Deliberation as Evidence Disclosure: A Tale of Two Protocol Types

Julian Chingoma , **Adrian Haret**
ILLC, University of Amsterdam
{j.z.chingoma, a.haret}@uva.nl

Abstract

We study a model inspired by deliberative practice, in which agents selectively disclose evidence about a set of alternatives prior to taking a final decision on them. We are interested in whether such a process, when iterated to termination, results in the objectively best alternatives being selected—thereby lending support to the idea that groups can be wise even when their members communicate with each other. We find that, under certain restrictions on the relative amounts of evidence, together with the actions available to the agents, there exist deliberation protocols in each of the two families we look at (i.e., simultaneous and sequential) that offer desirable guarantees. Simulation results further complement this picture, by showing how the distribution of evidence among the agents influences parameters of interest, such as the outcome of the protocols and the number of rounds until termination.

1 Introduction

Arguments that groups can be wise, i.e., that they can be trusted to find an objectively correct answer, date back to Condorcet [Condorcet, 1785; Elkind and Slinko, 2016] and are typically probabilistic in nature. However, work in this tradition often operates under the assumption that agents are independent, and thereby precluded from communicating with, or otherwise influencing, each other. At the same time, there is much enthusiasm amongst proponents of deliberative democracy about the idea that deliberation (which would involve communication) prior to taking a decision leads to better outcomes [Elster, 1998; Fishkin and Laslett, 2003; Pivato, 2019]. There is a question, then, as to whether *some* form of structured information exchange can be conducive to accurate beliefs and, down the line, to correct decisions.

This is the question we address here. In our paper we put forward a formal model in which alternatives are supported by (objective, unassailable) evidence that can be accessed by agents, who rank alternatives based on the evidence and ultimately take a collective decision. We take it that evidence may be unevenly distributed among the population, and this leads to varied, possibly incorrect, beliefs. The corrective to an ill-informed opinion, in our model, is communication, here formalized as *evidence disclosure*: we assume agents are truth-seeking, hence both open to changing their beliefs on the quality of alternatives, by absorbing evidence disclosed by others, as well as willing to inform others by selective disclosure of evidence. Formally, this information exchange is modelled by a *deliberation protocol*, which we think of as a process that takes place in rounds and that consists of rules about how information is disclosed and processed by the agents. The process stops when no one can, or wishes to, do anything to further change the status quo.

We distinguish between two main types of deliberation protocols: *simultaneous*, in which evidence is disclosed and absorbed by all agents at the same time; and *sequential*, in which agents take turns in speaking and releasing evidence. Our model is not probabilistic, i.e., we do not touch upon the process by which agents acquire evidence in the first place. Rather, we are interested in whether deliberative practices, as sketched above, are conducive to good decisions regardless of the distribution of evidence. Our goal, then, is to find a sweet spot where features of this distribution and the protocol in use are guaranteed to engender accurate beliefs in the agents, ensuring that they elect the right alternative.

Contributions. We put forward an evidence-based election model fitted with a general mechanism for deliberation, in which agents iteratively update their rankings in response to evidence disclosure from participating agents. We look at two types of deliberation protocols, one *simultaneous* and two *sequential*, meant to approximate, in broad terms, the dynamics of information exchange in a real-world context (e.g., the discussion board of a paper reviewing platform, or a boardroom of people taking turns to speak), and at two agent types, *lazy* and *keen*, distinguished by their readiness to disclose information. We analyse these protocols with respect to termination time and, most importantly, the conditions under which they lead to the optimal (i.e., supported by most evidence) alternative being the election winner. The formal analysis is complemented by simulations, for a finer-grained picture of how the protocols fare on profiles of varying structure.

A sobering, though perhaps not surprising, finding is that in many cases the optimal alternative can lose out to inferior alternatives: this can happen when the distribution of evidence is unbalanced, allowing for small (but very vocal!) sets of supporters for non-optimal alternatives to sway the remaining electorate in the wrong direction. This effect can be mitigated if the optimal alternative is endowed with (much) more evidence than its competitors, or if, as simulations suggest, the distributions of evidence are relatively similar in the way evidence is spread. Remarkably, we also find that by limiting the amount of information agents can put forward, and carefully engineering the order in which agents speak, desirable guarantees can be given for the sequential protocols.

Related Work. Research in *epistemic social choice* seeks out conditions and methods for accurate group decisions [Elkind and Slinko, 2016; Dietrich and Spiekermann, 2021; Dietrich and Spiekermann, 2020; Condorcet, 1785; List, 2018], with a large proportion of the literature dedicated to the study of jury theorems under varied assumptions about the agents [Condorcet, 1785; Ladha, 1992; List and Goodin, 2001; Owen *et al.*, 1989; Dietrich and Spiekermann, 2013; Grofman *et al.*, 1983; Dietrich and Spiekermann, 2022; Pivato, 2017; Michelini *et al.*, 2022]. Alongside it there is a significant AI literature looking at the effect of opinion dynamics on collective opinions [Auletta *et al.*, 2015; Brill *et al.*, 2016; Auletta *et al.*, 2019]. The idea that deliberation boosts the the truth-tracking ability of groups has been argued extensively in the *deliberative democracy* literature [Elster, 1998; Fishkin and Laslett, 2003; Landemore, 2013; Bächtiger *et al.*, 2018; Goodin and Spiekermann, 2018; Hartmann and Rafiee Rad, 2018], complemented with formal models aiming to untangle the effects of deliberation [List, 2007; Perote-Peña and Piggins, 2015; Fain *et al.*, 2017; Chung and Duggan, 2020; Ding and Pivato, 2021]. To the best of our knowledge, the model closest to our own is that of Ding and Pivato [2021], whose *parallel protocol* over binary decisions is similar to our simultaneous protocol. The dynamics of our model bears resemblance to that seen in iterative voting [Meir, 2017; Lev and Rosenschein, 2012], though the preference formation mechanism in our case is different.

Outline. In Section 2 we present our evidence-based deliberation model; in Section 3 we define the specific deliberation protocols we focus on (the impatient reader can jump to Example 1 to see the protocols in action); Section 4 outlines the main results; in Section 5 we provide simulation results; we conclude in Section 6.

2 The Model

We work with a finite set $A = \{a, b, c, \ldots\}$ of m *alternatives* and a finite set E of *global evidence*. Every $x \in A$ is associated with a finite set $E(x) \subseteq E$ of *evidence for x*. We make no assumption on any $e \in E(x)$ other than to say that e *supports* x. For simplicity, we require that every evidence item supports exactly one alternative, i.e., the family of sets $\{E(x)\}_{x \in A}$ forms a partition of E. Evidence, in our model, is objective and induces a ground-truth ranking $\succcurlyeq$ over A given by the amount of evidence supporting each alternative. We say that *x is (objectively) at least as good as y* if $|E(x)| \geq |E(y)|$. We typically assume that there is a unique *optimal* alternative, i.e., an alternative $x \in A$ such that $|E(x)| > |E(y)|$, for all $y \in A$.

