



HAL
open science

Poll-Confident Voters in Iterative Voting

Anaëlle Wilczynski

► **To cite this version:**

Anaëlle Wilczynski. Poll-Confident Voters in Iterative Voting. 33rd AAAI Conference on Artificial Intelligence (AAAI-19), Jan 2019, Honolulu, Hawaii, United States. pp.2205-2212, 10.1609/aaai.v33i01.33012205 . hal-02948562

HAL Id: hal-02948562

<https://hal.science/hal-02948562>

Submitted on 24 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Poll-Confident Voters in Iterative Voting

Anaëlle Wilczynski

Université Paris-Dauphine, PSL, CNRS, LAMSADE, Paris, France

anaelle.wilczynski@lamsade.dauphine.fr

Abstract

This article deals with strategic voting under incomplete information. We propose a descriptive model, inspired by political elections, where the information about the vote intentions of the electorate comes from public opinion polls and a social network, modeled as a graph over the voters. The voters are assumed to be confident in the poll and they update the communicated results with the information they get from their relatives in the social network. We consider an iterative voting model based on this behavior and study the associated “poll-confident” dynamics. In this context, we ask the question of manipulation by the polling institute.

Introduction

Strategic voting occurs in many scenarios, principally in political elections. Voters may manipulate, by not revealing their true preferences in the ballot they submit, in order to avoid an unfortunate outcome. Although one would desire to prevent this behavior, no voting rule is immune to voter manipulation (Gibbard 1973; Satterthwaite 1975), leading to consider other ways to grasp strategic voting, via for instance game-theoretical analysis. Manipulation in voting is commonly modeled as a strategic game where the players are voters, assumed to be rational, who look for the appropriate ballot to obtain the best possible outcome in the election, according to their preferences. Iterative voting (Meir 2017) is an iterated version of this game where the voters deviate by rounds. The deviations can be interpreted as the responses to a succession of polls, or as the description of changes in voting intentions considering tactical voting strategies.

The players are traditionally assumed to know all the others’ vote. However, this assumption appears highly unrealistic, especially for political elections where the set of voters is actually very large. From this observation, an important literature begins to develop in order to face with uncertainty in voting. Several approaches can be enumerated: the Bayesian approach assuming that the voters think in terms of probabilities over the possible voting profiles (Myerson and Weber 1993; Hazon et al. 2008), the knowledge-based approach via modal logic frameworks (Chopra, Pacuit, and Parikh 2004; Van Ditmarsch, Lang, and Saffidine 2012), the local-dominance approach with uncertainty thresholds (Meir, Lev,

and Rosenschein 2014; Meir 2015), the information function approach which assumes that only a given type of information is communicated to the voters (Endriss et al. 2016; Reijngoud and Endriss 2012), or the partial orders approach where the voters consider a set of possible profiles according to partial votes (Conitzer, Walsh, and Xia 2011; Dey, Misra, and Narahari 2016). The most part of these works considers a risk-averse behavior, in the sense that the voters manipulate if they are sure that their deviation does not produce a worse outcome. Even if it appears as a rational behavior, it does not really capture the behavior of real voters who might actually manipulate even if they are not guaranteed to obtain at least a better outcome in any case.

We argue that the voters adopt an elementary behavior regarding the information they dispose of. Inspired by political elections, we assume that the voters have two sources of information: the *public opinion polls* and the *social networks*. Opinion polls punctuate the election campaign in many countries, and represent an important part of the election process. Many studies aim at understanding how they affect voting (Brams 1982; Forsythe et al. 1993; Fredén 2017). Concerning social networks (Easley and Kleinberg 2010; Jackson 2008), they occupy an increasingly important place in our lives. For a particular citizen, they constitute a tool to have an idea of the population opinions, even if this vision can be biased. The social networks are a natural channel for acquiring information in context of uncertainty in voting (Chopra, Pacuit, and Parikh 2004; Clough 2007; Sina et al. 2015; Tsang and Larson 2016).

In a context where the polls offer a large part of the information available to the voters, a natural question is: what happens if the polling institute does not say the truth about the scores of the candidates? In fact, can the polling institute manipulate the election? This question is related to the problem of election control (Faliszewski and Rothe 2016), i.e., how an external agent can influence and manipulate the election. While usually this agent manipulates by adding / deleting candidates or votes, or even links of a social network (Sina et al. 2015), we suppose that the polling institute may lie during the communication of the results.

We investigate an *iterative voting* model where the voters trust the results of an opinion poll and have all the same prior assumption about the distribution of the votes, which is given by the poll. The voters update this belief about the

vote intentions of the society by observing the strategic deviations of their relatives in the social network. The strategic behavior of voters is then conditioned by personal *pivotal thresholds* modeling their willingness to deviate and how strong is their belief, in a same way as *non-myopic voters* (Obraztsova et al. 2016), or as in *local-dominance* (Meir, Lev, and Rosenschein 2014). All these parameters enable to define a new best-response dynamics, called *poll-confident*. Two configurations are investigated: a *local* dynamics with one initial poll, and a *global* dynamics where the results of several polls are announced along the election. Convergence results are provided, stating that local dynamics are guaranteed to converge for some classes of graphs representing the social network. However, in general, it is computationally hard to recognize convergent instances for both dynamics. Experiments are given for testing the practical convergence of the dynamics and the quality of their outcomes. Then, we study the computational complexity of manipulation from the polling institute, proving that it is hard even for simple graphs. Beyond the general hardness of the problem, we provide simple heuristics that are efficient in practice.

We present in the first section our model and the poll-confident dynamics in their local and global version. Then we analyze their convergence, before finally investigating the manipulation from the polling institute.

Poll-Confident Iterative Voting Model

Notations

Let $N = \{1, \dots, n\}$ be a set of n agents (or voters), and $M = \{x_1, \dots, x_m\}$ a set of m candidates. Each voter i has a strict preference order \succ_i over M . The whole preference profile is denoted by $\succ = (\succ_1, \dots, \succ_n)$. A voting rule computes the winner of the election. We consider single-winner elections and use, in case of ties, a deterministic tie-breaking rule, based on a linear order \triangleright over the candidates.

