Noname manuscript No.
(will be inserted by the editor)

Inferring True Voting Outcomes in Homophilic Social Networks

John A. Doucette, Alan Tsang, Hadi Hosseini, Kate Larson and Robin Cohen

Abstract We investigate the problem of binary opinion aggregation in a social network regarding an objective outcome. Agents receive independent noisy signals relating to the outcome, but may converse with their neighbors in the network before opinions are aggregated, resulting in incorrect opinions gaining prominence in the network. Recent work has shown that, in the general case, there is no procedure for inferring the correct outcome that incorporates information from the connections between agents (i.e. the structure of the social network).

We develop a new approach for inferring the true outcome that can benefit from the additional information provided by the social network, under the simple assumption that agents will more readily convert to the true opinion than to a false one, generating a homophilic effect for voters with the correct opinion. Our proposed approach is computationally efficient, and provides significantly more accurate inference in many domains, which we demonstrate via both simulated and real-world datasets. We also theoretically characterize the properties that are necessary for our approach to perform well. Finally, we extend our approach to directed social networks, and cases with many alternatives, and outline areas for future research.

1 Introduction

Social choice is the study of group decision making, a domain with ever increasing ramifications for artificial intelligence researchers. In this paper, we present a novel approach to social choice in the presence of a social network that connects agents together, causing the opinions of agents to change based on those of their peers. Often, we would like to aggregate the opinions of agents to determine the truth of some objective proposition, about which the agents have some prior knowledge. When opinions have propagated through the social network, however, the prior knowledge of some agents may be discarded in favor of the opinions of their neighbors. We propose a method for recovering the truth of an objective proposition from agents whose opinions may have been influenced in this fashion.

Address(es) of author(s) should be given

Consider for example a soccer game with spectators in the stadium acting as voters. The spectators are polled to determine whether a ball has crossed a line during the match. Either the ball has crossed, or it has not, but the opinions of individual voters regarding the truth of the matter may differ because of their differing perspectives on the event. Voters positioned far from the event may be unable to accurately assess the outcome compared to those positioned nearby. Condorcet's Jury Theorem [16] has long established the conditions under which the truth of such a matter may be recovered: by assuming that individual voters are more likely to observe the correct answer than the incorrect one, the theorem shows that as more votes are added, the chances of outcome that received more votes being correct increase, and that the correct outcome is always more likely to have received more votes. However, the Jury Theorem relies on the assumption that voters' opinions are *independently distributed* and drawn from a binomial process. In practice, however, voters' opinions may not be independently distributed. For example, the voters may talk among themselves before their opinions are aggregated, and may be more likely to speak with their friends, or with fans of the same team. This can distort the distribution of opinions, and introduce correlations into the voters' reports, preventing recovery of the true outcome. As a more extreme example, if an announcer states that the ball *did* cross the line, then voters who did not observe this may report this authoritative opinion rather than their own, because of the social interaction they have with the announcer.

Our new approach entails making a modest assumption about the nature of the process by which opinions spread throughout the social network between when agents observe the event, and when agents are queried for their opinions. By assuming that agents are more readily convinced to switch their opinion to the true outcome, than to a false one, we are able to produce a computationally efficient, intuitive, and effective method for inferring the true outcome *even after opinions have been altered throughout the social network.* Although this assumption limits the scope of our proposed model, it still captures a wide range of important processes. For instance, in scientific discourse it is usually easier to convince a peer of a true opinion than of a false one, since the arguments for the true opinion tend to be stronger (e.g. referring to compelling experiments that refute false opinions, but not true ones). In the sporting event example, agents that are better positioned to observe the truth have more convincing arguments simply by virtue of their positioning: one might be more inclined to believe a spectator seated adjacent to the goal line than one in the most distant bleachers.

The main contributions of this paper are the following:

- We describe and develop a new model for of agent conversations (i.e. opinion dynamics) in a social network, for which inference[1] is tractable (the Correct Conversation model).
- We provide a theoretical characterization of the model and investigate its properties on a variety of simulated and real world datasets and show that our model can effectively predict the ground truth more accurately than majority voting, under certain conditions.

[1] Throughout, we use *inference* in the sense of its common usage in statistics: inferring the value of a variable or parameter on the basis of a given set of data. For us, the variable of interest is usually the winner of an election, while the data are usually the votes.

The paper begins with a formalization of the problem domain, as well as a discussion of relevant background information and recent work on the topic in Section 2. Sections 3 and 4 present our new approach and formally characterize the problem domains for which it offers an advantage over existing approaches. Sections 5 extends the model to the cases of a directed or hierarchical social network. Finally, Sections 7 and 8 provide a complete discussion of related work, conclusions, and future directions.

2 Background

Differences of opinion or judgement can be settled by having the affected parties cast votes. There are many ways of combining votes, but in this paper we require only the familiar concept of *plurality* voting, in which voters consider a set of alternatives, and each casts a ballot for the single alternative they prefer most. The alternative with the most votes is declared the winner. Voting problems can be viewed on a continuum between objective and subjective extremes, corresponding respectively to voters expressing their observations of some objective truth about the world (e.g. whether a ball crosses a line or not), and voters expressing their subjective preferences (e.g. favourite colors).

In this paper, we focus on applications that lie toward the *objective* side of voting on decisions. The objective of a voting system in this domain is to infer the true outcome from the observed votes.

Perhaps the best known example of objective voting for preference aggregation is Sir Francis Galton's work in 1906, eliciting the weight of an ox by asking villagers to speculate on it. There has been considerable work since then, culminating in the recent developments in online systems such as Robovote.org [10].

However, The earliest research concerned with objective voting was that of the Marquee de Condorcet, summarized and translated into its modern form by Black [5]. Condorcet's Jury Theorem considered the problem of inferring the innocence or guilt of a defendant at trial via the votes of the jury. In Condorcet's model, each member of the jury observed the truth of the matter with a fixed probability $p > 0.5$.lead to more accurate predictions The Jury Theorem shows that, as the number of aggregated votes increases, the probability that the alternative receiving more votes is the true outcome will increase as well. Condorcet's model is readily represented as a Bayesian Network, shown in Figure 1. Individual voters are more likely to vote for the true outcome than the other outcome, but the voters do not influence each other, because there is assumed to be no social network in this model, or at least, no way for opinions to spread along such a network.

Inference in Condorcet's model is straightforward: a count of the votes for alternative c is a sufficient statistic for $P(W = c)$. However, real voting problems often do not conform to Condorcet's model. In particular, in real voting problems, voters often *interact* with one another, and influence each others' votes. If interactions between the voters have taken place, then the votes (i.e. $V_1 ... V_n$) will no longer be independently and identically distributed. Instead, they will exhibit some degree of correlation, depending on the exact nature of the interactions that took place between them.

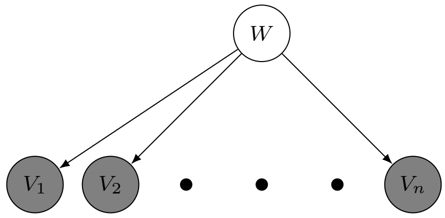


Fig. 1 A Bayesian Network showing Condorcet's model for the generation of votes. Votes (V) are observed (grey), and are influenced by the unobserved "true" outcome W. Votes do not influence each other.

Conitzer [14] proposed a model in which voters did influence each others' opinions. In this model, the probability of a voter expressing the correct opinion is proportionate to the product of the probability of voting a certain way given which outcome is correct, $g(V|W)$, and the probability of voting a certain way given the votes of ones' neighbours $h(V|N(V))$. The central result of this initial work by Conitzer was that a sufficient statistic for the likelihood of a given candidate being the winner, based on the observed votes *and the interactions between them*, $\mathcal{L}(W = c|V)$ is in fact $\prod_i g(V_i|c)$, exactly the same as the sufficient statistic for Condorcet's model (where $g(V_i = c|c) = p$). The proof of this follows directly from the observation that h does not depend on W, and so will contribute an identical quantity to the likelihood of both alternatives, allowing it to be factored out and discarded when maximizing the likelihood. This result is surprising because, intuitively, one might suppose that a known model for correlations between the votes would provide *some* predictive power. However, because the only assumption is that neighboring opinions are related to each other (and not necessarily to the truth), neighborhoods with many incorrect nodes are just as likely as those with many correct nodes, and so the composition of neighbourhoods conveys no additional information. Ultimately, our contribution takes the form of a refinement of this model, maintaining the intuitively nice model of opinions depending on the opinions of ones' neighbours and showing that under a modest additional assumption those interactions can offer additional information that leads to more accurate predictions.

2.1 Opinion Dynamics

The discussions of our model in subsequent sections depends to a large degree on the nature of the *opinion dynamics* present on the social network. Opinion dynamics describe the fashion in which opinions spread throughout a given social network.

Early work in opinion dynamics finds its roots in Granovetter's [22] study of weak ties . In this paper, he finds that many important social processes propagate through more casual associations within one's social network, the so-called *weak ties*. Opinion dynamics models the propagation of information through such networks. Early work in this field examine the penetration of novel ideas through specialized communities: for example, with the introduction of hybrid corn in rural communities [36] and the practice of prescribing antibiotics among physicians [13]. In these scenarios, agents are individual decision makers embedded in a social network. Each agent has a binary state: either they have adopted an idea (denoted +1), or they have not (denoted -1). Initially, agents in a social network begin in the -1 state, with a handful chosen as early adopters whose states are fixed at +1 and do not update. Each agent can observe the states of her neighbors. Because novel ideas are viewed as risky or costly, agents are reluctant to adopt an idea unless a certain fraction of their neighbors have also adopted the idea. As the simulation proceeds, each agent may choose to update their state from -1 to +1 if the adoption rate among her neighbors rises above a threshold; thus, an innovation diffuses throughout the social network. The process continues until agents cease updating their statues (i.e. reaches a stable state).

While the examples above depict the adoption of innovation as an irreversible choice, the same model can be extended to capture the dynamics of choice. For instance, consider a community of agents that are choosing to adopt one operating system or another. There is a benefit to having the same operating system as your friends, due to interoperability and compatibility issues. Each agent then, is playing a coordination game with her neighbors in the social network, with positive individual payoffs when two neighboring agents agree on choice A, a different positive payoff for converging on choice B, and the poorest payoff when they disagree. These payoffs may vary between agents. It can be shown that each agent's choice reduces to the threshold strategy above, and the social network may settle in a variety of stable states (equilibria) where agents may all converge on a single choice, or may remain split between the choices. In particular, this model is used to explain why some innovations which offer objectively better performance over their competitors can nonetheless still fail to achieve market dominance [33]. The abstract model underpinning these processes was also studied by Kempe et al. [29], who examined the problem of finding a minimal "target set" of vertices that, once activated, are sufficient to cause activation of all other vertices through a threshold-based process.

This process can be seen, then, as a proxy for voting. Agents in the network communicate and exchange information as they seek to coordinate on an outcome. Each agent acts on private information in the form of individual coordination payoffs, but also signals and reacts to information from neighbors. If we assume, in the spirit of Condorcet, that agents' pieces of private information are correlated as noisy signals drawn from an underlying truth, then as each agent acts in the coordination game to maximize her own utility, the community is also attempting to uncover the ground truth. The problem we study differs from innovation adoption in the particulars of how individual agents revise their states. While agents in innovation diffusion adopt simple threshold strategies, our agents act according to the Correct Conversation Model, detailed in the section below. This is because, while threshold strategies model the idea of *adopting* a new concept, the process we aim to model involves competing opinions that may be swayed though social interactions. In particular, we allow agents to switch from the "old" or "wrong" opinion to the "new" or "correct" one, as in threshold models. Moreover, we also allow agents to switch in the opposite direction from a "correct" opinion to an incorrect one, at a different rate, as might be expected in social discourse on a minor political issue, for example.