The evidence in E is held among a finite set $N = \{1, \ldots, n\}$ of *agents*, with $n \geqslant 3$, and across a sequence of discrete time steps indexed by $t \in \mathbb{N}$. We write $E_i^t(x)$ for the set of evidence that agent $i \in N$ has for alternative $x \in A$ at time $t \in \mathbb{N}$, with $E_i^t(x) \subseteq E(x)$. Each agent i uses the evidence at their disposal to form an *evidence order* $\succcurlyeq_i^t$ over A, where $x \succcurlyeq_i^t y$ holds if $|E_i^t(x)| \geqslant |E_i^t(y)|$, for any $x, y \in A$. Intuitively, agent i thinks x is *at least as good as y* at t if there is at least as much evidence supporting x as there is supporting y at t. We say agent i is *neutral* with respect to x and y at t, denoted $x \sim_i^t y$, if $x \succcurlyeq_i^t y$ and $y \succcurlyeq_i^t x$, and thinks x is *strictly better* than y, denoted $x \succ_i^t y$, if $x \succcurlyeq_i^t y$ and $y \not\succcurlyeq_i^t x$. The set $top(\succcurlyeq_i^t) = \{x \in A \mid x \succcurlyeq_i^t y \text{ for all } y \neq x \in A\}$ of *top alternatives of $\succcurlyeq_i^t$* consists of alternatives that agent i believes to be objectively best at t.

A *profile* $\succcurlyeq^t = (\succcurlyeq_1^t, \ldots, \succcurlyeq_n^t)$ at time t collects all evidence orders of agents in N at t. If $\boldsymbol{\alpha}$ is a list of alternatives in A (of arbitrary length), *the plurality winners $F_{PL}(\boldsymbol{\alpha}) \subseteq A$* is the set of alternatives that show up most often in $\boldsymbol{\alpha}$. A plausible example for $\boldsymbol{\alpha}$ is the list of top-ranked alternatives in $\succcurlyeq^t$, and we use $F_{PL}(\succcurlyeq^t)$ to denote the plurality winners of $\succcurlyeq^t$.

Presented with a set $C \subseteq A$ of alternatives up for consideration, an agent might judge, based on their evidence ranking $\succcurlyeq_i^t$, that there are alternatives deserving of their support at time t. We parse this by distinguishing between two types of agents, *lazy* and *keen*, and using the set $f_\bullet(\succcurlyeq_i^t, C)$ of *favored alternatives given C*, defined relative to the agent type:

$$f_{lazy}(\succcurlyeq_i^t, C) = \{x \in top(\succcurlyeq_i^t) \mid x \succ_i^t y, \text{ for all } y \in C\},$$

$$f_{keen}(\succcurlyeq_i^t, C) = \begin{cases} \emptyset, \text{ if } top(\succcurlyeq_i^t) = C, \\ \{x \in C \mid x \succ_i^t y \text{ for some } y \in C\} \cup \\ \{x \in A \setminus C \mid x \succcurlyeq_i^t y \text{ for some } y \in C\}, \\ \text{otherwise.} \end{cases}$$

Intuitively, the favored set contains alternatives that the agent is willing to disclose evidence in favor of, if the set of winners is C. Lazy and keen agents then differ in how they determine the set of alternatives they wish to lend their support to. Lazy agents support only top-ranked alternatives they consider strictly better to every element in C. In contrast, keen agents support any alternatives that move the result closer to their evidence order in the following manner: if C is exactly their top-ranked choices, they are happy and withhold support, otherwise, they support every $x \in C$ that they strictly prefer to some alternative in C, along with supporting those alternatives not in C that they weakly prefer to some member of C—this includes alternatives supported by lazy agents, but also potentially more. Disclosure is the engine behind deliberation, and the feature to which we now turn.

Deliberation occurs in rounds, with each round corresponding to a time t. During a round $t \geqslant 1$, each agent i: (1) decides whether to disclose evidence for alternatives in $f_\bullet(\succcurlyeq_i^{t-1}, C)$ based on a set $K \subseteq E$ of *public evidence items* and a set C of alternatives up for consideration, and (2) updates their evidence sets with any evidence disclosed throughout the round. As i's evidence sets get updated so does their evidence ranking, such that i starts the round with evidence ranking $\succcurlyeq_i^{t-1}$ and ends it with $\succcurlyeq_i^t$.

We write $D_i^t(x)$ for the set of evidence items supporting x that i discloses at round t, and make the following global assumptions. First, any evidence disclosed at round t gets absorbed into agents' evidence sets before the end of the round:

$E_i^t(x) = E_i^{t-1}(x) \cup \left(\bigcup_{j \in N} D_j^t(x)\right)$, for all $x \in A$, i.e., i's evidence set for x at the end of round t consists of i's evidence set for x at $t-1$ together with all the evidence for x disclosed throughout round t. Note that this condition applies only to i's evidence set at the end of round t, but leaves open the possibility (a possibility we will exploit) that i updates $E_i^{t-1}(x)$ incrementally. Note, as well, that $E_i^t(x)$ differs from $E_i^{t-1}(x)$ only if i at t finds out information about x that they did not know at $t-1$ (no double-counting of evidence).

Second, we require that when disclosing, agents do not repeat evidence that is already known. We say that an evidence item $e \in E$ is *private to i at t* if $e \in E_i^t(x)$, for some $x \in A$, and $e \notin K$. Agents disclose only evidence that is private. If C is a set of alternatives up for consideration, we write $u_i^t(C) = \{e \in E \mid \exists x \in f_\bullet(\succcurlyeq_i^{t-1}, C) \text{ s.t. } e \in E_i^{t-1}(x) \text{ and } e \notin K\}$ for the set of private evidence items in support of some alternatives that agent i is in support of with respect to C at t. If $u_i^t(C) \cap E(x) \neq \emptyset$, we say that *$i$ dissents on x at t* by disclosing evidence from $u_i^t(C)$.

We say that protocol P *terminates at round t* if there are no dissenters at t, and thus there is no change in agents' evidence sets. If P terminates at t, we say that the *final winners* are the plurality winners $F_{PL}(\succcurlyeq^t)$ of the profile at termination.

3 Deliberation Protocols

We now detail two types of deliberation protocols.

Simultaneous protocol. This protocol consists of one disclosure instance per round, in which all dissenting agents disclose *all* the evidence available to them, for all the alternatives on which they dissent.