We focus on the Plurality rule, and we denote Plurality associated with \triangleright by f . Plurality corresponds to a mapping $f : M^n \rightarrow M$, where each voter is asked to submit a ballot $\mathbf{b}_i \in M$, meant to represent her preferred candidate. Given a profile of ballots $\mathbf{b} \in M^n$, the score of each candidate x under Plurality is given by $\mathbf{s}_{\mathbf{b}}(x) = |\{i \in N : \mathbf{b}_i = x\}|$. The winner of Plurality in voting profile \mathbf{b} maximizes score $\mathbf{s}_{\mathbf{b}}$, i.e., $f(\mathbf{b}) \in \arg \max_{x \in M} \mathbf{s}_{\mathbf{b}}(x)$. By simplicity, we sometimes write $f(s)$ to designate the Plurality winner of a voting profile having s as a vector of scores.

We consider a *strategic game* where at each step, a single voter can change her vote if she thinks that she will be better off with this new strategy. The strategy profile at step t is denoted by \mathbf{b}^t . If different voters can deviate at the same step t , the deviator is arbitrarily chosen among them, unless we specify a particular *turn function* τ as part of the instance.

An *opinion poll* is undertaken from the initial voting profile \mathbf{b}^0 , where the voters are assumed to give their real preferences (the agents do not have enough information to manipulate yet), i.e., \mathbf{b}^0 is truthful. The result Δ^0 of this initial poll is communicated to the agents, via a vector of scores describing the Plurality score of each candidate.

Though the agents are aware of the initial poll, they do

not know all the deviations of the agents. We suppose that the agents are embedded into a *social network* $G = (N, E)$, which is a directed graph over the agents. We denote by $G[N']$ the subgraph induced by $N' \subseteq N$. We always care about the direction in the graph: a complete graph contains all the arcs, a clique refers to a complete subgraph, and a cycle to a closed directed path. The graph is *transitive* if the binary relation over N represented by E is transitive. An arc $(i, j) \in E$ means that agent i can observe the ballot of agent j at any time. We denote by $\Gamma(i)$ the set of agents that agent i can observe, i.e., $\Gamma(i) = \{j \in N : (i, j) \in E\} \cup \{i\}$. For a voting profile \mathbf{b} , the score of candidate x that agent i is able to observe is denoted by $\mathbf{s}_{\mathbf{b}}^i(x) = |\{j \in \Gamma(i) : \mathbf{b}_j = x\}|$. An instance of the game is a tuple $\mathcal{I} = (N, M, \succ, \mathbf{b}^0, G, \{p_i\}_{i \in N})$, where $\{p_i\}_{i \in N}$ are integer thresholds defining the strategic type of the voters.

Poll-confident dynamics

The voters have two sources of information about the current scores: the results of the initial poll and the current votes of their relatives, i.e., their direct successors in the network. The question is how they aggregate these informations. We assume that the voters base their belief on the results of the poll, updated with the votes of their relatives. The strength of this belief is conditioned by personal pivotal thresholds.

In fact, each voter has her own belief about the score of the candidates. We denote by $B_i^t : M \rightarrow \mathbb{N}$ the *believed score function* of agent i at the t^{th} step of the game. Initially, $B_i^0 = \Delta^0$ for all voters i since the voters trust the poll. We may see B_i^t and Δ^0 as m -tuples where the j^{th} coordinate represents the believed and the announced score of candidate x_j , respectively. At a step t of the game, the believed score function of the agents is updated only if a relative deviates.

Definition 1 (Score Belief Update). *The update of the believed score function of agent i at step $t + 1$, after the deviation of an agent j from candidate x to candidate y at step t is given by $B_i^{t+1} := B_i^t \otimes (j, x, y)$, where*

$$[B_i^t \otimes (j, x, y)](z) = \begin{cases} B_i^t(z) - 1 & \text{if } z = x \text{ and } j \in \Gamma(i) \\ B_i^t(z) + 1 & \text{if } z = y \text{ and } j \in \Gamma(i) \\ B_i^t(z) & \text{otherwise} \end{cases}$$

Note that for any voter i and any step t , $B_i^t = \Delta^0$ if $\Gamma(i) = \{i\}$ and i did not deviate, and $B_i^t = \mathbf{s}_{\mathbf{b}^t}$ if $\Gamma(i) = N$.

We define a *best response* for agent i at step t , w.r.t. her belief and her strategic behavior. A *pivotal threshold* p_i is associated with each voter i ; we denote it by p when p_i is the same for all the voters. The pivotal threshold of a voter represents her willingness to deviate to a new strategy on the basis of her belief about the scores. This can measure how much the voters are uncertain about their belief or represent the threshold from which they think that their vote is pivotal and matters in the election, notably by influencing the other votes. It also enables to model “sincere” voters, averse to manipulation. This threshold defines for each voter i a set of potential winners at step t , denoted by PW_i^t , representing the candidates that are still able to win the election, according to voter i . The believed winner of agent i at step t w.r.t. B_i^t is denoted by ω_i^t . We denote by B_i^{t, ω_i^t} the believed score

function of voter i at step t where the current vote of agent i is not taken into account, and by $\omega_i^{t,\setminus i}$ its associated winner.

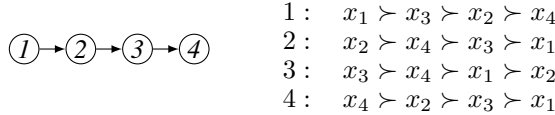
Definition 2 (Potential winner). *A candidate x belongs to PW_i^t for voter i at step t if, without considering the current vote of i , at most p_i votes are necessary to x for winning the election, i.e., $B_i^{t,\setminus i}(\omega_i^{t,\setminus i}) - B_i^{t,\setminus i}(x) + \mathbb{1}_{\{\omega_i^{t,\setminus i} \triangleright x\}} \leq p_i$.*

A best response at step t for voter i is defined as the candidate that i prefers among the potential winners $PW_i^t \setminus \{\omega_i^t\}$.

