note that the equilibrium opinion of stubborn nodes in the clique of u can drop by at most $\frac{k(k+1)}{\sigma} < \frac{n^2}{\sigma}$.

However, the equilibrium opinion of other nodes cannot decrease and that of j increases by at least $\frac{\frac{1}{2}\sqrt{d}}{d+\sqrt{d}} = \frac{n^2}{\sigma}$, by the choice of σ .

Hence, we have the contradiction $F(T') > F(T)$.

This completes the reduction proof for which the resistance parameter is chosen in the range $[0, 1]$. We next show to how to modify the proof for the case in which the resistance value is chosen in the interval $[\epsilon, 1]$ for some sufficiently small $\epsilon > 0$.

The key point is that when we view $F(T)$ as a function of the resistance parameters in the network G' constructed in the reduction, it is a continuous function. Define $\gamma := \theta - \theta' > 0$, where θ and θ' are defined as above.

One can choose $\epsilon > 0$ small enough such that the following holds. In the above proof, if we replace any 0 resistance value with ϵ , then we have (i) if T is a vertex cover of size k in G , then $F(T) \geq \theta - \frac{\gamma}{3}$; (ii) if G does not have a vertex cover of size k , then for any $T \subseteq V \cup V'$ of size k , $F(T) \leq \theta' + \frac{\gamma}{3}$. This completes the proof. \square

7 HEURISTIC ALGORITHMS FOR BUDGETED OPINION SUSCEPTIBILITY PROBLEM

7.1 Marginal Greedy

We propose the Marginal Greedy in Algorithm 3 which has similar framework as the greedy heuristic in [2] but employs the optimistic update strategy to approximate the equilibrium opinion vector $z(\alpha)$.

7.2 Batch Gradient Greedy

We also give a gradient-based heuristic, called **Batch Gradient Greedy (BGG)**, in Algorithm 4. The **while** loop in line 6 employs the optimistic update strategy to approximate the equilibrium opinion vector $z(\alpha)$ (as well as $r(\alpha)$) until it is far apart enough from s to enter the following procedures.

ALGORITHM 3: Marginal Greedy

Input: Innate opinions $s \in [0, 1]^V$; initial resistance vector $\alpha^{(0)}$; budget k ; interaction matrix P ; for each agent $i \in V$, upper u_i and lower l_i bounds for resistance.

Output: The optimal set T of agents with changed resistance and the corresponding resistance vector α with $\forall i \in T : \alpha_i \in \{l_i, u_i\}$ and $\forall i \in V \setminus T : \alpha_i = \alpha_i^{(0)}$.

- 1 (Technical step.) Randomly perturb each coordinate of s slightly.
- 2 Initialize the sum of equilibrium options $f \leftarrow |V|$ and the set of agents $T \leftarrow \emptyset$.
- 3 **for** $j \leftarrow 1$ to k **do**
- 4 Denote $\epsilon_\alpha := \min_{i \in V} \alpha_i^{(j)}$ and $\text{err}(t) := \frac{(1-\epsilon_\alpha)^t}{\epsilon_\alpha}$.
- 5 **for** each $v \in V \setminus T$ **do**
- 6 Set $\alpha^{(j)} \leftarrow \alpha^{(j-1)}$ and $\alpha_v^{(j)} \leftarrow u_v$ to its upper bound (and update ϵ_α).
- 7 Set $z \leftarrow (1, 1, \dots, 1)$ and $t \leftarrow 0$.
- 8 **while** $\exists i \in V : |s_i - z_i| \leq \text{err}(t)$ **do**
- 9 $z \leftarrow As + (I - A)Pz$, where $A = \text{Diag}(\alpha^{(j)})$.
- 10 $t \leftarrow t + 1$.
- 11 Set $L \leftarrow \{i \in T \cup \{v\} : z_i \leq s_i \wedge \alpha_i^{(j)} = u_i\}$ and $J \leftarrow \{i \in T \cup \{v\} : z_i > s_i \wedge \alpha_i^{(j)} = l_i\}$.
- 12 **if** $L \cup J \neq \emptyset$ **then**
- 13 **for** each $i \in L$ **do**
- 14 Set $\alpha_i^{(j)} \leftarrow l_i$ to its lower bound (and update ϵ_α).
- 15 **for** each $i \in J$ **do**
- 16 Set $\alpha_i^{(j)} \leftarrow u_i$ to its upper bound (and update ϵ_α).
- 17 $t \leftarrow 0$.
- 18 **if** $f > \sum_{i \in V} z_i$ **then**
- 19 Set $f \leftarrow \sum_{i \in V} z_i$.
- 20 Update the selected agent $v' \leftarrow v$ and the corresponding resistance vector $\alpha' \leftarrow \alpha^{(j)}$.
- 21 Update the set of selected agents $T \leftarrow T \cup \{v'\}$ and the corresponding resistance vector $\alpha^{(j)} \leftarrow \alpha'$.
- 22 **return** The set of agents T and resistance vector $\alpha^{(k)}$.

Observe that in Line 17, we introduce the batch approach to accelerate the algorithm. When dealing with a large-scale network and a large budget, we can set the batch size proportional to the budget, e.g., 1% of budget, to limit the times to run the outer **while** loop. However, a larger batch size would intuitively lead to a solution with worse quality, because the algorithm needs to commit to selecting more agents in each batch which might turn out to have less desirable partial derivatives later. Our experiment results in Section 9.2 agree with this intuition.

From lines 18 to 24, we consider a measure δ_i for each agent i , that is, the partial derivative times the change of resistance, to decide which agent to include in the batch in order to maximize the decrease of the equilibrium. Since we are using the optimistic update strategy to approximate z and r , we need to estimate its lower $\delta_i^{(l)}$ and upper $\delta_i^{(u)}$ bounds based on Lemmas 5.2 and 7.1.

Then in lines 25–28, we try to pick a subset of agents such that their minimal measure lower bound is greater than the maximal measure upper bound of the rest agents in $V \setminus T$, i.e., to make sure that we select a batch of agents with the greatest measure. Otherwise, we discard the subset and do one more update of r and z until we can find such a subset.

ALGORITHM 4: Batch Gradient Greedy (BGG)

Input: Innate opinions $s \in [0, 1]^V$; initial resistance vector $\alpha^{(0)}$; budget k ; batch size b , interaction matrix P ; for each agent $i \in V$, upper u_i and lower l_i bounds for resistance; precision ρ .

Output: The optimal set T of agents with changed resistance and the corresponding resistance vector α with $\forall i \in T : \alpha_i \in \{l_i, u_i\}$ and $\forall i \in V \setminus T : \alpha_i = \alpha_i^{(0)}$.

- 1 (*Technical step.*) Randomly perturb each coordinate of s slightly.
- 2 Initialize the resistance vector $\alpha \leftarrow \alpha^{(0)}$ and the set of agents $T \leftarrow \emptyset$.
- 3 Denote $\epsilon_\alpha := \min_{i \in V} \alpha_i$ and $\text{err}(t) := \frac{(1-\epsilon_\alpha)^t}{\epsilon_\alpha}$.
- 4 **while** $|T| < k$ **do**
 - 5 Set $z \leftarrow (1, 1, \dots, 1)$; $r \leftarrow (1, 1, \dots, 1)$; $t \leftarrow 0$.
 - 6 **while** $\exists i \in V : |s_i - z_i| \leq \text{err}(t)$ **do**
 - 7 $r \leftarrow \mathbf{1} + P^\top(I - A)r$ and $z \leftarrow As + (I - A)Pz$, where $A = \text{Diag}(\alpha)$.
 - 8 $t \leftarrow t + 1$.
 - 9 Set $L \leftarrow \{i \in T : z_i \leq s_i \wedge \alpha_i = u_i\}$ and $J \leftarrow \{i \in T : z_i > s_i \wedge \alpha_i = l_i\}$.
 - 10 **if** $L \cup J \neq \emptyset$ **then**
 - 11 **for each** $i \in L$ **do**
 - 12 Set $\alpha_i \leftarrow l_i$ to its lower bound (and update ϵ_α).
 - 13 **for each** $i \in J$ **do**
 - 14 Set $\alpha_i \leftarrow u_i$ to its upper bound (and update ϵ_α).
 - 15 $t \leftarrow 0$.
 - 16 Set $\delta^{(u)} \leftarrow (0, 0, \dots, 0)$; $\delta^{(l)} \leftarrow (0, 0, \dots, 0)$; $T' \leftarrow \emptyset$; $b' = \min\{b, k - |T|\}$.
 - 17 **while** $|T'| < b'$ **do**
 - 18 **for each** $i \in V \setminus T$ **do**
 - 19 (Compute upper $d_i^{(u)}$ and lower $d_i^{(l)}$ bound for the partial derivative.)
 - 20 $d_i^{(u)} \leftarrow [r_i + |V| \cdot \text{err}(t)] \cdot \frac{|s_i - z_i| + \text{err}(t)}{1 - \alpha_i}$, and $d_i^{(l)} \leftarrow [r_i - |V| \cdot \text{err}(t)] \cdot \frac{|s_i - z_i| - \text{err}(t)}{1 - \alpha_i}$.
 - 21 **if** $s_i \geq z_i$ **then**
 - 22 $\delta_i^{(u)} \leftarrow d_i^{(u)} \cdot (\alpha_i - l_i)$ and $\delta_i^{(l)} \leftarrow d_i^{(l)} \cdot (\alpha_i - l_i)$.
 - 23 **else**
 - 24 $\delta_i^{(u)} \leftarrow d_i^{(u)} \cdot (u_i - \alpha_i)$ and $\delta_i^{(l)} \leftarrow d_i^{(l)} \cdot (u_i - \alpha_i)$.
 - 25 Pick T' to be the set of b' agents i in $V \setminus T$ with the largest $\delta_i^{(l)}$.
 - 26 **if** $\max_{i \in V \setminus (T \cup T')} \delta_i^{(u)} > \min_{i \in T'} \delta_i^{(l)}$ **then**
 - 27 $r \leftarrow \mathbf{1} + P^\top(I - A)r$ and $z \leftarrow As + (I - A)Pz$, where $A = \text{Diag}(\alpha)$.
 - 28 $t \leftarrow t + 1$ and $T' \leftarrow \emptyset$.
 - 29 **for** $i \in T'$ **do**
 - 30 **if** $s_i \geq z_i$ **then**
 - 31 Set $\alpha_i \leftarrow l_i$ to its lower bound (and update ϵ_α).
 - 32 **else**
 - 33 Set $\alpha_i \leftarrow u_i$ to its upper bound (and update ϵ_α).
 - 34 Update the set of selected agents $T \leftarrow T \cup T'$.
- 35 **The code as Lines 5 to 15 without updating** r .
- 36 **return** *The set of agents* T *and resistance vector* α .