Symbol	Meaning
G	The social network.
V	The set of vertices (i.e. agents) in the social network.
E	The set of edges (i.e. connections) in the social network.
n	$\lvert V\rvert$, the number of agents in the social network, who participate in the vote.
m	$\lvert E\rvert$, the number of edges in the social network.
$E_{i,j}$	An indicator variable for the presence of an edge between i and j in G.
X_i	The initial opinion assigned to an agent i.
V_i	The final (report) opinion of a particular agent i.
C	The set of alternatives $\{\psi, \omega\}$.
λ	A particular element of C.
N_i	The neighbors of V_i; i.e. the set of vertices V_j such that $E_{i,j}$ is true.
W	The correct alternative, an element of C (i.e. the ground truth).
$g(V\mid W)$	The probability distribution of a vote, conditioned on W.
$h(V, N)$	The joint probability distribution of a vote and the votes of its neighbours.
$h'(V_i, V_j\mid W)$	The joint probability distribution of two votes, conditioned on W.

Table 1 The summary of the notation used in this paper.

3 Model

To capture the idea of interactions between voters, we model voters as vertices in a social graph. An edge between two vertices indicates that the voters in question may influence each others' opinions about the event of interest. Formally, we denote by $G = \{V, E\}$ a graph representing a social network with $|V| = n$ vertices (corresponding to voters in an election), and $|E| = m$ undirected edges. An edge $E_{i,j}$ between vertices V_i and V_j indicates that opinions can propagate between voters i and j. We will frequently treat $E_{i,j}$ as an indicator variable, true if and only if vertices i and j are connected. This is a shorthand for $I(E_{i,j} \in E)$. Similarly, we will frequently treat V_i as a variable taking on values in the set of alternatives or alternatives $C = \{\psi, \omega\}$, representing the observed vote of voter V_i. The true winner is represented by a random variable W, which takes on values from C. The related variable X_i is used to refer to the *initial* opinion of the agent corresponding to vertex V_i, which may be changed by social interactions prior to the reporting stage, and adopts values from the same set as V_i. We denote by N_i the set of neighbors of vertex V_i, that is: $N_i = \{V_j \in V | E_{i,j}\}$. The process we model is a vote held over the set of alternatives $C = \{\psi, \omega\}$ with true winner $W \in C$, with voters V, connected by edges E is the social network G. We assume that voters are honest, and report their true opinions without strategic behaviour. Table 1 summarizes our notation. Note that throughout, random variables and sets are represented by capital Latin letters, while the values a random variable may take on are represented by lowercase Greek letters. Constants and function names are represented by lowercase Latin letters

3.1 State of the Art

Conitzer [14] studied the question of whether inference along the lines of Condorcet's model could be conducted when voters had talked among themselves, provided that the set of interactions between each voter was known, and the social dynamics of the interaction were known. This assumed that the probability of an observed vote could be expressed in terms of two functions $g(V_i|W)$ (the probability distribution of a vote given the correct alternative) and $h(V_i, N_i)$ (the joint probability distribution of a vote and its neighbours).

Using this model, Conitzer showed that if h does not depend on which outcome is correct, then the Maximum Likelihood Estimator (MLE) for a given outcome being correct is proportionate to the fraction of votes cast for the outcome — exactly the estimator used by Condorcet, which does not depend on knowledge of the social network structure at all. This in turn [15] led to the proposal of a more complex model where opinions were not distributed at random among agents, but among edges (conversations) between agents in the social network. Although this second model produced a tractable estimator, the assumption of opinions arising from conversations *independently* of each other, even when the same agent participates in them, does not seem to be a natural representation of the spread of opinions in a social network.

We propose a model in the same spirit as Conitzer's original model, but with an h function that *does* depend on the correct outcome. The model we propose is termed the "Correct Conversation" (CC) Model, and makes a mild, and realistic, assumption that *pairs* of adjacent agents in the social network agree on the correct opinion more often than *pairs* of adjacent agents agree on an incorrect opinion, after opinion propagation has taken place. In the remainder of this section, we derive the Maximum Likelihood Estimator for this model, and provide several examples showing how the estimator works in practice, and how it can recover the correct opinions even after an unknown process has caused opinions to propagate through the social network.

3.2 Formal Statement of the Model

We now propose a formal statistical model for domains where voters' conversations are more likely to lead to the discovery of truth than of falsehood. Suppose, following Conitzer [14] that the likelihood function for the alternatives, given the votes, can be factored independently for each voter and into two functions, one representing the voter's innate or initial tendency to identify the truth (g), and the other representing the tendency for voters to agree with their neighbors (h). In our model, h as well as g is dependent on W, and the likelihood function $\mathcal{L}(W = \lambda|V)$ for an alternative $\lambda \in C$ being the winner, given observed votes V is given by:

$$\mathcal{L}(W = \lambda|V) \propto \prod_i g(V_i|W = \lambda)h(V_i, V_{N_i}|W = \lambda) \tag{1}$$

with strict equality holding when a uniform prior over the candidates is assumed (as is usual in voting). Note that voters' differing V_{N_i} values account for social network connections, even though the equation may at first appear similar to an i.i.d. distribution, and that this function describes the joint likelihood, and not a probability density function for individual opinions. Additionally, we suppose that the tendency for voters to agree with their neighbors in general can be factored into a tendency to agree with each neighbor individually. This makes the implicit assumption that social pressures are log-linear, meaning that agreeing with two neighbors should be a constant factor more likely than agreeing with one. Formally, this means h factors as:

$$h(V_i, V_{N_i}|W = \lambda) \propto \prod_{V_j \in V_{N_i}} h'(V_i, V_j|W = \lambda) \quad (2)$$

Finally, we place some constraints on the functions g and h' so that there is a bias toward the correct outcome both in terms of voters' innate opinions, and tendency to agree with their neighbors. The former is essentially identical to Condorcet's requirement, that is each voter's innate opinion must be right more often than not, while the latter represents our "Correct Conversation assumption" (voters who talk to one another are more likely to agree on the truth than on a falsehood):

$$g(V_i = \lambda|W = \lambda) > g(V_i \neq \lambda|W = \lambda) \quad (3)$$

$$h'(V_i = \lambda, V_j = \lambda|W = \lambda) > h'(V_i \neq \lambda, V_j \neq \lambda|W = \lambda) \quad (4)$$

Note that we do not require any constraint on the likelihood that adjacent voters will disagree[2].

As an additional explicit assumption, we require that these functions are identical for every voter. That is, the functions' results depend only on what the votes of an agent and its neighbors are, not which agents cast them.

The constraints implied by our model imply a certain set of assumptions about the nature of the opinions and the social network underlying them. If these assumptions hold, the model will work well. If they do not, it will not. As outlined above, the main assumptions are:

1. The correct conversations assumption, which states that voters who are adjacent in the social network who both have the same opinion are more likely to have the correct opinion than the incorrect one.
2. Condorcet's correctness assumption, which states that voters are more likely to have a correct opinion than an incorrect opinion.
3. Voters' opinions are distributed identically, though not necessarily independently. That is, a single, homogeneous, process is at work producing driving voters' opinions, but this process may take the opinion assigned to one voter into account when assigning opinions to other voters.

Note that our model makes no assumptions at all regarding *time*, and in fact, has no temporal component. We require only that the assumptions above hold for a particular set of opinions, and a particular social network, at the time that the opinions were gathered. Many possible processes can produce data consistent with these assumptions, but not all standard opinion dynamics processes will generate such data.

[2] Note also, that if other inequalities are known (e.g. $h'(V_i = \lambda, V_j = \lambda|W = \lambda) < h'(V_i \neq \lambda, V_j \neq \lambda|W = \lambda)$, our techniques can often be naturally extended to infer the true outcome by changing the interpretation of the counts described below. However, we believe these other cases are less likely to arise in practice.

3.3 Finding the Most Likely Alternative

Having described the likelihood function for the simplest version of the Correct Conversation model, we now explain how it may be used to infer the winner of a vote on the basis of the votes cast, and the observed social network structure. Finding the most likely winner of the election (i.e. the alternative most likely to be correct.) reduces to finding $\arg\max_{\lambda \in C} \mathcal{L}(W = \lambda | V)$. To facilitate this, we adopt the notation that $g(V_i = \lambda | W = \lambda) \propto p$, $g(V_i \neq \lambda | W = \lambda) \propto 1 - p = \dot{p}$, $h'(V_i = \lambda, V_j = \lambda | W = \lambda) \propto q$, $h'(V_i \neq \lambda, V_j \neq \lambda | W = \lambda) \propto \dot{q}$, and $h'(V_i = \lambda, V_j \neq \lambda | W = \lambda) \propto r$, where $q + \dot{q} + r = 1$. We can then write a more compact likelihood function as:

$$\mathcal{L}(W = \lambda | V) \propto p^x \dot{p}^{n-x} q^{2y} r^{2z} \dot{q}^{2(m-y-z)} \tag{5}$$

where n is the total number of voters, m is the total number of edges between voters in the social network, x is the number of votes observed for the alternative whose likelihood is under consideration, y is the number of *concordant edges* (i.e. edges between two voters that both vote for the same alternative) belonging to the alternative whose likelihood is under consideration, and z is the number of *discordant edges* (i.e. edges between voters who disagree). Note that r^{2z} has an identical value in the likelihood functions of both alternatives, and so can be ignored in most cases, though we will use it in some subsequent derivations.

It is clear that the alternative that maximizes this function will be the one with the highest values of x and y, but note that we do not know the precise values of p and q (if we did, we would already know which alternative was victorious, since these variables will have higher values for the true winner). We do have constraints on the values of p and q, which are embodied by Equations 3 and 4. Based on our assumptions we know that $p > \dot{p}$ and that $q > \dot{q}$. Therefore, there is an unambiguous winner when $x > \frac{n}{2}$ **and** $y > \frac{m-z}{2}$. That is, **an alternative that captures both a majority of the votes *and* a majority of the concordant edges, is certain to be the most likely winner.**

In the cases where $x > \frac{n}{2}$ and $y < \frac{m-z}{2}$, or where $x < \frac{n}{2}$ and $y > \frac{m-z}{2}$, however, identifying the winner will be more complex. Our Correct Conversation assumption was that $p > \dot{p}$ and $q > \dot{q}$, so one of the observed counts represents the unlikely event that more votes (or more concordant pairs) were observed for the incorrect alternative. Thus, we need to determine whether it is more likely that $p > \dot{p}$ but $x < \frac{n}{2}$ or $q > \dot{q}$ but $y < \frac{m-z}{2}$. Put another way, we should award that the election to whichever alternative is more likely to have lost one of the counts by a statistical fluke. Assuming without loss of generality that the true winner W has won a majority of concordant edges, but not of votes, and the existence of a uniform prior over parameter values, what we want to know is whether or not

$$\begin{aligned} &P(p > \dot{p} | x, n, \lambda = W) P(q > \dot{q} | y, z, m, \lambda = W) > \\ &\quad P(p > \dot{p} | x, n, \lambda \neq W) P(q > \dot{q} | y, z, m, \lambda \neq W) \end{aligned} \tag{6}$$

Both values can be expressed without difficulty for $\lambda = W$, and are symmetric for $\lambda \neq W$ (i.e. the values of x and $n - x$ will be swapped):

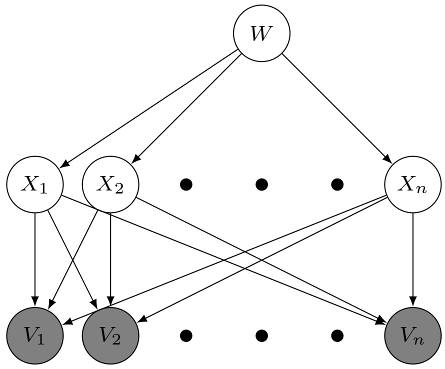


Fig. 2 A Bayesian Network showing a simple generative process for which the CC model gives a consistent estimate of the true winner, W. Votes (V) are observed (grey), and are influenced by the unobserved initial opinions X, which are in turn, influenced by W. Neither votes nor initial opinions influence each other directly. Note the similarity to Conitzer's independent conversations model.