Definition 3 (Best response). *A ballot approving candidate y is a best response for voter i at step t if $y \in PW_i^t \setminus \{\omega_i^t\}$ and for all $x \in PW_i^t$ such that $x \neq y$, $y \succ_i x$.*

A strategy profile is a Δ^0 -equilibrium iff for every agent there is no best response or the best response corresponds to her current vote. Observe that the classical iterative voting setting, where the current votes are common knowledge and the voters consider that they are pivotal if they are able to change the outcome, corresponds in our model to an instance where the social network is a complete graph and $p = 1$. In such a context, our formulated best response is a “direct” best response (Meir et al. 2010), and a Δ^0 -equilibrium is a Nash equilibrium, i.e., a stable state regarding unilateral deviations with complete information about the current votes (Aumann and Brandenburger 2014). When the social network is a complete graph, our best response corresponds to an NM-Plurality response (Obraztsova et al. 2016) and to a local-dominant strategy (Meir, Lev, and Rosenschein 2014) if \mathbf{b}^0 is truthful, where the distance metric between voting profiles is an ℓ_1 -norm. A sequence of best response deviations is illustrated in the following example.

Example 1. *Consider an instance where $n = m = 4$, with $x_1 \triangleright x_2 \triangleright x_3 \triangleright x_4$, $p_1 = 2$ and $p_2 = p_3 = p_4 = 1$. The social network G and the preferences \succ are as follows.*



Consider the following sequence of best responses. Each state \mathbf{b}^t is represented by $\mathbf{s}_{\mathbf{b}^t}$ and $f(\mathbf{b}^t)$. At each step t , an arrow designates the deviation performed by voter i mentioned below the arrow, believing the score function $B_i^{t,\setminus i}$ mentioned above the arrow. The deviations are in bold.

$$(1, 1, 1, 1):x_1 \xrightarrow[2]{(1,0,1,1)} (1, 0, 1, \mathbf{2}):x_4 \xrightarrow[3]{(1,1,0,1)} (1, 0, 0, \mathbf{3}):x_4 \xrightarrow[4]{(1,1,1,0)} (1, \mathbf{1}, 0, 2):x_4 \xrightarrow[3]{(1,2,0,0)} (\mathbf{2}, 1, 0, 1):x_1$$

From \mathbf{b}^0 , agent 2 changes her vote to candidate x_4 because she prefers x_4 to the winner x_1 . Agent 1 observes this move but has a pivotal threshold of 2: she still believes that x_1 can win so her best response is her current vote. Agent 3 did not observe any move, and thus believes in the initial scores, leading to her deviation to x_4 . This belief also holds for agent 4 who deviates to x_2 . Agent 3 observes this move while believing that agent 2 has not deviated and so she moves to x_1 . Now, no agent can deviate given the information she gets, so profile $\mathbf{b}^4 := (x_1, x_4, x_1, x_2)$ is an equilibrium.

The information contained in the poll, updated by the observations given by the network, as well as the pivotal

thresholds characterizing the best responses, enable to define a dynamics of the strategic game, that we call *poll-confident dynamics*. We say that the poll-confident dynamics converges if any sequence of deviations leads to a Δ^0 -equilibrium, and we say that the dynamics can cycle if there exists a sequence of deviations where we come back to a previous strategy profile. Such a poll-confident dynamics, defined according to one initial poll, is said to be *local*.

Global poll-confident dynamics

We define, analogously to the local poll-confident dynamics, a *global poll-confident dynamics* where several polls are communicated to the voters. In fact, each time the (local) poll-confident dynamics converges, the scores at the equilibrium are communicated via a poll. A sequence of deviations within global dynamics is a sequence of states $(\mathbf{b}^0, \mathbf{b}^{0,1}, \dots, \mathbf{b}^{0,t_0}, \mathbf{b}^{1,1}, \dots, \mathbf{b}^{1,t_1}, \dots, \mathbf{b}^{t,1}, \dots, \mathbf{b}^{t,t_t})$, where $\mathbf{b}^{i,j}$ is the strategy profile at i^{th} global step and j^{th} local step. For each $0 \leq k \leq t-1$, there is a poll Δ^{k+1} between state \mathbf{b}^{k,t_k} and state $\mathbf{b}^{k+1,1}$ which announces the scores of \mathbf{b}^{k,t_k} . Initial poll Δ^0 is given between initial state \mathbf{b}^0 and state $\mathbf{b}^{0,1}$. For all $0 \leq k \leq t$, sequence $(\mathbf{b}^{k,1}, \dots, \mathbf{b}^{k,t_k})$ represents the deviations of the local dynamics with initial poll Δ^k , which leads to the Δ^k -equilibrium \mathbf{b}^{k,t_k} . We assume that the voters do not keep in memory the history of the previous global steps. At each global step, the voters behave in the same way as in the initial one.

The global poll-confident dynamics can cycle if there exists a sequence of deviations $(\mathbf{b}^0, \mathbf{b}^{0,1}, \dots, \mathbf{b}^{0,t_0}, \dots, \mathbf{b}^{t,1}, \dots, \mathbf{b}^{t,t_t})$ such that there are $k < k'$ where $\mathbf{b}^{k',t_{k'}} = \mathbf{b}^{k,t_k}$ or $\mathbf{b}^{k',t_{k'}} = \mathbf{b}^0$, or if the associated local poll-confident dynamics cycles. A stable state regarding the global dynamics is called a *global equilibrium*.

Observation 1. *For a unanimous pivotal threshold $p = 1$, a state \mathbf{b} is a global equilibrium iff \mathbf{b} is a Nash equilibrium.*

A global equilibrium \mathbf{b} is either \mathbf{b}^0 or a Δ^t -equilibrium for some step t . In any case, its scores are announced either by Δ^0 or by Δ^{t+1} , leading to complete information about \mathbf{b} .

Convergence to Poll Equilibria

In this section, we study the convergence of the poll-confident dynamics and the quality of the equilibria. Despite some positive results, in the general case, it is difficult to recognize the instances for which the dynamics converges.

Convergence properties

The deviation of each agent depends on the information she has about the game. If she sees nobody deviating after her deviation, then a voter has no reason to deviate again.

Observation 2. *Any agent i such that $\Gamma(i) = \{i\}$ deviates at most once for any voting rule and any threshold p_i .*

Obs. 2 implies that the local dynamics converges for any voting rule if E is empty. A condition on the graph can be further derived to ensure the convergence of the dynamics.