7.3 Approximating the Derivative Vector

In Algorithm 4, we compute the partial derivative vector $d(\alpha) = \nabla f(\alpha) \in \mathbb{R}^V$ according to the equations in the proof of Lemma 4.2, where

$$d_i(\alpha) = \frac{\partial f(\alpha)}{\partial \alpha_i} = \frac{s_i - z_i(\alpha)}{1 - \alpha_i} \cdot \mathbf{1}^\top [I - (I - A)P]^{-1} e_i = \frac{s_i - z_i(\alpha)}{1 - \alpha_i} \cdot r_i(\alpha),$$

where we define $r(\alpha) = [I - P^\top(I - A)]^{-1} \mathbf{1} \in \mathbb{R}^V$.

LEMMA 7.1 (APPROXIMATION ERROR). *Suppose for some $\epsilon > 0$, for all $i \in V$, $\alpha_i \geq \epsilon$. For $t \geq 0$, define $r^{(t)} = \sum_{j=0}^t [P^\top(I - A)]^j \mathbf{1}$; we denote $r(\alpha) = r^{(\infty)}$.*

Then, for every $t \geq 0$, $\|r(\alpha) - r^{(t)}\|_1 \leq n \cdot \frac{(1-\epsilon)^{t+1}}{\epsilon}$.

PROOF. We use the Neumann series $[I - P^\top(I - A)]^{-1} = \sum_{j=0}^{\infty} [P^\top(I - A)]^j$.

We have $\|r(\alpha) - r^{(t)}\|_1 = \mathbf{1}^\top \sum_{j=t+1}^{\infty} [P^\top(I - A)]^j \mathbf{1}$.

We next prove, by induction, that for all $j \geq 0$, $\mathbf{1}^\top [P^\top(I - A)]^j \mathbf{1} \leq n(1 - \epsilon)^j$.

The base case $j = 0$ is trivial.

For the inductive step, assume that for some $j \geq 0$, $\mathbf{1}^\top [P^\top(I - A)]^j \mathbf{1} \leq n(1 - \epsilon)^j$. Since P is a row stochastic matrix, it follows that $\mathbf{1}^\top P^\top = \mathbf{1}^\top$. Hence, $\mathbf{1}^\top [P^\top(I - A)]^{j+1} \mathbf{1} = \mathbf{1}^\top (I - A) [P^\top(I - A)]^j \mathbf{1} \leq (1 - \epsilon) \mathbf{1}^\top [P^\top(I - A)]^j \mathbf{1}$, where the inequality holds because every entry of the row vector $\mathbf{1}^\top [P^\top(I - A)]^j$ is non-negative. The inductive step is completed by using the induction hypothesis.

Finally, we have

$$\|r(\alpha) - r^{(t)}\|_1 = \mathbf{1}^\top \sum_{j=t+1}^{\infty} [P^\top(I - A)]^j \mathbf{1} \leq n \sum_{j=t+1}^{\infty} (1 - \epsilon)^j = n \cdot \frac{(1 - \epsilon)^{t+1}}{\epsilon},$$

as required. \square

8 EXPERIMENTS FOR UNBUDGETED VARIANT

Experimental Setup. Our experiments run on a server with 2.1 GHz Intel Xeon Gold 6152 CPU and 64 GB of main memory. The server is limited to activate at most 24 threads by the administrator. The real network topologies we use in our experiment are shown in Table 1; we interpret each network as an undirected graph. The number n of nodes in the dataset networks ranges from about 1 million to 65 million; in each network, the number m of edges is around $2n-30n$.

Input Generation. For each dataset, we utilize the network topology and generate the input parameters as follows. The innate opinion s_i of each agent i is independently generated uniformly at random from $[0, 1]$. For each edge $\{i, j\}$ in the network, we independently pick w_{ij} uniformly at random from $[0, 1]$; otherwise, we set $w_{ij} = 0$. For $(i, j) \in V \times V$, we normalize $P_{ij} := \frac{w_{ij}}{\sum_{k \in V} w_{ik}}$. From Lemma 5.2, one can see that approximating the equilibrium opinions is more difficult when the resistance is low. However, since we still want to demonstrate that the resistance for each agent can have varied lower and upper bounds, we set the lower bound l_i of each agent i independently such that with probability 0.99, l_i equals 0.001, and with probability 0.01, it is picked uniformly at random from $[0.001, 0.1]$. Similarly, each upper bound u_i is independently selected such that with probability 0.99, it equals to 0.999, and with probability 0.01, it is chosen uniformly at random from $[0.9, 0.999]$.

8.1 Update Strategies Comparison

We compare the following three update strategies described in Section 5: conservative, opportunistic, and optimistic. For the three smaller networks (com-Youtube, com-LiveJournal, and

Table 1. Datasets Information for Different Networks

Name	Number n of Nodes	Number m of Edges	Source
Residence hall	217	2,672	[39]
Twitter	548	3,638	[19]
Hamsterster	1,788	12,476	[39]
Musae-twitch (PT)	1,912	31,299	[40]
Facebook (NIPS)	2,888	2,981	[39]
Advogato	5,042	39,227	[39]
Chess	7,115	55,779	[39]
Pretty Good Privacy	10,680	24,316	[39]
DBLP	12,495	49,563	[39]
Google+	23,613	39,182	[39]
Facebook (WOSN)	63,392	816,831	[39]
Catster	148,826	5,447,464	[39]
com-Youtube	1,134,890	2,987,624	[40]
com-LiveJournal	3,997,962	34,681,189	[40]
LiveJournal	10,690,276	112,307,385	[39]
com-Friendster	65,608,366	1,806,067,135	[40]

LiveJournal), we apply all three update strategies. For the largest network (com-Friendster), we report only the performance of the optimistic update strategy, as the other two update strategies are not efficient enough for such a large dataset.

Experimental Setup. For fair comparison among the update strategies, we always initialize $z = (1, 1, \dots, 1)$. To compare their performances, we plot a curve for each strategy. The curves have a common x -axis, which corresponds to the number of times that the vector z has been updated so far, i.e., the number of times line 5 (in both Algorithms 1 and 2) has been executed. Since line 5 is the most time-consuming part of the algorithms, it will be a suitable common reference. We use the term *iteration* to refer to each time z is updated.

Recall that in our local search method, initially all agents' resistance parameters are set to the upper bound. For each update strategy, the y -axis measures *progress ratio*, which is loosely reflected by the current fraction of agents whose resistance parameters are set (or will definitely later be set) to the lower bound. We will explain formally how the progress ratio for the y -axis is defined for each strategy as follows.

- **Conservative Strategy.** We run Algorithm 1 such that in line 7, L is non-empty only if for all $i \in V$, $|s_i - z_i| > \text{err}(t)$, in which case, we pick L to be the collection of all i 's such that $\alpha_i = u_i$ and $z_i \leq s_i - \text{err}(t)$.

For the y -axis, we plot the ratio of agents i for which currently $\alpha_i = l_i$, or we know definitely that α_i should be switched to l_i , i.e., currently $\alpha_i = u_i$ and $z_i \leq s_i - \text{err}(t)$.