Definition 1 (Simultaneous protocol P_{sim})**.** At round $t = 0$, set $K = \emptyset$ and $C = \emptyset$. For $t \geqslant 1$, start by setting $C = F_{PL}(\succcurlyeq^{t-1})$, for all $i \in N$. During round t each agent discloses all items of evidence in $u_i^t\left(F_{PL}\left(\succcurlyeq^{t-1}\right)\right)$. The round ends with both public knowledge and agents' evidence sets getting updated with all disclosed information, i.e., $K = K \cup \left(\bigcup_{i \in N} D_i^t(x)\right)$ and $E_i^t(x) = E_i^{t-1}(x) \cup K$, for all $i \in N$ and $x \in A$.

Intuitively, P_{sim} approximates a process in which disclosed information is pooled and presented to agents all at once (e.g., by posting it on a discussion board), after which agents relay their updated top-ranked choices to an aggregation mechanism. The plurality winners are announced to the group, possibly triggering a new round of disclosure.

Sequential Protocols. In *sequential protocols*, rounds are characterized by agents taking turns in nominating and disclosing evidence. A running tally is kept of the growing list of nominations, with each agent i being aware of who is winning based on the nominations of their predecessors. During their turn, an agent learns this information, discloses (or not), and may nominate (or not) their supported alternatives. As soon as evidence is disclosed, all agents update their evidence sets. Thus, in a sequential deliberation protocol the set of alternatives under consideration can change from one agent to other, as the list of nominations grows; similarly, since agents update their evidence sets incrementally, their evidence rankings can change multiple times during one round.

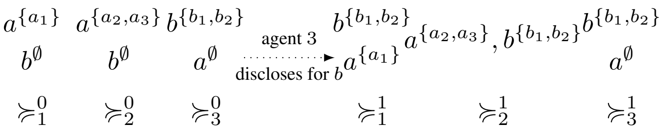


Figure 1: Revision of evidence rankings from $t = 0$ to $t = 1$ as a result of evidence disclosure with the simultaneous protocol P_{sim}. Evidence sets supporting an alternative are written as superscripts; higher means better in the evidence ranking.

How to handle intermediary winners within each round? Fix, first, an arbitrary ordering $\rhd = (\pi_1, \ldots, \pi_n)$ of the agents, given by a permutation of N, and denote by $\boldsymbol{\alpha}_{\pi_i}^t$ the vector of alternatives nominated by all agents up to, and including, agent π_i at round t. Vector $\boldsymbol{\alpha}_{\pi_i}^t$ is used differently by the following sequential protocols.

Definition 2 (Sequential Constant Protocol $P_{\textit{seq-con}}$)**.** At $t = 0$ set $K = \emptyset$ and $C = \emptyset$. At $t \geqslant 1$, start by setting $C = F_{PL}(\boldsymbol{\alpha}_{\pi_n}^{t-1})$, i.e., the first agent in $\rhd$ decides based on the previous round, and $E_i^t(x) = E_i^{t-1}(x)$, for all $i \in N$ and $x \in A$. Then, for every $i \geq 2$, set $C = F_{PL}(\boldsymbol{\alpha}_{\pi_{i-1}}^t)$, i.e., each agent $i \geqslant 2$ decides based on the plurality winners over the list $\boldsymbol{\alpha}_{\pi_{i-1}}^t$ of alternatives nominated by their predecessors. Nominations are handled thus: at the beginning of the round the list of nominations is empty, i.e., $\boldsymbol{\alpha}_0^t = ()$; agent π_i adds to $\boldsymbol{\alpha}_{i-1}^t$ alternatives on which they dissent; if there are no such alternatives, they add all their top-ranked alternatives. Each agent π_i discloses either one item of evidence in $u_{\pi_i}^t(C)$, or nothing if they do not dissent. As soon as π_i discloses, public knowledge and agents' evidence sets are updated: $K = K \cup D_{\pi_i}^t(x)$, and $E_i^t(x) = E_i^t(x) \cup K$, for all $i \in N$ and $x \in A$.

Definition 3 (Sequential Abstention Protocol $P_{\textit{seq-abs}}$)**.** At round $t = 0$, set $K = \emptyset$, $C = \emptyset$ At $t = 1$, agent π_1 kicks things off by nominating their top-ranked alternatives and by disclosing an item of evidence supporting each of them. Following this, agents decide based on the currently winning alternatives, i.e., for π_1 we set $C = F_{PL}(\boldsymbol{\alpha}_{\pi_n}^{t-1})$ while for $\pi_{i \geq 2}$, we set $C = F_{PL}(\boldsymbol{\alpha}_{\pi_{i-1}}^t)$. Nominations are handled differently to the sequential constant protocol. First, rather than the list resetting at the beginning of each round, nominations stay in place as we cycle through π. Then, agents nominate alternatives they dissent on, but *abstain* from nominating if they have nothing to dissent on. In another difference, agents disclose an item of evidence for each alternative they dissent on. Finally, updates occur, as for $P_{\textit{seq-con}}$, as soon as there is a disclosure event: $K = K \cup D_{\pi_i}^t(x)$ and $E_i^t(x) = E_i^t(x) \cup K$, for all $i \in N$ and $x \in A$.

Though slightly idealized, these protocols formalize rules of information exchange we could reasonably expect to see in real-life deliberations. The next example shows that it is not difficult to find cases where things go wrong.

Example 1. *Take alternatives $A = \{a, b\}$ supported by evidence $E(a) = \{a_1, a_2, a_3\}$ and $E(b) = \{b_1, b_2\}$, and a profile $\succcurlyeq^0 = (\succcurlyeq_1^0, \succcurlyeq_2^0, \succcurlyeq_3^0)$ with the initial evidence distribution depicted in Figure 1. Note that a is both the plurality winner at $t = 0$, and the optimal alternative, as $|E(a)| > |E(b)|$.*

Under P_{sim}, at $t = 1$ agent 3 discloses both b_1 and b_2. All agents update their evidence sets and 1 and 2 revise their evidence rankings accordingly, such that $\succcurlyeq^1 = (b \succ_1^1 a,\ a \sim_2^1 b,\ b \succ_3^1 a)$; it is the end of round $t = 1$ and b is the sole plurality winner. If agents are lazy no agent dissents, and deliberation stops at $t = 1$ with b as final winner. If agents are keen (i.e., want to see a winning as much as b), then agent 2 dissents for a at $t = 2$, at the end of which the profile is $\succcurlyeq^2 = (a \succ_1^2 b,\ a \sim_2^2 b,\ a \sim_3^2 b)$ with a back as the winner. Agents 2 and 3 support b but have no new evidence for it, so the protocol stops at $t = 2$ with a as the final winner.