Proposition 1. *If G is a directed acyclic graph (DAG), then the local dynamics converges from any initial state within $\mathcal{O}(n^2)$ steps, for any voting rule and any thresholds p_i .*

Proof. Suppose the dynamics cycles within the subset of agents $N' \subseteq N$. Obs. 2 implies that the outdegree of each agent i in N' in G is not equal to zero. Yet, a DAG contains at least one sink, i.e., a vertex of null outdegree, and a sub-graph of a DAG is still a DAG. Contradiction. A deviation of a sink can push at most $n - 1$ other agents to deviate. By iteratively removing a sink of G and counting the deviations caused by a sink in the new graph, we get that the sequence of deviations is of size $\mathcal{O}(n^2)$. \square

If the network does not contain cycles, then convergence is ensured. Actually, this is especially the presence of cycles in the graph that enables the dynamics to cycle. However, we can prevent cycles in the dynamics to occur if the cycles of the graph are in fact cliques, i.e., the network is transitive, and the dynamics converges in a complete graph.

Proposition 2. *If the local dynamics eventually converges from any initial state after $g(n, m)$ steps for a complete graph, then the dynamics converges from any initial state within $\mathcal{O}(n \cdot g(n, m))$ steps when the network is transitive.*

Proof. Let N' be a minimal subset of agents for who the dynamics cycles. All the agents i in a clique C of G have the same believed scores B_i^t at any step t : by transitivity, $\Gamma(i) = \Gamma(j)$ for all $i, j \in C$. So, since the dynamics is guaranteed to converge when all the current votes are known, $G[N']$ is not a clique. By Prop. 1, there must be a cycle in $G[N']$. But a cycle along agents $N'' \subseteq N$ in a transitive graph implies that $G[N'']$ is a clique. So, $G[N']$ includes a set of cliques. Since there is no visibility of the deviations between two disjoint connected components, $G[N']$ is connected. The agents in N' can be partitioned into the groups N'_1, \dots, N'_k w.r.t. their level of knowledge: each $G[N'_i]$ for $i \in [k]$ is a set of disjoint cliques and any agent in N'_i can only observe some agents in N'_j for $j \leq i$. Each agent in a group N'_i deviates according to the deviations of agents in $\bigcup_{j \leq i} N'_j$, because she cannot observe other deviations. So, a cycle in the dynamics within N' implies a cycle in the dynamics within $\bigcup_{j \leq i} N'_j$. By minimality of N' , N' is a single clique, contradiction. It is possible to see a transitive graph as a DAG whose nodes are cliques. So, any sequence of deviations is of length $\mathcal{O}(n \cdot g(n, m))$. \square

When the thresholds are heterogeneous, the dynamics may cycle even for a complete graph (Obraztsova et al.[2016]’s Thm. 3). However, from Meir et al.[2010]’s Thm. 3, the local dynamics converges from any initial state within $\mathcal{O}(nm)$ steps when $p = 1$ and the graph is complete.

Corollary 1. *The local dynamics is guaranteed to converge from any initial state within $\mathcal{O}(n^2m)$ steps when $p = 1$ and the graph is transitive.*

Moreover, this can be generalized to unanimous pivotal thresholds under some condition on the turn function τ , in the spirit of (Meir, Lev, and Rosenschein 2014)’s Prop. 6. A deviation of agent i from x to y is called a (1)-compromise move if $x \succ_i y$ and a (2)-opportunity move if $y \succ_i x$.

Proposition 3. *If the turn function τ always selects (2)-moves before (1)-moves, then the local dynamics converges*

within $\mathcal{O}(nm)$ steps from any initial state when p is unanimous and the graph is complete.

Proof. A (2)-move implies that the rank, in the deviator’s preferences, of the candidate approved in the new ballot is strictly better than the previous approved candidate. So, there are at most $n(m - 1)$ consecutive (2)-moves. By definition of τ , there exists a state t_0 where no agent has incentive to perform a (2)-move. One can prove by a complete induction that, for any step $t \geq t_0$, where an agent i deviates from x to y , the following holds: (I) the score of the winner does not decrease, (II) $\bigcup_{j \in N} PW_j^{t+1} \subseteq \bigcup_{j \in N} PW_j^t$, (III) $x \succ_i y$, and (IV) after this move, no agent can deviate to x . The details are omitted due to space limitations. \square

Corollary 2. *If τ selects (2)-moves before (1)-moves, then the local dynamics converges within $\mathcal{O}(n^2m)$ steps from any initial state when p is unanimous and the graph is transitive.*

However, for a general graph, knowing whether the local dynamics can cycle is difficult, even when $p = 1$.

Theorem 1. *Knowing whether the local dynamics can cycle is NP-hard, even for unanimous pivotal threshold $p = 1$.*

Proof. We perform a reduction from the NP-complete problem 2P1N-SAT (Yoshinaka 2005) defined as follows: given a set $\mathcal{C} = \{C_1, \dots, C_r\}$ of r clauses over a set $X = \{x_1, \dots, x_v\}$ of v variables, where each variable occurs twice as a positive literal and once as a negative literal, is \mathcal{C} satisfiable? The clauses can be indexed such that each first occurrence of a variable is a positive literal, because the original proof of NP-hardness relies on such an instance. We assume such a restriction. Each clause C_i contains r_i literals.

We construct an instance $\mathcal{I} = (N, M, \succ, \mathbf{b}^0, G, p = 1)$. The set M of candidates includes candidates a, d, y, z and clause-candidates c_i for each $i \in [r]$. The tie-breaking \triangleright is defined as: $d \triangleright y \triangleright z \triangleright c_r \triangleright c_{r-1} \triangleright \dots \triangleright c_1 \triangleright a$. The set N of agents includes agents A, D, Y and Z , and literal-agents L_j^i for each j^{th} literal of clause C_i . The preferences are as follows (all candidates not listed are ranked in arbitrary order within [...]):

$$\begin{aligned} L_j^i: & y \succ c_i \succ c_r \succ c_{i-1} \succ [\dots] \\ A: & a \succ y \succ d \succ [\dots] & Y: & y \succ z \succ d \succ [\dots] \\ D: & a \succ d \succ y \succ [\dots] & Z: & z \succ y \succ c_r \succ [\dots] \end{aligned}$$

We add $3v(r + 2) + r$ dummy voters, where $3v + 1$ among them rank c_i first for each candidate c_i , $3v$ rank d first and $3v$ rank z first. Every candidate, except a and d , obtains $3v + 1$ votes, and d has $3v$ votes. The initial winner is y , thanks to \triangleright . In the network G , all the agents of a same clause form a clique. There is an arc from each agent of clause C_i to each agent of clause C_{i-1} . There is an arc from each agent of clause C_1 to agents D and Z , and from agents Z and Y to the agents of clause C_r . Agent Z points to A and Y , and A points to D . There is an arc from agent L_j^i to agent L_k^ℓ if $i < \ell$ and the agents correspond to opposite literals. The graph construction is illustrated in Fig. 1.

