

Eliminating Majority Illusion is Easy

Jack Dippel¹, Max Dupré la Tour¹, April Niu¹, Sanjukta Roy², Adrian Vetta¹

¹McGill University

²Indian Statistical Institute, Kolkata

jack.dippel@mail.mcgill.ca

max.duprelatour@mail.mcgill.ca

yuexing.niu@mail.mcgill.ca,

sanjukta@isical.ac.in

adrian.vetta@mcgill.ca

Abstract

Majority illusion is a phenomenon in social networks wherein the decision by the majority of the network is not the same as one’s personal social circle’s majority, leading to an incorrect perception of the majority in a large network. We present polynomial-time algorithms which completely eliminate majority illusion by altering as few connections in the network as possible. Eliminating majority illusion ensures each neighbourhood in the network has at least a $1/2$ -fraction of the majority winner. This result is surprising as partially eliminating majority illusion is NP-hard. We generalize the majority illusion problem to an arbitrary fraction p and show that the problem of ensuring all neighbourhoods in the network contain at least a p -fraction of nodes consistent with a given preference is NP-hard, for nearly all values of p .

Extended version — <https://arxiv.org/html/2407.20187v1>

Introduction

Perceptions are liable to become distorted on a social network. An example of this is the *friendship paradox* (Feld 1991): on average, a person has fewer friends than their friends do.¹ Similarly, on average, a researcher is less productive than their co-authors (Eom and Jo 2014; Benvenuto, Laender, and Alves 2006). This network characteristic, whereby local observations can lead to erroneous conclusions regarding a global state, was dubbed *majority illusion* by Lerman et al. (Lerman, Yan, and Wu 2016). The nomenclature relates to situations where a minority viewpoint is perceived to be the majority viewpoint (Stewart et al. 2019).

One of the simplest manifestation of majority illusion concerns elections. Let’s model a two-party election by a graph $G = (V, E)$ where each node represents a voter and each edge connects two voters that are neighbours (or friends). Furthermore, imagine party preference is given by a colouring function $f : V \rightarrow \{B, R\}$. If there are more blue nodes than red nodes then we say blue is the *majority*. A node $v \in V$ is then under (*majority*) *illusion* if it has strictly more red neighbours than blue neighbours. There exist instances where *every* node is under illusion; see Figure 1.

Network topology can have a big impact on misinformation. Indeed, electoral networks are very susceptible to

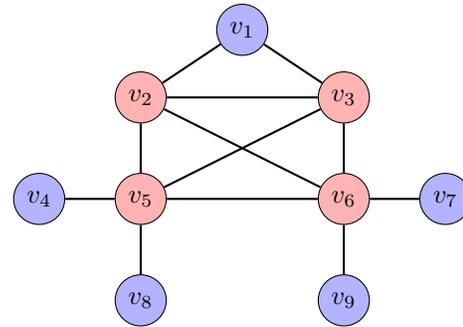


Figure 1: A network in which every node suffers illusion.

majority illusion because their underlying network topology has a social determinant. For example, *homophily*, the tendency for humans to connect with similar people, contributes to the creation of information bubbles and polarization (McPherson, Smith-Lovin, and Cook 2001).

Another topical example of majority illusion concerns vaccinations (Johnson et al. 2020). Here node colours may indicate vaccinated or unvaccinated. If a personal decision to vaccinate oneself is influenced by the actions of people in your local social network, then misinformation may have significant consequences. Moreover, a smaller number of influential high degree nodes may suffice to cause major distortions. An illusion can occur even when there are more than two alternatives and each node selects one alternative. If alternative *blue* is the majority winner, then a node under illusion has strictly less blue neighbors compared to red neighbors (neighbors selecting any of the non-winning alternatives). Indeed, this observation indicates the relevance of illusion to advertising, campaigning, and influence maximization (Becker, D’Angelo, and Ghobadi 2023; Zhou and Zhang 2021). Moreover, the effects of illusion have been studied in many disparate areas, including voting theory (Bara, Lev, and Turrini 2021; Castiglioni, Ferraioli, and Gatti 2020; Doucette et al. 2019; Wilder and Vorobeychik 2018), participatory budgeting and the alloca-

Intelligence (www.aaai.org). All rights reserved.

¹This is a simple consequence of the fact that a higher degree node appears in more neighbourhoods than a lower degree node.

tion of public goods (Kempe, Yu, and Vorobeychik 2020; Yu, Kempe, and Vorobeychik 2021).

These applications highlight the importance of creating social networks that give everyone fair and unbiased information (Johnson et al. 2020). That is the focus of this paper. Specifically, to determine whether or not majority illusion can be eliminated in a network. If so, can we do this optimally in polynomial time?

The Illusion Elimination Problem

To answer these questions we first study a class of *majority illusion elimination problems*. We are given a social network $G = (V, E)$, a non-negative integer k , and a voting function $f : V \rightarrow \{B, R\}$, where B (resp. R) denotes blue (resp. red) nodes and blue is the strict *majority*. Can we eliminate (majority) illusion, for every node, by altering (adding and/or deleting) at most k edges? That is, after the alteration, no node can have a strict majority of red nodes in its neighbourhood.

Observe that, since we may remove edges of G or add non-edges of G , there are three variants of the majority illusion elimination problems where we can only add edges, only remove edges, or both add and remove edges. Formally, this gives the following three decision problems, due to Grandi et al. (Grandi et al. 2023).

Problem 1. Majority Illusion Addition Elimination (MIAE): Given a graph $G = (V, E)$, $f : V \rightarrow \{B, R\}$ where blue is strict majority, and a non-negative integer k , determine if the addition of at most k edges can ensure each node has at least a $\frac{1}{2}$ -fraction of blue nodes in its neighbourhood.