For $P_{seq\text{-}con}$, let the agent ordering be $1 \rhd 2 \rhd 3$ with the agents being keen. At $t = 1$ agents 1 and 2 nominate a on their turns. At this point the list of nominations is (a, a), so 3 nominates, and discloses an evidence item for, b; based on the list of nominations (a, a, b), a is the round winner. To start round $t = 2$, agent 1 is tied between a and b. With a being the sole winner of the previous round, agent 1 supports b. However, they have no private evidence for b so they cannot dissent and hence simply nominate their top choices, a and b. This is followed by 2 nominating a and disclosing evidence for it, then 3 nominating b and disclosing evidence for it. At the end of round $t = 2$ the list of nominations is (a, b, a, b), with a and b as tied round winners. At $t = 3$ there is no more disclosure, and b is crowned final winner based on the sequence of nominations (a, b, a, b, b).

For $P_{seq\text{-}abs}$ the agent ordering is also $1 \rhd 2 \rhd 3$. Agent 1 starts by nominating a and disclosing a_1; agent 2 skips their turn, since they do not dissent; agent 3 nominates b and discloses an item of evidence for it. At the end of $t = 1$ the profile is $\succcurlyeq^1 = (a \sim_1^1 b,\ a \succ_2^1 b,\ b \succ_3^1 a)$. Round $t = 2$ starts with the list of nominations (a, b) built during the previous round. Assuming all agents are keen, we get: 1 skips; 2 dissents, nominates a and discloses for it, bringing the list of nominations to (a, b, a); agent 3 dissents again, nominates b and discloses their remaining item of evidence for it. At $t = 3$ the list of nominations is (a, b, a, b), so agent 1 skips their turn; 2 nominates a and discloses their last item of private evidence for it; 3 does not dissent, as they have no private evidence for b. The protocol stops with a as the final winner.

4 Results

Note that, since evidence is finite, there must be a round where no agent dissents: either via all agents being satisfied, or due to agents running out of private evidence to disclose. Thus, Protocols P_{sim}, $P_{seq\text{-}con}$ and $P_{seq\text{-}abs}$ terminate after a finite number of rounds. Ideally, though, deliberation not only ends, but also leads to the objectively best decisions: an outcome that, as illustrated by Example 1, is not guaranteed by our protocols. Thus, moving forward, our aim is to identify conditions under which our protocols produce optimal outcomes. To preface the results, we make the following two assumptions on the initial distribution of evidence:

(A_1) *Completeness*: $E(x) = \bigcup_{i \in N} E_i^0(x)$, for all $x \in A$.

(A_2) *Disjointness*: $E_i^0(x) \cap E_j^0(x) = \emptyset$, for $i \neq j \in N$, $x \in A$.

Completeness rules out situations in which deliberation has no chance of succeeding because of implausible draws of evidence (e.g., no one gets any evidence), and also motivates our notion of success: one consequence of completeness is that if all private information became public, agents would all converge to the same—objectively correct—ranking. However, since we can anticipate that the communication costs associated with such an outcome would be considerable, this is not the goal we set for a deliberation: correct decisions can be achieved with less information. Thus, to borrow an existing term from the literature [Ding and Pivato, 2021], we say a protocol P is *full-disclosure equivalent* if the final winners consist of the optimal alternative.

$$\begin{array}{ccccc}
 & & & & y^{|E(y)|} \\
 & & & & a_0^{|E(a_0)|} \\
x^{|E(y)|-\frac{k-1}{n-1}} & x^{|E(y)|-\frac{k-1}{n-1}} & & x^{|E(y)|-\frac{k-1}{n-1}} & x^{|E(y)|-1} \quad \dots \\
y^0, a_0^0, \dots, a_s^0 & y^0, a_0^0, \dots, a_s^0 & & y^0, a_0^0, \dots, a_s^0 & a_s^{|E(a_s)|} \\
 & & & & \\
\succcurlyeq_1^0 & \succcurlyeq_2^0 & \cdots & \succcurlyeq_{n-1}^0 & \succcurlyeq_n^0
\end{array}$$

Figure 2: Profile of Theorem 1, evidence amounts are superscripts.

Disjointness is worth wanting because without it, one can easily find distributions of evidence in which there is unanimous agreement on a sub-optimal alternative and deliberation never even kicks off. To get deliberation going, there needs to be at least one agent backing the optimal alternative, and assumptions A_1 and A_2 imply that there exists some agent at $t = 0$ that is in support of the better among a pair of two alternatives, i.e., if x is objectively better than y then x cannot be weakly Pareto dominated by y at $t = 0$.

One thing that should help the optimal alternative is the amount of evidence supporting it: intuitively, the more evidence (relative to non-optimal alternatives), the better. We make this intuition formal by deriving exact bounds on the amount of evidence needed.

Theorem 1. *If $|A| \geqslant 2$ and the initial evidence distribution satisfies A_1 and A_2, then, if x is the optimal alternative and y is the second-best, with $|E(y)| \geqslant 1$ and $n \geqslant 3$, every $P \in \{P_{sim}, P_{seq\text{-}con}, P_{seq\text{-}abs}\}$ is full-disclosure equivalent iff: (i) $|E(x)| \geqslant n \cdot |E(y)|$, when agents are lazy, and (ii) $|E(x)| > n \cdot |E(y)| - n$, when agents are keen.*

Proof. Take $A = \{x, y, a_0, \dots, a_s\}$ with $|E(x)| > |E(y)| \geqslant |E(a_0)| \geqslant \dots \geqslant |E(a_s)|$. We prove the claim for lazy agents (keen agents is analogous), and for P_{sim}.

("$\Rightarrow$") Suppose $|E(x)| < n \cdot |E(y)|$, rewritten as $|E(x)| = n \cdot |E(y)| - k$, for $k > 0$. Since $|E(x)| > |E(y)|$ we infer that $|E(y)| - {}^{k}/_{(n-1)} > 0$. Consider a scenario where at $t = 0$ agent n has all the evidence for y and $|E(y)| - 1$ amounts of evidence for x. Agent $i \in \{1, \dots, n-1\}$ gets $|E(y)| - k_i$ amounts of evidence for x, where $k_i \approx {}^{(k-1)}/_{n-1}$ and $k_1 + \dots + k_{n-1} = k - 1$. This works as an amount of evidence, since $|E(y)| - {}^{(k-1)}/_{(n-1)} > |E(y)| - {}^{k}/_{(n-1)} > 0$. Remaining alternatives are handled by giving all the evidence for them to agent n. See Figure 2 for an illustration of the resulting profile. At $t = 1$ agent n dissents and discloses their information for their favorite alternatives, of which y is one, and these alternatives become the plurality winners. Deliberation stops after this, with x not among the final winners.