We claim that \mathcal{C} is satisfiable iff the local dynamics can cycle in \mathcal{I} . One can prove that the only possible cycle within the dynamics involves the agents Y, Z and exactly one agent

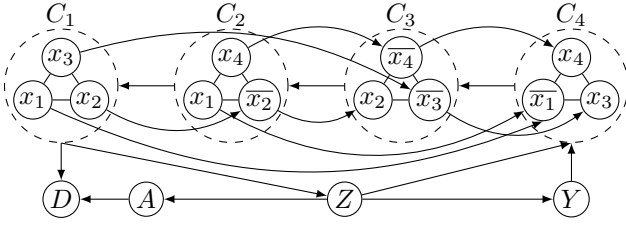


Figure 1: Construction of G for a 2P1N-SAT instance where $C_1 = (x_1 \vee x_2 \vee x_3)$, $C_2 = (x_1 \vee x_2 \vee x_4)$, $C_3 = (x_2 \vee x_3 \vee x_4)$ and $C_4 = (\bar{x}_1 \vee x_3 \vee x_4)$. The literals (we keep for the figure the literal name) in a same circle are in a same clause; the arcs from or to the circles concern every agent inside.

related to each clause. Moreover, if there are two agents corresponding to opposite literals among the deviating agents related to the clauses, then the cycle does not occur. The proof of the equivalence is omitted by lack of space. \square

Concerning the global poll-confident dynamics, they are not guaranteed to converge, even when the associated local dynamics always converges and for pivotal threshold $p = 1$.

Proposition 4. *The global dynamics may cycle even when G is empty and for unanimous pivotal threshold $p = 1$.*

Proof. Consider an instance where $n = 4$ and $m = 3$, with $x_1 \triangleright x_2 \triangleright x_3$, $p = 1$, an empty graph and the preferences: 1: $x_2 \succ x_1 \succ x_3$, 2: $x_3 \succ x_1 \succ x_2$, 3: $x_1 \succ x_2 \succ x_3$ and 4: $x_3 \succ x_1 \succ x_2$. Here is a cycle within 2 global steps (the polls are given in bold at the beginning of the lines):

$$[1,1,2] \quad (1, 1, 2):x_3 \xrightarrow{1} (2, 0, 2):x_1 \xrightarrow{3} (1, 1, 2):x_3$$

$$[1,1,2] \quad (1, 1, 2):x_3 \xrightarrow{1} (0, 2, 2):x_2 \xrightarrow{3} (1, 1, 2):x_3$$

In this case, the cycle mimics simultaneous moves. \square

For a general graph, we can prove that it is NP-hard to know whether the global dynamics can cycle, even for unanimous pivotal threshold $p = 1$, through another reduction from 2P1N-SAT. We omit this proposition by lack of space.

Experiments on the quality of poll equilibria

To conclude this section, we present some experiments over 10,000 generated instances with 100 voters and 10 candidates, under impartial culture for the preferences and random Erdős and Rényi[1959]’s graphs (see Fig. 2). We observe the frequency of convergence of the dynamics and the frequency of electing a Condorcet winner (for Condorcet domains) within the equilibria. We stop the iterative process only when an equilibrium is reached or a cycle is hit. The results are given w.r.t. the density of the graphs and the pivotal thresholds. We examine unanimous thresholds of value 1, 5 and 10 and heterogeneous ones uniformly distributed over the voters with values in $[1..5]$ or $[1..10]$.

The dynamics, both in their local and global version, mostly converge, especially for sparse or dense graphs. When $p = 1$ or $p = 10$, they almost always converge, contrary to the other thresholds for which we observe less

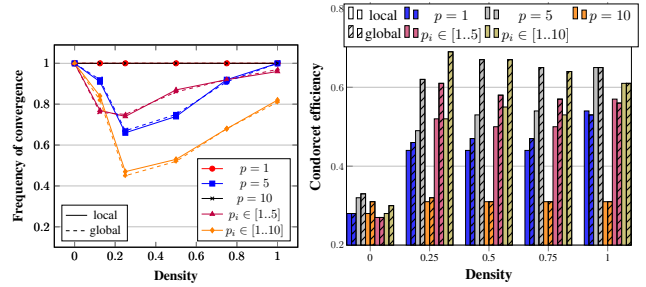


Figure 2: Frequency of convergence and Condorcet efficiency for local/global dynamics under different thresholds

convergent profiles, in particular when the density is around 0.25. This can be explained by the fact that when $p = 10$, most voters stay with their truthful ballot, and when $p = 1$ only a few candidates are potential winners. This is not the case for heterogeneous thresholds and $p = 5$, because the agents can have very various best strategies. Concerning the Condorcet efficiency, we observe a slight increase when the graph is denser, which can be explained by the gain of information. Moreover, the quality of the outcome is better for global dynamics, which appears natural since the voters regularly obtain more information about the current profile.

We have conducted the same experiments for graphs closer to real social networks, such as scale-free networks or networks with homophily where the agents are linked w.r.t. their preferences. However, it seems that the poll-confident dynamics behave as in random graphs of corresponding density. This highlights the role of the quantity of information possessed by the voters in the quality of poll equilibria.

Manipulation of the Opinion Poll

In our model, the voters base their belief on the results of the poll. If we consider the polling institute as an agent who has her own preferences over the candidates, then the question of manipulation from the polling institute naturally arises. The polling institute as an agent is denoted by π and her preferences are expressed via a linear order \succ_π over the candidates. We assume in this section that all the voters have the same pivotal threshold p .

Concretely, the manipulations of π must be restricted in order to satisfy some likelihood conditions. For instance, π could not announce that a candidate has no point if at least one voter has voted for it, otherwise this voter would know that the polling institute is lying. This credibility requirement is described more generally in the following definition.

Definition 4 (Likelihood condition). *The vector of scores Δ is a plausible result to the poll under strategy profile \mathbf{b} if $\Delta(x) \geq \max_{i \in N} s_b^i(x)$, for every candidate x .*

We assume that any manipulation performed by π satisfies the likelihood condition. Let the *manipulation margin* be the number of points which are available after having fulfilled the likelihood condition. We firstly ask whether it is possible for institute π to enforce the election of a given candidate x .