Problem 2. Majority Illusion Removal Elimination (MIRE): Given a graph $G = (V, E)$, $f : V \rightarrow \{B, R\}$ where blue is strict majority, and a non-negative integer k , determine if the subtraction of at most k edges can ensure each node has at least a $\frac{1}{2}$ -fraction of blue nodes in its neighbourhood.

Problem 3. Majority Illusion Elimination (MIE): Given a graph $G = (V, E)$, $f : V \rightarrow \{B, R\}$ where blue is strict majority, and a non-negative integer k , determine if the addition/subtraction of at most k edges can ensure each node has at least a $\frac{1}{2}$ -fraction of blue nodes in its neighbourhood.

An instance of a problem is denoted as (V, E, f, k) . Note that we can analogously define optimization version for each problem where the goal is to minimize the number of edges added and/or removed. We illustrate these three problems using the example we encountered in Figure 1. Figure 2 presents the three corresponding optimal solutions to eliminate majority illusion.

The assumption that the blue party is a strict majority is not necessary. However the assumption of strictness ensures that, for all three problems, it is always feasible to alter the graph to eliminate majority illusion. We remark that (Grandi et al. 2023) left unresolved the computational complexity of these three problems. Instead, rather than *completely* eliminating illusion at every node, they showed it to be NP-complete to *partially* eliminate illusion for a fraction of the nodes. In particular, they studied the following q -partial (*majority*) *illusion elimination problem*: Can at

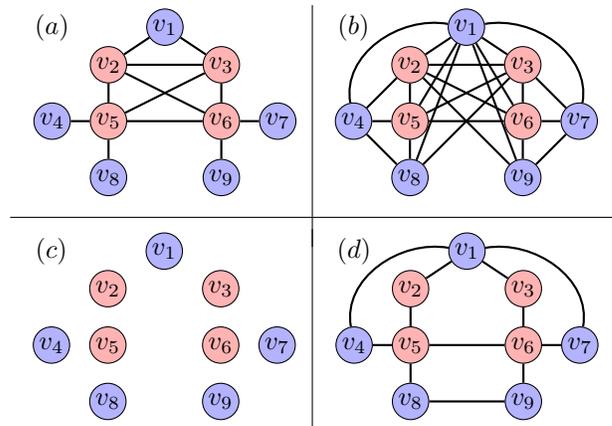


Figure 2: In (a) every node is under majority illusion. An optimal solution to the MIAE problem, adding 11 edges, is shown in (b); an optimal solution to the MIRE problem, requiring the removal of all 12 edges, is shown in (c); an optimal solution to the MIE problem, adding 3 edges and removing 3 edges, is shown in (d).

most q -fraction of the nodes be under illusion after altering (adding and/or deleting) at most k edges in G ? They proved this problem is NP-complete for $q \in (0, 1)$. Observe that the MIAE, MIRE, and MIE problems correspond to the case $q = 0$, that is, where no nodes are allowed to be left under illusion. Therefore, the computational complexity of the main problems of interest remained open and motivated our work.

Our Contributions

We make two contributions regarding illusion elimination in social networks. First, we prove that the majority illusion elimination problem is solvable in **polynomial time** for all three variants (MIAE, MIRE and MIE). Thus, completely eliminating majority illusion in a social network is easy. The result is surprising given, as stated before, that the problem of partially eliminating illusion is hard (Grandi et al. 2023).

In our second contribution, we generalize the problem of eliminating majority illusion and consider the problem of p -illusion. The majority illusion elimination problems assume the blue party has at least $\frac{1}{2}$ the votes. In contrast, suppose we make no assumption on the number of blue voters and instead ask if it is possible to ensure that every voter believes that the blue party has at least a p -fraction of the ballots, where $p \in [0, 1]$ is a rational number. Thus, we define the p -*Illusion problem*: Given a graph $G = (V, E)$, a colouring function $f : V \rightarrow \{B, R\}$, and a rational number $p \in [0, 1]$, can we modify at most k edges such that *every* node has at least p -fraction of blue nodes in its neighbourhood? This question induces three problems for p -Illusion² (analogous to majority illusion elimination), namely, p -IA, p -IR, and p -

²We emphasize the distinction between the p -I problem and the q -partial MIE problem (called q -majority illusion by (Grandi et al. 2023)). In p -I (and p -IA/ p -IR) **every** node must have at least a p -fraction of blue neighbors. In contrast, q -partial MIE only requires the elimination of illusion for a $(1 - q)$ -fraction of the nodes.

I, which as we prove are all polynomial time solvable for $p = \frac{1}{2}$. However, the majority case $p = \frac{1}{2}$ is very special. We prove that p -IA, p -IR, and p -I problems are all NP-complete for any rational $p \in [0, 1]$ except for $p \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$.

An overview of the rest of the paper is as follows. In Section we discuss related works. In Section , we present preliminary results and observations concerning illusion elimination. Our three polynomial time algorithms for MIAE, MIRE, and MIE are then given in Sections , , and , respectively. Finally, in Section we prove the NP-completeness of the p -IA, p -IR and p -I problems, for $p \notin \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$. (Due to space restrictions, missing proofs are deferred to the full version (Dippel et al. 2024).)

Related Work

Several studies have delved into the dynamics of decision-making within social networks and the manipulation of these networks. The article by Grandi et al. (Grandi et al. 2023) is most relevant to us as it directly addresses the phenomenon of majority illusion in social networks. The authors show that identifying the possibility of majority illusion by assigning colors to vertices is NP-hard in restricted graphs and present FPT algorithms with respect to the parameters maximum degree and tree width and a XP algorithm for maximum degree and clique width. Moreover, they show that q -partial majority illusion elimination is in fact W[1]-hard with respect to the number of edge modifications k . In a recent work Fioravantes et al. (2024) study the problem of eliminating majority illusion by altering the color of at most k vertices and show that, unlike our edge editing problem, it is NP-hard and W[2]-hard with respect to k . They present FPT algorithm parameterized by treewidth of the input graph and PTAS for planar graphs. The effect of adding and/or deleting edges in a social network has been investigated in several settings.