In Algorithm 1, the iterations (referring to each time z is updated) are grouped into *phases*, where a non-empty L in an iteration marks the end of a phase. Observe that at the end of a phase, for all $i \in L$, α_i is set to l_i and t is reset to 0. Hence, in the next iteration, no coordinate α_i is certain to be switched. Hence, the curve has a step-like shape, where each plateau occurs after the end of each phase.

Observe that initially $\epsilon_\alpha = \min_i u_i \geq 0.9$. Hence, it takes very few number of iterations to satisfy $\forall i \in V : |s_i - z_i| > \text{err}(t)$; we call this the phase 0. At the end of the phase 0, we

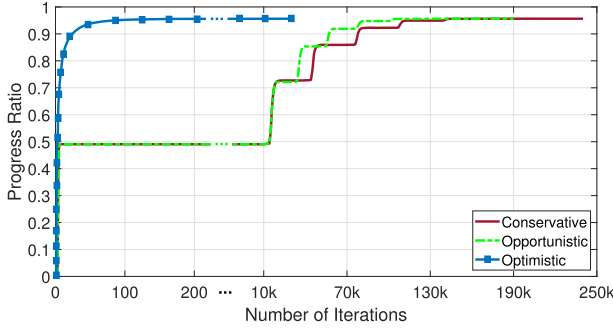


Fig. 2. Update strategies comparison on com-YouTube.

set some $\alpha_i = l_i$ and ϵ_α decreases significantly. Hence, subsequent phases have many more iterations.

Observe that we can stop the iterative process, when for all $i \in V$, $|s_i - z_i| > \text{err}(t)$, but there is no $i \in V$ such that $z_i < s_i$ and $\alpha_i = u_i$. This marks the end of the curve.

In each phase, we pick L of line 7 as the collection of all i 's such that $z_i \leq s_i - \text{err}(t)$ only when $\forall i \in V : |s_i - z_i| > \text{err}(t)$ (otherwise, we pick $L = \emptyset$). Then, we set $\alpha_i = l_i$ for each $i \in L$ and $t = 0$. We call such a phase a *conservative phase*.

- **Opportunistic Strategy.** We run Algorithm 1 similarly as before. Phase 0 is the same as the conservative strategy; we call a phase *conservative*, if it follows the conservative update strategy.

Starting from phase 1, we can perform it in an opportunistic manner as follows. Recall that at the beginning of a phase, t has just been reset to 0. At the t th iteration of that phase, we use $L(t)$ to denote the collection of i 's such that $z_i \leq s_i - \text{err}(t)$. For every 1,000 iterations, we compute an estimate slope(k) := $\frac{|L(1000k)| - |L(1000(k-1))|}{1000}$ of the slope; we keep track k_m of the maximum slope computed so far. After some iteration, if the estimated slope drops below some factor (we use 0.1 in our experiments) of k_m , then we end this phase. Intuitively, each additional iteration flips only a small number of coordinates α_i , and hence, one would like to end this phase. We call such a phase *opportunistic*.

Since typically the total number of phases is around 8, we run phases 1–6 opportunisticly, after which we run the remaining phases conservatively to make sure that all coordinates α_i that need to be changed will be flipped.

As in the conservative update strategy, for the y -axis, we plot the ratio of coordinates α_i that currently $\alpha_i = l_i$, or we know for sure should be switched to l_i , i.e., $\alpha_i = u_i$ and $z_i \leq s_i - \text{err}(t)$.

- **Optimistic Strategy.** We implement Algorithm 2, where in each iteration after z is updated, a coordinate α_i is (re)set to l_i if $z_i < s_i$, and (re)set to u_i if $z_i > s_i$. For the y -axis, we plot the ratio of coordinates that currently take their lower bounds. The curve ends when enough iterations are performed after some coordinate of α is last updated, in order to ensure that the estimate z is close enough to the equilibrium vector according to Lemma 5.2.

Experiment Results. Each of Figures 2–4 shows the plots for the three strategies in the corresponding network (com-YouTube, com-LiveJournal, or LiveJournal). Figure 5 shows the plot of the optimistic strategy in the com-Friendster network, where the other two strategies are not efficient enough for such a large network. As expected, the opportunistic strategy is slightly better than the conservative strategy. From the positions of the plateaus, we can see that the initial opportunistic

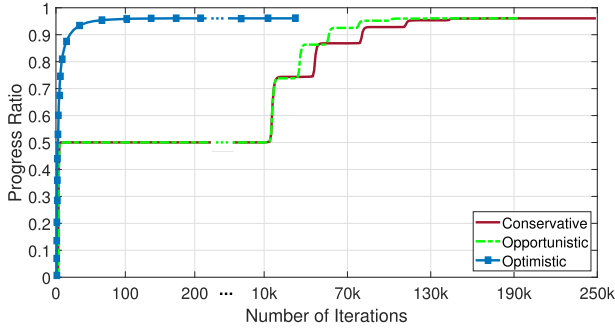


Fig. 3. Update strategies comparison on com-LiveJournal.

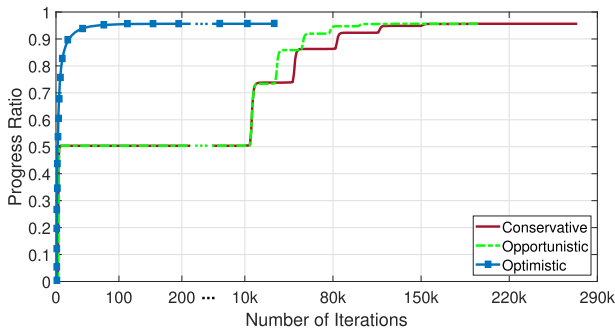


Fig. 4. Update strategies comparison on LiveJournal.

phases end sooner than their conservative counterparts. Hence, overall the opportunistic strategy performs slightly better than the conservative strategy; in increasing sizes of the three tested networks, the numbers of iterations taken by the opportunistic strategy are 79.2%, 77.9%, and 71.5%, respectively, of those taken by the conservative strategy.

On the other hand, the optimistic strategy can achieve the optimal resistance vector with much fewer number of iterations than the other two strategies. In increasing sizes of the three smaller networks, the numbers of iterations taken by the optimistic strategy are only 12.8%, 13.4%, and 12.4%, respectively, of those taken by the conservative strategy. Moreover, the optimistic strategy makes very few mistakes; in increasing sizes of the four networks, the numbers of times coordinates are flipped from lower bounds back to upper bounds are 1, 0, 13, and 168, which are negligible for networks with millions of nodes.

8.2 Running Time with Multiple Threads

We compare the actual running time using different number of threads for the optimistic strategy on only the three smaller networks, since the largest network takes too long using only one thread. Using all 24 available threads, running the optimistic strategy on the com-Friendster network already takes around 50 hours.

The three bar graphs in Figure 6 show the running time (measured in minutes) for running the optimistic strategy with different number of threads on the com-Youtube, com-LiveJournal, and LiveJournal networks. Since updating z (line 5 of Algorithm 2) is the most time-consuming part of the algorithm, the fact that it is readily parallelizable supports the empirical results that using

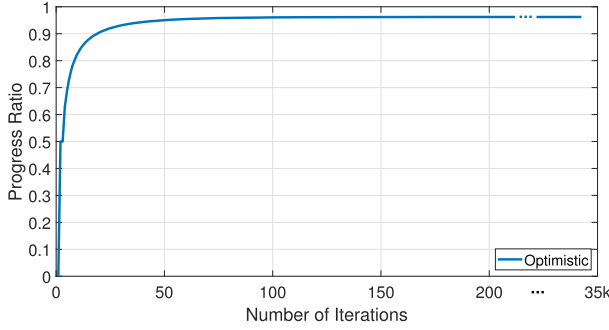


Fig. 5. Optimistic strategy on com-Friendster.

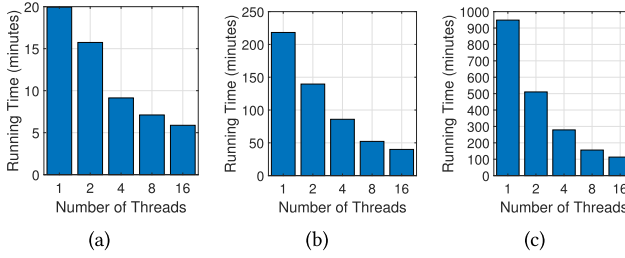


Fig. 6. Running times vs number of threads on (a) com-YouTube, (b) com-LiveJournal, and (c) LiveJournal.

multiple threads can greatly reduce the running time, where the effect is more prominent for larger networks.

9 EXPERIMENTS FOR BUDGETED VARIANT

Experimental Setup. We conduct the experiments on a server with 3.4 GHz Intel(R) Core(TM) i5-3570 CPU and 16 GB of main memory. The server can activate at most four threads. The real network topologies we use are also given in Table 1.