$$P(p > \dot{p}|x, n, \lambda = W) = \frac{\int_{0.5}^{1} p^x (1-p)^{n-x} dp}{\int_{0}^{1} p^x (1-p)^{n-x} dp} \tag{7}$$

$$P(q > \dot{q}|y, z, m, \lambda = W) = \frac{\int_0^{0.5} \int_0^q \eta d\dot{q}\ dq + \int_{0.5}^1 \int_0^{1-q} \eta d\dot{q} dq}{\int_0^{0.5} \int_0^q \eta d\dot{q}\ dq + \int_{0.5}^1 \int_0^{1-q} \eta d\dot{q}\ dq + \int_0^{0.5} \int_0^{\dot{q}} \eta dq\ d\dot{q} + \int_{0.5}^1 \int_0^{1-\dot{q}} \eta dq\ d\dot{q}} \tag{8}$$

where

$$\eta = P(q, \dot{q}|y, z, m) = q^y \dot{q}^z (1 - q - \dot{q})^{m-y-z} \tag{9}$$

3.4 Example Processes

While we have described a general class of models, in keeping with Conitzer's approach, we now provide a specific example of a process that satisfies the constraints associated with the Correct Conversation (CC) model. The initial model we consider is a generative process, which can be represented by the Bayesian Network depicted in Figure 2. In this process, each voter is assigned an initial opinion X_i, which depends on the true outcome W, and which represents their unaltered observation of the event. These initial opinions then propagate to their neighbors. The process appears similar to Conitzer's independent conversations model, but note that each X influences *many* final votes (all the neighbors of node V_i in the social network), whereas each edge $E_{i,j}$ in the independent conversations model influences exactly two final votes. The proposed process represents a situation where voters independently observe an event, discuss it with each of their neighbors (once, and simultaneously), and then arrive at a final opinion based on both their opinions and those of their neighbors.

We parametrize the model so that

$$P(X_i = \lambda | W = \lambda) = \mu$$

and

$$P(V_i = \lambda | X_{N_i}, X_i, W = \lambda) = \frac{I(X_i = \lambda) + \sum_{X_j \in X_{N_i}} I(X_j = \lambda)}{|N_i| + 1}$$

where I is an indicator function (that is, $P(V_i = \lambda | X_{N_i}, X_i, W = \lambda)$ is the fraction of V_i and its neighbours that agree on λ). Although direct inference for W in this process is possible, along much the same lines as in the independent conversations model, we will show that the process satisfies the assumptions of the Correct Conversation model, and thus, that the maximum likelihood estimator for the Correct Conversation model is a consistent estimator for W in this process. Specifically, we will show that using $g(V_i|W) = P(V_i|W)$, and $h(V_i, V_j|W) = P(V_i, V_j|W)$, a consistent estimator is obtained.

First, it is straightforward to show that the probability of any given voter expressing the correct final opinion is equal to μ (i.e. $P(V_i = W) = \mu$). This follows from the fact that

$$P(V_i = \lambda | X_{N_i}, X_i, W = \lambda) = \frac{\mathbb{E}[I(X_i = \lambda) + \sum_{X_j \in X_{N_i}} I(X_j = \lambda)]}{|N_i| + 1}$$

$$P(V_i = \lambda | X_{N_i}, X_i, W = \lambda) = \frac{(|N_i| + 1)\mu}{|N_i| + 1} = \mu$$

In turn, this implies that the expected fraction of voters who express the correct opinion will also be μ:

$$\frac{1}{n}\mathbb{E}[\sum_{V_i \in V} I(V_i = \lambda)] = \frac{n}{n}\mu = \mu$$

Therefore, the proposed generative process satisfies the Correct Conversation model's constraint that $g(V_i = \lambda | W = \lambda) > g(V_i \neq \lambda | W = \lambda)$ in expectation if $\mu > 0.5$.

It is similarly straightforward to show that the constraint $h'(V_i = \lambda, V_j = \lambda | W = \lambda) > h'(V_i \neq \lambda, V_j \neq \lambda | W = \lambda)$ is satisfied. Here it will be convenient to adopt the notation that $B_i = I(X_i = W)$, as an indicator for whether voter i's initial opinion was correct or not.

$$\mathbb{E}[I(V_i = W) \times I(V_k = W)] = \mathbb{E}[\frac{B_i + \sum_{X_j \in X_{N_i}} B_j}{|N_i| + 1} \times \frac{B_k + \sum_{X_l \in X_{N_k}} B_l}{|N_k| + 1}]$$

$$\mathbb{E}[I(V_i = W) \times I(V_k = W)] = \mathbb{E}[B_i B_k + rB_i^2 + \left(\sum_{X_l \in X_{N_k} \setminus X_i} B_i B_l \right) + rB_k^2$$

$$+ \left(\sum_{X_j \in X_{N_i} \setminus X_k} B_k B_j \right) + \left(\sum_{X_w \in X_{N_k} \cap X_{N_i}} B_w^2 \right) + \left(\sum_{X_j \in X_{N_i}} \sum_{X_l in X_{N_k} \setminus X_{N_i}} B_l B_j \right)$$

$$+\left(\sum_{X_l \in X_{N_k}} \sum_{X_j in X_{N_i} \setminus X_{N_k}} B_l B_j\right) - \left(\sum_{X_l \in X_{N_k} \setminus X_{N_j}} \sum_{X_j in X_{N_i} \setminus X_{N_k}} B_l B_j\right)]$$

Observing that the expectation of $B_i^2 = \mu$ (The indicator variable has value 1 a fraction μ of the time), and assuming that nodes are connected with probability r, and have an expected number of neighbors d, we derive:

$$\begin{aligned}\mathbb{E}[I(V_i = W) \times I(V_k = W)] =& \mu^2 + \mu r + d\mu^2 + \mu r + d\mu^2 + rd\mu \\ &+ d^2(1-r)\mu^2 + d^2(1-r)\mu^2 - d^2(1-r)^2\mu^2 \\ \mathbb{E}[I(V_i = W) \times I(V_k = W)] =& \mu(2r + rd) + \mu^2(1 + 2d + 2d^2(1-r)r)\end{aligned}$$

By identical derivation,

$$\mathbb{E}[I(V_i \neq W) \times I(V_k \neq W)] = (1-\mu)(2r + rd) + (1-\mu)^2(1 + 2d + 2d^2(1-r)r)$$

It follows that, if $\mu > 0.5$, then in expectation there will be more correct concordant edges than incorrect ones, so $q > \dot{q}$, as required by the CC model. However, as we show in the next section, although this process satisfies the motivational requirements of the CC model, the CC model would not outperform a direct count of the votes in expectation on problems generated from this process (it would perform approximately identically). In the next section, we characterize the processes for which the CC model *does* provide more accurate inference.

4 Characterization of the CC Model's Inference Advantages

In the previous section, we introduced the Correct Conversation model, and showed how a consistent estimator could be derived for the winner of a vote that incorporates social network structure. A key feature of the CC model is that it makes assumptions about the types of network dynamics that are present: adjacent voters must be more likely to agree on the correct alternative than not, after the opinion dynamics have taken place.

In this section we will provide a precise specification of situations in which social network structures will provide additional information, and thereby allow for more accurate inference of the correct outcome than a simple counting of votes.

4.1 Preliminaries

Throughout this section, for the sake of brevity, it will be useful to refer to indicator variables taking on the values $\{-1, 1\}$ in response to an agent's opinion being respectively incorrect or correct. We therefore define B_i as an indicator over voter's *final* (reported) opinions V_i:

$$B_i = \begin{cases} 1 & \text{if } V_i = W \\ -1 & \text{otherwise} \end{cases}$$

and Z_i as an indicator over voter's *initial* opinions X_i:

$$Z_i = \begin{cases} 1 & \text{if } X_i = W \\ -1 & \text{otherwise} \end{cases}$$

We also refer throughout to two random processes, $\mathcal{M}_x(W)$, which makes probabilistic assignments to voters' initial opinions X_i, and $\mathcal{M}_v(X, E, W)$, which makes probabilistic assignments to voters' final opinions V_i, incorporating their initial opinions and the social network structure. These are generalized random processes which do not necessarily satisfy the correct conversations assumption required by our model, and do not refer to any particular system. We will characterize family of models under which our model performs well, and show that it is exactly those which satisfy the correct conversations assumption. Note that, as in Section 3.4, we use X_i to refer to the initial opinion of an agent, but the process described in that section is not used here. A social planner's goal is to recover W, the original parameter of $\mathcal{M}_x$, and the correct alternative, by examining the votes V_i.

4.2 A Formal Characterization of the Effect of Counting Votes

The voting rule suggested by Condorcet's analysis is a simple majority rule: the true outcome is the one the largest number of voters voted for. Underlying this rule is the assumption that $\mathbb{E}[B_i] > 0$ for every agent. That is, regardless of how votes are correlated, a randomly selected voter, *ceteris paribus*, must be more likely to vote for the correct outcome than not.

Provided that M_v generates a final set of B_i values such that individuals in distant parts of the network have opinions that are independent from each other, and that the network is sufficiently large (i.e. $|V| \to \infty$, from the central limit theorem for weakly dependent random variables [4], we have a formula for distribution of the sum of a set of random variables, which follows a normal distribution. The mean of $\sum_i B_i$ is given by

$$\mu_V = \sum_i \mathbb{E}[B_i]$$

and the variance is given by

$$\sigma_V^2 = \sum_{i \in V} \sum_{j \in V} \sigma_{ij}^2$$

where σ_{ij} is the covariance between the two votes. A tolerance interval parametrized by α for this value is then given by

$$\mu_V \pm \alpha \sigma_V$$

and the majority rule will fail when the model returns a negative value, which, following from the above tolerance interval, happens with probability $\Phi^{-1}(\alpha)$ when[3]:

$$\mu_V - \alpha \sigma_V < 0$$

[3] Φ^{-1} being the inverse of the CDF of the normal distribution: $\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} dt$.

$$\mu_V < \alpha\sigma_V$$

$$\frac{\mu_V}{\sigma_V} < \alpha$$

It follows that, if applied to graphs and vote assignments that were drawn uniformly at random from a class where $P(E_{i,j})$ was constant, and $P(V_i = \lambda|W = \lambda)$ was also constant, the fraction of instances where S_{votes} will give the incorrect answer is given by

$$P(\text{Majority}(V) \neq W \mid W) = \Phi^{-1}(\frac{\mu_V}{\sigma_V})$$

4.3 A Formal Characterization of the Effect of Counting Both Votes and Concordant Pairs

The CC Model assumes that $P(V_i = \lambda|W = \lambda) > P(V_i \neq \lambda|W = \lambda)$, and also that if the edge $E(i, j)$ exists, then

$$P(V_i = \lambda, V_j = \lambda|W = \lambda) > P(V_i \neq \lambda, V_j \neq \lambda|W = \lambda)$$

Again, the model does not assume that the votes are uncorrelated (in fact, it explicitly assumes they *are* correlated). Inference for this model will first compute Majority(V), which will behave in exactly the same fashion as was described in the preceding subsection, but also computes a second function under the assumption that each alternative is correct:

$$\text{Edges}(V, E) = sign(\sum_i \sum_j (B_i + B_j)E_{i,j})$$

allowing the determination of the alternative with the larger number of *concordant pairs* in the network (i.e. the one on whom the largest number of pairs of voters agree). The CC model returns the alternative suggested by Majority(V) if that alternative also has Edges$(V, E) > 0$. Otherwise, it declares a tie, and decides using the MLE procedure discussed in the Section 3.