("$\Leftarrow$") If $|E(x)| \geqslant n \cdot |E(y)|$, then $|E_i^0(x)| \geqslant |E(y)|$, for at least one agent $i \in \{1, \dots, n\}$. This ensures that agent

i places x as their strictly top alternative at round 0, such that if x is not initially the winner then agent i discloses their evidence for x at $t = 1$. This disclosure persuades every other agent to put x on top of their evidence orders. There is no further disclosure for y, which ensures that x is the sole election winner at termination.

For $P_{seq\text{-}con}$ and $P_{seq\text{-}abs}$, the left-to-right argument from above works unmodified. For the right-to-left direction, construct a counterexample using the profile in Figure 2, with the agent ordering $1 \rhd \cdots \rhd n$. Agent n will disclose incrementally until they lift y to the top of everyone else's ranking. □

The bounds of Theorem 1 seem large: are they needed in practice? We return to this question in Section 5, but before that, we take a closer look at each protocol.

The Simultaneous Protocol. Can we get any *good* results with P_{sim} outside of the bounds given by Theorem 1 P_{sim}? Below we identify a particular situation that guarantees full-disclosure equivalence for two alternatives.

Theorem 2. *For $A = \{a, b\}$ with $a \in A$ as the optimal alternative and a complete and disjoint initial distribution of evidence, we have that if $a \notin F_{PL}(\succcurlyeq^0)$, then P_{sim} is full-disclosure equivalent, for both keen and lazy agents.*

Proof. There must be at least one agent j that strictly prefers a to b at $t = 0$, and thus dissents for a at $t = 1$. Let N^* be the set of dissenters for a at $t = 1$ (the set varies depending on the agent types). Note that $|E^0_i(a)| \geqslant |E^0_i(b)|$ for each $i \in N^*$, and $|E^0_i(a)| < |E^0_i(b)|$, for all $i \in N \setminus N^*$. Under P_{sim}, all agents in N^* disclose all their private evidence in support of a at $t = 1$. Assume agents are keen (the argument works similarly if agents are lazy). Note that for agents in $N \setminus N^*$, their private evidence for b at $t = 1$ remains unchanged as they absorb all disclosed evidence for a released by the dissenters in N^*. Disclosure by agents in N^*, we claim, convinces all agents in $N \setminus N^*$ to think a is strictly better than b at $t = 1$. Assume, towards a contradiction, that there is $\ell \in N \setminus N^*$ such that $b \succcurlyeq^1_\ell a$. The disjointness assumption implies that $|E^0_\ell(a)| + \sum_{j \in N^*} |E^0_j(a)| < |E^0_\ell(b)|$, and the completeness assumption implies that $|E(a)| = |E^0_\ell(a)| + \sum_{i \neq \ell \in N \setminus N^*} |E^0_i(a)| + \sum_{j \in N^*} |E^0_j(a)|$. It follows that $|E(b)| > |E(a)|$, contradicting the optimality of a. □

Theorem 2 points towards a feature that shows up in other contexts: the protocol favors initial underdogs, as their supporters rally for them and sometimes change the status quo. Note, however, that Theorem 2 fails to hold for more than two alternatives, as the rally can end up propping the entirely wrong alternative (see the Appendix for an example).

The Sequential Constant Protocol. First off we have that, in line with similar results in the literature [Hartmann and Rafiee Rad, 2020], the set of final winners with $P_{seq\text{-}con}$ is sensitive to the order in which agents communicate.

Example 2. *Take $N = \{1, \ldots, 5\}$, $A = \{a, b\}$, $E(a) = \{a_1, a_2, a_3\}$, $E(b) = \{b_1, b_2\}$, with initial evidence $E^0_i(a) = \{a_i\}$ and $E^0_i(b) = \emptyset$ for $i \in \{1, 2, 3\}$, $E^0_4(a) = \emptyset$, $E^0_4(b) = \{b_1\}$, $E^0_5(a) = \emptyset$ and $E^0_5(b) = \{b_2\}$. Take the agent ordering $1 \rhd 2 \rhd 3 \rhd 4 \rhd 5$. With $P_{seq\text{-}con}$ and either keen or lazy agents, at $t = 1$ we see agents 1, 2 and 3 nominating (but not disclosing for) a, then 4 and 5 nominating and disclosing for b, making b the unanimous winner at termination. If agents are ordered in reverse, a comes out as unique final winner.*

Example 2 raises the question of whether, given the evidence distribution, there exists an agent ordering that makes $P_{seq\text{-}con}$ full-disclosure equivalent: a problem faced by, e.g., the organizer of a debate tasked with determining the order in which people speak. The answer hinges on how much knowledge the organizer has about the distribution of evidence. In the case of two alternatives and full access to the initial distribution, it is straightforward to see that an agent ordering in which supporters of b (the worse alternative) speak first, in descending order of $|E^0_i(b)| - |E^0_i(a)|$, followed by supporters of a, guarantees a win for a. If the organizer knows only which alternative is optimal and what the evidence orders look like, but not necessarily the distribution of evidence among the agents, the following result offers a solution.

Theorem 3. *If $A = \{a, b\}$ with a as optimal, and the initial evidence distribution is complete, disjoint and with no tied agents, then an agent ordering where agents who put b at the top come first, followed by agents who put a at the top, guarantees full-disclosure equivalence for $P_{seq\text{-}con}$.*

Proof Sketch. The proof works by finding bounds for the maximum amount of evidence for b that can be released throughout the deliberation rounds. If $A, B \subseteq N$ are agents in N who put a and b at the top, respectively, then the protocol unfolds with agents in B nominating b, because they come first, followed by agents in A disclosing evidence for a until some agents in B are flipped to a, and potentially triggering disclosure for b. It is shown that the agents in b who disclose for b cannot flip all the agents in A, such that even if b regains majority support, then there are agents in A that can still disclose for a and, eventually, make a win. □

If the organizer knows less, e.g., only agents' rankings but not the optimal alternative, then very little can be done: Theorem 1 already implies that, for any agent ordering, we can find distributions of evidence that make either alternative win. Note, as well, that for $|A| = 2$ and keen agents, some of which can start out neutral, allowing supporters of the non-optimal alternative to begin the deliberation, as in Theorem 3, does not guarantee full-disclosure equivalence for $P_{seq\text{-}con}$.

Proposition 4. Take $A = \{a, b\}$ with a being the optimal alternative and a complete, disjoint profile. If neutral agents are present, then $P_{seq\text{-}con}$ is not full-disclosure equivalent, even for orderings $\rhd$ of a keen agent population where agents who put b at the top come first.