Input Generation. In each instance of the following experiments, we generate a setup of s, P, u, l , and $\alpha^{(0)}$ randomly in a similar way to that in Section 8. Particularly, the initial resistance $\alpha_i^{(0)}$ of each agent i is independently selected uniformly at random from $[l_i, u_i]$ after l_i and u_i are generated.

9.1 Agent Selection Strategies Comparison

As shown in Algorithms 3 and 4, we have two heuristic strategies, Marginal Greedy and BGG, to select new agent into T . We also run experiments on a trivial random node selection strategy, where it selects a new agent uniformly at random from $V \setminus T$ into T until the budget is satisfied.

To compare their performance, we give the average equilibrium opinion z_{avg} employing these three agent selection strategies on small networks in Figure 7. For fair comparison, we use the same setup of s, P, u, l , and $\alpha^{(0)}$ on the same network. In each graph, one curve represents one strategy. The curves share the same x -axis, which corresponds to the ratio of agents selected in T .

Observe that Marginal Greedy and BGG with batch size 1 have almost the same z_{avg} with the change of ratio of agents in T , which implies that they share similar performance. While the

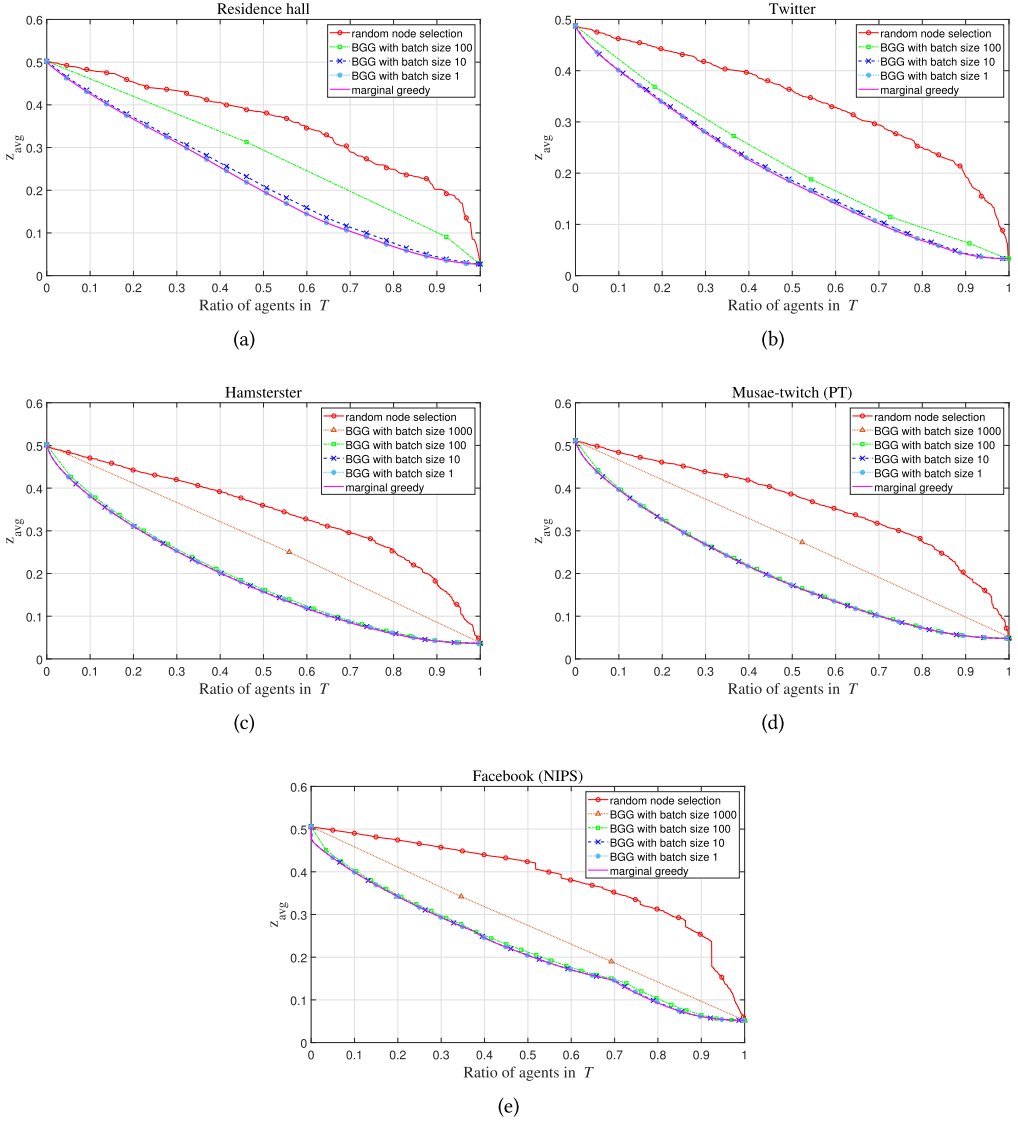


Fig. 7. Average equilibrium opinion on small networks.

random node selection strategy performs the worst among them. We also run experiments using BGG with different constant batch sizes. When the batch size is small enough relative to the number of agents in the network, BGG has similar performance to Marginal Greedy. We will show more results on choosing different batch sizes in the next section.

9.2 Batch Size Comparison

When the budget is proportional to the number n of agents, if we use constant batch size, the number of times to pick a batch would be $O(n)$ which can be too many on large-scale network. One way to solve this problem is to choose the batch size proportional to the number of agents (or the budget), e.g., $1\%n$. Then, the number of times to pick a batch would become $O(1)$.

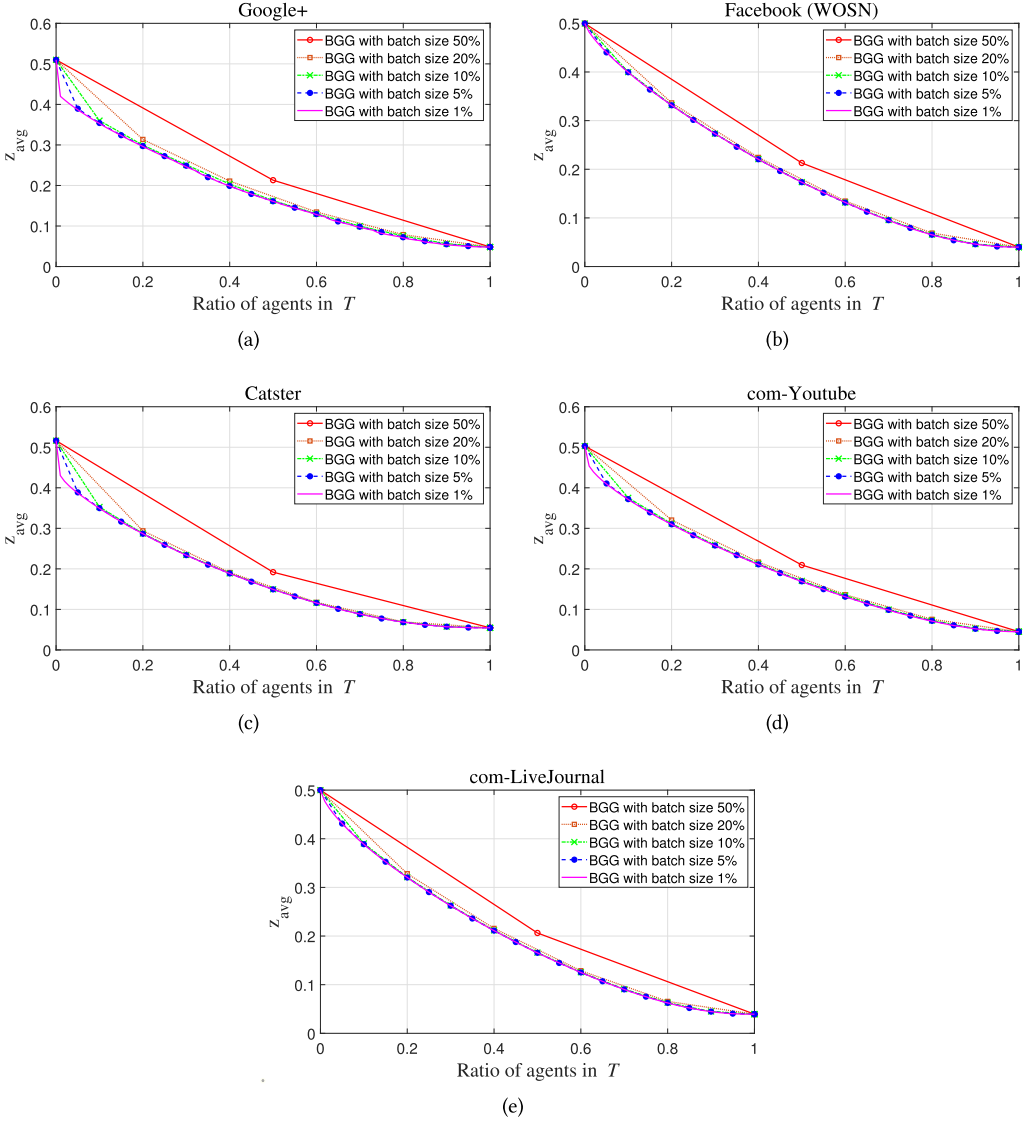


Fig. 8. Average equilibrium opinion on large networks.