Definition 5 (ELECTION ENFORCING). *Given candidate x*

and instance \mathcal{I} , is there a poll score vector Δ such that the local dynamics converges to a Δ -equilibrium electing x ?

This problem is computationally hard even when the network is a DAG. We denote by $PW(\Delta)$ the set of potential winners announced by Δ , w.r.t. the unanimous pivotal threshold p . The set $PW(\Delta)$ includes all the candidates for which the addition of p votes approving them is sufficient to win the election, according to Δ .

Theorem 2. ELECTION ENFORCING is NP-hard even when the social network is a DAG and $p = 1$.

Proof. We reduce from 3-SAT and consider an instance of r 3-clauses $\{C_1, \dots, C_r\}$ and v variables $\{x_1, \dots, x_v\}$.

We construct a set of agents N including: a set Y of r clause-agents Y_i for $i \in [r]$, sets Y' and L of $3r$ literal-agents Y'_{ij} and L_{ij} for $i \in [r]$ and $j \in [3]$, a set L' of agents L'_{ijk} for $i \in [r]$, $j \in [3]$ and $k \in [3]$ referring to 3 copies of the literals, sets Z^1 and Z^2 (whose union is denoted by Z) of respective agents Z^1_i and Z^2_j for $i \in [4r-1]$ and $j \in [11]$, a set X of agents X_i for $i \in [10]$, and agents X'_1 and X'_2 . The set of candidates M contains candidates z, x' and x , the set Cl of clause-candidates c_i for $i \in [r]$, and the set Lit of literal-candidates ℓ_{ij} for $i \in [r]$ and $j \in [3]$. In graph G , there is an arc from Y_i to L_{ij} and from Y'_{ij} to L'_{ijk} , for all $i \in [r]$, $j, k \in [3]$. There is an arc from L_{ij} to $L'_{ij'}$ if $i > i'$ and the j^{th} literal of i^{th} clause is the opposite of the j'^{th} literal of i'^{th} . In the sets Z^2 and X , there is one agent, say respectively Z^2_1 and X_1 , who is connected to all the agents in the set. Let ρ be a linear order over $Cl \cup Lit$ and ρ^{-1} its reverse order. The preferences are as follows ($i+1 = 1$ if $i = r$):

$$\begin{array}{ll} Y_i : z \succ c_i \succ x \succ \rho & Z^2_j : z \succ \rho^{-1} \succ x \\ Y'_{ij} : z \succ \ell_{ij} \succ x \succ \rho^{-1} & X_i : x \succ z \succ \rho \\ L_{ij} : c_i \succ \ell_{ij} \succ z \succ \rho^{-1} \succ x & X'_1 : x' \succ z \succ \rho \\ L'_{ijk} : \ell_{ij} \succ \ell_{i+1k} \succ z \succ \rho \succ x & X'_2 : x' \succ z \succ \rho^{-1} \end{array}$$

where ρ and ρ^{-1} are assumed to be defined without the candidates already mentioned in the preference order.

The turn function τ is such that $\{L, L'\} \succ_{\tau} \{Y, Y'\}$, and $L_{ij} \succ_{\tau} L'_{ij'}$ for $i' > i$. The tie-breaking rule \triangleright is such that $x \triangleright z \triangleright \ell_{11} \triangleright \dots \triangleright \ell_{13} \triangleright \dots \triangleright \ell_{r1} \triangleright \dots \triangleright \ell_{r3}$, and $p = 1$.

We claim that all the clauses are satisfiable iff polling institute π can enforce the election of x in the local dynamics.

At b^0 , each L_{ij} and L'_{ijk} vote respectively for c_i and ℓ_{ij} , and agents Y_i and Y'_{ij} respectively observe it. So, by the likelihood condition, each candidate $y \in Cl \cup Lit$ must have at least 3 points in the poll, i.e., $\Delta(y) \geq 3$. Each agent in X votes for x and this is visible for X_1 , thus $\Delta(x) \geq 10$. All the sets Y, Y' and Z vote for candidate z , but only agent Z^2_1 has a non-null outdegree for observing that, and thus $\Delta(z) \geq 11$, whereas the real score of z is $8r + 10$. X'_1 and X'_2 vote for x' so $\Delta(x') \geq 1$. To summarize, the margin of manipulation for institute π is $8r$. We can prove that the only strategy for π to make x win is to assign 8 more points to one literal-candidate associated with each clause, and the new points cannot be given to literal-candidates corresponding to opposite literals. The winner announced in this poll is z . We omit the details of the equivalence, by lack of space. \square

Beyond the computational hardness of manipulating the poll in the worst-case, we explore a heuristic perspective. For all experiments, we run 10,000 instances of 100 voters and 10 candidates, where the preferences are generated in impartial culture, and the graphs are randomly generated via Erdős and Rényi[1959]’s model for different densities.

Algm. 1 constructs a poll score vector where, as much as possible, there are only two best candidates. The intuition is that, if π wants to make a target candidate x elected then, as much as the manipulation margin allows, she predicts as the winner a “specter” candidate y , that π wants to show as a threat to the voters, and x as a potential winner. More precisely, after having fulfilled the likelihood condition (l. 2-4), points are added to x until it becomes a potential winner (l. 5-6). Then, we rise the score of y until one more point to y would remove x from the set of potential winners (l. 7-8), in order to quickly increase the gap with other candidates while keeping x in $PW(\Delta)$. At this point, y is the winner in Δ and x is at the limit of $PW(\Delta)$. Now we simultaneously increase the scores of x and y (l. 9-11) in order to further eliminate other candidates from $PW(\Delta)$. If there remains one point at the end, we assign it to x if that does not make it the winner, or otherwise to the last ranked candidate in Δ by safety (l. 12-14). The process stops if at some step the manipulation margin is not sufficient.