The key feature of the CC model is that it sums over the edges of the graph, which, if the CC assumption holds, should give an estimate of the truth in much the same way as the majority rule does for the votes, but incorporating information about the structure of the social network.

Edges(V, E) effectively sums up some subset of the random variables B_i; therefore, like in the majority rule, we can derive a distribution for this sum, a tolerance interval for the distribution, and thus determine the fraction of instances where it is greater than zero for the correct alternative. As with the majority rule, the sum will be distributed according to a Normal Distribution, such that:

$$\mu_E = \sum_i \sum_j \mathbb{E}[E_{i,j}(B_i + B_j)])$$

and

$$\sigma_E^2 = \sum_i \sum_j \sum_k \sum_l E_{i,j} E_{k,l} \sigma^2_{(i,j),(l,k)}$$

where $\sigma^2_{(i,j),(l,k)}$ is the co-variance between the pairs of nodes (i,j) and (k,l).

The derivation of the tolerance interval and fraction of correct recommendations for Edges(V,E) is then identical to that of the majorty rule.

4.4 Counting Edges is Often an Affine Transformation of Counting Votes

In this subsection, we will derive the surprising result that, for a substantial family of graph and vote generating processes, the statistics computed by the majority rule and Edges(V,E) are in fact affine transformations of each other, and have functionally identical probabilities of producing the correct answer. The proofs of our lemmas and theorem are found in Appendix B.

We rely on the idea of *anonymity* for the processes $\mathcal{M}_x$ and $\mathcal{M}_v$. That is respectively: changing the enumeration of the voters must not change the probability of any particular assignment of the initial opinions X; and changing the names of the voters, but leaving the values of X and E unchanged, must not change the probability of any particular assignment of final opinions V. More formally, for an anonymous $\mathcal{M}_x$, then for any permutation over the numbers from 1 to n, $\pi([n])$, $P(X_i = \lambda | W = \lambda) = P(X_{\pi(i)} = \lambda | W = \lambda)$, and likewise for $\mathcal{M}_v$. That is, changing the names of voters must not change the probabilities of them ending up with a given opinion. We first show that the mean of Edges(V,E) can be expressed as a function of the mean of the majority rule.

Lemma 1 *When averaged over the set of all possible graphs (or all permutations of a given graph), $\mu_E \propto \mu_V$, provided that $\mathcal{M}_x$ and $\mathcal{M}_v$ are anonymous.*

We now consider the relation between the co-variances $\sigma^2_{i,j}$ and $\sigma^2_{(i,j),(k,l)}$.

Lemma 2 *When averaged over the set of all possible graphs (or all permutations of a given graph), $\sigma_E \propto \sigma_V$, provided that $\mathcal{M}_x$ and $\mathcal{M}_v$ are anonymous.*

This leads in turn to the theorem:

Theorem 1 *When averaged over the set of all possible graphs (or all permutations of a given graph),*

$$\Phi^{-1}(\frac{\mu_E}{\sigma_E}) \approx \Phi^{-1}(\frac{\mu_V}{\sigma_V})$$

provided that $\mathcal{M}_x$ and $\mathcal{M}_v$ are anonymous.

Proof Let θ be the expected number of neighbours for a voter. From the proofs of the two preceding lemmas, we have that

$$\mu_E = 2\theta\mu_V(|V| - 1)$$

and,

$$\sigma_E = 2\theta\sigma_V\sqrt{(|V| - 1)|V|}$$

It follows immediately that

$$\lim_{|V|\to\infty} \frac{\mu_E}{\sigma_E} = \frac{\mu_V}{\sigma_V}$$

Thus, for large enough $|V|$,

$$\Phi^{-1}(\frac{\mu_E}{\sigma_E}) \approx \Phi^{-1}(\frac{\mu_V}{\sigma_V}) \quad \square$$

The implication of this theorem is that, when applied to a randomly sampled social network generated such that any two voters have a constant probability of being linked[4], where voters were randomly assigned initial opinions via some process $\mathcal{M}_x$ independently, and then reached a final opinion that incorporates the initial opinions of their neighbors according to a process $\mathcal{M}_v$, counting concordant edges has exactly the same chance of selecting the correct outcome as simply picking the winner by counting the votes, and ignoring the social network structure entirely (discounting the decisions made during ties, for now).

Although this theorem characterizes a class of problems for which social network structure does not matter, it also points the way toward problems where it *does*. In particular, as shown in the proof in the appendix, the theorem holds only when two voters are equally likely to be connected whether their opinions are correct or not, that is, when the generated graph has an equal degree of homophilia for voters holding each opinion. However, there also exists a large class of problems for which this assumption need not hold. We now characterize this set of problems.

4.5 Homophily and the CC Model

A key simplifying assumption in our analysis so far was the idea that voters are just as likely to be neighbors whether they ultimately end up agreeing on the correct answer or not. Even if voters aggregate their opinions on the basis of their neighbors, with a positive bias, the chance of both votes being correct is simply a function of either vote being correct in the first place.

However, Theorem 1 does reveal an important feature that *would* differentiate between the performance of the CC model and more straightforward approaches: *homophily* in the social network. It must not just be more probable that two neighbors agree than that they disagree, but that nodes agreeing are disproportionately likely to be clustered together and tightly connected with each other. There are many generative processes which are capable of producing highly homophilic vote distributions on a social network, and we turn our attention to these now.

An example homophilic process for generating votes is given by Figure 3. In this process, each voter is assigned an initial opinion, $X_{0,i}$ which is correct with probability p. Then, a voter, Swap, is sampled from the set of voters with probability proportionate to

[4] Note, this class is *not* limited to problems where linkages do not depend on other linkage decisions that have been made. Scalefree networks, for example, have the same probability of any two agents being connected in the final graph *a priori*, before any connections have been added. This is all that is required.

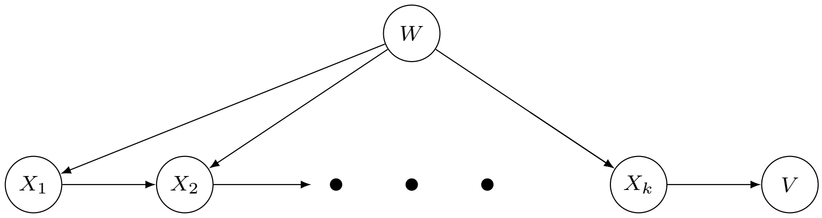


Fig. 3 A Bayesian Network showing a simple generative process for which the CC model gives a *more* consistent estimate of the true winner, W, than simply counting the votes. Votes (V) are observed, and are influenced by a series of unobserved sets of intermediate opinions X, which are in turn, influenced by W and by the social network structure (implicit in each X).

$$P(\text{Swap} = i) \propto 1 + \sum_{j \in N_i} I(X_{0,j} = W \wedge X_{0,i} \neq W)$$

and the opinion $X_{0,\text{Swap}}$ of the sampled voter is flipped to the opposite alternative. This leads to a new profile of opinions X_1, and a new swap is sampled according to this new profile, and flipped as well. The process continues until k swaps have been made in total.

It is easy to see that this process will produce highly homophilic graphs, and the mechanism has strong parallels with innovation diffusion mechanisms mentioned in Section 2 insofar as agents with neighbors who have "adopted" the correct opinion are more likely to "adopt" it themselves. The voters who currently have the wrong opinion, but have many neighbors with the correct opinion will be sampled preferentially. While flipping the opinion of such a node will increase the discriminatory power of the majority rule by $\frac{1}{|V|}$, it also increases the discriminatory power of counting the concordant edges by far more than $\frac{1}{|E|}$. For example, in a scale free network, flipping a hub with many neighbors who are already voting for W would generate many new correct and concordant edges.

5 Asymmetric Influence in Social Network

A natural extension of our Correct Conversation model with symmetric social influence (i.e. V_i and V_j mutually influence each other if adjacent) is to consider social structures wherein individuals may have one-sided, non-reciprocal social influence over their associates. These may include boss-employee relationships, the reach of popular media, and unusually stubborn and steadfast individuals.

In this setting, we need to expand the social network model since pairwise interactions are not necessarily symmetric; that is, for each pair of nodes v and u in graph $G = \langle V, E \rangle$, $(u, v) \in E$ represents a directed edge from u to v, and does not imply $(v, u) \in E$. Let the in-neighbors of v be $\overleftarrow{N_v} = \{u \mid (u, v) \in E\}$ and the out-neighbors of v be $\overrightarrow{N_v} = \{u \mid (v, u) \in E\}$. The sets of votes by in- and out-neighbors of v are represented by $A_{\overleftarrow{N_v}}$ and $A_{\overrightarrow{N_v}}$ respectively. The in-neighbors of v represent those individuals who are influenced by v's opinion, and the out-neighbors of v represent the individuals who influence v's opinion; these sets need not be disjoint.

Incorporating the objective truth of voters' opinions into their ability to influence their in-neighbors, function $f(A_v, A_{N_v}|W)$ is decomposed into three components: $g(A_v|W)$, $\overleftarrow{h}(A_v, A_{\overleftarrow{N_v}}|W)$, and $\overrightarrow{h}(A_v, A_{\overrightarrow{N_v}}|W)$.[5] Expanding the social influence model to independent pairwise factors, we have

$$\overleftarrow{h}(A_v, A_{\overleftarrow{N_v}}|W) = \prod_{A_u \in A_{\overleftarrow{N_v}}} h'(A_u, A_v|W)$$

and,

$$\overrightarrow{h}(A_v, A_{\overrightarrow{N_v}}|W) = \prod_{A_u \in A_{\overrightarrow{N_v}}} h'(A_v, A_u|W)$$

Here, $h'(A_v, A_u|W)$ represents the probability that the configuration A_v and A_u was achieved after v attempted to influence u's opinion ($(u, v) \in E$), given the correct choice W. Based on the pairwise interactions, we define the following distinct interaction cases:

I **Enlightened pair**: Let $q = h'(A_v = A_u = \lambda|W = \lambda)$ be the probability that two vertices with an interaction agree on the correct outcome.

II **Unenlightened pair**: Let $\dot{q} = h'(A_v = A_u \neq \lambda|W = \lambda)$ be the probability that two vertices with an interaction agree on the incorrect outcome.

III **Successful Resistance**: Let $r = h'(A_v \neq \lambda, A_u = \lambda|W = \lambda)$ denote the probability that node u was unconvinced by the arguments of node v toward the incorrect alternative.

IV **Failed to Enlighten**: Let $\dot{r} = h'(A_v = \lambda, A_j \neq \lambda|W = \lambda)$ denote the probability that node u was unconvinced by the arguments of v toward the correct alternative.

If we make the CC assumption that interactions increase the likelihood of finding the correct outcome, then it follows that $q > \dot{q}$, and $r > \dot{r}$. We can calculate the probability of a voting profile, proportionate to:

$$\mathcal{L}(A_V|W) \propto \prod_{v \in V} g(A_v|W) \Big[\prod_{u \in \overleftarrow{N_v}} h'(A_v, A_u|W) \prod_{u \in \overrightarrow{N_v}} h'(A_u, A_v|W) \Big]$$

And selecting the most likely winner is equivalent to:

$$\arg\max_{\lambda \in C} \mathcal{L}(A_V|W = \lambda)$$

Then, let x be the number of votes for c, y be the number of edges where both end points voted for c (case I), $\dot{y}$ be the number of edges where both endpoints voted for c' (case II), and z and $\dot{z}$ be the number of edges that disagree according to case III and case IV respectively. Then, the likelihood for c is maximized as

$$p^x \dot{p}^{n-x} q^y \dot{q}^{\dot{y}} r^z \dot{r}^{\dot{z}}$$

[5] Despite the state of a node being influenced only by its out-neighbors, we need to examine both $\overleftarrow{h}$ and $\overrightarrow{h}$, because we are evaluating the likelihood of observing the configuration of the entire directed graph.