Figure 8 gives the average equilibrium opinion when using BGG with different batch sizes on large graphs. We can see that the batch sizes from 1% to 10% get similarly good performance. The result of batch size 20% is slightly worse but acceptable. While the batch size 50% performs the worst among them and is not a good choice. Thus, it is recommended to select 10% of the number of agents (or the budget) or less as the batch size to balance the speed and performance.

9.3 Running Time Comparison

We compare the actual running time of Marginal Greedy and BGG using different methods to compute (or approximate) the equilibrium on only small networks, since the results on larger networks are too time-consuming to collect.

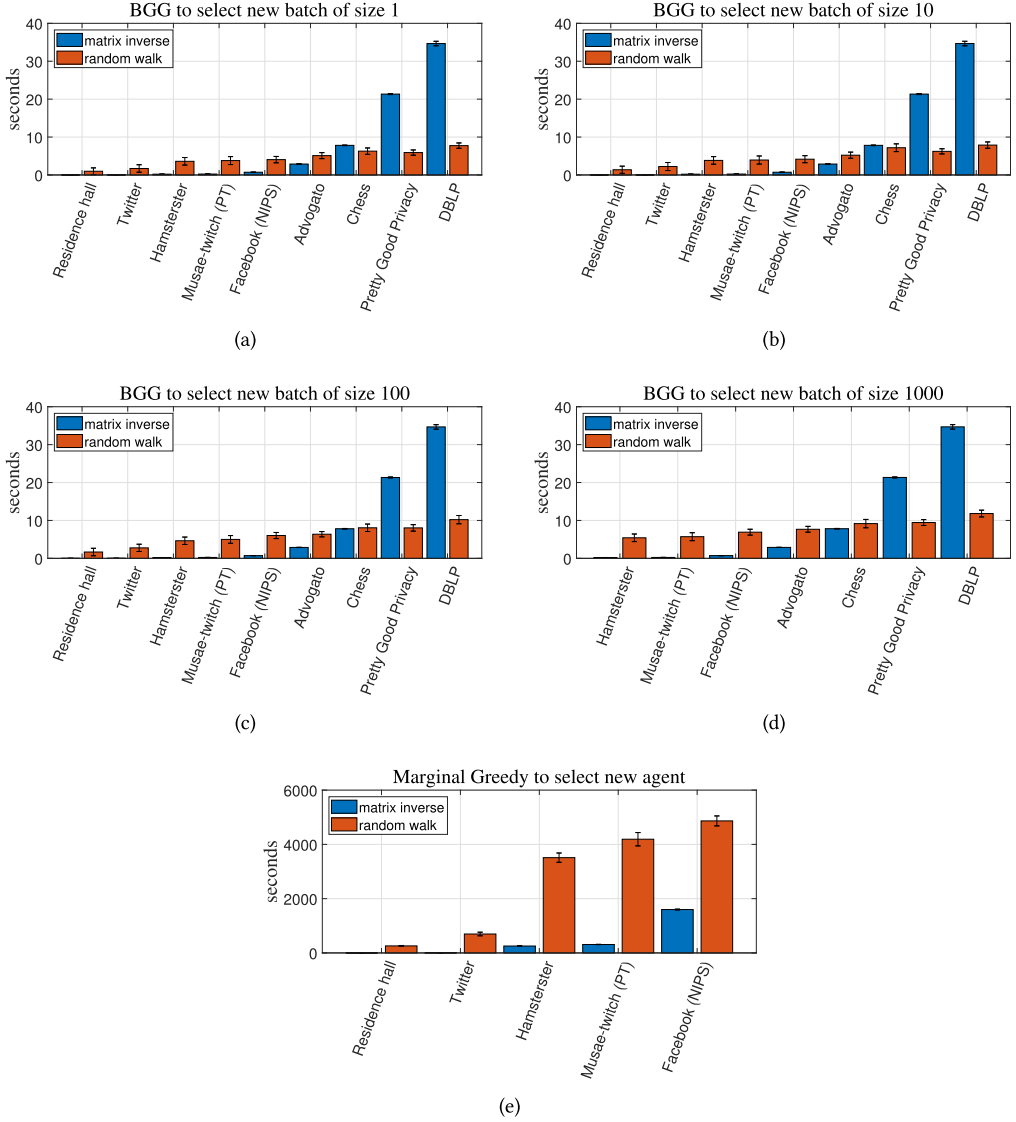


Fig. 9. Average and standard deviation of running time to select new agent or batch.

Figure 9(a)–(d) show the running time of BGG. We compare the efficiency of computing the equilibrium opinion vector $z(\alpha)$ using the matrix inverse $z(\alpha) = [I - (I - A)P]^{-1}As$ and approximating $z(\alpha)$ using the random walk recurrence $z^{(0)} \in [0, 1]^V$ and $z^{(t+1)} := As + (I - A)Pz^{(t)}$, where $A = \text{Diag}(\alpha)$. For each network, we run 30 different random setups of s, P, u, l , and $\alpha^{(0)}$. Then, we record the running time to select the first batch of different batch sizes in each setup and give the average and standard deviation in the bar graphs. Observe that the matrix inverse is faster for networks with less than a few thousands agents. But roughly starting from Chess, the matrix inverse requires more time to select a new batch than the random walk recurrence, which implies that the complexity of matrix inverse is much higher.

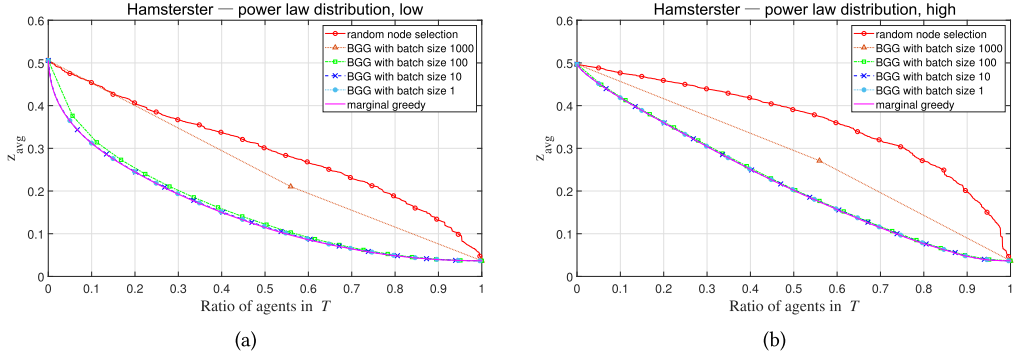


Fig. 10. Average equilibrium opinion on Hamsterster with $\alpha^{(0)}$ generated from power law distribution.

While in Figure 9(e), we show the running time results using Marginal Greedy. We run only one random setup of s, P, u, l , and $\alpha^{(0)}$ on the same network and collect the running time to select the first 30 agents. Then, we report the average and standard deviation of running time to select a new agent. Together with Figure 9(a)–(d), we see that Marginal Greedy is significantly more time-consuming than BGG since Marginal Greedy has to compute the equilibrium opinion vector for adding each candidate before selecting the best one. Considering they have similar performance as shown in Section 9.1, BGG with proper batch size is much more efficient than Marginal Greedy.

9.4 Resistance Generation from Power Law Distribution

We run experiments with the initial resistance vector $\alpha^{(0)}$ generated from the power law distribution instead of the uniform distribution to see how the heuristic algorithms respond to different distributions. Particularly, each coordinate $\alpha_i^{(0)}$ is independently generated from $[l_i, u_i]$ with probability density function proportional to x^{-2} . Observe that the initial resistance vector generated in this way would have most of its coordinates being low values. Figure 10(a) shows the results using the above setup (denoted as power law distribution, low) on Hamsterster where s, P, u, l are unchanged. We also run experiments using a resistance vector with most of its coordinates having high values. For each coordinate $\alpha_i^{(0)}$, we first independently generate a α_i from $[l_i, u_i]$ with probability density function proportional to x^{-2} , and then compute $\alpha_i^{(0)} = u_i - \alpha_i + l_i$. The corresponding results (denoted as power law distribution, high) on Hamsterster are given in Figure 10(b). We can see that Marginal Greedy and BGG with batch size 1 still have similar performance under different resistance distributions.

Figure 11 gives the average equilibrium opinion on Google+ with initial resistance vector $\alpha^{(0)}$ generated from different distributions. We only run experiments with batch size 1% since the performance is good enough. We can conclude that under the same s, P, u, l and budget, the obtained average equilibrium opinion would be higher if the agents of the network tend to have higher resistance.