In *voting theory*, for example, Castiglioni et al. (Castiglioni, Ferraioli, and Gatti 2020) study election control in social networks through the addition or removal of connections between individuals. Their study proposes methods for strategically manipulating network connections to influence election outcomes. Liu et al. (Liu et al. 2023) introduce a metric called *influence assortment* to quantify the phenomenon “information gerrymandering” in elections, where the strategic distribution of individuals within social networks can influence voting outcomes. Faliszewski et al. (Faliszewski et al. 2022) examine techniques for controlling elections through social influence. By manipulating the spread of information and opinions within social networks, they consider strategies for influencing voter behavior and election outcomes and provide insights into the potential for social influence to shape electoral processes.

In *information diffusion*, by analyzing the impact of link additions on information diffusion and exposure diversity, Becker et al. (Becker, D’Angelo, and Ghobadi 2023) show approximation algorithms to achieve maximum fairness in reaching all nodes or communities. Empirically, they show that adding few edges greedily to maximize spread surpasses all fairness-tailored algorithms in terms of ex-post fairness. Wilder and Vorobeychik (Wilder and Vorobeychik 2018) in-

vestigate opinion diffusion and campaigning strategies on society graphs, how opinions spread through networks, and its effectiveness. They focus on the dynamics of opinion formation and diffusion in social networks, with the aim of understanding influence dynamics and designing effective communication strategies.

Kempe et al. (Kempe, Yu, and Vorobeychik 2020) explore methods for inducing equilibria in networked public goods games by modifying the structure of the network. Through mathematical modeling and simulations, their research investigates strategies for promoting cooperation and achieving desirable outcomes in public goods dilemmas. Their work also proffers insights into the role of network topology in shaping collective behavior.

Preliminaries

We begin by defining some basic notations. Given a graph $G = (V, E)$, let B denote the set of blue nodes and R the set of red nodes in G such that the node set $V(G) = B \cup R$. The edge set $E(R, B)$ denotes the edges in $E(G)$ with one endpoint in R and other in B . For any node $v \in V$, let $N_E(v)$ denote the neighbourhood of v in G . Let $r(v) = |\{u \in R : u \in N_E(v)\}|$ be the number of red nodes in the neighbourhood of v , and $b(v) = |\{u \in B : u \in N_E(v)\}|$ be the number of blue nodes in its neighbourhood.

Definition 1 (p -Illusion). *Let $G = (V, E)$ be a graph such that $\frac{|B|}{|V|} > p$. Then, a node $v \in V$ is under p -illusion if it has strictly less than p -fraction of blue neighbors, i.e., $\frac{b(v)}{|N_E(v)|} < p$. When $p = 1/2$, this is known as majority illusion.*

Next, we describe basic properties that optimal solutions to our three *majority* illusion elimination problems must possess. For the algorithms, we assume blue denotes the majority, i.e., $\frac{|B|}{|V|} > \frac{1}{2}$. For a node $v \in V$, if $r(v) > b(v)$ then we say v is in *deficit*; if $r(v) < b(v)$ then we say v is in *surplus*. So the task is to add/remove as few edges as possible so that no node is in deficit. The solution $E' \subseteq V \times V$ is the set of edges that are modified. That is, the edges $E' \setminus E$ are added to G and the edges $E \cap E'$ are removed from G .

Structural Properties

First we make the trivial observation that it is never of benefit to add an edge between two red nodes.

Observation 1. *In an optimal solution E' to an instance (V, E, f, k) of MIE or MIAE, there are no edges of $E' \setminus E$ between two red nodes.*

Analogously, it is never of benefit to remove an edge between two blue nodes.

Observation 2. *In an optimal solution E' to an instance (V, E, f, k) of MIE or MIRE, there are no edges of $E \cap E'$ between two blue nodes.*

We conclude this section with three more observations for the three problems. The first, for the addition variant, states that for any red node, it is never beneficial to connect it to more new blue neighbours than is absolutely necessary.

Lemma 1. *In an optimal solution E' to an instance (V, E, f, k) of MIAE, there are exactly $\max(r(v) - b(v), 0)$ edges of E' incident to any red node v .*

Similarly, for the removal variant, for any blue node, it is never beneficial to disconnect it from more red neighbours than is absolutely necessary.

Lemma 2. *In an optimal solution E' to an instance (V, E, f, k) of MIRE, there are exactly $\max(r(v) - b(v), 0)$ edges of E' incident to any blue node v .*

Finally, for the general majority illusion elimination problem, the following extends the previous observations.

Lemma 3. *There exists an optimal solution E' to an instance (V, E, f, k) of MIE with no edges of $E' \setminus E$ incident to any red node, and no edges of $E \cap E'$ incident to any blue node.*

Majority Illusion Addition Elimination

We begin by presenting an integer programming formulation of the Majority Illusion Elimination Problem (MIE) restricted to adding edges (MIAE).

A First Attempt

Given a graph $G = (V, E)$ where $V = B \cup R$, we define an auxiliary graph $G^* = (V, E^*)$, called the *availability graph*. Motivated by Observation 1, its edges are the edges in the complement of G excluding those between two red nodes. That is, $E^* = \overline{E}(R, B) \cup \overline{E}(G[B])$ where $\overline{E}(G)$ denotes the non-edges of G . Thus, any optimal solution for MIAE adds edges from E^* . Now let

$$b^*(v) = \begin{cases} \max(r(v) - b(v), 0) & \text{if } v \in R \\ b(v) - r(v) & \text{if } v \in B \end{cases}$$

Next consider the following integer program.