Conversely, the likelihood for c' is

$$p^{n-x}\dot{p}^{x}q^{\dot{y}}\dot{q}^{y}r^{\dot{z}}\dot{r}^{z}$$

Therefore, c wins if following holds true:

$$p^{x}\dot{p}^{n-x}q^{y}\dot{q}^{\dot{y}}r^{z}\dot{r}^{\dot{z}} > p^{n-x}\dot{p}^{x}q^{\dot{y}}\dot{q}^{y}r^{\dot{z}}\dot{r}^{z}$$

Which can be simplified to

$$p^{2x-n}q^{y-\dot{y}}r^{z-\dot{z}} > \dot{p}^{2x-n}\dot{q}^{y-\dot{y}}\dot{r}^{z-\dot{z}}$$

This rule unambiguously selects a winner if $2x > n$, $y > \dot{y}$ and $z > \dot{z}$; or $2x < n$, $y < \dot{y}$ and $z < \dot{z}$. That is, when one alternative has a majority of votes, a majority of both Case I over Case II, and a majority of Case III over Case IV edges. Again, this can be computed in time proportional to the size of the social network.

If the counts of votes and the counts of edges suggest different winners, the ambiguity can be resolved by determining whether it is more probable that the estimators an alternative has not won are incorrect, or that the estimators an alternative has won are. This is accomplished using approximate numeric methods, but with a joint estimate over the probabilities of $r, \dot{r}, q$ and $\dot{q}$ together, since these four values must sum to one, and so are interdependent.

Empirical results demonstrating the effectiveness of this model are shown in the next section.

6 Empirical Evaluations on Common Networks

We empirically evaluate the performance of the Correct Conversation model that we have described on simulated elections held within various social networks. We focus on scenarios with binary outcomes, but evaluate our model against a naive voting process, across several families of random graphs, including both directed and undirected graphs.

To generate a problem instance, we generate a social network with n vertices (the voters), according to one of several (parameterized) graph generation algorithms. We then assign an initial opinion to each voter, with n_c of the n voters starting with the correct opinion, and the remainder receiving the incorrect opinion. A model of influence dynamics (described below) is then applied to the graph, so that opinions of voters tend to change to match those of their neighbors. The final product is a social network where each vertex is assigned a vote, and where the votes are the result of initial opinions modified by discussion between the voters.

Unless otherwise specified, we examine our models on graphs of sizes $n = 40$, with $n_c \in [0, n/2]$ and $k \in [0, 3n/4]$ (where k is the number of Markovian steps in the opinion propagation process). The ranges of n_c and k were chosen to highlight regions of interest; higher values of n_c and k produce situations where most of the population already agree on the correct opinion and no performance gain can be achieved. While we do not vary the number of voters n, we do not expect it impact the qualitative results. The specifics of the graph models are described below. The simulation was implemented using Python version 3.3.2. Each data point (i.e. each separate shaded region in each figure) is the result of 1,000 replications at the corresponding combination of parameter settings, with a new randomly generated graph, and new randomly simulated opinion dynamics.

The integrals involved in computing the probabilities in the event that the counts of votes and concordant pairs do not agree do not have a closed form solution. This need not preclude a principled tie-breaking as both equations (7 and 8) are readily solved via Markov-Chain Monte-Carlo methods. The implementation used to break ties is a simple Metropolis-Hastings algorithm [24, 30], with likelihood functions $x \log p + (n-x) \log \dot{p}$, and $y \log q + z \log \dot{q} + (m-y-z) \log (1-q-\dot{q})$, and rejecting any invalid parameter values automatically. Subsequent points in the chain are sampled from a Gaussian with mean equal to the current point, and variance of 0.1. We use a burn-in period of 1,000 steps, and an average computed from every 20th step in the chain, with 10,000 points used in the average. In practice, a tie-breaking system could sample until sufficient precision to separate the two probabilities has been obtained. The source code is provided in Appendix A. It is essential to handle tie-breaking correctly, because the marginal cases are the ones where using a more sophisticated inference algorithm is most likely to provide an advantage. In cases where a large majority favour a particular alternative, it is essentially certain that there will be a greater proportion of concordant pairs for that alternative as well.

6.1 Influence Dynamics

Let $N(V_i)$ denote those voters whose opinions have a direct influence on v when the latter updates. In undirected graphs, $N(V_i)$ is the neighbors of V_i, and V_i itself (i.e. $N_i \cup V_i$). In directed graphs, $N(V_i)$ is the in-neighbors of V_i and V_i itself. The inclusion of V_i in $N(V_i)$ models a memory effect in the voters, similar to the Friedkin and Johnsen's [20] model of interpersonal influence.

We consider a process by which voters revise their opinions one by one, with parameter k controlling the total number of voters whose opinions have changed and/or changed back (i.e. the total number of Markovian steps in the opinion propagation process). The process also incorporates the idea of Correct Conversation, where each discussion has a higher chance of swaying a voter toward the correct state than the incorrect state. Voters who hold an incorrect opinion are more likely to be swayed to the correct position when many of their friends already hold the correct opinion. On the other hand, a voter who already holds the correct opinion is unlikely to be swayed away from that position (independent of the number of friends with incorrect opinions).

This simulation models situations where voters are convinced only by arguments for the truth (or, with some small constant probability, may revert to a false viewpoint), and where the probability of an interaction with any given neighbor is constant and independent over time (so voters with more neighbors that vote correctly wait less time to talk to someone with the truth). These dynamics are captured by the process by which we select the voters whose opinions will be changed — voters who currently hold the incorrect opinion are more likely to be picked (proportionally to the number of correct neighbors they have), while voters who currently hold the correct opinion are selected with some constant probability.

Formally, our influence dynamics iteratively picks k voters whose opinions will be changed by flipping the candidate they support. Let $P(V_i)$ denote the probability a voter V_i is selected:

$$P(V_i) = \begin{cases} \frac{1}{Z}, & \text{if } V_i = W \\ \frac{1}{Z}(|\{u|u \in N(v), V_u = W\}| + 1), & \text{if } V_i \neq W \end{cases}$$

where Z is a normalization constant such that the probabilities over all voters sum to 1. These probabilities are recomputed after each voter opinion update, and so a voters' opinion may change multiple times over the course of a simulation.

We selected this process because it is a very simple model that satisfies the asymmetry requirement of our correct conversations assumption. The theoretical results of the previous chapter suggest that we should observe an advantage for our inference method on any data generated by a process which satisfies this assumption, so the experiments serve primarily as a "sanity check" in support of our theoretical results. By comparison, many other opinion dynamics processes treat all opinions symmetrically. In these cases, our MLE should return the correct result exactly as often as a simple counting of the votes when no such bias is present.

6.2 Aggregation

Aggregation occurs after influence dynamics have had a chance to affect the opinions of k voters. We consider two methods of aggregating opinions from nodes in a social network: (1) The naive model, in which the aggregated opinion is the majority opinion from the votes at the vertices, and (2) the Correct Conversations model, where we examine both the majority opinion from the vertices, and the distribution of opinions along the edges of the network, as described above. We are interested in comparing the fractions of simulations where each model produces the correct opinion (i.e. the recall rates of each model).

We measured the performance of our models in terms of the improvement in ability to accurately predict the correct winner. We compare performance of our model to the performance of the naive model, by defining acc_{naive} to be the accuracy of the naive model, acc_{ra} to be the accuracy of our model, and $\Delta_{sg,naive} = acc_{ra} - acc_{naive}$ to be the *Performance Improvement* from using our model instead of the naive model. For instance, a value of 0.05 indicates that our model is 5% more accurate (in absolute difference) in deciding the outcome of the election.

6.3 Results

We first examine the performance of our models on two undirected graph models. In undirected graphs, an edge (u, v) represents ongoing communication between voters u and v, which can allow them to influence each others' opinions. The Erdös-Rényi (RE) random graph [17] is a model that incorporates a minimum number of assumptions. Given a connection probability pr, an Erdös-Rényi random graph is generated by connecting distinct vertices u, v with probability pr.

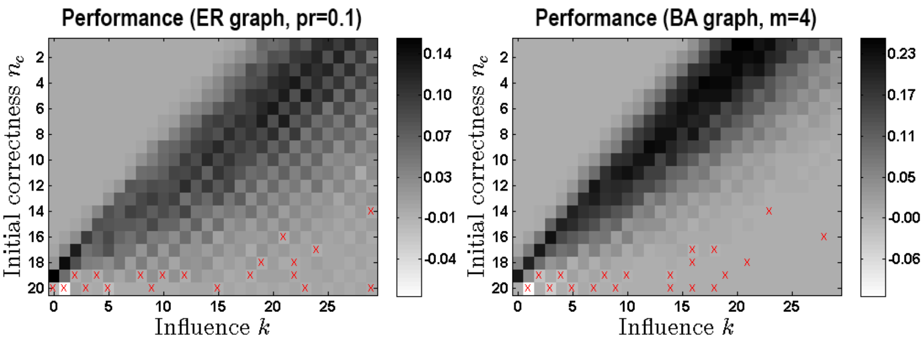
Many human generated networks exhibit a property called *scale-free*. They are characterized by an exponential distribution on node degrees: this leads to a network having very few distinctively high-connected "hubs" and numerous sparsely connected nodes. Scale-free networks can be generated via preferential attachment mechanisms such as the Barabási-Albert (BA) random graph [2]. In this model, new vertices are added to an existing graph in a way so that they are 'drawn' toward high-degree vertices. A Barabási-Albert random graph with attachment parameter m is generated by repeatedly adding vertices to a network, and connecting the new vertex with m existing vertices, each chosen with probability proportional to their respective degrees.

Figure 4 shows a heat map of the Performance Improvement $\Delta_{sg,naive}$ between our Correct Conversation model and naive voting, on both the Erdös-Rényi (a) and the Barabási-Albert (b) random graphs. Darker regions correspond to data points where our model performs better than the naive model. The data points where the Correct Conversation model performs *worse* than the naive model by more than 0.0 are marked with a red X.

In both scenarios, the lighter region in the top left of each plot represents scenarios where neither model gives the correct answer, because almost all voters begin with the incorrect opinion and have little opportunity to change. In this case, the performance of both models is exactly zero. In the lower right portion of each plot (i.e. for large values of n_c and k) similar region exists where, after opinion dynamics, most voters in the network have the correct opinion, so both models almost always give the correct answer. In some of the more extreme parameter settings the Naive model actually performs better, but this effect is very small (typically around 1-2%), and not significantly different from zero. In between however, there is a critical band where the Correct Conversation enjoys a considerable advantage: around 15% in ER graphs, and up to 24% in the more structured BA graphs. The results in the ER graphs is more scattered, likely owing to its absence of structure compared to the BA graphs. Interestingly, in the bottom left, just beside the critical band are several scenarios that elicits the worst performance from our model. In these scenarios, almost half of the voters initially begin with the correct opinion, and just enough opinions are revised to produce a small majority. This allows naive voting to produce the correct result, but does not produce sufficient structure in the edges (i.e. concordant edges) for our model to exploit to greater effect, since the correct voters are essentially randomly distributed throughout the network.

Next, we investigate the performance of our models on directed graphs. In a directed graph, a directed edge (u, v) represents the ability for v to affect the opinion of u; however, the reverse is not necessarily true. The Erdös-Rényi random graph [17] extends naturally to the directed case. Given a connection probability pr, any directed edge from u to v exists with probability pr. The Barabási-Albert model, however, is more challenging to extend to directed graphs. Each new edge added to the graph may be oriented one of two ways: they may allow influence to flow from the older vertices to the new vertex, or the other way around. The first case is more natural, as new vertices represent less experienced members of the network, and it stands to reason they will be influenced by the more well connected, older members. The result is a strongly hierarchical network; it is an acyclic, connected directed graph, with the first node as the sink.