10 TECHNICAL PROOFS

LEMMA 10.1 (LEMMA 4.1 RESTATED). *Suppose $\alpha \in [0, 1]^V$ has at least one non-zero coordinate and P is an irreducible interaction matrix. Denote $A := \text{Diag}(\alpha)$. Moreover, the inverse $M = [I - (I - A)P]^{-1}$ exists, and every entry of M is positive. Moreover, for every $k \in V$, we have*

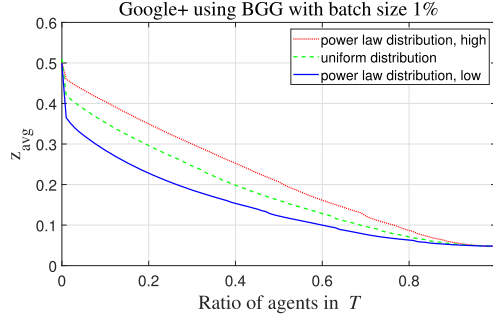


Fig. 11. Average equilibrium opinion on Google+ with $\alpha^{(0)}$ generated from different distributions.

- (1) $(PM)_{kk} = \frac{M_{kk}-1}{1-\alpha_k} > 0$;
- (2) $(PM)_{kj} = \frac{M_{kj}}{1-\alpha_k} > 0$, for each $j \neq k$.

PROOF. Observe that P corresponds to an irreducible random walk. Suppose that $\alpha_{i_0} > 0$. Hence, $(I - A)P$ represents a diluted random walk, where at the beginning of each step, the measure at node i_0 will suffer a factor of $(1 - \alpha_0) < 1$. The irreducibility of the random walk P means that every state is reachable from any state. Hence, starting from any measure vector, eventually the measure at every node will tend to 0. This means that $(I - A)P$ has eigenvalues with magnitude strictly less than 1. Therefore, we can consider the following Neumann series of a matrix:

$$M = [I - (I - A)P]^{-1} = I + \sum_{k=1}^{\infty} [(I - A)P]^k,$$

which implies that the inverse M exists, and every entry of M is positive; in particular, for every $k \in V$, $M_{kk} > 1$.

By the definition of M , we have $[I - (I - A)P]M = I$. We fix some $k \in V$. By considering the (k, k) -the entry, i.e., the dot product between the k th row of $[I - (I - A)P]$ and the k th column of M , we have

$$M_{kk} - \sum_{i \in V} (1 - \alpha_k) P_{ki} M_{ik} = 1.$$

Hence, we have $(PM)_{kk} = \sum_{i \in V} P_{ki} M_{ik} = \frac{M_{kk}-1}{1-\alpha_k}$.

Similarly, for $j \neq k$, by considering the dot product between the k th row of $[I - (I - A)P]$ and the j th column of M , we have

$$M_{kj} - \sum_{i \in V} (1 - \alpha_k) P_{ki} M_{ij} = 0.$$

Hence, we have for $j \neq k$,

$$(PM)_{kj} = \sum_{i \in V} P_{ki} M_{ij} = \frac{M_{kj}}{1 - \alpha_k},$$

as required. \square

LEMMA 10.2 (LEMMA 4.3 RESTATED). Referring to Lemma 4.2. For any $\alpha \in (0, 1)^V$ and any $i \in V$, denote $M = [I - (I - A_{-\{i\}})P]^{-1}$. Then, $\frac{\partial f(\alpha)}{\partial \alpha_i}$ has the same sign in $\{-, 0, +\}$ as $s_i - \sum_{j \neq i} M_{ij} \alpha_j s_j$, which is independent of α_i .

PROOF. Using the Sherman–Morrison formula, we consider

$$\begin{aligned} X &:= [I - (I - A)P]^{-1} = [I - (I - A_{-\{i\}} - \alpha_i \mathbf{e}_i \mathbf{e}_i^T)P]^{-1} \\ &= [I - (I - A_{-\{i\}})P + \alpha_i \mathbf{e}_i \mathbf{e}_i^T P]^{-1} \\ &= M - \frac{\alpha_i}{1 + \alpha_i \mathbf{e}_i^T P M \mathbf{e}_i} M \mathbf{e}_i \mathbf{e}_i^T P M. \end{aligned}$$

Observe that $\mathbf{e}_i^T P M \mathbf{e}_i = (PM)_{ii}$ and $(M \mathbf{e}_i \mathbf{e}_i^T P M)_{ij} = M_{ii}(PM)_{ij}$ for each $j \in V$. Then, by Lemma 4.1, we have

$$\begin{aligned} X_{ii} &= M_{ii} - \frac{\alpha_i M_{ii}(PM)_{ii}}{1 + \alpha_i (PM)_{ii}} = M_{ii} - \frac{\alpha_i M_{ii}(M_{ii} - 1)}{1 + \alpha_i (M_{ii} - 1)} \\ &= \frac{(1 - \alpha_i + \alpha_i M_{ii} - \alpha_i M_{ii} + \alpha_i) M_{ii}}{1 - \alpha_i + \alpha_i M_{ii}} = \frac{M_{ii}}{1 - \alpha_i + \alpha_i M_{ii}}; \end{aligned}$$

and for $j \neq i$,

$$\begin{aligned} X_{ij} &= M_{ij} - \frac{\alpha_i M_{ii}(PM)_{ij}}{1 + \alpha_i (PM)_{ii}} = M_{ij} - \frac{\alpha_i M_{ii} M_{ij}}{1 + \alpha_i (M_{ii} - 1)} \\ &= \frac{(1 - \alpha_i + \alpha_i M_{ii} - \alpha_i M_{ii}) M_{ij}}{1 - \alpha_i + \alpha_i M_{ii}} = \frac{(1 - \alpha_i) M_{ij}}{1 - \alpha_i + \alpha_i M_{ii}}. \end{aligned}$$

By Lemma 4.2, we know $\frac{\partial f(\alpha)}{\partial \alpha_i}$ and $s_i - z_i(\alpha)$ have the same sign in $\{-, 0, +\}$. Recall that $z(\alpha) = [I - (I - A)P]^{-1} A s = X A s$. Applying the above results, we have

$$\begin{aligned} s_i - z_i(\alpha) &= s_i - \sum_{j \in V} X_{ij} \alpha_j s_j \\ &= s_i - \frac{M_{ii} \alpha_i s_i}{1 - \alpha_i + \alpha_i M_{ii}} - \sum_{j \neq i} \frac{(1 - \alpha_i) M_{ij} \alpha_j s_j}{1 - \alpha_i + \alpha_i M_{ii}} \\ &= \frac{s_i(1 - \alpha_i + \alpha_i M_{ii}) - M_{ii} \alpha_i s_i - (1 - \alpha_i) \sum_{j \neq i} M_{ij} \alpha_j s_j}{1 - \alpha_i + \alpha_i M_{ii}} \\ &= \frac{(1 - \alpha_i)(s_i - \sum_{j \neq i} M_{ij} \alpha_j s_j)}{1 - \alpha_i + \alpha_i M_{ii}}. \end{aligned}$$

Since $\alpha_i \in (0, 1)$, we conclude that $\frac{1 - \alpha_i}{1 - \alpha_i + \alpha_i M_{ii}} > 0$. Thus, $\frac{\partial f(\alpha)}{\partial \alpha_i}$, $s_i - z_i(\alpha)$ and $s_i - \sum_{j \neq i} M_{ij} \alpha_j s_j$ have the same sign. \square

LEMMA 10.3 (LEMMA 4.6 RESTATED). For any $i, k \in V$ such that $i \neq k$, let $M = [I - (I - A_{-\{i\}})P]^{-1}$ and $R = [I - (I - A_{-\{i, k\}})P]^{-1}$. Then for any $j \in V$, we have

- (1) $M_{jk} = \frac{R_{jk}}{1 + \alpha_k R_{kk} - \alpha_k}$,
- (2) $M_{jh} = R_{jh} - \frac{\alpha_k R_{jk} R_{kh}}{1 + \alpha_k R_{kk} - \alpha_k}$, for each $h \neq k$.