Proof. Consider the counterexample in Figure 3 with the b-first ordering $1 \rhd \ldots \rhd 33$. Agents $1 - 7$ nominate b, agent 8 follows by nominating and disclosing for a; agents $9 - 14$, initially neutral, now support a and only nominate a as their top-ranked alternative, as they have no evidence for it; agent 15 nominates a and discloses an evidence item for it; the remaining agents just nominate a. Deliberation ends at $t = 2$ with b as the unanimous winner, even though a is optimal. □

$$\begin{array}{ccccc}
b^{\{b_i\}} & b^{\{b_j^1,b_j^2,b_j^3,b_j^4\}} & a^{\{a_8\}} & a^{\emptyset} & a^{\{a_\ell\}} \\
a^{\emptyset} & a^{\emptyset} & b^{\emptyset} & b^{\emptyset} & b^{\emptyset} \\
\succcurlyeq^0_{i\in\{1,2,3\}} & \succcurlyeq^0_{j\in\{4,\dots,7\}} & \succcurlyeq^0_8 & \succcurlyeq^0_{k\in\{9,\dots,14\}} & \succcurlyeq^0_{\ell\in\{15,\dots,33\}}
\end{array}$$

Figure 3: Initial evidence distribution for profile in Proposition 4.

Whether there exists an agent ordering that guarantees full-disclosure equivalence for keen agents that may be neutral remains an open question, and highlights the care that must be taken in ensuring good results for deliberation. Luckily, the $P_{seq\text{-}abs}$ protocol, to which we turn to next, nicely complements the mixed bag of results obtained for $P_{seq\text{-}con}$.

The Sequential Abstention Protocol. For $P_{seq\text{-}abs}$ we show that when $|A| = 2$, a balance is maintained between the evidence that has been disclosed for each alternative, regardless of the agent ordering. We write $K^t(a) = \{e \in E(a) \mid e \in K \text{ at time } t\}$ for the public evidence supporting $a \in A$ at t.

Lemma 5. For $A = \{a, b\}$ with $a \in A$ as the optimal alternative, and a complete and disjoint initial evidence distribution, under $P_{seq\text{-}abs}$ it holds that $||K^t(a)| - |K^t(b)|| \leqslant 1$ for every round t, ordering $\rhd$, and agent type.

Proof sketch. For every $a \in A$, the number of nominations at end of round t is $|K^t(a)|$. For any $a \in A$ (both keen and lazy agents), with $P_{seq\text{-}abs}$ there is no nomination, or disclosure, in favour of a at any round $t + 1$ where $C = \{a\}$, which would be the case if $|K^t(a)| - |K^t(b)| = 1$ for $b \in A$. □

The following result states that for any agent ordering of keen agents, $P_{seq\text{-}abs}$ works as desired for two alternatives.

Theorem 6. *For $A = \{a, b\}$ with a being the optimal alternative, and an initial evidence distribution that is complete and disjoint, it holds that $P_{seq\text{-}abs}$ is full-disclosure equivalent for any ordering $\rhd$ of keen agents.*

Proof. For $P_{seq\text{-}abs}$, consider any round t starting with either $C = \{b\}$, or $C = \{a, b\}$ as the current winning alternative set. The process terminates if all n agents abstain on their turn when $C = \{b\}$, i.e., if none of the agents dissent for a, and thus nominate a. For $C = \{b\}$, or $C = \{a, b\}$, no agent dissents on their turn at t, then we know that for every agent, the evidence they are privy to, publicly and privately, is either in favour of b, or equal between a and b. From Lemma 5, $\big||K^t(a)| - |K^t(b)|\big| \leqslant 1$ for all rounds t. So if no agent discloses then it must be the case that each agent's amount of private evidence supporting a at t is at most as large as their private evidence for b (otherwise they would dissent for a). This implies that the total amount of evidence known by all agents, both privately and publicly, that supports b, surpasses that for a, contradicting the optimality of a. □

With $A = \{a, b\}$ and lazy agents, full-disclosure equivalence no longer holds under $P_{seq\text{-}abs}$ for any agent ordering $\rhd$, as the following example illustrates.

Example 3. *Take $A = \{a, b\}$ and $N = \{1, 2, 3\}$ with $|E_1^0(a)| = 0$ and $|E_1^0(b)| = 1$ for agent 1, while $|E_j^0(a)| = 1$ and $|E_j^0(b)| = 0$ for agents $j \in \{2, 3\}$. Note that a is the optimal alternative with $|E(a)| = 2 > 1 = |E(b)|$. Then observe that, under protocol $P_{seq\text{-}abs}$ with the ordering $1 \rhd 2 \rhd 3$ and all three agents being lazy, after the nomination and disclosure for b by agent 1, both agents 2 and 3 will be neutral between a and b, so will both abstain during their turns due to being lazy, and with b winning via previous nominations.*

Also, the positive result of Theorem 6 does not hold for $m \geqslant 3$ alternatives (see the Appendix for an example). In this case, more conditions are required on the agent ordering in order to guarantee full-disclosure equivalence. We leave to future work the task of characterising these exact requirements.

5 Evidence Gap and Spread

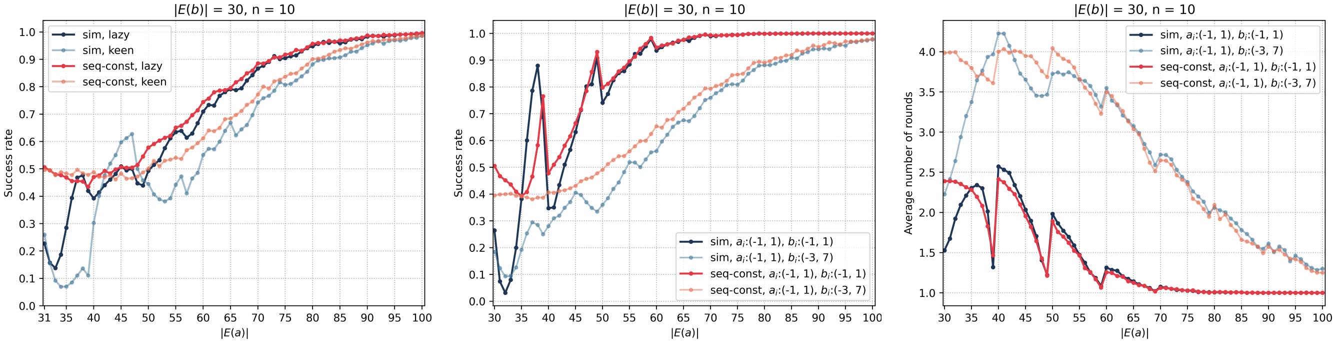
We want to understand how (i) the evidence gap in favor of the optimal alternative, herewith a, and (ii) the spread of evidence among the agents, influences the final outcome of a deliberation. Recall, from Theorem 1, that $|E(a)|$ needs to be around n times larger than $|E(b)|$ for the protocols to guarantee good results. What happens if $|E(a)|$ is below this bound? The numerical simulations presented in this section indicate that, under mild assumptions on the initial distribution of evidence, good results occur even for smaller values of $|E(a)|$.