Figure 5(a) and (b) show the Performance Improvement on the directed Erdös-Rényi and the hierarchical Barabási-Albert graphs respectively. Performance on directed ER graphs qualitatively mirrors that of undirected ER graphs, with a slightly higher improvement peaking at 18%.



(a) (b)

Fig. 4 *Performance Improvement of Correct Conversations on Undirected Graphs. (a) Erdös-Rényi random graph with connection probability $pr = 0.1$, and (b) Barabási-Albert random graph with connectivity parameter $m = 4$.*

Performance on the hierarchical BA graph, however, tells an entirely different story. In this plot, the Correct Conversation model peaks at a mere 3% improvement, with the majority of the data points reporting a win for naive voting, by as much as 12%. It is the strict hierarchical nature of the dBA network that causes this unfortunate effect. The earliest vertices in the network are preferentially admired by new vertices, and if they are seeded with the incorrect opinion, they can have a magnified effect on the remainder of the graph. Even worse, these early vertices themselves are influenced by few others (with the earliest vertex having no out-neighbors). Therefore, even under the Correct Conversation model, these vertices are unlikely to be revised to the correct opinions. This explains the poor performance seen on the strongly hierarchical graphs. It also explains why we gain a small advantage in the upper left region: In this region, it is unlikely any reasonable model will aggregate and produce the correct opinion. However, because of the magnified influence of earlier vertices, a fortuitous seeding of correct opinions to them may cause herding to produce a correct aggregate opinion by chance.

To test this hypothesis of the detrimental effect of a strongly hierarchical network, we decide to reverse the orientation of this directed BA model to produce the reversed BA model. Each newly added edge will be oriented toward the newer vertex, meaning that new vertices will influence earlier ones. Under this scheme preferential attachment is preserved, but perversely, highly connected hubs will be sources rather than sinks. Rather than being influential, they will be influenced by numerous other vertices, aggregating their opinions locally. It bears remarking that such a network is quite artificial. These "upper aggregation hubs" do not have a natural counterpart in networks (excepting perhaps for surveillance authorities).

Figure 5(c) shows the effect of the reversed BA model. Qualitatively, it replicates the pattern we have seen so far, but with vast improvement in performance. Performance improvement peaks at an astounding 86%, with a thicker critical band than any of the preceding (or indeed, succeeding) models. Only a few data points (once again in the lower left region) report performance loss. The presence of the aggregation hubs vastly improves the performance of the Correct Conversation model, while the dominance of sink nodes do the opposite.

Real social networks, however, are neither strongly hierarchical, nor reversed. But rather, they are a mix of the two types. Bollobas et al. [7] modified the preferential attachment model to produce this type of hybrid directed scale-free graph. In this model, a new vertex may be added to the graph as an in-neighbor or out-neighbor of one existing vertex. Moreover, the existing graph may be made more dense by adding a new directed edge between existing vertices – the head of the vertex is sampled with weight according to out-degree (plus a small fixed value); the tail is similarly sampled with respect to in-degrees. Graph parameters control the frequency with which each of the three steps occurs. The process continues until a desired number of vertices have been added to the graph.[6] The authors provide parameters that fit their model to the graph of the internet, and we utilize these parameters for simulating a real social network. Figure 5(d) contains the results of this experiment.

Once again, we observe the similar division into three regions: the upper left and lower right showing a tie between the models, and a critical band in between showing improvement for our model, peaking at 22%. Real social networks avoid the "sink" node pitfall of strongly hierarchical model, and behave more like the BA or ER graphs that we have observed before. It should be noted that the critical band is thinner than in previous models, meaning while the performance gain is significant, it occurs in a slightly more restricted set of conditions.

Finally, we also test our model using a real physical social network. The Kapferer Tailor Shop dataset [27] is a series of observations gathered in a tailor shop in Zambia (Northern Rhodesia) by Bruce Kapferer, an early pioneer in social network analysis. The data describes the shift in worker interactions during a series of negotiations for higher wages. This represents the sort of scenario where social network structure may have a profound effect on the outcome of social choice mechanisms. The dataset is divided into four sections, accounting for two time periods (spaced 7 months apart), and contain both "instrumental interactions" (based on work and aid rendered) and "social interaction" (friendship and socioemotional). The former is described as a directed network, while the latter is described as an undirected network; both are rendered as separate adjacency matrices. We focus on the data gathered in the second time period, near the climax of the negotiation period. The data corresponding to these two networks are labelled `KAPFTI2` and `KAPFTS2` respectively in the original dataset. We combine both datasets into one network by taking the union of edges from both sets. However, since the meaning of edge orientation is not described in the dataset, we consider both the described directed graph, and the directed graph with edges reversed. Figure 6 outlines our results.

Therefore, we observe that the Correct Conversation performs well on simulated networks, providing improvements of up to 24% in accuracy over naive voting in certain scenarios. Moreover, the structure that our model relies on is also present in real world interaction networks, allowing us to gain up to 37% increase in accuracy on the Kapferer Tailor Shop dataset.

[6] Actually, to generate a graph on n vertices, the process terminates only when we attempt to add the $n+1$th vertex.

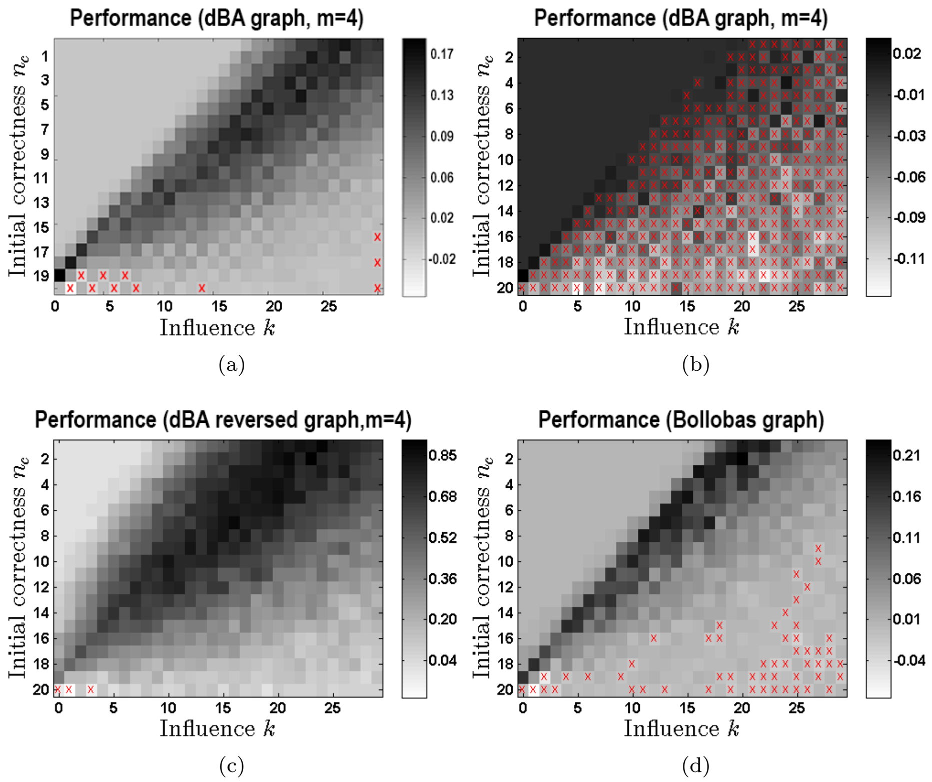


Fig. 5 *Performance Improvement of Correct Conversations on Directed Graphs. (a) Directed Erdös-Rényi random graph with connection probability $pr = 0.1$. (b) & (c) Hierarchical (and reversed hierarchical) directed Barabási-Albert random graph with connectivity parameter $m = 4$. (d) Bollobas directed scale-free random graph.*

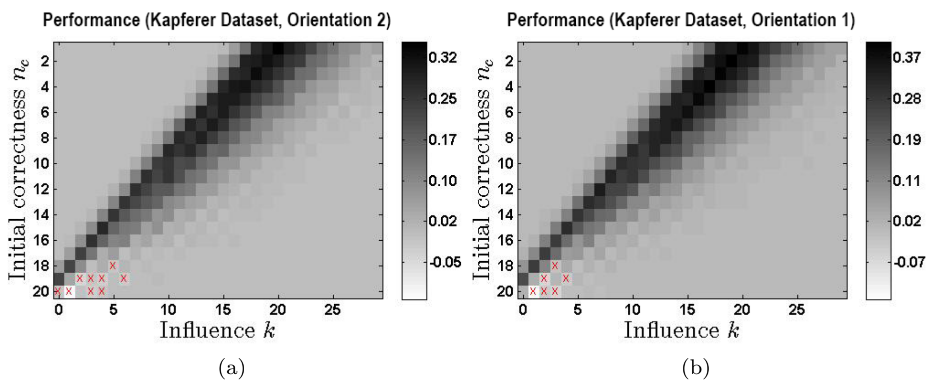


Fig. 6 *Performance Improvement of Correct Conversations on Real Social Networks from the Kapferer Tailor Shop Dataset. (a) and (b) are based on two possible orientations of the social network.*

7 Related Work

Our approach to the problem of incorporating information about social networks structure into voting systems is new, but builds on a rich existing literature studying different facets of the problem.

Our work is most similar to that of Conitzer [15, 14], which adopts the maximum-likelihood approach to voting, and attempts to incorporate information about the likelihood of different local neighborhoods casting certain combinations of final ballots. The 2012 paper, as discussed earlier, assumes that the probabilities of observing a set of votes from a given neighborhood are independent of the true outcome, and shows that no useful information can be derived from network structure on this basis. We assume the opposite, and find that useful information can be derived. The 2013 paper proposes the independent conversations model, in which *edges* rather than vertices are assigned opinions, and shows that useful inference can be conducted on the basis of the social network structure. However, the assumption that "opinions" are generated on the edges rather than the vertices causes this model to represent a different class of problems than the ones we consider.

Other authors have also drawn inspiration from Conitzer's model. Procaccia et al. [32] propose an extension of Conitzer's pairwise conversations model. This extension allows for the possibility of several alternatives to be modelled by adopting a random-utility-theoretic system, in which conversations sample a Gaussian distributed utility for each alternative, and then rank them based on the sample. In this paper, we argued that the pairwise conversation was not suitable for modelling some social interactions, and instead extended Conitzer's earlier vertex-centric model. Mosell et al. [31] have considered the closely related problem of finding aggregation rules for social networks that are generated under conditions similar to our own. However, whereas we consider a single voting rule, and a stochastic process for opinion propagation, they consider a deterministic propagation process and characterize voting rules for which the correct answer can be recovered. Auletta et al. [3] characterize algorithms for a similar process. Elkind et al. [8] characterize the use of bribery and other manipulations in opinion formation on similar networks.

Voting in social networks has also been considered by Boldi, Bonchi, Castillo and Vigra [6]. Here, they focus on offering a pragmatic alternative to direct voting, made available by the proliferation of social media, called *transitive proxy voting.* Motivated by low voter turn out, especially in online environments, their proposed system allows voters to either vote directly or delegate their ballot to a trusted associate in their social network. This associate can, in turn, transitively entrust those ballots to another person in the network. The system may incorporate a dampening factor that effectively reduces the weight of a ballot with each successive act of delegation. The authors analyze theoretic and empirical properties of the system, and compare it to Google PageRank. Tsang and Larson [39] also study voting on social networks, with a focus on strategic elements, rather than on opinion propagation and recovery. Watts [41] examines models of correct voting, but assumes subjective correctness (i.e. voting in one's own interests), rather than objective correctness (as we do). Homophily, which is central to our results, is shown to also play a major role when correctness is subjective. Ismaili and Perny [25] examined multiagent coordination on a social network from a utility-maximizing perspective, rather than using the techniques of social choice proper. Fish et al. [19] examined the problem of extracting the structure of social networks by observing the votes. Grandi [21] provides an extensive survey of these and other social choice topics associated with social networks.