Algorithm 1: Margin rebalance on two candidates

Input: \mathcal{I} , state \mathbf{b} , target $x \in M$, specter $y \in M$

Output: Δ : communicated scores of the poll from \mathbf{b}

```

1 margin  $\leftarrow n$ ;
2 foreach  $z \in M$  do
3    $\Delta(z) \leftarrow \max_{i \in N} \mathbf{s}_{\mathbf{b}}^i(z)$ ;
4   margin  $\leftarrow$  margin  $- \Delta(z)$ ;
5 while margin  $> 0$  and  $x \notin PW(\Delta)$  do
6    $\Delta(x) \leftarrow \Delta(x) + 1$ ; margin  $\leftarrow$  margin  $- 1$ ;
7 while margin  $> 0$  and  $x$  remains in  $PW(\Delta)$  if  $\Delta(y)$ 
   is increased by 1 do
8    $\Delta(y) \leftarrow \Delta(y) + 1$ ; margin  $\leftarrow$  margin  $- 1$ ;
9 while margin  $> 1$  do
10   $\Delta(x) \leftarrow \Delta(x) + 1$ ;  $\Delta(y) \leftarrow \Delta(y) + 1$ ;
11  margin  $\leftarrow$  margin  $- 2$ ;
12 if margin = 1 then
13   if  $p > 1$  then  $\Delta(x) \leftarrow \Delta(x) + 1$ ;
14   else  $\ell \leftarrow \arg \min_{z \in M} \Delta(z)$ ;  $\Delta(\ell) \leftarrow \Delta(\ell) + 1$ ;
15 return  $\Delta$ ;
```

Algm. 2 is a heuristic based on Algm. 1 where we choose, as a specter candidate y , the candidate that is the most disliked compared to target candidate x . We aim at creating a situation where only two candidates, x and y , are favorites in the election. One of them, the “specter” y , is mostly disliked by the population but is announced the winner with a slight lead over x . One could think that, in reaction, a large part of the electorate will report her ballot to the other candidate x .

We present experiments in Fig. 3, where the targets are the Condorcet winner (for Condorcet domains), the Borda

Algorithm 2: Heuristic for ELECTION ENFORCING

Input: \mathcal{I} , state \mathbf{b} , target $x \in M$ **Output:** Δ : communicated scores of the poll from \mathbf{b}

- 1 Sort the candidates $M \setminus \{x\}$ in decreasing order of the number of voters who prefer x to the candidate;
 - 2 **foreach** *specter* $y \in M \setminus \{x\}$ **do**
 - 3 $\Delta \leftarrow \text{Algm. 1}(\mathcal{I}, \mathbf{b}, x, y)$;
 - 4 **if** $f(\Delta) = y$ and $x \in PW(\Delta)$ **then return** Δ ;
 - 5 **return** $\text{Algm. 1}(\mathcal{I}, \mathbf{b}, x, \arg \max_{z \neq x} \max_{i \in N} s_{\mathbf{b}}^i(z))$;
-

winner, the truthful winner (the winner of \mathbf{b}^0) and the best candidate in \triangleright . The frequency of election of the target candidate is given in a context of poll manipulation via Algm. 2 or no manipulation. For the sparsest graphs, this frequency is at least twice higher with Algm. 2 than without manipulation. In global dynamics, there is manipulation from π at each global step, that is why the results are slightly better.

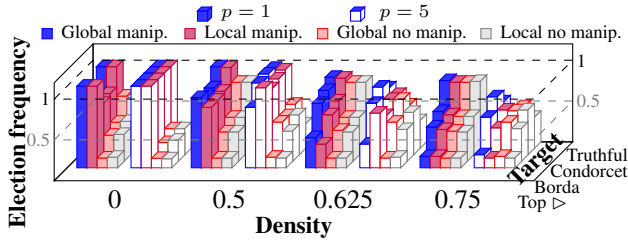


Figure 3: Algm. 2 in global/local dynamics ($p = 1$ or 5)

From a different perspective, instead of enforcing the election of a specific candidate, institute π could try to perform best responses at each global step, with immediate benefits in the associated local dynamics, in the same myopic spirit as a voter in iterative voting. In other words, π manipulates at each global step in order to make elected her best possible candidate at the end of the next local dynamics.

However, as Thm. 2 states, such a best response is hard to compute, even for a DAG. Consequently, we restrict the manipulation of π to a simple move exposed in Algm. 1: trying to favor only two candidates. We derive Algm. 3 where a manipulated poll score vector Δ is built by Algm. 1 for every couple of candidates (x, y) , and the associated sequence of local deviations is tested. We choose the poll score that leads to an equilibrium, in the tested local dynamics, electing the best candidate for π . In order to efficiently test the sequence of local deviations, we restrict to cases where the dynamics is guaranteed to converge after a polynomial number of steps, i.e., when the network is acyclic or transitive (Prop. 1 and cases of Cor. 1 and 2). We thus consider spanning subgraph G' of network G that is either empty, acyclic or transitive. Since Maximal Acyclic Subgraph and Maximum Transitive Subgraph are NP-complete problems, we use simple classical approximations for computing these subgraphs.

The results, presented in Fig. 4, are good for π . These results show the average of the rank in \succ_{π} of the final winner when the dynamics converges (lower is better). For both

Algorithm 3: Restricted poll manipulation move

Input: \mathcal{I} , turn function τ , preferences \succ_{π} , state \mathbf{b} , acyclic / transitive spanning subgraph G' of G **Output:** Δ : communicated scores of the poll from \mathbf{b}

- 1 **foreach** *target* $x \in M$ **do**
 - 2 **foreach** *specter* $y \in M \setminus \{x\}$ **do**
 - 3 $\Delta_y \leftarrow \text{Algm. 1}(\mathcal{I}, \mathbf{b}, x, y)$;
 - 4 $R(y) \leftarrow f(\Delta_y\text{-eq.})$ in \mathcal{I} with $G \leftarrow G'$ and τ ;
 - 5 **return** Δ_y for which $R(y)$ is the best in \succ_{π} ;
-

types of dynamics, the rank with manipulation is clearly better than without manipulation, with a large gap for sparse graphs. The gap decreases with the increase of the density. For global dynamics, the results are very good for π , especially for sparse graphs until density 0.4. Like Algm. 2, the results deteriorate with the increase of the density. This is related to the likelihood condition, making the manipulation margin decreasing with the increase of the density, until the complete graph, where no manipulation is possible for π since every agent has complete information.