$$\max \sum_{e \in E^*(G^*)} -y_e \quad (\text{IP-1a})$$

$$\text{s.t.} \sum_{u \in N_{E^*}(v) \cap B} y_{uv} = b^*(v) \quad \forall v \in R \quad (1a-1)$$

$$\sum_{u \in N_{E^*}(v) \cap B} -y_{uv} + \sum_{u \in N_{E^*}(v) \cap R} y_{uv} \leq b^*(v) \quad \forall v \in B \quad (1a-2)$$

$$y_e \in \{0, 1\} \quad \forall e \in E^* \quad (1a-3)$$

Theorem 4. *The integer program (IP-1a) is the MIAE problem.*

Proof sketch. The variable y_e is set to 1 if and only if the edge e is added to the solution. Since, we want to minimize the number edges added, the objective function is correct. The left hand side (LHS) of Equation (1a-1) encodes the number of edges added between a red node v and blue nodes B . By Lemma 1 and definition of $b^*(v)$, the equality in (1a-1) holds true for an optimal solution. Next, in the LHS of Equation (1a-2), the first summand is the number of blue neighbours added to the blue node v and the second summand is the number of red neighbours added to a blue

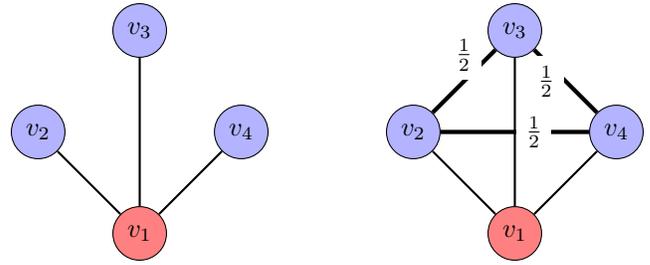


Figure 3: Left: the input graph G . Right: depicts a non-integral optimal solution to the LP relaxation of (IP-1a).

node v . Any feasible solution to the MIAE problem must satisfy the property that the number of red nodes does not exceed the number of blue nodes in the neighbourhood of any blue node. Thus, simple calculations show that (1a-2) is valid. Finally, (1a-3) is valid due to Observation 1. \square

Therefore to solve the MIAE problem, we can simply solve the integer program (IP-1a). But running an integer programming solver is not polynomial time in the worst case, so we need a more efficient method. Observe that (IP-1a) has size polynomial in the size of the graph. Thus, its linear program relaxation can be solved in polynomial time. So, if the optimal solution to the LP relaxation is always integral then, by Theorem 4, we would have an efficient algorithm for the MIAE problem. Unfortunately, this property is not guaranteed: setting $X = B$ and $F = \emptyset$ violates the constraint $1 - y_{v_3 v_4} + 1 - y_{v_3 v_2} + 1 - y_{v_2 v_4} \leq \lfloor \frac{3}{2} \rfloor$ for a non-integral \mathbf{y} for the vertices v_2, v_3 , and v_4 in Figure 3.

A Second Attempt

We desire an integer program that (i) exactly models the MIAE problem, (ii) has an integral LP relaxation, and (iii) the LP relaxation is solvable in polynomial time.

Towards this goal we make two alterations to the integer program (IP-1a). The first is we alter the constraint (1a-2) for each blue node v by adding $|N_{E^*}(v) \cap B|$ to both sides. This transformation is evidently innocuous. For ease of exposition, we define a new function $b'(v)$ to account for this modification as follows:

$$b'(v) = \begin{cases} \max(r(v) - b(v), 0) & \text{if } v \in R \\ |N_{E^*}(v) \cap B| + b(v) - r(v) & \text{if } v \in B \end{cases}$$

Thus, the RHS of Equations (1a-1) and (1a-2) change to $b'(v)$ and LHS of (1a-2) changes. Critically, we also add the following new set of constraints to the integer program. For each $X \subseteq B \cup R$ and for each $F \subseteq \delta_{E^*}(X)$ where $\delta_{E^*}(X)$ is the set of edges in E^* that have one endpoint in X :

$$\sum_{\substack{e: e \in E^*(G^*[B]) \\ e \in E^*(G^*[X]) \cup F}} (1 - y_e) + \sum_{\substack{e: e \in E^*(R, B) \\ e \in E^*(G^*[X]) \cup F}} y_e \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b'(v) + |F| \right) \right\rfloor \quad (1b-4)$$

We will show in Theorem 5 that these new constraints are satisfied by any feasible solution to the MIAE problem.

Moreover, we will prove that adding this set of new constraints will eliminate all non-integral solutions.

Putting all this together our new integer program is

$$\begin{aligned}
\max \quad & \sum_{e \in E^*(G^*)} -y_e & \text{(IP-1b)} \\
\text{s.t.} \quad & \sum_{u \in N_{E^*}(v) \cap B} y_{uv} = b'(v) & \forall v \in R \\
& \sum_{u \in N_{E^*}(v) \cap B} 1 - y_{uv} + \sum_{u \in N_{E^*}(v) \cap R} y_{uv} \leq b'(v) & \forall v \in B \\
& \sum_{\substack{e: e \in E^*(G^*[B]) \\ e \in E^*(G^*[X]) \cup F}} (1 - y_e) + \sum_{\substack{e: e \in E^*(R, B) \\ e \in E^*(G^*[X]) \cup F}} y_e \\
& \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b'(v) + |F| \right) \right\rfloor & \forall X \subseteq B \cup R \\
& & \forall F \subseteq \delta_{E^*}(X) \\
& y_e \in \{0, 1\} & \forall e \in E^*
\end{aligned}$$

Our first step now is to verify that (IP-1b) solves the MIAE problem.

Theorem 5. *The integer program (IP-1b) is the MIAE problem.*

Again, Theorem 5 implies that we can solve the MIAE problem by solving the integer program (IP-1b). We will do this efficiently by solving its linear program relaxation. To achieve this successfully, we must show that the LP relaxation is integral and also that it can be solved in polynomial time despite the fact that there are an exponential number of constraints of type (1b-4). Towards this, we modify (IP-1b) to show it is total dual integral.