What kind of assumptions? One common feature of bad outcomes in Example 1 and similar cases is that evidence for a starts out being distributed roughly equally, with all agents getting a similar share, whereas evidence for b is heavily skewed towards a few agents. These agents become exceptionally motivated to reveal evidence, and end up derailing the final outcome. The hypothesis we test is that such cases, with an unbalanced initial distribution, spell trouble.

Setup. We look at the case of two alternatives $A = \{a, b\}$, with a as the optimal alternative for complete and disjoint initial distributions of evidence. For given $|E(x)|$, $x \in A$, number of agents n and agent type, we randomly generate a vector $\boldsymbol{x} = (x_1, \dots, x_n)$, where $\boldsymbol{x}$ is an integer partition of $|E(x)|$, i.e., $x_i \in \mathbb{N}$, $i \in \{1, \dots, n\}$ and $x_1 + \cdots + x_n = |E(x)|$. For each $\boldsymbol{x}$ we can specify a set of parameters that control its variance. Each x_i stands for the amount of evidence for x that agent i has at $t = 0$. Doing this for every alternative in A creates an initial profile $\succcurlyeq^0$, and we run P_{sim} and $P_{seq\text{-}con}$ on the generated instance. For each combination of parameters we run 5000 instances.[1] The *success rate* is the fraction of instances where the optimal alternative a is the final winner.

Success rate as function of evidence gap. The *evidence gap for a over b* is defined as $|E(a)| - |E(b)|$, telling us not just that a is objectively better than b, but also by how much. For the simulations we set $n = 10$ agents, the amount of evidence for b to $|E(b)| = 30$, and varied $|E(a)|$ from 31 to 100. Results are shown in Figure 4a. As expected, the success rate of all protocols gets better as the evidence gap for a grows: more evidence for a relative to b means a has a better

[1] We leave $P_{seq\text{-}abs}$ out because, at least in the case of keen agents, it is full-disclosure equivalent (see Theorem 6).



(a) Impact of evidence gap

(b) Impact of variance in initial distribution

(c) Number of rounds to termination

Figure 4: Simulation results. Note the drops in success rate in Figure 4b, when evidence amounts are multiples of the number of agents and the variance is small, an interesting artefact of the numbers involved. For instance, for $n = 10$, $|E(a)| = 40$, $|E(b)| = 30$, we expect most agents get around 3 items of evidence for b and 4 for a: a starts out on top of most agents' rankings, albeit by a narrow margin. At the same time, we can also expect a small number of agents who start out favoring b over a: because they start out in the minority, these agents disclose at the first round and, due to the fragile advantage of a, it becomes very likely that all the a-supporters get flipped at this point, allowing for a victory for b. This intuition is confirmed by Figure 4c, showing that deliberation in these cases usually stops after about one round.

chance of being the final winner. Interestingly, we see near perfect success rate around $|E(a)| = 90 = 3 \cdot |E(b)|$, (surprisingly) smaller than the amount $n \cdot |E(b)| = 300$ suggested by Theorem 1. In performing these simulations, all evidence distributions $\boldsymbol{x}$ are set to similar variance, for each $x \in A$. The assumption here is that whatever process operates in the background to provide agents with evidence, it works in similar ways for all alternatives. But what happens if it does not?

Success rate and variance. For the second batch of simulations we keep the variance in the evidence distribution for a constant and small, such that all agents receive similar amounts of evidence for a, while increasing the variance in the distribution for b. Intuitively, this is in line with the hypothesis, outlined above, that unbalanced distributions lead to bad results because they allow some agents to hold large amounts of evidence for b. Results are shown in Figure 4b, and they bear out this intuition. The numbers in parentheses indicate by how much x_i, the amount of evidence i has for x, is allowed to deviate from the average $^{|E(x)|}/_{n}$. Thus, for $n = 10$ agents and $|E(a)| = 30$, an equal distribution of evidence sees every agent get $^{30}/_{10} = 3$; a pair $(-1, 1)$ means that each agent actually has in between 2 and 4 items. Thus, larger spreads lead to vectors of evidence with higher variance, and cases where the distribution for b is based on a higher spread have lower success rate.

Number of rounds to termination. Finally, we look at the impact that variance in the initial distribution has on the length of the process. For each batch of 500 deliberation rounds, we plot the average number of rounds to termination. As before, results are plotted against a growing amount of evidence for a, while $|E(b)|$ is kept fixed. Results are presented in Figure 4c and show that unbalanced distributions lead to longer deliberation sessions, especially for values of $|E(a)|$ close to $|E(b)|$. As the evidence gap between a and b grows larger there is a growing chance that profiles are close to consensus for a, with less need for drawn out discussions.

6 Conclusion

We have put forward an election model that incorporates deliberation between agents, where deliberation, formalized as three protocols ($P_{sim}, P_{seq\text{-}con}, P_{seq\text{-}abs}$), is based on selective disclosure of private evidence. We have found that even under severe restrictions of the evidence distribution (i.e., completeness and disjointness) there is ample space for things to go wrong, as motivated lobbies for sub-optimal alternatives can steer the rest of the electorate away from the truth. Future work may reasonably seek to relax these assumptions: in real-life scenarios, where agents have access to similar sources, we can expect considerable overlap in evidence sets. However, our work points to significant pitfalls that can occur even in the restricted setting in which this does not happen, the pitfalls being symptomatic of what can go wrong in a debate. In Section 4 we found theoretical guarantees if the evidence gap for the optimal alternative is large enough; if not, we need careful orchestration of the rules of debate: sequential protocols can be successful by regulating the order in which agents speak, putting limits on how much they can say, and by exploiting the behavioral assumptions leading them to disclose (i.e., whether keen or lazy). Simulations in Section 5 show that, on average, the optimal alternative has a better chance of winning, even when not supported by an overwhelming amount of evidence, if evidence is distributed in a similar way across agents. The latter finding lends support to the idea that diversity is good for decision making [Hong and Page, 2004; Page, 2007]: in our context, few agents monopolizing all evidence for an alternative indicates a non-diverse crowd.

Looking forward, our model can be nudged even closer to reality by: allowing agents to also reject evidence; adding a probabilistic model of the acquisition of evidence; or, in following the lead of the literature on learning in social networks [Acemoglu *et al.*, 2010; Golub and Jackson, 2010; Golub and Sadler, 2016; Bikhchandani *et al.*, 2021], by restricting communication to a local neighborhood.

Acknowledgments

We thank the reviewers for their useful comments.

References

[Acemoglu *et al.*, 2010] Daron Acemoglu, Asuman Ozdaglar, and Ali ParandehGheibi. Spread of (mis)information in social networks. *Games and Economic Behavior*, 70(2):194–227, 2010.