In particular, the quantity in Lemma 4.3 can be rewritten as follows:

$$s_i - \sum_{j \neq i} M_{ij} \alpha_j s_j = s_i - \sum_{j \neq i, k} R_{ij} \alpha_j s_j - \frac{\alpha_k R_{ik}}{1 + \alpha_k R_{kk} - \alpha_k} \left(s_k - \sum_{j \neq i, k} R_{kj} \alpha_j s_j \right).$$

PROOF. Using the Sherman–Morrison formula, we have

$$\begin{aligned} M &= \left[I - (I - A_{-\{i,k\}})P + \alpha_k \mathbf{e}_k \mathbf{e}_k^T P \right]^{-1} \\ &= R - \frac{\alpha_k}{1 + \alpha_k \mathbf{e}_k^T P R \mathbf{e}_k} R \mathbf{e}_k \mathbf{e}_k^T P R. \end{aligned}$$

We can compute that $\mathbf{e}_k^T P R \mathbf{e}_k = (PR)_{kk}$ and $(R \mathbf{e}_k \mathbf{e}_k^T P R)_{jh} = R_{jk} (PR)_{kh}$ for $j, h \in V$. Then, we have

$$M_{jh} = R_{jh} - \frac{\alpha_k R_{jk} (PR)_{kh}}{1 + \alpha_k (PR)_{kk}}.$$

By Lemma 4.1, we obtain

$$M_{jh} = R_{jh} - \frac{\alpha_k R_{jk} R_{kh}}{1 + \alpha_k R_{kk} - \alpha_k} \text{ for } j, h \in V \text{ and } h \neq k$$

and

$$M_{jk} = R_{jk} - \frac{\alpha_k R_{jk} (R_{kk} - 1)}{1 + \alpha_k R_{kk} - \alpha_k} = \frac{R_{jk}}{1 + \alpha_k R_{kk} - \alpha_k} \text{ for } j \in V.$$

as required. \square

LEMMA 10.4 (LEMMA 4.7 RESTATED). *Suppose $\alpha \in (0, 1)^V$, recall that $A_{-\{i,k\}} := \text{Diag}(\alpha_{-\{i,k\}})$, and P corresponds to an irreducible interaction matrix. For any $i, k \in V$ such that $i \neq k$, let $R = [I - (I - A_{-\{i,k\}})P]^{-1}$, then $R_{ii} = \max_{j \in V} R_{ji}$. Moreover, $R_{ii} = R_{ki}$ if and only if $P_{kk} + P_{ki} = 1$.*

PROOF. We have $[I - (I - A_{-\{i,k\}})P]R = I$. By considering the dot product between each row of $[I - (I - A_{-\{i,k\}})P]$ and column i of R , we have

$$\begin{aligned} R_{ii} - \sum_{h \in V} P_{ih} R_{hi} &= 1, R_{ki} - \sum_{h \in V} P_{kh} R_{hi} = 0, \\ \text{and } R_{ji} - \sum_{h \in V} (1 - \alpha_j) P_{jh} R_{hi} &= 0, \text{ for } j \neq i, k. \end{aligned}$$

After rearranging, we have

$$\begin{aligned} R_{ii} &= 1 + \sum_{h \in V} P_{ih} R_{hi}, R_{ki} = \sum_{h \in V} P_{kh} R_{hi}, \\ \text{and } R_{ji} &= (1 - \alpha_j) \sum_{h \in V} P_{jh} R_{hi}, \text{ for } j \neq i, k. \end{aligned}$$

Now it suffices to show that for $j \neq i, k$, the above R_{ji} cannot be the maximum among them, and $R_{ki} \leq R_{ii}$.

First, we show that R_{ji} cannot be the maximum. Since $\sum_{h \in V} P_{jh} = 1$ and $\alpha_j \in (0, 1)$, we have

$$R_{ji} = (1 - \alpha_j) \sum_{h \in V} P_{jh} R_{hi} \leq (1 - \alpha_j) \max_{h \in V} R_{hi} < \max_{h \in V} R_{hi}.$$

Thus, R_{ji} cannot be the maximum.

Next, we show that $R_{ki} \leq R_{ii}$ by contradiction. Suppose $R_{ki} > R_{ii}$, then R_{ki} is the unique maximum in the i th column of R . Since $\sum_{h \in V} P_{kh} = 1$ and $R_{ki} = \sum_{h \in V} P_{kh} R_{hi}$, it must be the case that $P_{kk} = 1$. This means P corresponds to a random walk with absorbing state k , which contradicts that P is irreducible. Therefore, we have $R_{ki} \leq R_{ii}$, and hence $R_{ii} = \max_{h \in V} R_{hi}$.

Observe that we already know $R_{ji} < R_{ii}$ for $j \neq i, k$, and $R_{ki} = \sum_{h \in V} P_{kh} R_{hi}$. Hence, $R_{ki} = R_{ii}$ implies that $P_{kk} + P_{ki} = 1$.

Conversely, $P_{kk} + P_{ki} = 1$ implies that $R_{ki} = P_{kk} R_{ki} + P_{ki} R_{ii}$. As argued above, we must have $P_{kk} \neq 1$, which implies $R_{ki} = R_{ii}$. \square

11 CONCLUSION AND FUTURE WORK

In this work, we have introduced a novel formulation of social influence that focuses on interventions at the level of susceptibility using a well-established opinion dynamics model. We give a solid theoretical analysis of the unbudgeted variant of the opinion susceptibility problem, and designed scalable local search algorithms that can solve the problem optimally on graphs with millions of nodes. We also prove that the budgeted variant is NP-hard, and provide scalable heuristics that we evaluate experimentally. We believe that our techniques for the unbudgeted variant will lead to insights for the analysis of the budgeted variant of the problem. We leave the task of providing theoretical guarantees for greedy algorithms on the budgeted variant as future work.

Other Models Involving Inertia. We remark that the notion of susceptibility is similar to inertia in other iterative update rules considered in the literature. For instance, Como and Fagnani [14] considered an iterative update rule that replaces the innate opinion s with the previous expressed opinion $z^{(t)}$ in Equation (2) in Section 3. The resulting rule becomes $z^{(t+1)} := P_\alpha z^{(t)}$, where the resistance parameters α are absorbed into the row stochastic matrix $P_\alpha := A + (I - A)P$. In this case, observe that an equilibrium z_α corresponds to an 1-eigenvector, which is unique if P is irreducible and at least one $\alpha_i > 0$. Therefore, the sum of opinions becomes $\mathbf{1}^\top z_\alpha = n\pi_\alpha^\top z^{(0)}$, where $\pi_\alpha^\top = \pi_\alpha^\top P_\alpha$ is the stationary distribution (also known as the centrality vector) of the random walk corresponding to P_α .

It would be interesting to investigate how does this affect the hardness of the influence optimization under this new update rule. We would like to point out that Amelkin and Singh [5] have analyzed how the centrality vector of a graph changes as edges are modified, and they showed that some associated optimization problem is NP-hard.

REFERENCES

- [1] Rediet Abebe, Lada A. Adamic, and Jon M. Kleinberg. 2018. Mitigating overexposure in viral marketing. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*. AAAI, 241–248.
- [2] Rediet Abebe, Jon M. Kleinberg, David C. Parkes, and Charalampos E. Tsourakakis. 2018. Opinion dynamics with varying susceptibility to persuasion. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 1089–1098.
- [3] Daron Acemoglu and Asuman E. Ozdaglar. 2011. Opinion dynamics and learning in social networks. *Dynamic Games and Applications* 1, 1 (2011), 3–49.
- [4] Joseph P. Allen, Maryfrances R. Porter, and F. Christy McFarland. 2006. Leaders and followers in adolescent close friendships: Susceptibility to peer influence as a predictor of risky behavior, friendship instability, and depression. *Development and Psychopathology* 18, 1 (2006), 155–172. DOI : <https://doi.org/10.1017/S09545794060060093>
- [5] Victor Amelkin and Ambuj K. Singh. 2019. Fighting opinion control in social networks via link recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 677–685.
- [6] Solomon E. Asch. 1955. Opinions and social pressure. *Scientific American* 193, 5 (1955), 31–35. DOI : <https://doi.org/10.1038/scientificamerican1155-31>
- [7] Jonah Berger and Chip Heath. 2007. Where consumers diverge from others: Identity signaling and product domains. *Journal of Consumer Research* 34, 2 (2007), 121–134. Retrieved from <http://www.jstor.org/stable/10.1086/519142>
- [8] Srinadh Bhojanapalli, Behnam Neyshabur, and Nati Srebro. 2016. Global optimality of local search for low rank matrix recovery. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*. 3873–3881.
- [9] David Bindel, Jon M. Kleinberg, and Sigal Oren. 2015. How bad is forming your own opinion? *Games and Economic Behavior* 92, C (2015), 248–265.
- [10] T.-H. Hubert Chan, Zhibin Liang, and Mauro Sozio. 2019. Revisiting opinion dynamics with varying susceptibility to persuasion via non-convex local search. In *Proceedings of the Conference on World Wide Web*. ACM, 173–183.
- [11] T.-H. Hubert Chan and Chui Shan Lee. 2021. On the hardness of opinion dynamics optimization with L_1 -budget on varying susceptibility to persuasion. arXiv:2105.04105. Retrieved from <https://arxiv.org/abs/2105.04105>
- [12] Robert B. Cialdini. 1993. *Influence: Science and Practice* (3rd. ed.). HarperCollins College Publishers.
- [13] Robert B. Cialdini. 2001. The science of persuasion. *Scientific American* 284, 2 (2001), 76–81. DOI : <https://doi.org/10.1038/scientificamerican0201-76>