A TDI System

For ease of exposition, we begin by imposing a change of variables. Specifically, let

$$x_e = \begin{cases} y_e & \text{if } e \in E^*(R, B) \\ 1 - y_e & \text{if } e \in E^*(G^*[B]) \end{cases}$$

Then LP relaxation of (IP-1b) simplifies to

$$\begin{aligned}
\max \quad & \sum_{e \in E^*(G[B])} x_e - \sum_{e \in E^*(R, B)} x_e - |E^*(G[B])| \\
\text{s.t.} \quad & \sum_{e: e \in \delta_{E^*}(v)} x_e = b'(v) & \forall v \in R \\
& \sum_{e: e \in \delta_{E^*}(v)} x_e \leq b'(v) & \forall v \in B \\
& \sum_{e: e \in E^*(G^*[X])} x_e + \sum_{e: e \in F} x_e \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b'(v) + |F| \right) \right\rfloor \\
& & \forall X \subseteq B \cup R, \forall F \subseteq \delta_{E^*}(X) \\
& 0 \leq x_e \leq 1 & \forall e \in E^*
\end{aligned}$$

Our goal now is to prove that (LP-1c) produces an integral solution. To prove that the system of linear equations is integral, we use the concept of total dual integrality (TDI).

Definition 2. *Let $\mathbf{Ax} \leq \mathbf{b}$ be a linear system where \mathbf{A}, \mathbf{b} have rational entries. The linear system is **Total Dual Integral (TDI)** if for any integer valued \mathbf{c} such that the linear program $\max\{\mathbf{c}^T \mathbf{x} : \mathbf{Ax} \leq \mathbf{b}\}$ has an optimal solution, there is an integer optimal solution to the dual.*

A proof that a linear program is TDI allows us to use the following result to show that it is integral.

Theorem 6. *(Edmonds and Giles 1977) If a linear system $\mathbf{Ax} \leq \mathbf{b}$ is TDI, and \mathbf{b} is integral, then $\mathbf{Ax} \leq \mathbf{b}$ is integral.*

We will prove that (LP-1c) is total dual integral (TDI) to show its integrality. To do this, we turn our attention to simple b -matchings.

Theorem 7. *(Schrijver 2003) The following linear system (representing a simple b' -matching polytope in the graph G^*) is TDI:*

$$\begin{aligned}
& \sum_{e: e \in \delta(v)} x_e \leq b'(v) & \forall v \in B \cup R \\
& \sum_{e: e \in E^*(G^*[X])} x_e + \sum_{e: e \in F} x_e \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b'(v) + |F| \right) \right\rfloor \\
& & \forall X \subseteq B \cup R, \forall F \subseteq \delta_{E^*}(X) \\
& 0 \leq x_e \leq 1 & \forall e \in E^*(G^*)
\end{aligned}$$

(b' -matching LP)

This TDI system is very similar to our system. We require one more result.

Theorem 8. *(Cook 1983) Let a linear system $\mathbf{Ax} \leq \mathbf{b}$ be TDI and let $\mathbf{A}'\mathbf{x} \leq \mathbf{b}'$ denote a linear system obtained by adding inequality $-\mathbf{a}^T \mathbf{x} \leq -\beta$ to $\mathbf{Ax} \leq \mathbf{b}$, where $\mathbf{a}^T \mathbf{x} \leq \beta$ is an inequality in $\mathbf{Ax} \leq \mathbf{b}$. Then $\mathbf{A}'\mathbf{x} \leq \mathbf{b}'$ is TDI.*

Corollary 9. *The linear system of (LP-1b) is TDI.*

Thus, using Theorem 6 we get the following.

Corollary 10. *The linear program (LP-1b) is integral.*

A Polynomial Time Algorithm

We showed that (LP-1b) is integral. However, it has exponentially many constraints. Thus, to show it can be solved in polynomial time, we apply a separation oracle.

Definition 3 (SEPARATION PROBLEM). *Given a convex set $P \subseteq \mathbb{R}^n$ and a vector $\mathbf{y} \in \mathbb{Q}^n$, either determine that $\mathbf{y} \in P$ or find a vector $\mathbf{d} \in \mathbb{Q}^n$ such that $\mathbf{d}^T \mathbf{x} < \mathbf{d}^T \mathbf{y}$ for all $\mathbf{x} \in P$.*

The separation problem for the b -matching polytope can be solved in polynomial time.

Theorem 11. *(Padberg and Rao 1982; Letchford, Reinelt, and Theis 2008) For undirected graphs $G, \mathbf{w} : E(G) \rightarrow \mathbb{N} \cup \{\infty\}$ and $\mathbf{b} : V(G) \rightarrow \mathbb{N}$, the SEPARATION PROBLEM for the \mathbf{b} -matching polytope of (G, \mathbf{w}) can be solved in time $O(n^4)$ where n is the number of nodes in G .*

We show analogous result for (LP-1b).

Theorem 12. *The SEPARATION PROBLEM for (LP-1b) can be solved in time $O(n^4)$ where n is number of nodes.*

Using the next result, we can finally show a polynomial time algorithm for MIAE and consequently, for $\frac{1}{2}$ -IA.

Theorem 13. (Grötschel, Lovász, and Schrijver 1988) Let $n \in \mathbb{N}$ and $\mathbf{c} \in \mathbb{Q}^n$. Let $P \subseteq \mathbb{R}^n$ be a rational polytope, and let $\mathbf{x}_0 \in \mathbb{Q}^n$ be a point in the interior of P . Let $T \in \mathbb{N}$ such that $\text{size}(\mathbf{x}_0) \leq \log T$ and $\text{size}(\mathbf{x}) \leq \log T$ for all vertices \mathbf{x} of P . Given $n, \mathbf{c}, \mathbf{x}_0, T$, and a polynomial-time oracle for the SEPARATION PROBLEM for P , a node \mathbf{x}^* of P attaining $\max\{\mathbf{c}^T \mathbf{x} : \mathbf{x} \in P\}$ can be found in time polynomial in $n, \log T$, and $\text{size}(\mathbf{c})$.