[Auletta *et al.*, 2015] Vincenzo Auletta, Ioannis Caragiannis, Diodato Ferraioli, Clemente Galdi, and Giuseppe Persiano. Minority Becomes Majority in Social Networks. In *Proc. of WINE 2015*, pages 74–88, 2015.

[Auletta *et al.*, 2019] Vincenzo Auletta, Diodato Ferraioli, Valeria Fionda, and Gianluigi Greco. Maximizing the Spread of an Opinion when Tertium Datur Est. In *Proc. of AAMAS 2019*, pages 1207–1215, 2019.

[Bächtiger *et al.*, 2018] André Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark E. Warren. *The Oxford Handbook of Deliberative Democracy*. Oxford University Press, 2018.

[Bikhchandani *et al.*, 2021] Sushil Bikhchandani, David Hirshleifer, Omer Tamuz, and Ivo Welch. Information Cascades and Social Learning. NBER Working Papers 28887, National Bureau of Economic Research, Inc, June 2021.

[Brill *et al.*, 2016] Markus Brill, Edith Elkind, Ulle Endriss, and Umberto Grandi. Pairwise Diffusion of Preference Rankings in Social Networks. In *Proc. of IJCAI 2016*, pages 130–136, 2016.

[Chung and Duggan, 2020] Hun Chung and John Duggan. A formal theory of democratic deliberation. *American Political Science Review*, 114(1):14–35, 2020.

[Condorcet, 1785] Marie J. A. N. de Caritat, Marquis de Condorcet. *Essai sur l'Application de L'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. Imprimerie Royale, 1785.

[Dietrich and Spiekermann, 2013] Franz Dietrich and Kai Spiekermann. Epistemic Democracy With Defensible Premises. *Economics & Philosophy*, 29(1):87–120, 2013.

[Dietrich and Spiekermann, 2020] Franz Dietrich and Kai Spiekermann. Jury Theorems. In *The Routledge Handbook of Social Epistemology*, pages 386–396. Routledge, 2020.

[Dietrich and Spiekermann, 2021] Franz Dietrich and Kai Spiekermann. Social Epistemology. In *The Handbook of Rationality*, pages 579–590. MIT Press, 2021.

[Dietrich and Spiekermann, 2022] Franz Dietrich and Kai Spiekermann. Deliberation and the Wisdom of Crowds. *Documents de travail du Centre d'Économie de la Sorbonne*, 2022.

[Ding and Pivato, 2021] Huihui Ding and Marcus Pivato. Deliberation and epistemic democracy. *Journal of Economic Behavior & Organization*, 185:138–167, 2021.

[Elkind and Slinko, 2016] Edith Elkind and Arkadii Slinko. Rationalizations of Voting Rules. In *Handbook of Computational Social Choice*, pages 169–196. Cambridge University Press, 2016.

[Elster, 1998] Jon Elster. *Deliberative Democracy*. Cambridge University Press, 1998.

[Fain *et al.*, 2017] Brandon Fain, Ashish Goel, Kamesh Munagala, and Sukolsak Sakshuwong. Sequential Deliberation for Social Choice. In *Proc. of WINE 2017*, pages 177–190, 2017.

[Fishkin and Laslett, 2003] James S. Fishkin and Peter Laslett. *Debating Deliberative Democracy*. Blackwell, 2003.

[Golub and Jackson, 2010] Benjamin Golub and Matthew O. Jackson. Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*, 2(1):112–149, February 2010.

[Golub and Sadler, 2016] Benjamin Golub and Evan Sadler. Learning in Social Networks. In *The Oxford Handbook of the Economics of Networks*, pages 503–542. Oxford University Press, 2016.

[Goodin and Spiekermann, 2018] Robert E. Goodin and Kai Spiekermann. *An Epistemic Theory of Democracy*. Oxford University Press, 2018.

[Grofman *et al.*, 1983] Bernard Grofman, Guillermo Owen, and Scott L. Feld. Thirteen Theorems in Search of the Truth. *Theory and Decision*, 15(3):261–278, 1983.

[Hartmann and Rafiee Rad, 2018] Stephan Hartmann and Soroush Rafiee Rad. Voting, deliberation and truth. *Synthese*, 195(3):1273–1293, 2018.

[Hartmann and Rafiee Rad, 2020] Stephan Hartmann and Soroush Rafiee Rad. Anchoring in Deliberations. *Erkenntnis*, 85(5):1041–1069, 2020.

[Hong and Page, 2004] Lu Hong and Scott E. Page. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proc. of the National Academy of Sciences*, 101(46):16385–16389, 2004.

[Ladha, 1992] Krishna K. Ladha. The Condorcet Jury Theorem, Free Speech, and Correlated Votes. *American Journal of Political Science*, pages 617–634, 1992.

[Landemore, 2013] Hélène Landemore. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton University Press, 2013.

[Lev and Rosenschein, 2012] Omer Lev and Jeffrey S. Rosenschein. Convergence of Iiterative Voting. In *Proc. of AAMAS 2012*, pages 611–618, 2012.

[List and Goodin, 2001] Christian List and Robert E. Goodin. Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy*, 9(3):277–306, 2001.

[List, 2007] Christian List. Group Deliberation and the Transformation of Judgments: An Impossibility Result. 2007. LSE PSPE Working Paper No. 4, Available at SSRN: https://ssrn.com/abstract=1078127.

[List, 2018] Christian List. Democratic Deliberation and Social Choice: A Review. In *The Oxford Handbook of Deliberative Democracy*, pages 463–489. Oxford University Press, 2018.

[Meir, 2017] Reshef Meir. Iterative Voting. In *Trends in Computational Social Choice*, chapter 4, pages 69–86. AI Access, 2017.

[Michelini *et al.*, 2022] Matteo Michelini, Adrian Haret, and Davide Grossi. Group Wisdom at a Price: Jury Theorems with Costly Information. In *Proc. of IJCAI 2022*, pages 419–425, 2022.

[Owen *et al.*, 1989] Guillermo Owen, Bernard Grofman, and Scott L. Feld. Proving a distribution-free generalization of the Condorcet Jury Theorem. *Mathematical Social Sciences*, 17(1):1–16, 1989.

[Page, 2007] Scott E. Page. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press, 2007.

[Perote-Peña and Piggins, 2015] Juan Perote-Peña and Ashley Piggins. A model of deliberative and aggregative democracy. *Economics & Philosophy*, 31(1):93–121, 2015.

[Pivato, 2017] Marcus Pivato. Epistemic democracy with correlated voters. *Journal of Mathematical Economics*, 72:51–69, 2017.

[Pivato, 2019] Marcus Pivato. Realizing Epistemic Democracy. In *The Future of Economic Design*, pages 103–112. Springer, Cham, 2019.