Other researchers have also proposed methods for utilizing network analysis tools for social choice. Tosatto and van Zee [38] propose imputing a social network between the voters based on similarity between their ballots. They consider an election where a set of m policies are being voted upon, and each ballot is an m dimensional vector that votes *yea* or *nay* on each issue. Similarity between voters can be calculated as the number of issues that they agree on. This allows the authors to impute a weighted social network on the voters. They then propose using a modified degree centrality metric for obtaining an outcome similar to that produced by a Median Voter Rule [5]. Wilder and Vorobeychik [42] consider the question of whether a skilled adversary can manipulate an electoral outcome by making use of social network information, a compliment to our approach which could perhaps be extended to uncover such manipulation. Faliszewski et al. [18] demonstrate a similar capability through a detailed model of the effects of campaigning on the electorate. Izsak et al. [26] use network analysis tools to study the synergies between *candidates* rather than between voters.

Other social choice research that incorporate social network components include Salehi-Abari and Boutilier's *Empathetic Social Choice* [34], where an agent's utility not only depends on their own intrinsic utility for the elected outcome, but also on that of their friends. Related work examined this approach for group recommendation [35]. This body of literature differs from our approach as their is focused on utility maximization with subjective preferences, where each voter may have different preferences on the alternatives, while our approach assumes an objective "best" alternative that is preferred by all voters (noisiness aside).

Another related line of work concerns developments toward expanding the range of noise models for which voting rules may be applied beyond those consistent with the assumptions of MLE, but without considering social network structures explicitly [11,43]. As we have shown, by using information about the structure of the social network, we are able to improve the accuracy of our voting mechanism.

Beyond social choice proper, researchers have taken an interest in the question of opinion formation on social networks. Viswanath et al. [40] consider the use of machine learning methods to detect the presence of attempts to manipulate opinions through a social network. Halberstam and Knight [23] use empirical data to show the importance of homophily and other factors in the formation of opinions on social networks like twitter. Acemoglu et al. [1] consider a model of dynamic network formation, and show that the propagation of correct opinions can depend on the strategic behaviours of the agents, and on the structure of the network. Katakis et al. [28] develop a system to recommend political votes based on both user information, and on social connections. More recently, Brill et al. [9] provided a detailed analysis of the termination conditions of opinion diffusion processes on social networks with an eye toward applications in social choice.

8 Conclusions and Future Work

In conclusion, we have proposed the Correct Conversation model of opinion aggregation for agents interacting in a social network: agents formulate some initial opinion about a proposition — an opinion that either does or does not accurately reflect ground truth — and then converse with each other within the social network. The central assumption of our model is that it is easier for two agents to agree upon a truthful proposition than a falsehood, and we show that our model may be used to recover the ground truth in certain types of networks.

Our work is chiefly motivated by Conitzer's [14] question in his titular paper: should social network structure be taken into account in elections? While he shows that network structure does not matter when using a simplistic model, in a follow-up paper [15], he proposes an alternative model where individual conversations (i.e. the edges of the social network) act as noisy, independent samples of the underlying ground truth. While this conversation based model admits maximum likelihood solutions, it also permits unnatural configurations. For instance, a voter may express contradictory opinions in conversations with each of her peers. Our model makes a more natural assumption that each voter holds some particular opinion that is a noisy sample of the ground truth, and conversations between voters may affect this opinion. It is the first model to support tractable maximum likelihood estimations that incorporate information about the structure of the social network.

This form of social influence is reminiscent of the field of opinion dynamics. Early pioneers in this field studied the adoption of new technologies such as hybrid corn [37] and antibiotics [12]. While opinion dynamics is concerned with the long term evolution of opinions in a community, the Correct Conversation model incorporates only a single round of conversations between our voters, and a voters' final opinion is affected only by her immediate neighbors in her social network. This represents a relatively limited window of interactions between voters corresponding to the deliberation period or the lead-up to election day.

In our paper, we have offered both theoretical guarantees on the robustness and limitations of our mechanism and simulations of its effectiveness on artificial and real data sets. In particular, when the Correct Conversation model is able to recover the ground truth even when the majority of agents have been misled in networks that exhibit homophily – the idea that people tend to connect to peers that are similar to themselves. Homophily is a natural human tendency and is present in many social networks. As such, our technique may be useful in bolstering the performance of social choice mechanisms operating in these settings of social influence.

As social media becomes a increasingly ubiquitous part of our lives, the need to quantify the impact of social influence becomes increasingly urgent. While cries of false representation in news outlets are often exaggerated, the coverage from traditional media often come from biased (but known) perspectives. By comparison, the influence of peers in social networks is much more subtle, and their impact, much more difficult to discern. Our paper offers a tool toward countering these forms of social influence for a decision making agency, with applications in crowdsourcing domains, news reliability verification, and detecting and countering online bullying, because it provides a way to detect signs that opinions have been influenced by others, even when we do not observe the opinion dynamics directly. Such a mechanism may also find use evaluating technologies and reviewing products, distinguishing fashionable gimmicks from truly innovative developments.

Future work in this direction could focus on the extension of the Correct Conversation model to more complex graph models; ranked voting rules suitable for use with many alternatives, rather than the binary alternatives case we consider here; and a wider range of opinion dynamics.

References

1. Daron Acemoglu, Kostas Bimpikis, and Asuman Ozdaglar. Dynamics of information exchange in endogenous social networks. *Theoretical Economics*, 9(1):41–97, 2014.
2. Reka Albert and Albert-Laszlo Barabasi. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, 2002.
3. Vincenzo Auletta, Ioannis Caragiannis, Diodato Ferraioli, Clemente Galdi, and Giuseppe Persiano. Minority becomes majority in social networks. In *International Conference on Web and Internet Economics*, pages 74–88. Springer, 2015.
4. Patrick Billingsley. Probability and measure. wiley series in probability and mathematical statistics, 1995.
5. Duncan Black. *The theory of committees and elections*. Springer, 1958.
6. Paolo Boldi, Francesco Bonchi, Carlos Castillo, and Sebastiano Vigna. Voting in social networks. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 777–786. ACM, 2009.
7. Béla Bollobás, Christian Borgs, Jennifer Chayes, and Oliver Riordan. Directed scale-free graphs. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 132–139. Society for Industrial and Applied Mathematics, 2003.
8. Robert Bredereck and Edith Elkind. Manipulating opinion diffusion in social networks. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*, pages 894–900, 2017.
9. Markus Brill, Edith Elkind, Ulle Endriss, and Umberto Grandi. Pairwise diffusion of preference rankings in social networks. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, pages 130–136, 2016.
10. Ioannis Caragiannis, Swaprava Nath, Ariel D Procaccia, and Nisarg Shah. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152, 2017.
11. Ioannis Caragiannis, Ariel D Procaccia, and Nisarg Shah. When do noisy votes reveal the truth? In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 143–160. ACM, 2013.
12. James Coleman, Elihu Katz, and Herbert Menzel. The diffusion of an innovation among physicians. *Sociometry*, pages 253–270, 1957.
13. James Coleman, Herbert Menzel, and Elihu Katz. *Medical Innovations: A Diffusion Study*. Bobbs Merrill, 1966.
14. Vincent Conitzer. Should social network structure be taken into account in elections? *Mathematical Social Sciences*, 64(1):100–102, 2012.
15. Vincent Conitzer. The maximum likelihood approach to voting on social networks. In *Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton-13)*, pages 1482–1487. Allerton Retreat Center, Monticello, IL, USA, 2013.
16. Marie Jean Antoine Nicolas de Caritat et al. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. L'imprimerie royale, 1785. [Translation found in Black, 1958].
17. Paul Erdős and Alfréd Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.
18. Piotr Faliszewski, Rica Gonen, Martin Koutecký, and Nimrod Talmon. Opinion diffusion and campaigning on society graphs. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 219–225, 2018.
19. Benjamin Fish, Yi Huang, and Lev Reyzin. Recovering social networks by observing votes. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 376–384. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

20. Noah E Friedkin and Eugene C Johnsen. Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193–206, 1990.
21. Umberto Grandi. Social choice and social networks. *Trends in Computational Social Choice*, pages 169–184, 2017.
22. Mark S Granovetter. The strength of weak ties. In *Social networks*, pages 347–367. Elsevier, 1977.
23. Yosh Halberstam and Brian Knight. Homophily, group size, and the diffusion of political information in social networks: Evidence from twitter. *Journal of Public Economics*, 143:73–88, 2016.
24. W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
25. Anisse Ismaili and Patrice Perny. Computational social choice for coordination in agent networks. *Annals of Mathematics and Artificial Intelligence*, 77(3-4):335–359, 2016.
26. Rani Izsak, Nimrod Talmon, and Gerhard J. Woeginger. Committee selection with intraclass and interclass synergies. In *Proceedings of the 32nd AAAI Conference*, pages 1071–1078, 2018.
27. Bruce Kapferer. *Strategy and transaction in an African factory: African workers and Indian management in a Zambian town.* Manchester University Press, 1972.
28. Ioannis Katakis, Nicolas Tsapatsoulis, Fernando Mendez, Vasiliki Triga, and Constantinos Djouvas. Social voting advice applications—definitions, challenges, datasets and evaluation. *IEEE Transactions on cybernetics*, 44(7):1039–1052, 2014.
29. David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.
30. Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
31. Elchanan Mossel, Joe Neeman, and Omer Tamuz. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multi-Agent Systems*, 28(3):408–429, 2014.
32. Ariel D Procaccia, Nisarg Shah, and Eric Sodomka. Ranked voting on social networks. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 2040–2046. AAAI Press, 2015.
33. Everett M Rogers. *Diffusion of innovations.* Simon and Schuster, 2010.
34. Amirali Salehi-Abari and Craig Boutilier. Empathetic social choice on social networks. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 693–700. International Foundation for Autonomous Agents and Multiagent Systems, 2014.
35. Amirali Salehi-Abari and Craig Boutilier. Preference-oriented social networks: Group recommendation and inference. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 35–42. ACM, 2015.
36. David Strang and Sarah Soule. Diffusion in organizations and social movements: From hybrid corn to poison pills. *Annual Review of Sociology*, 24:265–290, 1998.
37. David Strang and Sarah A Soule. Diffusion in organizations and social movements: From hybrid corn to poison pills. *Annual review of sociology*, pages 265–290, 1998.
38. Silvano Colombo Tosatto and Marc Van Zee. Bridging social network analysis and judgment aggregation. In *International Conference on Social Informatics*, pages 17–33. Springer, 2014.
39. Alan Tsang and Kate Larson. The echo chamber: Strategic voting and homophily in social networks. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 368–375. International Foundation for Autonomous Agents and Multiagent Systems, 2016.
40. Bimal Viswanath, Muhammad Ahmad Bashir, Mark Crovella, Saikat Guha, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. Towards detecting anomalous user behavior in online social networks. In *USENIX Security Symposium*, pages 223–238, 2014.
41. Alison Watts. The influence of social networks and homophily on correct voting. *Network Science*, 2(1):90–106, 2014.
42. Bryan Wilder and Yevgeniy Vorobeychik. Controlling elections through social influence. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 265–273. International Foundation for Autonomous Agents and Multiagent Systems, 2018.

43. Lirong Xia. Generalized decision scoring rules: Statistical, computational, and axiomatic properties. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 661–678. ACM, 2015.