Corollary 14. MIAE and the $\frac{1}{2}$ -ILLUSION ADDITION problem can be solved in polynomial time.

Remark 1. If there are more than two color of nodes and blue has strict majority, then we can identify all the other colors as red. Thus, our algorithms will work in such cases.

Majority Illusion Removal Elimination

The Majority Illusion Removal Elimination (MIRE) Problem is the Majority Illusion Elimination (MIE) Problem restricted to removing edges. Given $G = (V, E)$ where $V = B \cup R$, we define an auxiliary graph $G^* = (V, E^*)$, called the *removable graph*. Its edges are the edges in G excluding those between two blue nodes. That is, $E^* = E(R, B) \cup E(G[R])$. Now let

$$b^*(v) = \begin{cases} \max(r(v) - b(v), 0) & \text{if } v \in B \\ b(v) - r(v) & \text{if } v \in R \end{cases}$$

Consider the following integer program analogous to MIAE.

$$\max \sum_{e \in E^*(G^*)} -y_e \quad (\text{IP-2})$$

$$\text{s.t.} \quad \sum_{u \in N_{E^*}(v) \cap R} y_{uv} = b^*(v) \quad \forall v \in B \quad (2-1)$$

$$\sum_{u \in N_{E^*}(v) \cap R} -y_{uv} + \sum_{u \in N_{E^*}(v) \cap B} y_{uv} \leq b^*(v) \quad \forall v \in R \quad (2-2)$$

$$y_e \in \{0, 1\} \quad \forall e \in E^* \quad (2-3)$$

Theorem 15. The integer program (IP-2) is the MIRE problem.

As one might expect, our strategy for solving the MIRE problem will mirror the approach for the MIAE problem. We will modify (IP-2) to create a linear program (LP-2) which is symmetric to (LP-1b). In this case, we define function $b^i(v)$ to account for adding $|N_{E^*}(v) \cap R|$:

$$b^i(v) = \begin{cases} \max(r(v) - b(v), 0) & \text{if } v \in B \\ |N_{E^*}(v) \cap R| + b(v) - r(v) & \text{if } v \in R \end{cases}$$

Finally, we add an analogous set of new constraints to the integer program.

$$\begin{aligned} & \sum_{\substack{e: e \in E^*(G^*[R]) \\ e \in E^*(G^*[X]) \cup F}} (1 - y_e) + \sum_{\substack{e: e \in E^*(R, B) \\ e \in E^*(G^*[X]) \cup F}} y_e \\ & \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b^i(v) + |F| \right) \right\rfloor \quad \begin{array}{l} \forall X \subseteq B \cup R \\ \forall F \subseteq \delta_{E^*}(X) \end{array} \end{aligned}$$

Finally, we get the following LP relaxation.

$$\begin{aligned} \max \quad & \sum_{e \in E^*(G^*)} -y_e \quad (\text{LP-2}) \\ \text{s.t.} \quad & \sum_{u \in N_{E^*}(v) \cap R} y_{uv} = b^i(v) \quad \forall v \in B \\ & \sum_{u \in N_{E^*}(v) \cap R} (1 - y_{uv}) + \sum_{u \in N_{E^*}(v) \cap B} y_{uv} \leq b^i(v) \quad \forall v \in R \\ & \sum_{\substack{e: e \in E^*(G^*[R]) \\ e \in E^*(G^*[X]) \cup F}} (1 - y_e) + \sum_{\substack{e: e \in E^*(R, B) \\ e \in E^*(G^*[X]) \cup F}} y_e \\ & \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b^i(v) + |F| \right) \right\rfloor \quad \begin{array}{l} \forall X \subseteq B \cup R \\ \forall F \subseteq \delta_{E^*}(X) \end{array} \\ & 0 \leq y_e \leq 1 \quad \forall e \in E^* \end{aligned}$$

Corollary 16. The linear program (LP-2) is integral.

Corollary 17. The SEPARATION PROBLEM for linear program (LP-2) can be solved in $O(n^4)$ time.

Corollary 18. MIRE and the $\frac{1}{2}$ -ILLUSION REMOVAL problem can be solved in polynomial time.

Majority Illusion Elimination

Finally, we arrive at the general Majority Illusion Elimination (MIE) problem. Given $G = (V, E)$, to solve the general problem, we will create new graphs $G^a = (R \cup B, E^a)$ and $G^r = (R \cup B, E^r)$. Specifically:

- Create G^a from G by removing all edges of $G[R]$.
- Create G^r from G by adding all edges of the complement of $G[B]$. The remarkable fact now is that an optimal solution to the general problem on $G = (V, E)$ can be reduced to an optimal solution to the MIAE problem on G^a plus an optimal solution to the MIRE problem on G^r .

Theorem 19. There is an optimal solution E^* to the MIE problem on G composed of the union of the optimal solution E^a to the MIAE problem on G^a and the optimal solution E^r to the MIRE problem on G^r .

Proof idea. Observe that in G^r the blue nodes form a clique. Consequently, no blue node is under illusion, and only red nodes can be under illusion. Therefore, any optimal solution E^r to MIRE of G^r will be a subset of the edges removed in some optimal solution E^* to MIE on G . Similarly, in G^a the red nodes form an independent set. Consequently, no red node is under illusion, and only blue nodes can be under illusion. Therefore, the only edges added in an optimal solution E^a to the MIAE problem will be between two blue nodes. Furthermore, all the red (or blue) nodes in G^r (resp. in G^a) are under exactly the same illusion as the red (resp. blue) nodes in G . It can be shown that majority illusion in G can be eliminated using these two solutions E^a and E^r , and it is optimal. \square

Theorem 19 implies we can solve the MIE problem in polynomial time. We reduce it to a MIAE problem and a

MIRE problem, both of which can be solved in polynomial time by the methods developed in Sections and .