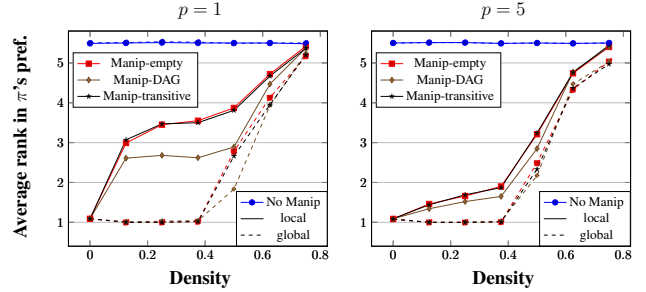


Figure 4: Algm. 3 in global/local dynamics ($p = 1$ or 5)

Conclusion

We have studied a best response dynamics where the voters aggregate the informations from opinion polls and social networks, and adopt a strategic behavior conditioned by pivotal thresholds. We showed the convergence of the dynamics for some classes of graphs but in general it is difficult to recognize instances with cycles. However, it turns out that practically, the dynamics mostly converges. The quality of the equilibria depends on the density of the network: better outcomes are found in dense graphs (there is more information). The equilibrium analysis allows to underline the bias produced by partial information and the dependency on the information sources, raising the question of election control.

Actually, manipulation of the poll can be hard to compute, even for simple sparse graphs. However, simple heuristics, based on the idea of announcing a “specter” candidate (that is mostly disliked) as winner, are very efficient. The manipulation is less beneficial in dense graphs where the knowledge of the voters is close to be complete. These heuristics are not too demanding regarding the network structure, but

they need to know the preferences and thresholds. Nevertheless, this is not a strong assumption that the polling institute knows the preferences of the voters since she collects them. Moreover, we have conducted experiments (not presented here) where the polling institute computes her strategy with thresholds that do not correspond to the real ones, and the results are similar to those presented in the article.

In addition to classical perspectives such as the study of other voting rules, coalitional manipulation and more sophisticated heuristics, one could think about voters who are not memory-less in the global dynamics or relaxing the assumption of a poll on the entire electorate. The links with opinion diffusion in networks could also be examined.

References

- Aumann, R., and Brandenburger, A. 2014. Epistemic conditions for Nash equilibrium. In *The Language of Game Theory: Putting Epistemics into the Mathematics of Games*. World Scientific. 113–136.
- Brams, S. J. 1982. Strategic information and voting behavior. *Society* 19(6):4–11.
- Chopra, S.; Pacuit, E.; and Parikh, R. 2004. Knowledge-theoretic properties of strategic voting. In *Proceedings of the 9th European Conference on Logics in Artificial Intelligence (JELIA-04)*, 18–30.
- Clough, E. 2007. Strategic voting under conditions of uncertainty: A re-evaluation of Duverger’s law. *British Journal of Political Science* 37(2):313–332.
- Conitzer, V.; Walsh, T.; and Xia, L. 2011. Dominating manipulations in voting with partial information. In *Proceedings of the 25th Conference on Artificial Intelligence (AAAI-2011)*, 638–643.
- Dey, P.; Misra, N.; and Narahari, Y. 2016. Complexity of manipulation with partial information in voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*, 229–235.
- Easley, D., and Kleinberg, J. 2010. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press.
- Endriss, U.; Obratsova, S.; Polukarov, M.; and Rosenschein, J. S. 2016. Strategic voting with incomplete information. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*, 236–242.
- Erdős, P., and Rényi, A. 1959. On random graphs I. *Publicationes Mathematicae (Debrecen)* 6:290–297.
- Faliszewski, P., and Rothe, J. 2016. Control and bribery in voting. In Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D., eds., *Handbook of Computational Social Choice*. Cambridge University Press. chapter 7, 146–168.
- Forsythe, R.; Myerson, R. B.; Rietz, T. A.; and Weber, R. J. 1993. An experiment on coordination in multi-candidate elections: The importance of polls and election histories. *Social Choice and Welfare* 10(3):223–247.
- Fredén, A. 2017. Opinion polls, coalition signals and strategic voting: Evidence from a survey experiment. *Scandinavian Political Studies* 40(3):247–264.
- Gibbard, A. 1973. Manipulation of voting schemes: a general result. *Econometrica* 41(4):587–601.
- Hazon, N.; Aumann, Y.; Kraus, S.; and Wooldridge, M. 2008. Evaluation of election outcomes under uncertainty. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-08)*, 959–966.
- Jackson, M. O. 2008. *Social and Economic Networks*. Princeton University Press.
- Meir, R.; Polukarov, M.; Rosenschein, J. S.; and Jennings, N. R. 2010. Convergence to equilibria in plurality voting. In *Proceedings of the 24th Conference on Artificial Intelligence (AAAI-10)*, 823–828.
- Meir, R.; Lev, O.; and Rosenschein, J. S. 2014. A local-dominance theory of voting equilibria. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC-14)*, 313–330.
- Meir, R. 2015. Plurality voting under uncertainty. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI-15)*, volume 15, 2103–2109.
- Meir, R. 2017. Iterative voting. In Endriss, U., ed., *Trends in Computational Social Choice*. AI Access. chapter 4, 69–86.
- Myerson, R. B., and Weber, R. J. 1993. A theory of voting equilibria. *American Political Science Review* 87(01):102–114.
- Obratsova, S.; Lev, O.; Polukarov, M.; Rabinovich, Z.; and Rosenschein, J. S. 2016. Nonmyopic voting dynamics: An optimistic approach. In *Proceedings of the 10th Multi-disciplinary Workshop on Advances in Preference Handling (MPREF-16)*.
- Reijngoud, A., and Endriss, U. 2012. Voter response to iterated poll information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-12)*, 635–644.
- Satterthwaite, M. A. 1975. Strategy-proofness and Arrow’s conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic theory* 10(2):187–217.
- Sina, S.; Hazon, N.; Hassidim, A.; and Kraus, S. 2015. Adapting the social network to affect elections. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-15)*, 705–713.
- Tsang, A., and Larson, K. 2016. The echo chamber: Strategic voting and homophily in social networks. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-16)*, 368–375.
- Van Ditmarsch, H.; Lang, J.; and Saffidine, A. 2012. Strategic voting and the logic of knowledge. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-12)*, 1247–1248.
- Yoshinaka, R. 2005. Higher-order matching in the linear lambda calculus in the absence of constants is NP-complete. In *Proceedings of the 16th International Conference on Rewriting Techniques and Applications (RTA-05)*, 235–249.