- [14] Giacomo Como and Fabio Fagnani. 2016. From local averaging to emergent global behaviors: The fundamental role of network interconnections. *Systems & Control Letters* 95 (2016), 70–76.
- [15] Ayn E. Crowley and Wayne D. Hoyer. 1994. An integrative framework for understanding two-sided persuasion. *Journal of Consumer Research* 20, 4 (1994), 561–574. Retrieved from <http://www.jstor.org/stable/2489759>.
- [16] Abhimanyu Das, Sreenivas Gollapudi, Arindham Khan, and Renato Paes Leme. 2014. Role of conformity in opinion dynamics in social networks. In *Proceedings of the 2nd ACM Conference on Online Social Networks*. ACM, New York, NY, 25–36. DOI : <https://doi.org/10.1145/2660460.2660479>
- [17] Abhimanyu Das, Sreenivas Gollapudi, and Kamesh Munagala. 2014. Modeling opinion dynamics in social networks. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining*. ACM, New York, NY, 403–412. DOI : <https://doi.org/10.1145/2556195.2559896>
- [18] Abhimanyu Das, Sreenivas Gollapudi, Rina Panigrahy, and Mahyar Salek. 2013. Debiasing social wisdom. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 500–508. DOI : <https://doi.org/10.1145/2487575.2487684>
- [19] Abir De, Sourangshu Bhattacharya, Parantapa Bhattacharya, Niloy Ganguly, and Soumen Chakrabarti. 2019. Learning linear influence models in social networks from transient opinion dynamics. *ACM Transactions on the Web* 13, 3 (2019), 16:1–16:33. DOI : <https://doi.org/10.1145/3343483>
- [20] Morris H. DeGroot. 1974. Reaching a consensus. *Journal of the American Statistical Association* 69, 345 (1974), 118–121. Retrieved from <http://www.jstor.org/stable/2285509>
- [21] T. E. Dielman, Deborah D. Kloska, Sharon L. Leech, John E. Schulenberg, and Jean T. Shope. 1992. Susceptibility to peer pressure as an explanatory variable for the differential effectiveness of an alcohol misuse prevention program in elementary schools. *Journal of School Health* 62, 6 (1992), 233–237.
- [22] Pedro Domingos and Matt Richardson. 2001. Mining the network value of customers. In *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 57–66. DOI : <https://doi.org/10.1145/502512.502525>
- [23] William N. Evans, Wallace E. Oates, and Robert M. Schwab. 1992. Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy* 100, 5 (1992), 966–991.
- [24] B. J. Fogg. 2002. Persuasive technology: Using computers to change what we think and do. *Ubiquity* 2002, December, Article 5 (Dec. 2002), 32 pages. DOI : <https://doi.org/10.1145/764008.763957>
- [25] Marieke L. Franssen, Edith G. Smit, and Peeter W. J. Verlegh. 2015. Strategies and motives for resistance to persuasion: An integrative framework. *Frontiers in Psychology* 6 (Aug. 2015), 1201.
- [26] Noah E. Friedkin and Eugene C. Johnsen. 1999. Social influence networks and opinion change. *Advances in Group Processes* 16, 1 (Jan. 1999), 1–19.
- [27] Rong Ge, Furong Huang, Chi Jin, and Yang Yuan. 2015. Escaping from saddle points—online stochastic gradient for tensor decomposition. In *Proceedings of the 28th Conference on Learning Theory*. Vol. 40, PMLR , 797–842.
- [28] Rong Ge, Jason D. Lee, and Tengyu Ma. 2016. Matrix completion has no spurious local minimum. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*. 2973–2981.
- [29] Javad Ghaderi and R. Srikant. 2013. Opinion dynamics in social networks: A local interaction game with stubborn agents. In *Proceedings of the 2013 American Control Conference*. IEEE, 1982–1987.
- [30] Aristides Gionis, Evimaria Terzi, and Panayiotis Tsaparas. 2013. Opinion maximization in social networks. In *Proceedings of the 13th SIAM International Conference on Data Mining*. SIAM, 387–395.
- [31] Philip Hartman. 1959. On functions representable as a difference of convex functions. *Pacific Journal of Mathematics* 9, 3 (1959), 707–713. Retrieved from <https://projecteuclid.org/443/euclid.pjm/1103039111>.
- [32] Yansong Hu and Christophe Van den Bulte. 2014. Nonmonotonic status effects in new product adoption. *Marketing Science* 33, 4 (July 2014), 509–533. DOI : <https://doi.org/10.1287/mksc.2014.0857>
- [33] Wijnand IJsselstein, Yvonne de Kort, Cees Midden, Berry Eggen, and Elise van den Hoven. 2006. Persuasive technology for human well-being: Setting the scene. In *Proceedings of the 1st International Conference on Persuasive Technology for Human Well-Being*. Springer-Verlag, Berlin, 1–5. Retrieved from <http://dl.acm.org/citation.cfm?id=1986822.1986823>
- [34] Prateek Jain and Purushottam Kar. 2017. Non-convex optimization for machine learning. *Foundations and Trends in Machine Learning* 10, 3–4 (2017), 142–336.
- [35] Chi Jin, Rong Ge, Praneeth Netrapalli, Sham M. Kakade, and Michael I. Jordan. 2017. How to escape saddle points efficiently. In *Proceedings of the 34th International Conference on Machine Learning*. Vol. 70, PMLR, 1724–1732.
- [36] Yogesh V. Joshi, David J. Reibstein, and Z. John Zhang. 2016. Turf Wars: Product line strategies in competitive markets. *Marketing Science* 35, 1 (2016), 128–141.
- [37] Maurits Kaptein, Panos Markopoulos, Boris Ruyter, and Emile Aarts. 2009. Can you be persuaded? Individual differences in susceptibility to persuasion. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I*. Springer-Verlag, Berlin, 115–118. DOI : https://doi.org/10.1007/978-3-642-03655-2_13

- [38] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 137–146. DOI : <https://doi.org/10.1145/956750.956769>
- [39] Jérôme Kunegis. 2017. KONECT: Konec Network Dataset. Retrieved April 2017 from <http://konect.uni-koblenz.de>.
- [40] Jure Leskovec and Andrej Krevl. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. Retrieved June 2014 from <http://snap.stanford.edu/data>.
- [41] William J McGuire. 1964. Inducing resistance to persuasion. Some contemporary approaches. *Advances in Experimental Social Psychology* 1 (1964), 191-229.
- [42] Kevin M. Murphy and Andrei Shleifer. 2004. Persuasion in politics. *American Economic Review* 94, 2 (2004), 435–439.
- [43] Cameron Musco, Christopher Musco, and Charalampos E. Tsourakakis. 2018. Minimizing polarization and disagreement in social networks. In *Proceedings of the 2018 World Wide Web Conference*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 369–378. DOI : <https://doi.org/10.1145/3178876.3186103>
- [44] Jacob R. Neiheisel and Sarah Niebler. 2015. On the limits of persuasion: Campaign ads and the structure of voters’ interpersonal discussion networks. *Political Communication* 32, 3 (2015), 434–452. DOI : <https://doi.org/10.1080/10584609.2014.958258>
- [45] Dohyung Park, Anastasios Kyrillidis, Constantine Caramanis, and Sujay Sanghavi. 2017. Non-square matrix sensing without spurious local minima via the Burer-Monteiro approach. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*. Vol. 54, PMLR, 65–74.
- [46] Laurence Steinberg and Kathryn C. Monahan. 2007. Age differences in resistance to peer influence. *Developmental Psychology* 43, 6 (2007), 1531–1543. DOI : <https://doi.org/10.1037/0012-1649.43.6.1531>
- [47] Ju Sun, Qing Qu, and John Wright. 2016. A geometric analysis of phase retrieval. In *Proceedings of the 2016 IEEE International Symposium on Information Theory*. IEEE, 2379–2383.
- [48] Ju Sun, Qing Qu, and John Wright. 2017. Complete dictionary recovery over the sphere I: Overview and the geometric picture. *IEEE Trans. Information Theory* 63, 2 (2017), 853–884.
- [49] Zakary L. Tormala and Richard E. Petty. 2002. What doesn’t kill me makes me stronger: The effects of resisting persuasion on attitude certainty. *Journal of Personality and Social Psychology* 83, 6 (2002), 1298.
- [50] Zakary L. Tormala and Richard E. Petty. 2004. Source credibility and attitude certainty: A metacognitive analysis of resistance to persuasion. *Journal of Consumer Psychology* 14, 4 (2004), 427–442.
- [51] Melanie Wakefield, Brian Flay, Mark Nichter, and Gary Giovino. 2003. Effects of anti-smoking advertising on youth smoking: A review. *Journal of Health Communication* 8, 3 (2003), 229–247.
- [52] Ercan Yildiz, Asuman Ozdaglar, Daron Acemoglu, Amin Saberi, and Anna Scaglione. 2013. Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation* 1, 4, Article 19 (Dec. 2013), 30 pages. DOI : <https://doi.org/10.1145/2538508>

Received November 2020; revised February 2021; accepted May 2021