However, we can exploit Theorem 19 to solve the MIE directly using a single linear program. This is because the proof works by essentially reducing the problem to two independent b -matching problems within the same graph! In particular, take $G = (B \cup R, E)$ and define functions

$$\begin{aligned} b^a(v) &= |N_{E^*}(v) \cap B| + b(v) - r(v) & \forall v \in B \\ b^r(v) &= |N_{E^*}(v) \cap R| + b(v) - r(v) & \forall v \in R \end{aligned}$$

Our linear program is then

$$\max \sum_{e \in E^*(G^*)} -y_e \quad (\text{LP-3})$$

$$\text{s.t.} \quad \sum_{u \in N_{E^*}(v) \cap B} (1 - y_{uv}) \leq b^a(v) \quad v \in B(G^*) \quad (3-1)$$

$$\sum_{\substack{e: e \in E^*(G^*[B]) \\ e \in E^*(G^*[X]) \cup F}} (1 - y_e) \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b^a(v) + |F| \right) \right\rfloor \quad \forall X \subseteq B, \forall F \subseteq \delta_{E^*}(X) \quad (3-2)$$

$$\sum_{u \in N_{E^*}(v) \cap R} (1 - y_{uv}) \leq b^r(v) \quad v \in R(G^*) \quad (3-3)$$

$$\sum_{\substack{e: e \in E^*(G^*[R]) \\ e \in E^*(G^*[X]) \cup F}} (1 - y_e) \leq \left\lfloor \frac{1}{2} \left(\sum_{v \in X} b^r(v) + |F| \right) \right\rfloor \quad \forall X \subseteq R, \forall F \subseteq \delta_{E^*}(X) \quad (3-4)$$

$$0 \leq y_e \leq 1 \quad e \in E^*(G^*)$$

We show the following to finally obtain our desired result.

Corollary 20. *The linear program (LP-3) is integral.*

Corollary 21. *The SEPARATION PROBLEM for linear program (LP-3) can be solved in $O(n^4)$ time.*

Corollary 22. *MIE and the $\frac{1}{2}$ -ILLUSION problem can be solved in polynomial time.*

Hardness

In this section we present variants of the MIE, MIAE and MIRE problems.

Problem 4. *Let $p \in [0, 1], p \in \mathcal{Q}$. We define the p -IA problem as follows: Given a graph $G = (V, E)$, $f : V \rightarrow \{B, R\}$, and integer k , determine if the addition of at most k edges can ensure each node has at least a p -fraction of blue nodes in its neighbourhood.*

Problem 5. *Let $p \in [0, 1], p \in \mathcal{Q}$. We define the p -IR problem as follows: Given a graph $G = (V, E)$, $f : V \rightarrow \{B, R\}$, and integer k , determine if the subtraction of at most k edges can ensure each node has at least a p -fraction of blue nodes in its neighbourhood.*

Problem 6. *Let $p \in [0, 1], p \in \mathcal{Q}$. We define the p -I problem as follows: Given a graph $G = (V, E)$, $f : V \rightarrow \{B, R\}$, and integer k , determine if the addition/subtraction of at most k edges can ensure each node has at least a p -fraction of blue nodes in its neighbourhood.*

For $p = \frac{1}{2}$, if we add the requirement that blue be the strict majority, these extended problems correspond to the original MIAE, MIRE and MIE problems for which we have provided polynomial-time algorithms in the previous sections of this paper. Our algorithms can be extended to handle the case when blue is not the true majority, showing that the problems are solvable in polynomial time when $p = \frac{1}{2}$. Regardless, to show NP-hardness for other values of p , we provide a polynomial time reduction from the following NP-complete, satisfiability problem.

Definition 4. *XSAT for l -CNF $_+^l$ is the problem of determining whether a boolean formula in conjunctive normal form composed of n variables appearing exactly l times as positive literals within clauses of size exactly l has an assignment which evaluates to TRUE such that each clause has exactly one variable assigned TRUE.*

Theorem 23. (Schmidt 2011) *XSAT for l -CNF $_+^l$ is NP-complete for any $l \geq 3$*

Theorem 24. *The p -IA problem and p -I problem are NP-hard for $p \in \mathbb{Q}, 0 < p < \frac{1}{2}, p \neq \frac{1}{3}$.*

Theorem 25. *The p -IR problem and p -I problem is NP-hard for $p \in \mathbb{Q}, \frac{1}{2} < p < 1, p \neq \frac{2}{3}$.*

Theorem 26. *The p -IA problem is NP-hard for $p \in \mathbb{Q}, \frac{1}{2} < p < 1, p \neq \frac{2}{3}$.*

Theorem 27. *The p -IR problem is NP-hard for $p \in \mathbb{Q}, 0 < p < \frac{1}{2}, p \neq \frac{1}{3}$.*

Conclusion

We've settled the q -ILLUSION ELIMINATION problems studied in (Grandi et al. 2023) for the special $q = 0$ case. While the previous paper showed hardness for $q = (0, 1)$ and a trivial solution for $q = 1$, we present here the surprising result that fully determining and eliminating illusion can be solved in polynomial time. The three problems, MIAE, MIRE, and MIE, have simple linear programs with integer solutions.

We also define a natural extension to all three problems. For any $p \in [0, 1]$, the p -IA, p -IR, and p -I problems ask if it is possible to alter a network to ensure the neighbourhood of every node has a p -fraction of blue nodes by adding edges, removing edges, or both respectively. These problems reduce to MIAE, MIRE and MIE when $p = \frac{1}{2}$, and just as in those problems, we ask that the total number of blue nodes be such that the problems are solvable regardless of the initial network given. We show that these problems are hard for $p \in (0, \frac{1}{3}) \cup (\frac{1}{3}, \frac{1}{2}) \cup (\frac{1}{2}, \frac{2}{3}) \cup (\frac{2}{3}, 1)$.