A Code

```
import random
import math
import copy

# The metropolis-hastings algorithm
def metro_hastings_step(num_params, params_list, likelihood_function
    , step_size):

    old_p = copy.deepcopy(params_list)
    old_l = likelihood_function(old_p)
    for p in range(0, num_params):
        params_list[p] = random.gauss(old_p[p], step_size)
    new_l = likelihood_function(params_list)
    if new_l == float("-inf") or new_l >= old_l :
      params_list = old_p
    elsif math.exp(new_l - old_l) <= random.random():
        params_list = old_p
    return params_list

#Markov Chain Monte Carlo
def metro_hastings_run(params_list, num_variable_params,
    likelihood_function, stat_function, burnin_period=1000, samples
    =10000, steps_per_sample=20):
    for i in range(0,burnin_period):
        params_list = metro_hastings_step(num_variable_params,
    params_list, likelihood_function, 0.1)
    for i in range(0, samples):
        for j in range(0, steps_per_sample):
            params_list = metro_hastings_step(num_variable_params,
    params_list, likelihood_function, 0.1)
        stat_function(params_list)


#Problem specific likelihood function for estimating
# probability that p > 0.5, given the observations we
# have seen.
def p_likelihood(params_list):
    p = params_list[0]      # unknown
    x = params_list[1]      # number of votes supporting alternative
    n = params_list[2]      # number of voters total
    if p < 0:
        p = 0
        params_list[0] = p
        return float("-inf")
    if p > 1:
        p = 1
        params_list[0] = p
        return float("-inf")
    return x * math.log(p) + (n-x)*math.log(1-p)

# statistics collector for p.
def p_stats_maker():
```

```
    counter=[0]
    norm = [0]
    def p_stat_func(params_list):
        if params_list[0] >= 0.5:
            counter[0] += 1
        norm[0]+= 1
    def p_stat_print():
        print(counter[0] / norm[0])
    def p_stat_return():
        return(counter[0] / norm[0])
    return [p_stat_func, p_stat_print, p_stat_return]

# Returns the probability that p>0.5 given simulation results
def probP(numVertexVoteSupport, numVertices):
    stats = p_stats_maker()
    metro_hastings_run([0.5, numVertexVoteSupport, numVertices], 1, p_likelihood, stats[0])
    return stats[2]()

#Problem specific likelihood function for estimating
# probability that q > q_bar, given the observations we
# have seen.
def q_likelihood(params_list):
    q = params_list[0]          # unknown
    q_bar = params_list[1]      # unknown
    y = params_list[2]          # number of concordant edges supporting alternative
    y_bar = params_list[3]      # number of concordant edges supporting opponent
    m = params_list[4]          # number of edges total
    if q < 0 or q_bar < 0 or q > 1 or q_bar > 1 or (1-q-q_bar) < 0:
        return float("-inf")
    else:
        return y*math.log(q) + y_bar*math.log(q_bar) +(m-y-y_bar)*math.log(1-q-q_bar)

# statistics collector for q.
def q_stats_maker():
    counter=[0]
    norm=[0]
    def q_stat_func(params_list):
        if(params_list[0] > params_list[1]):
            counter[0] += 1
        norm[0] += 1

    def q_stat_print():
        print(counter[0] / norm[0])
    def q_stat_return():
        return(counter[0] / norm[0])
    return [q_stat_func, q_stat_print, q_stat_return]

def probQ(numConcordSupport, numConcordOppose, numDiscordant):
    stats = q_stats_maker()
    metro_hastings_run([0.25, 0.25, numConcordSupport, numConcordOppose, numDiscordant], 2, q_likelihood, stats[0])
    return stats[2]()

# Problem specific likelihood function for estimating
# the probability that q > q_bar AND r > r_bar=(1-q-q_bar-r),
# given observations y, y_bar, z, z_bar
```

```python
def qr_likelihood(params_list):
    q = params_list[0]          # unknown
    q_bar = params_list[1]      # unknown
    r = params_list[2]          # unknown
    y = params_list[3]          # number of concordant edges supporting alternative
    y_bar = params_list[4]      # number of concordant edges supporting opponent
    z = params_list[5]          # number of discordant edges supporting alternative
    z_bar = params_list[6]      # number of discordant edges supporting opponent
    if q < 0 or q_bar < 0 or r < 0 or (1-q-q_bar-r) < 0 or q > 1 or q_bar > 1 or r > 1 or (1-q-q_bar-r) > 1:
        return float("-inf")
    else:
        return y*math.log(q) + y_bar*math.log(q_bar) + z*math.log(r) + z_bar*math.log(1-q-q_bar-r)

# statistics collector for qr.
def qr_stats_maker():
    counter=[0]
    norm=[0]
    def qr_stat_func(params_list):
        if((params_list[0] > params_list[1]) and (params_list[2] > 1-params_list[0]-params_list[1]-params_list[2])):
            counter[0] += 1
        norm[0] += 1

    def qr_stat_print():
        print(counter[0] / norm[0])

    def qr_stat_return():
        return(counter[0] / norm[0])
    return [qr_stat_func, qr_stat_print, qr_stat_return]

def probQR(numConcordSupport, numConcordOppose, numDiscordSupport, numDiscordOppose):
    stats = qr_stats_maker()
    metro_hastings_run([0.25, 0.25, 0.25, numConcordSupport, numConcordOppose, numDiscordSupport, numDiscordOppose], 3, qr_likelihood, stats[0])
    return stats[2]()
```

B Proofs

Lemma 3 *When averaged over the set of all possible graphs (or all permutations of a given graph), $\mu_E \propto \mu_V$, provided that $\mathcal{M}_x$ and $\mathcal{M}_v$ are anonymous.*

In expectation, for a randomly generated graph and vote profile according to a process from the family described above, there are $|V|(|V|-1)P(V_i = W \wedge V_j = W \wedge E_{i,j} = 1)$ concordant edges in the graph that vote for the correct answer, and $|V|(|V|-1)P(V_i \neq W \wedge V_j \neq W \wedge E_{i,j} = 1)$ that vote for the incorrect answer.

It should be apparent then that the expected value contributed to μ_E by an edge $E_{i,j}$ is given by:

$$P(V_i = W \wedge V_j = W \wedge E_{i,j} = 1|W) - P(V_i \neq W \wedge V_j \neq W \wedge E_{i,j} = 1|W)$$

Now, notice that $E_{i,j}$ is independent of V and W (since it is used to generate V, and V is not used to generate it). Then we can write

$$P(V_i = W \wedge V_j = W \wedge E_{i,j} = 1|W) = P(E_{i,j} = 1) \cdot P(V_i = W \wedge V_j = W|W)$$

and, by definition of co-variance,

$$P(E_{i,j} = 1) \cdot P(V_i = W \wedge V_j = W|W) = P(E_{i,j} = 1) \cdot (P(V_i = W|W)P(V_j = W|W) + \sigma_{ij}^2)$$

so

$$P(V_i = W \wedge V_j = W \wedge E_{i,j} = 1|W) = P(E_{i,j} = 1) \cdot ((P(V_i = W|W)P(V_j = W|W) + \sigma_{ij}^2)$$

and

$$P(V_i \neq W \wedge V_j \neq W \wedge E_{i,j} = 1|W) = P(E_{i,j} = 1)(P(V_i \neq W|W)P(V_j \neq W|W) + \sigma_{ij}^2)$$

From this we can derive the expected contribution of a single pair of voters to the statistic computed by S_{edges}.

$$\mathbb{E}[E_{i,j}(V_i + V_j)] = P(E_{i,j} = 1)(P(V_i = W|W)P(V_j = W|W) - P(V_i \neq W)P(V_j \neq W))$$

$$\mathbb{E}[E_{i,j}(V_i+V_j)] = P(E_{i,j} = 1)(P(V_i = W|W)P(V_j = W|W) - (1-P(V_i = W))(1-P(V_j = W)))$$

$$\mathbb{E}[E_{i,j}(V_i + V_j)] = P(E_{i,j} = 1)(P(V_i = W|W) + P(V_j = W|W) - 1)$$

but of course, $P(V_i = c) = \frac{\mathbb{E}[V_i]+1}{2}$ (since $\mathbb{E}[V_i] = (1)P(V_i = 1) + (-1)(1 - P(V_i = 1))$), so

$$\mathbb{E}[E_{i,j}(V_i + V_j)] = P(E_{i,j} = 1)(\mathbb{E}[V_i] + \mathbb{E}[V_j])$$

If the prior probability that $P(E_{i,j} = 1) = \theta \forall i, j$, (i.e. we consider the expected performance of the method over all graphs, or all graphs of a family), then it follows immediately that

$$\mu_E = \sum_i \sum_j \mathbb{E}[E_{i,j}(V_i + V_j)])$$

$$\mu_E = \theta \sum_i \sum_{j \neq i} \mathbb{E}[V_i] + \mathbb{E}[V_j]$$

If we further assume that the processes $\mathcal{M}_x$ and $\mathcal{M}_v$ are *anonymous* (i.e. they do not assign especial importance to a particular voter based only on that voter's name), then $\mathbb{E}[V_i] = \mathbb{E}[V_j]$ over the set of all possible social graphs, so:

$$\mu_E = 2\theta(|V| - 1) \sum_i \mathbb{E}[V_i]$$

$$\mu_E = 2\theta(|V| - 1)\mu_V$$

This demonstrates that the means differ by a factor of $2\theta(|V| - 1)$. □

Lemma 4 *When averaged over the set of all possible graphs (or all permutations of a given graph), $\sigma_E \propto \sigma_V$, provided that $\mathcal{M}_x$ and $\mathcal{M}_v$ are anonymous.*

By definition of co-variance, the co-variance between two pairs of votes is given by:

$$\sigma^2_{(i,j),(k,l)} = \mathbb{E}[(V_i + V_j)(V_k + V_l)] - \mathbb{E}[V_i + V_j]\,\mathbb{E}[V_k + V_l]$$

Since the sum of expectations is the expectation of the sum, we can rewrite this as

$$\sigma^2_{(i,j),(k,l)} = \mathbb{E}[V_iV_k + V_iV_l + V_jV_k + V_jV_l] - (\mathbb{E}[V_i] + \mathbb{E}[V_j])(\mathbb{E}[V_k] + \mathbb{E}[V_l])$$

$$\sigma^2_{(i,j),(k,l)} = \mathbb{E}[V_iV_k] + \mathbb{E}[V_iV_l] + \mathbb{E}[V_jV_k] + \mathbb{E}[V_jV_l]$$
$$-(\mathbb{E}[V_i]\,\mathbb{E}[V_k] + \mathbb{E}[V_i]\,\mathbb{E}[V_l] + \mathbb{E}[V_j]\,\mathbb{E}[V_k] + \mathbb{E}[V_j]\,\mathbb{E}[V_l])$$

but $\mathbb{E}[V_iV_j] = \mathbb{E}[V_i]E[V_j] + \sigma^2_{i,j}$, so :

$$\sigma^2_{(i,j),(k,l)} = \sigma^2_{i,k} + \sigma^2_{i,l} + \sigma^2_{j,k} + \sigma^2_{j,l}$$

It follows then that

$$\sigma^2_E = \sum_i \sum_j \sum_k \sum_l E_{i,j}E_{k,l}(\sigma^2_{i,k} + \sigma^2_{i,l} + \sigma^2_{j,k} + \sigma^2_{j,l})$$

and under the assumptions that the prior probability that $P(E_{i,j} = 1) = \theta \forall i, j$, and anonymity, that

$$\sigma^2_E = \theta^2 \sum_i \sum_j \sum_k \sum_l (\sigma^2_{i,k} + \sigma^2_{i,l} + \sigma^2_{j,k} + \sigma^2_{j,l})$$

$$\sigma^2_E = \theta^2 \sum_i \sum_j 4|V|(|V| - 1)\sigma^2_{i,j}$$

$$\sigma^2_E = 4\theta^2(|V| - 1)|V|\sigma^2_V$$

This demonstrates that the variances differ by a factor of $4\theta^2(|V| - 1)|V|$. Note that this implies the standard deviations differ by $\sqrt{(4\theta^2(|V| - 1)|V|)} = 2\theta\sqrt{((|V| - 1)|V|)}$ $\square$.