In terms of future work, our p -IA, p -IR, and p -I problems have two p -values for which there is no polynomial time algorithm nor a hardness result: $p = \frac{1}{3}$ and $p = \frac{2}{3}$. Interestingly, if we add edge weights to the linear programs used to solve MIAE, MIRE and MIE, we can create linear programs to solve the p -IA, p -IR, and p -I problems. All values of p except for $p \in \{0, \frac{1}{2}, 1\}$ result in a program that is no longer totally dual integral. However, $p \in \{\frac{1}{3}, \frac{2}{3}\}$ are the only other values for which the program is half integral. Therefore, it is entirely possible that these problems are solvable in polynomial time for $p \in \{\frac{1}{3}, \frac{2}{3}\}$; this would be an intriguing result.

Acknowledgments

AV is supported by grant number NSERC 2022-04191.

References

- Bara, J.; Lev, O.; and Turrini, P. 2021. Predicting voting outcomes in presence of communities. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021)*, 151–159.
- Becker, R.; D’Angelo, G.; and Ghobadi, S. 2023. Improving fairness in information exposure by adding links. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 14119–14126.
- Benevenuto, F.; Laender, A.; and Alves, B. 2006. The h -index paradox: your coauthors have a higher h -index than you do. *Scientometrics*, 106: 469–474.
- Castiglioni, M.; Ferraioli, D.; and Gatti, N. 2020. Election control in social networks via edge addition or removal. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 1878–1885.
- Cook, W. 1983. Operations that preserve total dual integrality. *Operations Research Letters*, 2(1): 31–35.
- Dippel, J.; la Tour, M. D.; Niu, A.; Roy, S.; and Vetta, A. 2024. Eliminating Majority Illusion is Easy. *CoRR*, abs/2407.20187.
- Doucette, J.; Tsang, A.; Hosseini, H.; Larson, K.; and Cohen, R. 2019. Inferring true voting outcomes in homophilic social networks. *Autonomous Agents and Multi-Agent Systems*, 33: 298–329.
- Edmonds, J.; and Giles, R. 1977. A Min-Max Relation for Submodular Functions on Graphs. In Hammer, P.; Johnson, E.; Korte, B.; and Nemhauser, G., eds., *Studies in Integer Programming*, volume 1 of *Annals of Discrete Mathematics*, 185–204. Elsevier.
- Eom, Y.; and Jo, H. 2014. Generalized friendship paradox in complex networks: The case of scientific collaboration. *Scientific Reports*, 4: 4603.
- Faliszewski, P.; Gonen, R.; Koutecký, M.; and Talmon, N. 2022. Opinion diffusion and campaigning on society graphs*. *Journal of Logic and Computation*, 32(6): 1162–1194.
- Feld, S. 1991. Why your friends have more friends than you do. *Am. J. Sociol.*, 96: 1464–1477.
- Fioravantes, F.; Lahiri, A.; Lauerbach, A.; Sieper, M. D.; and Wolf, S. 2024. Eliminating Majority Illusion. *HAL*, hal-04574647.
- Grandi, U.; Kanesh, L.; Lisowski, G.; Sridharan, R.; and Turrini, P. 2023. Identifying and Eliminating Majority Illusion in Social Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(4): 5062–5069.
- Grötschel, M.; Lovász, L.; and Schrijver, A. 1988. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag.
- Johnson, N.; Velásquez, N.; Restrepo, N.; Leahy, R.; Gabriel, N.; El Oud, S.; Zheng, M.; Manrique, P.; Wuchty, S.; and Lupu, Y. 2020. The online competition between pro- and anti-vaccination views. *Nature*, 582(7811): 230–233.
- Kempe, D.; Yu, S.; and Vorobeychik, Y. 2020. Inducing Equilibria in Networked Public Goods Games through Network Structure Modification. In *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems*, 611–619.
- Lerman, K.; Yan, X.; and Wu, X.-Z. 2016. The “Majority Illusion” in Social Networks. *PLOS ONE*, 11(2): 1–13.
- Letchford, A.; Reinelt, G.; and Theis, D. 2008. Odd Minimum Cut Sets and b -Matchings Revisited. *SIAM Journal on Discrete Mathematics*, 22(4): 1480–1487.
- Liu, X.; Kato, S.; Ren, F.; Su, G.; Zhang, M.; and Gu, W. 2023. Information Gerrymandering in Elections. In Wu, S.; Yang, W.; Amin, M.; Kang, B.-H.; and Xu, G., eds., *Knowledge Management and Acquisition for Intelligent Systems*, 83–97. Springer Singapore.
- McPherson, M.; Smith-Lovin, L.; and Cook, J. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1): 415–444.
- Padberg, M.; and Rao, M. 1982. Odd Minimum Cut-Sets and b -Matchings. *Mathematics of Operations Research*, 7(1): 67–80.
- Schmidt, T. 2011. *Computational complexity of SAT, XSAT and NAE-sat for linear and mixed horn CNF formulas*. Südwestdeutscher Verlag für Hochschulschriften.
- Schrijver, A. 2003. *Combinatorial Optimization: Polyhedra and Efficiency*, volume B. Springer-Verlag.
- Stewart, A.; Mosleh, M.; Diakonova, M.; Arechar, A.; Rand, D.; and Plotkin, J. 2019. Information gerrymandering and undemocratic decisions. *Nature*, 573(7772): 117–121.
- Wilder, B.; and Vorobeychik, Y. 2018. Controlling Elections through Social Influence. In *Proceedings of the 17th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 265–273. International Foundation for Autonomous Agents and Multiagent Systems.
- Yu, S.; Kempe, D.; and Vorobeychik, Y. 2021. Altruism design in networked public goods games. *arXiv:2105.00505*.
- Zhou, X.; and Zhang, Z. 2021. Maximizing influence of leaders in social networks. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2400–2408.